

**IDENTIFICAÇÃO E CLASSIFICAÇÃO DE LITOFÁCIES COM O USO
DA TEORIA BAYESIANA DE RECONHECIMENTO DE PADRÕES**

IGOR LOPES SANTANA BRAGA

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE - UENF
LABORATÓRIO DE ENGENHARIA E EXPLORAÇÃO DE PETRÓLEO - LENEP**

**MACAÉ - RJ
DEZEMBRO - 2005**

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

IDENTIFICAÇÃO E CLASSIFICAÇÃO DE LITOFÁCIES COM O USO DA TEORIA BAYESIANA DE RECONHECIMENTO DE PADRÕES

IGOR LOPES SANTANA BRAGA

Tese apresentada ao Centro de
Ciência e Tecnologia da Universidade
Estadual do Norte Fluminense, como
parte das exigências para obtenção
do título de Mestre em Engenharia de
Reservatório e de Exploração

Orientador: Luiz Geraldo Loures, Ph.D.

MACAÉ - RJ
DEZEMBRO - 2005

IDENTIFICAÇÃO E CLASSIFICAÇÃO DE LITOFÁCIES COM O USO DA TEORIA BAYESIANA DE RECONHECIMENTO DE PADRÕES

IGOR LOPES SANTANA BRAGA

Tese apresentada ao Centro de Ciência e Tecnologia da Universidade Estadual do Norte Fluminense, como parte das exigências para obtenção do título de Mestre em Engenharia de Reservatório e de Exploração

Aprovada em 16 de dezembro de 2005

Comissão Examinadora:

Prof. Fernando Sergio de Moraes, Ph.D. (LENEP/UENF)

Prof. Roseane Marchezi Misságia, Ph.D. (LENEP/UENF)

Prof. Jadir da Conceição da Silva, Ph.D. (E&P-UNBC/PETROBRAS)

Prof. Luiz Geraldo do Carmo Lucchesi Loures, Ph.D. (LENEP/UENF)
(orientador)

AGRADECIMENTOS

Agradeço a todos que contribuíram para a realização deste trabalho. Em especial ao meu pai, por sempre me incentivar em assuntos relacionados a geociência. A minha mãe por toda uma vida dedicada a educação, não somente de seu filho, mas a de milhares de brasileiros. A minha esposa por toda a compreensão e carinho oferecidos durante a realização deste trabalho. A Zezé por ter sido uma segunda mãe por boa parte de minha vida e a todos familiares pelo suporte e incentivo a minha vida profissional.

Ao meu orientador Luiz Geraldo (Gera) por todas as contribuições fornecidas durante toda a realização deste trabalho e por sua amizade.

Aos professores do LENEP/CCT/UENF que foram fundamentais para a minha formação durante estes dois anos. Agradeço também aos colegas de classe que direta ou indiretamente ajudaram-me durante este percurso, em especial a Raquel Teixeira por valiosas discussões. Ao PRH-20 pelo suporte financeiro oferecido, a título de bolsa e recursos para a participação em congressos. A todas as empresas que oferecem licenças de uso de seus software que são de extrema importância para o desenvolvimento científico do país.

RESUMO

Este trabalho apresenta o desenvolvimento de uma metodologia para a identificação e classificação de litofácies utilizando a teoria bayesiana de reconhecimento de padrões. O algoritmo Expectation-Maximization é utilizado para realizar a interpretação automática de perfis de poços. Os poços interpretados formam o dado de treino, necessário para a classificação de dados sísmicos, utilizando um classificador de máxima probabilidade a posteriori (MAP). Estas duas técnicas formam uma estrutura de trabalho que combinam múltiplas fontes de informação para a realização do objetivo do trabalho.

Análise de incerteza, baseada em dados de perfil de poços, é realizada para prever quais atributos sísmicos fornecem os resultados classificatórios de maior sucesso.

O produto final da metodologia consiste no mapeamento no volume do reservatório das litofácies identificadas, e associada a cada classificação, uma medida de incerteza. Realizaram-se testes em dados sintéticos e dados reais. Compararam-se os resultados da identificação de litofácies, com a interpretação de um geólogo.

ABSTRACT

This work presents the development of a methodology for identification and classification of lithofacies, based on the bayesian theory of pattern recognition. The Expectation-Maximization algorithm is used to interpret wells, in a non-supervised framework. The interpreted wells are used as training data, necessary to classify the seismic volume, using a maximum a posterior probability classifier (MAP). This two techniques combine a variety of information sources to accomplish the objective of the work.

Uncertainty analyses, based on well log data, is used to predict which seismic attributes will give the best classification results.

The final product of this methodology is the mapping in the reservoir volume of de identified lithofacies and associated with each sample classified, a measure of uncertainty. Tests are realized in synthetic and real data. The results are evaluated and compared with the interpretation of a geologist.

SUMÁRIO

AGRADECIMENTOS.....	ii
RESUMO.....	iii
ABSTRACT.....	iv
1 Introdução	1
1.1 Objetivo	2
1.2 Estado da Arte	3
1.3 Estrutura da tese	6
2 Relações entre atributos sísmicos e propriedades da rocha	8
2.1 Introdução	8
2.2 Modelo efetivo do meio	9
2.3 Relações de V_p , V_s , densidade e porosidade.....	12
2.3.1 Relação $V_p - V_s$	14
2.3.2 Relação de Gardner ($V_p - \rho$)	16
2.3.3 Relação de Wyllie's ($V_p - \phi$)	17
2.3.4 Relação de Han (V_p , V_s , porosidade e volume de argila).....	18
2.4 Efeito da saturação	19

2.4.1	Aplicação da substituição de fluido	20
2.5	Fator de escala	22
2.6	Conclusões	24
3	Teoria de Decisão Bayesiana	25
3.1	Introdução	25
3.2	Teoria de Decisão – caso de duas classes	26
3.2.1	Estendendo o problema para mais de duas classes	27
3.3	Rede Bayesiana	29
3.4	Classificação por k-vizinhos mais próximos	31
3.5	Classificador robusto baseado em redes Bayesiana	32
3.6	Processo de classificação de litofácies	33
3.7	Conclusões	34
4	Identificando fácies litológicas em poços	36
4.1	Introdução	36
4.2	Formas tradicionais de identificação de litofácies através de perfis.....	36
4.2.1	Perfis tipicamente utilizados para a determinação de litofácies.....	37
4.2.2	Determinação de litofácies através da combinação de perfis.....	40
4.3	Definindo litofácies de forma automatizada com o uso do algoritmo EM.....	41
4.3.1	Algoritmo EM	42
4.3.2	Identificação do número ótimo de componentes	45
4.3.3	Passo a passo do uso do Algoritmo EM para a identificação de litofácies.....	46
4.4	– Conclusão	48
5	Quantificando incertezas.....	49
5.1	Introdução	49
5.2	Erro de Bayes	50

5.3	Reduzindo incertezas	51
5.4	Utilizando o erro de Bayes para a escolha dos atributos mais significativos.....	52
5.5	– Conclusão	53
6	Testes com dados sintéticos	54
6.1	Introdução	54
6.2	Fuxograma da metodologia	54
6.3	Descrição do modelo sintético	55
6.4	Teste do algoritmo EM com dados de poço sintético	59
6.5	Avaliação do Bayesian Information Criteria (BIC)	62
6.5.1	Estimativa inicial de μ quando não é conhecido o valor central dos atributos das litofácies.....	63
6.6	Escolhendo os atributos que forneçam a melhor qualidade de classificação	64
6.7	Classificando a seção de atributos sísmicos com o classificador MAP.....	69
6.7.1	Teste com as seções sintéticas com 5% de ruído	71
6.7.2	Classificação utilizando outros atributos com maior nível de incerteza.....	71
6.7.3	Teste com as seções sintéticas com 10% de ruído	72
6.8	Conclusão	76
7	Aplicação da metodologia desenvolvida em um conjunto de dados reais	77
7.1	Introdução.....	77
7.2	Descrição do Campo	77
7.3	Fontes de informação.....	79
7.4	Análise de incerteza dos dados	83
7.5	Classificando poços a partir de outros poços	84
7.6	Aplicação em dados sísmicos utilizando litofácies definidas por intérprete.....	93
7.7	Aplicação em dados sísmicos utilizando litofácies definidas aplicando EM.....	96

7.7.1	Análise petrofísica e comparação das litofácies identificadas com as interpretadas.....	98
7.7.2	Análise de incertezas para a escolha dos parâmetros com menor risco de classificação.....	105
7.7.3	Classificação de seção sísmica com as litofácies identificadas automaticamente.....	107
7.8	Conclusões	111
8	Conclusões	112
8.1	Recomendações para trabalhos futuros	114
	Referências Bibliográficas	115

Capítulo 1

INTRODUÇÃO

O conhecimento das heterogeneidades internas de reservatórios de petróleo é de grande importância para os trabalhos de desenvolvimento e produção de campos de hidrocarbonetos. A otimização de custos de perfuração e da vida útil dos reservatórios está diretamente relacionada ao grau de conhecimento existente. A falta de detalhamento dos reservatórios é a principal causa da baixa recuperação de hidrocarbonetos, podendo resultar em desvios de possíveis alvos e varredura ineficiente dos reservatórios (Mukerji et al. ,2001).

A partir da década de 80 ocorreu grande avanço tecnológico nas áreas de aquisição sísmica, processamento sísmico e ciências da computação. Estes fatos associados ao alto custo de exploração de petróleo, nas novas fronteiras de água profundas e a competitividade na indústria de exploração, desencadearam uma corrida tecnológica para a descrição das propriedades internas do reservatório, visando melhor recuperação de petróleo, levando ao desenvolvimento de novas tecnologias.

A identificação de litofácies, juntamente com a estimativa de propriedades petrofísicas (porosidade, saturação e volume de argila), definem a etapa de

caracterização de reservatório, sendo na atualidade um grande desafio da indústria de E&P. Concomitantemente, surgem uma variedade de atributos sísmicos e a física de rocha, objetivando a interpretação quantitativa do reservatório, trabalho integrado entre geólogos, geofísicos, engenheiros de petróleo e cientistas da computação (Barnes, 2001). Entretanto, a caracterização de reservatórios é uma tarefa que lida com incertezas, e quantificar essas incertezas é de extrema importância para a tomada de decisão da estratégia operacional a ser aplicada a um determinado campo.

1.1 Objetivo

O objetivo principal deste trabalho é o desenvolvimento de uma metodologia para classificação e inferência de fácies litológica, com base em dados petrofísicos e atributos sísmicos, tendo como resultados finais:

- I. classificação das fácies presentes em função da resposta petrofísica;
- II. o mapeamento, no volume do reservatório, das fácies classificadas na etapa anterior, a partir de dados de atributos sísmicos;
- III. uma medida de incertezas do modelo de fácies obtido.

Duas abordagens foram analisadas: i – o uso de um classificador de máxima probabilidade a posteriori em conjunto com dados de treinamento previamente definidos (aprendizado supervisionado) e ii- a aplicação do algoritmo (EM) “Expectation Maximization” (Dampster et al. , 1977) para encontrar padrões naturais ocultos nos dados e assim definir um conjunto de dados de treinamento de forma totalmente automatizada (aprendizado não supervisionado).

A metodologia desenvolvida foi testada em dados sintéticos para análise de resolução e sensibilidade a diferentes formas de ruído. Após testes com dados sintéticos a metodologia foi novamente testada, porém, em dados reais, visando uma avaliação de sua aplicabilidade prática e de seu comportamento em situação real onde as fontes de incertezas relacionadas a ruído não estão sob controle.

1.2 Estado da Arte

Através de novas técnicas de processamento geofísico e inversão elástica de AVO (Amplitude Versus Off-set) é possível extrair atributos elásticos como V_p , V_s , ρ , I_p e I_s , (Mahmoudian e Margrave, 2004). Estes parâmetros elásticos caracterizam a resposta sísmica do meio geológico, e são fontes de informação valiosa para a caracterização de reservatório.

Atributos elásticos, isoladamente, são de difícil interpretação, não tendo significado quantitativo do ponto de vista do conteúdo de hidrocarbonetos presente e da capacidade produtora. Modelos de física de rocha tornam-se elementos chave para a estimativa de propriedades petrofísicas, tais modelos relacionam os parâmetros petrofísicos com parâmetros elásticos, como por exemplo o modelo de Han et al. 1986, que relaciona V_p ou V_s com porosidade e volume de argila. Exemplos de estimativa de parâmetros petrofísicos utilizando modelos de física de rocha, podem ser encontrados em Loures e Moraes (2002) .

Para cada litologia é necessária a calibração do modelo de física de rochas a ser aplicado, ou seja, para cada situação litológica do reservatório (cada fácies) haverá um modelo de física de rochas especificamente calibrado. Sendo assim, a obtenção da distribuição espacial das litofácies, torna-se uma etapa da caracterização de reservatório a ser realizada antes das etapas de inferência de porosidade, volume de argila, saturação e pressão.

A identificação de litofácies convencional é baseada na interpretação de seções 2D sísmicas, onde o intérprete busca características distintas que tornem um grupo de reflexões diferente das reflexões adjacentes (Sheriff e Geldart, 1995). Estas reflexões são então correlacionadas com as informações dos perfis de poços. Este processo de identificação pode levar meses até que um volume de dados 3D seja interpretado e depende exclusivamente da experiência do intérprete para um bom resultado.

Outra abordagem bastante comum é a utilização de técnicas de interpolação (kriging e co-kriging). Esta só fornece resultados de boa qualidade quando existe um número expressivo de poços dentro do reservatório (Corvi et al., 1992).

Recentes metodologias foram apresentadas com o objetivo de identificar e classificar litofácies. Elas fazem uso de algoritmos computacionais para agrupar dados com características comuns (clusterização) ou classificam novas amostras com base em um padrão pré-determinado (classificação). Estas técnicas têm o objetivo de tornar tal tarefa automática, fazendo com que o computador “simule” o raciocínio humano, porém com uma velocidade de processamento altamente superior.

Neste cenário, pode-se citar alguns trabalhos:

Em Takahashi (2000), são expostos diversos fatores que geram incertezas na estimativa de propriedades das rochas, sendo feita extensa análise do comportamento dos diversos tipos de rochas em relação a vários atributos elásticos e petrofísicos. Além disso, propõe-se uma forma de quantificar incertezas associadas a estimativas com o uso do erro de Bayes. Tendo sido este trabalho de grande contribuição para a análise e a quantificação de incerteza envolvidas no processo de classificação de litofácies, segundo a teoria de decisão bayesiana.

Avesth et al. (2001) realizou estudo integrado de análise de perfis, física de rocha e inversão de AVO, onde a partir de um poço interpretado cuidadosamente, as litofácies existentes foram identificadas. Outros poços são classificados utilizando análise discriminante. No volume sísmico, é realizada inversão de AVO a fim de obter atributos de refletividade a zero-offset e gradiente. A partir dos poços foram criadas funções de distribuição de probabilidade para os dois. O volume sísmico foi classificado segundo a abordagem bayesiana, sendo seu resultado final o de um volume sísmico onde as litofácies eram interpretadas devido a sua probabilidade de ocorrência. Neste trabalho foi também discutido como certos atributos contém uma quantidade maior de informação sobre uma determinada litofácies em relação a outros atributos. Embora esta “quantidade de informação” não tenha sido

quantificada, este trabalho introduziu a idéia de fluxo de trabalho para a classificação de litofácies em dados sísmicos.

Mukerji et. al. (2001) fez uso dos atributos impedância acústica e impedância elástica e um classificador baseado em distância de Mahalanobis, para classificar litofácies. Ele realizou análise do comportamento de diversas litofácies em certos domínios, juntamente com estudos de influência da presença de fluidos distintos nas rochas.

O uso de estatística multivariada foi aplicado por Bosch et al. (2002) que utilizou três técnicas diferentes para discriminar litologias: agrupamento hierárquico, classificador gaussiano e análise discriminante. Os dados foram obtidos sob condições laboratoriais e as propriedades físicas utilizadas foram: densidade, velocidade da onda P, velocidade da onda S, condutividade elétrica e susceptibilidade magnética. Rochas de diferentes tipos foram analisadas em dois grupos, o primeiro com características físicas bem distintas (gabro, dolomita, mármore e arenito) onde a tarefa de discriminação é mais simples e o segundo onde as rochas possuíam características semelhantes (granito, gnais, gabro e basalto). Tal trabalho demonstra que é possível realizar discriminação de litologias quando possuímos informações suficientes sobre o comportamento físico das rochas e estas informações são utilizadas em conjunto.

Tratando-se de trabalhos nacionais, seu número, nesta área, é extremamente escasso, podendo-se citar Cunha (2002), onde é utilizada rede neural do tipo "multilayer perceptron" para classificar poços com base em dados de poços e testemunhos. Porém, é constatado problema de super especialização de classes (definição de litofácies de forma muito detalhada, através de testemunhos, onde esta escala de detalhamento só pode ser percebida por análises laboratoriais) o que dificulta a classificação de dados de perfil, havendo a necessidade de agrupar subcategorias em classes com menor detalhamento. Embora o trabalho citado acima apresente bons resultados, ele não foi estendido a dados sísmicos, limitando sua aplicação na indústria de exploração de petróleo.

O trabalho de Xuelei e Lei, (2004) aplica o critério de seleção de modelos BIC (Bayesian Information Criteria) utilizado em conjunto com o algoritmo EM para a determinação do número ótimo de agrupamentos (fácies distintas) em dados de poços. Este método, isoladamente, exige que o número de categorias distintas seja previamente definida pelo usuário. A combinação do EM e do BIC fornece uma estrutura de agrupamento não supervisionado.

1.3 Estrutura da tese

A estrutura geral deste trabalho apresenta-se da seguinte forma:

No capítulo 2 é apresentada a base de física de rocha necessária para o entendimento de como as variações de atributos sísmicos podem servir de indicadores de certos tipos litológicos.

No capítulo 3 é feita uma introdução à Teoria Bayesiana de Decisão e aos fundamentos na construção de um classificador de máxima probabilidade a posteriori com o uso de redes bayesianas. São apresentados os aspectos teóricos, e também a exemplos práticos de aplicação desta técnica para a classificação de litofácies.

No capítulo 4 são descritos tipos de perfis de poços utilizados como fonte de informação para o trabalho, sendo também apresentados métodos convencionais para a identificação de fácies em perfis de poços. Um novo método para a identificação de fácies é proposto, o qual utiliza aprendizado não supervisionado (algoritmo EM). Tal procedimento cria automaticamente o conjunto de dados de treinamento utilizados pelo classificador de máxima probabilidade a posteriori.

No capítulo 5 é feita uma descrição da forma de quantificação das incertezas envolvidas no processo de classificação de litofácies, com a apresentação do conceito de erro de Bayes, utilizado como métrica para quantificar tais incertezas. É demonstrado como quantificar a diminuição do nível de incerteza devido ao acréscimo de informação.

No capítulo 6 são apresentados testes com dados sintéticos, com a finalidade de testar e avaliar a sensibilidade do método. São criados perfis de V_p , V_s e densidade e gradualmente é acrescentado ruído gaussiano para testar a “robustez” do classificador MAP.

No capítulo 7 são apresentados diferentes testes com atributos obtidos por inversão de AVO (seções de V_p e V_s). Testes com diferentes dados de treinamento são realizados: um com as litofácies definidas por um geólogo intérprete (aprendizado supervisionado) e outro de forma totalmente automatizada com as litofácies sendo identificadas com o uso do EM. Os resultados são analisados e comparados.

O trabalho é concluído no capítulo 8. É feita uma análise da metodologia proposta e de sua aplicabilidade prática.

Capítulo 2

RELAÇÃO ENTRE ATRIBUTOS ELÁSTICOS E PROPRIEDADES DA ROCHA

2.1 Introdução

Ondas elásticas possuem informações de rochas e fluídos em sub-superfície na forma de tempo de trânsito, amplitude e variação de fase. Inicialmente, o principal objetivo da sísmica de exploração de petróleo era o imageamento da sub-superfície, permitindo a interpretação geológica e o mapeamento de estruturas que poderiam servir de armadilhas para hidrocarbonetos. Com o avanço do poder de processamento dos computadores e das novas técnicas de processamento e interpretação, os dados sísmicos podem ser analisados para a determinação de litologias, porosidade e fluidos (Wang, 2001). Para isso, dados sísmicos são invertidos para parâmetros físicos do meio, como propriedades elásticas e densidade.

Estudar as relações entre atributos sísmicos e propriedades do meio é o objetivo da física de rocha. Estudos de física de rocha são de extrema importância para técnicas de discriminação de litologias, monitoramento 4D de reservatórios e detecção direta de hidrocarbonetos com análise de AVO; o instrumento fundamental que faz o elo entre os dados sísmicos e as propriedades dos reservatórios.

A resposta sísmica é sensível às propriedades do meio poroso, como aspectos texturais, propriedades físicas dos minerais, fluidos, pressão, temperatura, efeito diagenético, entre outros. Física de rocha descreve as rochas através de suas propriedades físicas e procura relacioná-las à resposta sísmica observada, através de modelos teóricos ou empíricos. Uma das aplicações da física de rocha é a análise de sensibilidade de atributo sísmicos a variações de propriedades litológicas. Isso é feito a partir de análises de dados de perfis de poços e experimento em laboratórios.

Neste capítulo, apresentam-se fundamentos básicos da física de rocha necessários para o entendimento da relação entre propriedades do meio poroso e assinatura sísmica. Serão apresentadas as principais relações teóricas e empíricas entre propriedades elásticas e parâmetros litológicos do meio poroso.

2.2 Modelo efetivo do meio

Rochas são agregados de grãos minerais. O espaço vazio dentro de uma rocha, gerado pelo seu processo de criação ou por eventos posteriores a sua criação é denominado de poro. A razão entre o volume de poros existentes em uma rocha e seu volume total é denominado porosidade (ϕ). Os poros de uma rocha encontram-se preenchidos por algum fluido (gás, água, óleo). Para uma melhor compreensão das relações entre propriedades da rocha e seus atributos elásticos, surgiu a necessidade da criação de modelos teóricos que descrevam o comportamento efetivo de seus constituintes.

Os modelos efetivos do meio são modelos criados para descrever as propriedades de uma rocha, baseada nos seus vários minerais constituintes, textura, fluido e efeitos diagenéticos. Os modelos efetivos do meio seguem basicamente duas abordagens: em uma a rocha é considerada um bloco elástico perturbado por buracos e são chamados de “modelos de inclusão”; na segunda, a rocha é descrita por grãos que permanecem em contato com outros e são chamados de “modelos de contato”. Independente da abordagem, esses modelos precisam especificar três tipos de informação: i) a fração de volume de cada constituinte mineral; ii) os módulos elásticos de cada fase; iii) informações geométricas de como cada fase está relacionada com a outra. Na prática as informações geométricas nunca foram

incorporadas adequadamente aos modelos teóricos, devido a complexidade geométrica dos poros das rochas, tais informações sempre foram simplificadas e aproximadas. O domínio K (módulo de incompressibilidade) - ϕ tem sido utilizado para o desenvolvimento de modelos teóricos que descrevam fatores que controlem as variações nos módulos elásticos das rochas (Mavko et al., 1998).

Quando fornecemos somente informações sobre as frações dos constituintes e módulos elásticos, desconsiderando as informações geométricas, é possível somente descrever os limites superiores e inferiores dos módulos elásticos e da velocidade de uma rocha composta. Uma importante aplicação do uso desses limites superior e inferior é na análise de física de rocha a partir de dados de perfis de poços. Os perfis de poços comumente trazem pouca informação sobre a forma geométrica do poro e dos grãos minerais, porém trazem informações sobre a fração de cada constituinte.

Os modelos mais clássicos que descrevem o limite superior e inferior dos módulos elásticos são denominados de Voigt e Reuss (Mavko et. al., 1998). Voigt define o limite superior e Reuss o limite inferior. Abaixo são apresentados esses modelos

$$M_v = \sum_{i=1}^n f_i M_i, \quad (2.1)$$

onde

M_v Módulo elástico efetivo de uma mistura de n fases constituintes;

f_i fração de volume do i-ésimo constituinte;

M_i módulo elástico do i-ésimo constituinte.

A expressão (2.1) é conhecida como limite superior de Voigt.

$$\frac{1}{M_r} = \sum_{i=1}^n \frac{f_i}{M_i}, \quad (2.2)$$

onde

M_r Módulo elástico efetivo de uma mistura de n fases constituintes

A expressão (2.2) é conhecida com limite inferior de Reuss. O M nas expressões (2.1) e (2.2) pode representar qualquer módulo elástico das rochas (módulo de incompressibilidade K ou módulo cisalhante μ). O módulo de incompressibilidade é definido como a deformação volumétrica que um corpo sofre ao serem aplicadas tensões, da mesma forma que o módulo de cisalhamento é a razão entre uma tensão cisalhante e sua deformação. Outro modelo que fornece limites inferiore e superiore para os módulos elásticos é o limite de Hashin e Shiktriman (Hashin e Shtrikman, 1963). Este modelo é conhecido por fornecer o menor intervalo entre os limites inferior e superior dos módulos elásticos, sendo considerado um modelo mais refinado. A Figura 2.1 ilustra o comportamento de uma rocha quando quatro propriedades (pressão, textura, diagênese, fluido) sofrem alteração. A linha em preto representa o limite superior (Voigt) e a linha em azul o limite inferior de Hashin e Shiktriman. Caso alguma amostra ultrapasse os limites, significaria uma falta de coerência entre algum valor do módulo elástico de alguma fase, utilizados para o cálculo do valor efetivo.

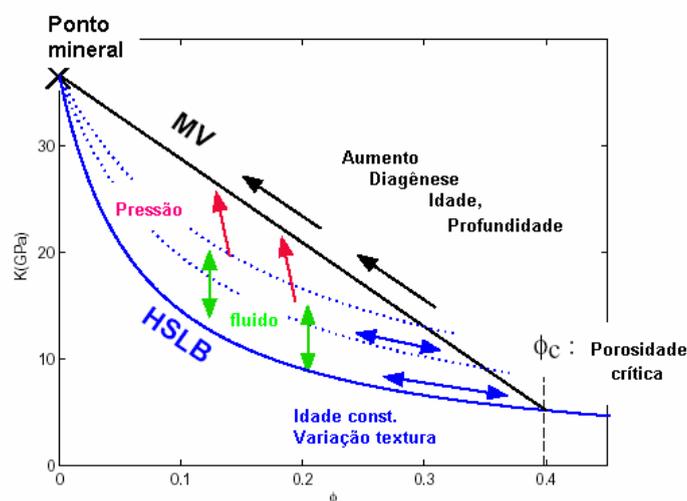


Fig. 2.1 – Comportamento elástico das rochas no domínio K - ϕ , com relação a mudança das seguintes propriedades; i) Pressão (setas vermelhas); ii) Mudança de fluido (setas verde); iii) Diagênese (setas pretas); iv) Textura (setas azuis). A linha MV representa o limite superior modificado de Voigt, e a curva HSLB o limite inferior de Hashin e Shiktriman (Retirado e modificado de Takahashi, 2000)

O modelo efetivo do meio pode ser descrito de forma ilustrativa através do uso do gráfico K- ϕ . A Figura 2.1 representa esquematicamente como as propriedades litológicas influenciam as amostras neste domínio. Quando a porosidade é controlada pela diagênese, a relação K- ϕ tende a seguir o modelo de Voigt, já quando a alteração de porosidade é controlada por variações texturais, K apresenta menos sensibilidade a porosidade (com exceção de situações de baixa porosidade) e a variação tende a seguir o limite inferior de Hashin e Shiktriman. Em relação à pressão efetiva, percebe-se que a variação é quase vertical, pois quando existe aumento da pressão, a rocha fica mais rígida e a porosidade sofre pequeno decréscimo. A mudança do fluido do poro de uma rocha não afeta a porosidade, mas afeta o K total da rocha. Abaixo do ponto de porosidade crítica ϕ_c é considerada estado de suspensão (onde os grãos estão dispersos em algum fluido). O ponto mineral é o ponto de porosidade zero.

Numa análise de dados petrofísicos através do gráfico K- ϕ obtém-se um modelo do comportamento das diversas fácies de uma área em estudo.

Nas próximas seções serão apresentados alguns modelos de física de rocha.

2.3 Relações de Vp, Vs, densidade e porosidade

As velocidades da onda-p e onda-s em um meio homogêneo e isotrópico são dadas por:

$$V_p = \left[\frac{k + \frac{4}{3}m}{\rho} \right]^{\frac{1}{2}} \quad (2.3)$$

e

$$V_s = \left[\frac{m}{\rho} \right]^{\frac{1}{2}}, \quad (2.4)$$

onde

V_p = velocidade da onda P (onda compressional),

V_s = velocidade da onda S (onda cisalhante),

K = módulo de bulk da rocha ou incompressibilidade,

μ = módulo de cisalhamento ou rigidez,

ρ = densidade da rocha.

As equações (2.3) e (2.4) fornecem relações fundamentais entre as velocidades sísmicas e porosidade e volume de argila, propriedades litológicas importantes na caracterização de reservatório.

Devido ao fato de material fluido não sofrer cisalhamento, o módulo de cisalhamento é insensível a alterações no mesmo (Castagna e Batzle, 1993). Desta forma, enquanto V_p é sensível a variações de fluido na rocha, V_s é praticamente insensível. As únicas variações de V_s devido a mudança no fluido são devido às mudanças na densidade volumétrica do material. Por este motivo, a razão entre a V_p/V_s é um bom indicador direto de hidrocarboneto, principalmente de gás. A razão V_p/V_s também possui boa capacidade de discriminar litologias. Entre as rochas reservatório a dolomita possui o maior valor de V_p , seguida de calcários e arenitos. Essa ordem é inversa para V_s , por isso, essas rochas possuem razão V_p/V_s distintas, podendo essa informação ser utilizada para a diferenciação das mesmas. A Tabela 1 apresenta alguns valores típicos de velocidades de minerais puros (Castagna e Batzle, 1993).

Tabela 1 – Valores típicos de propriedades de minerais puros (Castagna e Batzle, 1993).

Mineral	Densidade(g/cm³)	V_p (km/s)	V_s (km/s)	V_p/V_s
Calcita	2.71	6.53	3.36	1.94
Dolomita	2.87	7.05	4.16	1.70
Halita	2.16	4.50	2.59	1.74
Muscovita	2.79	5.78	3.33	1.74
Quartz	2.65	6.06	4.15	1.46
Anidrita	2.96	6.01	3.37	1.78

2.3.1 Relação Vp – Vs

Os perfis de Vs são necessários em várias tarefas na geofísica de reservatório, como análise de AVO e modelagem sísmica. Entender a dependência entre Vp e Vs é de extrema importância quando é necessário realizar estimativas de Vs em dados legados, que foram prospectados em épocas que a ferramenta sônica com fonte dipolar (ferramenta que mede o tempo de trânsito da onda P e da onda S) não era muito comum. Existem diversos modelos empíricos que descrevem a dependência de Vp e Vs. A relação de Pickett (Castagna e Backus, 1993) obtém um bom ajuste em calcários nas velocidades acima de 1,5 Km/s. Esta relação pode ser escrita como:

$$V_p = 1.9V_s . \quad (2.5)$$

A relação conhecida como “Mudrock line” (Castagna et al., 1985) ajustam-se adequadamente em arenitos:

$$V_s = 0.862V_p - 1.1724. \quad (2.6)$$

A expressão apresentada por Castagna et al. (1995) foi rescrita para os folhelhos, da seguinte forma:

$$V_s = 0.7696V_p - 0.86735. \quad (2.7)$$

Para rochas calcáreas, a seguinte expressão foi apresentada por Castagna e Batzle (1993)

$$V_s = -0.0550V_p^2 + 1.01677V_p - 1.03049 \quad (2.8)$$

Greenberg e Castagna (1992) introduziram uma relação para a estimativa de Vs a partir de Vp em rochas multiminerálicas e saturadas com água, baseado em regressões polinomiais de relações entre Vp e Vs:

$$V_s = \frac{1}{2} \left\{ \left[\sum_{i=1}^L f_i \sum_{j=0}^{N_i} a_{ij} V_p^j \right] + \left[\sum_{i=1}^L f_i \left(\sum_{j=0}^{N_i} a_{ij} V_p^j \right)^{-1} \right]^{-1} \right\}, \quad (2.9)$$

onde L é o número de minerais constituintes,

f_i fração de volume do i-ésimo constituinte,

a_{ij} os coeficientes de regressão empíricas,

N_i a ordem do polinômio do i-ésimo constituinte,

V_p e V_s as velocidades em Km/s.

Como na expressão (2.9) os dados de entrada devem ser proveniente de uma rocha 100% saturada em água, em aplicações reais, para uma melhor estimativa de V_s , a expressão (2.9) é combinada com a expressão 2.12 (Gassman, 1951) que realiza a substituição de fluido.

A Figura 2.2 mostra o “crossplot” de um conjunto de dados e as relações empíricas explicitadas acima. Os dados de V_s foram estimados utilizando a metodologia de Greenberg e Castagna (1992). As curvas em roxo, azul escuro, azul claro e laranja representam, respectivamente, a relação de Picket, Castagna para arenitos, Castagna para folhelhos e Castagna para os calcários. pode-se perceber um bom ajuste da estimativa, comparando-a com os modelos empíricos.

Um parâmetro que possui uma relação direta com a razão V_p/V_s é denominado razão de Poisson (σ). A razão de Poisson é definida como a fração negativa de mudança na largura de um corpo dividida pela fração de deformação de comprimento quando aplicamos tensões uniaxiais em um material isotrópico. A razão de Poisson pode ser formulada por:

$$\sigma = \frac{.5(V_p/V_s)^2 - 1}{(V_p/V_s)^2 - 1}, \quad (2.10)$$

$\sigma=0$ corresponde a uma razão V_p/V_s de 1.41. Material fluído possui $\sigma=0.5$ e razão V_p/V_s de infinito (Castagna e Batzle, 1993).

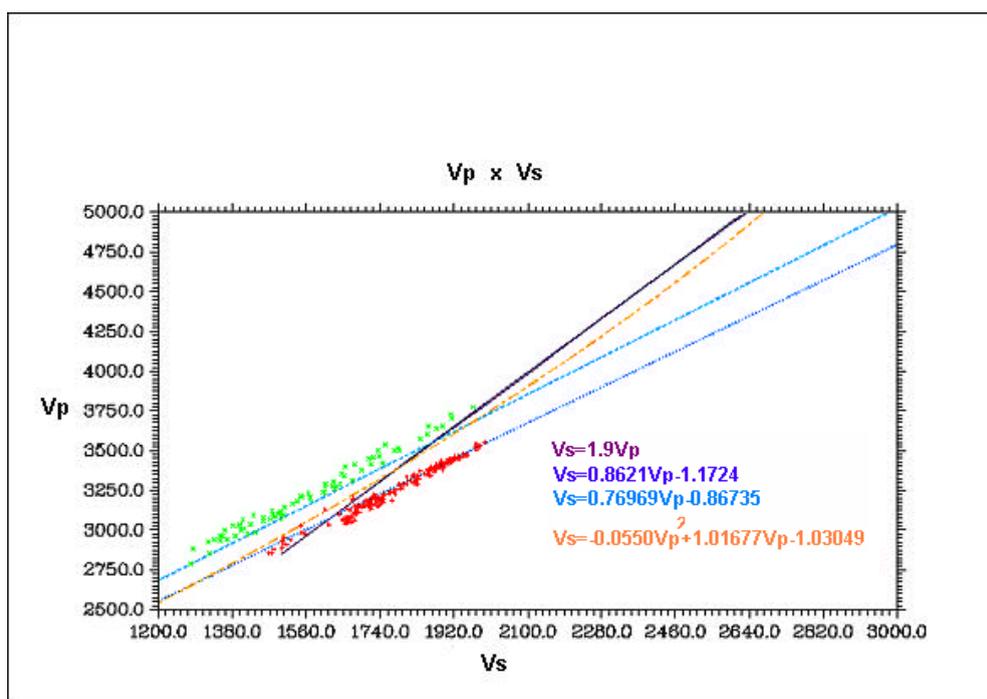


Fig 2.2 – Amostra de arenito em vermelho e folhelho em verde, demonstrando a coerência com as relações empíricas e a estimativa de Vs utilizando Greenberg e Castagna (1992).

2.3.2 Relação de Gardner (V_p - r)

Na teoria, velocidade nem sempre cresce junto com o aumento da densidade (Wang, 2001). A velocidade de uma rocha é mais dependente do módulo de bulk (K) e módulo de cisalhamento (μ) do que da densidade. Pode-se tomar como exemplo o caso da anidrita comparando-a com a dolomita. Anidritas possuem maior densidade total porém uma menor velocidade do que dolomitas. A relação de Gardner (Gardner, 1974) relaciona velocidade com densidade, expressa por:

$$r = aV_p^b, \quad (2.11)$$

onde $a=1741$ e $b= 0.25$.

A relação de Gardner se ajusta a uma variedades de rochas, de diferentes velocidades e porosidades. Por isso, a relação de Gardner é limitada por dois fatores, primeiramente a relação costuma subestimar V_p em rochas reservatório saturadas em água, em segundo, superestimar V_p em argilas saturadas em água.

Para litologias específicas podem ser levadas em consideração modificações dos coeficientes de Gardner conforme a Tabela 2.

Tabela 2 – Valores dos coeficientes para várias litologias retirados de Gardner (1974)

Litologia	A	B	Range de Vp (km/s)
Folhelho	1.75	0.265	1.5-5.0
Arenito	1.66	0.261	1.5-6.0
Calcário	1.50	0.225	3.5-6.4
Dolomita	1.74	0.252	4.5-7.1
Anidrita	2.19	0.160	4.6-7.4

2.3.3 Relação de Wyllie's (Vp-f)

Uma das relações que relaciona Vp com ϕ é apresentado por Wyllie et al. (1956), conhecida como “time-average equation” que pode ser expressa por:

$$\frac{1}{V_p} = \frac{(1-f)}{V_{p_{ma}}} + \frac{f}{V_f}, \quad (2.12)$$

onde

Vp = Velocidade da onda-p,

Vp_{ma} = Velocidade da onda-p na matriz e

V_f = velocidade da onda p no fluído.

A equação (2.12) pode ser interpretada como a propagação da onda por uma fração do tempo na matriz e uma fração no fluído. Esta representação só é fisicamente válida para comprimentos de onda menores que o tamanho dos poros/grãos, embora seja uma forma conveniente para estimar porosidade a partir de perfis de velocidade. A equação (2.12) fornece bons resultados quando utilizada em rochas consolidadas e tende a fornecer resultados errôneos quando aplicada a zonas de gás ou rochas pouco consolidadas. Os efeitos de física de rocha no domínio Vp- ϕ são similares aos efeitos do domínio K- ϕ (Figura 2.1) devido ao fato da alta

correlação entre K e V_p . Isto faz com que seja possível analisar os efeitos de pressão, textura, substituição de fluido e diagênese da mesma forma nos dois domínios.

2.3.4 Relações de Han (V_p , V_s , porosidade e volume de argila)

Velocidade e impedância das rochas decrescem com aumento da porosidade. Tal relação, contudo só é válida estatisticamente, pois as propriedades sísmicas das rochas também são afetadas pelo formato dos poros (Kuster e Toksoz, 1974). Rochas com baixas porosidade e poros achatados podem apresentar velocidades inferiores a rochas com alta porosidade e poros esféricos, devido a uma maior compressibilidade dos poros achatados em relação aos poros esféricos.

A sensibilidade à argila é muito forte nos atributos sísmicos. O efeito da argila depende mais da posição da partícula, do que do tipo de argila. Caso a argila faça parte da matriz da rocha e caso o tipo de argila seja mais incompressível que o quartzo, a velocidade crescerá. Caso a argila esteja preenchendo os poros da rocha, a não ser pela mudança de densidade, pouca será a influência nos atributos sísmicos. Um dos estudos mais detalhados das relações entre velocidade, porosidade e volume de argila foi realizado por Han et al. (1986) com amostras laboratoriais de 80 arenitos. A relação é formulada por:

$$V_p = a + a_1 f + a_2 j \quad (2.13)$$

e

$$V_s = b + b_1 f + b_2 j \quad (2.14)$$

onde, ϕ é o volume de argila e a e b são os coeficientes de regressão, proporcionais a pressão. A tabela 3 demonstra os coeficientes para arenitos saturados com água e com gás a 40 Mpa.

Tabela 3 – Valores dos coeficientes da relação de Han et al. (1986).

Pressão	a	a1	a2	b	b1	b2
Saturado em água						
40 Mpa	5.9	-6.93	-2.18	-3.52	-4.91	-1.89
Saturado em gás						
40 Mpa	5.41	-6.35	-2.7	-3.57	-4.57	-1.83

2.4 Efeito da saturação

O tipo de fluido (gás, água ou óleo) presente nos poros influencia de forma variada as propriedades sísmicas de uma rocha. Existe uma variedade de equações teóricas, que calculam os efeitos das propriedades, quando há substituição do fluido nas rochas. Uma das equações mais usuais (Gassman, 1951) calcula o módulo de incompressibilidade de uma rocha saturada por um certo fluido, sabendo o módulo de incompressibilidade da matriz, do arcabouço e do fluido. Esta equação forma a base das técnicas de substituição de fluidos. A equação de Gassman é escrita por:

$$K^* = K_d + \frac{(1 - K_d / K_m)^2}{\frac{\phi}{K_f} + \frac{1 - \phi}{K_m} - \frac{K_d}{K_m^2}}, \quad (2.15)$$

onde, K^* é o módulo de incompressibilidade de uma rocha saturada com um fluido de módulo K_f , K_d é o módulo de bulk do arcabouço, K_m o módulo de bulk da matrix e ϕ porosidade.

2.4.1 Substituição de fluido

A sensibilidade de V_p às propriedades do fluido e a não sensibilidade de V_s fornecem valiosas informações sobre a rocha. Suponha o seguinte objetivo: diferenciar um conjunto de amostras em duas categorias: i) rochas reservatório saturadas em água; e ii) rochas reservatório saturadas em óleo. Caso os atributos escolhidos para a realização de tal tarefa não fossem sensíveis ao fluido, alcançar tal objetivo seria uma tarefa impossível. De uma maneira oposta se o nosso foco for discriminar duas amostras em tipos litológicos diferentes, e não dispomos de informações sobre a saturação da rocha, deve-se utilizar atributos que sofram o mínimo de influência às propriedades dos fluidos presente e sejam sensíveis as propriedades litológicas.

Uma maneira de prever o comportamento de certas rochas à substituição de fluido é através de técnicas de substituição de fluido. A Fig 2.3 ilustra amostras de um arenito saturado em água (azul) e as mesmas amostras, após a substituição para óleo. Observe o diferente comportamento no gráfico K- V_p para as duas situações de saturação:

A Figura 2.4 representa o mesmo conjunto de dados apresentado na Figura 2.3, no domínio $\mu \times V_s$. Nota-se claramente a falta de sensibilidade às propriedades do fluido.

Para realizar a substituição de fluido, deve-se conhecer K_f e K_m . Proceda-se da seguinte maneira, calcula-se de K_d , sabendo a densidade da matriz, através da expressão:

$$K_d = r_d (V_p^2 - 4/3 V_s^2) \quad (2.16)$$

onde,

$$r_d = (1 - f) r_m \quad (2.17)$$

V_p , V_s e F são as velocidades e porosidade medida. Como o módulo de cisalhamento não sofre alteração, tem-se:

$$m^* = m_d \quad (2.18)$$

O cálculo de ρ^* (densidade da rocha saturada) pode ser obtido através da seguinte expressão:

$$r^* = r_d + fr_f, \quad (2.19)$$

onde ρ_f é a densidade do fluido. Com o cálculo dos módulos elásticos da rocha saturada e da densidade da rocha saturada, calcula-se as velocidades da rocha saturada.

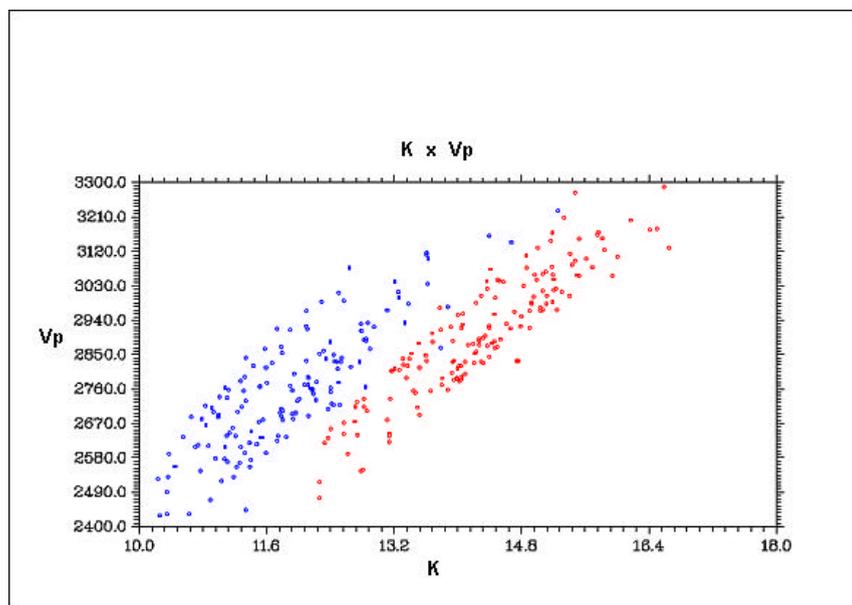


Fig. 2.3 – Amostras de arenito saturado em água (azul) e saturado em óleo (vermelho) após a substituição de fluido.

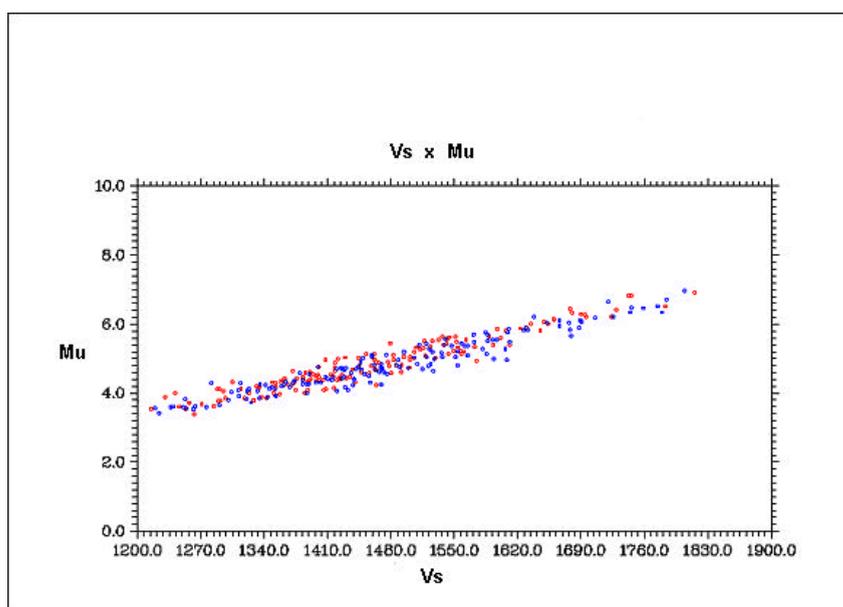


Fig. 2.4 – Arenito da Bacia de Campos. Pontos em vermelho representam a rocha saturada em água, azul em óleo

2.5 Fator de escala

A escala em que os diferentes métodos de prospecção trabalham são muito diferentes. As medidas de perfis de poços possuem resolução de menos de 0,3 metros enquanto que a resolução sísmica dificilmente ultrapassa 15 metros. Esse fato deve-se às diferentes frequências de levantamento de cada método. Na perfilagem de poços a frequência fica em torno de 10KHz já na sísmica este valor é está entre 10-80 Hz (Dewar e Pickford, 2001). A determinação da assinatura sísmica através de dados de perfis é denominada de “upscaling”. O upscaling de propriedades é uma tarefa crítica para a interpretação de eventos na sísmica quando os dados de perfis são utilizados como referência. Finas camadas em sub-superfície não são detectados pela sísmica. A realização de upscaling é uma forma de prever a visibilidade de tal evento.

Neste trabalho foi adotado o uso da “Backus averaging” chamado também de filtro de Backus (Backus, 1962) para a realização do “upscaling”. A importância de realizar o upscaling neste trabalho, deve-se ao fato que os dados que serão classificados encontram-se na escala sísmica, e os dados que serão utilizados como dados de treinamento, encontram-se na escala de perfis de poços. Como o padrão de treinamento deve representar, de forma fiel, o conhecimento sobre uma categoria, estes dados devem estar na mesma escala.

O filtro de Backus consiste em realizar a média aritmética da densidade para uma determinada janela de tamanho n , conforme a expressão (2.20)

$$\mathbf{r}_{eff} = \frac{1}{n} \sum_{i=1}^n \mathbf{r}_i, \quad (2.20)$$

onde ρ_{eff} é a densidade efetiva,

ρ_i a densidade da i -ésima profundidade ,

n o tamanho da janela,

e a realização de média geométrica para os módulos elásticos expressões (2.21) e (2.22) para o mesmo tamanho de janela.

$$\left(\frac{1}{K_{eff}} \right) = \frac{1}{n} \sum_{i=1}^n \frac{1}{K_i}, \quad (2.21)$$

$$\left(\frac{1}{\mathbf{m}_{eff}} \right) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\mathbf{m}_i}. \quad (2.22)$$

Tendo a densidade e os módulos elásticos na escala da sísmica, outras propriedades elásticas podem ser calculadas na mesma escala, como por exemplo, impedância P, impedância S, razão de Poisson, etc. A Figura 2.5 ilustra algumas propriedades petrofísicas na escala de perfis de poços (linha preta) e após a aplicação do filtro de Backus para a realização do upscaling (linha vermelha). Da esquerda para direita os perfis são: impedância P, Módulo de cisalhamento, módulo de bulk, lambda (primeiro parâmetro de Lamé)*densidade e módulo de cisalhamento * densidade.

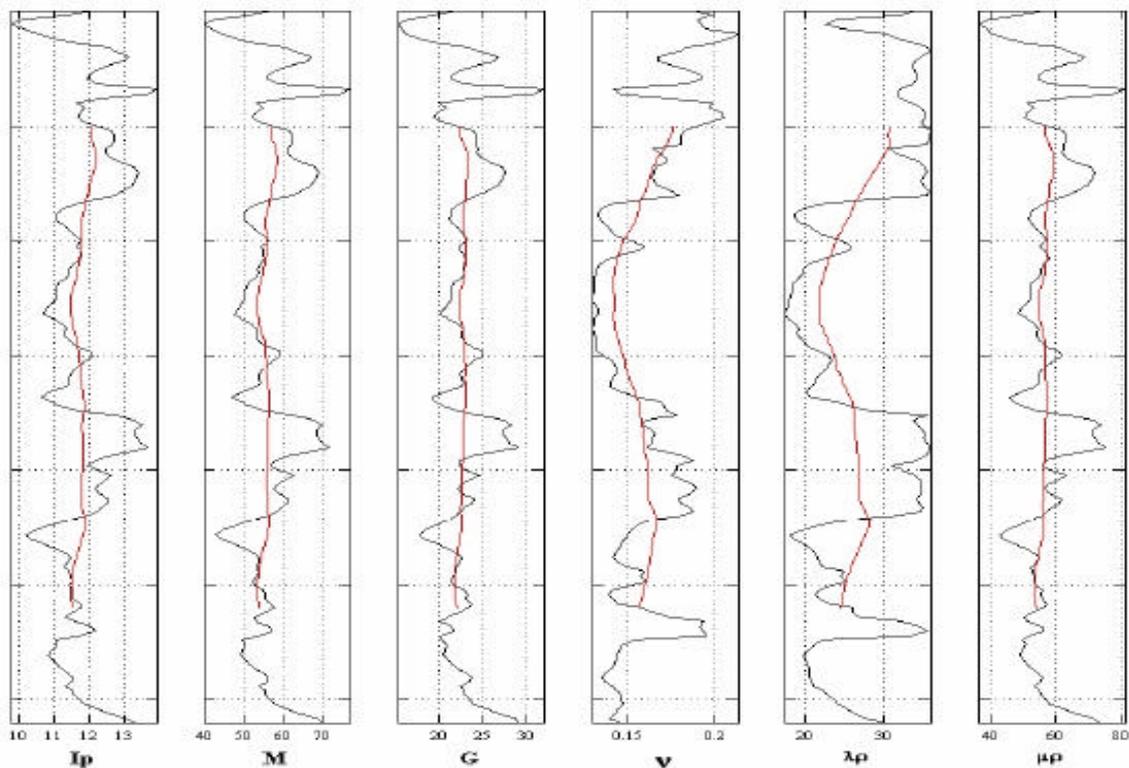


Fig 2.5 – Propriedades petrofísicas na escala de perfis de poços (linha preta) e após a aplicação do filtro de Backus para a realização do upscaling dos dados (linha vermelha). Da esquerda para direita os perfis são: impedância P, Módulo de cisalhamento, módulo de bulk, lambda*densidade e módulo de cisalhamento*densidade. (Retirado de Rasolosovoahangry, 2002)

2.6 Conclusões

Neste capítulo foi apresentado o conceito de modelo efetivo do meio e como as propriedades litológicas influenciam as propriedades elásticas do meio poroso. Os modelos de física de rocha relevantes a este trabalho foram apresentados, com ênfase nas relações entre V_p , V_s e densidade e das influências da porosidade, volume de argila e saturação. O conceito de “upscaling” e o filtro de Backus também foram abordados.

Capítulo 3

TEORIA DE DECISÃO BAYESIANA

3.1 Introdução

A teoria de decisão bayesiana é uma ferramenta que auxilia o processo de escolha, faz uso de probabilidades para a escolha de hipóteses e quantifica o risco existente em tomar uma decisão.

O reconhecimento de padrões probabilísticos cria regras de classificação utilizando processos estocásticos que transcrevem nosso conhecimento adquirido através de funções de distribuições de probabilidades e utilizam o Teorema de Bayes para a tomada de decisão. Segundo Devroye et. al. (1996), discriminação ou reconhecimento pode tratar de quantidades discretas ou contínuas para categorizar observações com características comuns, tendo como objetivo estimar ou prever a natureza desconhecida de uma observação.

Observação é a coleção de medidas numéricas gerada a partir de um observador ou equipamento, sendo a natureza desconhecida de uma observação x , denominada de w (categoria ou classe). Em discriminação busca-se a função $g(x)$ que mapeie o espaço paramétrico n -dimensional e representa o conhecimento sobre

uma classe, dadas as observações. O mapeamento g do espaço é denominado classificador ou regra de classificação e sua forma varia de acordo com o problema em questão. Comumente, especialistas trabalham com um classificador interno e pessoal baseado em conhecimentos adquiridos através de suas experiências pessoais (ex.: um geólogo intérprete indentificando litofácies utilizando os perfis de poço).

A magnitude e a complexidade de x pode desafiar até o melhor especialista, por isso, descrever g de forma simplesmente matemática dificilmente será à prova de falhas; por tanto, deve-se estar preparado para lidar com as imperfeições dos classificadores. Para a tomada de decisão, é necessário, quantificar a imperfeição de nossa regra quando da observação de x , $g(x) \neq w$, analisando a classificação da categoria juntamente com a medida de incerteza.

3.2 Teoria de decisão – caso de duas classes

Suponha a tarefa de escolher a qual categoria um objeto x pertence, existindo como opção as categorias w_1 ou w_2 . Expressa-se como $P(w_1)$ como a probabilidade a priori de w_1 e $P(w_2)$ como a probabilidade a priori de w_2 , antes de qualquer observação sobre x . Na ausência de conhecimento prévio sobre w_1 ou w_2 , a única forma de definir $P(w_1)$ e $P(w_2)$ é torná-los equiprováveis, $P(w_1) = P(w_2)$ e caso não exista nenhuma outra opção de escolha, temos que $P(w_1) + P(w_2) = 1$. Como pode-se observar, neste caso, uma tomada de decisão teria um alto grau de incerteza, pois o erro ao classificar x como w_1 seria $P(\text{erro } w_1) = 1 - P(w_1) = 0.5$, da mesma forma que $P(\text{erro } w_2) = 1 - P(w_2) = 0.5$.

Para melhorar o poder de decisão, deve-se fazer uso de observações $\{x_1, \dots, x_2, \dots, x_n\}$ que podem existir através de experimentos ou medições sobre w_1 e w_2 e definirmos $P(x|w_l)$, $l = 1, \dots, 2$, como a *probabilidade condicional* de X dado w_1 e w_2 . Supondo que se conhece $P(w_l)$ e $P(x|w_l)$ para $l = 1, 2$, faz-se uso do teorema de Bayes para aumentar o conhecimento sobre x .

Define-se $P(w_l|x)$ como a probabilidade a posteriori de w_l dado um conjunto de observações $X = \{x_1, x_2, \dots, x_n\}$, assim, temos então:

$$P(\mathbf{w}_l | x) = \frac{P(x | \mathbf{w}_l)P(\mathbf{w}_l)}{P(x)}, \quad (3.1)$$

$$P(x) = \sum_{l=1}^2 P(x | \mathbf{w}_l)P(\mathbf{w}_l).$$

A expressão acima é conhecida como o teorema de Bayes. Assim, o critério de decisão pode ser definido como:

$$\text{Decidir} \begin{cases} \mathbf{w}_1 \text{ caso } P(\mathbf{w}_1 | x) > P(\mathbf{w}_2 | x) \\ \mathbf{w}_2 \text{ caso } P(\mathbf{w}_2 | x) > P(\mathbf{w}_1 | x). \end{cases} \quad (3.2)$$

E a probabilidade de erro da decisão:

$$P(e) = \begin{cases} P(\mathbf{w}_1 | x) \text{ caso } \mathbf{w}_2 \text{ escolhido} \\ P(\mathbf{w}_2 | x) \text{ caso } \mathbf{w}_1 \text{ escolhido} \end{cases} \quad (3.3)$$

E a média de erro (ou risco médio) é dada por:

$$P(e) = \int_{-\infty}^{\infty} p(e | x)p(x)dx. \quad (3.4)$$

3.2.1 Estendendo o problema para mais de duas classes

Denomina-se $\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_k\}$ de um conjunto de possíveis classes ou categorias. Define-se $\{a_1, a_2, \dots, a_n\}$ um conjunto finito de ações ou escolhas a serem tomadas e $?(a_i | \mathbf{w}_l)$ como a perda envolvida no caso de tomar-se uma decisão a_i quando o verdadeiro estado da natureza for \mathbf{w}_l e X um vetor de parâmetros (atributos) de tamanho n .

Suponha que x é observado e uma ação a_i seja tomada. Se o verdadeiro estado da natureza for \mathbf{w}_l , $i \neq l$ ocasionará uma perda $?(a_i | \mathbf{w}_l)$ e o risco condicional em tomar a decisão a_i será:

$$R(\mathbf{a}_i | X) = \sum_{l \neq i}^k \mathbf{I}(\mathbf{a}_i | \mathbf{w}_l) P(\mathbf{w}_l | X) \quad (3.5)$$

O que se busca é uma regra de decisão que minimize o risco total:

$$R = \int R(\mathbf{a}(x) | x) P(x) dx, \quad (3.6)$$

sendo que a minimização do risco total é alcançada pela seleção da ação a_i onde $R(a_i | x)$ seja mínimo. Caso uma ação a_i seja tomada e o verdadeiro estado da natureza seja w_l , então a ação ou decisão está correta se $i = l$ e errada caso $i \neq l$. Definindo a função de perda como do tipo *zero-um*, tem-se:

$$\mathbf{I}(\mathbf{a}_i | \mathbf{w}_l) = \begin{cases} 0 & \text{caso } i = l \\ 1 & \text{caso } i \neq l, \end{cases} \quad (3.7)$$

e o risco condicional torna-se:

$$\begin{aligned} R(\mathbf{a}_i | X) &= \sum_{l \neq i}^l \mathbf{I}(\mathbf{a}_i | \mathbf{w}_l) P(\mathbf{w}_l | X) \\ &= \sum_{l \neq i} P(\mathbf{w}_l | X) \\ &= 1 - P(\mathbf{w}_i | X). \end{aligned} \quad (3.8)$$

Para a minimização do risco, maximiza-se $P(\mathbf{w}_i | X)$, tornando a regra de decisão:

decidir w_i caso $P(\mathbf{w}_i | X) > P(\mathbf{w}_1 | X)$, o que é equivalente a:

$$g(x) = \arg \max P(\mathbf{w}_i) P(x | \mathbf{w}_i), \quad (3.9)$$

ou a:

$$g(x) = \arg \min P(\mathbf{w}_l)P(x | \mathbf{w}_l) \quad (3.10)$$

O termo $P(x)$, do teorema de Bayes, pode ser desconsiderado por ser uma constante de normalização, não influenciando na tomada de decisão. Um classificador que faça uso da expressão (3.9) é conhecido como *classificador de máxima probabilidade a posteriori* (MAP).

3.3 Rede Bayesiana

Rede Bayesiana é a forma gráfica de representar um problema composto de decisão, baseado na teoria de decisão bayesiana. Consiste em um grafo orientado em que cada nó é identificado com informações quantitativas de probabilidade (Russell e Norvig, 2004).

A rede bayesiana é especificada da seguinte forma:

- I. Um conjunto de variáveis aleatórias constitui os nós da rede. As variáveis podem ser discretas ou contínuas.
- II. Um conjunto de vínculos orientados ou setas conecta os pares de nós. Se houver uma seta do nó A para o B, A será denominado pai de B.
- III. Cada nó possui uma distribuição de probabilidade condicional $P(A|B)$, que quantifica o efeito dos pais sobre o nó.

A topologia da rede especifica os relacionamentos de independência condicional que são válidos no domínio. Como exemplo, é idealizada a seguinte tarefa; i) especificar qual a probabilidade de uma amostra laboratorial ser um arenito, baseados em dois atributos distintos, Os atributos são: a) impedância P e b) impedância S.

Assume-se a independência condicional entre I_p e I_s e a existência de informação a priori.

A Figura 3.1 ilustra a topologia da rede.

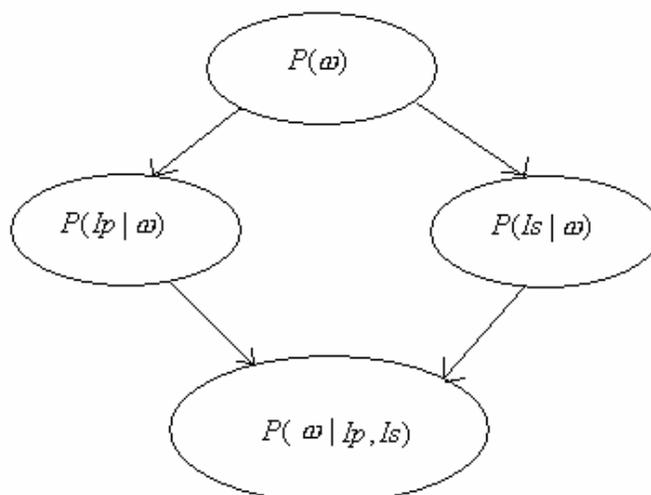


Fig 3.1 – Rede Bayesiana para o cálculo da probabilidade de uma amostra ser arenito, baseado em I_p , I_s e informação a priori

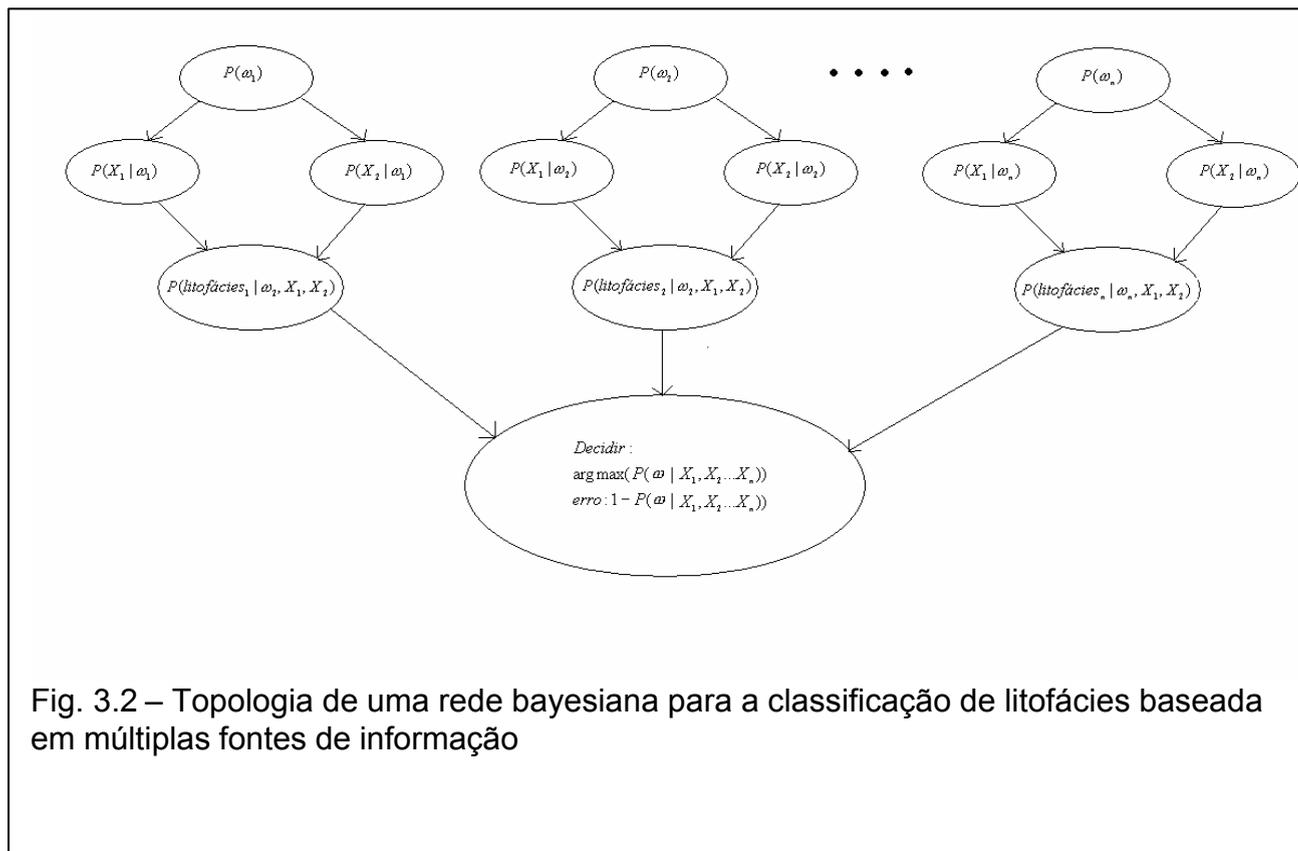
Pode-se interpretar o resultado da rede bayesiana acima, da probabilidade de ser arenito, como a probabilidade conjunta total de todos os nós da rede, ou seja, a probabilidade conjunta de I_p , I_s e informação a priori.

Na rede acima, $P(\omega)$ é uma informação totalmente independente das demais, já a probabilidade de ser arenito é dependente de todas as fontes de informação.

As redes bayesiana fornecem uma estrutura versátil para a incorporação de novas fontes de informação que auxiliem a tarefa de classificação. Para incorporar novas informações à rede, basta criar um novo nó e avaliar a relação de dependência entre os outros nós.

Para que a rede bayesiana acima se torne um classificador $g(x)$ é necessário somente um pequeno ajuste em sua topologia, bastando replicar a estrutura acima para o número de classes existentes no problema e relacionar o último nó de cada estrutura à regra de classificação, eq. (3.9).

A Figura 3.2 ilustra a topologia de uma rede bayesiana para a classificação de litofácies baseada em múltiplas fontes de informação.



3.4 – Classificação por k-vizinhos mais próximos

O conceito de k-vizinhos é baseado na relação em que categorias semelhantes sempre estarão mais próximas entre si do que qualquer outra, e utiliza a distância euclidiana como métrica para definir tal proximidade.

A Distância Euclidiana definida por:

$$d(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}, \quad (3.11)$$

Onde p são pontos de dados do espaço paramétrico, e x e y suas coordenadas.

- I. Inicialmente, calcula-se a distância entre todos os padrões de treinamento;
- II. Verifica-se a quais classes pertencem os k padrões mais próximos;
- III. A classificação é feita associando-se o padrão de teste à classe que for mais freqüente entre os k padrões mais próximos.

- IV. Caso a resposta desejada seja em termos de probabilidade, divide-se o número de k-vizinhos w_i por k total.

O método de classificação por k-vizinhos na prática apresenta bons resultados de classificação, embora demonstre grande desvantagem no quesito performance.

Para que seja avaliada a distância entre a amostra e os k-vizinhos mais próximos, é necessário o cálculo de todas as distâncias entre a amostra e o conjunto de treinamento, ocasionando um consumo de memória e processamento computacional proporcional ao tamanho do dado de treinamento.

Outro problema existente, consiste na escolha ideal do valor de k. Um k muito pequeno não conterá quaisquer pontos de dado; se for muito grande, poderá incluir todos os padrões de treinamento (Russell e Norvig, 2004).

A classificação por kvizinhos mais próximos não é considerada um método bayesiano, e sim um método baseado em instâncias. Devido aos bons resultados de classificação fornecido pelo método de k-vizinhos mais próximos, para este trabalho foi incorporado sua funcionalidade a Rede Bayesiana.

3.5 – Classificador robusto baseado em redes Bayesiana

Nesta seção é incorporada a técnica de k-vizinhos mais próximos a uma Rede Bayesiana de decisão. Espera-se que com a introdução de um novo método de classificação à estrutura da rede, forneça-se uma resposta final que incorpore os resultados dos dois métodos simultaneamente. Isto é facilmente alcançado devido a facilidade de adaptação das redes bayesianas.

Novamente será alterada a estrutura da rede bayesiana, acrescentando mais um nó que represente a probabilidade de uma categoria, dada a classificação por k-vizinhos mais próximos. O resultado final é a

combinação de dois métodos em uma única estrutura com o objetivo de classificar de litofácies.

A Figura 3.3 representa a estrutura da rede utilizada neste trabalho.

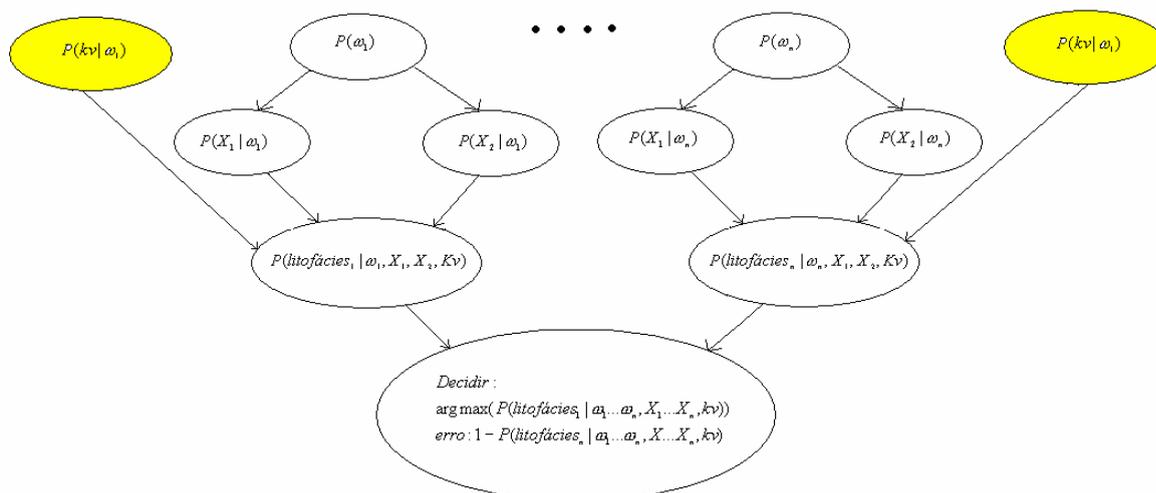


Fig 3.3 – Estrutura da rede bayesiana utilizada para a classificação de litofácies. O classificador gerado pela rede é uma combinação da decisão bayesiana e o método de k-vizinhos mais próximos. O nó em amarelo representa a inclusão do método de k-vizinhos mais próximos.

3.6 – Processo de classificação de litofácies

Foi desenvolvido o processo de classificação de litofácies com o uso da rede bayesiana, conforme apresentado na seção anterior. Supondo-se que as observações a serem classificadas sejam provenientes de seções de atributos sísmicos (ex. V_p , V_s , I_p , I_s ...), no qual serão denominados de dados de entrada. Para fins de exemplo, considera-se que existam na área apenas duas litofácies: litofácies1 e litofácies2.

O processo de classificação consiste em quatro etapas:

- I) Gerar funções de distribuição de probabilidade (pdfs) a partir de dados de treinamento (considera-se dado de treinamento, perfis de poços interpretados com as litofácies definidas). Será considerada ausência de informação a priori sobre cada litofácies. Consequentemente, as distribuições a priori $P(w_i) \quad i = 1,2$, serão uniformes e de valor 0.5;
- II) Pressuposto o dado de treinamento, geram-se as distribuições de cada classe dados seus atributos $P(Vp|w_i)$ e $P(Vs|w_i)$;
- III) Observando uma amostra, é avaliada a probabilidade de ser uma determinada litofácies, com o uso do método de k-vizinhos mais próximos, a informação a priori e as distribuições de probabilidades de cada litofácies;
- IV) Classifica-se em litofácies1 ou litofácies2, segundo o critério de máxima probabilidade a posteriori e associada a decisão, uma medida de incerteza é fornecida.

3.7 - Conclusões

Neste capítulo, demonstrou-se como utilizar redes bayesianas para construir o classificador de máxima probabilidade a posteriori para a classificação de litofácies. Sendo a estrutura final do classificador a combinação de dois métodos de classificação, que avaliam simultaneamente a probabilidade de uma amostra pertencer a um determinado tipo litológico - baseado em múltiplas fontes de informação. Tal estrutura de classificação pode ser aplicada para classificar volumes sísmicos, otimizando o tempo total de interpretação.

Para que seja possível o uso do classificador, é necessário um conjunto de dados de treinamento para gerar as distribuições de probabilidade, estes dados são perfis de poços que foram interpretados previamente por um geólogo intérprete.

No próximo capítulo, discuti-se formas convencionais de interpretação de litofácies em poços perfilados. Como é desejado da metodologia que seja a mais automatizada possível, é proposta a interpretação de perfis utilizando o algoritmo EM - técnica de aprendizado não supervisionado para a identificação automática de litofácies.

Capítulo 4

Identificando fácies litológicas em poços

4.1 – Introdução

Neste capítulo são discutidas as formas tradicionais de identificação litológica utilizando perfis de poços. Apresenta-se o funcionamento básico dos perfis de raio gama, nêutron, sônicos e densidade. É apresentada uma forma alternativa de identificação litológica utilizando técnica de agrupamento não supervisionado (algoritmo EM) que automatiza o processo de interpretação.

4.2 – Formas tradicionais de identificação de litofácies através de perfis

Os perfis de poços são comumente utilizados para a interpretação geológica de uma área. O seu menor custo em relação aos testemunhos, aliado ao fato de fornecer informações de forma contínuas e de qualidade uniforme, fazem com que seu uso seja extremamente atraente. Cada perfil responde de uma forma específica a um determinado tipo de rocha. A interpretação de litologias é realizada combinando as respostas dos perfis aliadas com o conhecimento do intérprete para a área de estudo.

4.2.1 – Perfis tipicamente utilizados para a determinação de litofácies

Apresenta-se abaixo o funcionamento básico dos perfis tipicamente utilizados para a determinação de litologias:

I – Raio gama: Segundo Brock (1986) esta é uma medida da radiação natural da formação. Os elementos radioativos geralmente concentram-se nas argilas. Enquanto folhelhos e areias argilosas apresentam níveis radioativos mais elevados, areias limpas e rochas carbonáticas apresentam baixos níveis de radiação. Devido a estas propriedades o uso do perfil de raio gama pode ser utilizado para diferenciar entre rochas reservatórias em potencial e folhelhos. Uma outra aplicação do perfil de raio gama é o cálculo de volume de argila (V_{sh}) como expressado abaixo:

$$GR_i = \frac{GR - GR_{sand}}{GR_{shale} - GR_{sand}} \quad (4.1)$$

e

$$V_{sh} = \frac{\rho_b \times GR_i}{\rho_{shale}} \quad (4.2)$$

onde:

Gr_i = índice de raio gama,

GR = Valor de raio gama registrado pelo equipamento,

Gr_{sand} = Valor de raio gama registrado na zona de areia mais limpa,

Gr_{shale} = Valor de raio gama registrado na zona de folhelho,

ρ_b = densidade total,

ρ_{shale} = densidade da argila,

II – Perfil nêutron

A ferramenta de nêutron utiliza uma fonte radioativa como Américo-Berílio para bombardear a formação com uma partícula neutra de elevada energia (nêutrons). Ao atingir os átomos tanto da formação como do fluido os nêutrons perdem energia. A quantidade de energia perdida está relacionada com o tamanho do átomo, quanto menor o átomo maior a energia transferida na colisão. Os nêutrons

são quase do mesmo tamanho do núcleo do átomo de hidrogênio. Quanto maior o número de hidrogênio nos poros da rocha, maior o número de colisões, fazendo com que os nêutrons percam energia e sejam capturados, em consequência a contagem no receptor é reduzida. Os poros das rochas encontram-se normalmente com água, óleo ou gás, todos contêm hidrogênio. Água e óleo possuem quase a mesma quantidade de hidrogênio, não podendo ser diferenciados através deste perfil, enquanto o gás possui uma concentração bem menor. Após a calibração da ferramenta para uma litologia de referência, geralmente arenitos ou calcários, é possível, através da contagem de hidrogênio presente nos poros da rocha, estimar a porosidade da rocha. Nas argilas, devido a grande quantidade de água adsorvida, este perfil fornece valores superestimados da porosidade.

III – Perfil Densidade

O princípio da medida do perfil de densidade baseia-se em um receptor e uma fonte de raios gama que bombardeiam a formação. Raios gama são partículas de alta velocidade que ao colidirem com um átomo podem perder energia (Espalhamento Compton) e expulsar elétrons. Neste processo o raio gama muda sua trajetória, e esta interação é registrada no receptor. A densidade total está relacionada com a densidade do elétron pela expressão:

$$\rho_e = 2\rho_b \frac{\sum Z}{M}, \quad (4.3)$$

onde ρ_e , Z e M são respectivamente as densidades do elétron, $\sum Z$ o somatório do número atômico e M o peso molecular do composto. Desta relação é possível obter a densidade da formação tendo o resultado da contagem do receptor.

IV – Perfis sônicos

A velocidade e a atenuação de ondas acústicas nas rochas dependem de diversas características da rocha como porosidade, densidade e saturação. Quando uma onda percorre a formação ela sofre os efeitos destes fatores e ao medir essas

variações podem ser feitas análises quantitativas e qualitativas sobre a rocha. O perfil sônico é um registro do tempo necessário para uma onda acústica percorrer uma distância pré-definida através da formação. Este tempo de trânsito Δt é comumente medido em microssegundos por pé ($\mu\text{sec/ft}$) e é recíproco à velocidade de propagação:

$$\Delta t = \frac{10^6}{\text{Velocidade}}. \quad (4.4)$$

Uma utilização comum dos perfis sônicos é o cálculo da porosidade, quando a litologia é conhecida (expressão 2.12). O cálculo da porosidade através do perfil sônico somente produz bons resultados quando a rocha é monominerálica e saturada e ausente da presença de gás. Outro fator que influencia o cálculo da porosidade por este método é a presença de argila na rocha. Os perfis sônicos são necessários também quando deseja-se criar seções sísmicas sintéticas para integrar ou calibrar dados de poços com dados sísmicos.

IV - Porosidade

A porosidade pode existir de diversas formas: porosidade intergranular, que é o espaço livre entre grãos da rocha; porosidade intragranular, o espaço livre dentro dos grãos da rocha; porosidade de fratura: porosidade existente entre superfícies fraturadas; ou ainda a porosidade vulgular: porosidade formada por bolhas de gás no processo de resfriamento da rocha. Nas rochas sedimentares, a porosidade formada durante o processo deposicional é denominada porosidade primária, enquanto a porosidade gerada após o processo deposicional, causada por processos geoquímicos como a dissolução de grãos é denominada de porosidade secundária.

Segundo Hearst et al. (2001), a porosidade pode ser obtida de diversas formas, podem ser citados métodos laboratoriais, métodos óticos ou mais recente através de ressonância magnética nuclear. É comum obter a porosidade através de perfis convencionais como densidade, sônico e nêutron.

4.2.2 – Determinação de litofácies através da combinação de perfis

Pode-se considerar a resposta das ferramentas de densidade, nêutron e sônica como sendo função de porosidade e litologia:

$$\begin{aligned}\rho_b &= f(\phi, \text{lito logia...}), \\ \phi_n &= f(\phi, \text{lito logia...}), \\ \Delta t &= f(\phi, \text{lito logia...}),\end{aligned}\tag{4.5}$$

Então torna-se natural a combinação destas ferramentas duas a duas para eliminarmos os efeitos da porosidade e obtermos informação sobre a litologia. O geólogo intérprete faz uso de gráficos do tipo “crossplot” e análise direta dos perfis para a identificação de litofácies. Um dos crossplots mais populares entre os analistas de perfis é o *densidade versus nêutron*. Este gráfico pode ser utilizado para determinação da matriz da rocha reservatória quando assumido que a matriz é mono-mineral. Abaixo é apresentado um exemplo deste gráfico:

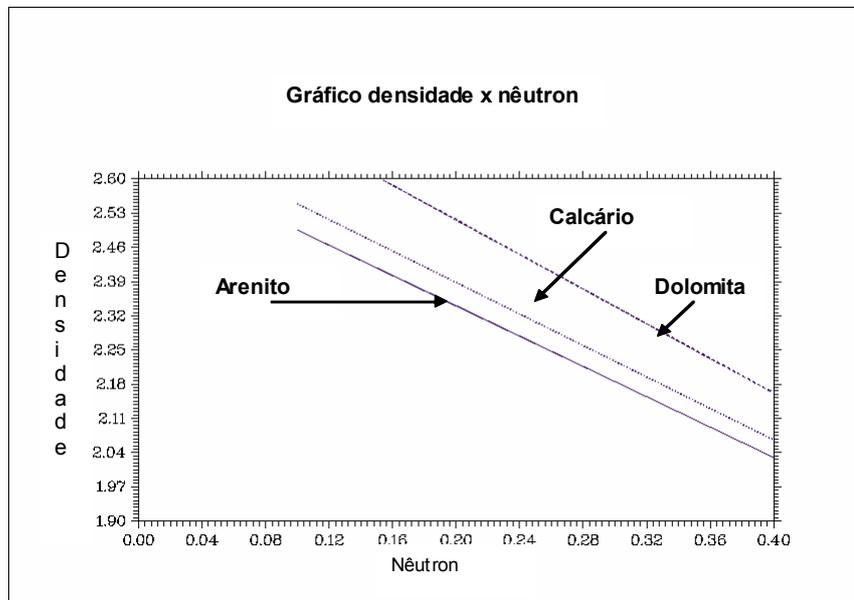


Fig 4.1 – Gráfico densidade X nêutron utilizada para discriminar possíveis rochas reservatórias

Alguns problemas de interpretação do gráfico acima podem ocorrer quando a formação contém pequenas quantidades de elementos que absorvem nêutron como boro. Wiley e Patchett (1990) examinaram um calcário onde as amostras localizavam entre a linha do calcário e a linha das dolomitas, porém através de análises laboratoriais a rocha mostrou-se uma calcita com 2% de argila e 10-47 ppm de boro. Esse tipo de discrepância afeta a interpretação não só deste tipo de gráfico, mas como qualquer análise gráfica. Conseqüentemente todas as análises gráficas, quando possível, devem ser validadas com a ajuda de amostra de calha e testemunhos para a determinação da litologia. Outra característica deste gráfico, é o fato de não podermos criar uma linha teórica que represente as argilas, devido a diferente composição mineral. Os diversos tipos de argila ocupam espaços distintos neste domínio. Devido a este fato faz-se uso previamente do perfil de raio gama para separar as rochas em duas categorias, rochas reservatórias e rochas não reservatórias.

Segundo Ushirobira (2004) um dos critérios mais utilizados para a determinação de litofácies é baseado no afastamento entre a curva de porosidade neutrônica com escala reversa e a curva de densidade. Fazendo a diferença entre as duas curvas, separa-se os arenitos e os folhelhos com o seguinte critério: caso o valor seja negativo, considera-se arenito, caso positivo folhelho. Combinando esse critério com outras informações, como o raio gama é refinada a classificação. As variações entre as interpretações vão depender da experiência do interprete.

Existem diversos métodos gráficos, além dos apresentados acima que são comuns na identificação de litologias. O objetivo da descrição apresentada até agora foi puramente introdutório para que possa ser apresentada uma nova técnica para a identificação de litofácies. Um estudo mais aprofundado sobre o assunto pode ser encontrado em Hearst et al. (2001).

4.3 – Definindo litofácies de forma automatizada com o uso do algoritmo EM

Acadêmicos e pesquisadores freqüentemente deparam-se com a necessidade de definirem grupos com características semelhantes entre dados

observados. Seja qual for o ramo de atuação da pesquisa, o agrupamento de amostras em classes distintas tem como objetivo fazer com que os indivíduos de um mesmo grupo sejam mais semelhantes entre si, do que de qualquer indivíduo de outro grupo, maximizando a homogeneidade de indivíduos de um mesmo grupo e maximizando a heterogeneidade entre grupos.

A análise de “clusters” é um ramo da estatística multivariada que define um conjunto de técnicas com este objetivo. O algoritmo EM (Expectation-Maximization) (Dampster, 1977) é uma técnica de agrupamento não supervisionado para distinguir várias categorias em uma coleção de amostras. O problema é não supervisionado, pois não conhecemos os “rótulos” ou nome de cada categoria.

4.3.1 –Algoritmo EM

O método EM assume que os dados observados (perfis de poços) estão distribuídos conforme uma mistura de funções de probabilidades (pdfs). Tratando-se de dados contínuos, multidimensionais, uma escolha natural de distribuição para o desenvolvimento são funções Gaussianas multivariadas, definidas por:

$$p(X) = \sum_{l=1}^C G(X | \mu_l, \Sigma_l) \omega_l, \quad (4.6)$$

onde $l = 1, \dots, C$, é o número de componentes da mistura, μ e Σ são respectivamente vetor de médias e matriz de covariância, ω_l o peso associado a cada componente. Denomina-se $\theta_C = \{\mu_1, \dots, \mu_C, \Sigma_1, \dots, \Sigma_C, \omega_1, \dots, \omega_C\}$ o conjunto de parâmetros da mistura. O objetivo do EM é de estimar θ_C baseado em um conjunto finito de amostras $\{x_1, \dots, x_2, \dots, x_n\}$, utilizando algum método de aprendizagem. O método de aprendizagem utilizado para estimar os parâmetros é conhecido como maximização do logaritmo da função verossimilhança $L(\theta_C)$ definido por:

$$L(\theta_C) = \ln \prod_{i=1}^n p(X_i) = \sum_{i=1}^n \ln \sum_{l=1}^C G(X_i | \mu_l, \Sigma_l) \omega_l \quad (4.7)$$

Pelo princípio da maximização do logaritmo da função verossimilhança, os parâmetros que maximizam a expressão (4.7) são encontrados igualando as derivadas parciais dos parâmetros de interesse a zero, desta forma obtém-se:

$$\mu_l = \frac{\sum_{i=1}^n \omega_i x_i}{\sum_{i=1}^n \omega_i}, \quad (4.8)$$

$$\Sigma_l = \frac{\sum_{i=1}^n \omega_i (x_i - \mu_l)(x_i - \mu_l)^T}{\sum_{i=1}^n \omega_i}, \quad (4.9)$$

$$\omega_l = \frac{\sum_{i=1}^n \omega_i}{n}, \quad l = 1, 2, \dots, C. \quad (4.10)$$

A demonstração analítica deste processo pode ser encontrada em Lu e Minas (2003). O algoritmo alterna em dois estágios. No primeiro, “Expectation” ou esperança os valores esperados ω_i são calculados para cada amostra, para todos os C componentes:

$$\omega_i = G_l(X_i | \mu_l, \Sigma_l) \omega_l. \quad (4.11)$$

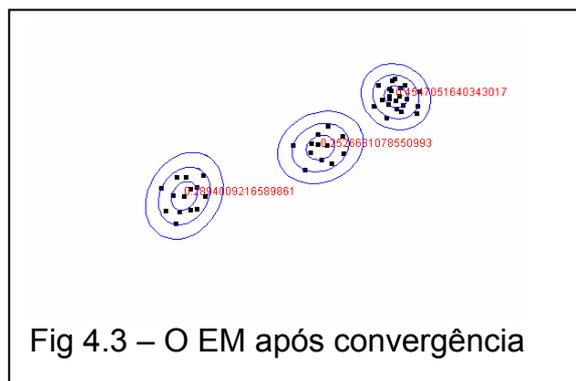
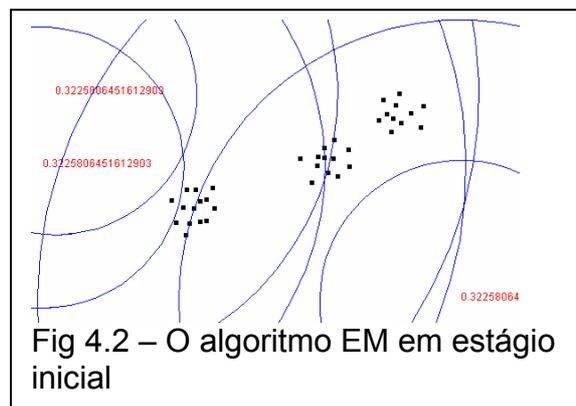
O segundo estágio, “maximization” ou maximização, consiste em maximizar os valores dos parâmetros conforme as expressões (4.8), (4.9) e (4.10). A princípio não são conhecidos os valores dos parâmetros $\theta_C = \{\mu_1, \dots, \mu_C, \Sigma_1, \dots, \Sigma_C, \omega_1, \dots, \omega_C\}$, portanto uma estimativa inicial deve ser fornecida para cada parâmetro. A estimativa inicial deve ser cautelosa, pois alguns problemas podem acontecer.

I – Pode acontecer de um componente, vir a cobrir uma única amostra, desta forma sua variância tenderá a zero e sua probabilidade a infinito.

II – Caso as estimativas iniciais dos parâmetros sejam idênticas ou muito próximas, as componentes podem se fundir.

O EM, a cada interação, aumenta a probabilidade dos dados em relação a cada componente e o critério de parada acontece quando o logaritmo da verossimilhança convergir ou quando atingir um número de interações previamente fixado.

As figuras 4.2 e 4.3 ilustram o funcionamento do EM em um domínio fictício. Existem três agrupamentos distintos que representam o comportamento de três litofácies distintas obtida por dois diferentes perfis de poços. A Figura 4.2 simula o EM em seu estágio inicial.



A Figura 4.3 é uma representação gráfica (em duas dimensões) do EM após a convergência. Desta forma, ao agrupar-se observações de perfis de poços devido as suas respostas comuns, rochas com mesmas características pertencerão ao mesmo

grupo, o que corresponde ao raciocínio de um intérprete de perfis, ao classificar litofácies com base em padrões de semelhança entre os dados.

4.3.2 - Identificação do número ótimo de componentes

O EM espera que o número de componentes C sejam previamente definidos, porém isto é uma informação não conhecida. Antes de aplicar o algoritmo não se tem a informação do número de litofácies existente nos dados. Para contornar esse problema executa-se o algoritmo para $\{C = 2, \dots, \text{número máximo de possíveis fácies}\}$ e aplica-se algum critério de escolha de modelos. O critério utilizado neste trabalho é conhecido como BIC - Bayesian Information Criteria (Schwarz, 1978), definido por:

$$BIC = -2L(\theta_c) + \ln(n)(C - 1) + C(d + (d + 1)/2), \quad (4.12)$$

onde d é a dimensão do e n o número de amostras. Denomina-se de dimensão, a quantidade de informações independentes utilizado pelo EM. Caso sejam utilizados três perfis distintos para a interpretação do poço, $d=3$.

O primeiro termo pode ser interpretado como uma medida de falta de ajuste entre distribuições, quando estimadores de máxima verossimilhança são utilizados. O segundo e terceiro termos, são medidas da complexidade do modelo que penalizam o primeiro, devido a tendência do mesmo favorecer modelos mais complexos (“overfitting”). O que se procura é o modelo que forneça o valor mínimo da expressão (4.12). A derivação da expressão (4.12) pode ser encontrada em Xuelei e Lei (2004).

A Figura 4.4 exemplifica os resultados obtidos ao aplicar o EM em um conjunto com o BIC. O mínimo do gráfico abaixo ($C=5$) representaria o número ideal de fácies que estariam bem definidas pelos dados de perfis.

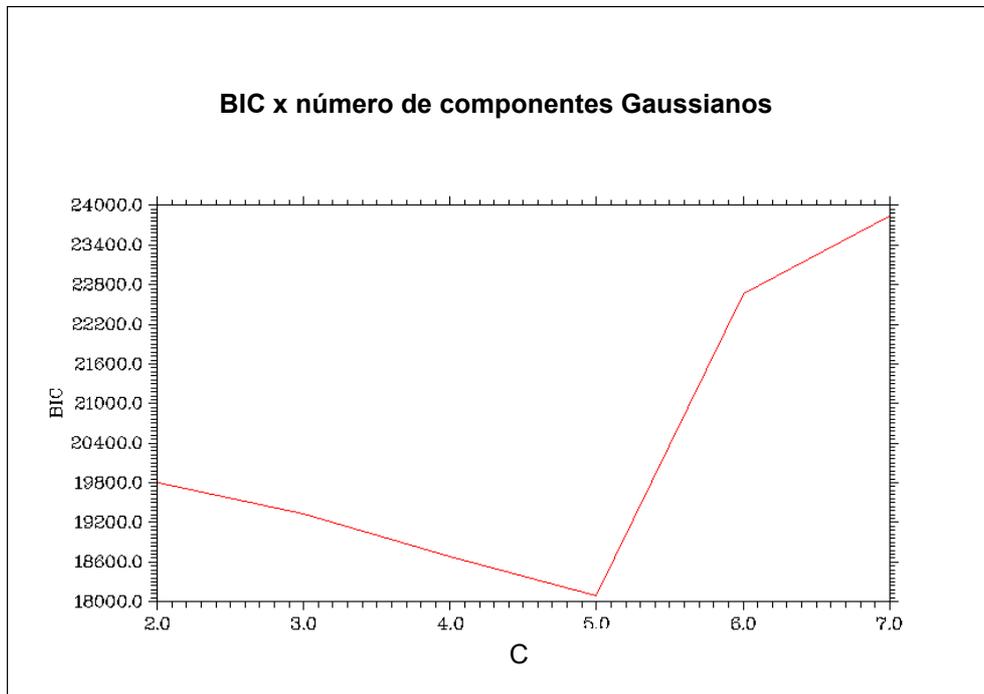
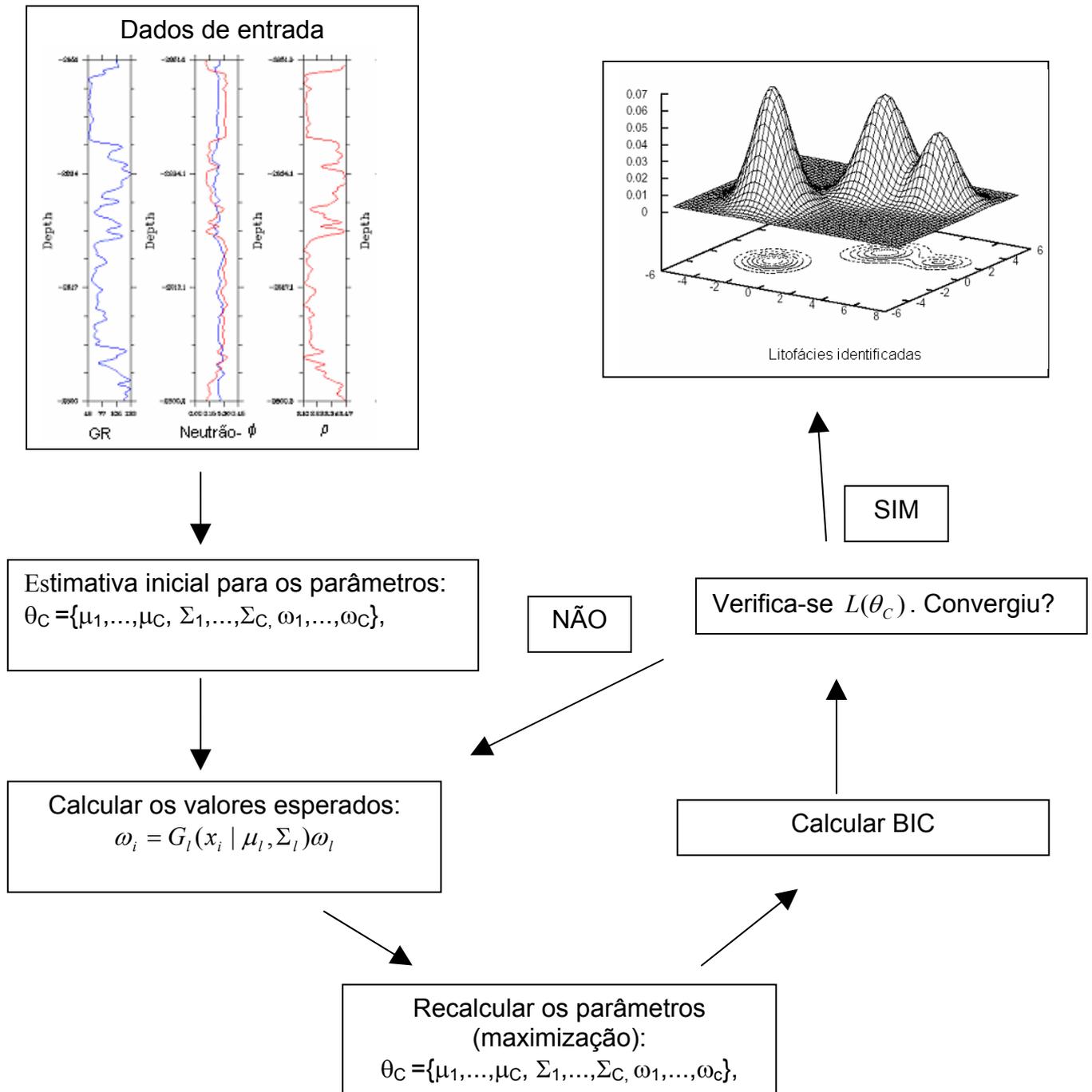


Fig. 4.4 - Exemplo do resultado do BIC obtidos após a execução do EM em dados de poço. O mínimo do gráfico indica o número ideal de litofácies definida pelo dado.

O mínimo neste gráfico ($C=5$) representa o número ideal de fácies que está bem definida pelos dados. Neste caso, tem-se cinco grupos distintos entre os dados e, conseqüentemente, cinco litofácies identificadas.

4.3.3 – Passo a passo do uso do Algoritmo EM para a identificação de litofácies

Nesta seção é apresentado através de um fluxograma, o “passo a passo” de como aplicar o E.M, em conjunto com o BIC, para a identificação automática de litofácies. Os dados de entrada para o E.M são os perfis de poços convencionalmente utilizados para a identificação de fácies em poços (perfis raio gama, nêutron, e densidade). Outros perfis quando disponíveis também podem ser utilizados.



O EM interage até que o logaritmo da verossimilhança convirja. Ao final, o que se obtém são os dados agrupados pela semelhança de seus atributos.

4.4 - Conclusão

Neste capítulo foi apresentada a forma convencional utilizada pela indústria do petróleo para a identificação de litofácies em poços. Uma breve descrição foi realizada sobre os perfis mais usuais, utilizados para tal tarefa. Foi apresentado o desenvolvimento teórico do algoritmo EM, um método que pode ser utilizado para a identificação automática de litofácies. Foi exposto, que deve-se fazer uso do EM em conjunto com o BIC para que o número ótimo de litofácies seja definido em função dos dados. A grande vantagem da utilização de mecanismos automáticos para a interpretação de poços é a grande economia de tempo em relação aos métodos convencionais.

Capítulo 5

Quantificando Incertezas

5.1 – Introdução

Este capítulo explora como quantificar e diminuir incertezas envolvidas no processo de classificação segundo a teoria de decisão Bayesiana. Incertezas existem em qualquer tarefa de reconhecimento de padrão. Conhecer e saber quantificar as fontes de incertezas garantem resultados realísticos sobre o objetivo, auxiliando a interpretação do resultado obtido.

Segundo Takahashi (2000) na geofísica de exploração surgem incertezas, devido a causas inevitáveis, tais como condições impróprias de levantamento, dados insuficientes comparados à heterogeneidade do alvo, largura de banda restrita da onda sísmica, conhecimentos insuficientes de física de rocha e várias outras casualidades. Uma forma simples de diminuir incertezas é aumentarmos a quantidade de informação sobre um determinado problema. A classificação de litofácies será feita tendo como dado de entrada, um conjunto de atributos sísmicos obtidos através de inversão de AVO. Para aumentar o nível de sucesso no processo de classificação, utiliza-se a maior quantidade de informação disponível para classificar novos dados.

A base da teoria de decisão Bayesiana é o uso de distribuições de probabilidades para tomada de decisão. A seguir é demonstrado como quantificar o risco em tomar uma ação e como isso influenciará no resultado final de classificação. Esta forma de quantificar incertezas será a chave para a escolha dos atributos sísmicos que fornecerão uma classificação com o menor risco possível.

5.2 – Erro de Bayes

O erro de bayes é uma medida de sobreposição entre pdf's. A Figura 5.1 ilustra duas distribuições de probabilidade, criadas a partir de dados sintéticos que serão detalhados no próximo capítulo. A pdf em vermelho representa a probabilidade $P(V_p | \text{litofácies}=\text{arenito})$ e a pdf em azul a probabilidade $P(v_p | \text{litofácies} = ! \text{arenito})$, o símbolo (!) representa uma negação. A área em amarelo representa o quanto as duas distribuições estão sobrepostas. Esta medida pode ser quantificada por:

$$p(\text{erro}) = \int_{V_p > V_{p_d}} p(\omega_1 | V_p) P(V_p) dV_p + \int_{V_p < V_{p_d}} p(\omega_2 | V_p) P(V_p) dV_p, \quad (5.1)$$

onde V_{p_d} é o valor de V_p no qual as duas distribuições se interceptam (aproximadamente em $V_p=2397$ m/s). O valor de V_{p_d} é considerado a fronteira de decisão entre classificar uma nova amostra como arenito ou não. Quanto maior a sobreposição entre as distribuições, maior a incerteza envolvida no processo de classificação. O erro de Bayes fornece uma medida de erro médio de classificação, esta medida será útil na escolha dos atributos que forneçam o melhor resultado de classificação.

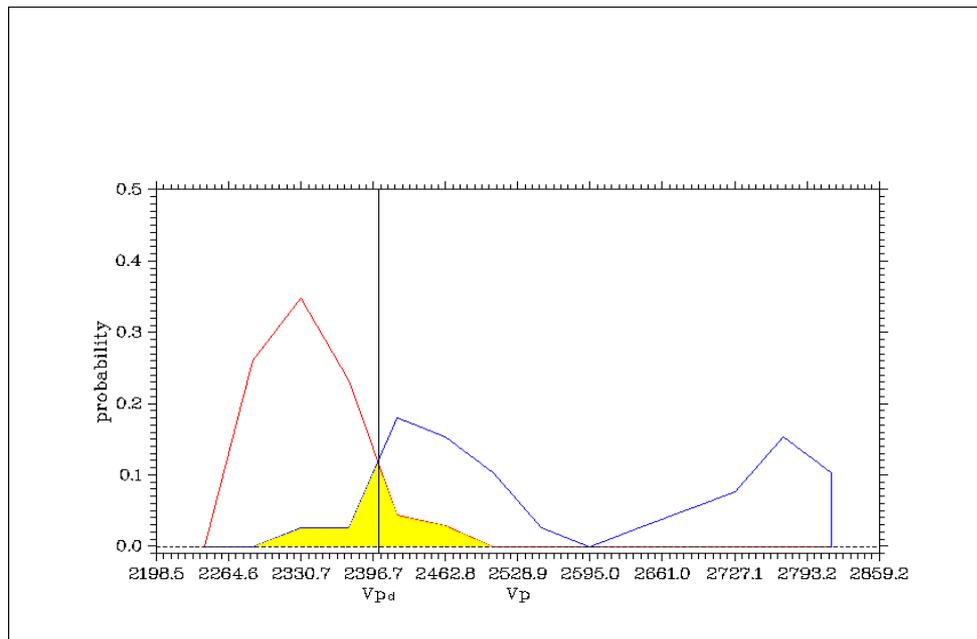
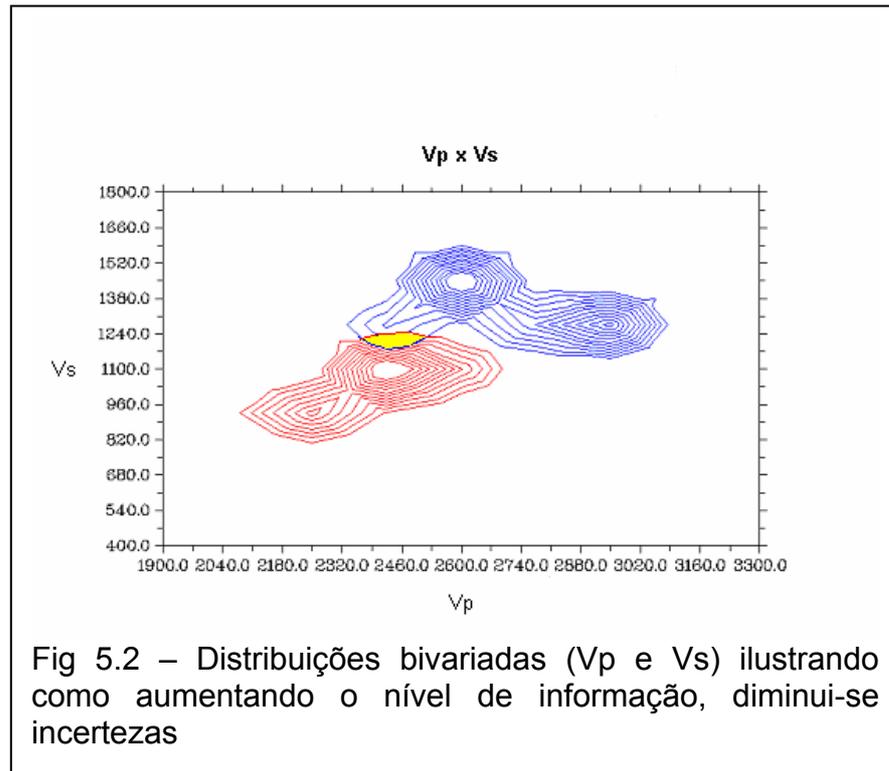


Fig 5.1 – A área sombreada representa a sobreposição entre as pdf's, conhecida como erro de Bayes. *Erro de Bayes*= 0.013430

5.3 – Reduzindo incertezas

Para a obtenção de resultados de classificação com o máximo de sucesso, é necessário reduzir as incertezas. Uma forma razoável de realizar tal tarefa é aumentar o nível de informação sobre nosso objetivo. A Figura 5.1 apresenta a relação de determinada litologia com V_p , contudo sabemos que não existe uma relação única de V_p com uma certa rocha. Diferentes rochas podem apresentar valores iguais ou bem próximos de V_p como o quartzo ($V_p=6.05$ m/s) e a anidrita ($v_p=6.01$) (Castagna et al.,1993), tornando impossível distinguir as duas rochas utilizando somente este atributo. Para reduzir essa incerteza, ou seja, aumentar o poder de discriminação, deve-se combinar outros atributos para melhorar o nível de informação.

A Figura 5.2 mostra duas distribuições com o uso de V_p e V_s . Novamente a pdf em vermelho representa a probabilidade $P(V_p, V_s | \text{litofácies}=\text{arenito})$ e a pdf em azul a probabilidade $P(V_p, V_s | \text{litofácies} \neq \text{arenito})$. Observa-se como a área de sobreposição entre distribuições é muito menor que a ilustrada pela Figura 5.1. A inserção de um novo atributo (V_s) ao problema de discriminação diminuiu o erro de Bayes, o que ocasiona melhora no poder de discriminação do classificador.

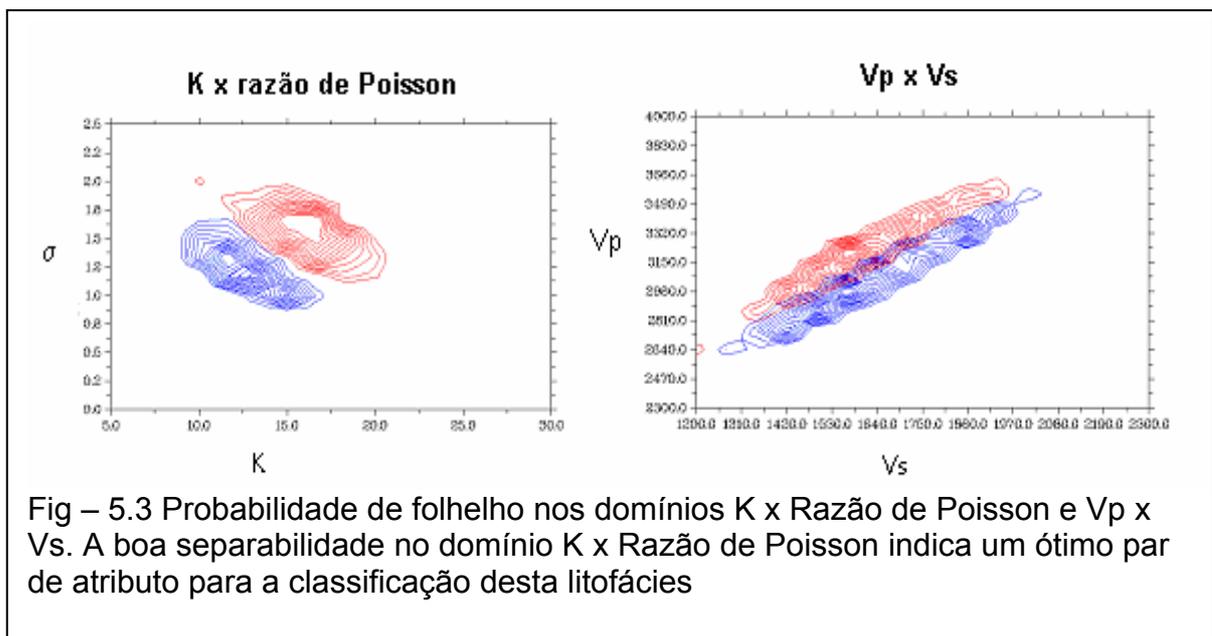


5.4 – Utilizando o erro de Bayes para a escolha dos atributos mais significativos

Quanto maior a sobreposição entre as funções de distribuição de probabilidade dos atributos de rochas distintas, menor o nível de informação existente para sua classificação; e quanto menor a sobreposição existente, melhor será o resultado da classificação. Sendo assim, para que sejam escolhidos os atributos que garantirão o melhor resultado, deve-se analisar combinações de atributos que melhor representem uma determinada rocha. Com a finalidade de escolher os atributos que viabilizarão a melhor classificação é realizado um teste que consiste em combinar todos os parâmetros que possam ser obtidos por inversão de AVO. A cada combinação, é calculado o erro de Bayes. Os atributos que possuírem o menor erro, serão o que fornecerão o menor risco de classificação e deverão ser utilizado como dado de entrada no classificador.

A Figura 5.3 ilustra como certos pares de atributos apresentam melhor separabilidade que outros, analisando-os para uma determinada litofácies. As

distribuições são de um folhelho de um reservatório turbidítico. A distribuição em vermelho, do lado esquerdo, representa a probabilidade de ser folhelho dados imcompressibilidade e razão de Poisson, $P(K, \sigma | \text{litofácies} = \text{folhelho})$; e a distribuição em azul a probabilidade de não ser folhelho, $P(K, \sigma | \text{litofácies} = ! \text{folhelho})$. Do lado direito tem-se $P(Vp, Vs | \text{litofácies} = \text{folhelho})$ em vermelho e $P(Vp, Vs | \text{litofácies} = ! \text{folhelho})$ em azul. Nota-se que no domínio K x Razão de Poisson as distribuições praticamente não se sobrepõem, indicando um ótimo par de atributos para o processo de classificação desta litofácies.



5.5 – Conclusão

Neste capítulo demonstrou-se como é possível usar o erro de Bayes como medida de incerteza no processo de classificação de litofácies. Também mostrou-se que para a obtenção da classificação com o menor risco possível, deve-se escolher os atributos utilizados no processo classificatório com cautela, sempre escolhendo os pares de atributos que tenham a maior separabilidade, ou seja, o par de atributo que tenha a maior quantidade de informação em relação a uma determinada litofácies.

Capítulo 6

Testes com dados sintéticos

6.1 – Introdução

Neste capítulo, aplica-se a metodologia apresentada neste trabalho a um conjunto de dados sintéticos. O teste tem a finalidade de avaliar a sensibilidade da metodologia à presença de ruídos nos dados, em condições controladas. São simuladas as etapas do processo de classificação de litofácies seguindo o formato idealizado para a realização desta tarefa. Na seção seguinte ilustra-se a metodologia de forma completa através de um fluxograma.

6.2 – Fluxograma da metodologia

A Figura 6.1 ilustra o fluxograma da metodologia proposta. Através dele pretende-se apresentar uma visão geral da metodologia. Como neste teste os dados utilizados como dado de entrada são sintéticos (dados que serão classificados) não existe a necessidade de realizar a etapa de “Análise para rotular as litofácies identificadas”, pois já é sabido quais são os diferentes rótulos (tipo litológico) que existem no conjunto de dados.

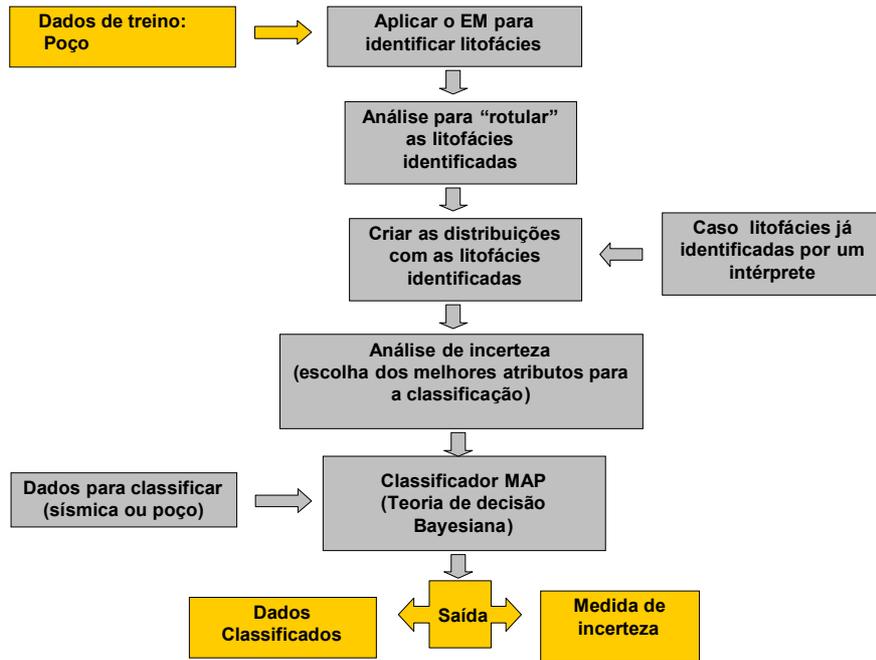


Fig 6.1 – Fluxograma da metodologia idealizada para a classificação de litofácies

6.3 Descrição do modelo sintético

O modelo gerado para a realização dos testes é formado por três litofácies diferentes, organizadas em um arranjo de quatro camadas conforme a Figura 6.2. O modelo consiste em seções sísmicas de 25 traços, cada traço com 110 observações. Os valores dos seus atributos foram retirados de Castagna et al.(1993) representando três litologias distintas conforme descrita na Tabela 4. Foram geradas seções dos atributos V_p , V_s e densidade, já que todos os outros atributos podem ser gerados a partir destes três. Criou-se a princípio um modelo de valores constantes (sem ruído) e, gradualmente, foram adicionados ruídos gaussianos de 5% e 10%. Para simular o poço foi retirada um traço aproximadamente no meio da seção (traço número 13), do modelo com 5% de ruído.

Tabela 4 – Valores dos atributos do modelo sintético.

Litofácies	Densidade(gm/cm ³)	Vp (Km/s)	Vs (Km/s)
1 - Arenito	2.02	2.44	1.30
2 - Arenito/folhelho interlaminado	2.20	2.75	1.20
3 - Folhelho	2.25	2.30	0.95

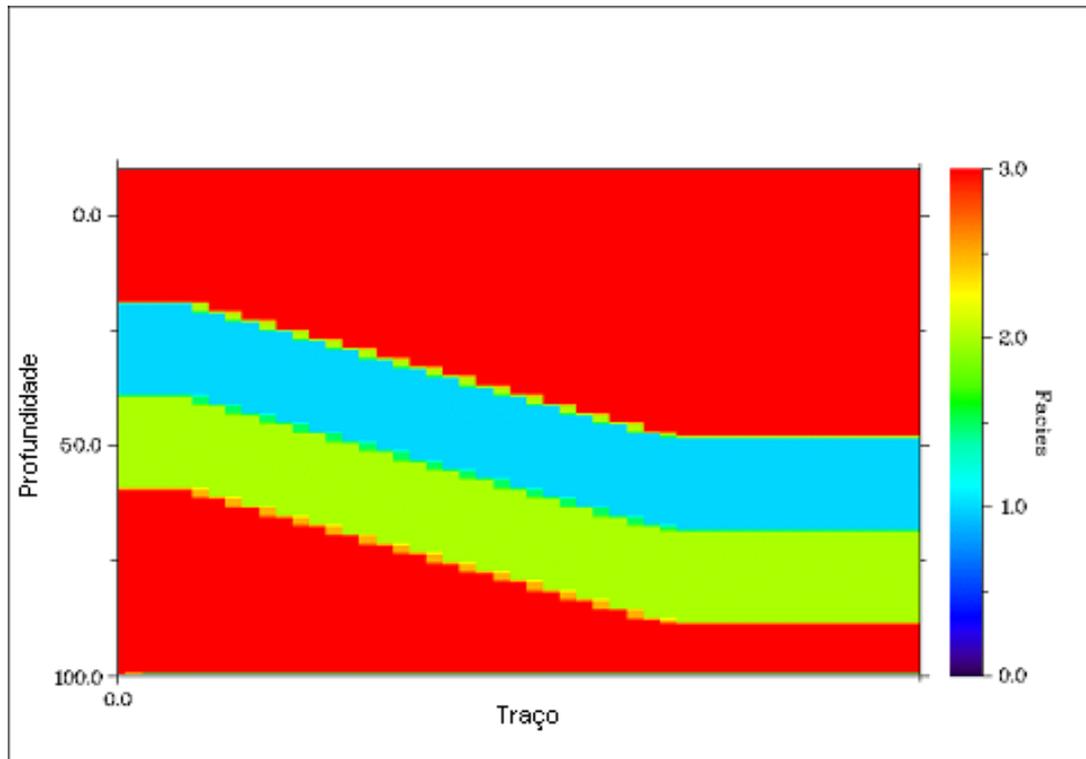
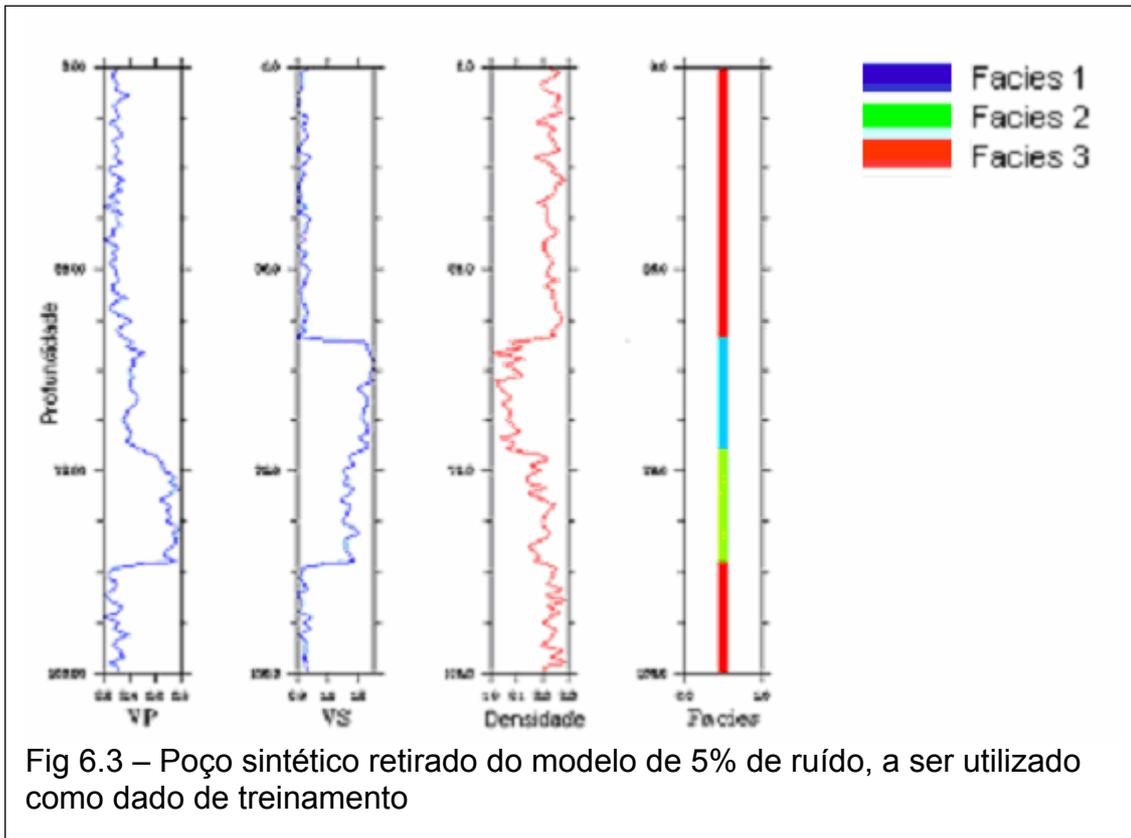
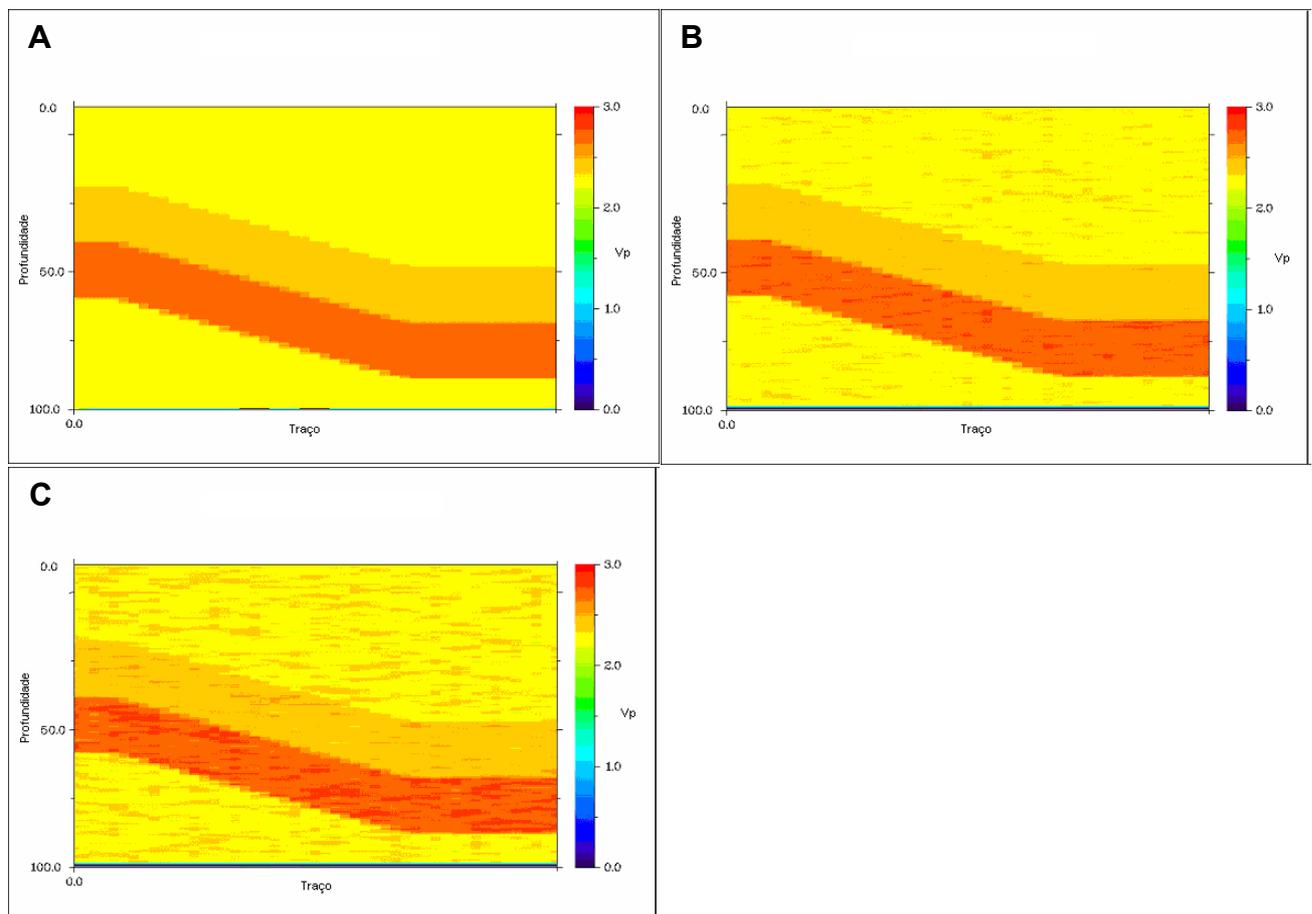


Fig 6.2 – Modelo sintético, sem ruído, de quatro camadas utilizado como referência para o teste.



As figuras 6.4 e 6.5 ilustram as seções de atributos para cada litofácies do modelo sintético, estas seções serão os dados que serão classificados.



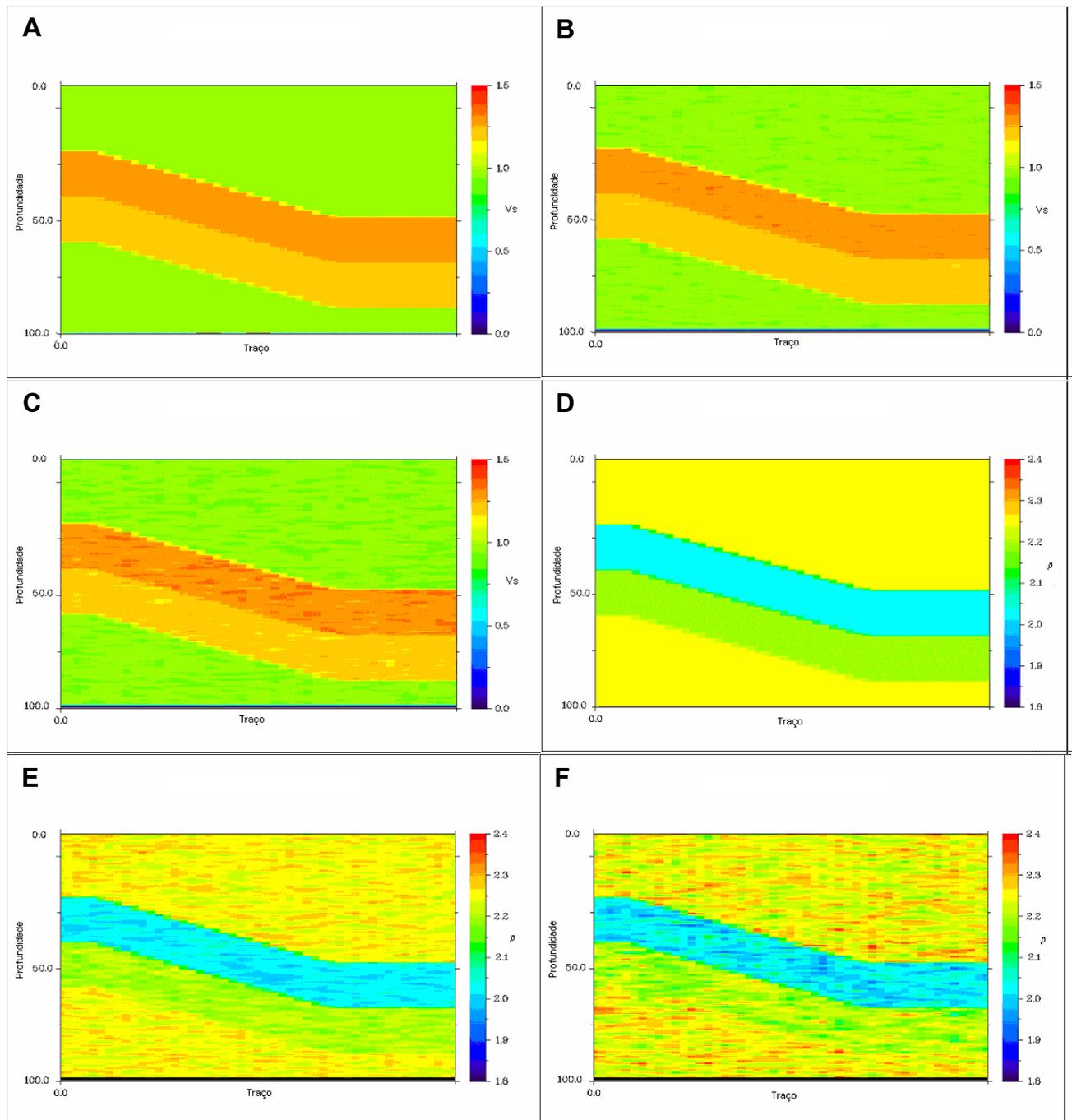


Fig 6.5 – Seções sintéticas de Vs (A,B e C) e densidade (D,E e F). As figuras A e D, B e E e C e F respectivamente foram geradas com ruídos de 0%, 5% e 10%.

6.4 – Teste do algoritmo EM com dados de poço sintético

Os testes para verificar a eficácia do algoritmo EM em identificar padrões distintos, foram realizados utilizando o poço sintético. A estimativa inicial dos parâmetros $\theta_3 = \{\mu_1, \dots, \mu_3, \Sigma_1, \dots, \Sigma_3, \omega_1, \dots, \omega_3\}$ foi selecionada conforme a seguinte regra:

- a) Para as médias μ_1, \dots, μ_3 foram escolhidos os valores centrais do modelo sintético conforme descrito na Tabela 4.
- b) Para a matriz de covariância, assume-se que os atributos são independentes, neste caso a matriz assume a forma de matriz diagonal, e as variâncias foram calculadas assumindo um desvio padrão de 30% do valor central e elevando-o ao quadrado.
- c) Para o cálculo do peso $\omega_1, \dots, \omega_3$, já que não existe nenhuma informação que faça acreditar que uma litofácies seja mais provável de acontecer do que outra, serão atribuídos pesos iguais para cada componente.

É esperado somente testar a estabilidade do EM à presença de ruídos, e como é sabido que o modelo possui três fácies distintas, foi fixado e somente executado o EM para $C=3$. Os ruídos gaussianos acrescentados são acumulativos, já que o poço sintético é proveniente do modelo de 5% de ruído.

A Tabela 5 resume o resultados do EM para cada teste realizado.

Tabela 5 – Resultado do EM aplicado ao poço sintético

Ruído no dado	Índice de acerto
0%	100%
5%	97,2%
10%	75,2%
15%	57,7%

As figuras abaixo ilustram os resultados do EM.

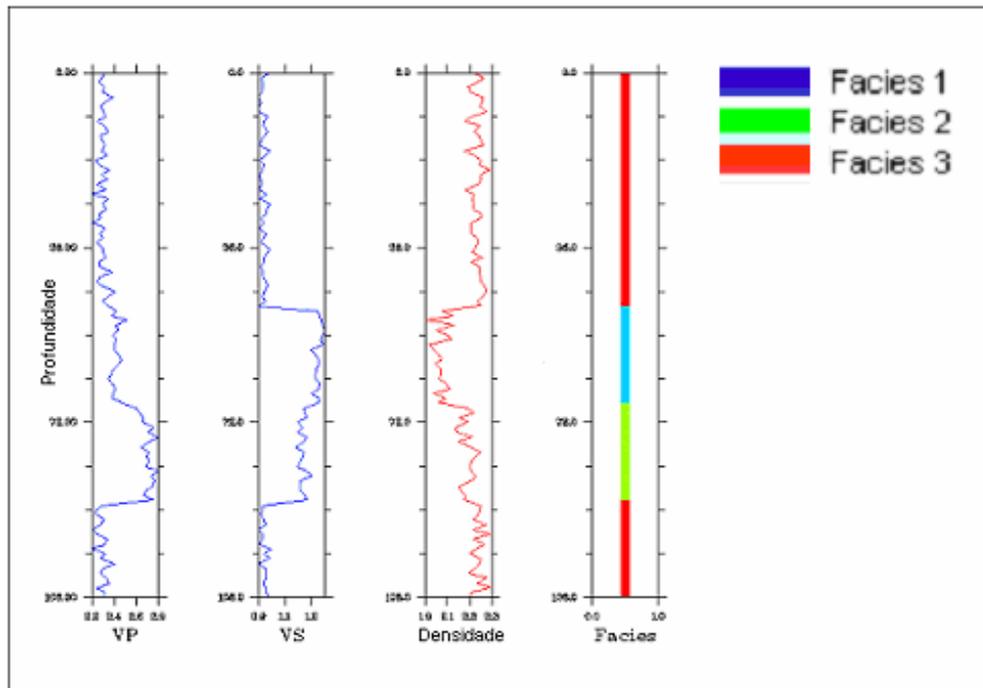


Fig 6.6 – Resultados do EM com poço sintético e ruído de 0%

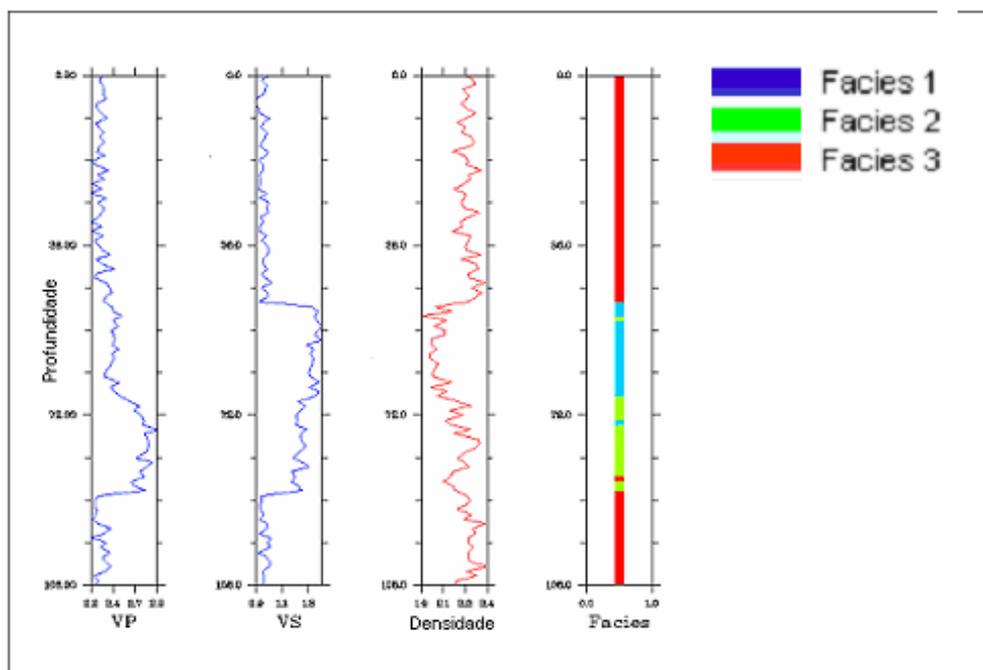


Fig 6.7 – Resultados do EM com poço sintético e ruído de 5%

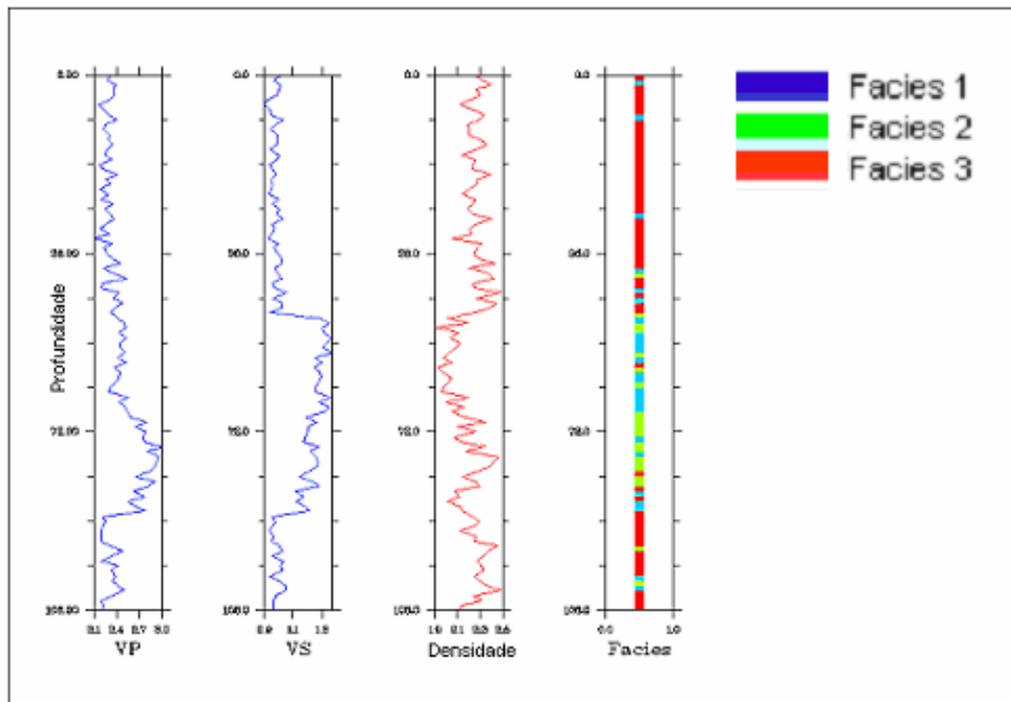


Fig 6.8 – Resultados do EM com poço sintético e ruído de 10%

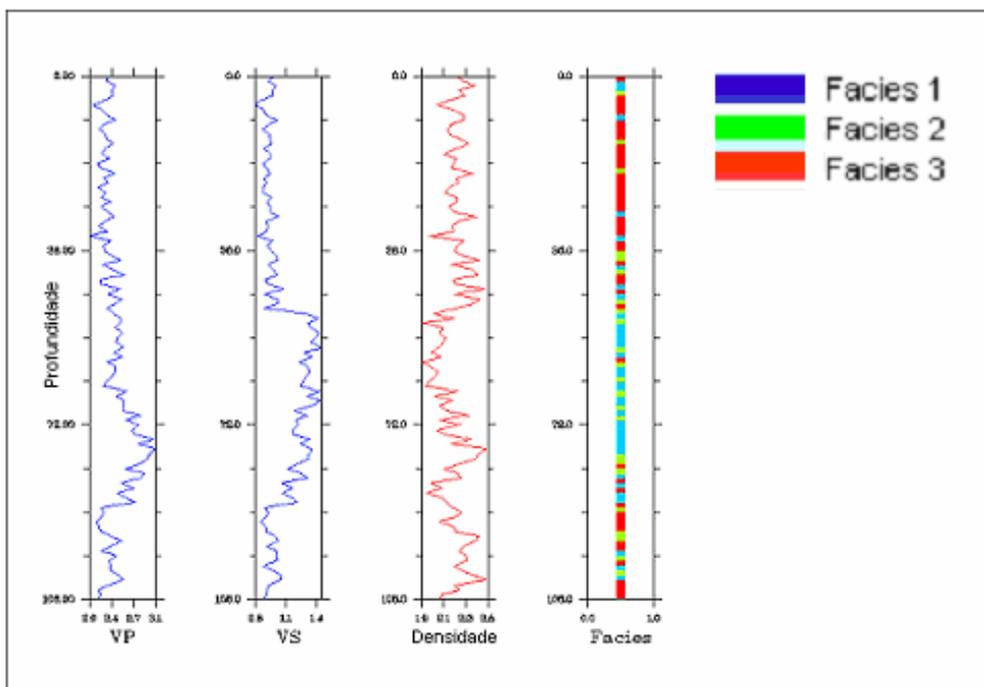


Fig 6.9 – Resultados do EM com poço sintético e ruído de 15%

O resultado do grupamento realizado pelo EM é a identificação das distintas litofácies existentes nos dados. As litofácies identificadas com o uso do dado de poço formaram o conjunto de dado de treinamento a ser utilizado pelo classificador MAP para a classificação do dado sísmico. A última coluna das figuras 6.6 à 6.9 é uma representação gráfica do número de amostras agrupadas erroneamente. Pode-se perceber que a qualidade do agrupamento é proporcional a quantidade de ruído no dado.

6.5 – Avaliação do Bayesian Information Criteria (BIC)

O BIC pode ser utilizado em conjunto com o EM como um indicativo de número ótimo de componente, quando não se sabe exatamente quantas litofácies existem nos dados. Para avaliar a capacidade do BIC em indicar o número ideal de componentes, foi executado o EM para $C=2, \dots, 6$ e a cada execução foi calculado o BIC. Como existem apenas três litofácies em nosso dado sintético, é esperado que o valor mínimo do BIC seja para a interação $C=3$. A Figura 6.10 mostra o resultado do teste, confirmando que a expectativa.

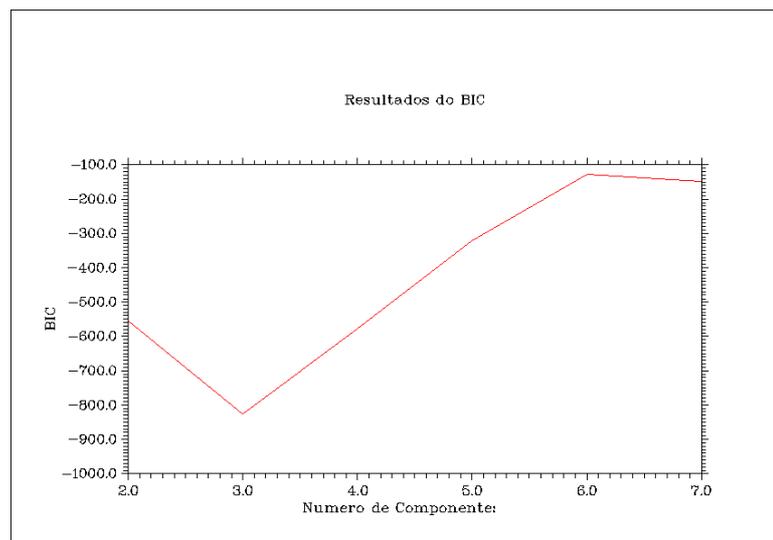


Fig 6.10 – Resultado do BIC indicando corretamente três litofácies distintas no dado sintético

A curva no gráfico acima com o mínimo em $C=3$ indica que o “Bayesian Information Criteria” indicou corretamente o número de litofácies existentes nos dados.

6.5.1 – Estimativa inicial de μ quando não é conhecido o valor central dos atributos das litofácies

Quando não existe nenhuma informação quanto aos valores de μ , , deve-se procurar formas alternativas para a realização da estimativa inicial deste parâmetro. Uma boa alternativa é criar distribuições de probabilidade para cada atributo que será utilizado como entrada do EM e selecionar os valores máximos da distribuição. As distribuições ilustradas pelas figuras 6.11 e 6.12, foram criadas a partir do poço sintético, podemos observar três máximos na distribuição, no qual seriam bons candidatos a estimativa inicial do parâmetro média.

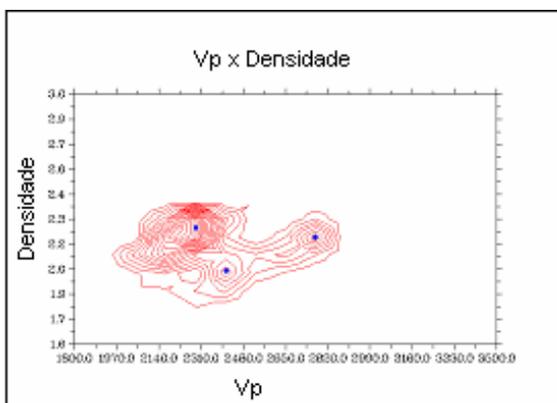


Fig 6.11 – Distribuição de Vp e densidade gerada para realizar a estimativa inicial de μ

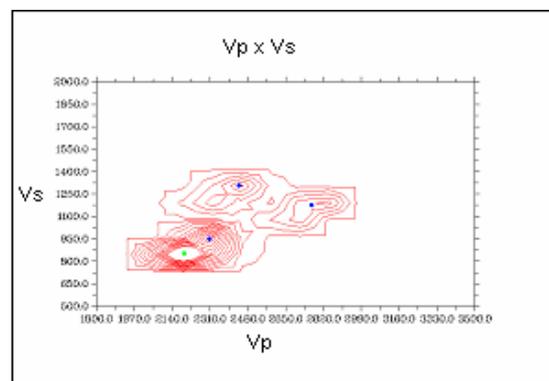


Fig 6.12 – Distribuição de Vp e Vs gerada para realizar a estimativa inicial de μ

Os pontos em azul nas figuras 6.11 e 6.12 representam fortes candidatos a estimativa inicial da média; o ponto em verde seria uma opção secundária. Desta forma é possível realizar a estimativa inicial de forma coerente.

6.6 – Escolhendo os atributos que forneçam a melhor qualidade de classificação

Foi discutido, no capítulo anterior, que a qualidade da classificação depende basicamente do quanto as distribuições dos parâmetros para cada litofácies, se sobrepõem. Quanto maior a sobreposição entre distribuições, maior o grau de incerteza envolvido do processo classificatório. Foi visto que se pode medir a quantidade de sobreposição entre distribuições de probabilidade com o uso do erro de Bayes.

Para a escolha dos pares de atributos que minimizem o erro de classificação, foi calculado o erro de Bayes para todas as combinações de parâmetros comumente obtidos por inversão de AVO (ex. V_p , V_s , Impedância P, Impedância S, etc.) e é escolhido o par que forneça o menor erro de Bayes. Os gráficos a seguir (figuras 6.13 a 6.16) apresentam os resultados obtidos ao aplicarmos este critério de escolha no poço sintético. Cada grupo de cor representa um parâmetro. No eixo Y temos o erro de Bayes correspondente a cada teste para cada par de parâmetro. As barras que representam o maior erro de Bayes ocorrem justamente quando é testado um atributo com ele mesmo (ex. V_p e V_p) o que comprova que ao ser acrescida mais informação no processo, consegue-se diminuir as incertezas. O teste foi aplicado separadamente para cada litofácies e depois feita a média do erro de Bayes. Os pares de atributos que forneceram o menor erro de Bayes médio, para todas as litofácies, foram V_p e K (incompressibilidade). Estes resultados não representam uma situação geral para o uso destes dois atributos em situação real, pois devido a heterogeneidade de cada reservatório, cada caso deve ser analisado separadamente.

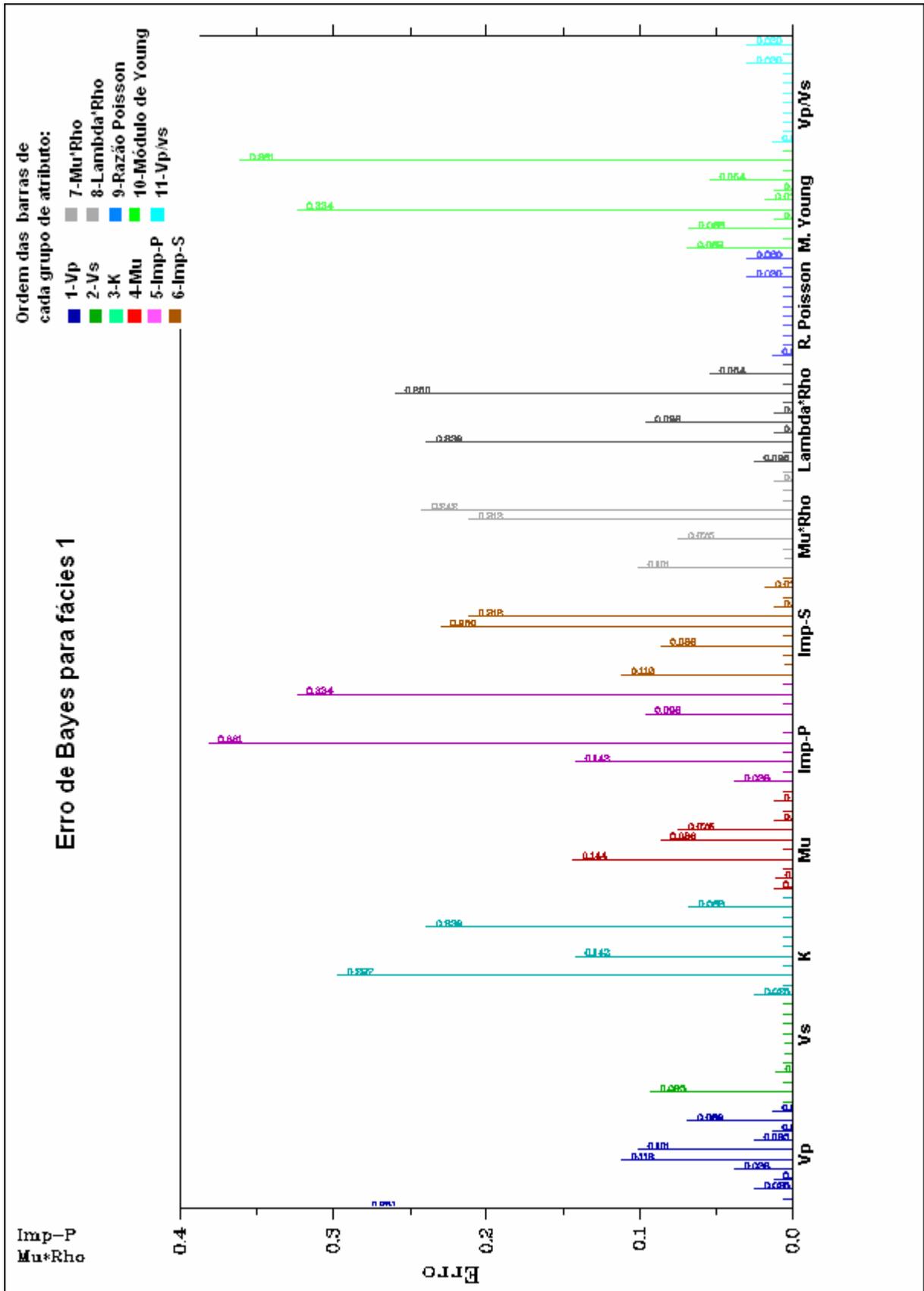


Fig 6.13 – Erro de Bayes para cada par de atributos, os atributos Impedância P e μ^*Rho apresentam o menor erro para litofácies 1

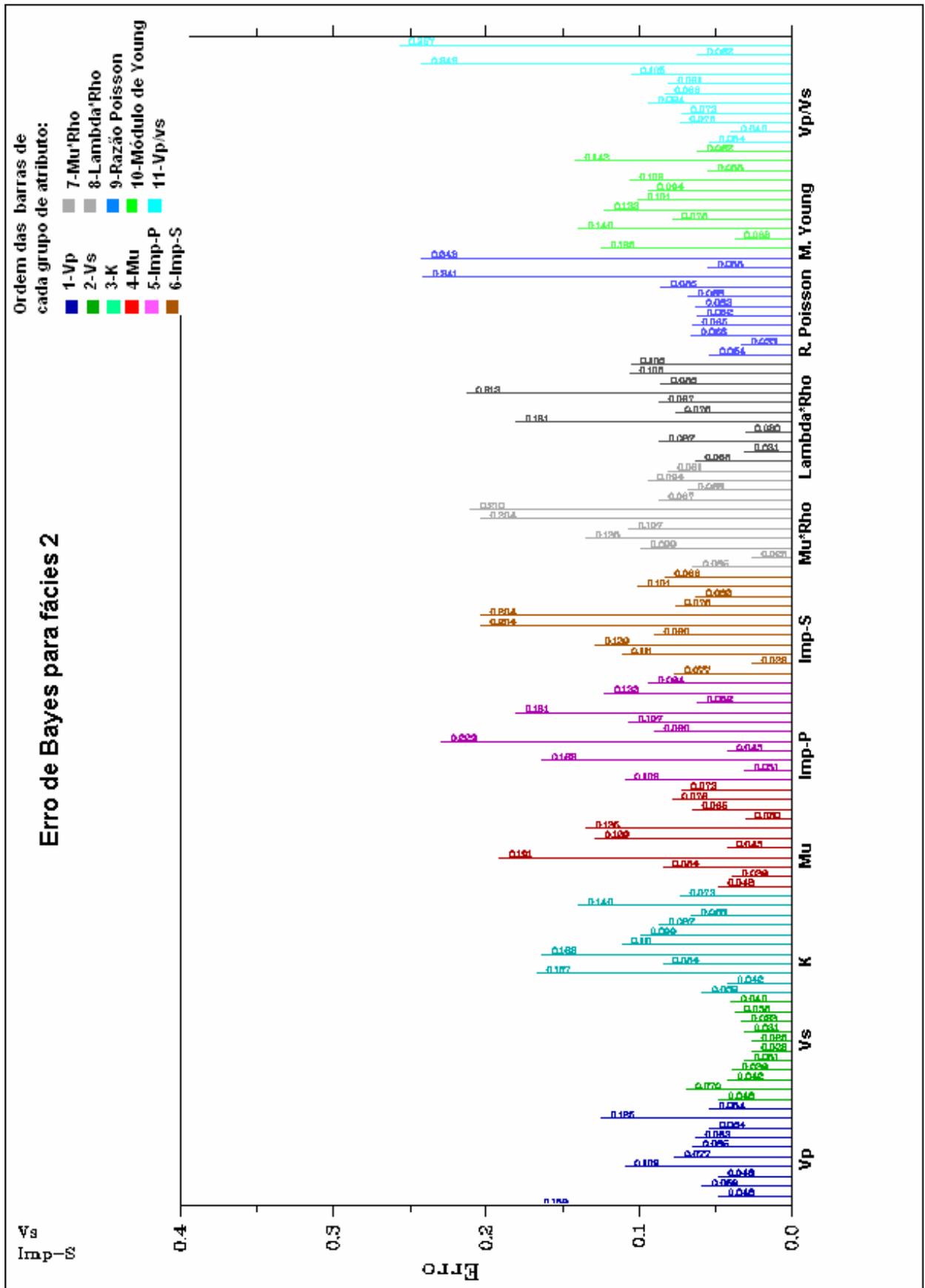


Fig 6.14 – Erro de Bayes para cada par de atributos. Os atributos Impedância S e Vp apresentam o menor erro para litofácies 2

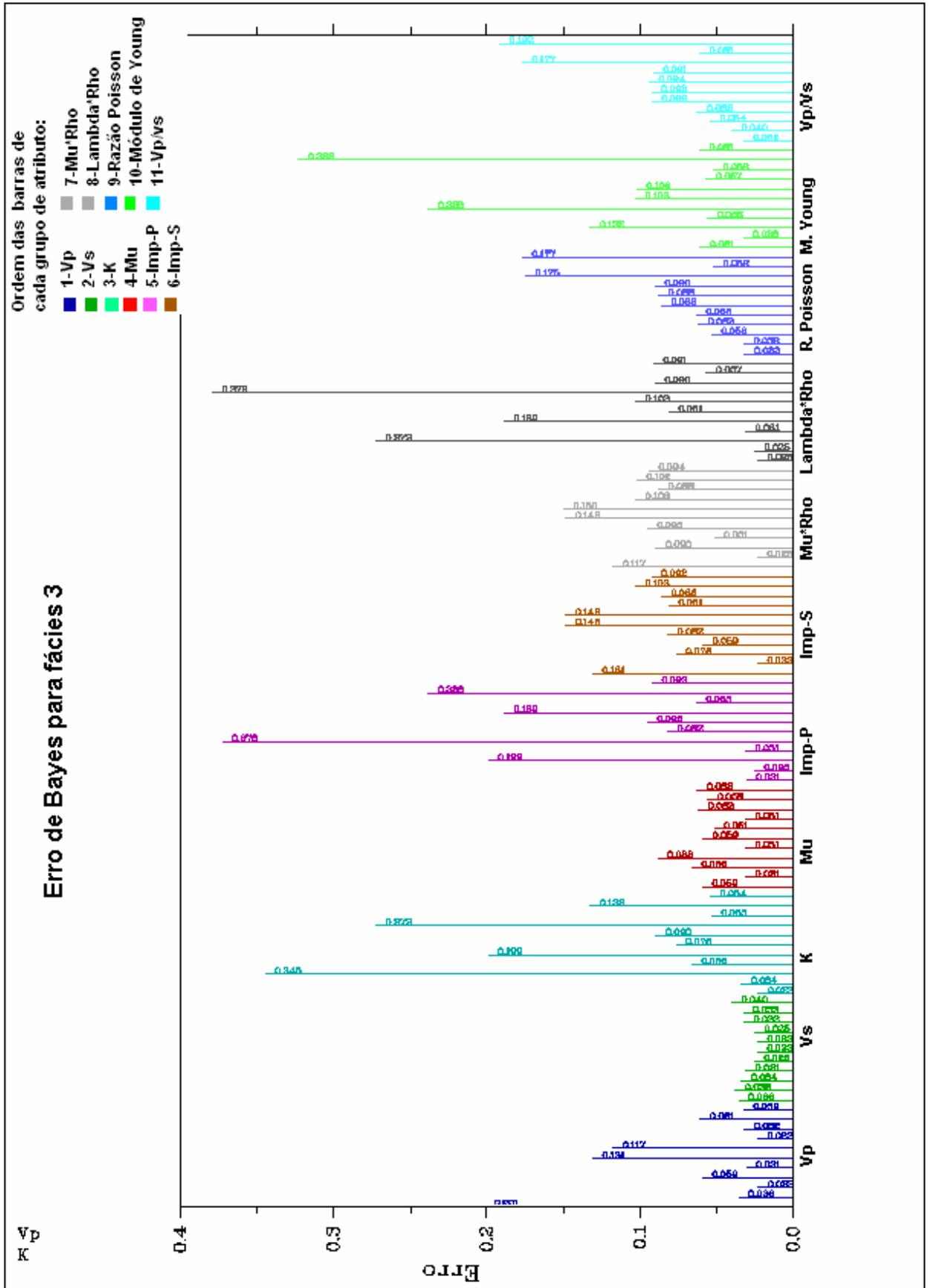


Fig 6.15 – Erro de Bayes para cada par de atributos, os atributos K e Vp apresentam o menor erro para litofácies 3

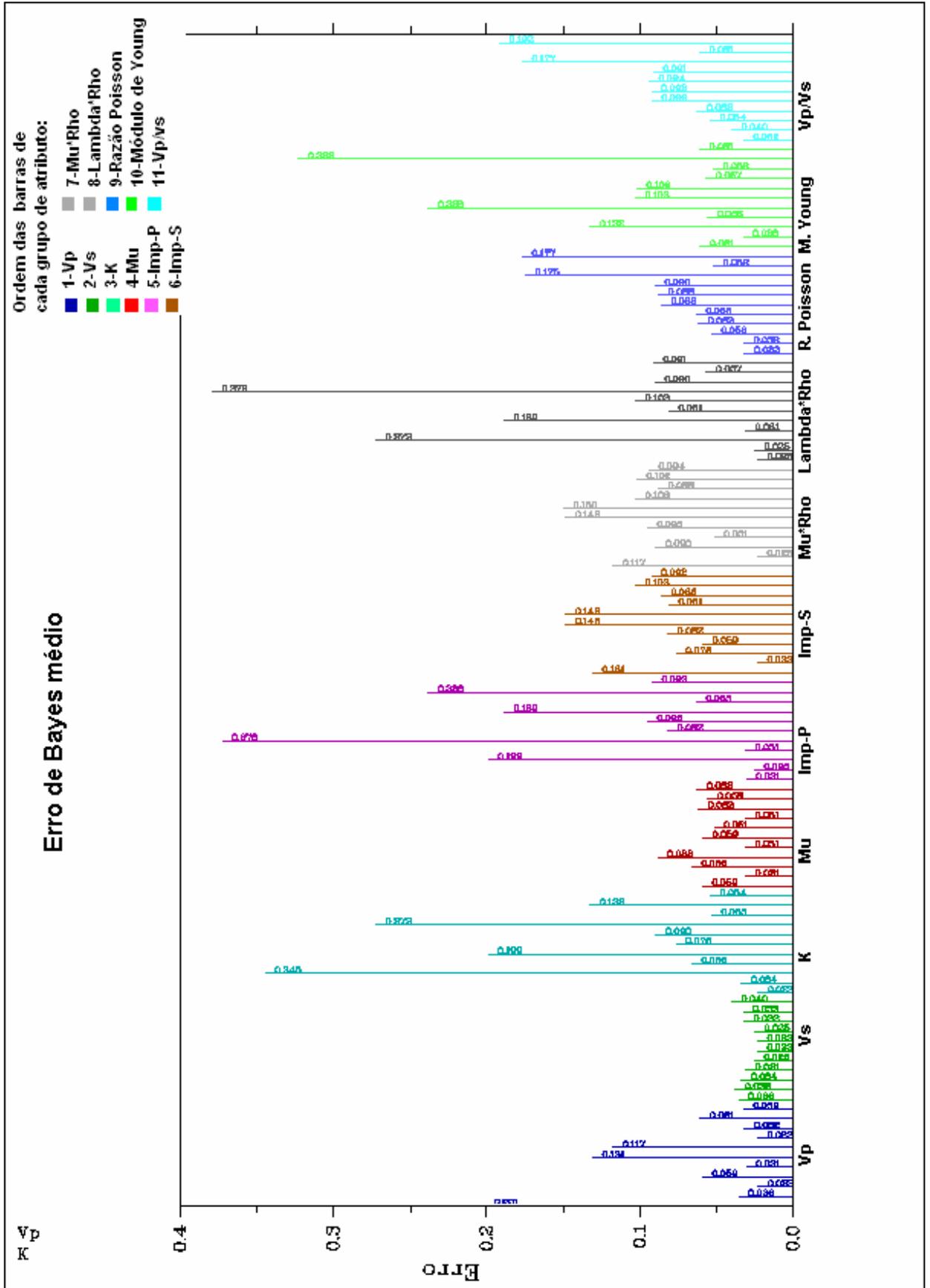


Fig 6.16 – Erro de Bayes para cada par de atributos, os atributos K e Vp apresentam o menor erro médio

6.7 – Classificando a seção de atributos sísmicos com o classificador MAP

Conforme visto na seção 6.4, pode-se fazer uso do EM para definir as litofácies existentes nos dados de poço e assim gerar o dado de treinamento, utilizado para a classificação das seções sísmicas. Nesta seção, aplica-se o classificador MAP nas seções sintéticas geradas, com o intuito de analisar a qualidade da classificação do método. É utilizado o poço sintético, como dado de treinamento e as seções de atributos como dado de entrada. Avaliam-se duas situações; i) a presença de ruído no dado e ii) a classificação com atributos que apresentam um erro de Bayes maior do que o par de parâmetro escolhido. Na seção anterior, foi visto que os parâmetros V_p e K (incompressibilidade) são os atributos que fornecem a menor incerteza de classificação para este conjunto de dados. Assim, são utilizados estes dois pares e compara-se com outro par que apresente um nível de incerteza maior. Avalia-se o resultado das duas classificações.

Outro fator considerado neste teste é a escolha do parâmetro k utilizado pelo método de k -vizinhos mais próximos (parte integrante do classificador). Não existem regras que fixem o valor para este parâmetro, porém o valor de k sempre deve ser um valor ímpar para não ocorrer empate na decisão.

6.7.1 – Teste com as seções sintéticas com 5% de ruído

É apresentado inicialmente o teste de classificação com o modelo de 5% de ruído, pois com o modelo sem ruído não apresentou erro algum de classificação. Primeiramente foi realizado o teste excluindo o nó da rede Bayesiana que é composto do método de k -vizinhos mais próximos e com o par de parâmetro que apresentou a melhor separabilidade entre as litofácies (V_p e K -incompressibilidade). Para informação a priori, considera-se todas as litofácies equiprováveis. O índice de acerto obtido foi igual a 92,7%. A Figura 6.17 apresenta o resultado do teste.

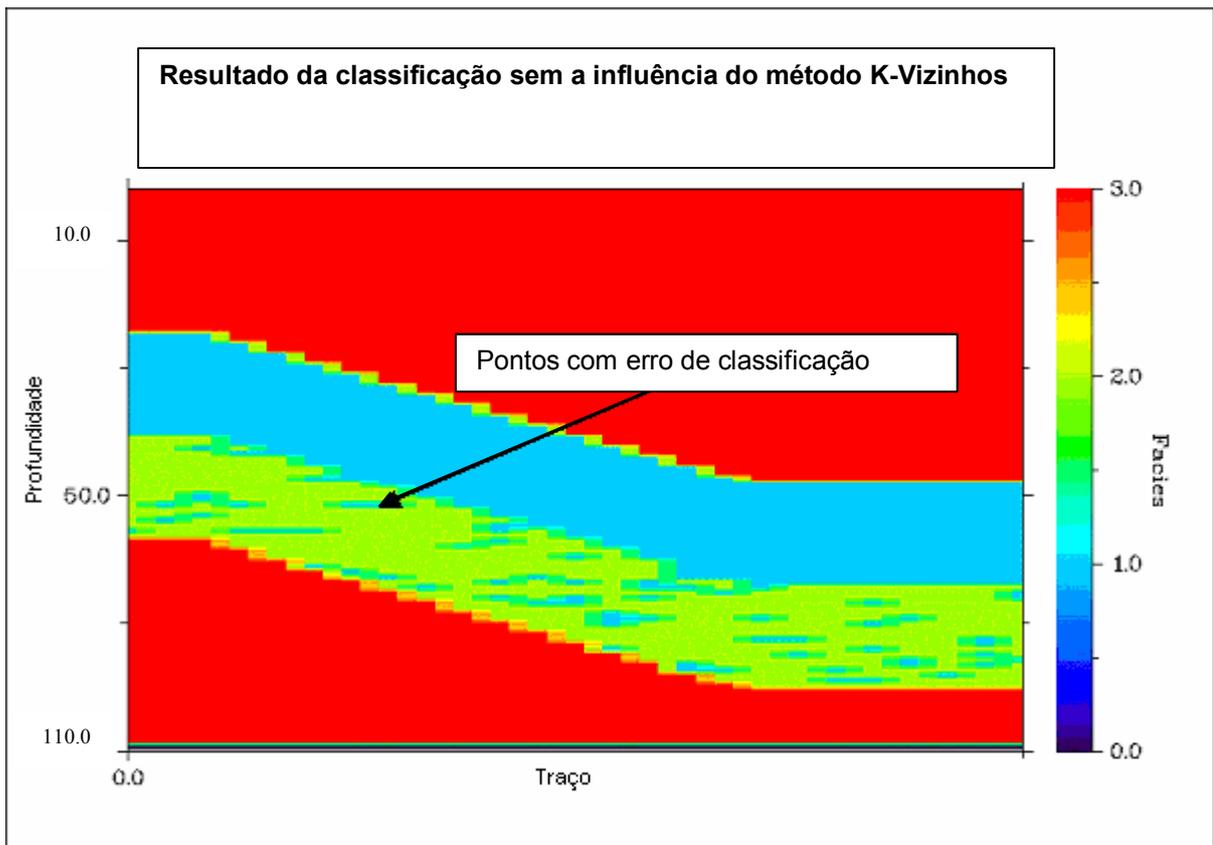


Fig 6.17 – Resultados do classificador MAP (sem a influência do método de k-Vizinhos) com as seções de atributos sísmicos sintéticas com ruído de 5%

O próximo teste foi realizado com a estrutura da rede bayesiana completa. O índice de acerto neste caso foi igual a 98,9%. Pode-se perceber que a presença desta nova informação ajuda a estabilizar a tarefa de classificação. A Figura 6.18 ilustra o resultado. É visível a melhora considerável no resultado da classificação com a inclusão do método de k-vizinhos, já que foram classificados um número muito menor de amostras de forma incorreta.

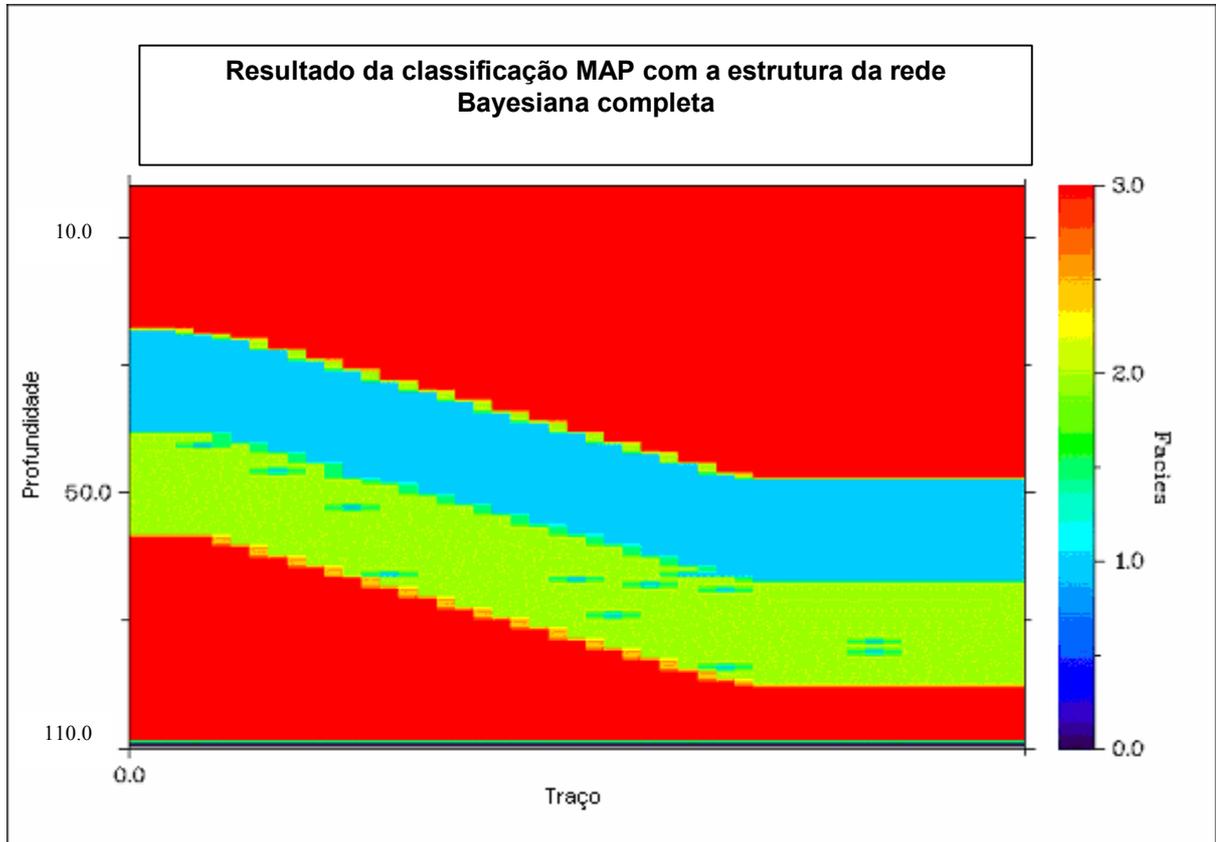


Fig 6.18 – Resultados do classificador MAP com a estrutura da rede Bayesiana completa e seções de atributos sísmicos sintéticos com ruído de 5%

6.7.2 – Classificação utilizando outros atributos com maior nível de incerteza

Na seção 6.6, foi verificado que os atributos V_p e K apresentaram o menor erro de Bayes entre todos os atributos testados. Nesta seção, foi realizada a classificação com os atributos V_p e V_s , que apresentaram um erro de Bayes um pouco superior. O intuito deste teste é comprovar a influência da escolha dos atributos no processo de classificação. É esperado que o resultado desta classificação seja menos eficaz do que o anterior. A Figura 6.19 ilustra a resultado deste teste.

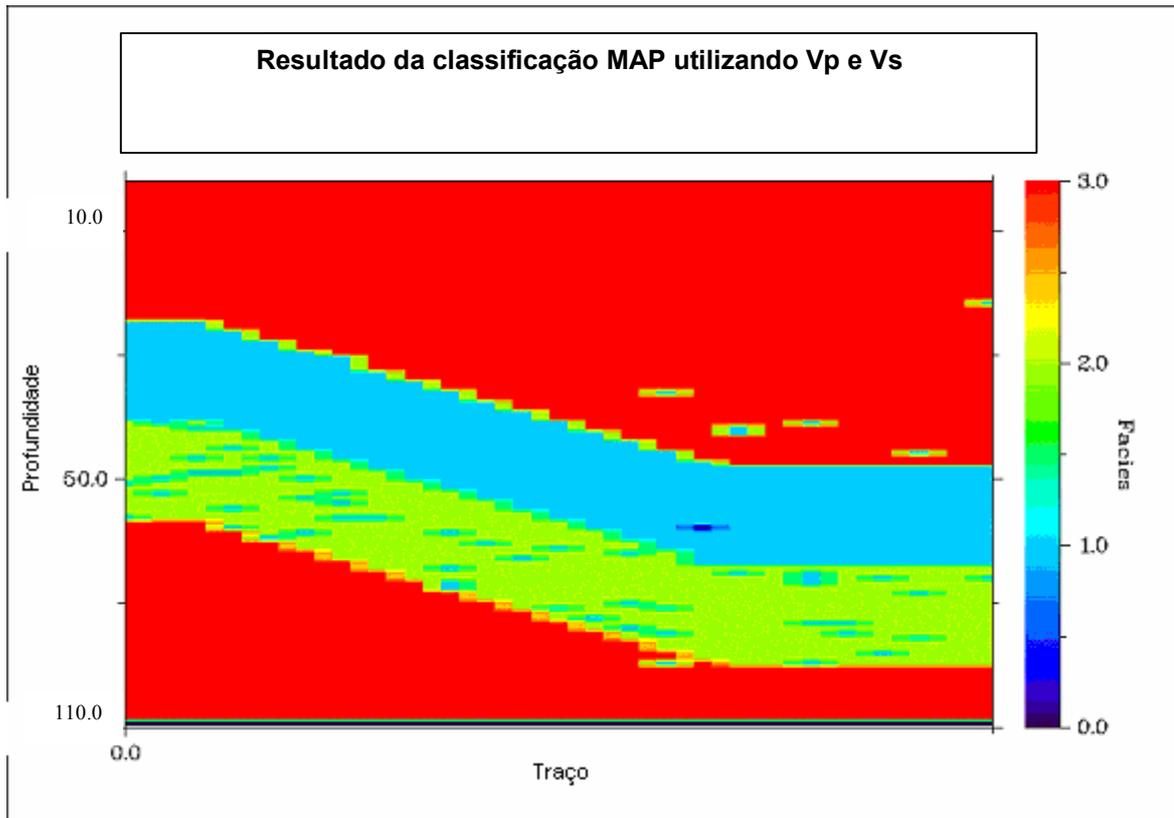


Fig 6.19 – Resultados do classificador MAP com as seções de Vp e Vs sintéticas com ruído de 5%

Nota-se nos resultados da Figura 6.19, uma piora considerável no processo de classificação, comprovando que o erro de Bayes é uma medida confiável para avaliarmos as incertezas envolvidas na escolha dos parâmetros utilizados na classificação de novas amostras.

6.7.3 – Teste com as seções sintéticas com 10% de ruído

O classificador na sua forma completa foi aplicado ao modelo de atributos com 10% de ruído. A finalidade deste teste é somente avaliar a sensibilidade no ruído nos dados de entrada. O poço sintético é o mesmo da seção 6.7.1 (5% de ruído). A Figura 6.20 mostra o resultado da classificação com Vp e K, a figura 6.21 a classificação com Vp e Vs.

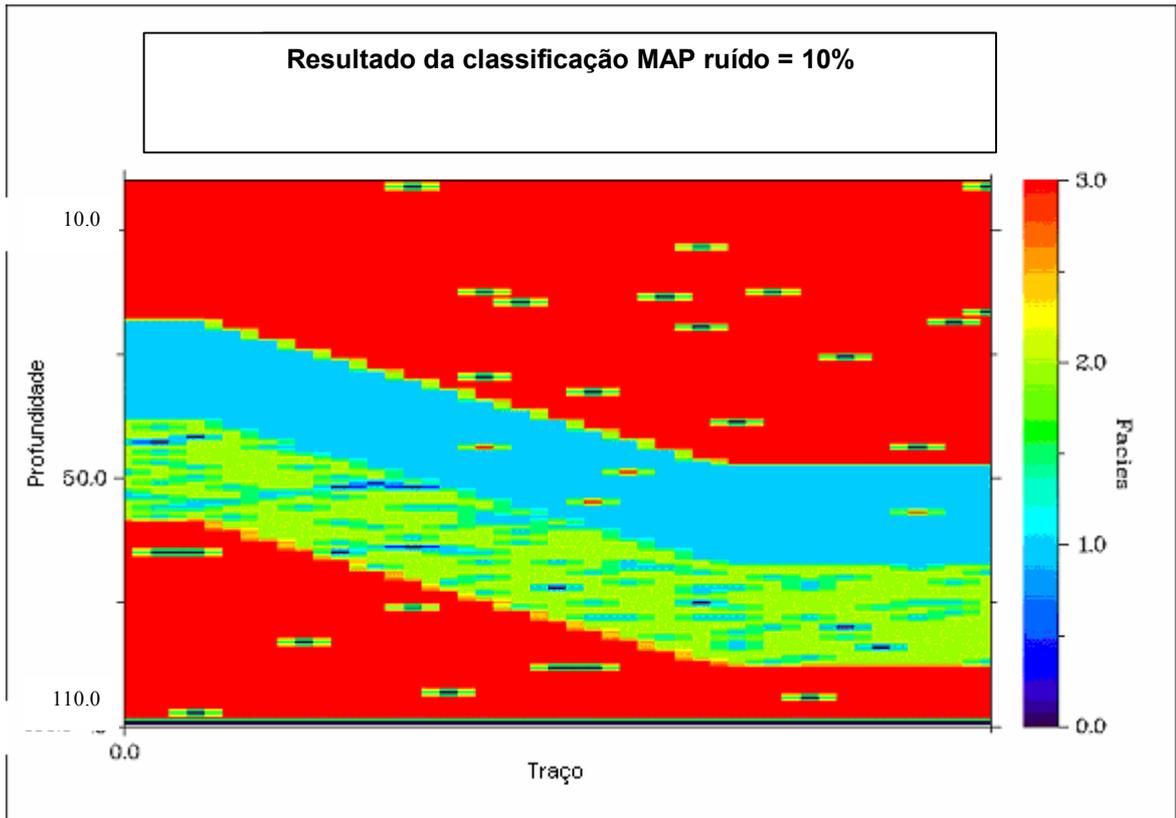


Fig 6.20 – Resultados do classificador MAP com as seções de Vp e K sintéticas com ruído de 10%

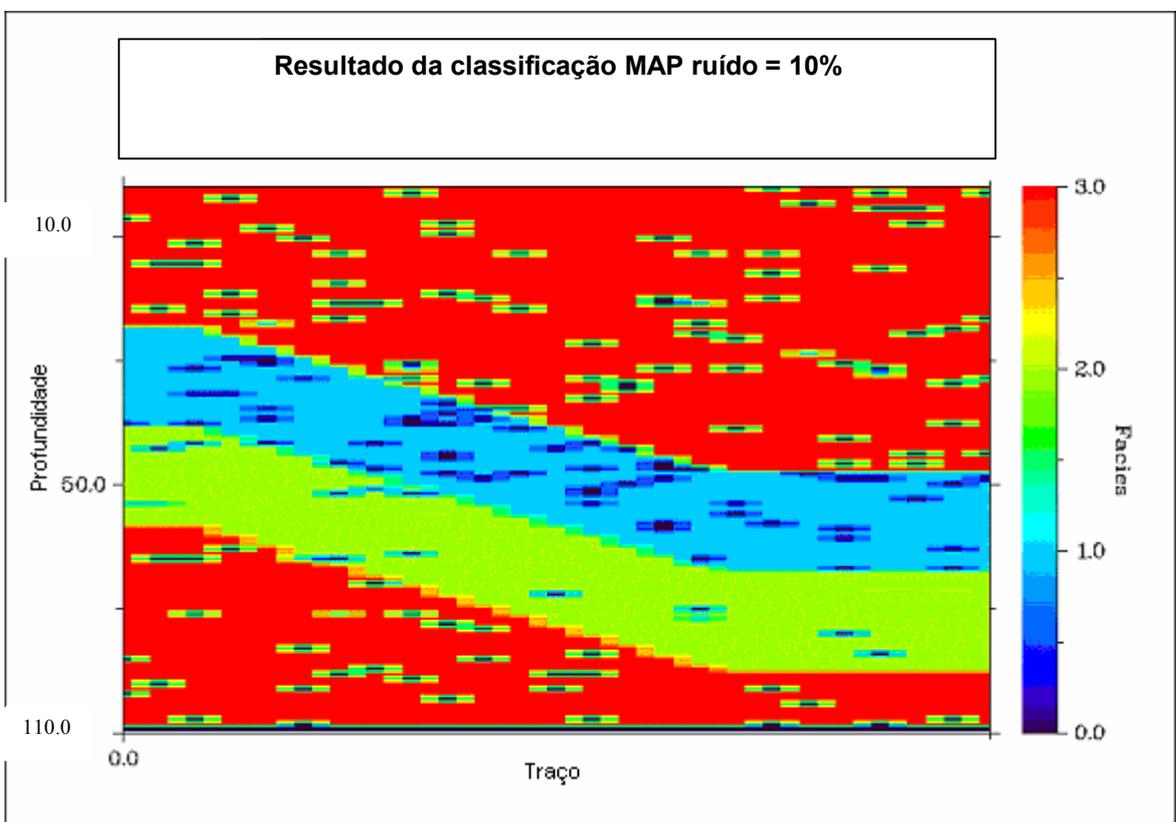


Fig 6.21 – Resultados do classificador MAP com as seções de Vp e Vs sintéticas com ruído de 10%

Nota-se que a qualidade da classificação é proporcional a quantidade de ruído existente nos dados de entrada. Embora exista um número considerável de amostras classificadas erroneamente, ainda é possível distinguir com facilidade as camadas do modelo. Neste teste fica ainda mais evidente a influência da escolha dos parâmetros no processo de classificação.

As figuras 6.22 a 6.24 representam a probabilidade de cada litofácies dado o modelo de 10% de ruído e os atributos V_p e V_s . Elas ilustram o quanto de incerteza existe na classificação de cada litofácies.

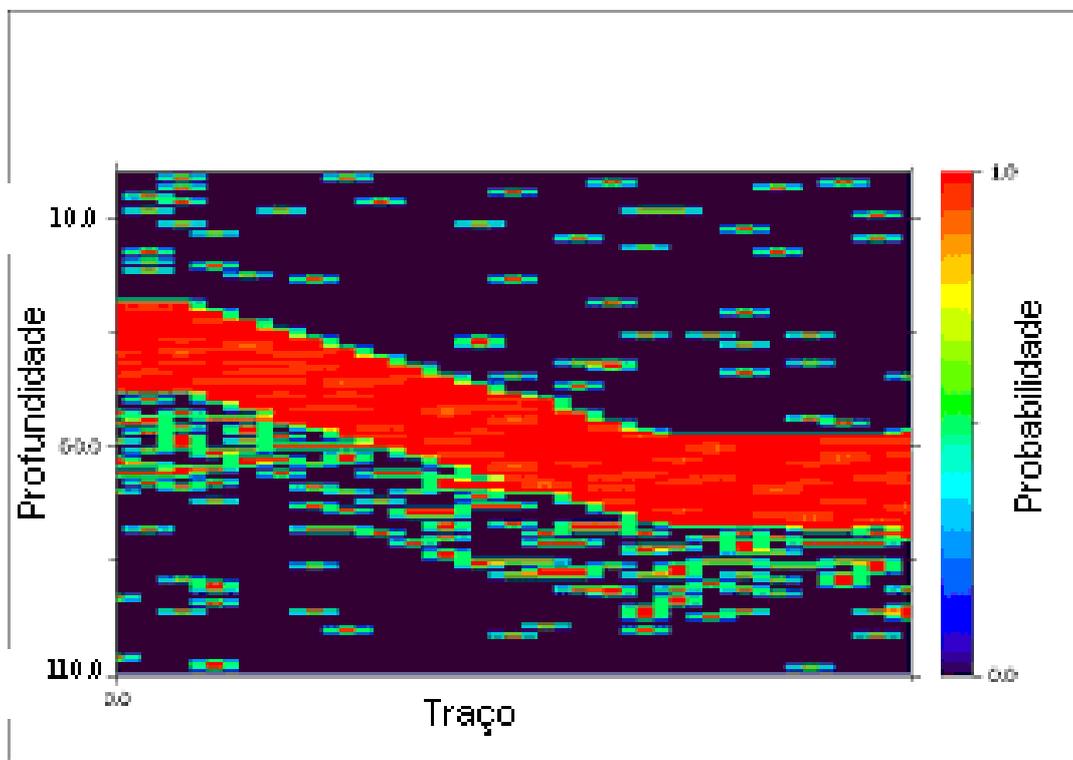


Fig 6.22 – Probabilidade de litofácies 1 para o modelo com 10% de ruído e V_p e V_s como atributos de entrada

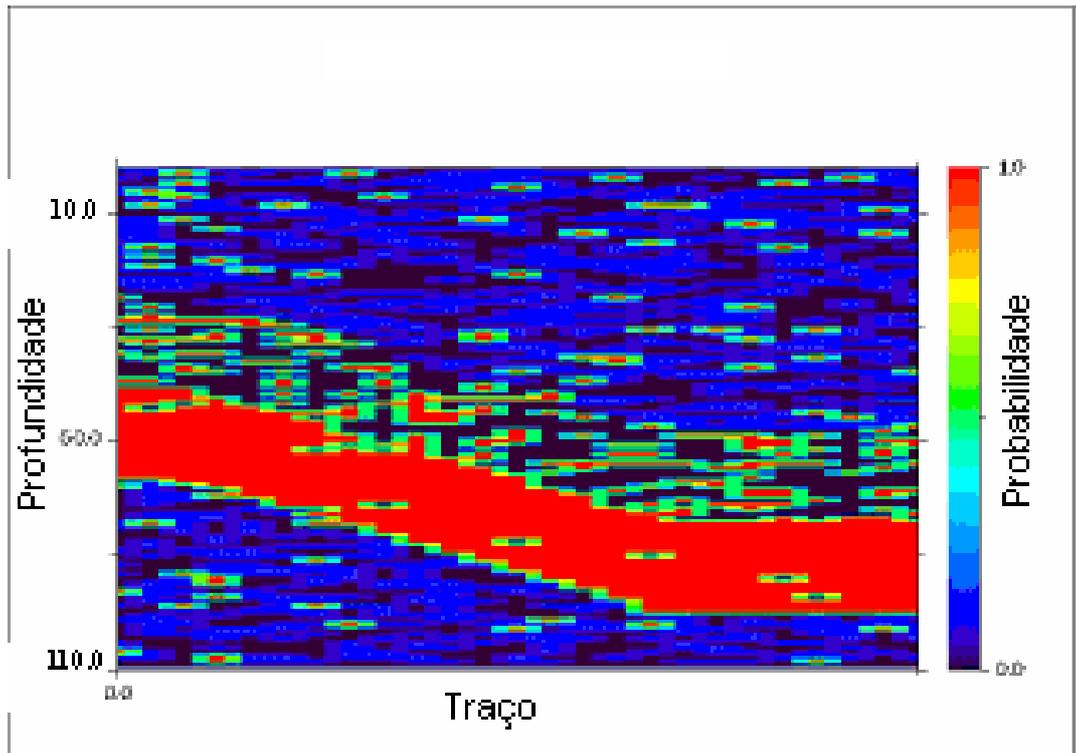


Fig 6.23 – Probabilidade de litofácies 2 para o modelo com 10% de ruído e Vp e Vs como atributos de entrada

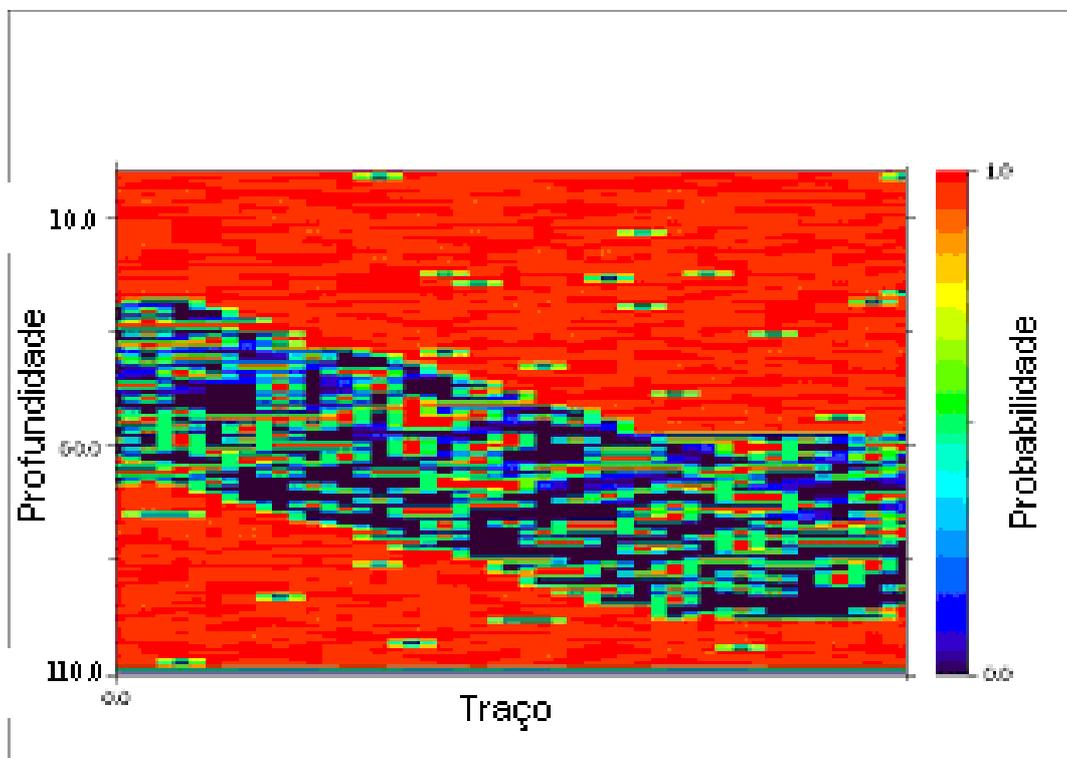


Fig 6.24 – Probabilidade de litofácies 3 para o modelo com 10% de ruído e Vp e Vs como atributos de entrada

6.8 – Conclusão

Neste capítulo a metodologia proposta foi aplicada em um modelo sintético. As etapas realizadas seguem o fluxograma proposto. A única etapa não realizada do fluxograma foi a análise petrofísica dos dados por se tratarem de dados sintéticos, onde suas características já são conhecidas. Avaliou-se como a escolha dos parâmetros de entrada influenciam a qualidade da classificação e como o erro de Bayes é uma medida consistente de incerteza. Outro fator avaliado foi a presença de ruídos no dado de entrada e como o resultado é proporcional ao nível de ruído.

A combinação do resultado do classificador MAP junto com os mapas de probabilidade de cada litofácies formam o resultado final da metodologia. Em uma aplicação real os dois resultados devem ser analisados em conjunto para avaliar o grau de incerteza na classificação.

No próximo capítulo será aplicada a metodologia em um conjunto de dados reais de um reservatório turbidito de um campo de petróleo maduro.

Capítulo 7

Aplicação da metodologia desenvolvida em um conjunto de dados reais

7.1 – Introdução

Este capítulo demonstra a aplicação da metodologia desenvolvida em um conjunto de dados de um reservatório arenito/folhelho, no Campo Morena.

A aplicação da metodologia foi dividida de duas formas. Primeiramente foram utilizados, como dado de treinamento, os poços interpretados por um geólogo intérprete, onde as litofácies já encontravam-se identificadas (aprendizado supervisionado) excluindo a etapa do algoritmo EM. Segundo, foi aplicado a metodologia na sua forma completa, conforme idealizada. Os resultados são comparados e discutidos.

7.2 – Descrição do Campo

Segundo Ferreira et al. (2002) o Campo de Morena foi descoberto em março de 1984 e está localizado entre lâmina d'água de 300 a 700 m. A produção atual consiste em 49.000 barris de óleo por dia. O campo produz óleo de reservatórios de

turbiditos do santoniano-cretáceo e do eoceno. Existe ainda uma ocorrência minoritária de óleo nos arenitos do turoniano. São reconhecidos quatro elementos arquiteturais; i) sistema de preenchimento de canyons, onde as principais litofácies existentes são arenito maciço grosso e arenito maciço médio; ii) depósitos de escorregamento, onde são encontrados folhelho deformado e arenito deformado; iii) sistema de canais, onde as litofácies presentes são semelhantes às encontradas no depósito de preenchimento de canyons porém são um pouco menos grossas, denotando menor competência do transporte; iv) sistemas de canais divergentes, onde a deposição foi realizada por correntes mais diluídas e as litofácies resultantes foram arenitos maciços de finos a médios com estratificação plano/paralela. As figuras 7.1 e 7.2 mostram respectivamente uma seção geológica perpendicular e paralela ao eixo do canyon.

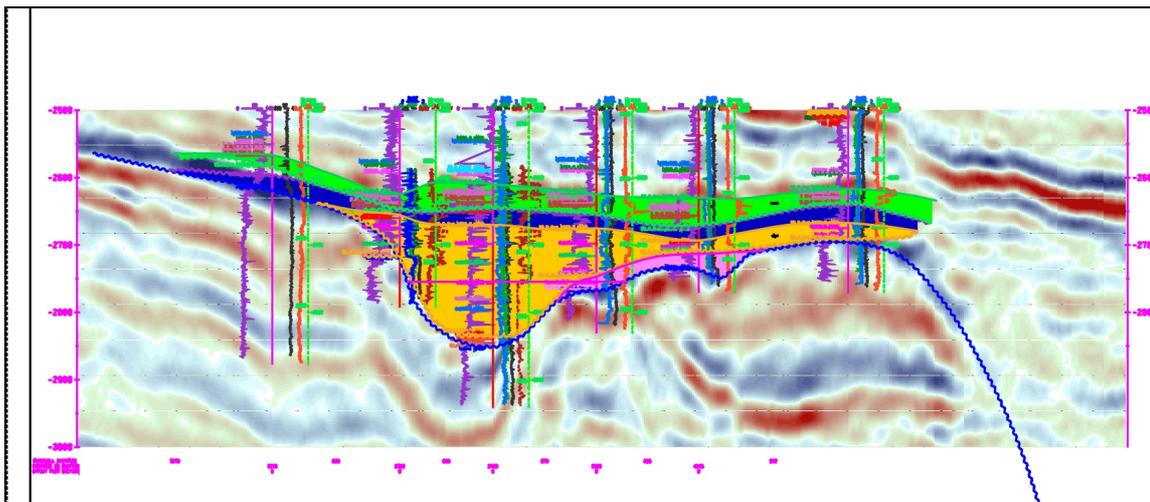


Fig 7.1 – Seção perpendicular ao eixo do canyon, retirado de Ferreira et al. (2002)

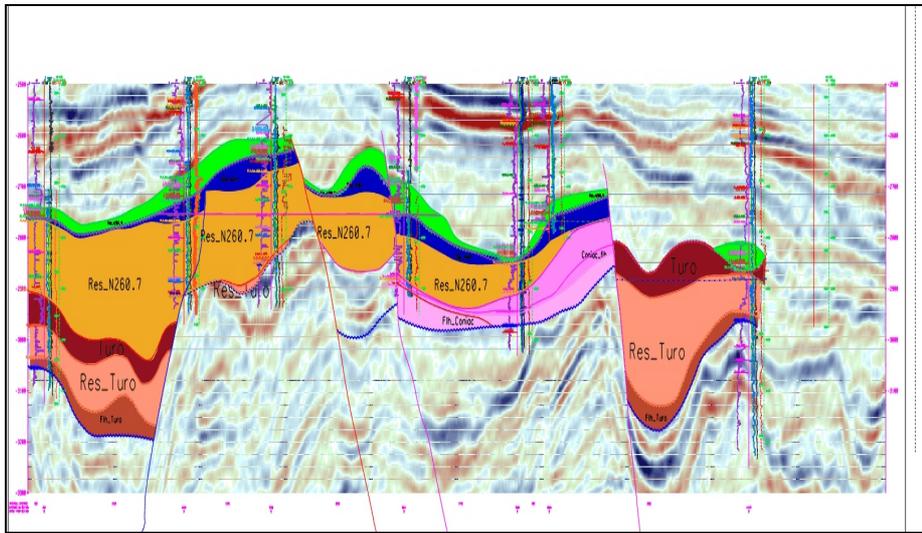


Fig 7.2 – Seção paralela ao eixo do canyon retirado de Ferreira et al., (2002)

7.3 – Fontes de informação

Para o teste, foram disponibilizados cinco poços, aqui denominados de w1, w2, w3, w4 e w5, além de uma seção sísmica (i_line 1) de um levantamento marítimo 3D. Os poços w1 à w4 estão localizados próximos a área de estudo, porém não possuem perfis sônicos de onda cisalhante. O poço w5 está localizado um pouco afastado da área de estudo e possui perfil sônico bipolar (V_p e V_s). O poço w5 foi utilizado como base para a estimativa de V_s nos outros poços. A metodologia utilizada para a estimativa foi a de Greenberg e Castagna (1992). A seção sísmica foi invertida para a obtenção de uma seção sísmica de V_p e uma de V_s . Os poços foram processados e as fácies foram definidas por um intérprete. Foram definidas quatro litofácies distintas na área de estudo conforme abaixo:

- I. Litofácies 1 – Arenito limpo – reservatório de boa qualidade.
- II. Litofácies 2 – Arenito com presença de argila – reservatório.
- III. Litofácies 3 – Arenito cimentado – não reservatório.
- IV. Litofácies 4 – Folhelho – não reservatório.

A diferença de classificação entre as litofácies aqui descritas e a da seção anterior, deve-se a primeira classificação ser baseada em testemunhos, e a segunda classificação ser baseada nos perfis. Para este teste utiliza-se as litofácies definidas segundo os perfis.

As figuras 7.3 - 7.6 mostram os quatro poços w1 à w4 que serão utilizados como dado de treinamento. A Figura 7.7 mostra o ajuste obtido com o poço w5 na estimativa de V_s .

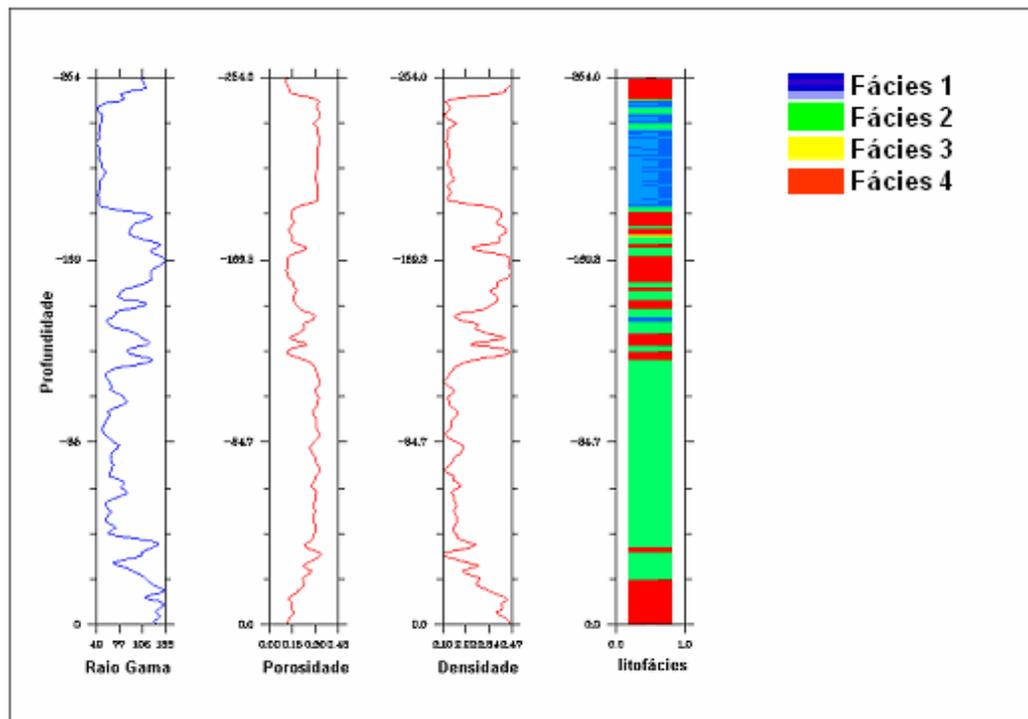


Fig 7.3 – Poço w1. Os perfis da esquerda para a direita são: raio gama, porosidade, densidade e interpretação das litofácies.

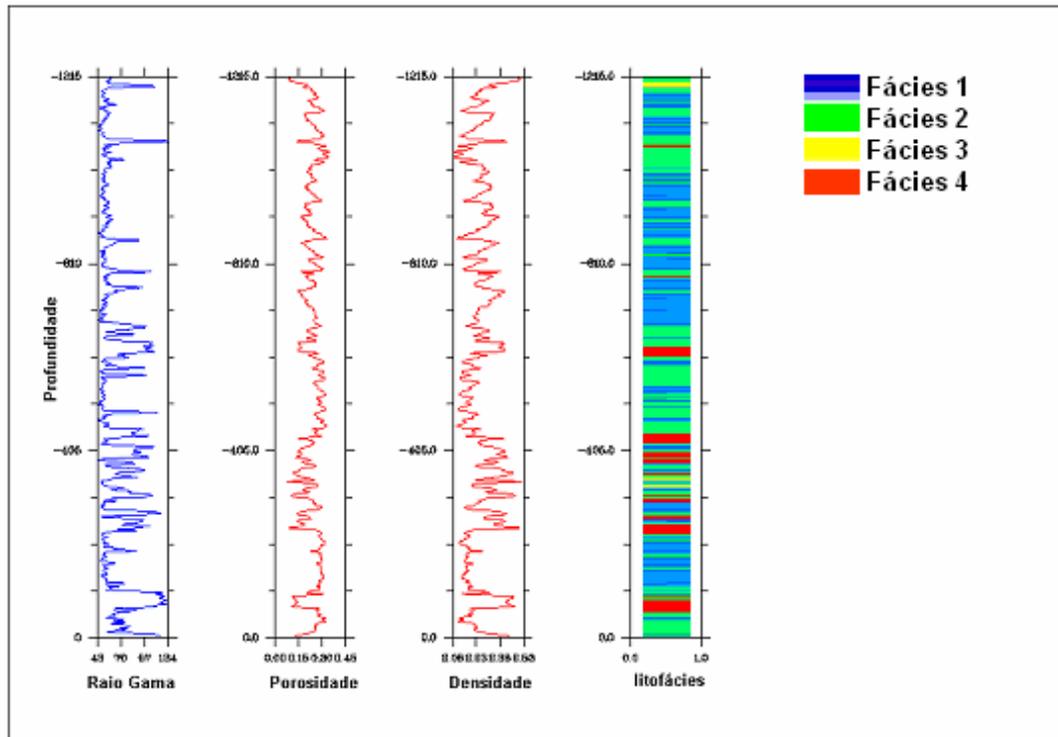


Fig 7.4 – Poço w2. Os perfis da esquerda para a direita são: raio gama, porosidade, densidade e interpretação das litofácies.

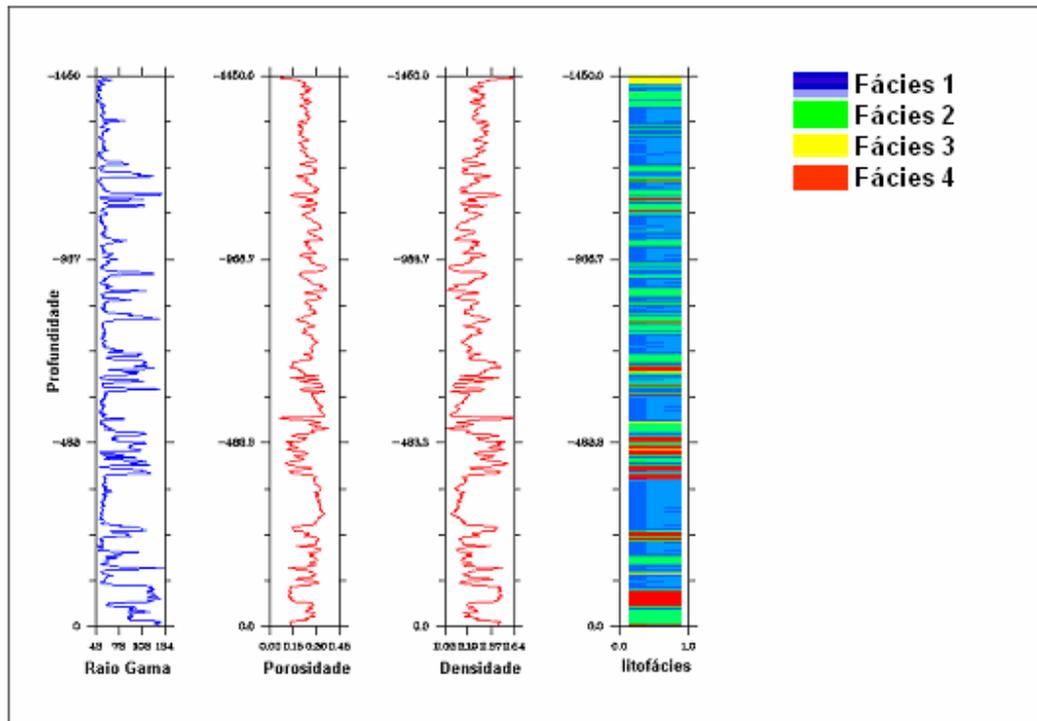


Fig 7.5 – Poço w3. Os perfis da esquerda para a direita são: raio gama, porosidade, densidade e interpretação das litofácies.

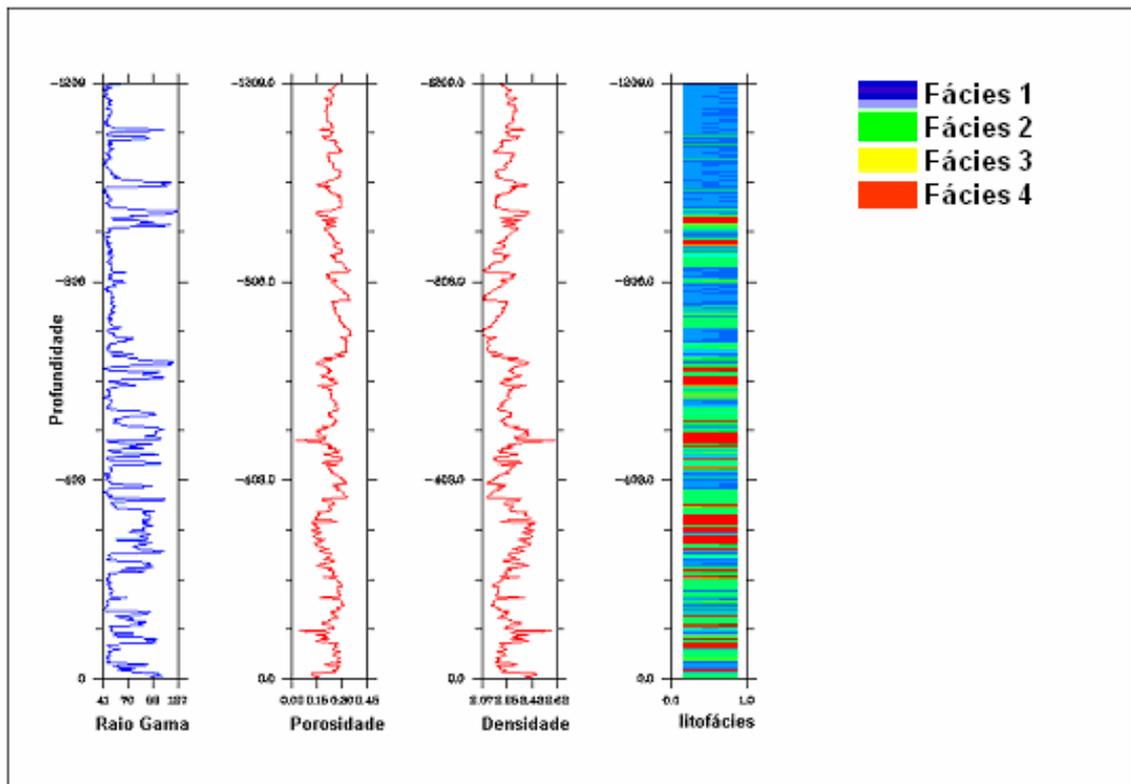


Fig 7.6 – Poço w4. Os perfis da esquerda para a direita são: raio gama, porosidade, densidade e interpretação das litofácies.

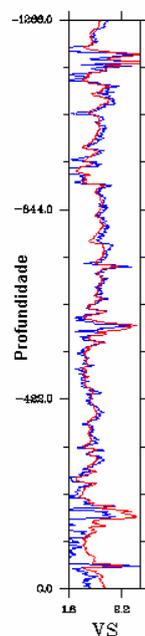


Fig 7.7 – Estimativa de Vs (vermelho) comparado com Vs do poço w5 utilizado como referência para a estimativa de Vs nos demais poços

7.4 – Análise de incerteza dos dados

Seguindo o fluxograma da metodologia proposta, quando já existem dados disponíveis e interpretados, a próxima etapa consiste em criar distribuições de probabilidade para cada atributo e analisar qual par de atributos fornece o menor risco de classificação. Com o uso dos poços w1 a w4, foi realizada a análise de incerteza com base no erro de Bayes. Os resultados encontram-se resumidos na Figura 7.8.

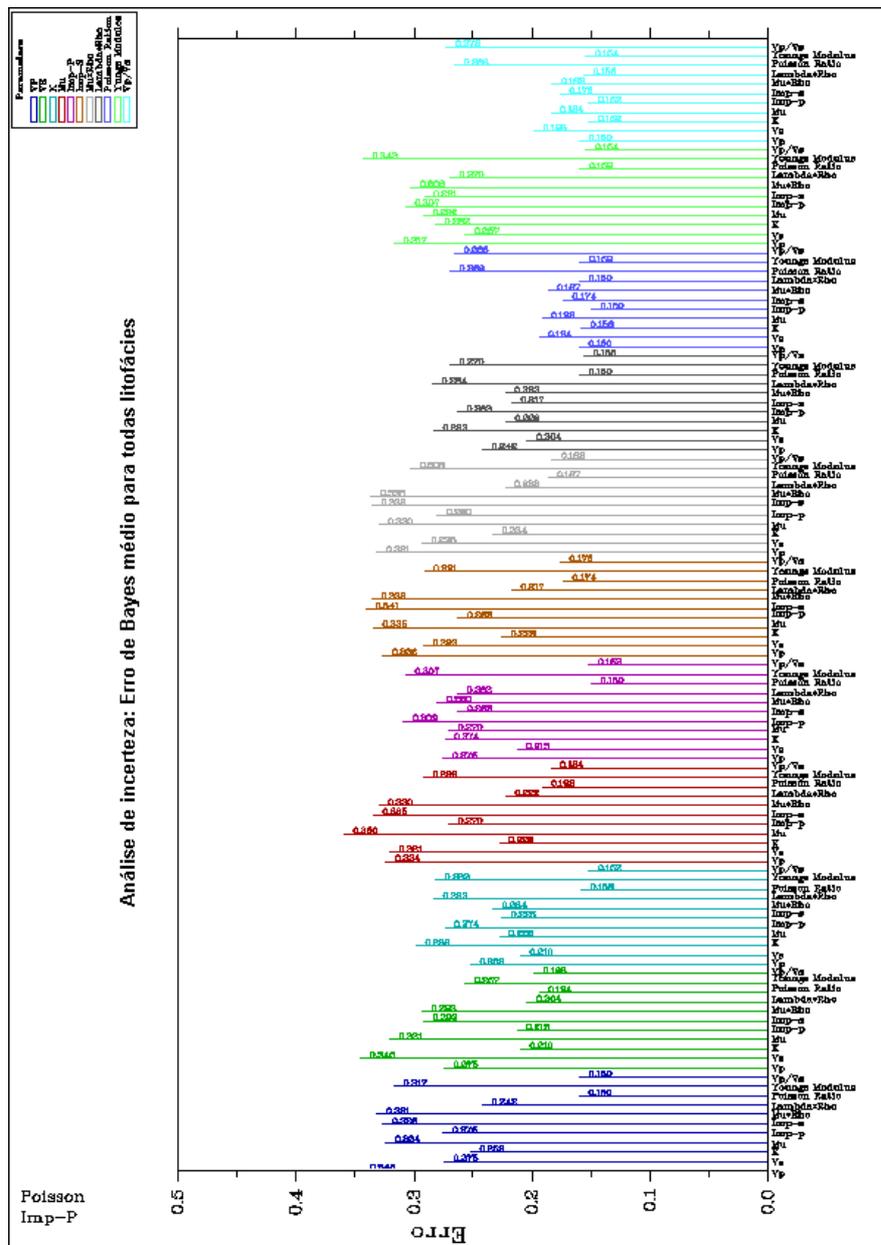


Fig 7.8 – Resultados da análise de incerteza realizada com dados definidos pela interpretação por geólogo. O par de parâmetros com o menor risco de classificação é impedância-P e razão de Poisson.

Embora os atributos impedância P e razão de Poisson apresentem melhor resultado para a classificação, existem apenas seções de Vp e Vs para a realização do teste. Com isso, foram classificados os poços com esses dois atributos, para que seja feita uma avaliação da qualidade do resultado.

7.5 – Classificando poços a partir de outros poços

Nesta seção, é demonstrado como é possível fazer uso do classificador de máxima probabilidade a posteriori para classificar poços que ainda não foram interpretados, baseados em poços que já o foram. Os poços interpretados formarão o conjunto de dados que será utilizado pelo classificador como dado de treinamento, este tipo de abordagem, quando existe interferência humana para a definição do dado de treinamento é denominado “aprendizado supervisionado”.

Para cada poço classificado, retira-se o mesmo das amostras do dado de treinamento e realiza-se a classificação, com base nos três poços restantes. Foi repetido este procedimento para todos os poços. Com base no teste realizado na seção anterior, foram utilizados os atributos impedância-P e razão de Poisson. Foi comparado o resultado com a classificação utilizando Vp e Vs. As figuras 7.9 a 7.16 apresentam os resultados da classificação para os poços w1 a w4.

As colunas das figuras 7.9, 7.11, 7.13 e 7.15 são respectivamente, o perfil de impedância-P, o perfil de razão de Poisson e o perfil de litofácies com dois traços. O primeiro é a interpretação definida por um geólogo intérprete, a segunda o resultado da classificação obtida pelo classificador. As colunas das figuras 7.10, 7.12, 7.14 e 7.16 são respectivamente, o perfil de Vp, o perfil de Vs e o perfil de litofácies com dois traços. O primeiro é a interpretação definida por um geólogo intérprete, a segunda o resultado da classificação obtida pelo classificador.

Nota-se que embora Vp e Vs apresentem bons resultados, a qualidade é inferior aos resultados obtidos com impedância-p e razão de Poisson, conforme esperado.

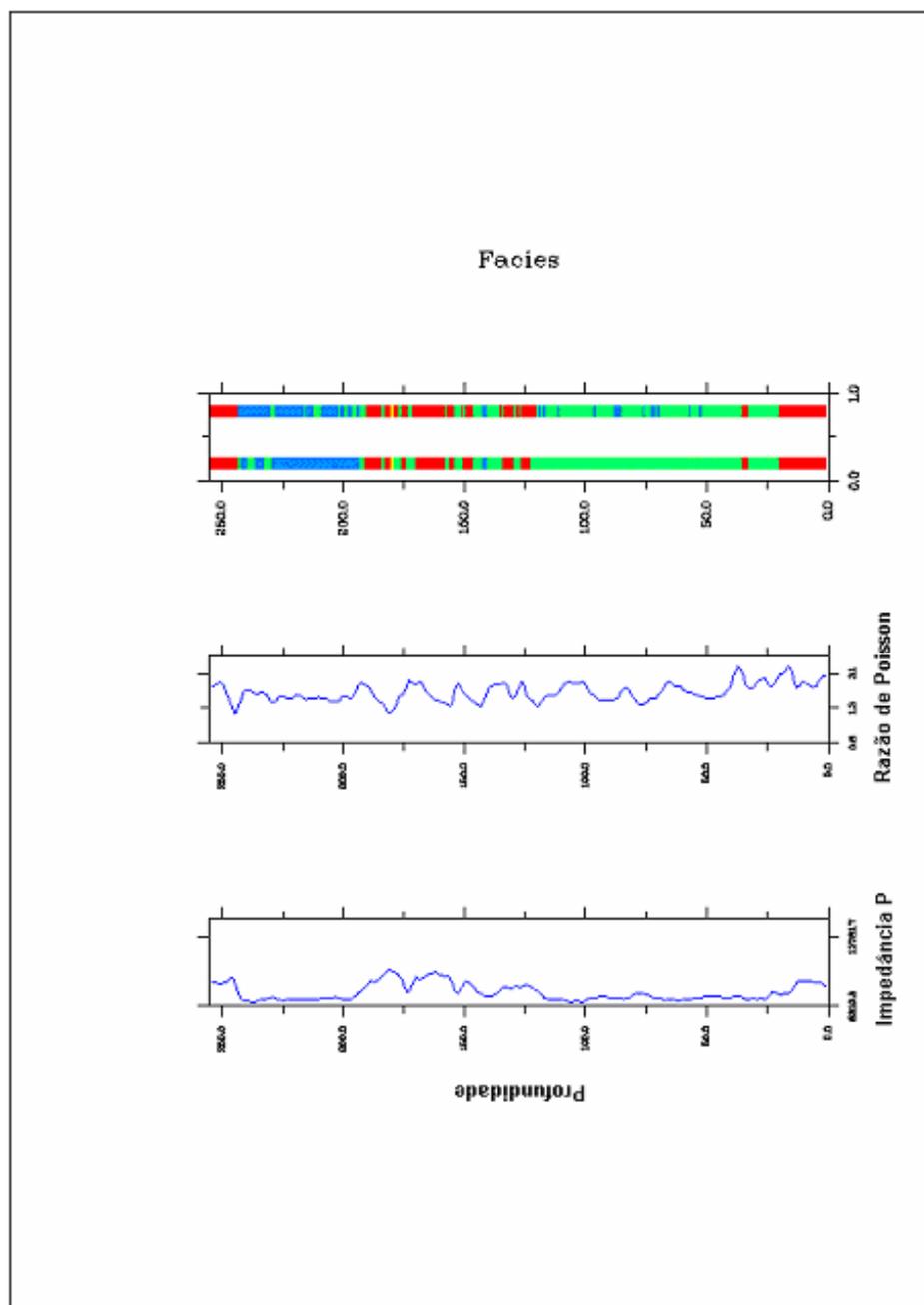


Fig 7.9 – Classificação do poço w1 com impedância-P e razão de Poisson. O índice de acerto obtido foi de 83,9%.

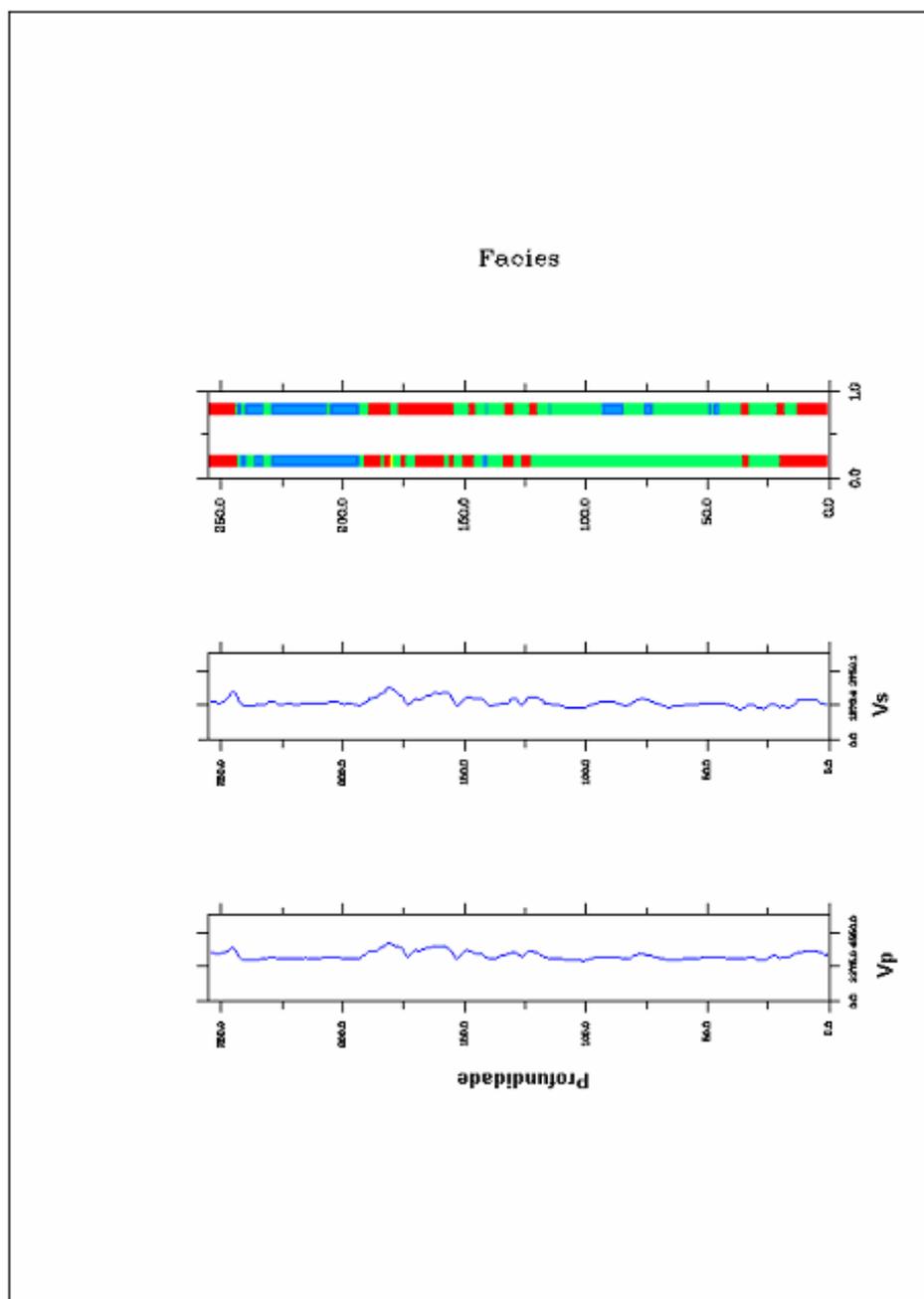


Fig 7.10 – Classificação do poço w1 com V_p e V_s . O índice de acerto obtido foi de 79,3%.

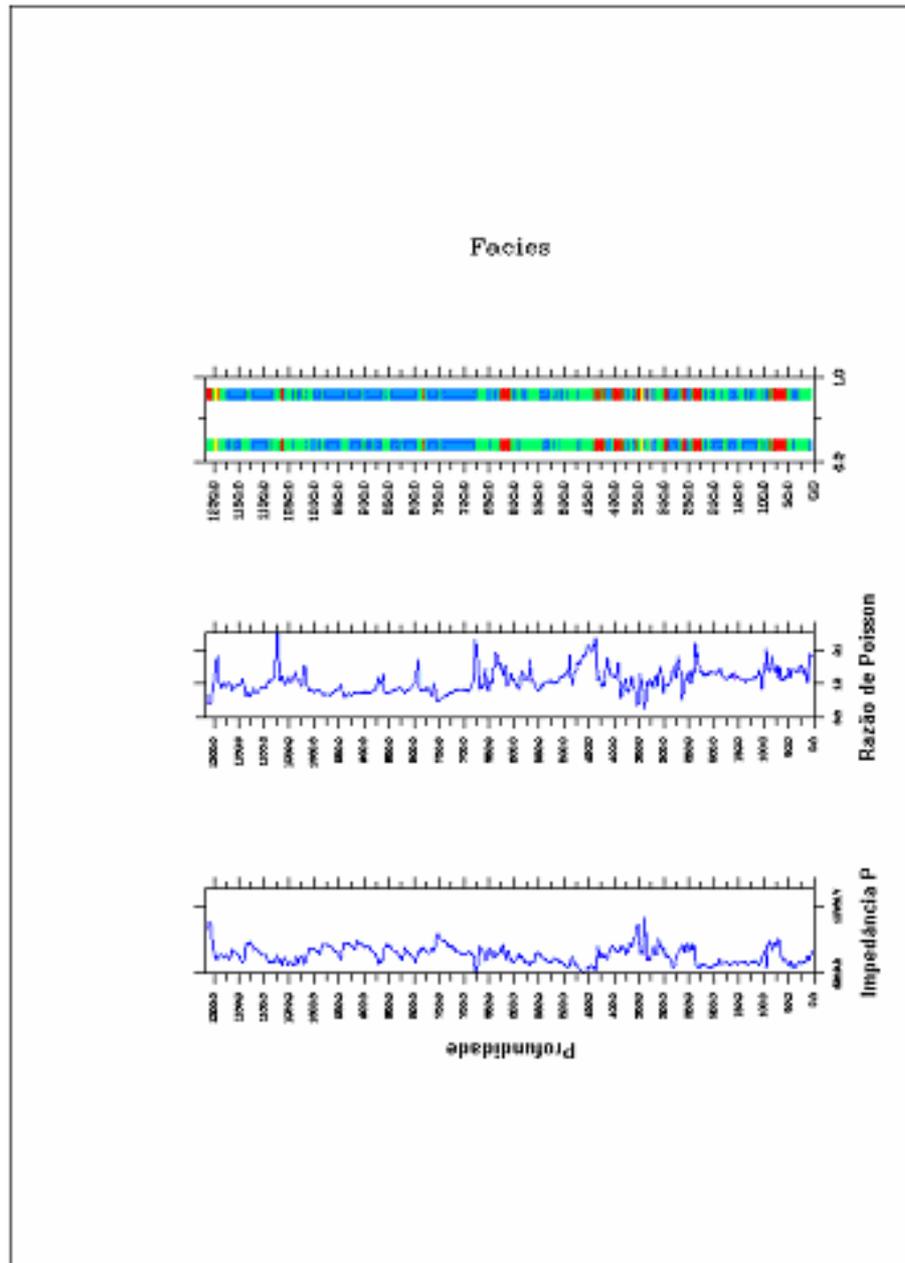


Fig 7.11 – Classificação do poço w2 com impedância-P e razão de Poisson. O índice de acerto obtido foi de 83,3%.

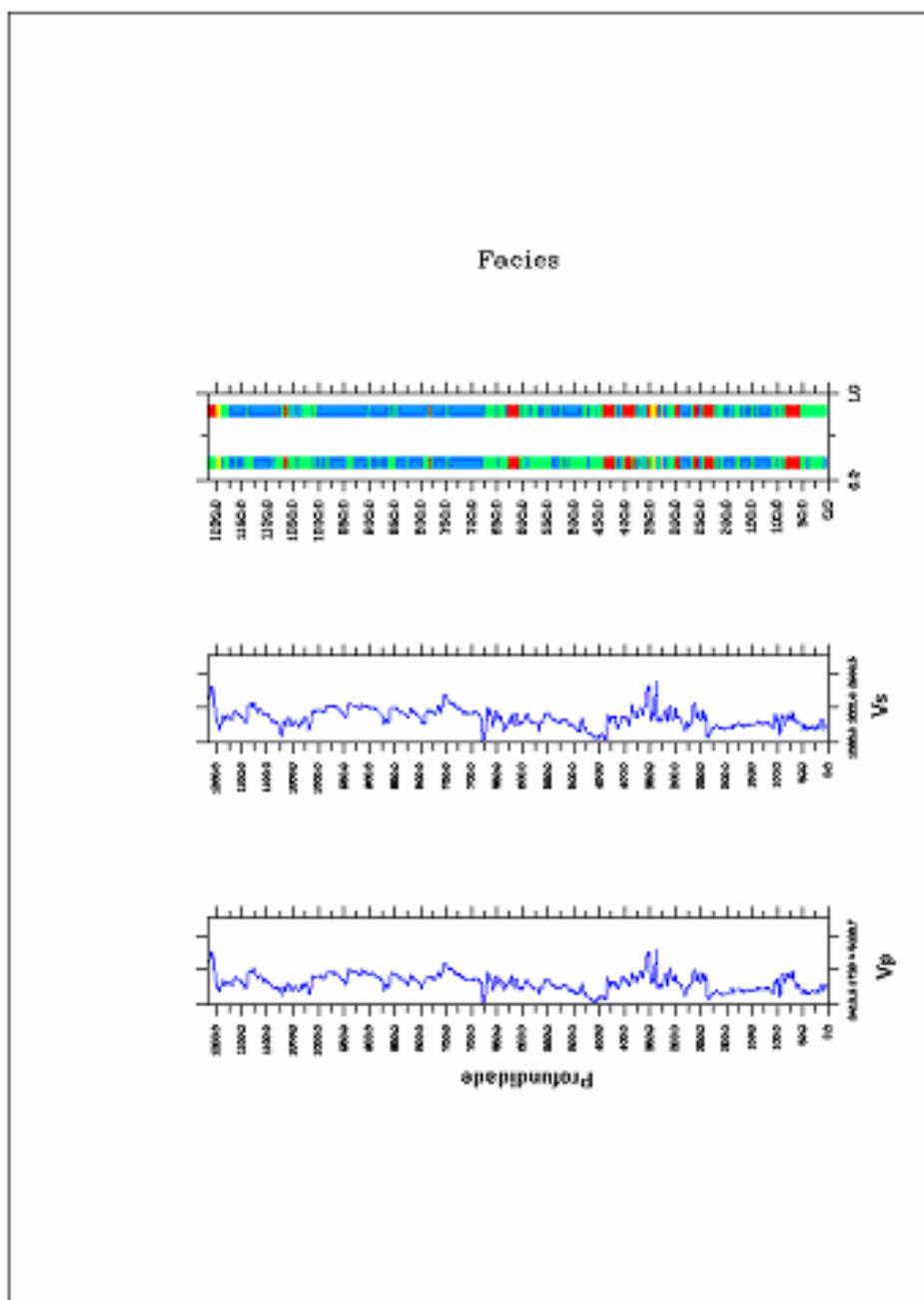


Fig 7.12 – Classificação do poço w2 com Vp e Vs. O índice de acerto obtido foi de 77,3%.

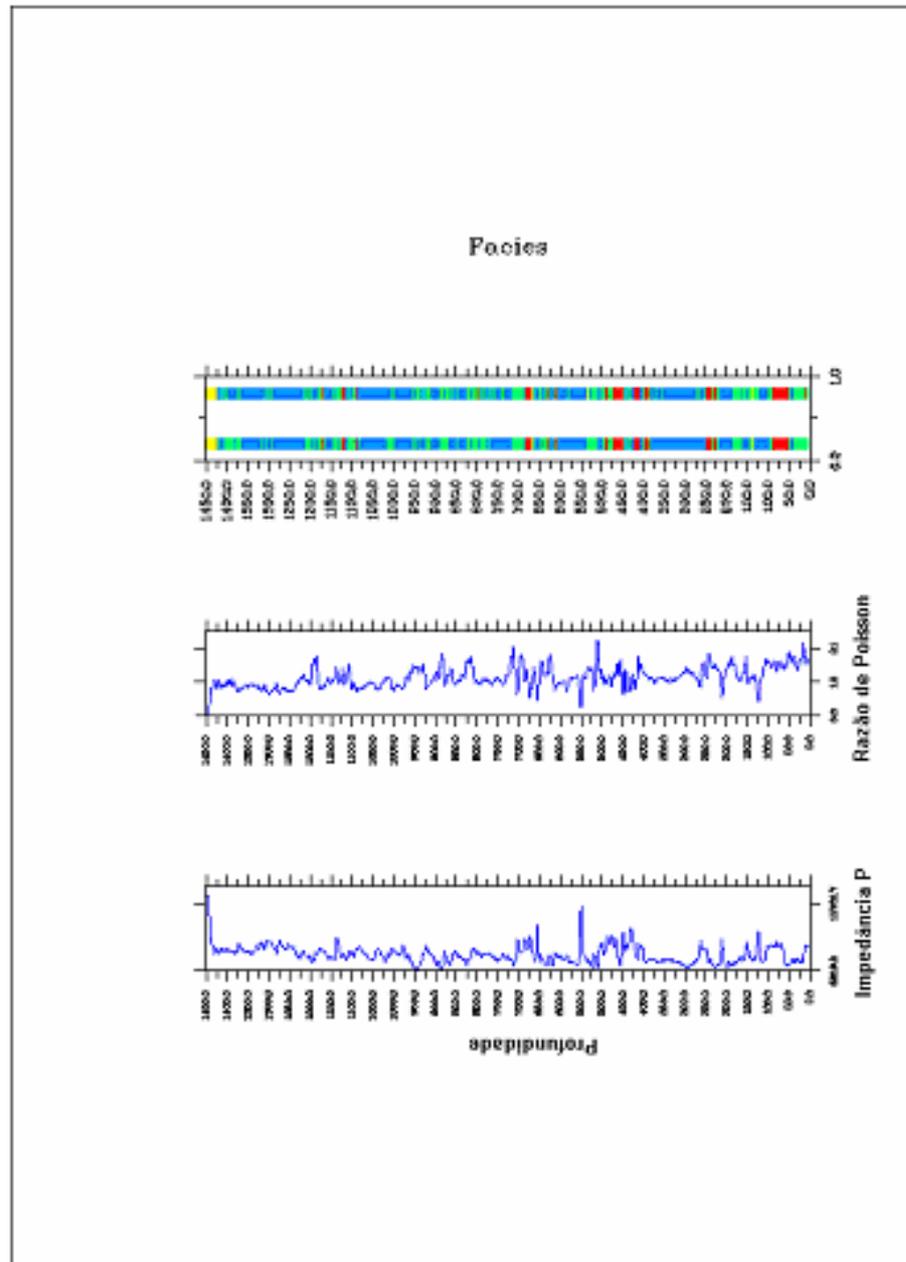


Fig 7.13 – Classificação do poço w3 com impedância-P e razão de Poisson. O índice de acerto obtido foi de 85%

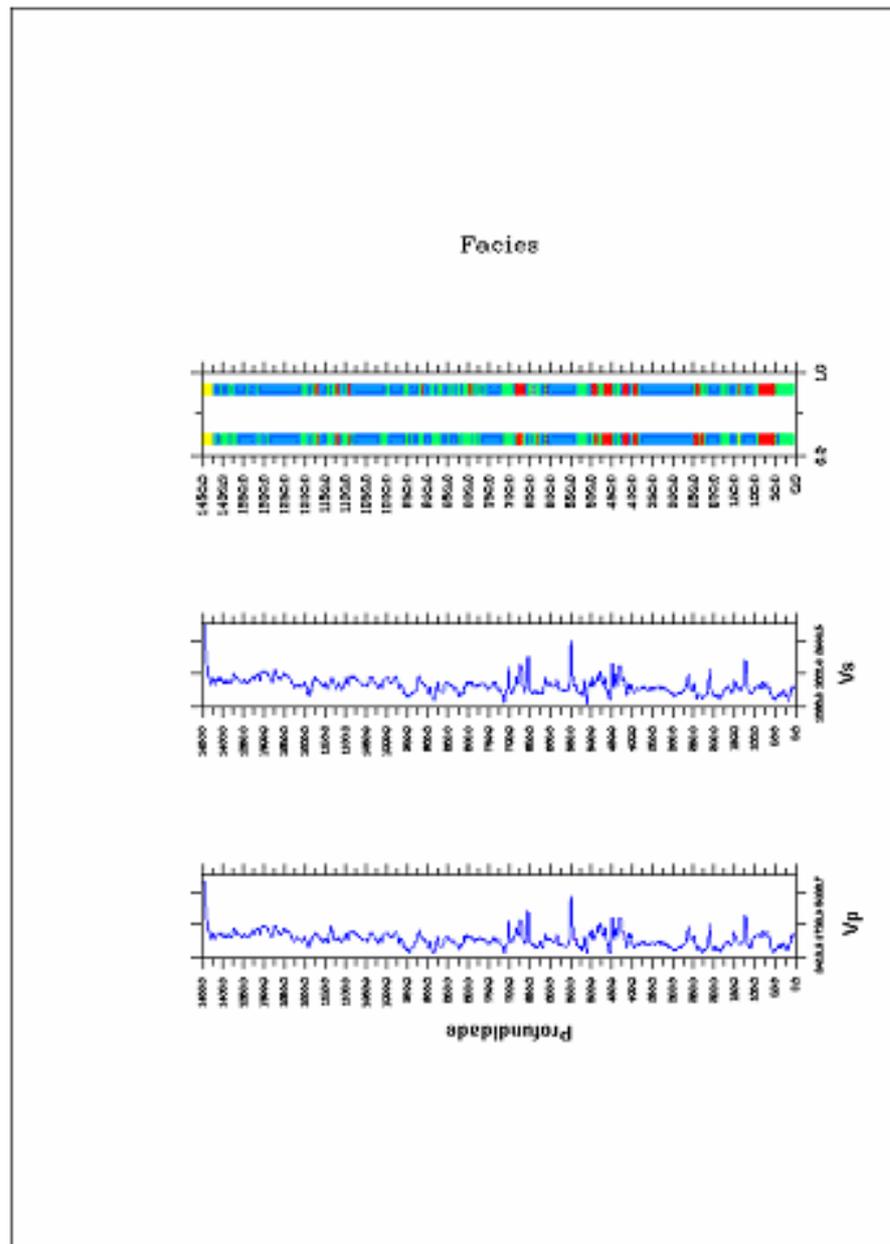


Fig 7.14 – Classificação do poço w3 com Vp e Vs. O índice de acerto obtido foi de 82,2%

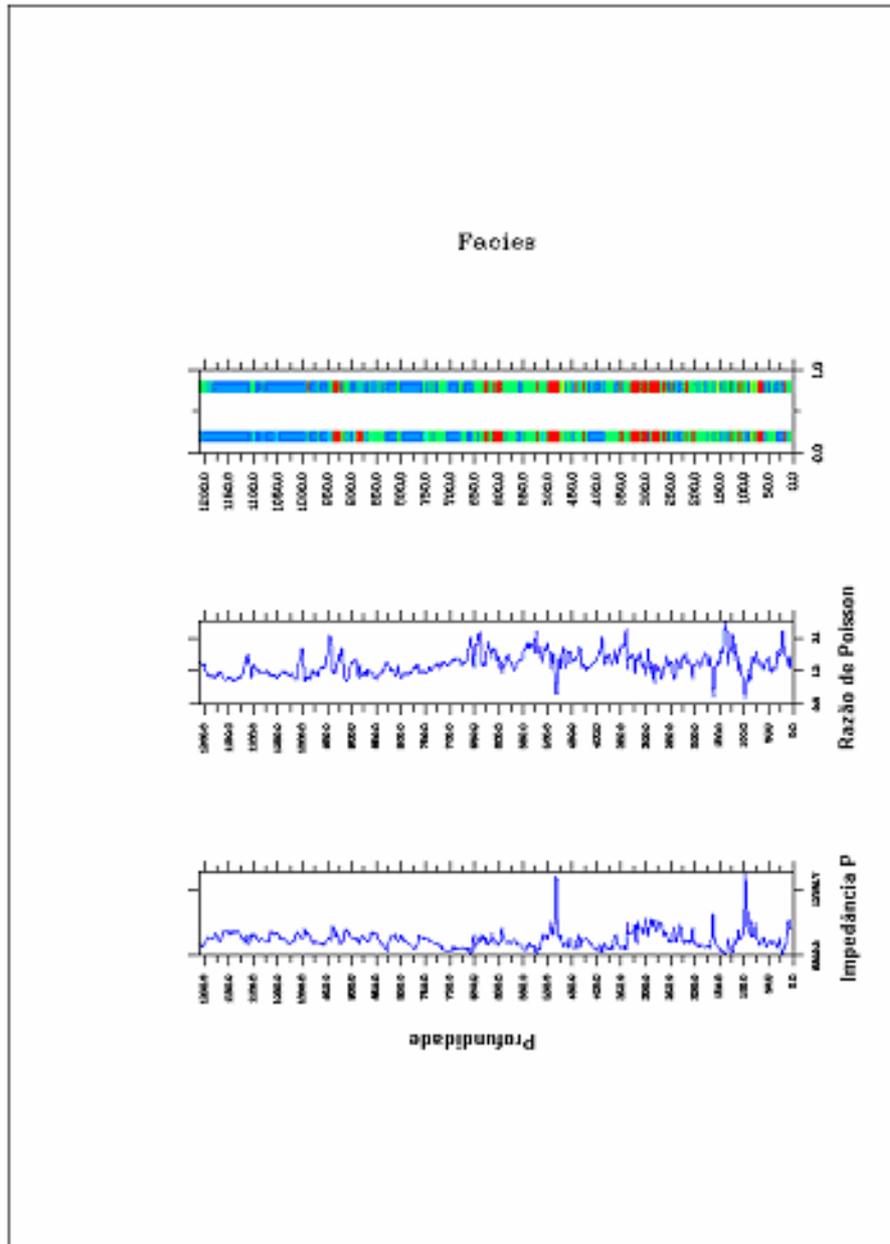


Fig 7.15 – Classificação do poço w4 com impedância-P e razão de Poisson. O índice de acerto obtido foi de 77,6%.

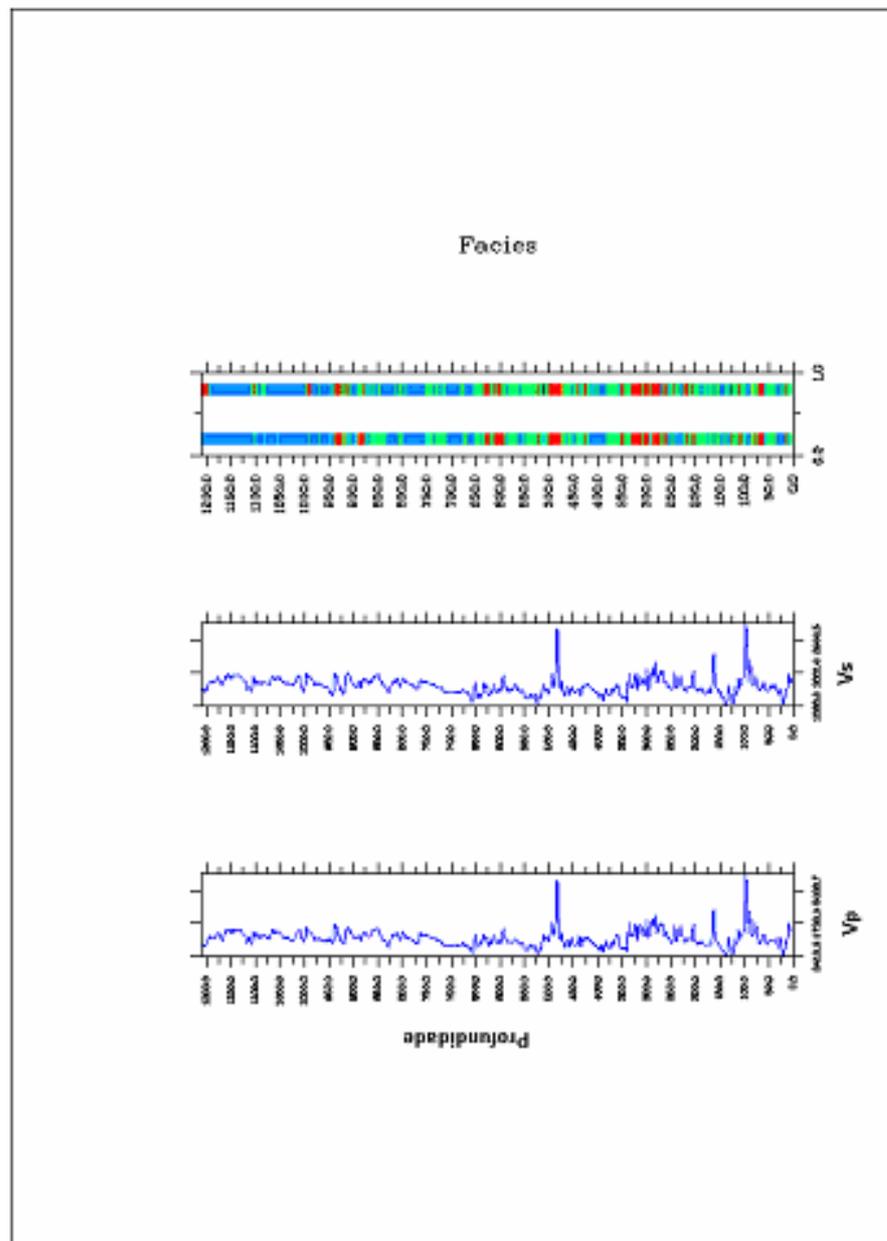


Fig 7.16 – Classificação do poço w4 com Vp e Vs. O índice de acerto obtido foi de 74,5%

7.6 – Aplicação em dados sísmicos utilizando litofácies definidas por intérprete

Utilizando os poços de w1 a w4 como dado de treinamento foi classificada uma linha sísmica. Foram utilizados os parâmetros V_p e V_s . O dado sísmico foi “mutado” na base e no topo do reservatório.

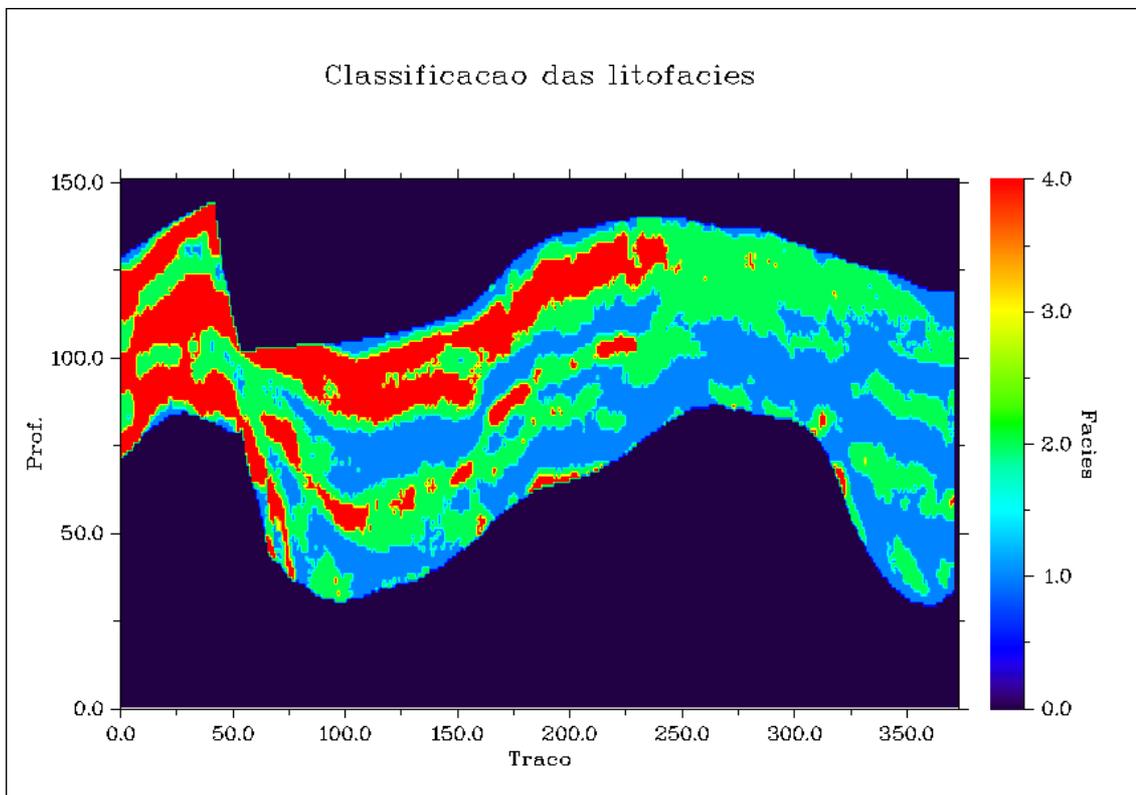


Fig 7.17 – Classificação da i_line 1 com V_p e V_s .

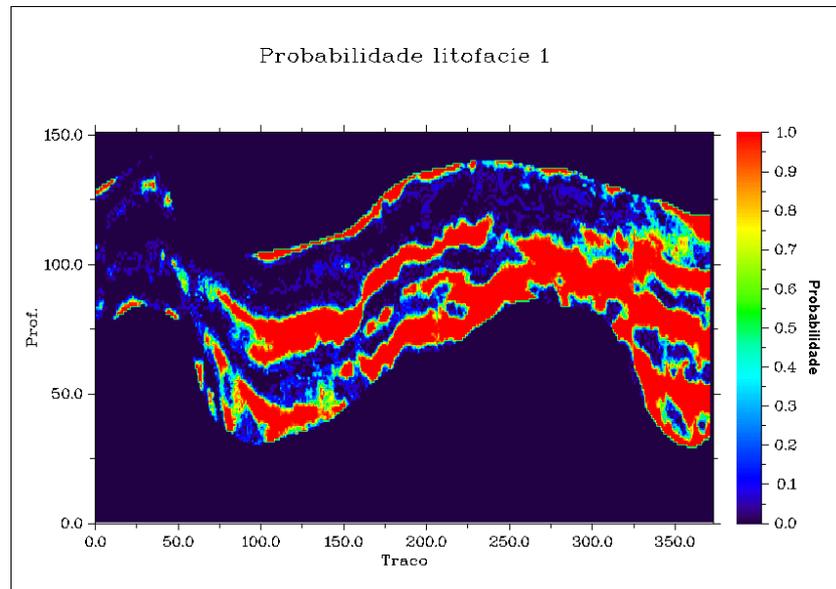


Fig 7.18 – Medida de incerteza para a litofácies 1 na i_line 1

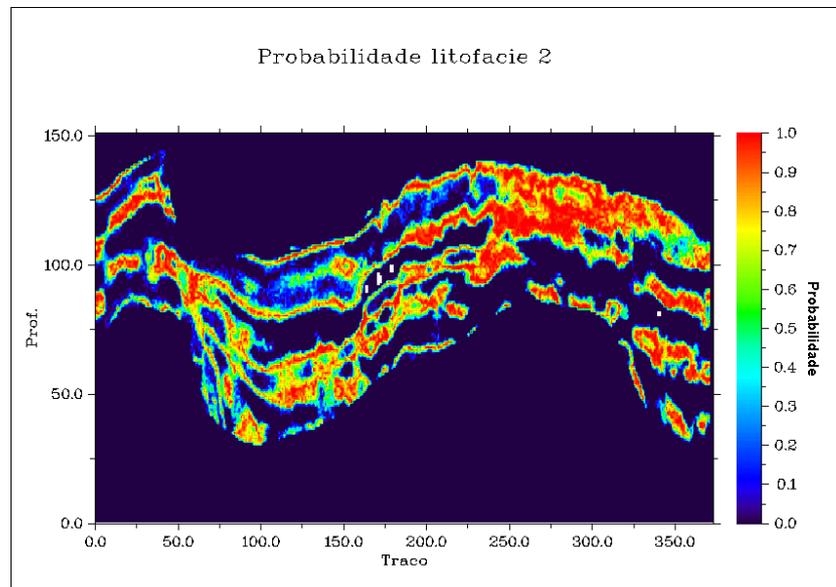


Fig 7.19 – Medida de incerteza para a litofácies 2 na i_line 1

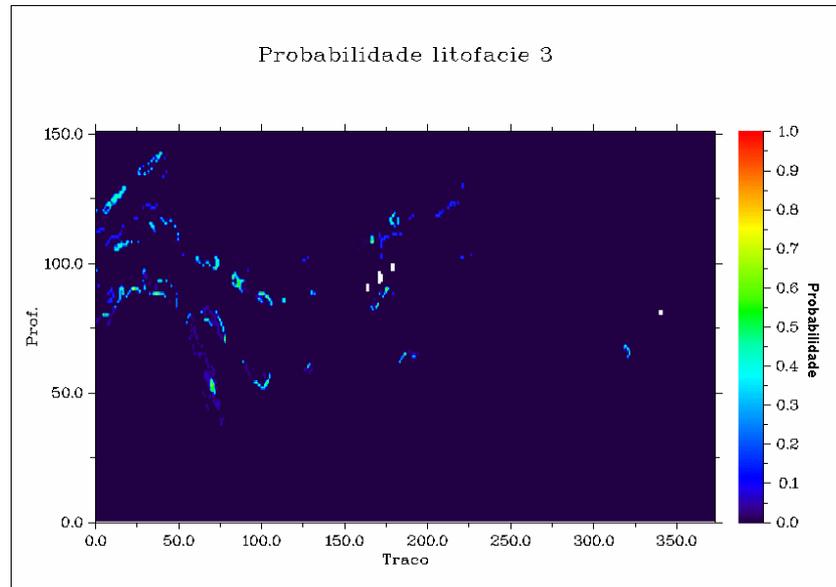


Fig 7.20 – Medida de incerteza para a litofácies 3 na i_line 1.

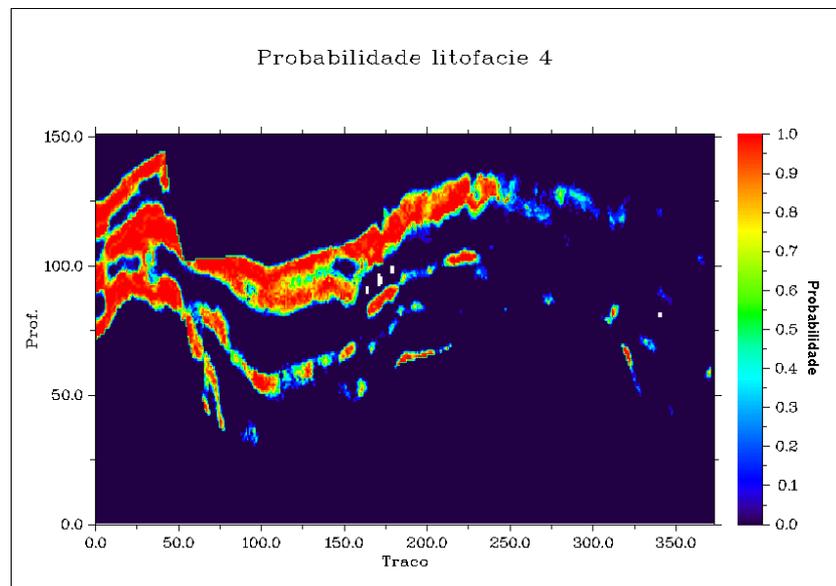


Fig 7.21 – Medida de incerteza para a litofácies 4 na i_line 1

As figuras acima apresentam o resultado final da classificação, a Figura 7.17 mostra as litofácies identificadas nesta linha e as figuras 7.18 à 7.21 as medidas de incerteza associadas a classificação de cada litofácies.

7.7 – Aplicação em dados sísmicos utilizando litofácies definidas aplicando EM

A metodologia foi aplicada no mesmo dado sísmico utilizado na seção anterior. A primeira etapa foi realizada utilizando o algoritmo EM nos quatro poços disponibilizados (w_1, w_2, w_3 e w_4). Os perfis utilizados como entrada foram os perfis de raio gama, porosidade e densidade. Foram identificados seis agrupamentos naturais nas amostras, que representam seis litofácies distintas. Para a estimativa inicial dos parâmetros do EM foi seguido o critério utilizado nas seções 6.4 e 6.5.1. Abaixo são apresentados os poços interpretados de forma automatizada através do uso do EM.

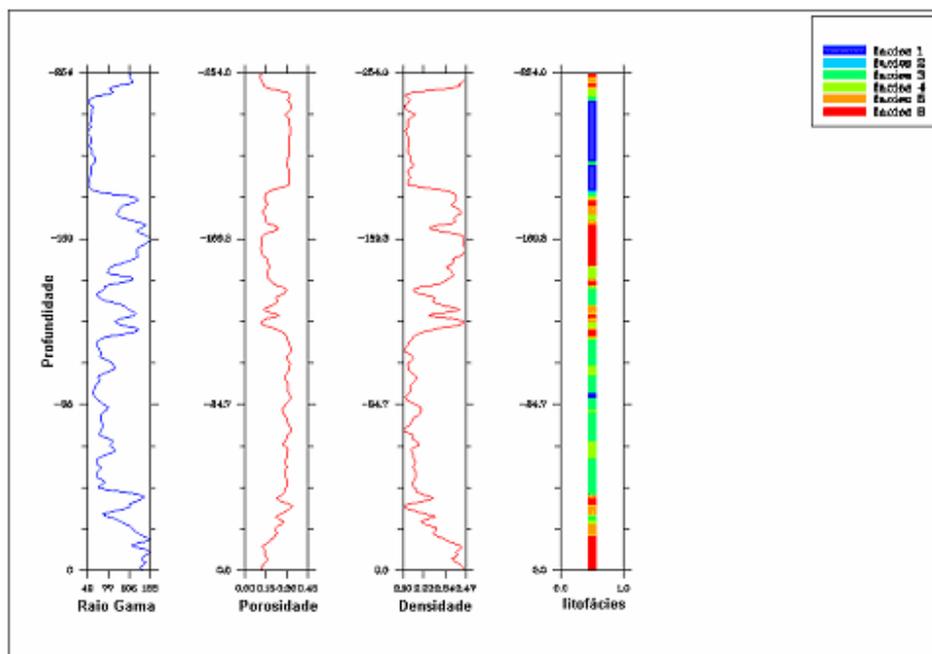


Fig. 7.22 – Poço w1 classificado automaticamente com o uso do EM.

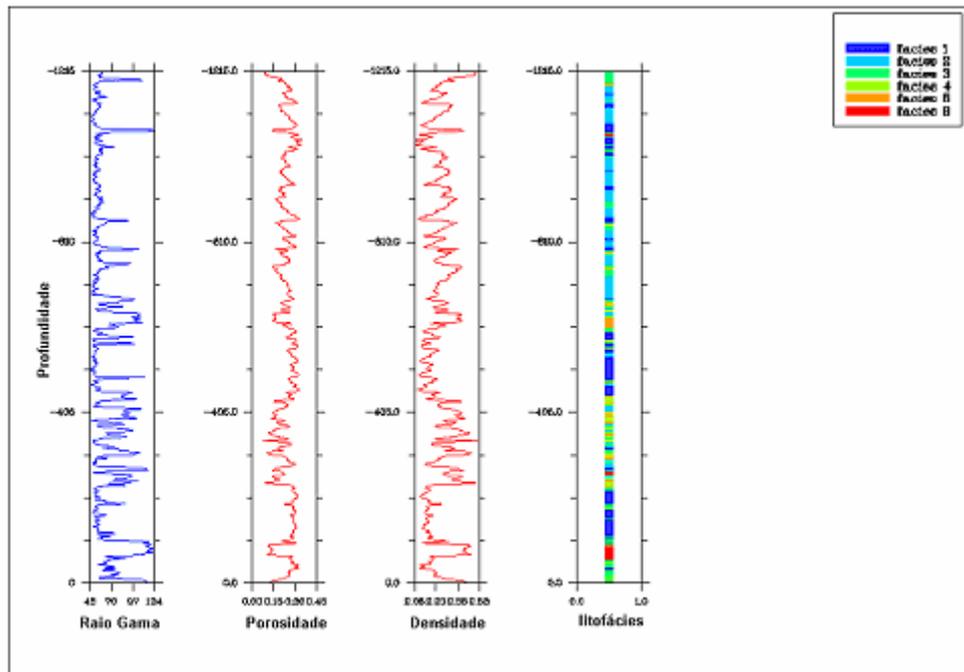


Fig. 7.23 – Poço w2 classificado automaticamente com o uso do EM.

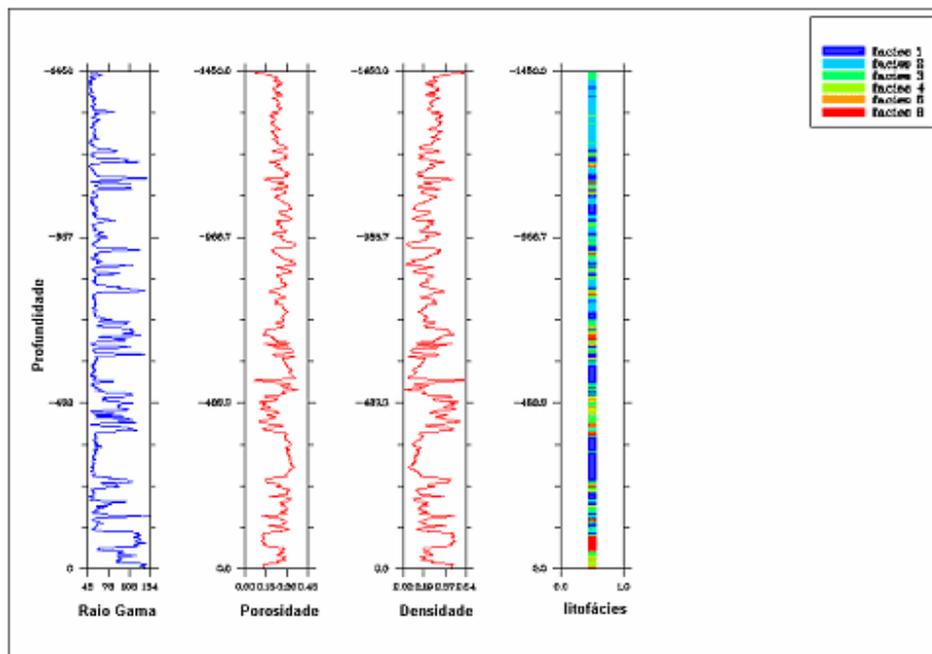


Fig. 7.24 – Poço w3 classificado automaticamente com o uso do EM.

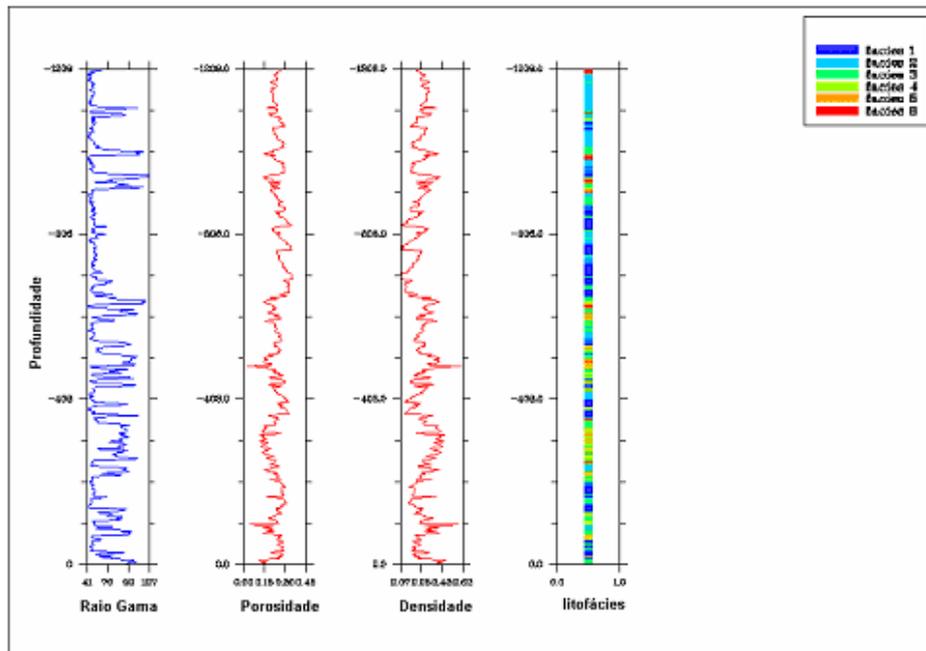


Fig. 7.25 – Poço w4 classificado automaticamente com o uso do EM.

7.7.1 – Análise petrofísica e comparação das litofácies identificadas com as interpretadas

Esta etapa consiste em realizar uma análise petrofísica e de física de rocha para “rotular” as litofácies identificadas. Primeiramente é explorado o comportamento das litofácies no domínio V_p - V_s . A Figura 7.26 é um “crossplot” das litofácies identificadas juntamente com alguns modelos teóricos. A Figura 7.27 representa as litofácies definidas por intérprete.

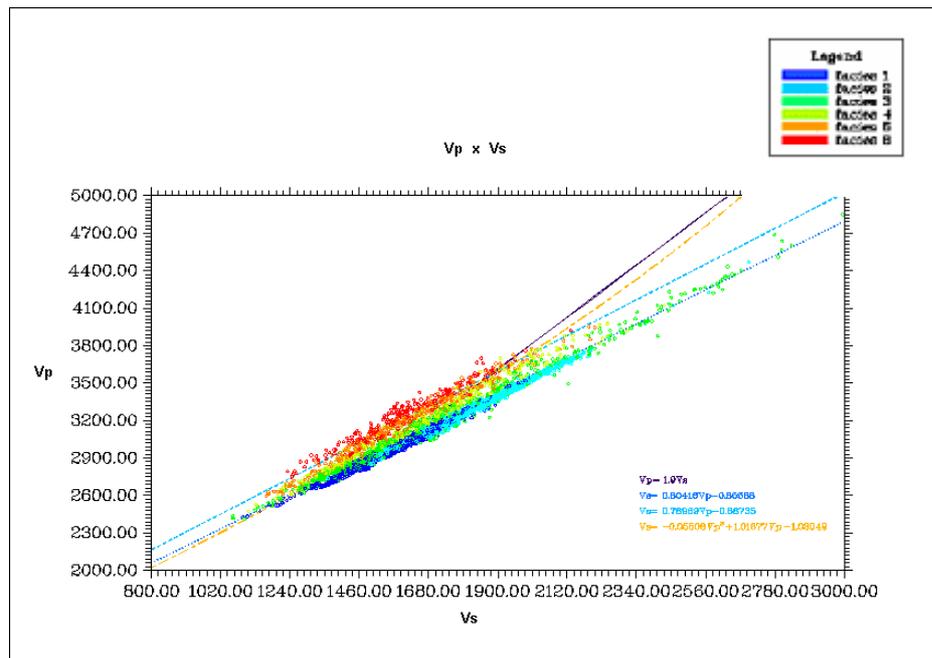


Fig 7.26 – Litofácies identificadas automaticamente no domínio V_p e V_s

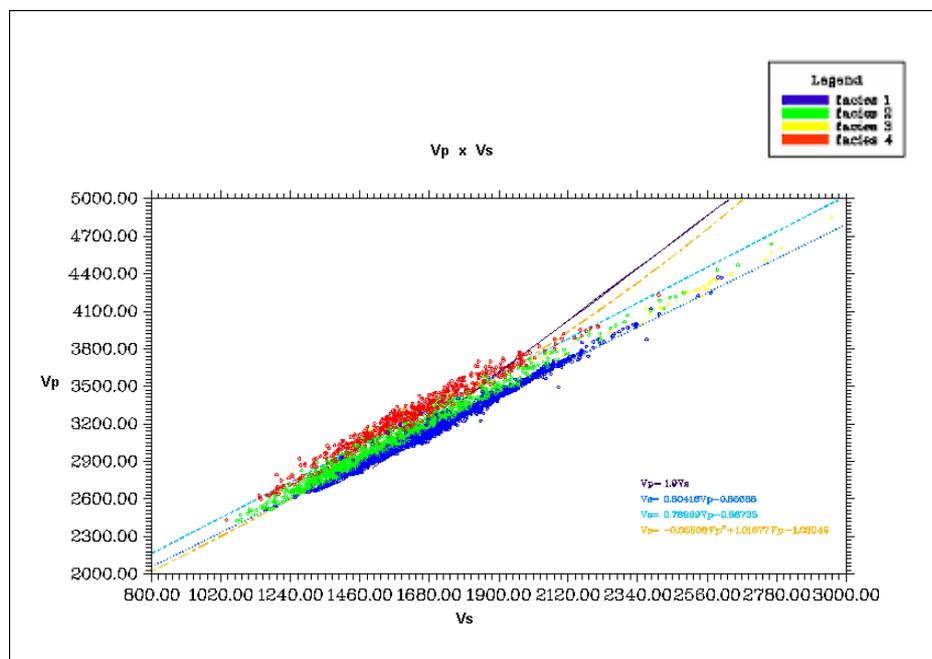


Fig 7.27 – Litofácies identificadas por intérprete no domínio V_p e V_s

Comparando os resultados da interpretação com os resultados do EM no domínio Vp e Vs, observa-se que:

- a) As litofácies 5 e 6 (identificadas automaticamente) seguem a tendência da linha teórica dos folhelhos e representam a litofácies 4 definida por intérprete.
- b) As litofácies 1 e 2 (identificadas automaticamente) seguem a tendência da linha teórica das areias e representam a litofácies 1 definida por intérprete.
- c) As litofácies 3 e 4 (identificadas automaticamente) localizadas entre a linha teórica das areias e do folhelhos representam o arenito com presença de argila (areia suja) definido como litofácies 2, pelo intérprete.
- d) A litofácies 3 definida pelo intérprete como arenito cimentado, possui um número pequeno de amostras no dado de treino, que embora possua propriedades distintas (alto valor de Vp e Vs) estas não foram suficientes para serem consideradas uma categoria distintas pelo EM, desta forma, acabaram sendo incorporadas pela classe da litofácies 4 (identificada automaticamente).

Foi analisado o comportamento das litofácies identificadas no domínio K (incompressibilidade) e porosidade.

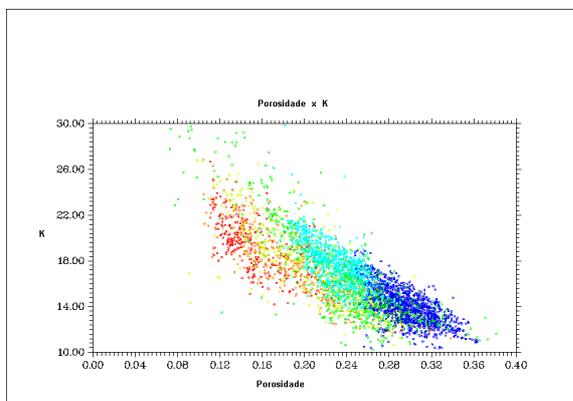


Fig 7.28 – Litofácies identificadas automaticamente no domínio K - Porosidade

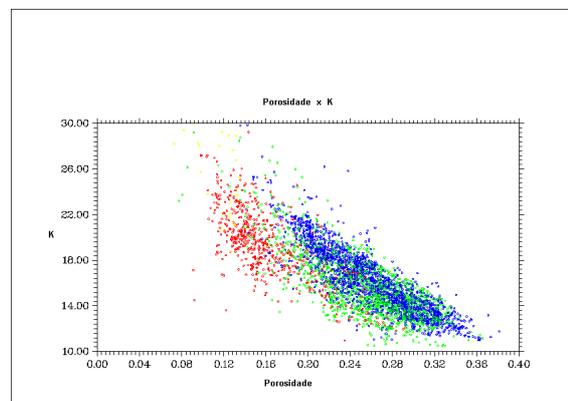


Fig 7.29 – Litofácies identificadas por intérprete no domínio K - Porosidade

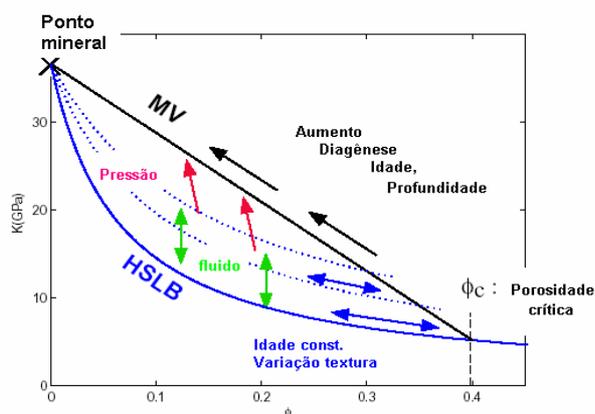


Fig 7.30 – Tendência das rochas siliciclásticas no domínio K-Porosidade (Takahashi, 2000)

Pode-se perceber, analisando os gráficos acima, que os agrupamentos gerados pela etapa de identificação de litofácies é bastante coerente com a classificação feita pelo intérprete, porém devido ao fato do agrupamento definir categorias baseadas em agrupamentos naturais, existentes nos dados, a identificação de forma automatizada gerou subdivisões das litofácies identificadas por intérprete.

Observando a Figura 7.30, nota-se que esta subdivisão foi basicamente devido às variações texturais e diagenéticas existentes dentro de uma mesma categoria.

Com a finalidade de conhecer qual o atributo que teve a maior contribuição para a discriminação (subdivisões geradas em relação ao dado interpretado e o reconhecido automaticamente), foram gerados gráficos dos três atributos utilizados como entrada do EM (raio gama, densidade e porosidade). As figuras 7.31 a 7.36 ilustram essa análise.

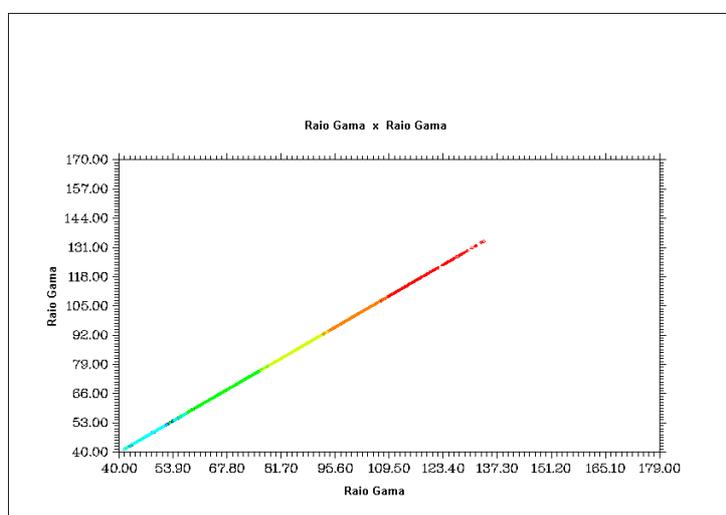


Fig 7.31 – Litofácies identificadas automaticamente no domínio raio gama x raio gama

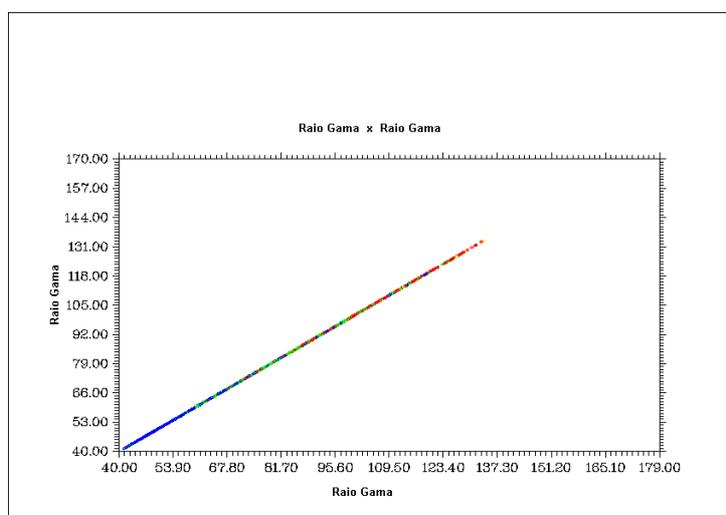


Fig 7.32 – Litofácies interpretadas no domínio raio gama x raio gama

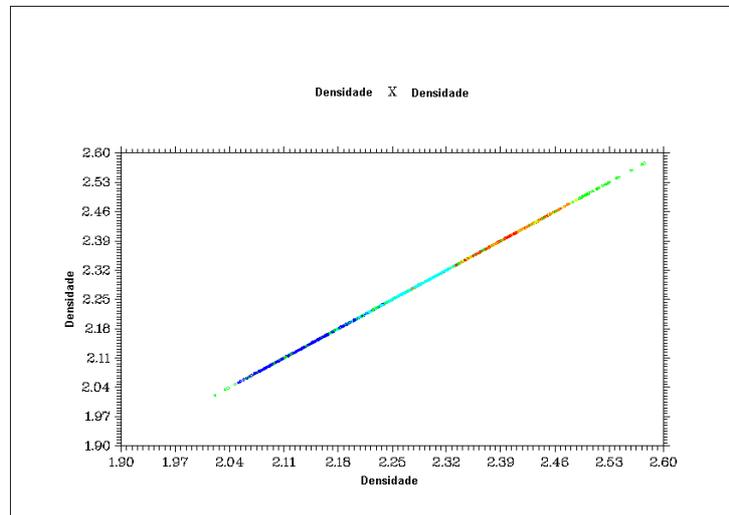


Fig 7.33 – Litofácies identificadas automaticamente no domínio densidade x densidade

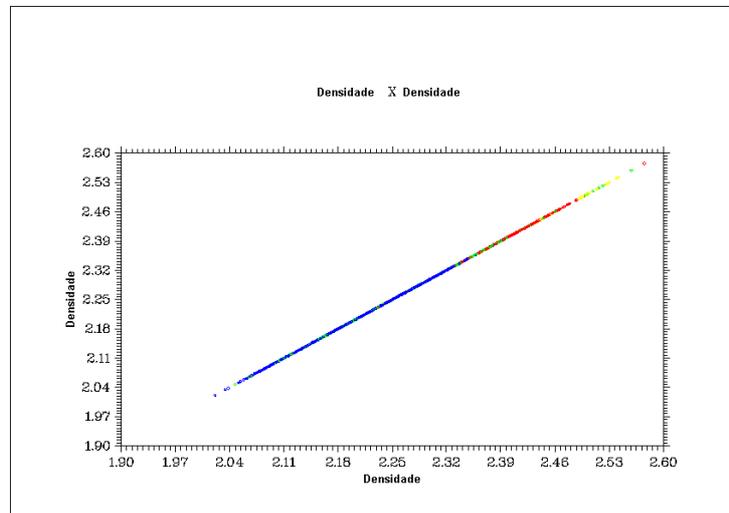


Fig 7.34 – Litofácies interpretadas no domínio densidade x densidade

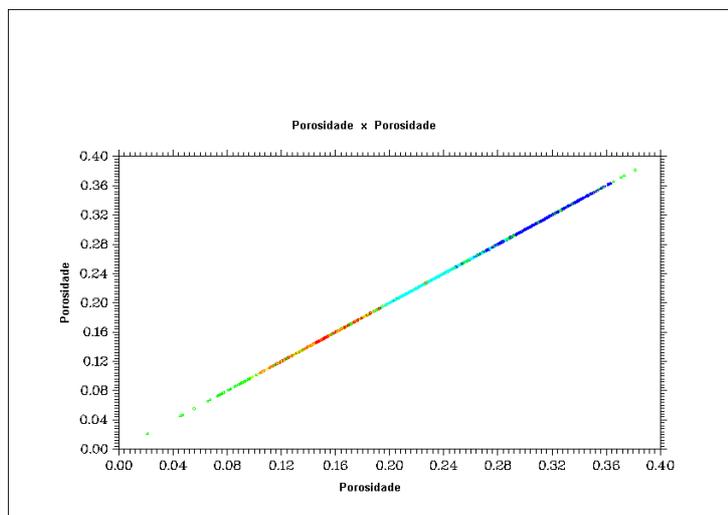


Fig 7.33 – Litofácies identificadas automaticamente no domínio porosidade x porosidade

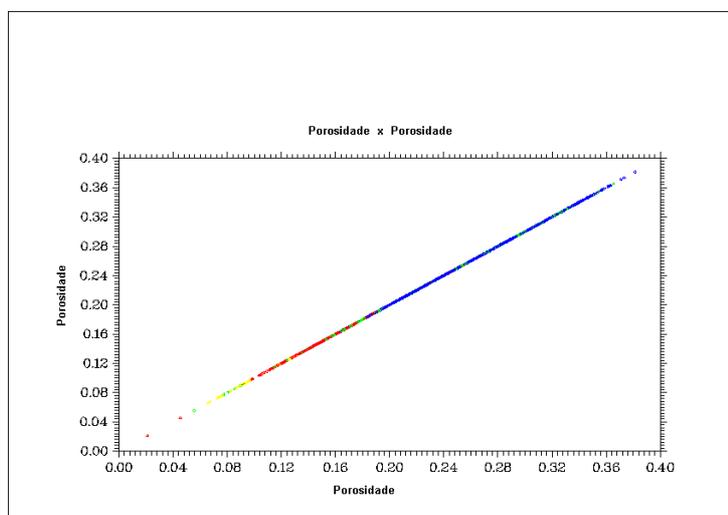


Fig 7.34 – Litofácies interpretadas no domínio porosidade x porosidade

Com relação às litofácies 1 e 2 identificadas automaticamente (que corresponde a litofácies 1 interpretada) pode-se observar que embora possuam valores intervalares iguais para o raio gama (aproximadamente entre 40 e 55 API), possuem valores distintos para densidade (aproximadamente entre 2,06 e 2,19 para litofácies 1 e entre 2.20 g/cc e 2.34 g/cc para litofácies 2) e porosidade (aproximadamente entre 27% e 36% para litofácies 1 e entre 19% e 27% para litofácies 2). Sendo assim a subdivisão gerada pelo agrupamento não supervisionado, neste caso, subdividiu a litofácies 1 interpretada em relação a existência de duas faixas distintas de valores para densidade e porosidade.

Observando as litofácies 3 e 4 identificadas automaticamente (que correspondem a litofácies 2 interpretada) observa-se que existem duas faixas de valores bem distintas para o raio gama (aproximadamente entre 55 e 78 API para litofácies 3 e entre 78 a 94 para litofácies 4), já para os valores de densidade existem valores distintos para litofácies 3 (entre 2.47 g/cc e 2.55 g/cc) e uma faixa de valores sobrepostos (entre 2.42 g/cc e 2.47 g/cc) para as litofácies 3, 4. Em relação a porosidade existe uma faixa de valores (aproximadamente entre 10% e 19%) onde existe sobreposição das litofácies 3, 4, 5 e 6. Pode-se concluir que, para a distinção destas duas fácies (litofácies 3 e 4) os atributos que mais influenciaram na discriminação, em ordem de influência, foram respectivamente raio gama, densidade e porosidade.

As litofácies 5 e 6 identificadas automaticamente (que correspondem a litofácies 4 interpretada) possuem uma clara distinção de valores para o raio gama parcial sobreposição para densidade e porosidade.

7.7.2 – Análise de incertezas para a escolha dos parâmetros com menor risco de classificação

Embora, para a seção sísmica em estudo, existam apenas dados invertidos de V_p e V_s , foi realizada a análise de incerteza para todos os pares de atributos. O

7.7.3 – Classificação de seção sísmica com as litofácies identificadas automaticamente

Foi aplicado à seção de V_p e V_s , da área de estudo, o classificador de máxima probabilidade a posteriori. Os dados de poços utilizados como dados de treinamento foram tempo de trânsito da onda P e o tempo de trânsito da onda S dos poços w1 a w4. Os tempos de trânsito foram transformados em V_p e V_s e filtrados utilizando filtro de Backus para uma janela de 5 metros. Após a preparação dos dados de treinamento, a seção sísmica foi classificada.

A Figura 7.36 apresenta o resultado final da classificação. As figuras 7.37 a 7.41 apresentam as medidas de incerteza relacionadas a classificação de cada litofácies (probabilidade de cada litofácies).

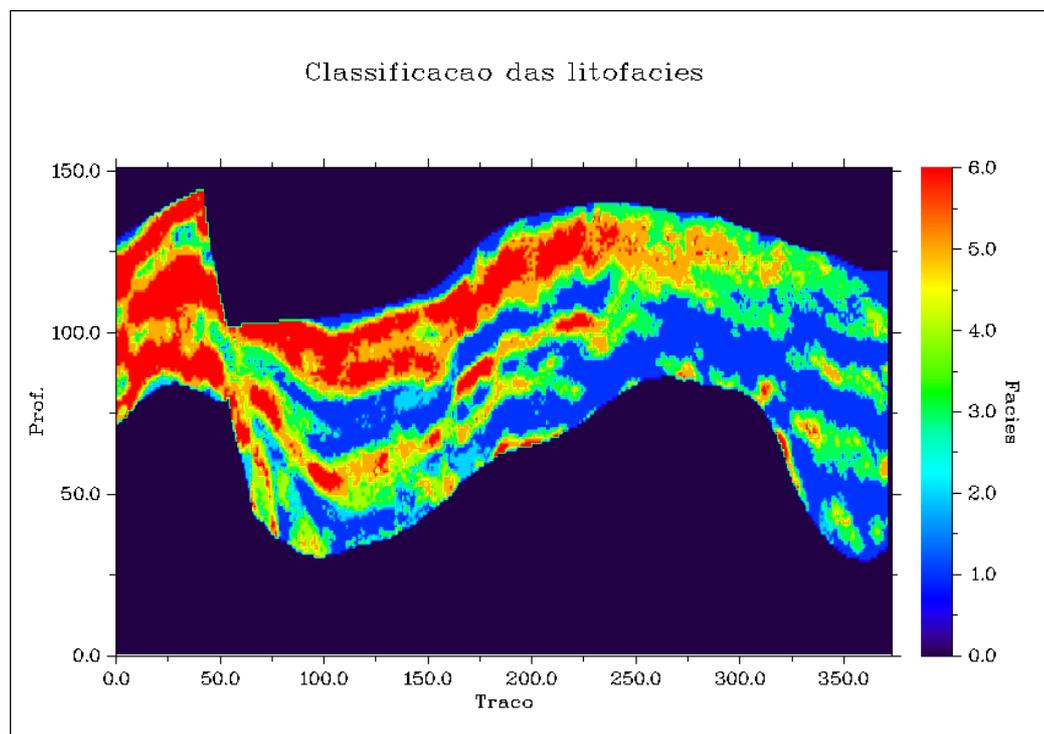


Fig 7.36 – Resultado da classificação da seção sísmica com as litofácies identificadas automaticamente

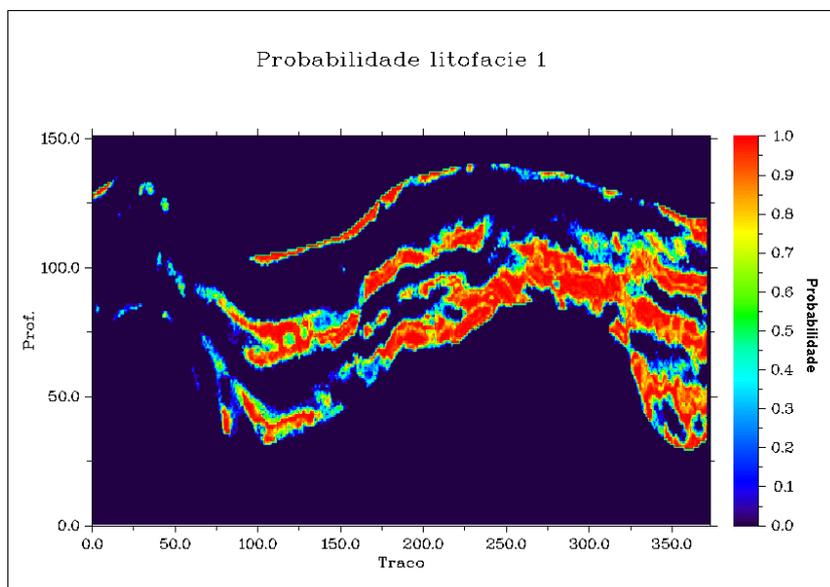


Fig 7.37 – Medida de incerteza para a litofácies 1 identificada automaticamente

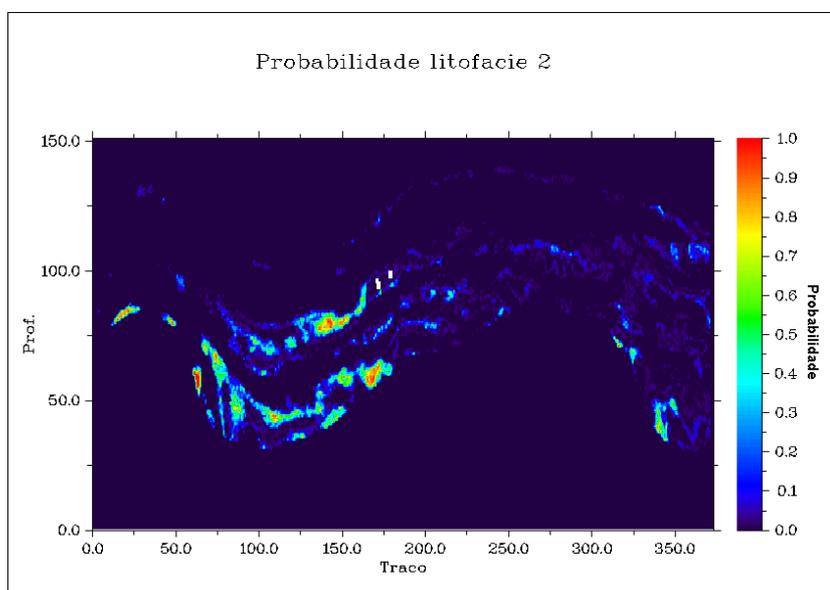


Fig 7.38 – Medida de incerteza para a litofácies 2 identificada automaticamente

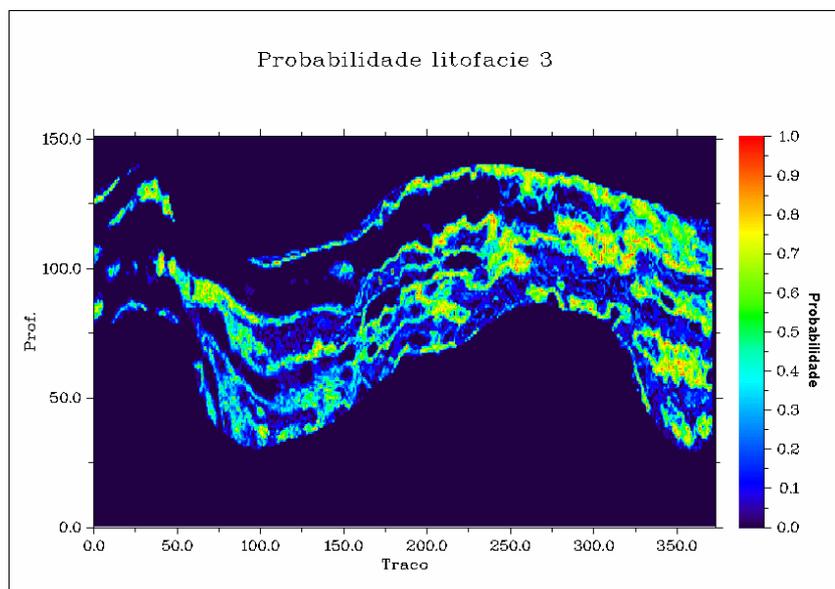


Fig 7.39 – Medida de incerteza para a litofácies 3 identificada automaticamente

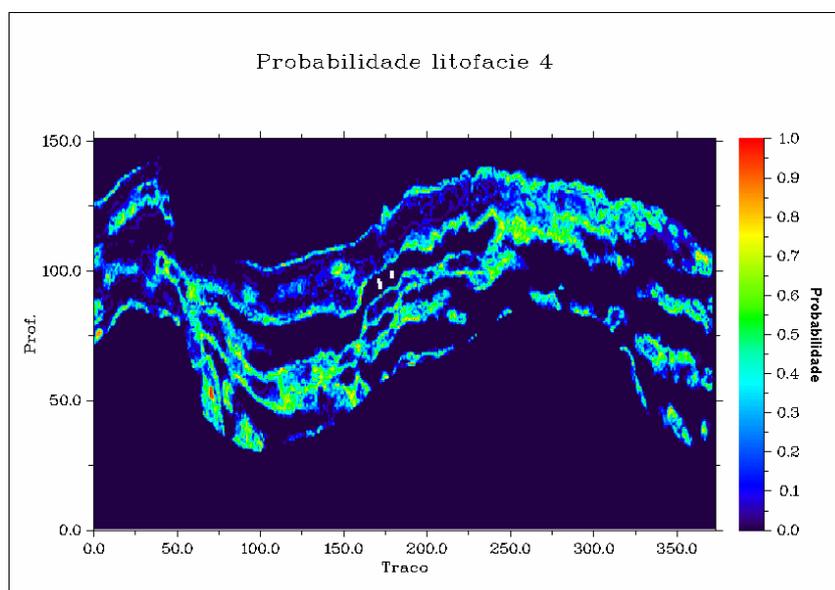


Fig 7.40 – Medida de incerteza para a litofácies 4 identificada automaticamente

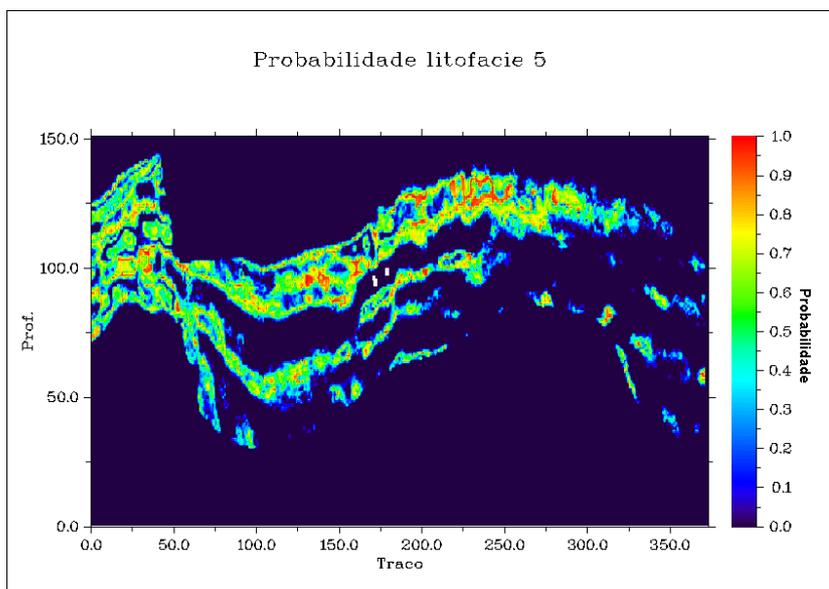


Fig 7.41 – Medida de incerteza para a litofácies 5 identificada automaticamente

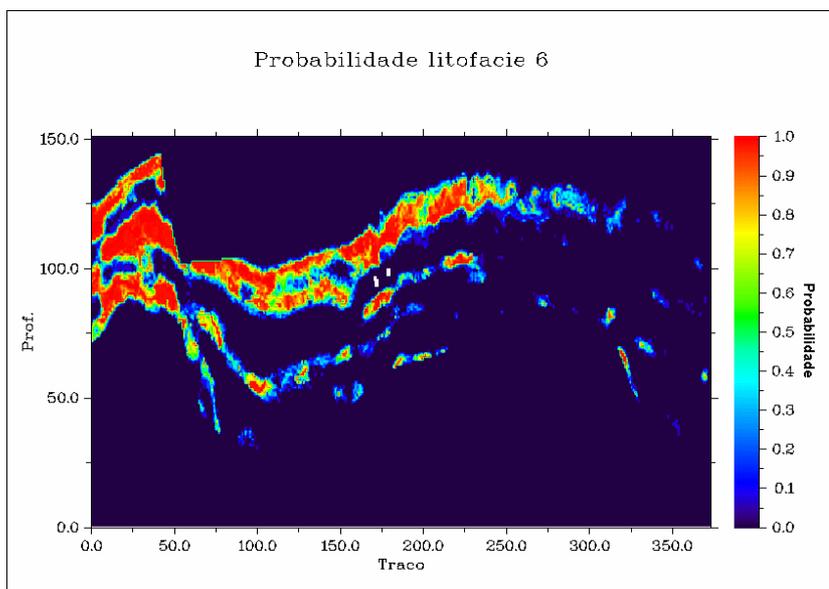


Fig 7.42 – Medida de incerteza para a litofácies 6 identificada automaticamente

Os resultados mostram a classificação das litofácies identificadas juntamente com as medidas de incerteza associadas à classificação de cada litofácies. O resultado aqui apresentado, é o resultado final da metodologia.

7.8 – Conclusões

Neste capítulo foi demonstrada a aplicação da metodologia em um conjunto de dados reais. Pode-se concluir, com base na classificação dos poços interpretados, que os resultados obtidos são de ótima qualidade (acima de 80% de índice de sucesso, na média).

Em relação a identificação automática de litofácies utilizando o algoritmo EM, foi demonstrado que as seis litofácies reconhecidas automaticamente, comparado com as quatro litofácies identificadas pelo intérprete, são na verdade subdivisões das litofácies interpretadas, exceto o fato da litofácies 3 (arenito cimentado) cujo número de amostras existentes no dado serem escassas não sendo possível sua detecção como uma categoria distinta pelo algoritmo EM. A análise petrofísica realizada demonstrou coerência entre as litofácies interpretadas, as identificadas de forma automática e os modelos de física de rocha.

Em relação à análise de incerteza, comprova-se que a escolha dos atributos com base na análise é uma etapa fundamental da metodologia para a diminuição dos riscos de classificação.

Na classificação sísmica, pode-se perceber que todas as litofácies identificadas no poço foram classificadas na seção. Percebe-se a presença notória das litofácies 1 e 6 na seção classificada.

Capítulo 8

Conclusões

A metodologia apresentada fornece uma estrutura integrada de trabalho que incorpora diversas fontes de informação e apresenta como resultado final, a classificação de litofácies em conjunto com uma medida de incerteza associada a cada amostra classificada. Os resultados obtidos com os dados reais são de ótima qualidade (acima de 80% de acerto na média) quando avaliado o método com dados de poços. Estes resultados são encorajadores para que a metodologia desenvolvida se torne parte integrante dos processos de caracterização de reservatório.

A metodologia desenvolvida possibilita que litofácies sejam identificadas em dados de poço ou em dados sísmicos. Alguns detalhes merecem ser ressaltados sobre a metodologia.

Com relação ao dado de treinamento, existem duas formas de criá-los:

- a) Com o uso de aprendizado supervisionado – onde o dado utilizado como dado de treinamento é proveniente de poços que foram interpretados por geólogos intérpretes.

- b) Aprendizado não supervisionado – onde o dado de treinamento é gerado pelo algoritmo EM, identificando automaticamente as litofácies existentes com base em padrões naturais dos dados.

Com relação às duas diferentes formas de gerar o dado de treinamento conclui-se que:

- i) O uso do EM torna o processo classificatório totalmente automatizado, o que gera ganho de produtividade na caracterização de reservatórios. A desvantagem desta abordagem é quando existe um número muito limitado de amostras de uma determinada litofácies (caso de uma camada muito fina de uma determinada litologia), o EM não a reconhece como uma classe distinta, e as amostras são incorporadas por outra classe semelhante.
- ii) O uso de poços interpretados como dado de treinamento, dificilmente deixará de conter a interpretação de uma determinada litofácies, mesmo que ela se apresente em camadas muito finas, porém o tempo de interpretação para uma grande quantidade de poços é relativamente extenso.

Quando são classificados poços baseados em outros poços como treinamento, não existe a necessidade de realizar qualquer manipulação no dado. No caso da classificação de seções (ou cubos) sísmicos, o dado de treinamento deverá estar na mesma escala do dado sísmico, o “upscaling” deve ser realizado conforme descrito no capítulo 2 seção 2.4.

A criação do classificador de máxima probabilidade a posteriori, utilizando redes bayesianas, possibilita uma fácil integração de novas fontes de informação. No caso da rede elaborada para este trabalho, além de incorporar fontes de informação diversas, existe a agregação do método de k-vizinhos mais próximos, que eleva o índice de acerto do classificador, apesar de elevar consideravelmente o tempo de processamento quando seções sísmicas são classificadas.

Em relação a análise de incerteza para a escolha dos atributos utilizados no processo classificatório, o erro de Bayes demonstrou ser uma medida coerente para a realização de tal tarefa.

8.1 Recomendações para trabalhos futuros

Os resultados obtidos neste trabalho inspiram sua continuidade da seguinte forma:

- a) Após utilizar a metodologia proposta para identificar possíveis rochas reservatórias, utilizar substituição de fluido para saturar o dado de treinamento em óleo e em água e re-classificar as áreas anteriormente classificadas como rochas reservatórias, excluindo as demais litofácies, para a detecção de possíveis alvos exploratórios.
- b) Certas metodologias, como Loures e Moraes (2002), necessitam que modelos de física de rocha sejam calibrados com a finalidade de realizar inferência de porosidade e volume de argila. Com o uso da metodologia desenvolvida neste trabalho, como uma etapa anterior a etapa de inferência, possibilitará a calibração destes modelos para cada litofácies distinta, aumentando a confiabilidade da processo de inferência.
- c) O protótipo gerado em linguagem C, para a realização deste trabalho, pode ser paralelizado para aumentar a eficiência da metodologia.
- d) Uma etapa da metodologia que é realizada sem nenhum método de inteligência artificial é a “rotulação” das classes identificadas. Sugere-se que seja criada uma outra Rede Bayesiana, que com base em modelos de física de rocha, nomeie automaticamente os diferente tipos litológicos.

REFERÊNCIA BIBLIOGRÁFICA

- Avseth, P., Mukerji, T., Jørstad, A., Mavko, G., Veggeland, T. (2001) *Seismic reservoir mapping from 3-D AVO in a North Sea turbidite system*, *Geophysics*, 66 (4):1157–1176.
- Backus, G.E. (1962) *Long-wave elastic anisotropy produced by horizontal layering*, *J. Geophys.*, 67:4427-4440.
- Barnes, E.A. (2001) *Seismic attributes in your facies*, *CSEG Recorder*: 41-47.
- Bosh, M., Zamora, M., Utama, W. (2002) *Lithology discrimination from physical rock properties*, *Geophysics*, 67 (2):573–581.
- Castagna, J.P., Batzle, M.L, Eastwood, R.L. (1985) Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks, *Geophysics*, 50:571-581.
- Castagna, J.P., Batzle, M.L, Kan, T.K. (1993) *Rock Physics: the link between rock properties and AVO response; in Offset-dependent reflectivity: theory and practice of AVO analysis*, *Soc. Expl. Geophys*: 135-171.
- Corvi, P., Heffer, K., King, P., Tyson, S., Verly, G. (1992) *Reservoir characterization using expert knowled data and statistics*, *Oilfield review*, 4, (1).
- Cunha, E.C. (2002) *Identificação de Litofácies de Poços de Petróleo Utilizando um Método Baseado em Redes Neurais Artificiais*, Tese (Mestrado em Ciência da Computação), Campina Grande, Paraíba, Brasil, Universidade Federal de Campina Grande, 118 p.
- Dampster, A.N., Laird, Rubin, D. (1977) *Maximum likelihood estimation from incomplete data via the EM algorithm*, *J. Royal Statistical Soc. B*, 39:1–38.
- Dewar, J., Pickford, S. (2001) *Rock physics for the rest of us –an informal discussion*, *CSEG Recorder*, Calgary, Core Laboratories Company:43-49.

Doyen, M.P. (1998) *Porosity from seismic data: a geo-statistical approach*, Geophysics, 53 :1263.

Ferreira, I.M., Malagutti, D.M., Thedy, S.R., Castro, F.C.C. (2002) *Exploitation Evolution of a Mature Oil Field from Campos Basin Brazil*, SPE 77670.

Gassman, F. (1951) *Elastic waves through a packing of spheres*, Geophysics, 16:673-685.

Greenberg, M.L., Castagna, J.P. (1992), *Shear-Wave Velocity Estimation in Porous Rocks: Theoretical Formulation, Preliminary Verification and Application*, Geophysical Prospecting 40:195-209.

Han, D., Nur, A., Morgani, D. (1986) *Effects of porosity and clay content on wave velocities in sandstones*, Geophysics, 51. (2):2093-2107.

Hashin, Z., Shtrikman, S. (1963) *A variational approach to the elastic behaviour of multiphase materials*, J. Mech. Phys. Solids, 11:127-140.

Hearst, R., Joseph, N.H., Paillet, L.P.F. (2004) *Well logging for physical properties second edition*, Wiley-England.

Joseph, H. F., Rolph, A., Tatham, R., Black, W. (1998) *Multivariate Data Analysis* 5. ed.:Prentice Hall.

Kuster, G.T., Toksöz, N.M. (1974) *Velocity and attenuation of seismic waves in two-phase media, I. Theoretical formulations*, Geophysics, 39:607-618.

Loures, L., Moraes, F. (2002), *Reservoir Porosity Inference*, Society of Exploration Geophysicist Annual Meeting, Salt Lake City.

Lu, Ye., Minas, E. S., (2003) *Clustering on unobserved data using mixture of Gaussians*, York University, Department of computer science, Technical Report CS-2003-08.

Mahmoudian, F., Margrave, G. F. (2004), *AVO inversion of multicomponent data for P and S impedance*, CSEG National Convention.

Mavko, G., Mukerj, T. e Dvorkin, J., 1998, *The Rock Physics Handbook*, Cambridge University Press, 329 pp.

Mukerj, T. I, Jørstad, A., Avseth, P., Mavko, G., Granli, J.R. (2001) *Mapping lithofacies and pore-fluid probabilities in a North Sea reservoir: Seismic inversions and statistical rock physics*, Geophysics, 66 (4):988–1001.

Pelletier, H., Gundersen, J. (2004) *Application of Rock Physics to an Exploration Play, A Case Study from the Brazeau River 3D*, CSEG National Convencion.

Rasolovoahangy, R.E. (2002) *Rock Physics Of LowPorosity Sandstones For Seismic Reservoir Characterization*, Tese (Doutorado em Geofísica) – Stanford - CA, Estados Unidos, Stanford University, 134 p.

Russel, S., Norvig, P. (2003) *Artificial Intelligence*, 2.ed., Prentice Hall.

Schwarz, G. (1978) Estimating the Dimension of a Model, *Annals of Statistics*, 6:461–464.

Sheriff, E.R., Geldart, P.L. (1995) *Exploration Seismology*, 2.ed., Cambridge: Cambridge Press.

Takahashi, I., (2000) *Quantifying Information and Uncertainty of Rock Property Estimation From Seismic Data*, Stanford University.

USHIROBIRA C.A. (2004) *Construção de modelo de porosidade de reservatório*, Tese (Mestrado em Engenharia de Exploração e Produção de Petróleo) – Campos do Goytacazes – RJ, Brasil, LENEP.

Xuelei, H., Lei, X. (2004) *Investigation on Several Model Selection Criteria for Determining the Number of Cluster*, Department of Computer Science and Engineering, The Chinese University of Hong Kong.

Wang, Z., (2001) *Y2K Tutorial: Fundamentals of seismic rock physics*, *Geophysics*, 66 (2):398-412.

West, B.P., Steve, R.M., Eastwood, J.E., Rossen, C. (2002) *Interactive seismic facies classification using textural attributes and neural networks*, The Leading Edge.

Willie, M.R.J., Gregory, A. R., Gardner, L.W. (1956) *elastic wave velocities in heterogeneous and porous media*, *Geophysics*, 21:41-70.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)