** UNIVERSIDADE ESTADUAL DE CAMPINAS**

**FACULDADE DE ENGENHARIA MECÂNICA E  
 INSTITUTO DE GEOCIÊNCIAS**

**Modelamento geostatístico de propriedades petrofísicas e classificação de reservatório por aprendizado de máquina não-supervisionado em um reservatório carbonático naturalmente fraturado na Bacia de Campos, Brasil.**

Orientador: Prof. Dr. Alexandre Campane Vidal

Autor: João Fernando Sampaio Ralha

**Campinas**

**2020**

**Sumário**

[Abstract 7](#_Toc49705803)

[Índice de Figuras 8](#_Toc49705804)

[Índice de Tabelas 11](#_Toc49705805)

[**1.** **Introdução** 13](#_Toc49705806)

[**2.** **Contexto Geológico** 18](#_Toc49705807)

[**2.1 Bacia de Campos** 18](#_Toc49705808)

[**2.2 Grupo Macaé** 19](#_Toc49705809)

[**2.3** **Campo B** 20](#_Toc49705810)

[**3.** **Metodologia** 21](#_Toc49705811)

[**3.1.** **Análise e apresentação de base de dados disponível** 21](#_Toc49705812)

[**3.2.** **Processamento de atributos sísmicos para realce de descontinuidades.** 28](#_Toc49705813)

[**3.3.** **Estabelecimento do modelo estrutural.** 34](#_Toc49705814)

[**3.4.** **Definição de Camadas e *Upscaling* de L*ogs* de Poços** 36](#_Toc49705815)

[**3.5.** **Modelamento de Propriedades** 38](#_Toc49705816)

[**3.6.** **Classificação de Reservatório e Cálculos de Volume de Hidrocarbonetos.** 40](#_Toc49705817)

[**3.6.1.** **Clusterização de células por aprendizado de máquina não supervisionado.** 42](#_Toc49705818)

[**4.** **Resultados e Discussão** 45](#_Toc49705819)

[**4.1.** **Processamento Sísmico e Descontinuidades** 46](#_Toc49705820)

[**4.2.** ***Upscaling* e Modelamento de Propriedades** 52](#_Toc49705821)

[**4.2.1.** **Modelos de Raio Gama (GR).** 53](#_Toc49705822)

[**4.2.2.** **Modelos de Perfil Sônico (DT)** 56](#_Toc49705825)

[**4.2.3.** **Modelos de Densidade (RHOB)** 58](#_Toc49705828)

[**4.2.4.** **Modelos de Porosidade Efetiva (PHIE)** 60](#_Toc49705831)

[**4.2.5.** **Modelos de saturação de água (SW)** 62](#_Toc49705833)

[**4.2.6.** **Correlações entre propriedades de rocha modeladas** 64](#_Toc49705835)

[**4.3.** **Classificação de Reservatório** 67](#_Toc49705844)

[**4.4.** **Cálculos de Volume** 73](#_Toc49705845)

[**5.** **CONCLUSÕES** 76](#_Toc49705846)

[**6.** **REFERÊNCIAS BIBLIOGRÁFICAS** 77](#_Toc49705847)

Abstract

Índice de Figuras

[**Figura 1 - Localização da Bacia de Campos (modificado de Rangel e Martins, 1998)** 14](#_Toc32229521)

[**Figura 2 - Carta estratigráfica da Bacia de Campos simplificada. (Modificada de Guardado et al., 2000, e parcialmente atualizada após Winter et al., 2007, in Melani, 2015).** 15](#_Toc32229522)

[**Figura 3 - Extensões em superfície do reservatório e do volume sísmico no domínio da profundidade. Números ao redor da extensão do volume sísmico correspondem à crosslines e inlines.** 18](#_Toc32229523)

[**Figura 4 - Horizontes carregados - De cima para baixo: Formação Outeiro, Formação Imbetiba, Formação Quissamã e Grupo Lagoa Feia. Cubo no entorno representa os limites do volume sísmico com profundidades de 1798m até 3076m** 19](#_Toc32229524)

[**Figura 5 - Localização dos 47 poços estudados, ícones representam as coordenadas em que os poços têm sua maior profundidade, não sua localização em superfície.** 19](#_Toc32229525)

[**Figura 6 - Pickett Plot utilizado na determinação do fator de cimentação (m = 1,77). De Melani et al. 2015.** 22](#_Toc32229526)

[**Figura 7 - Fluxo de trabalho para geração de volumes de ant-tracking.** 24](#_Toc32229527)

[**Figura 8 - Visualização do fluxo de trabalho utilizado identificando uma descontinuidade.** 27](#_Toc32229528)

[**Figura 9 - Exemplo de fluxo de trabalho para modelamento estrutural. Os dados de interpretações de estruturas e horizontes são inicialmente carregados (a) para então horizontes serem modificados de acordo com sua relação com estruturas que os interceptam (b) e finalmente é gerado um modelo volumétrico de zonas com base nas superfícies modeladas (c). Alterado de Schlumberger (2015).** 29](#_Toc32229529)

[**Figura 10 - Horizonte do topo Quissamã. (a) Mostra o horizonte originalmente carregado. (b) Mostra o horizonte modificado pelas interseções com estruturas. Áreas em azul representam superfícies de feições estruturais.** 31](#_Toc32229530)

[**Figura 11 – (a) Modelo volumétrico de zonas com camadas definidas e (b) Zona do Quissamã entre os horizontes do Quissamã e do Lagoa Feia isolado do resto do modelo.** 32](#_Toc32229531)

[**Figura 12 – Diagrama do funcionamento do algoritmo de upscaling de dados de poço para a escala do modelo. Para cada célula do grid colunar é atribuído um valor para cada log correspondente ao valor médio do log no espaço amostral correspondente à célula.** 32](#_Toc32229532)

[**Figura 13 - Grid colunares de porosidade efetiva (PHIE) cortados para a região entre o topo do Quissamã (superfície em roxo) e o contato óleo-água (superfície em preto).** 33](#_Toc32229533)

[**Figura 14 - Ilustração da obtenção de valores residuais para métodos de Krigagem com Deriva Externa (aqui denominado como regression-kriging). De: Hengl (2009).** 35](#_Toc32229534)

[**Figura 15 - Atributo de Ant-Tracking no crossline 1744. Refletor entre 2200m e 2400m de profundidade corresponde ao horizonte do topo do Quissamã. Refletor entre 2600m e 2800m de profundidade corresponde ao horizonte do topo do Lagoa Feia. Nota-se a ocorrência de forma disseminada de descontinuidades detectadas pelo fluxo de trabalho definido.** 42](#_Toc32229535)

[**Figura 16 - Atributo de Ant-Tracking nos cortes em profundidade de 2300 e 2500m. 2300m sendo a profundidade aproximada do topo do Quissamã e 2500m a profundidade aproximada do contato óleo-água. Polígono em azul representa os limites do reservatório em superfície.** 43](#_Toc32229536)

[**Figura 17 - Distribuição de descontinuidades para todas as descontinuidades. Estruturas predominantemente Leste-Oeste com mergulhos para Norte e Sul.** 43](#_Toc32229537)

[**Figura 18 - Distribuição de descontinuidades com mais de 200m de extensão em qualquer direção. Nota-se a manutenção da tendência estrutural observada para o conjunto de todas as descontinuidades.** 44](#_Toc32229538)

[**Figura 19 - Variação de Ant-Tracking com Latitude. Não se observa tendência clara na distribuição entre as duas variáveis. Cada ponto no gráfico corresponde à uma célula no modelo. Consideradas apenas células dentro do reservatório (Entre Quissamã e contato óleo-água).** 45](#_Toc32229539)

[**Figura 20 - Delimitação de Zonas Norte e Sul do reservatório, cada uma com 234.109 células. Fronteira entre as zonas encontra-se deslocada para o sul devido à maior espessura do reservatório na Zona Sul. Cada ponto no gráfico corresponde à uma célula no modelo. Pontos plotados independentemente de sua posição vertical. Consideradas apenas células dentro do reservatório (Entre Quissamã e contato óleo-água).** 46](#_Toc32229540)

[**Figura 21 - Distribuição no espaço para os pontos de altos valores de Ant-Tracking. -0.91 corresponde ao valor médio de Ant-Tracking no reservatório adicionado ao valor do desvio padrão de Ant-Tracking no reservatório, -0.78 representa o valor médio somado a três vezes o desvio padrão. Nota-se a maior frequência com que estes pontos ocorrem na zona sul em relação à zona norte. Cada ponto no gráfico corresponde à uma célula no modelo. Pontos plotados independentemente de sua posição vertical. Consideradas apenas células dentro do reservatório (Entre Quissamã e contato óleo-água)** 46](#_Toc32229541)

[**Figura 23 - Distribuição de dados de Saturação de Água. Dados após upscaling honram bem os dados de poço. Modelagem por KDE tem como efeito a possível superestimação de valores medianos de saturação e suprimindo valores de baixa saturação. Efeitos do modelamento indicam que os cálculos de volume possam acabar subestimando o volume de hidrocarbonetos HCPV.** 48](#_Toc32229542)

[**Figura 24 - Distribuição de dados de Porosidade Efetiva. Dados de upscaling e modelados por GRFS honram de maneira satisfatória seus dados de input. Resultando num modelo confiável para a distribuição de valores de porosidade ao longo do reservatório.** 49](#_Toc32229543)

[**Figura 25 – Distribuição de dados de raio gama. Embora o algoritmo de GRFS tenha honrado de maneira satisfatória a distribuição dos dados após upscaling, o método de upscaling acaba por subestimar valores de baixo GR entre 30 e 40 gAPI.** 49](#_Toc32229544)

[**Figura 26 – Distribuição de dados de densidade. Assim como para dados de porosidade efetiva, tanto o algoritmo de GRFS quanto o método de upscaling honram seus dados de input.** 50](#_Toc32229545)

[**Figura 27 – Distribuição de dados de resistividade. Assim como para os dados de raio gama, o algoritmo de GRFS honra de maneira satisfatória os dados de upscaling, enquanto o método de upscaling acaba por subestimar valores intermediários entre 5 e 10 ohm.m** 50](#_Toc32229546)

[**Figura 28 – Modelo de saturação de água (SW). Valores distribuem-se de maneira similar para diferentes níveis de saturação de água. Algoritmo de KDE resulta em transições de valores suavizadas. Nota-se os maiores valores de saturação próximos ao topo e base do reservatório, evidenciando as regiões de fronteira do reservatório, assim como a presença de uma região na porção sudoeste do reservatório com valores muito baixos de saturação.** 53](#_Toc32229547)

[**Figura 29 - Modelo de porosidade efetiva (PHIE). Valores distribuídos em camadas horizontais contendo maiores valores. Observa-se associação entre maiores valores de porosidade nos mesmos locais onde ocorrem menores valores de saturação de água (SW).** 54](#_Toc32229548)

[**Figura 30 - Modelo de raio gama (GR). Observa-se dependência entre os valores modelados e profundidade, com valores próximos ao topo da Formação Quissamã apresentando um aumento nos valores de raio gama observados.** 55](#_Toc32229549)

[**Figura 31 – Modelo de resistividade (ILD). Observa-se predominância de valores baixos. A ocorrência de picos com valores bastante elevados na região sudoeste do modelo explica os altos valores de CV observados na Tabela 6.** 56](#_Toc32229550)

[**Figura 32 – Modelo de Densidade (RHOB). Valores aproximadamente constantes em torno da média de aproximadamente 2400g/cm³ explicam baixos valores de CV observados na Tabela 6. Ao mesmo tempo, raras variações são observadas na região sudoeste, com quedas nos valores de Densidade associadas a valores de alta porosidade, alta resistividade e baixa saturação de água.** 57](#_Toc32229551)

[**Figura 33 – Distribuições par a par das propriedades modeladas. Histogramas de distribuição das propriedades em casos em que a propriedade é comparada a ela mesma. Valores intermediários de correlação entre SW-PHIE e SW-RHOB observados na Tabela 7 podem ser visualizados. Nota-se ainda a distribuição concentrada em baixos valores observada para a propriedade de resistividade (ILD).** 58](#_Toc32229552)

[**Figura 34 – Matriz de Distâncias e respostas à cada variável do algoritmo de mapas auto organizáveis (SOM). Observa-se a relação entre altas respostas de porosidade e resistividade com baixas respostas de saturação de água e densidade.** 60](#_Toc32229553)

[**Figura 35 – Valores para os índices de silhueta e Davies-Bouldin computados para diferentes valores de K no algoritmo K-Means. O valor máximo para índice de silhueta e mínimo para Davies-Bouldin no valor de 3 classes indica provável separação dos dados em 3 diferente classes.** 61](#_Toc32229554)

[**Figura 36 – Classificações da matriz de distâncias. Valores entre 0 e 1 para o mapa de probabilidades de GMM são a probabilidade de cada neurônio pertencer a classe 2.** 61](#_Toc32229555)

[**Figura 37 – Mapas de distribuição das classificações para K-Means e GMM e propriedades de porosidade efetive e saturação de água para profundidade de 2392m. Linha laranja indica as seções transversais da Figura 39.** 64](#_Toc32229556)

[**Figura 38 - Seções transversais para as distribuições das classificações obtidas de K-Means e GMM. Seções correspondem as linhas indicadas nos mapas da Figura 38.** 65](#_Toc32229557)

[**Figura 39 – Mapa e seção transversal de distribuição de probabilidades de cada célula pertencer a Classe 0 ou 1 na classificação obtida do método GMM.** 66](#_Toc32229558)

Índice de Tabelas

[**Tabela 1 - Profundidades dos horizontes carregados. O horizonte ‘OWC’ representa o contato óleo-água (Oil Water Contact) enquanto o horizonte ‘Água Até’ representa as profundidades mais rasas com conteúdo de água dentro do reservatório.** 18](#_Toc32229559)

[**Tabela 2 - Relação de disponibilidade de logs para cada poço estudado. A coluna de saturação de água diz respeito as 4 curvas determinadas por Melani et al. (2015), visto que, para os poços em que elas estão disponíveis, as 4 curvas encontram-se disponíveis.** 20](#_Toc32229560)

[**Tabela 3 - Fatores de cimentação e saturação para cada um dos cenários de cálculo de saturação de água. De Melani et al. (2015)** 21](#_Toc32229561)

[**Tabela 4 - Definição de parâmetros para o atributo de ant-tracking e valores utilizados no processamento do atributo. Os valores tomados para cada parâmetro seguem os valores sugeridos na literatura para reservatórios fraturados. Baseado em Schlumberger (2015).** 28](file:///C:\Users\Ralha\Repos\campob\dissertacao\Cálculos%20de%20volume%20e%20classificação%20de%20reservatório%20baseados%20em%20propriedades%20petrofísicas%203D.docx#_Toc32229562)

[**Tabela 5 - Distribuição dos valores de Ant-Tracking nas diferentes Zonas do volume sísmico. Embora a Zona Sul mostre um valor média de Ant-Tracking superior tanto à Zona Norte quanto a região não reservatório do Quissamã, a diferença observada na média é menor do que o desvio padrão destas medidas.** 47](#_Toc32229563)

[**Tabela 6 - Valores estatísticos para a distribuição dos valores modelados. CV é o coeficiente de variação.** 51](#_Toc32229564)

[**Tabela 7 – Valores de correlação entre as propriedades modeladas. Tabela tem valores repetidos omitidos.** 52](#_Toc32229565)

[**Tabela 8 - Distribuição estatística das propriedades modeladas entre as classes obtidas com o algoritmo de Modelo de Mistura Gaussiana, GMM.** 63](#_Toc32229566)

[**Tabela 9 – Distribuição estatística das propriedades modeladas entre as classes obtidas com o algoritmo de K-Means.** 63](#_Toc32229567)

[**Tabela 10 – Volumes obtidos para cada classe de acordo com os classificadores de K-Means e GMM. Os valores percentuais indicam a relação entre o volume referido e o obtido para a mesma classe do K-Means. Vb (Bulk Volume) é o volume total das células de cada classe, Vp (Pore Volume) o volume poroso enquanto HCPV (Hydrocarbon Pore Volume) é o volume de hidrocarbonetos.** 67](#_Toc32229568)

[**Tabela 11 – Relação dos casos de cálculos de volume. Casos que trazem valores de Prop para saturação de água (SW) ou porosidade efetiva (PHIE) utilizam os valores modelados para cada célula para o cálculo. [1] O Caso 3 utiliza o valor constante de 0.581 para saturação de água, valor correspondente ao valor médio no reservatório. [2] O Caso 2 utiliza o valor constante de 0.198 para porosidade efetiva, valor corresponde ao valor médio no reservatório. [3] Valores de N/G de 0 ou 1 para cada célula baseados na classificação em poços de Melani et al. (2015), cada célula possui valor atribuído igual ao ponto mais próximo nos poços classificados. [4] Valores de corte baseados no percentil de 50% para ambas as propriedades. [5] Valores deorte baseados no percentil de 75% para ambas as propriedades. [6] Valores de corte baseados no percentil de 90% para ambas as propriedades.** 67](#_Toc32229569)

[**Tabela 12 – Valores de volumes obtidos para cada caso descrito na Tabela 11. Vb (Bulk Volume) é o volume total das células de cada classe, Vp (Pore Volume) o volume poroso enquanto HCPV (Hydrocarbon Pore Volume) é o volume de hidrocarbonetos e tHCPV é o volume de hidrocarbonetos antes de aplicação das razões N/G. Valores percentuais indicam a variação de cada caso em relação ao Caso 0. Casos 7 e 8 apresentam valores de volume total (Vb) diferentes pois não incluem as células da Classe 0.** 68](#_Toc32229570)

1. **Introdução**

Este estudo visa desenvolver modelos de propriedades geométricas e petrofísicas, além de propriedades derivadas de atributos sísmicos, de um reservatório naturalmente fraturado (Chamado aqui de “Campo B”) inserido no intervalo de idade do albiano da Bacia de Campos, possibilitando cálculos de volume e classificação de zonas de interesse no reservatório por meio da inferência de propriedades em zonas inter-poços e o uso de métodos de aprendizado de máquina não supervisionado. O Campo B mostra-se como de especial interesse para este estudo por se tratar de um reservatório naturalmente fraturado inserido em uma anticlinal, tornando o controle estrutural um fator de alta relevância na definição das extensões finais do modelo, e consequentemente no volume total.

Na indústria de exploração em petróleo e gás, o conceito de modelamento de reservatório é entendido como a integração de múltiplas fontes de dados em diferentes resoluções e qualidades com o objetivo de desenvolver uma representação acurada da geologia em subsuperfície. A combinação de conceitos de geoestatística com o aumento de capacidades de processamento e bases de dados cada vez maiores dá início ao processo de técnicas estatísticas sendo utilizadas para a inferência e estimativa de propriedades geológicas (Ringrose & Bentley, 2014a).

Nota-se conjuntamente ao desenvolvimento de novas tecnologias, um aumento constante na capacidade de um único geocientista em prover informações quantitativas e qualitativas sobre a geologia em subsuperfície. Possibilitando a realização de modelos preditivos de alta complexidade (Trabanou et al, 2004). A complexidade de dado processo se mostra na quantidade de variáveis, na variabilidade da resposta a estas variáveis, na qualidade, extensão e resolução dos dados disponíveis bem como na subjetividade associada a interpretações e decisões do operador (Hall, 2013; Anderson & Newrick, 2008).

A construção de modelos geológicos deve ter um objetivo claro desde o seu início, pois as considerações e requerimentos são tão diferentes quanto seus objetivos. Ainda que a maior parte dos modelos somente busque a visualização de inferências sobre a geologia local. O cálculo de modelos para estimativas de volume de reservatório é majoritariamente ligado a propriedades estatisticamente aditivas como saturação de água e porosidade, assim, grids voltados ao cálculo de volumes devem ter extensões claras e bem definidas e não requerem alto nível de detalhamento quanto à sua resolução (Ringrose & Bentley, 2014c).

Em sua maioria, técnicas em modelamento geoestatístico dependem de modelos de natureza gaussiana, mesmo simulações classificadas como bayesianas ainda dependem de modelos gaussianos multivariáveis para ajustes. Modelos geológicos, no entanto, consistem em diversos componentes que não podem ser descritos satisfatoriamente por variáveis aleatórias, distribuições normais e suas probabilidades associadas, fazendo com que o estado da arte em técnicas de modelamento seja indefinido para um grande número de processos (Caers et al., 2010). Devido a isto, os modelos geoestatístico gerados foram iterativamente desenvolvidos tentando otimizar a distribuição estatística das propriedades modeladas de acordo com a distribuição dos dados de entrada. Neste estudo utiliza-se um algoritmo de simulação condicional para a etapa de modelamento geostatístico, aplicando algoritmos de krigagem para dados na porção estruturada dos variogramas e simulações gaussianas para dados na porção não estruturada, permitindo melhor desempenho na estimativa de dados nas porções não estruturadas dos variogramas (Hengl et al., 2007; Hansen et al., 2006) e evitando as feições suavizadas características de algoritmos baseados unicamente em métodos de krigagem (Schlumberger, 2015).

Reservatórios naturalmente fraturados como o Campo B são definidos como reservatórios onde fraturas tem um impacto significativo na migração, armazenamento e extração de hidrocarbonetos (Warren & Root, 1963). Lyons & Plisga (2016) dividem reservatórios fraturados de acordo com o sistema principal de fluxo e de armazenamento em três tipos de reservatório: Tipo I, onde a matriz não contribui de maneira significativa para o fluxo tão pouco quanto para o armazenamento de hidrocarbonetos; Tipo II, onde o armazenamento é primariamente na matriz enquanto o fluxo é controlado majoritariamente pelo sistema de fraturas: e um Tipo 3, onde a matriz é um controle tanto no armazenamento quanto no fluxo, com fraturas atuando de maneira suplementar. Reservatórios dos tipos I e II comumente apresentam alta heterogeneidade quanto a parâmetros petrofísicos, dificultando a caracterização de reservatório e definição de zonas de interesse.

O Campo B é definido na literatura disponível como um reservatório de alto controle estrutural no fluxo de hidrocarbonetos, mas com um controle estratigráfico no armazenamento, caracterizando-se assim como um reservatório fraturado de tipo II, uma descrição detalhada e referenciada sobre os controles estruturais no Campo B é desenvolvida no Capitulo 2 – Contexto Geológico.

À fim de honrar a complexidade estrutural do reservatório, um fluxo de trabalho baseado em atributos sísmicos é desenvolvido no estudo para caracterizar feições estruturais no reservatório.

Atributos sísmicos caracterizam-se por qualquer característica quantitativa calculadas com base nos valores de amplitude originais de um volume sísmico (Taner et al, 1979; Bjorlykke, 1989). Atributos sísmicos, de maneira geral, procuram realçar ou suprimir alguma característica observada no volume sísmico de amplitudes original. Seja o realce de descontinuidades, de refletores ou um aprimoramento da razão sinal-ruído, entre outros (Bahorich & Farmer, 1995; Brown, 1999). Com base nisto, este estudo busca caracterizar a complexidade estrutural do Campo B desenvolvendo atributos sísmicos que realcem e discretizem feições estruturais no reservatório previamente a qualquer outro processo.

O processo de caracterização estrutural de um reservatório é um processo tradicionalmente repleto de subjetividades relacionadas à interpretação do usuário e aos dados disponíveis, atributos relativos à geometria de fraturas em testemunho e/ou log de poço são comumente utilizados (Costa et al., 2018; De Lima et al., 2019) devido a limitação de escala dos dados baseados em sísmica de reflexão. No entanto, dados relativos à geometria de fratura não estão disponíveis para o Campo B, limitando a capacidade dos atributos sísmicos utilizados de representar fraturas sub-sísmicas, ainda que fraturais nesta escala não influenciem significativamente modelos volumétricos (Ringrose & Bentley, 2014c). Ao mesmo tempo, altos valores para o atributo de ant-tracking são associados a altos graus de fraturamento em testemunho (Fang et al., 2016), possibilitando a utilização do atributo para identificação de zonas fraturadas.

A utilização de fluxos de trabalho baseados em atributos sísmicos, ao evitar interpretações estruturais manuais, possibilita também a redução do viés interpretativo associado ao processo (Cox & Seitz, 2007). Baseado nisto, Ralha (2016) utiliza em estudo prévio sobre o Campo B um fluxo de trabalho para o processamento de atributos sísmicos para o realce de descontinuidades baseado nos atributos de variância e ant-tracking. Ao mesmo tempo, outros autores sugerem a utilização de filtros de suavização estrutural previamente ao cálculo do atributo de variância à fim de reduzir os efeitos de ruídos nas estruturas caracterizadas (Randen et al., 2001; Zhao et al., 2015; Basir et al., 2013).

Técnicas de aprendizado de máquina não supervisionado, por sua vez, mostram-se como uma alternativa eficiente para a classificação de reservatório, utilizando a grande capacidade de reconhecimento de padrões dos algoritmos de aprendizado de máquina para caracterizar heterogeneidades onde não existe classificação prévia (Taner et al, 2001). Os métodos de aprendizado de máquina não supervisionados mostram-se efetivos para estimativas de fácies derivadas de dados petrofísicos (Al-Mudhafar et al., 2015; Aragão & Sava, 2019). Kuznetsova et al. (2016), apresenta resultados comparáveis à métodos de aprendizado de máquina supervisionado para classificação de fácies em logs de poço, sugerindo que as heterogeneidades encontradas no processo de classificação de reservatório conseguem ser representadas por processos não supervisionados, possibilitando a aplicação destes algoritmos em regiões sem qualquer tipo de classificação realizada previamente.

Kuroda et al. (2012) e Leite & Filho (2010) realizam estudos aplicando técnicas de aprendizado de máquina não supervisionado como Mapas Auto-Organizáveis e K-Means para classificação de fácies em logs de poço na Bacia de Campos. O fluxo de trabalho desenvolvido na etapa de classificação deste estudo baseia-se no estudo de Kuroda et al. (2012), aplicando um método similar para modelos volumétricos utilizando os métodos de K-Means e Mistura Gaussiana.

A etapa final de cálculo de volumes porosos e de hidrocarbonetos para o reservatório inclui a estimativa de volume poroso e volume de hidrocarbonetos para cada célula do modelo geológico, utilizando os valores obtidos na etapa de modelamento para os cálculos e as classificações obtidas via aprendizado de máquina para diferenciar zonas de interesse e estimar razões *net-to-gross*.

O algoritmo de Mistura Gaussiana, ao apresentar saídas probabilísticas para cada classe, permite quantificar as probabilidades de cada ponto pertencer a cada classe (Viroli & McLachlan, 2019). Como o cálculo de volumes, objetivo final deste estudo, passa pela estimativa da razão Net-to-Gross, também medida entre 0 e 1 (Egbele et al., 2015), modelos de Mistura Gaussiana são utilizados para estimar as razões nos cálculos de volume.

Os resultados deste trabalho incluem: Atributos sísmicos para identificação automática de falhas e fraturas utilizados em conjunto à horizontes previamente interpretados para delimitação do reservatório; Modelos tridimensionais para a área do reservatório baseados em inferências de dados de poço para regiões intra-poço e atributos sísmicos; Classificação do reservatório em diferentes classes; Correlação de classes com litologias; Estimativa de volumes baseadas em classificações probabilísticas de reservatório; e cálculo de volume poroso e volume de hidrocarbonetos para as diferentes zonas.

A base de dados deste estudo inclui horizontes no domínio da profundidade para o topo e base do reservatório além de outras unidades geológicas estratigraficamente próximas, volume 3D de sísmica de reflexão no domínio da profundidade recortado para a área do reservatório e logs de poço com diferentes disponibilidades de logs para 47 poços diferentes interceptando ou próximos ao reservatório, dados fornecidos pela Agência Nacional de Petróleo – ANP.

1. **Contexto Geológico**

## **2.1 Bacia de Campos**

A Bacia de Campos, localizada na região sudeste do Brasil, na costa dos estados do Rio de Janeiro e Espírito Santo (Figura 1), ocupa uma área de 115 mil km² desde a linha de costa (Bruhn et al., 1998). Reservatórios de hidrocarbonetos da Bacia de Campos têm como geradoras as rochas do Grupo Lagoa Feia (Bruhn et al., 2003). A Bacia de Campos é considerada a bacia brasileira de maior interesse econômico, com mais de 30 anos de exploração continua e mais de 1600 poços, compreendendo mais de 75% da produção e das reservas brasileiras de hidrocarbonetos (ANP, 2013).



**Figura 1 - Localização da Bacia de Campos (modificado de Rangel e Martins, 1998)**

Bacia de margem passiva, a Bacia de Campos tem sua evolução relacionada à quebra do Gondwana (Guardado et al. 1989; Dias et al., 1990. Definem-se assim 2 estilos estruturais principais à bacia: (1) falhamentos de blocos relacionados à fase *rift*, compondo *horsts*, *grabens* e semi-*grabens*, envolvendo o embasamento e sedimentos do pré-sal e (2) estruturas geradas pela movimentação de sal, em geral falhas lístricas, afetando sedimentos do pós-sal (Spadini, 1992; Figueiredo & Mohriak, 1984; Guardado et al., 1989; Dias et al., 1990).

Winter et al. (2007) dividem a estratigrafia da Bacia de Campos em Formação Cabiúnas, Grupo Lagoa Feia, Grupo Macaé e Grupo Campos (Figura 2). O Campo B, estudado neste trabalho, encontra-se inserido no intervalo do Grupo Macaé.



**Figura 2 - Carta estratigráfica da Bacia de Campos simplificada. (Modificada de Guardado et al., 2000, e parcialmente atualizada após Winter et al., 2007, in Melani, 2015).**

## **2.2 Grupo Macaé**

Composto pelas Formações: Goitacás, Quissamã, Outeiro, Imbetiba e Namorado (Winter et al, 2007), o Grupo Macaé representa porção inferior da sequência pós-sal, com deposição carbonática em mar raso (Robaina et al, 1991). Dominado por carbonatos marinhos de águas rasas, o Grupo Macaé é caracterizado por uma associação de calcarenitos, calcirruditos e calcilutitos, depositados em ambiente de moderada a alta energia (Franz, 1987).

Segundo Spadini (2008), reservatórios carbonáticos albianos mostram grande variação de porosidade e permeabilidade. Altas permeabilidades correspondem à porosidade intergranular deposicional, enquanto baixas permeabilidades das rochas refletem predominância de microporosidade. Bruhn et al. (2003) define que as acumulações com idades albianas de petróleo têm controle estrutural, por falhas e dobras, e também estratigráfico, dado por variação lateral de fácies, com calcarenitos e calcirruditos gradando para calcarenitos, calcissiltitos e calcilutitos ricos em lama.

* 1. **Campo B**

O Campo B é um reservatório fraturado carbonático do albiano, produtivo a quase 40 anos. A cerca de 85km da costa, cobre uma área de 14km². Composto predominantemente por calcarenitos e calcirruditos da Formação Quissamã distribuídos em bancos de areias carbonáticas de direção NE (Franz, 1987), o reservatório ocupa um bloco falhado, inserido em uma anticlinal, associado a duas falhas principais de crescimento que ocorrem a oeste e leste-sudeste do campo. (Franz, 1987).

Tomaso et al (2013) dividem o Campo B em duas zonas principais de acordo com intensidade de fraturamento: (1) área sul, onde predominam presença de fraturas e microporosidade; e área norte, região não fraturada com preservação de porosidade primária, havendo compartimentalização do reservatório em zona sul e zona norte.

Fraturas desempenham um importante papel na produção do campo, controlando e criando as condições necessárias para a produção de óleo (Franz, 1987). A origem destas fraturas no Campo B está intimamente ligada aos eventos causadores dos falhamentos regionais, sendo possível inferir ao menos dois estágios principais de fraturamento no reservatório. Primeiramente as fraturas foram totalmente cimentadas por calcita espática. Depois ocorrendo a reabertura e alargamento do sistema de fraturas, tornando estas falhas via preferencial de fluxo de fluido.

Tomaso et al (2007) ainda definem um limite superior para o reservatório no topo da formação Quissamã, com sua base sendo o contato óleo/água, não sendo possível identificar contato geológico relacionado.

1. **Metodologia**

Divide-se o estudo em 8 fases distintas: (1) análise e apresentação de base de dados disponível, (2) processamento de atributos sísmicos para realce de descontinuidades, (3) estabelecimento de modelo estrutural, (4) definição de camadas e *upscaling* de *logs* de poços, (5) modelamento de propriedades, (6) classificação de reservatório e cálculos de volume de hidrocarbonetos.

* 1. **Análise e apresentação de base de dados disponível**

O volume sísmico 3D disponível para o estudo caracteriza-se por um volume pós-empilhamento convertido para o domínio do tempo por interpolações de velocidades de trânsito obtidas de *logs* sônicos de poço. Anteriormente à conversão de domínio sabe-se que o volume sísmico passou por processo de *normal moveout* e ganho em tempos maiores, embora não exista registro do histórico de processamento do volume no cabeçalho do arquivo *segy*.

Já convertido para o domínio da profundidade, o volume cortado tem uma área superficial de aproximadamente 28km² (Figura 3) e profundidades aproximadamente entre 1,9km e 4km, apresentando extensão vertical de pouco mais de 2km. Sendo composto por 296 *inlines* com espaçamento de 25m e 311 *crosslines* com espaçamento de 12.5m, o volume possui 8.5m de profundidade por amostra (4ms no dado em tempo) com 218 amostras por traço. Totalizando pouco mais de 20 milhões de amostras.

Horizontes inicialmente interpretados no domínio do tempo e convertidos para o domínio da profundidade conjuntamente ao volume sísmico também fazem parte da base de dados, estes horizontes representam os topos das Formações Outeiro, Imbetiba, Quissamã e Lagoa Feia (Figura 4) assim como o contato óleo-água na base do reservatório e o limite superior de ocorrência de água no reservatório. Assume-se que o topo de uma formação corresponde à base da unidade que a sobrepõe. A Tabela 1 traz as profundidades em que os horizontes são encontrados.

Os dados de poço disponibilizados referem-se à poços que interceptam ou localizam-se próximos ao reservatório (Figura 5) e encontram-se em situações diferentes de disponibilidade de logs e profundidades em que os dados são disponibilizados. A Tabela 2 traz a relação de disponibilidade de dados para cada um dos 47 poços estudados.



**Figura 3 - Extensões em superfície do reservatório e do volume sísmico no domínio da profundidade. Números ao redor da extensão do volume sísmico correspondem à crosslines e inlines.**

**Tabela 1 - Profundidades dos horizontes carregados. O horizonte ‘OWC’ representa o contato óleo-água (Oil Water Contact) enquanto o horizonte ‘Água Até’ representa as profundidades mais rasas com conteúdo de água dentro do reservatório.**



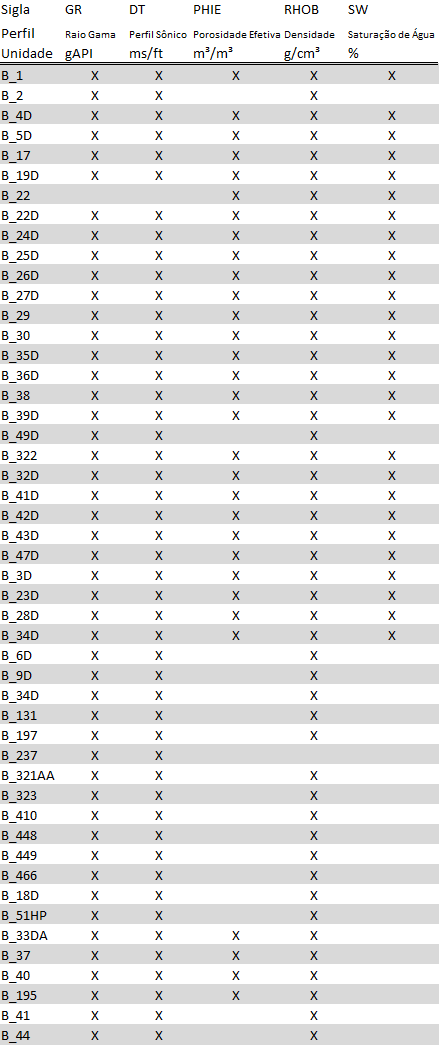


**Figura 4 - Horizontes carregados - De cima para baixo: Formação Outeiro, Formação Imbetiba, Formação Quissamã e Grupo Lagoa Feia. Cubo no entorno representa os limites do volume sísmico com profundidades de 1798m até 3076m**



**Figura 5 - Localização dos 47 poços estudados, ícones representam as coordenadas em que os poços têm sua maior profundidade, não sua localização em superfície.**

**Tabela 2 - Relação de disponibilidade de logs para cada poço estudado. A coluna de saturação de água diz respeito as 4 curvas determinadas por Melani et al. (2015), visto que, para os poços em que elas estão disponíveis, as 4 curvas encontram-se disponíveis.**



Os dados de porosidade efetiva (PHIE) e saturação de água (SW) são modelados neste estudo, pois são as propriedades que influenciam os cálculos de volume desenvolvidos. Modela-se também os dados de densidade (RHOB), perfil sônico (DT) e raio gama (GR) para fornecer maiores dados em tentativas de classificação automática de fácies. Os dados de porosidade efetiva, densidade, resistividade e raio gama foram fornecidos de maneira direta pela Agência Nacional de Petróleo (ANP). Já os dados de saturação de água são representados por 4 curvas de saturação calculadas por Melani et al. (2015) utilizando a equação de Archie:

|  |  |  |
| --- | --- | --- |
|  |  | (3.1.1) |
|  |  |  |

onde *Phi* é a porosidade efetiva, é o valor calculado de saturação de água, representa a resistividade da água, é a resistividade real da rocha, *a* é o fator de tortuosidade, *m* é o fator de cimentação e *n* é o fator de saturação. As quatro curvas assumem um valor constante de 1 para o fator de tortuosidade (a) devido a não existência de dados laboratoriais. Os valores de fator de cimentação (m) e saturação (n) são diferentes para cada curva (Tabela 2).

**Tabela 3 - Fatores de cimentação e saturação para cada um dos cenários de cálculo de saturação de água. De Melani et al. (2015)**



O cenário C1 utiliza valores empíricos tradicionais de m e n para reservatórios carbonáticos (Glover, 2011). C2 utiliza uma análise de *Pickett plot* para definir m e define n baseado no menor valor medido em laboratório para um reservatório análogo (Elias & Steagall, 1996). C3 assume o valor médio de medidas de laboratório para testemunhos de 2 poços do Campo B para m e o valor médio de todas as medidas de laboratório para um reservatório análogo para n. C4 assume m e n como os maiores valores medidos em laboratório para um reservatório análogo.

O método da análise de Pickett plot utilizado no cenário C2 consiste num gráfico log-log de resistividade e porosidade efetiva. Pontos de mesma saturação de água () neste gráfico podem ser caracterizados por linha retas paralelas na distribuição dos valores e o fator de cimentação (m) é definido pelo coeficiente angular destas linhas (Figura 6).

****

**Figura 6 - Pickett Plot utilizado na determinação do fator de cimentação (m = 1,77). De Melani et al. 2015.**

O fato de cimentação (m) é entendido como um modelo da influência do sistema de poros na resistividade da rocha, uma vez que se assume que a matriz é não condutora. Relacionando assim o fator de cimentação à permeabilidade da rocha, uma vez que altos valores de permeabilidade reduzem o fator de cimentação. O fator de saturação (n) modela o comportamento da condutividade de fluidos no espaço poroso da rocha e relaciona-se com a molhabilidade da rocha. Rochas em ambientes aquosos, para baixos valores de saturação de água, formam uma camada fina de água ao longo das paredes dos poros, tornando a rocha condutiva. Ao mesmo tempo, rochas em meio à óleo são caracterizadas por gotas descontinuas de água no espaço poroso, reduzindo a condutividade da rocha (Ellis & Singer, 2008).

* 1. **Processamento de atributos sísmicos para realce de descontinuidades.**

Os atributos processados nesta etapa são processados utilizando fluxos de trabalho disponíveis no software *Schlumberger Petrel*, com licença disponibilizada ao laboratório de informática do Departamento de Engenharia do Petróleo – DEP, da Universidade Estadual de Campinas. O Campo B é descrito na literatura como um reservatório fraturado que se encontra inserido em uma anticlinal de eixo com direção aproximadamente NW, assim a caracterização estrutural do reservatório mostra-se como importante etapa na definição de sua geometria e limites, fatores diretamente relacionados ao cálculo de volumes no reservatório.

O atributo de variância é largamente considerado como um atributo que realça a descontinuidade de volumes sísmicos e sua utilização na caracterização estrutural de reservatório é sugerida por diversos autores (Pedersen et al, 2005; Fletcher et al, 2011; Cox & Seitz, 2005; Silva et al, 2005; Chopra & Marfurt, 2007; Rijks & Jauffred, 1991; Bruno et al, 1998).

O atributo de *ant-tracking* mostra-se como uma técnica de realce de valores contínuos em um volume e sua utilização com *inputs* de atributos de realce de descontinuidades é aceito como satisfatório em processos de classificação estrutural (Pedersen et al, 2005; Cox & Seitz, 2005; Silva et al, 2005), especialmente em litologias com elevadas heterogeneidades associadas como reservatório fraturados e/ou carbonáticos (CSM, 2012; Chahine et al, 2014; Zeng & Kerans, 2003).

Define-se assim para este estudo um fluxo de trabalho caracterizado por filtro de suavização estrutural inicial aplicado às amplitudes no domínio da profundidade e utilizados como input para cálculo de atributo de variância, que por sua vez é utilizado como input no algoritmo de *ant-tracking* (Figura 7).



**Figura 7 - Fluxo de trabalho para geração de volumes de ant-tracking.**

O filtro de suavização estrutural é um filtro aplicado diretamente a amplitudes sísmicas que aplica suavização por um filtro gaussiano paralelo à orientação de refletores (Schlumberger, 2015; Hale, 2009). A aplicação do filtro de suavização estrutural aprimora, em especial para litologias carbonáticas, a continuidade de refletores bem como colabora com a supressão de ruídos (Randen et al., 2001).

O principal parâmetro de configuração do filtro é o espaço amostral para filtragem em cada traço, o espaço amostral pode ser definido tanto verticalmente quanto nas direções de *crossline* e *inline* e tem, no software Schlumberger Petrel, valor padrão de 1.5 amostras e é o valor utilizado neste estudo. Ou seja, para cada amostra no volume original, a suavização leva em conta 1.5 amostras em cada direção e sentido.

O filtro Gaussiano em 3 dimensões é definido pela equação:

|  |  |  |
| --- | --- | --- |
|  |  | (3.2.1) |

onde representa o valor de amplitude após a aplicação do filtro, x, y e z representam os espaços amostrais em cada direção e representa o desvio padrão das amplitudes dentro deste espaço amostral (Ponte., 2019).

O atributo de variância é calculado em três dimensões e representa a variância traço a traço da amplitude original num mesmo espaço vertical de tempo/profundidade ou paralelamente à um refletor. O cubo de variância caracteriza-se pelo cálculo de variâncias normalizadas das amplitudes convertidas para uma mesma fase em espaços amostrais definidos pelo usuário (Randen et al., 2000; Van Bemmel & Pepper, 2000). Ou seja, para cada amostra, o seu valor no cubo de variância será a variância estatística dos módulos das amplitudes no espaço amostral bidimensional definido normalizada para valores entre 0 e 1. Tal qual:

|  |  |  |
| --- | --- | --- |
|  |  | (3.2.2) |
|  |  |  |

Onde V é o valor do atributo de variância em cada amostra, é o desvio padrão das amplitudes dentro do espaço amostral selecionado (com sendo a variância estatística), A é amplitude de cada amostra e N é o total de amostras no espaço amostral. Os valores padrões utilizados para o espaço amostral são de 3 *inlines* e 3 *crosslines* em cada direção, totalizando 36 amostras de amplitude para cada valor de variância.

O atributo de *ant*-*tracking* baseia-se em conceitos de inteligência de enxame mimetizando colônias reais de formigas. O conceito de inteligência de enxame se refere ao comportamento coletivo de grupos de insetos. Formigas no mundo real utilizam feromônios para marcar o caminho percorrido em sua busca por comida. Analogamente, o atributo de *ant*-*tracking* utiliza-se de uma grande quantidade de agentes distribuídos ao longo de um volume de descontinuidades. Estes agentes, representando formigas no modelo, designam feromônios às áreas correspondentes a altos valores de descontinuidade e, por um processo iterativo, traçam o menor caminho entre seu ponto de origem (seu ninho) e altos valores de variância (Pedersen et al, 2005). Os pontos de origens das formigas são definidos dividindo o volume de variâncias em sub volumes de mesmo tamanho para cada formiga e tomando como a origem os máximos locais de variância, garantindo assim que o caminho traçado corresponde a descontinuidades (Chen et al, 2018).

A definição especifica do algoritmo de *ant*-*tracking* é protegida por patente concedida à companhia *Schlumberger*. Chen et al (2018), no entanto, sugerem um algoritmo similar caracterizado pela utilização de amplitudes como altitudes em um modelo imitando um modelo topográfico e formigas navegando o relevo virtual em busca do menor caminho possível. Esta alternativa, embora bem definida matematicamente por funções de gradientes locais, limita-se a dados sísmicos bidimensionais, utiliza como entrada o volume sísmico de forma direta e, diferentemente do algoritmo de *ant*-*tracking*, não possui extensa validação de uso estabelecida na literatura científica.

O atributo de *ant*-*tracking* caracteriza-se por uma série de parâmetros como: distância limite entre formigas, desvio de trilha, tamanho de passo, passos ilegais permitidos, passos legais necessários e critério de parada. A tabela 4 traz a definição destes parâmetros e os valores utilizados para cada um deles. A escolha dos parâmetros afeta a sensibilidade da resposta encontrada no volume de *ant*-*tracking* aos valores de variância. No entanto, como o fluxo de trabalho definido busca a supressão de ruídos não relacionados a feições estruturais previamente ao cálculo de variâncias, a definição de parâmetros para o volume passa a ser uma decisão entre o realce de feições regionais ou descontinuidades em menor escala. Assim, os parâmetros utilizados baseiam-se nas configurações do software Petrel para reservatórios fraturados (Schlumberger, 2015; Fang et al, 2016).

O cubo de *ant*-*tracking* resultante do processo é composto por valores entre -1 e 1 para cada amostra do volume original, onde -1 representa um volume inteiramente homogêneo e continuo enquanto 1 representa um volume teórico inteiramente descontinuo (Chahine et al., 2015). Ao mesmo tempo, medidas de grau de fraturamente obtidas de testemunhos correlacionam-se com valores de ant-tracking de -0.8 até 0, enquanto que valores muito próximos de 1 indicam refletores distribuídos de maneira caótica e normalmente são observados apenas em depósitos de sal ou regiões de intensa deformação (Godfrey & Bachrach, 2008).

O processo de individualização de estruturas é tradicionalmente realizado manualmente com base em interpretações do operador, consumindo grande quantidade de tempo e inserindo um alto grau de subjetividade às interpretações estruturais. Em especial para reservatórios fraturados, como o estudado, devido ao alto número de feições a serem identificadas. Fang et al. (2016) sugerem que o limite de detecção de estruturas individuais no volume de *ant*-*tracking* seja de 200m, assim, apenas descontinuidades que tenham extensão superior a 200m em alguma direção são consideradas no processo de modelamento, embora altos valores de *ant*-*tracking* ainda apresentam alta correlação com densidade de fraturas mesmo em regiões onde as descontinuidades não podem ser precisamente discretizadas devido à resolução do atributo.

Com um volume de realce de descontinuidades como o volume de *ant*-*tracking* é possível caracterizar áreas do volume sísmico como continuas ou descontinuas de forma objetiva e veloz, permitindo realizações iterativas dos processos para melhor definição de parâmetros. Assim dando maior robustez e um maior nível de controle de qualidade aos dados gerados. A figura 8 traz o fluxo de trabalho estabelecido aplicado à uma seção arbitrária do volume sísmico disponibilizado para este estudo.

****

**Figura 8 - Visualização do fluxo de trabalho utilizado identificando uma descontinuidade.**



**Tabela 4 - Definição de parâmetros para o atributo de ant-tracking e valores utilizados no processamento do atributo. Os valores tomados para cada parâmetro seguem os valores sugeridos na literatura para reservatórios fraturados. Baseado em Schlumberger (2015).**

* 1. **Estabelecimento do modelo estrutural.**

Uma vez que descontinuidades sejam individualizadas pelo processo de *ant*-*tracking* é possível integrar estas descontinuidades num modelo estrutural do reservatório. A definição de um modelo estrutural é o processo da combinação de dados de interpretação de estruturas e horizontes para formar um modelo volumétrico entre as superfícies interpretadas (Figura 9).



**Figura 9 - Exemplo de fluxo de trabalho para modelamento estrutural. Os dados de interpretações de estruturas e horizontes são inicialmente carregados (a) para então horizontes serem modificados de acordo com sua relação com estruturas que os interceptam (b) e finalmente é gerado um modelo volumétrico de zonas com base nas superfícies modeladas (c). Alterado de Schlumberger (2015).**

A primeira etapa no estabelecimento de um modelo estrutural é a definição da geometria do *grid* onde o modelo será gerado, em seguida estabelece-se relações entre as estruturas interpretadas em suas intersecções e finalmente definem-se as relações entre estruturas e horizontes em suas interseções.

Neste estudo, o modelo estrutural gerado utiliza de *grids* de mesma geometria do volume sísmico, cobrindo a área total com disponibilidade de dados de sísmica e, consequentemente, de *ant*-*tracking*. O grid utilizado para o modelamento de estruturas utiliza células cúbicas de 100m de lado, totalizando pouco menos de 40 mil células. Enquanto o *grid* utilizado para modelar os horizontes é composto por células bidimensionais paralelas ao horizonte de 25m de comprimento na direção dos *inlines* e 50m na direção dos *crosslines,* totalizando pouco mais de 22 mil células por horizonte.

A etapa de relações entre interseções estrutura-estrutura define-se pela classificação para cada intersecção de qual estrutura encontra-se truncada por qual. Ao mesmo tempo, a etapa de relações entre interseções estrutura-horizonte define-se pelo estabelecimento de direções de movimentação em cada lado da estrutura.

Uma vez que o modelo inclui 411 estruturas diferentes, o processamento de relações em interseções é feito de maneira automática. Para as interseções estrutura-estrutura a estrutura de maior área superficial é considerada como a dominante, com a estrutura de menor área sendo truncada. Justifica-se este processo pelo fato de estruturas de maior área serem associadas a maiores valores de *ant*-*tracking*, refletindo assim em uma região provavelmente associada a maiores níveis de fraturamento. Para as interseções estrutura-horizonte, os horizontes são modificados pelas estruturas de acordo com as propriedades de mergulho e direção de movimentação obtidas na extração automática de descontinuidades (Figura 10). Com as superfícies modeladas, tem-se definidas as delimitações do modelo, sendo possível gerar um modelo de zonas para a área estudada e a definição de volumes de rocha em cada zona. A resolução inicial do modelo volumétrico é a mesma do *grid* de estruturas, sendo refinado posteriormente para o modelamento de propriedades ao longo das zonas.



**Figura 10 - Horizonte do topo Quissamã. (a) Mostra o horizonte originalmente carregado. (b) Mostra o horizonte modificado pelas interseções com estruturas. Áreas em azul representam superfícies de feições estruturais.**

* 1. **Definição de Camadas e *Upscaling* de L*ogs* de Poços**

Devido a diferença de resolução entre os dados de poço e o modelo construído, antes que o modelo possa ser populado com valores para cada propriedade, é necessária que seja realizado o *upscaling* das propriedades de poço para a escala do modelo.

O primeiro passo para realização do *upscaling* é a definição do *grid* em que as propriedades serão modeladas. A resolução horizontal do *grid* é mantida igual a do *grid* de horizontes, com células de 25m por 50m, a resolução vertical do *grid* é determinada pelo número de camadas no modelo. As camadas do modelo proporcionam os limites verticais do modelo e podem ser estabelecidas para cada zona. Neste estudo opta-se pela utilização de camadas paralelas aos horizontes e com 5 metros de espessura para capturar com maior detalhe variações verticais nos dados, visto que a fonte principal de dados a serem modelados são logs de poço (Figura 11). O *grid* após a delimitação das camadas possui pouco menos de 12 milhões de células.



**Figura 11 – (a) Modelo volumétrico de zonas com camadas definidas e (b) Zona do Quissamã entre os horizontes do Quissamã e do Lagoa Feia isolado do resto do modelo.**

Com o *grid* definido, um segundo grid unidimensional vertical é definido com células correspondentes a cada camada. As células deste *grid* são posicionadas ao longo do caminho de cada um dos poços estudados e um processo de média móvel ao longo das células é utilizado para o *upscaling* dos dados (Figura 12). Ou seja, para cada *log* em cada poço os valores atribuídos a cada célula correspondem a média dos valores deste *log* no espaço amostral correspondente à célula. O resultado deste processo é uma série de grid colunares para cada propriedade a ser modelada espalhadas pelo volume do modelo de camadas (Figura 13).



**Figura 12 – Diagrama do funcionamento do algoritmo de upscaling de dados de poço para a escala do modelo. Para cada célula do grid colunar é atribuído um valor para cada log correspondente ao valor médio do log no espaço amostral correspondente à célula.**



**Figura 13 - Grid colunares de porosidade efetiva (PHIE) cortados para a região entre o topo do Quissamã (superfície em roxo) e o contato óleo-água (superfície em preto).**

* 1. **Modelamento de Propriedades**

O modelamento de propriedades trata-se da utilização de técnicas de geoestatística para a população do modelo com valores para as propriedades de poço após o *upscaling* para a escala do modelo de camadas. Previamente a seleção de uma técnica de modelamento é necessário análise dos dados de input do modelamento, uma vez que cada técnica parte de uma série de pressupostos sobre o dado original.

Dado que o modelo estrutural e de camadas definidos anteriormente buscam o estabelecimento do modelo físico do reservatório, assume-se que o modelo físico é conhecido. Ao mesmo tempo, tem-se a informação a priori de que o modelamento é realizado em rocha reservatório, assim assume-se que existe auto correlação espacial das variáveis.

Técnicas de Krigagem podem ser definidas por estimativas de valores em um volume embasadas numa ponderação de todas as amostras, tendo que o peso de cada amostra é obtido com a condição restritiva de que a soma dos pesos seja igual a 1 e a variância das estimativas seja mínima (Oliveira, 1997). Tal que, para um conjunto de n amostras {,, ...,}:

|  |  |  |
| --- | --- | --- |
|  |  | (3.5.1) |
|  |  |  |

onde é o valor estimado para cada localização S do grid ou modelo, é um peso desconhecido para cada i-ésima amostra . No caso de interpolações baseadas no inverso da distância entre pontos, o peso depende apenas da distância entre cada local S e as amostras. Para métodos de krigagem os pesos são obtidos a partir de estimativas originadas da regressão linear entre os pontos em um variogramas (ESRI, 2019).

A Krigagem com Deriva Externa (KDE) é um algoritmo derivado da Krigagem Ordinária por Deutsch & Journel (1992) e Wackernagel (1995) mostra-se como ferramenta de grande utilidade para estimativas em reservatórios onde evidencia-se auto correlação espacial dos atributos a serem modelados (Hengl et al., 2007: Zanão, 2008). Segundo Hengl (2009) principal diferencial do algoritmo de Krigagem com Deriva Externa em relação a Krigagem ordinária se dá na prévia utilização de uma regressão linear entre as amostras para então utilizar os erros da regressão linear (chamados de residuais ou deriva) como valores para a krigagem (Figura 14).

No entanto, a utilização do algoritmo de Função Gaussiana Aleatória de Simulação (Gaussian Random Function Simulation - GRFS) incorpora algoritmos de KDE com simulações condicionais no campo não estruturado dos variogramas, honrando de maneira mais efetiva variações locais no dado de entrada (Schlumberger, 2015; Hansen et al., 2006).



**Figura 14 - Ilustração da obtenção de valores residuais para métodos de Krigagem com Deriva Externa (aqui denominado como regression-kriging). De: Hengl (2009).**

.

Assumindo que os dados de entrada do modelamento estejam sob distribuição normal, o algoritmo de GRFS mostra-se superior a algoritmos de Krigagem pura pois é um algoritmo estacionário no domínio espacial do dado, mantendo valores de média e variância constantes ao longo do modelo, dando realce a variâncias locais (Daly & Caers, 2010; Dongas, 2016; Benabbou et al., 2015). Foram utilizados nos modelos variogramas com direção azimutal de 30 graus, direção paralela ao eixo da anticlinal em que o reservatório é incluso e aproximadamente paralela à direção dos *crosslines*. Os variogramas também são do tipo esférico e são caracterizados por distância máxima de anisotropia de 3500m horizontalmente (aproximadamente metade da menor dimensão do modelo) e 10m na direção vertical (equivalente a 2 células no grid colunar) Ao mesmo tempo, a distribuição dos dados de poço comparada com a distribuição dos dados após *upscaling* e dos dados após o modelamento é utilizada como controle de qualidade dos dados de modelamento.

Propriedades geométricas como espessura do reservatório e volume de célula também são computadas. Propriedades geométricas são medidas simples de distâncias e volumes de cada célula com referência a outro atributo do modelo. Também é computada uma propriedade baseada no volume de *ant*-*tracking* tomando valores médios do atributo sísmico para cada célula e a partir desta propriedade baseada em ant-tracking avalia-se a distribuição de descontinuidades no modelo.

* 1. **Classificação de Reservatório e Cálculos de Volume de Hidrocarbonetos.**

Uma vez que as células do modelo são populadas com valores para cada uma das propriedades, é possível classificar o modelo de acordo com qualquer parâmetro de classificação baseado nestas propriedades.

Tradicionalmente, zonas de interesse em reservatórios são definidas pela definição binária de zonas dos logs de poço em zonas de *net pay* ou não, ou seja, uma classificação binária sobre a viabilidade de produtividade de uma porção do log de poço. Esta classificação é feita atribuindo valores de corte mínimos para cada um dos logs utilizados na estimativa (geralmente porosidade e saturação de água) e considerando como não viável qualquer área em que os valores para um dos logs seja inferior ao valor de corte definido para esta medida. Embora largamente utilizado, este método tem suas limitações devido à natureza booleana das classificações assim como introduz um grau de subjetividade na definição dos valores de corte (Cobb et al, 1998).

Dada a existência de um modelo espacial para as propriedades de saturação de água e porosidade como o obtido no processo de modelamento, este estudo busca realizar a classificação de zonas de interesse de maneiras distintas que aproveitem a disponibilidade deste modelo. Em seguida, calcula-se os valores de volume de hidrocarbonetos em espaço poroso (*Hydrocarbon Pore Volume* – HCPV) para as zonas definidas por cada um dos métodos utilizados, dado que o cálculo de volume é definido por:

|  |  |  |
| --- | --- | --- |
|  |  | (3.6.1) |
|  |  |  |

onde HCPV é o volume de hidrocarbonetos, n é o total de células do modelo, (*bulk volume*) é o volume total da célula, é o valor da propriedade de porosidade efetiva para a célula, é o valor da propriedade de saturação de água para a célula e (*pay volume*) é o volume da célula considerado como *net* *pay*. A relação entre é chamada de razão *Net-to-Gross,* ou N/G*.* Nos métodos de classificação booleanos tradicionais, a razão N/G é definida como 0 ou 1 dependendo dos valores de corte definidos previamente, ou seja, cada célula contribui de maneira completa ou não contribui para o cálculo de volumes. Entre os métodos sugeridos neste estudo, alguns deles utilizam razões N/G booleanas enquanto outros usam razões variáveis de 0 a 1. Para reduzir o número de saídas dos cálculos de volume, na etapa de avaliação dos métodos sugeridos, os valores de saturação de água sempre correspondem à curva do cenário C2, pois esta apresenta maior grau de correlação com os dados de testemunho do Campo B (Melani et al, 2015).

Cada um dos métodos apresentados tem então cálculos de HCPV realizados com base em sua razão N/G. Os volumes são então comparados com cálculos de HCPV obtidos com base nas classificações de perfis de poço em *net pay* ou não realizadas por Melani et al. (2015). Para o cálculo de volumes baseados nos estudos de Melani et al. (2015) realiza-se o *upscaling* das classificações de poço para um grid colunar assim como descrito em 3.4, exceto que, para manter a natureza binária da propriedade, o valor da célula é a moda no espaço amostral e não a média. O modelo é populado por uma propriedade onde cada célula tem um valor de razão N/G booleano correspondente ao valor da célula mais próxima no grid colunar.

* + 1. **Clusterização de células por aprendizado de máquina não supervisionado.**

Dado que o objetivo da definição de valores de corte é delimitar áreas do reservatório que representem maior viabilidade de produtividade, a separação de partes do reservatório (neste caso, as células) em viáveis ou não se trata de um problema de classificação discreta. Assim, é possível diminuir o fator de subjetividade na definição de valores de corte pela clusterização das células do modelo em diferentes classes baseadas nos valores modelados para cada propriedade.

Tendo em mente que o método K-Means é escalável para grandes bases de dados e produz resultados esféricos condizentes com o princípio geoestatístico de autocorrelação espacial (Romary et al, 2012), além de ser comumente usados para classificação de dados de reservatório em diferentes eletrofácies (Schlumberger, 2015), utiliza-se do método para separação das células em diferentes classes.

O método de clusterização K-Means é um método de classificação não supervisionado que classifica os dados de entrada em K diferentes classes. O método K-Means trata-se de um método iterativo, inicialmente escolhendo K pontos aleatórios dentro do domínio dos dados (sendo K o número de classes) e designando uma classe a cada ponto de acordo com a menor distância euclidiana do ponto aos centroides. Estes pontos aleatórios são tidos como os centroides iniciais de cada classe e tem sua posição iterativamente atualizada tentando maximizar a distância entre os centroides e minimizar a variância entre os pontos pertencentes à mesma classe (Lloyd, 1982).

Matematicamente, o algoritmo do K-Means é definido por uma etapa inicial de atribuição de classes a cada ponto equivalente à construção de um Diagrama de Voronoi utilizando os centroides iniciais como pontos de referência, tal que, para cada amostra x num conjunto {:

|  |  |  |
| --- | --- | --- |
|  |  | (3.6.3.1) |

tal que a resposta para cada x é um vetor V de comprimento k:

|  |  |  |
| --- | --- | --- |
|  |  | (3.6.3.2) |

e a classe C assinalada a cada ponto se dá na forma:

|  |  |  |
| --- | --- | --- |
|  |  | (3.6.3.3) |

representa uma variável booleana indicando se cada ponto pertence a cada classe k. k é o número de classes e representa a distância euclidiana entre elemento x e cada i-ésimo centroide . Como V é um vetor com k-1 elementos 0 e apena um único elemento 1, a posição do elemento 1 será a classe assinalada à amostra x. Em seguida, a posição de cada centroide é atualizada e, para cada classe, o novo centroide é posicionado na posição média de todos os pontos pertencentes à aquela classe. Estas duas etapas são repetidas iterativamente até que nenhum ponto troque de classe entre duas iterações, momento em qual o algoritmo é considerado como convergente, ou até que se alcance um número máximo de iterações pré-definido.

A escolha do número de classes para classificação dos dados permanece como principal fonte de subjetividade no método. Busca-se ao mesmo maximizar a variância entre as diferentes classes e ao mesmo tempo minimizar a variância dentro de cada classe. Quanto maior o número de classes, menor será a variância intra-classe de cada classe e maior será a variância inter-classe, no entanto, quanto maior o número de classes, menos significativa é a classificação obtida. Por exemplo, com N amostras, ao considerarmos N classes, a variância intra-classe é 0 e a variância inter-classe é a variância total da amostra, mas um modelo com N classes não introduz qualquer informação nova. Pode-se reduzir a subjetividade desta medida ao se computar o algoritmo com diferentes valores de K e registrar os valores em cada iteração para os índices de Davies-Bouldin e de silhueta. O índice de Davies-Bouldin tem como parâmetros a relação inversa entre a distância entre os centroides de cada classe e a variância interna de cada classe, assim, para valores menores que 1, o índice representa boa separação entre as classes (Kuroda et al, 2012). O índice de silhueta é uma média da relação para cada ponto da distância entre o ponto e os pontos de outras classes e a distância entre o ponto e os pontos de sua mesma classe, assim, quanto maior o índice de silhueta de um classificador, melhor os pontos estão representados dentro de suas classes.

Embora o método K-Means ofereça uma classificação robusta das células em diferentes classes, o K-Means ainda tem a limitação de oferecer classificações discretas para um número pré-determinado de classes, portando, seu uso na estimação de N/G é limitado a valores booleanos.

Outra limitação do K-Means, diretamente relacionada à sua natureza discreta, é a incapacidade do algoritmo de classificar pontos de natureza ambígua (Raykov et al., 2016), ou seja, para pontos próximos a fronteira de 2 classes, embora o K-Means consiga determinar qual o centroide mais próximo, o algoritmo não fornece qualquer tipo de informação sobre os outros centroides.

Modelos de mistura gaussiana (Gaussian Mixture Models – GMM), embora caracterizados por funções de densidade de probabilidades de diferentes subpopulações dentro de uma maior população, podem ser utilizados como modelos de classificação com usos similares ao K-Means. Considerando K classes numa distribuição de valores como K subpopulações dentro do conjunto de valores, é possível utilizar o GMM para obter a função de densidade de probabilidade de cada classe e, consequentemente, a probabilidade de cada ponto pertencer a cada classe. Em seguida, classifica-se cada ponto de acordo com a classe de maior probabilidade. Assumindo que se conhece o número de classes ou subpopulações em uma distribuição, o GMM estima máxima verossimilhança por um algoritmo de máxima expectativa (VanderPlas, 2016).

Assim, este estudo utiliza o K-Means como forma de iterativamente encontrar o número ótimo de agrupamentos nos dados, em seguida utilizando GMM para a classificação em K diferentes classes dos dados modelados.

Devido ao GMM ter como saídas as probabilidades de cada ponto pertencer a cada classe, é possível estabelecer valores não booleanos para a razão N/G. Assim, calcula-se também HCPV considerando a probabilidade de cada ponto pertencer a classe de maior porosidade e menor saturação de água como a razão N/G da célula. De maneira similar, também se calcula HCPV tomando a razão N/G de cada célula como a probabilidade de o ponto em questão não pertencer a classe de menor porosidade e maior saturação de água.

Para aplicação dos métodos aos dados modelados, inicialmente foi desenvolvida função em Python para conversão dos dados de cada propriedade de rocha (Porosidade Efetiva, Perfil Sônico, Densidade e Raio Gama) e fluído (Saturação de Água) do formato nativo .grdecl do software Schlumberger Petrel para arrays da biblioteca open source NumPy (Oliphant, 2006) e convertidos para tabelas .csv com a utilização das biblioteca Pandas (McKinney, 2010; van der Walt et al., 2011).

Na aplicação dos métodos K-Means e GMM aos dados e no cálculo dos índices de silhueta e Davis-Bouldin também foram desenvolvidos scripts em Python, utilizando as bibliotecas citadas anteriormente e também a biblioteca Scikit-Learn (Pedregosa et al., 2011). Utiliza-se também as bibliotecas Matplotlib (Hunter, 2007) e Seaborn para visualizações.

Em seguida, toma-se a classificação discreta das células com base nas propriedades de rocha em diferentes classes como uma medida booleana de N/G, com células pertencentes à classe de maiores porosidades médias consideradas com N/G de 1, e células pertencentes às demais classes com N/G de 0. Ao mesmo tempo, calcula-se também o volume das células de cada classe, possibilitando a utilização de outras classes com N/G em 1.

1. **Resultados e Discussão**

Nesta seção do estudo os resultados obtidos dos métodos apresentados anteriormente são apresentados na mesma ordem em que os métodos são apresentados. Ao mesmo tempo, também acompanham estes resultados discussões sobre as informações geradas e seu impacto nas etapas seguintes.

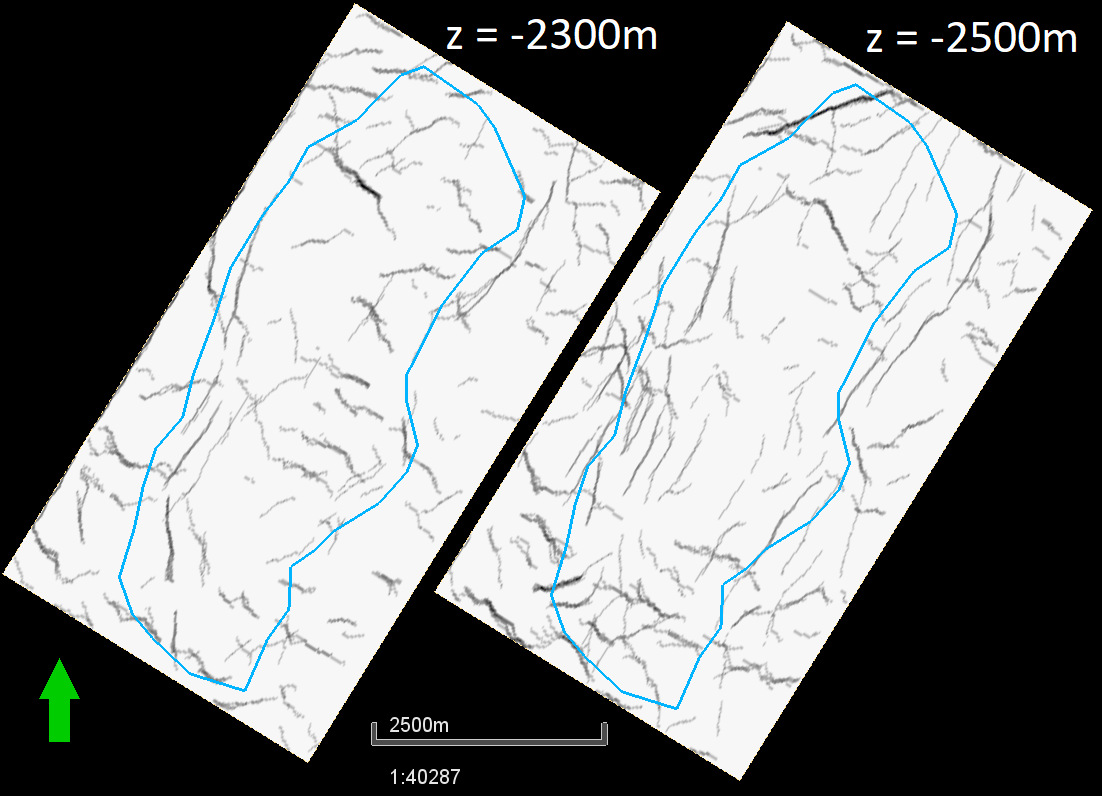
* 1. **Processamento Sísmico e Descontinuidades**

A etapa de processamento de atributos sísmicos tem como resultado o volume de *ant*-*tracking* bem como a individualização de estruturas identificadas no volume de *ant*-*tracking*, sendo possível observar que descontinuidades ocorrem de maneira disseminada no volume sísmico analisado (Figura 15; Figura 16).

Individualizando as descontinuidades obtidas pelo método do *ant*-*tracking*, observa-se uma predominância de mergulhos sub-verticais a verticais com direção azimutal aproximadamente E-W (Figura 17). Filtrando as descontinuidades obtidas mantendo apenas descontinuidades com extensão mínima de 200 metros em qualquer direção é possível observar a manutenção das tendências estruturais observadas (Figura 18), fornecendo evidência de que as estruturas mantem geometria similar independentemente da escala em que ocorrem.



**Figura 15 - Atributo de Ant-Tracking no crossline 1744. Refletor entre 2200m e 2400m de profundidade corresponde ao horizonte do topo do Quissamã. Refletor entre 2600m e 2800m de profundidade corresponde ao horizonte do topo do Lagoa Feia. Nota-se a ocorrência de forma disseminada de descontinuidades detectadas pelo fluxo de trabalho definido.**



**Figura 16 - Atributo de Ant-Tracking nos cortes em profundidade de 2300 e 2500m. 2300m sendo a profundidade aproximada do topo do Quissamã e 2500m a profundidade aproximada do contato óleo-água. Polígono em azul representa os limites do reservatório em superfície.**



**Figura 17 - Distribuição de descontinuidades para todas as descontinuidades. Estruturas predominantemente Leste-Oeste com mergulhos para Norte e Sul.**

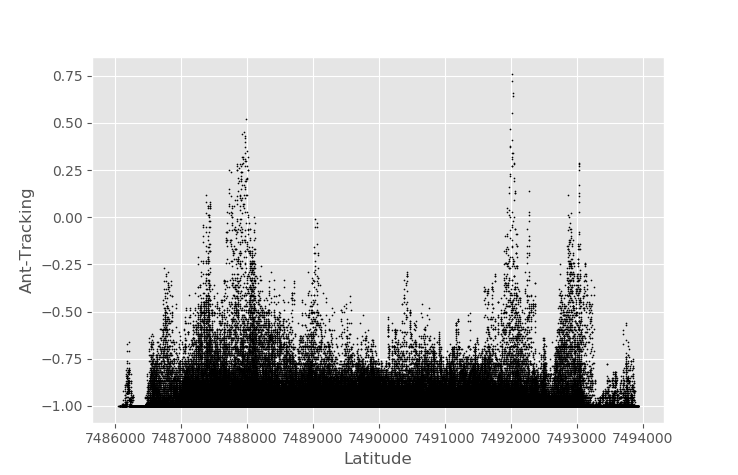


**Figura 18 - Distribuição de descontinuidades com mais de 200m de extensão em qualquer direção. Nota-se a manutenção da tendência estrutural observada para o conjunto de todas as descontinuidades.**

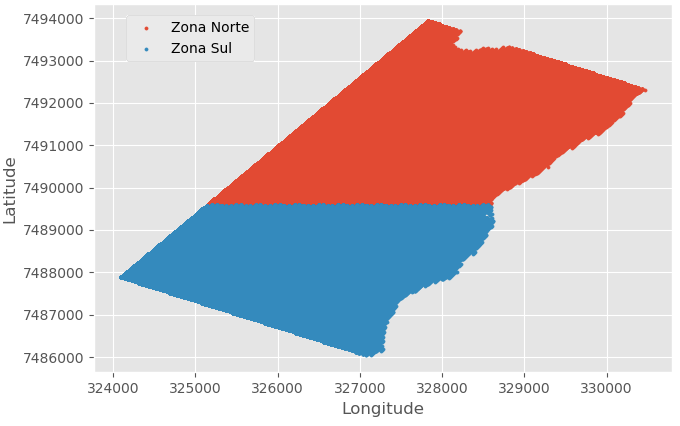
Franz, et al (1987) descreve o Campo B como inserido em uma anticlinal de eixo NE-SW localizada entre duas grandes falhas regionais paralelas ao eixo de dobra. Nota-se que a maior parte das descontinuidades identificadas na utilização do método de *ant-tracking* são superfície perpendiculares ao eixo da anticlinal (Figura 17; Figura 18), fornecendo possível indicador de que o evento de dobramento originador do anticlinal não é também originador das falhas e fraturas do Campo B. Corroborando assim com a teoria para a evolução estrutural de reservatórios da Bacia de Campos proposta por Franz (1987) onde são identificados ao menos dois estágios de fraturamento.

O estudo de Tomaso et al (2013) divide o Campo B em duas zonas distintas de acordo com intensidade de fraturamento, sendo estas uma zona sul mais fraturada e com microporosidade; e uma zona norte, com porosidade primária e pouco fraturada. Considerando esta divisão e assumindo o Ant-Tracking como uma medida do grau de fraturamento da rocha, observa-se a variação dos valores de Ant-Tracking de acordo com as coordenadas de cada célula (Figura 19). Também se separa o reservatório em zona sul e zona norte utilizando a mediana das coordenadas como fronteira (Figura 20) e observando a distribuição espacial dos os valores de Ant-Tracking nas 2 zonas (Tabela 5; Figura 21), assegurando assim que as duas zonas tenham o mesmo número de células.

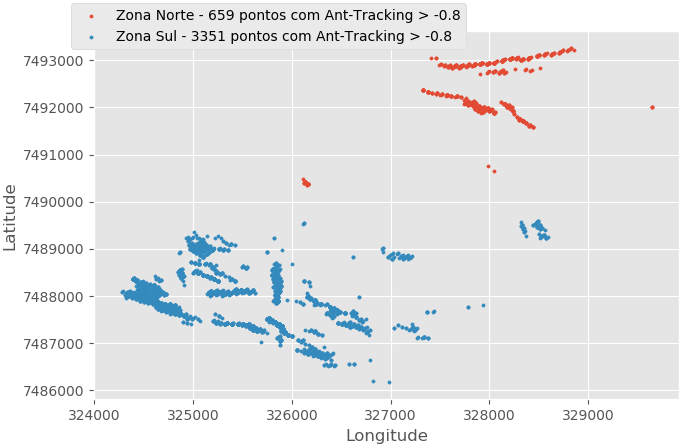
Não é possível identificar uma relação direta entre a variação de valores de Ant-Tracking e Latitudes e tão pouco é possível observar valores de Ant-Tracking suficientemente maiores em média na Zona Sul em relação à outras zonas (Figura 19; Tabela 5). No entanto, considerando que o atributo de ant-tracking tem valores entre -1 e 1 para cada célula tal que -1 representa um volume continuo enquanto 1 representa volume inteiramente descontinuo e que medidas de grau de fraturamento obtidas de testemunhos correlacionam-se com valores superiores a -0.8 (Godfrey & Bachrach, 2008), observa-se a distribuição destes pontos ao longo do reservatório (Figura 21; Figura 22).



**Figura 19 - Variação de Ant-Tracking com Latitude. Não se observa tendência clara na distribuição entre as duas variáveis, embora a concentração de altos valores em picos indique a existência de feições de descontinuidade nestas latitudes. Cada ponto no gráfico corresponde à uma célula no modelo. Consideradas apenas células dentro do reservatório (Entre Quissamã e contato óleo-água).**

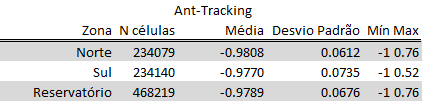


**Figura 20 - Delimitação de Zonas Norte e Sul do reservatório, cada uma com 234.109 células. Fronteira entre as zonas encontra-se deslocada para o sul devido à maior espessura do reservatório na Zona Sul. Cada ponto no gráfico corresponde à uma célula no modelo. Consideradas apenas células dentro do reservatório (Entre Quissamã e contato óleo-água).**



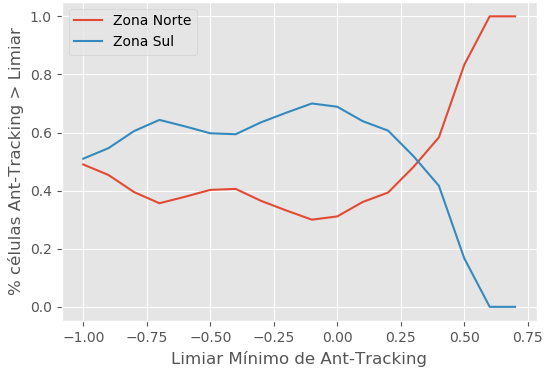
**Figura 21 - Distribuição no espaço para os pontos de altos valores de Ant-Tracking. Cada ponto no gráfico corresponde à uma célula no modelo. Pontos plotados independentemente de sua posição vertical. Consideradas apenas células dentro do reservatório (Entre Quissamã e contato óleo-água)**

**Tabela 5 - Distribuição dos valores de Ant-Tracking nas diferentes Zonas do volume sísmico. Embora a Zona Sul mostre um valor médio de Ant-Tracking superior tanto à Zona Norte quanto a região não reservatório do Quissamã, a diferença observada na média é menor do que o desvio padrão destas medidas.**



A Zona Sul mostra uma quantidade de células (3351 células) associadas a zonas de fraturamento consideravelmente maior que a Zona Norte (659). De maneira similar, na Zona Norte células com altos valores de ant-tracking ocorrem primariamente associadas a feições lineares de grande escala enquanto na Zona Sul estes valores ocorrem de maneira disseminada ao longo do volume. O valor médio de ant-tracking aproximadamente igual para as duas zonas também fornece evidência de que os altos valores na zona norte se encontram concentrados ao redor das mesmas grandes estruturas.

Os resultados obtidos corroboram com as observações de Tomaso et al. (2013) descrevendo uma Zona Sul com predominância de fraturas e uma região norte não fraturada; sendo discutidos em maior detalhe e levando em conta os resultados das etapas posteriores no Capitulo 5 – Conclusões. Observa-se ainda que ao assumir diferentes *cut-offs* mínimos de ant-tracking e analisar a que zona cada célula remanescente pertence é possível representar de maneira simples este zoneamento (Figura 22). As células com valores de ant-tracking associados às regiões de fratura (-0.8 a 0) encontram-se primariamente na Zona Sul enquanto os valores associados a regiões de intenso faturamento (valores superiores a 0 e se aproximando a 1) concentram-se na Zona Norte, refletindo as observações da literatura sobre faturamento disseminado ocorrendo na Zona Sul e a Zona Norte sendo largamente não fraturada exceto em regiões associadas a estruturas de larga escala.



**Figura 22 - Porcentagem de células localizadas na Zona Norte ou Sul para cada valor mínimo de Ant-Tracking (Limiar). Nota-se que para Ant-Tracking de -1,0 os valores se distribuem igualmente entre as duas zonas, refletindo a separação das duas zonas pela Latitude mediana do reservatório. Para valores entre -0.8 e 0, valores que se correlacionam zonas de fraturamento, 60% a 75% dos valores encontram-se na Zona Sul. Para os poucos (n = 151) valores positivos de Ant-Tracking, estes concentram-se fortemente na Zona Norte. Assim indicando que as descontinuidades ocorrem de maneira mais disseminada pela Zona Sul enquanto concentram-se em uma área de faturamento intenso na Zona Norte.**

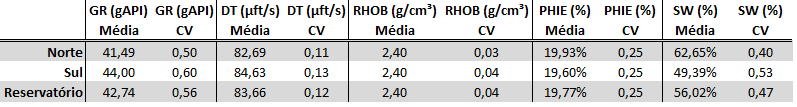
* 1. ***Upscaling* e Modelamento de Propriedades**

Utilizando o método de *upscaling* por média móvel apresentado anteriormente (Figura 12), os dados de petrofísica de poço foram transformados da escala de poço para a escala do modelo estrutural, populando células que interceptam os poços e criando grids colunares. A distribuição estatística dos dados antes e depois da transformação é utilizada para validar os resultados do *upscaling*. Em seguida é feito o modelamento geoestatístico destas propriedades para as zonas inter-poços, populando o modelo de reservatório com valores para as medidas petrofísicas de poço.

As propriedades modeladas no estudo são propriedades de rocha como porosidade efetiva (PHIE), densidade (RHOB), perfil sônico (DT) e raio gama (GR), além de modelos de saturação de água (SW) também baseados em dados de petrofísica de poço. As propriedades foram modeladas utilizando o algoritmo de GRFS (Gaussian Random Function Simulation, GRFS), onde valores na porção estruturada dos variogramas são modelados usando técnicas de krigagem e valores na porção não estruturada dos variogramas são definidos por uma simulação condicional baseada na distribuição dos dados de entrada. Os resultados são em seguida validados comparando a distribuição estatística dos dados modelados aos dados do *upscaling* e os dados originais de poço.

Os dados modelados são apresentados em mapas e perfis paralelos à direção dos *crosslines*, cortados para o reservatório (sobre o contato óleo-água). Para as visualizações em mapa, os dados apresentados correspondem aos dados na superfície do topo da Formação Quissamã, não representando uma fatia de profundidade, o polígono em vermelho representa os limites em superfície do reservatório e a linha laranja representa a seção onde é traçada a visualização em perfil. Os resultados para o modelamento e upscaling de cada propriedade são apresentados separadamente e depois discute-se correlações entre os dados e observações gerais sobre os resultados. A tabela 6 apresenta estatísticas básicas sobre os dados modelados para diferentes zonas.

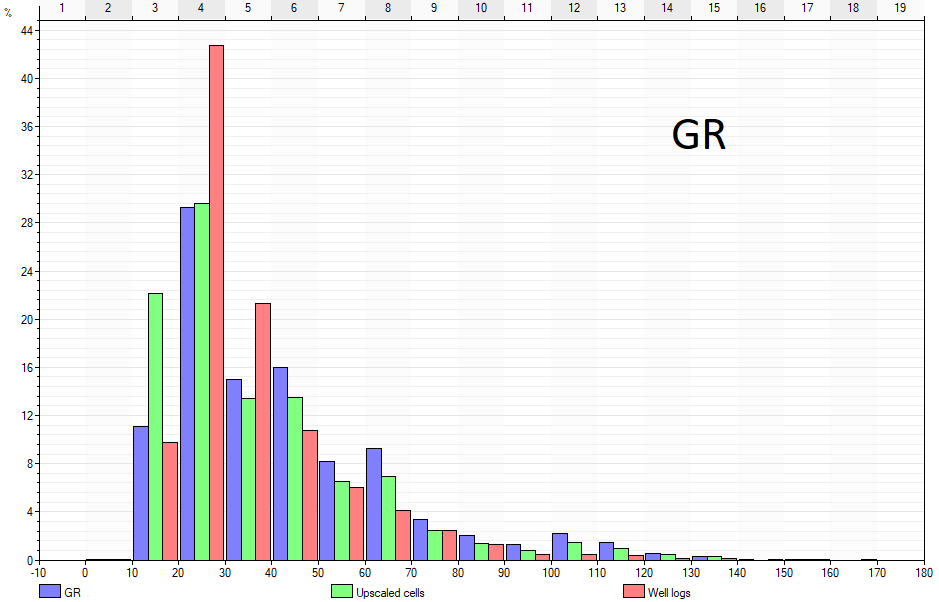
**Tabela 6 - Valores estatísticos para a distribuição dos valores modelados. CV é o coeficiente de variação.**



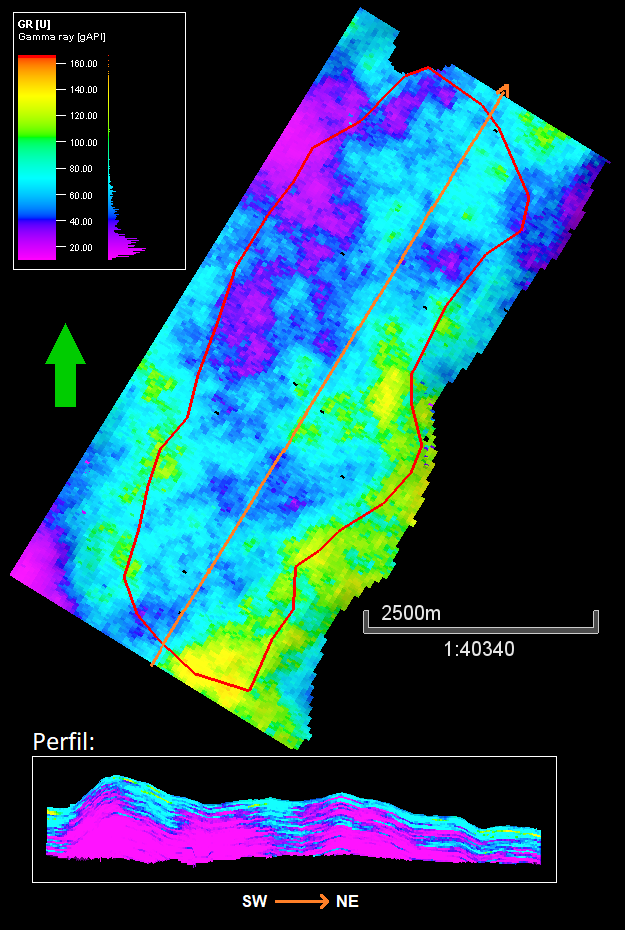
* + 1. **Modelos de Raio Gama (GR).**

O upscaling dos dados de GR, quando comparado aos dados de poço subestima a quantidade de células com valores entre 30 e 40 gAPI, no entanto, a distribuição dos dados de poço é mantida de maneira geral (Figura 23). Nas visualizações em mapa e perfil é possível observar uma transição de valores em torno de 30 gAPI em células mais profundas próximas ao contato óleo-água para valores em torno de 60 gAPI para células mais próximas ao horizonte do topo do Quissamã. Os valores médios de GR são similares na porção sul e norte do reservatório com coeficientes de variação similares nas 2 zonas (Tabela 6). No entanto, é possível diferenciar a variância nas duas zonas na visão em perfil, onde a porção norte mostra valores em torno de 60 gAPI intercalados com valores em torno de 30 gAPI enquanto a porção sul apresenta transição mais relacionada com a profundidade (Figura 24).

Os resultados do modelamento de dados de raio gama sugerem a ocorrência de uma transição de fácies, com fácies folhelho ou granulometrias mais finas cobrindo e se intercalando às sequências carbonáticas do Quissamã, em especial em regiões próximas ao topo da Formação, sugerindo início de transição para a Formação Outeiro.



**Figura 23 - Distribuição de dados de raio gama (RHOB). Os dados após upscaling subestimam valores em torno de 30 gAPI mas honram de maneira geral a distribuição original dos dados de poço. Dados modelados (em azul) mantem a distribuição dos dados de entrada (em vermelho).**

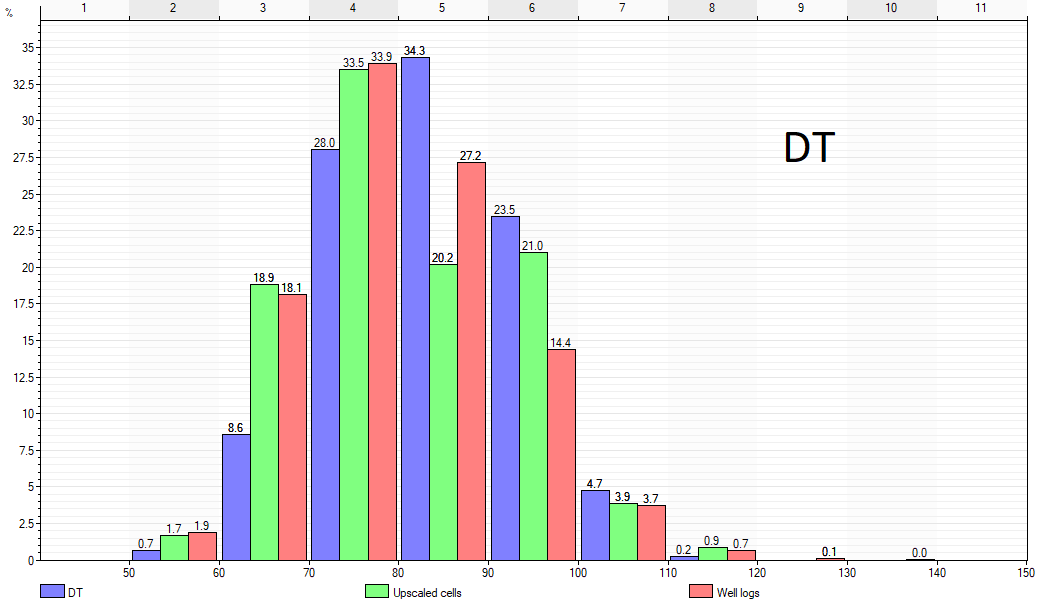


**Figura 24 - Modelo de raio gama (GR) para o reservatório. Observa-se relação entre os valores modelados e profundidade, com valores próximos ao topo da Formação Quissamã apresentando um aumento nos valores de raio gama observados, sugerindo transição dos carbonatos da Formação Quissamã para fácies de menor granulometria como folhelhos.**

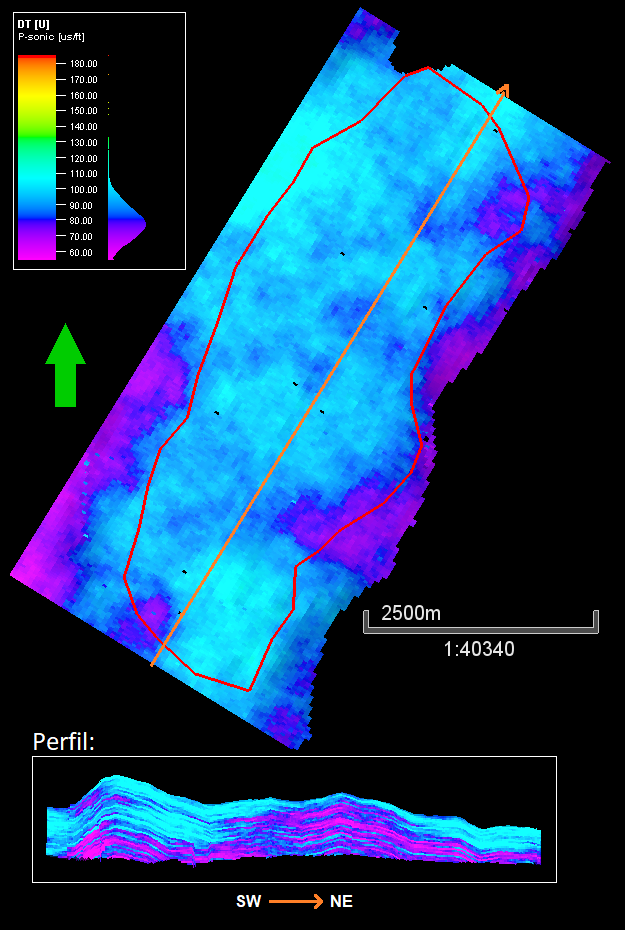
* + 1. **Modelos de Perfil Sônico (DT)**

Os dados de perfil sônico apresentam uma distribuição aproximadamente normal com média em torno de 80 μs/ft tanto para os dados de entrada como para os dados após upscaling e modelados (Figura 25). Nos modelos em mapa e perfil se observa assim como nos modelos de raio gama uma transição de baixos valores próximos ao contato óleo-água para altos valores próximos ao topo do Quissamã, bem como variância nos valores da zona norte relacionada a intercalação de camadas de altos e baixos valores (Figura 26).

Os valores obtidos para o modelo variam primariamente entre 60 e 100 μs/ft, valores consideravelmente superiores aos valores típicos entre 40 e 50 μs/ft para minerais carbonáticos como calcita e dolomita, sendo valores esperados para filosilicatos como feldspatos ou argilas. O perfil sônico, no entanto, é altamente influenciado pela porosidade, assim provavelmente explicando os valores observados. O valores em torno de 100 μs/ft podem ser explicados por relativos aumentos de porosidade ou por transições para fácies compostas primariamente por folhelhos. Nas regiões de alto DT no topo do reservatório é provável que isto seja explicado por uma transição para folhelhos, enquanto no alto estrutural na porção sul do reservatório o aumento nos valores de DT provavelmente é explicado por zonas de alta porosidade.



**Figura 25 - Distribuição de dados do perfil sônico (DT). Os dados após upscaling honram de maneira geral a distribuição original aproximadamente normal dos dados de poço. Dados modelados (em azul) mantem a distribuição dos dados de entrada (em vermelho).**

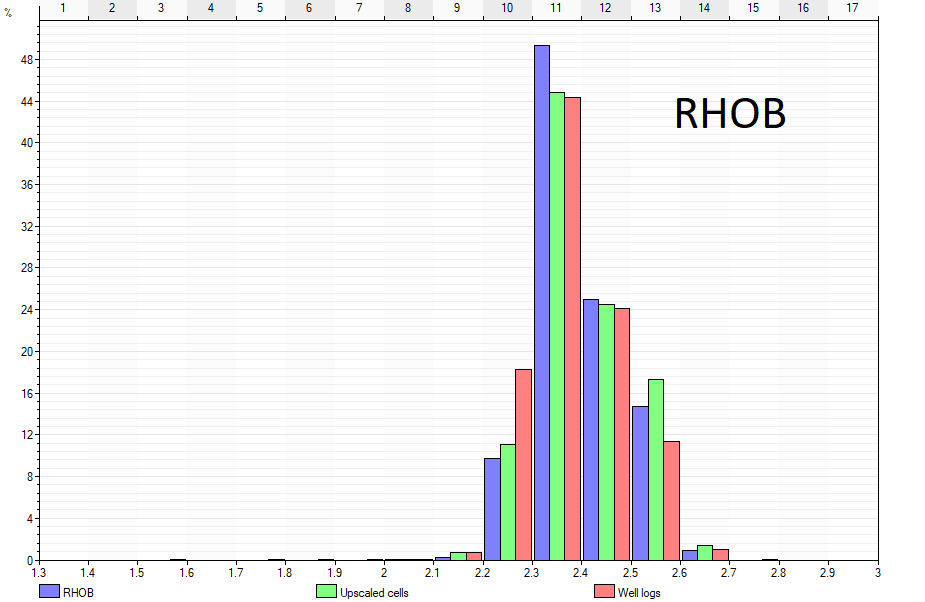


**Figura 26 - Modelo de dados de perfil sônico (DT) para o reservatório. Observa-se também relação entre os valores modelados e profundidade, com valores próximos ao topo da Formação Quissamã apresentando um aumento nos valores de DT observados, podendo associar-se tanto com variações em litologia quanto variações em porosidade.**

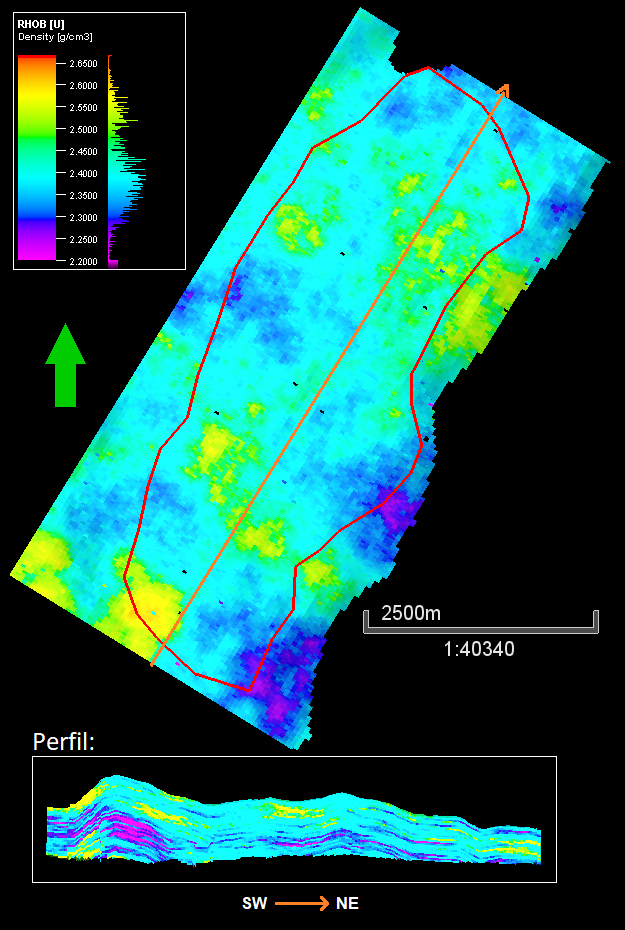
* + 1. **Modelos de Densidade (RHOB)**

Os modelos de Densidade apresentam valores principalmente concentrados entre 2,3 g/cm³ e 2,5 g/cm³, com valores mínimos de 2,1 g/cm³ e máximos de 2,7g/cm³ (Figura 27). Assim como os dados derivados de perfil sônico, dados de densidade são largamente dependentes da porosidade da rocha, sendo especialmente difícil a inferência de fácies pelo uso exclusivo dos dados de densidade devido a densidade aproximadamente similar da maior parte dos minerais encontrados comumente em reservatórios carbonáticos da Bacia de Santos. Assim como para os outros modelos, a variância dos dados na zona norte é similar a variância na zona sul (Figura 28).

Na visualização em perfil é possível notar uma região de baixa densidade associada ao alto estrutural na porção sudoeste do reservatório, sugerindo uma região de alta porosidade em meio ao reservatório. Outras camadas de baixos valores de densidade podem ser observadas ao longo do perfil, em especial na região nordeste do modelo, com pequenas intercalações de zonas de baixa densidade com alta densidade.



**Figura 27 - Distribuição de dados de densidade (RHOB). Os dados após upscaling honram de maneira geral a distribuição original e concentram-se em valores entre 2,3 g/cm³ e 2,5 g/cm³. Dados modelados (em azul) mantem a distribuição dos dados de entrada (em vermelho).**



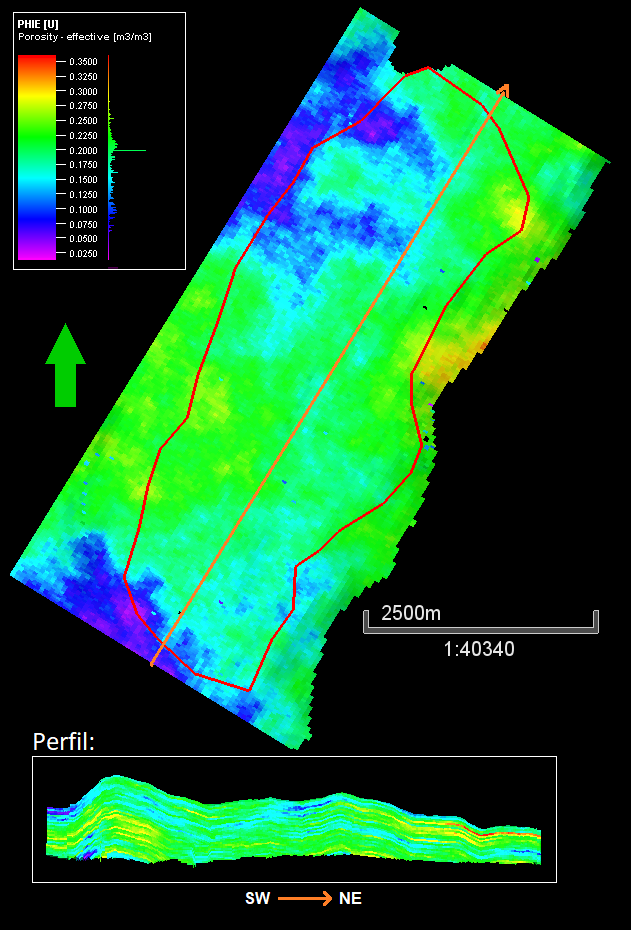
**Figura 28 - Modelo de dados de densidade (RHOB) para o reservatório. Observa-se a concentração de valores de baixa densidade associados ao alto estrutural na porção sudoeste do reservatório, provavelmente caracterizando uma zona de alta porosidade.**

* + 1. **Modelos de Porosidade Efetiva (PHIE)**

Os dados de porosidade efetiva mostram valores distribuídos de forma aproximadamente normal centrados em valores em torno de 20%, com valores mínimos próximos de 0 e máximos próximos de 50%, e assim como para os outros modelos, o método de *upscaling* honra esta distribuição (Figura 29). Na visualização em mapa é possível observar valores primariamente em torno de 20% na superfície do Quissamã com algumas zonas de baixa porosidade. Na visão em perfil é possível notar duas zonas principais de alta porosidade, uma zona associada ao alto estrutural da porção sudoeste do reservatório e outra zona na porção nordeste associada a camadas de muito alta porosidade. O modelo de porosidade efetiva também ilustra a importância da visualização em perfil, dado que a visualização em mapa indica a ocorrência de 2 zonas de muito baixa porosidade, mas o perfil revela que estes valores não se propagam em outras profundidades. Ao mesmo tempo, as camadas de alta porosidade (35%) na porção norte do reservatório resultam na diferença de valores médios entre as Zonas Sul e Norte observada na Tabela 6.



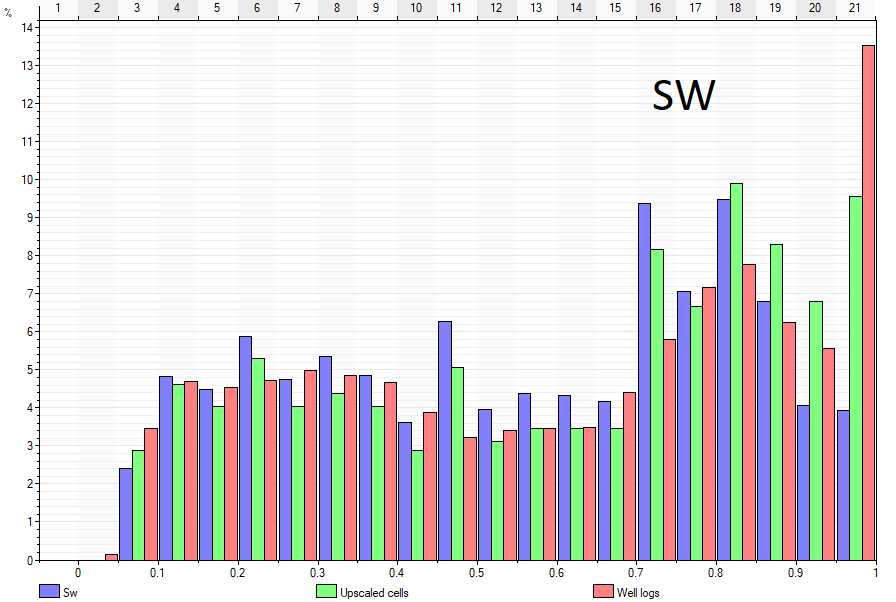
**Figura 29 - Distribuição de dados de porosidade efetiva (PHIE). Os dados após upscaling mantem a distribuição normal concentrada em valores em torno de 20%. Dados modelados (em azul) mantem a distribuição dos dados de entrada (em vermelho).**



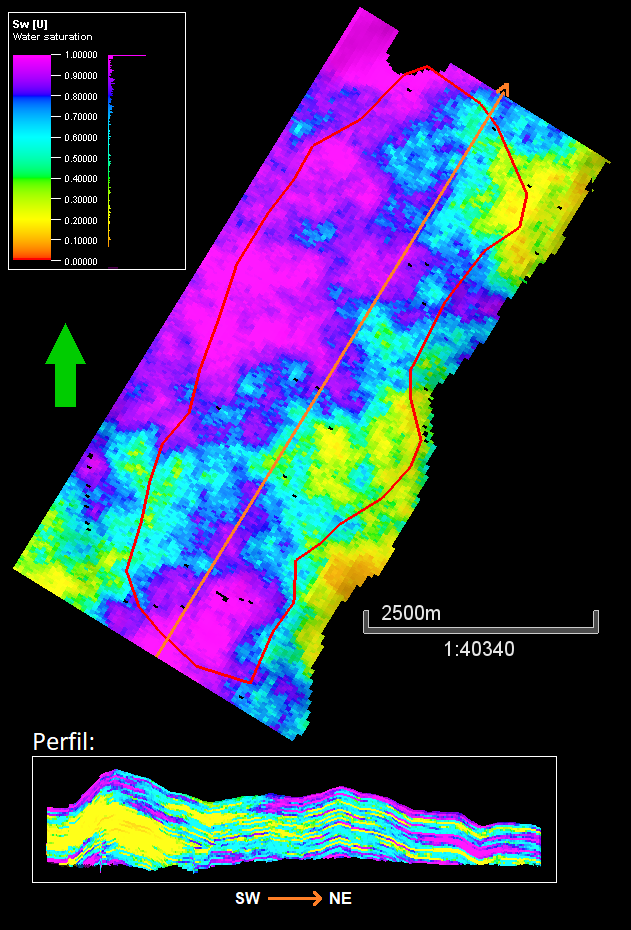
**Figura 30 - Modelo de dados de porosidade efetiva (PHIE) para o reservatório. Observa-se a concentração de valores de alta porosidade associados ao alto estrutural na porção sudoeste do reservatório e a um pico nos valores na porção nordeste.**

* + 1. **Modelos de saturação de água (SW)**

A distribuição dos dados de saturação de água ocorre de maneira bastante heterogenia, ainda assim os dados modelados mantem a distribuição original dos dados de poço, subestimando valores próximos de 100% (Figura 30). Nota-se na visualização em mapa e perfil que os valores de saturação de água tendem a aumentar bastante próximos ao topo do Quissamã, indicando o fim da zona reservatório. Os valores de baixa saturação ocorrem primariamente associados ao alto estrutural na porção sudoeste do reservatório, mostrando uma grande zona de baixa saturação (Figura 31). Na porção norte do reservatório as regiões de baixa saturação ocorrem de maneira intercalada com regiões de alta saturação.



**Figura 31 - Distribuição de dados de saturação de água (SW). Os dados modelados (azul) mantem a distribuição heterogênea dos dados de entrada (em vermelho), embora subestime o número de células cujo espaço poroso é inteiramente saturado por água.**



**Figura 32 - Modelo de dados de porosidade efetiva (PHIE) para o reservatório. Observa-se a concentração de valores de alta porosidade associados ao alto estrutural na porção sudoeste do reservatório e a um pico nos valores na porção nordeste.**

* + 1. **Correlações entre propriedades de rocha modeladas**

A análise dos valores modelados para as propriedades de rocha (GR, DT, RHOB, PHIE) permite inferências sobre a organização dos dados modelados em diferente fácies. A distribuição das propriedades de rocha não mostra grande variação entre as zonas norte e sul do reservatório (Figura 33), indicando que embora dentro das zonas ocorram prováveis variações faciológicas (como observado nos modelos em mapa e seção), a composição das rochas é similar. Picos locais na distribuição das propriedades podem indicar possíveis agrupamentos de dados e, consequentemente, a ocorrência de diferentes fácies. Observa-se isto em especial para os modelos de GR e RHOB, enquanto PHIE mostra diversos pequenos picos e DT mostra distribuições aproximadamente normais. Os valores de GR sugerem 3 principais agrupamentos, um principal em torno de 25 gAPI, um segundo agrupamento, mais pronunciado na zona norte, em torno de 40 gAPI e um terceiro em torno de 60 gAPI. Já a distribuição dos valores de RHOB indica dois agrupamentos, um em torno de 2,3 g/cm³ e outro em torno de 2,5 g/cm³.



**Figura 33 – Funções de densidade de probabilidade para cada propriedade de rocha modelada para as zonas norte e sul. Observa-se o comportamento similar para as 4 propriedades nas 2 zonas, com DT mostrando comportamento com maiores valores médios e maior heterogeneidade na zona sul.**

Gráficos de dispersão e de densidade de pontos para cada par de propriedades são gerados possibilitando a observação de possíveis agrupamentos de dados e tendências entre as diferentes zonas (Figura 34).

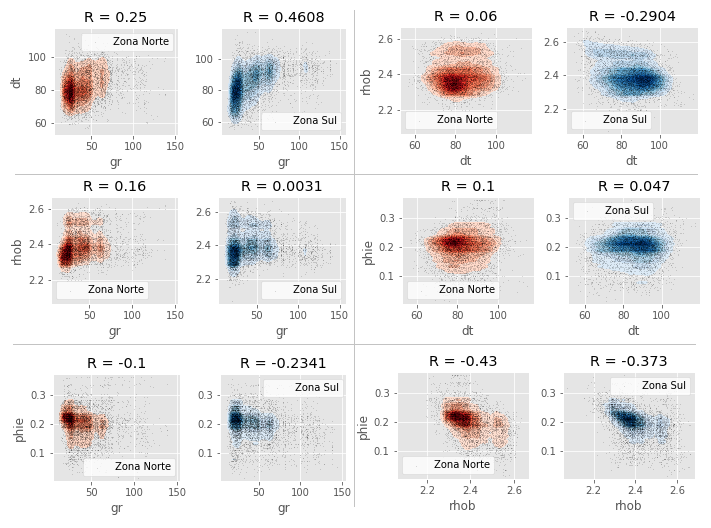
A dispersão dos dados de GR quando comparados aos dados de DT mostram maior correlação entre os valores na zona sul do reservatório, a ocorrência de camadas com valores alternados para DT na zona norte provavelmente explica o menor coeficiente de correlação observado, dado que os mais altos valores para GR associam-se primariamente com regiões próximas ao topo do reservatório. Ao mesmo tempo, é possível observar 3 diferentes agrupamentos de dados no gráfico de densidade de pontos, com grupos com respostas mais altas para ambas as propriedades possivelmente indicando porções com conteúdo argiloso no reservatório.

O mesmo pode ser notado ao se observar a dispersão dos dados de GR quando comparados aos dados de densidade, sendo que para os valores de RHOB não existe correlação significativa com os valores de raio gama para nenhuma das duas zonas. O mesmo agrupamento em 3 diferentes *clusters* de pontos no eixo de GR pode ser observado no gráfico de densidade de pontos, indicando também a ocorrência de porções com conteúdo argiloso para as células com altos valores das 2 propriedades. Este agrupamento em 3 diferentes *clusters* pode ser observado em todos os gráficos em que raio gama é utilizado como uma das variáveis como reflexo da função de densidade de probabilidades para os dados de GR apresentada na Figura 33. De maneira similar, dois agrupamentos no eixo de RHOB podem ser observados em todos os gráficos em que RHOB é utilizado como uma das variáveis.

Nota-se também o comportamento similar entre as duas zonas mesmo em casos com valores de coeficiente de correlação significativamente diferentes. Os gráficos com DT como uma das suas variáveis mostram maior variação, com as amostras convergindo para valores de maior DT na zona sul, refletindo a diferença entre os valores médios para as zonas observada na Figura 33.

Os gráficos de dispersão entre PHIE e GR e entre PHIE e RHOB mostram os mesmos agrupamentos vistos nos demais gráficos em que RHOB ou GR são utilizados como variáveis, com valores de PHIE ocorrendo de forma disseminada entre todos os agrupamentos. Ao mesmo tempo, o gráfico de dispersão entre PHIE e RHOB mostra o comportamento esperado entre as duas variáveis, com relevante correlação negativa entre os valores. No entanto, um número significativo de amostras apresenta valores altos ou baixos para as duas propriedades simultaneamente, indicando a ocorrência de intervalos porosos em distintas litologias ou processos diagenéticos reduzindo porosidades.

De maneira geral, a correlação entre as propriedades modeladas fornece evidências de que existem camadas provavelmente não propicias como reservatório dentro do espaço do reservatório (baixa porosidade, conteúdo significativo de argila e/ou cimentação afetando o espaço poroso). A fim de determinar o número ótimo de agrupamentos necessários para melhor representar estatisticamente as heterogeneidades do reservatório, análises baseadas em aprendizado de máquina não supervisionado são realizadas na secção ‘4.3 – Classificação de Reservatório’.



**Figura 34 – Gráficos de dispersão e densidade de amostras para cada par de propriedades divididos entre as zonas norte e sul do reservatório. R indica o coeficiente de correlação entre as duas propriedades.**

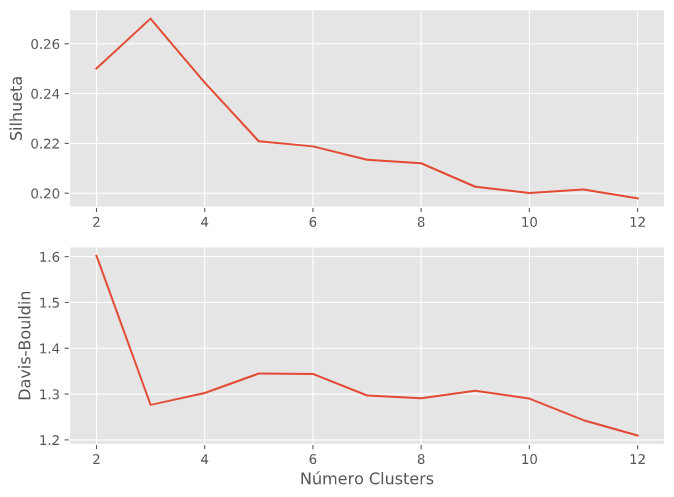
* 1. **Classificação de Reservatório**

Com os dados modelados, busca-se a separação dos dados em diferentes agrupamentos estatisticamente semelhantes para classificar diferentes zonas do reservatório. Para realizar esta separação foram utilizados modelos de aprendizado de máquina não supervisionado baseados no algoritmo de Modelo de Mistura Gaussiana (*Gaussian Mixture Model* – GMM).

A escolha do GMM como o algoritmo para esta função justifica-se pois, embora K-Means seja comumente utilizado em problemas de clusterização aplicados à dados de poço (Schlumberger, 2015 & Romary et al, 2012), sua saída se dá em classificações discretas sem qualquer consideração sobre ambiguidade. O método de GMM, no entanto, fornece resultados similares ao K-Means, porém possibilita a consideração de probabilidades de cada amostra pertencer a cada uma das classes geradas.

Em ambos os métodos, o número de classes para a separação continua sendo um parâmetro a ser determinado pelo usuário. Os resultados observados na secção ‘4.2.6 – Correlações entre propriedades’ indicam diferentes números de agrupamentos nos dados modelados, variando de 2 a 6 agrupamento dependendo de quais variáveis são escolhidas para visualização dos dados. Dado isto, para cada número entre 2 e 10 classes (K), computou-se o algoritmo de GMM com base nas propriedades modeladas de rocha (DT, RHOB, GR e PHIE) e a variância interna e externa das classes foi considerada utilizando os índices de Davies-Bouldin e de Silhueta, estes tendo, respectivamente, seus valores mínimos e máximos conforme se torna mais significativo estatisticamente o agrupamento dos dados.

O máximo local do índice de silhueta para um K de 3 classes acompanhado do mínimo local do índice de Davies-Bouldin para o mesmo valor de K representam forte evidência de que as propriedades modeladas para o reservatório podem ser divididas em três classes diferentes (Figura 35). Simultaneamente, os valores para ambos os índices indicam que conforme aumenta-se o número de classes, rapidamente é reduzida a representatividade da classificação realizada, indicando que os agrupamentos em 6 grupos observados em alguns gráficos não representam grande variância entre estes grupos.



**Figura 35 – Valores para os índices de silhueta e Davies-Bouldin computados para diferentes valores de K no algoritmo GMM. O valor máximo para índice de silhueta e mínimo para Davies-Bouldin no valor de 3 classes indica provável separação dos dados em 3 diferentes classes.**

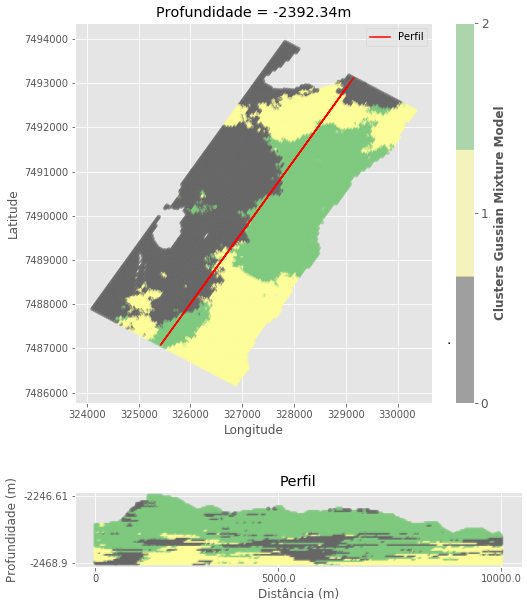
Definido o número de classes para a classificação dos dados, computa-se os algoritmos de GMM com K = 3, separando assim todas as células do modelo em 3 diferentes classes ou *clusters*. A distribuição espacial das classes permite (Figura 36), combinada com as observações feitas sobre a distribuição de cada propriedade, algumas inferências sobre a geologia do reservatório.

Diversas das propriedades modeladas indicam uma transição vertical de fácies no reservatório, isto sendo refletido nas classificações realizadas pelo algoritmo de GMM (Figura 36). As visualizações em mapa das classificações são feitas obtendo uma fatia de profundidade constante na profundidade que define a profundidade mediana das células do reservatório, assim como o perfil é traçado em uma linha que também divide o reservatório em duas regiões com o mesmo número de células.

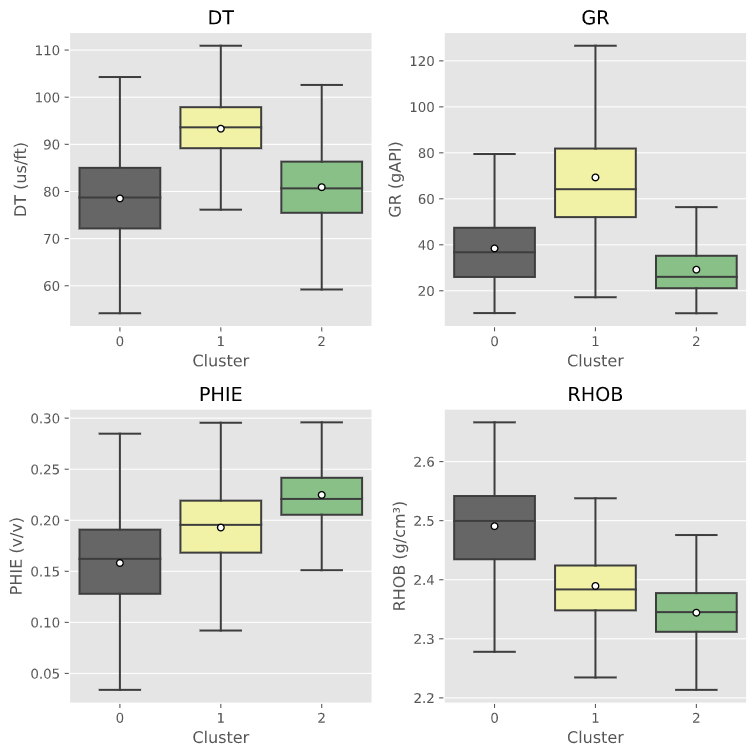
Para dar significado a transição das fácies classificadas, é preciso comparar como as propriedades se distribuem entre cada uma das fácies. Para este fim, *box plots* para cada propriedade são criados distinguindo a distribuição das propriedades em cada classe (Figura 37). Embora valores máximos e mínimos para cada uma das propriedades mostrem-se de forma disseminada entre as 3 classes modeladas, é possível observar clara separação das propriedades, esta separação mostra-se mais evidente para os dados de porosidade efetiva (PHIE) e densidade (RHOB), onde cada classe é centrada em diferentes valores. Já os valores do perfil sônico (DT) e raio gama (GR) não mostram uma separação tão evidente entre as classes 0 e 2, mas descreve bem as células da Classe 1.

As diferenças de PHIE entre as classes fornecem uma importante fonte de informações sobre a geologia do reservatório. Valores superiores a 20% de porosidade efetiva concentram-se nos Clusters 1 e 2, provavelmente correspondendo a intervalos compostos pelos calcarenitos e calcirruditos da formação Quissamã. A resposta elevada de GR e reduzida de RHOB no Cluster 2, no entanto, indica a presença de conteúdo argiloso, provavelmente correspondendo a intervalos com algum nível de cimentação ou uma transição para calcissiltitos e calcilutitos. O Cluster 0 mostra valores de PHIE maiores e de RHOB menores quando comparados aos valores dos demais Clusters, no entanto, seus valores de GR e DT são similares aos observados para o Cluster 2, ao mesmo tempo, segundo Spadini (2008), reservatórios albianos da Bacia de Campos tendem a mostrar predominância de microporosidade em zonas de baixa porosidade, assim, é possível que o Cluster 2 possa corresponder a um intervalo de calcário com predominância de microporosidade, explicando a baixa resposta de GR.

Desta forma, interpreta-se o Cluster 2 como “Calcarenito Poroso”, o Cluster 1 como “Calcarenito Cimentado” e o Cluster 0 como “Calcário com Microporosidade”. Devido a natureza não discreta do GMM, pode-se ainda considerar a probabilidade de cada célula pertencer a cada Cluster. Uma vez que as células do Cluster 0 mostram mais baixa porosidade e foram interpretadas como intervalo com microporosidade, pode-se, de maneira inversa, computar a probabilidade de cada célula do reservatório não pertencer a este Cluster, ou seja, de pertencer ao Cluster 1 ou ao Cluster 2 (Figura 38). Ao interpretarmos os Cluster 1 e 2 como zonas de interesse, a probabilidade de cada célula pertencer a um destes Clusters representa a probabilidade de cada célula fazer parte da zona de interesse no reservatório, possibilitando a utilização dos valores de probabilidade calculados na estimativa de volumes de hidrocarbonetos.

****

**Figura 36 - Mapa e perfil da distribuição das classes modeladas utilizando o algoritmo de Gaussian Mixture Model com K=3. Observa-se a clara relação vertical entre as classes modeladas, embora valores de classe 0 ocorram de maneira menos comum em todas as profundidades do reservatório.**



**Figura 37 - Box plots para cada propriedade para cada cluster modelado pelo GMM.**

Uma imagem contendo texto, mapa

Descrição gerada automaticamente

**Figura 38 – Mapa e perfil com a probabilidade de cada célula pertencer ao Cluster 1 ou ao Cluster 2.**

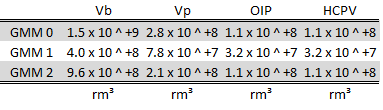
* 1. **Cálculos de Volume**

Utilizando a equação 3.6.1. para o cálculo de volume poroso de hidrocarbonetos, pode-se definir diferentes casos de cálculos de volume com diferentes valores para os parâmetros da equação. Utilizando a mesma equação também são calculados os volumes presentes em cada classe obtidas do classificador GMM. Para cada classe, a equação utiliza N/G de 1 para células pertencentes à classe em questão e N/G de 0 para outras células (Tabela 10). O Caso 0, ou controle, obtém o valor de volume atribuindo a cada célula o valor de N/G do ponto mais próximo nas classificações de net pay para poços do Campo B realizadas por Melani et al. (2015) e seus valores de porosidade e saturação de água correspondentes. Em seguida, são calculados diversos casos apresentados na Tabela 11 e seus resultados são comparados entre si e com o Caso 0 de controle na Tabela 12.

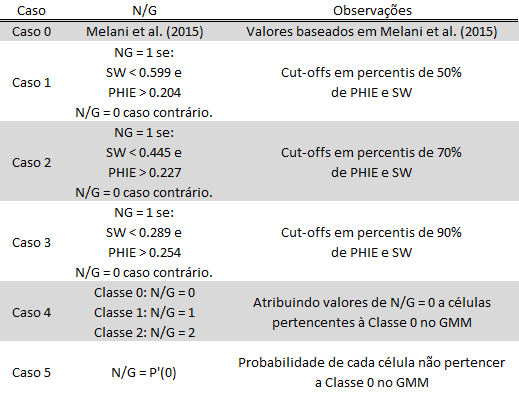
Os Casos 1, 2 e 3 utilizam valores de N/G derivados de valores de corte (cut-offs) baseados em diferentes percentis de PHIE e SW. Para o Caso 1, é atribuído N/G de 1 para todas as células que tenham PHIE superior que o valor de percentil de 50% para todo o reservatório e, ao mesmo tempo, tenham valor de SW menor que o valor de percentil de 50%. De maneira equivalente, os Casos 2 e 3 fazem o mesmo com percentis de 70% e 90%, respectivamente.

Levando em conta as interpretações realizadas sobre as classificações realizadas pelo GMM, o Caso 4 cálcula o valor de HCPV presente nas Classes 1 e 2 do GMM, atribuindo valores booleanos de N/G de acordo com a classe em cada célula se encontra. Uma vez que a Classe 0 foi interpretada como um intervalor com presença de microporosidade, atribui-se N/G de 0 para as células atribuídas a tal classe. O Caso 5, no entanto, utiliza a probabilidade de cada célula não pertencer a Classe 0 como N/G, atribuindo valores contínuos entre 0 e 1 para cada célula.

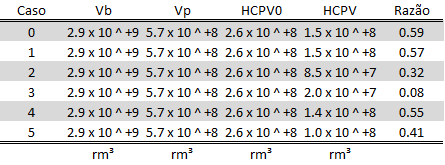
**Tabela 9 – Volumes obtidos para cada classe de acordo com os classificadores de K-Means e GMM. Os valores percentuais indicam a relação entre o volume referido e o obtido para a mesma classe do K-Means. Vb (Bulk Volume) é o volume total das células de cada classe, Vp (Pore Volume) o volume poroso enquanto HCPV (Hydrocarbon Pore Volume) é o volume de hidrocarbonetos.**



**Tabela 10 – Relação dos casos de cálculos de volume. Casos que trazem valores de Prop para saturação de água (SW) ou porosidade efetiva (PHIE) utilizam os valores modelados para cada célula para o cálculo. [1] O Caso 3 utiliza o valor constante de 0.581 para saturação de água, valor correspondente ao valor médio no reservatório. [2] O Caso 2 utiliza o valor constante de 0.198 para porosidade efetiva, valor corresponde ao valor médio no reservatório. [3] Valores de N/G de 0 ou 1 para cada célula baseados na classificação em poços de Melani et al. (2015), cada célula possui valor atribuído igual ao ponto mais próximo nos poços classificados. [4] Valores de corte baseados no percentil de 50% para ambas as propriedades. [5] Valores deorte baseados no percentil de 75% para ambas as propriedades. [6] Valores de corte baseados no percentil de 90% para ambas as propriedades.**



**Tabela 11 – Valores de volumes obtidos para cada caso descrito na Tabela 11. Vb (Bulk Volume) é o volume total das células de cada classe, Vp (Pore Volume) o volume poroso enquanto HCPV (Hydrocarbon Pore Volume) é o volume de hidrocarbonetos e tHCPV é o volume de hidrocarbonetos antes de aplicação das razões N/G. Valores percentuais indicam a variação de cada caso em relação ao Caso 0. Casos 7 e 8 apresentam valores de volume total (Vb) diferentes pois não incluem as células da Classe 0.**



Seguindo o comportamento observado em 4.3., o K-Means estima maiores volumes na Classe 1 (alvo), ao mesmo tempo que estima menores volumes na Classe 0 (descarte). Da mesma forma, o Caso 7, calculado com base no K-Means, mostra maiores valores de Vp e HCPV quando comparado ao Caso 8, baseado no GMM. O Caso 9, baseado nas probabilidades de cada célula não pertencerem a Classe 0 no GMM mostra valor de HCPV similar ao obtido no Caso 8, indicando que a introdução de valores não booleanos para N/G não afetou significativamente os cálculos de volume. Isto se dá devido a maior parte das células terem probabilidades próximas a 0 ou 1 (Figura 39).

Nota-se também que com exceção dos Casos 4, 5 e 6, que utilizam os percentis dos valores das propriedades para definir valores de corte, todos os casos apresentam valores de HCPV consideravelmente maiores quando comparados ao Caso 0 baseado na aproximação das classificações de poço realizadas por Melani et al. (2015).

A comparação entre os valores obtidos nos Casos 1, 2 e 3, todos com N/G de 1, também traz informações a respeito do efeito da variação dos valores de porosidade efetiva e saturação de água nos cálculos de HCPV. Enquanto o Caso 1 utiliza os valores individuais de cada célula tanto para porosidade efetiva quanto para saturação de água, o Caso 2 atribui a cada célula o valor médio de porosidade efetiva no reservatório e o Caso 3 faz o mesmo para o valor médio de saturação de água. Observa-se assim que a utilização do valor médio de porosidade no Caso 2 resulta num aumento pouco significativo do volume poroso (Vp) quando comparado ao Caso 1, indicando uma distribuição pouco variável ao longo do reservatório. O aumento do espaço poroso no Caso 2 também resulta num aumento do valor de HCPV calculado. Para o Caso 3, a utilização de valores médios de saturação de água não afeta o volume poroso. O valor de HCPV calculado aumenta em relação ao Caso 1 de maneira similar ao Caso 2, indicando também uma distribuição pouco variável dos valores de saturação de água. Este comportamento observado nos Casos 2 e 3 corresponde com os baixos valores de CV na zona do reservatório obtidos na etapa de modelamento (Tabela 6).

A comparação dos valores obtidos dos classificadores de K-Means e GMM (Casos 7, 8 e 9) com os valores obtidos nos Casos 1, 2 e 3, mostra que a definição de valores de N/G afeta de maneira significativa os volumes finais obtidos. Atribuir a todas as células valores de 1 para N/G pode levar a grandes superestimações do conteúdo de hidrocarbonetos disponível no reservatório. Ao mesmo tempo, a comparação dos Casos 7, 8 e 9 com os casos obtidos de valores de corte para as propriedades (Casos 4, 5 e 6) indica que a definição de valores de corte pode levar a subestimações de volumes de hidrocarbonetos até para casos considerando um percentil de 50% como o Caso 4.

1. **CONCLUSÕES**

A combinação dos algoritmos de Krigagem com Deriva Externa (KDE) e Gaussian Randon Function Simulation (GRFS) permite a modelagem de dados de poço para um modelo volumétrico tridimensional de maneira satisfatória. Embora o algoritmo de KDE evite transições bruscas entre valores devido à suavização dos resultados, o GRFS tende a honrar melhor a distribuição estatística dos dados de entrada. A limitação do GRFS de necessitar dados de entrada distribuídos de forma aproximadamente normal é uma limitação clara do método. É possível, no entanto, que transformações prévias à execução do algoritmo possam possibilitar seu uso para propriedades distribuições não normais.

Quanto ao upscaling dos dados, o método desenvolvido neste estudo mostrou bons resultados, honrando de maneira geral a distribuição original dos dados de entrada. Exceções são notadas para picos muito altos ou baixos na curva de distribuição dos valores, aonde o método tende a subestimar altos valores e sobrestimar baixos valores.

Quanta à etapa de processamento sísmico, o fluxo de trabalho utilizado passando por atributos de realce de estruturas e realce de descontinuidades permitiu delinear um bom número de descontinuidades e construir um modelo estrutural do reservatório que, em conjunto com as propriedades modeladas, fornece base para os cálculos de volume.

Na etapa de classificação, os algoritmos de K-Means e GMM mostram-se como efetivas técnicas para classificação de células no modelo. O método GMM em especial mostra-se capaz de classificar as células não só de acordo com suas informações litológicas (porosidade, densidade, raio gama), mas também de acordo com as propriedades dos fluídos em seu espaço poroso (saturação de água e resistividade).

A principal vantagem do GMM em relação ao K-Means, no entanto, é a possibilidade de realizar classificações não discretas atribuindo à cada célula um valor de probabilidade. Neste estudo, porém, esta capacidade adicional do GMM não mostrou-se muito relevante, tendo em vista que a maior parte das células foi classificada com probabilidades muito próximas de 1 para sua respectiva classe.

Nota-se, no entanto, que não necessariamente o comportamento observado pelos classificadores será sempre observado. Embora a técnica de Mapas Auto-Organizáveis (Self-Organizing Maps – SOM) tenha se mostrado como bastante efetiva como técnica de redução de dimensionalidade para classificação ao reduzir um modelo de milhões de células para um mapa de 58x58 pixels, é possível que diferentes resultados fossem observados, em especial para as probabilidades do método GMM, caso as classificações fossem realizadas para o modelo completo em seu espaço original.

1. **REFERÊNCIAS BIBLIOGRÁFICAS**

Anderson, P., and R. Newrick, 2008, Strange but true stories of synthetic seismograms: CSEG Recorder, 12, December, 51–56.

ANP. 2013. Boletim da Produção de Petróleo e Gás Natural. Agência Nacional do Petróleo, Gás Natural e Biocombustíveis, Superintendência de Desenvolvimento e Produção (SDP). 26p.

Aragao, O., & Sava, P. 2019. Elastic wavefield tomography with probabilistic petrophysical clustering. In SEG Technical Program Expanded Abstracts 2019 (pp. 1420-1424). Society of Exploration Geophysicists.

Bahorich, M. and Farmer, S. 1995. 3D Seismic Discontinuity for Faults and Stratigraphic

Features: The Coherence Cube. The Leading Edge 14 (10)

Basir, H. M., Javaherian, A., & Yaraki, M. T. 2013. Multi-attribute ant-tracking and neural network for fault detection: a case study of an Iranian oilfield. Journal of Geophysics and Engineering

Benabbou, A., Daly, C., Macé, L., Levannier, A., & Buchholz, C. 2015. An unstructured depositional grid for property modeling. In: Petroleum Geostatistics 2015.

Bentley, M., & Smith, S. 2008. Scenario-based reservoir modelling: the need for more determinism and less anchoring. Geological Society, London, Special Publications, 309(1), 145-159. doi:10.1144/sp309.11

Bjorlykke, K. 1989. Sedimentology and petroleum geology.

Brown, A.R. 1999. Interpretation of Three-Dimensional Seismic Data, 9, 528, fifth edition. Tulsa, Oklahoma: Investigations in Geophysics, Soc. of Exploration Geophysicists.

Bruhn C.H.L. 1998. Petroleum Geology of Rift and Passive Margin Turbidite Systems: Brazilian and Worldwide Examples. In: AAPG International Conference and Exhibition, Rio de Janeiro, Brazil. Part 2, Course 6.

Bruhn, C.; Gomes, J.; Lucchese Jr., C.; Johann, P. 2003. Campos Basin: Reservoir Characterization and Management – Historical Overview and Future Challenges. Offshore Technology Conference, Houston, Tx. OTC-15220.

Bruno, P. P. G., Cippitelli, G., & Rapolla, A. 1998. Seismic study of the Mesozoic carbonate basement around Mt. Somma–Vesuvius, Italy. Journal of volcanology and Geothermal Research

Caers, J., Park, K., & Scheidt, C. 2010. Modeling Uncertainty of Complex Earth Systems in Metric Space. Handbook Of Geomathematics

Caers, J. 2011. Modeling uncertainty in the earth sciences.

Chopra, S. & Marfurt, K. J. 2007. Seismic attributes for prospect identification and reservoir characterization.

Chahine, C., Truelove, L., & Volcan, M. H. 2014. Enhanced delineation of reservoir compartmentalization from advanced preand post-stack seismic attribute analysis.

Chen, L., Xiao, C., Li, X., Wang, Z., & Huo, S. 2018. A seismic fault recognition method based on ant colony optimization. Journal of Applied Geophysics, 152, 1-8.

Cobb, W. M., & Marek, F. J. 1998. Net pay determination for primary and waterflood depletion mechanisms. In SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.

Costa, L. A. N., Maschio, C., & Schiozer, D. J. 2018. A new methodology to reduce uncertainty of global attributes in naturally fractured reservoirs. Oil & Gas Science and Technology–Revue d’IFP Energies nouvelles, 73, 41.

Cox, T., & Seitz, K. 2007. Ant tracking seismic volumes for automated fault interpretation. In CSPG CSEG Convention, Calgary, Alberta, Canada.

CSM E&P Solutions. 2012. Automated Data Conditioning – Ant Tracking Workflow.

Daly, C., & Caers, J. 2010. Multi-point geostatistics–an introductory overview. First Break, 28(9), 39-47.

Deutsch, C. V. & Journel, A. G. 1992. GSLIB Geostatistical Software Library and User's Guide Oxford University Press, New York

Deutsch, C. V. 1999. Reservoir modeling with publicly available software. Computers & Geosciences.

Dias, J.L.; Scarton, J.C.; Esteves, F.R.; Carminatti, M.; Guardado, L.R. 1990. Aspectos da Evolução Tectono-Sedimentar e a Ocorrência de Hidrocarbonetos na Bacia de Campos. In: Origem e Evolução de Bacias Sedimentares, Raja Gabaglia, G.P. and Milani, E.J., eds., Petrobras, Rio de Janeiro, Brasil.

Dongas, J. M. 2016. Development and characterization of a geostatic model for monitoring shallow CO2 injection.

Egbele, E., Ezuka, I., & Onyekonwu, M. 2005. Net-to-gross ratios: Implications in integrated reservoir management studies. In Nigeria Annual International Conference and Exhibition. Society of Petroleum Engineers.

Elias V.L.G.; Steagall, D.E. 1996. The Impact of the Values of Cementation Factor and Saturation Exponent in the Calculation of Water Saturation for Macae Formation. SCA Conference Paper Number 9611

Ellis, Darwin V.; Singer, Julian M. 2008. Well Logging for Earth Scientists (Second ed.). Springer. p. 692. ISBN 1-4020-3738-4.

ESRI. 2019. How Kriging Works. Disponível em: http://desktop.arcgis.com/en/

arcmap/10.3/tools/3d-analyst-toolbox/how-kriging-works.htm. Acesso em: 17/03/2019.

De Lima, A., Fourno, A., Noetinger, B., & Schiozer, D. J. 2019. Characterization and Modeling of the Fault Network of a Brazilian Pre-Salt Reservoir and Upscaling Results. In SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers.

Fang, J., Zhou, F., & Tang, Z. 2017. Discrete fracture network modelling in a naturally fractured carbonate reservoir in the Jingbei oilfield, China. Energies, 10(2), 183.

Fletcher, M., Dornhaus, A., & Shin, M. C. 2011. Multiple ant tracking with global foreground maximization and variable target proposal distribution. In Applications of Computer Vision (WACV), 2011 IEEE Workshop on (pp. 570-576). IEEE.

Franz, E.P. 1987. Análise da Geologia de Reservatório dos Carbonatos da Formação Macaé do Campo de Enchova – Bacia de Campos, Brasil. Dissertação de Mestrado, Departamento de Geologia da Escola de Minas, Universidade Federal de Ouro Preto, Ouro Preto. 120p.

Figueiredo, A.M.F.; Mohriak, W.U. 1984. A Tectônica Salífera e as Acumulações de Petróleo da Bacia de Campos. In: SBG, Congresso Brasileiro de Geologia, 33, Anais, Rio de Janeiro, p.1380- 1394.

Guardado, L.R., Gamboa, L.A.P., Lucchesi, C.F. 1989. Petroleum Geology of the Campos Basin: A Model for Producing Atlantic Type Basin. In: Edwards, J.D. and Santogrossi, P.A., eds., Divergent/Passive Margin Basins. AAPG Memoir, 48, p.3-79.

Godfrey, R., & Bachrach, R. 2008. Seismically Guided Fracture Characterization.

Hale, D. 2009. Structure-oriented smoothing and semblance. CWP report, 635, 261-270.

Hall, M., & Hall, M. 2013. Well-tie workflow. Agile. Disponível em: https://agilescientific.com/blog/2013/4/25/well-tie-workflow.html. Acesso em: 20/10/2018.

Hampson, D. & Russell, B. 2004. AVO theory. Hampson-Russell Software Services Ltd, 22-26.

Hansen, T. M., Journel, A. G., Tarantola, A., & Mosegaard, K. 2006. Linear inverse Gaussian theory and geostatistics. Geophysics, 71(6).

Hengl, T., Heuvelink, G. B., & Rossiter, D. G. (2007). About regression-kriging: From equations to case studies. Computers & geosciences, 33(10), 1301-1315.

Hengl, T. 2009. A practical guide to geostatistical mapping.

Hengl, T. Herrera, R. H., Fomel, S., & van der Baan, M. 2014. Automatic approaches for seismic to well tying.

Hunter, J.D., 2007. Matplotlib: A 2D graphics environment. Computing in science & engineering.

Krief, M., Garati, J., Stellingwerff, J. & Ventre, J. 1990. A petrophysical interpretation using the velocities of P- and S- waves (Full-wave sonic). The log analyst.

Kuroda, M. C., Vidal, A. C., Leite, E. P., & Drummond, R. D. 2012. Electrofacies characterization using self-organizing maps. Revista Brasileira de Geofísica.

Kuznetsova, A. Demyanov, V., & Christie, M. "Hierarchical Geological Realism for Reliable Reservoir Prediction." 78th EAGE Conference and Exhibition 2016.

Leite E.P., Filho C.R.S. 2010. Mapas Auto-Organizáveis aplicados ao mapeamento do potencial mineral na região de Serra Leste, Província Mineral de Carajás, Pará. Rev. Brasileira de Geofísica. 28(3):397-409.

Lloyd, S. 1982. Least squares quantization in PCM. IEEE transactions on information theory, 28(2), 129-137.

Lyons, W. C., & Plisga, G. J. (2016). Standard handbook of petroleum and natural gas engineering, Third Edition. Elsevier.

McKinney, W. 2010. Data structures for statistical computing in python. In Proceedings of the 9th Python in Science Conference.

Melani, L., Vidal, A. C., Filho, I. M., & Schuab, F. B. 2015. The impact of Archie's parameters in the calculation of water saturation for carbonate reservoir, Campos Basin, Brazil. In SEG Technical Program Expanded Abstracts 2015 (pp. 2984-2989). Society of Exploration Geophysicists.

Nelson, R.A. 2001. Geologic Analysis of Naturally Fractured Reservoirs. 2 nd Ed. Woburn, Massachusetts, USA. Gulf Professional Publishing. 332p.

Newrick, R. 2012. Well tie basics. Well tie perfection, 52, 104-107

Pedersen, S. I., Skov, T., Randen, T., & Sønneland, L. 2005. Automatic fault extraction using artificial ants. In Mathematical Methods and Modelling in Hydrocarbon Exploration and Production (pp. 107-116). Springer Berlin Heidelberg

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J., 2011. Scikit-learn: Machine learning in Python. Journal of machine learning research

Ponte J. P. P. 2019. Aplicação de métodos computacionais na análise de imagens referentes a dados geológicos. Dissertação de Mestrado. Universidade Estadual de Campinas.

Pyrcz, M. J., & Deutsch, C. V. 2014. Geostatistical reservoir modeling. Oxford university press.

Oliphant, T.E., 2006. A guide to NumPy (Vol. 1, p. 85). Trelgol Publishing.

Oliveira, M.L. 1997. Análise de Incertezas Envolvidas na Modelagem de Reservatório no Contexto Geoestatístico. Dissertação de Mestrado. Universidade Estadual de Campinas.

Ralha, J.F.S. 2016. Mapeamento de Falhas e Fraturas por Reflexão Sísmica em um Campo Produtivo da Bacia de Campos. Trabalho de Conclusão de Curso. Universidade Estadual de Campinas.

Randen, T., Monsen, E., Signer, C., Abrahamsen, A., Hansen, J. O., Sæter, T., & Schlaf, J. (2000). Three-dimensional texture attributes for seismic data analysis. In SEG Technical Program Expanded Abstracts 2000. Society of Exploration Geophysicists.

Randen, T., Pedersen, S. I., & Sonneland, L. 2001. Automatic extraction of fault surfaces from three-dimensional seismic data. In Annual International Meeting, Society of Exploration Geophycists

Raykov, Y.P., Boukouvalas, A., Baig, F. and Little, M.A., 2016. What to do when k-means clustering fails: A simple yet principled alternative algorithm.

Rijks, E. J. H., & Jauffred, J. C. E. M. 1991. Attribute extraction: An important application in any detailed 3-D interpretation study. The Leading Edge, 10(9), 11-19.

Ringrose P. & Bentley M. 2014a. Introduction to Modeling. In: Reservoir Model Design. Springer, Dordrecht

Ringrose, P., & Bentley, M. 2014b. Model Purpose. In: Reservoir Model Design, Springer, Dordrecht

Robaina, L.E.S; Formoso, M.L.L.; Spadini, A.R. 1991. Carbonatos Reservatório da Formação Macaé, Bacia de Campos, RJ, Brasil. Parte I – Contribuição à Geoquímica dos Carbonatos. Geochim. Brasil.

Romary, T., Rivoirard, J., Deraisme, J., Quinones, C., & Freulon, X. 2012. Domaining by clustering multivariate geostatistical data. In Geostatistics Oslo 2012. Springer, Dordrecht.

Schlumberger. 2015. Petrel Seismic-to-Evaluation Software Documentation.

Silva, C. C., Marcolino, C. S., & Lima, F. D. 2005. Automatic fault extraction using ant tracking algorithm in the Marlim South Field, Campos Basin. In 9th International Congress of the Brazilian Geophysical Society.

Spadini, A.R. 1992. Processos Deposicionais e Ciclicidade em Carbonatos Albianos de Plataforma Rasa da Bacia de Campos. Dissertação de Mestrado, Instituto de Geociências UFRJ

Spadini, A.R. 2008. Carbonate Reservoirs in Brazilian Sedimentary Basins. 19th World Petroleum Congress, Spain.

Stewart, R. R., Huddleston, P. D., & Kan, T. K. 1984. Seismic versus sonic velocities: A vertical seismic profiling study. Geophysics, 49(8), 1153-1168.

Taner, M.T., Koehler, F., and Sheriff, R.E. 1979. Complex Seismic Trace Analysis. Geophysics 44 (6): 1041.

Taner, M.T., Walls, J.D., Smith, M., Taylor, G., Carr, M.B., & Dumas, D. 2001, "Reservoir characterization by calibration of self‐organized map clusters," SEG Technical Program Expanded Abstracts : 1552-1555.

Trabanou, J.R.; Chabernaud, T.; Goswami, J.; Li, J.; Horkowitz, J.; Jacobsen, S.; Neville, T.; Seydoux, J. 2004. How Certain Are We About Reservoir Parameter Uncertainties? 45 th Annual Logging Symposium SPWLA.

Tomaso, L.; Benac, P.; Sarturi, J.; González, G.; de Ribet, B. 2013. Seismic Fracture Characterization Workflow and Support for the Geological Model: Albian Carbonate Reservoir, Campos Basin, Brazil. 13th International Congress of the Brazillian Geophysical Society, Rio de Janeiro, Brazil.

Van Bemmel, P. P., & Pepper, R. E. 2000. U.S. Patent No. 6,151,555. Washington, DC: U.S. Patent and Trademark Office.

VanderPlas, J., 2016. Python data science handbook: essential tools for working with data. O'Reilly Media, Inc.

Van Der Walt, S., Colbert, S. C., & Varoquaux, G. 2011. The NumPy array: a structure for efficient numerical computation. Computing in Science & Engineering

Viroli, C., & McLachlan, G. J. (2019). Deep gaussian mixture models. Statistics and Computing, 29(1), 43-51.

Wackernagel, H. 1995. Multivariate geostatistics: an introduction with applications.

White, R. E., & Simm, R. 2003. Tutorial: Good practice in well ties. First Break, 21(10).

Winter, W.R.; Jahnert, R.J.; França, A.B. 2007. Bacia de Campos. Boletim de Geociências da Petrobras, Rio de Janeiro

Warren, J. E., & Root, P. J. (1963). The behavior of naturally fractured reservoirs. Society of Petroleum Engineers Journal, 3(03), 245-255.

Zanão, R. 2008. Técnicas Geostatísticas Aplicadas na Modelagem de Saturação de Óleo. Dissertação de Mestrado. Universidade Estadual de Campinas.

Zeng, H., & Kerans, C. 2003. Seismic frequency control on carbonate seismic stratigraphy: A case study of the Kingdom Abo sequence, west Texas. AAPG bulletin, 87(2), 273-293.

Zhao, T., Jayaram, V., Roy, A., & Marfurt, K. J. 2015. A comparison of classification techniques for seismic facies recognition.