



Universidade Federal  
do Rio de Janeiro

---

Escola Politécnica

## QUANTIFICAÇÃO E VISUALIZAÇÃO DE INCERTEZAS NA MODELAGEM DE RESERVATÓRIOS DE PETRÓLEO

Thiago Guimarães Saraiva

Projeto de Graduação apresentado ao curso de Engenharia de  
Petróleo da Escola Politécnica da Universidade Federal do Rio de  
Janeiro apresentado como requisito necessário à aprovação da  
respectiva disciplina.

Orientadores: Virgílio José Martins Ferreira Filho

2015.

Saraiva, Thiago Guimarães.

Quantificação e visualização de incertezas na modelagem de reservatórios de petróleo/ Thiago Guimarães Saraiva. – Rio de Janeiro: UFRJ/ Escola Politécnica, 2015.

X, 80 p.: il.; 29,7 cm.

Orientador: Virgílio José Martins Ferreira Filho

Projeto de Graduação – UFRJ/ Escola Politécnica/

Curso de Engenharia do Petróleo, 2015.

Referências Bibliográficas: p. 80.

2.Definições e modelagem de incertezas 2.Quantificação de incertezas. 3.Visualização de incertezas. 4.Análise de incertezas. 5.Aplicação.

I. José Martins Ferreira Filho, Virgílio. II. Universidade Federal do Rio de Janeiro, Escola Politécnica, Curso de Engenharia de Petróleo. III. Título.

QUANTIFICAÇÃO E VISUALIZAÇÃO DE INCERTEZAS NA MODELAGEM DE  
RESERVATÓRIOS DE PETRÓLEO

Thiago Guimarães Saraiva

PROJETO DE GRADUAÇÃO SUBMETIDO AO CORPO DOCENTE DO CURSO  
DE ENGENHARIA DE PETRÓLEO DA ESCOLA POLITÉCNICA DA  
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS  
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE  
ENGENHEIRO DE PETRÓLEO

Examinada por:

---

Prof. Virgílio José Martins Ferreira Filho.

---

Prof. Paulo Camargo Silva

---

Dr. Regis Kruel Romeu

RIO DE JANEIRO, RJ – BRASIL

MÊS de 2015

# DEDICATÓRIA

---

Dedico esse projeto de graduação a Deus que me deu o privilégio de pôr as pessoas certas no meu caminho para estar realizando essa conquista.

Dedico também aos meus pais e familiares que muito me apoiaram em diversos sentidos.

# AGRADECIMENTOS

---

Agradeço primeiramente a Deus que abriu todas as minhas portas.

Ao meu Pai que se sacrificou muito ao longo de todos esses anos para me dar uma educação de qualidade que me permitiu ingressar nesta universidade de nome e honra. Por ter me apoiado e me fornecido todos os recursos necessários nesses cinco anos e meio de muita luta, noites mal dormidas e engarrafamentos.

À minha Mãe que se privou de sua carreira, de sua educação para garantir a minha. Isso é um privilégio, muitos não têm uma mãe que não foi um terço do que a minha foi, que não fez um terço do que a minha fez por mim, que não renunciou um terço do que a minha renunciou por mim. Se ela tivesse sido uma mãe ausente, talvez eu não estaria me formando, já teria desistido há muito tempo.

Ao Professor Virgílio. Graças a ele tive a oportunidade de aprender sobre o tema que escrevo nesse projeto.

À ANP pelos recursos e incentivo à pesquisa.

Ao CENPES (Centro de Pesquisas da Petrobras) pela excelente oportunidade de estágio e aprender com profissionais de excelência.

# RESUMO

---

Resumo do Projeto de Graduação apresentado à Escola Politécnica/ UFRJ como parte dos requisitos necessários para a obtenção do grau de Engenheiro de Petróleo.

## QUANTIFICAÇÃO E VISUALIZAÇÃO DE INCERTEZAS NA MODELAGEM DE RESERVATÓRIOS DE PETRÓLEO

Thiago Guimarães Saraiva

March/2016

Orientador: Virgílio José Martins Ferreira Filho

Curso: Engenharia de Petróleo

Em todas as áreas, seja nas ciências ou no cotidiano, tomada de decisões são necessárias, contudo nunca se tem uma informação perfeita e completa sobre determinado fenômeno físico que está ocorrendo em um determinado sistema. Às vezes o excesso de incerteza, falta de conhecimento sobre algo, pode levar a uma decisão ruim, causando assim prejuízos de todos os tipos possíveis. Daí a necessidade de se considerar a modelagem de incerteza em uma tomada de decisão.

Esse trabalho visa primeiramente transcorrer sobre as principais definições de incerteza por vários autores. E depois de discutir algumas das diferentes concepções, são apresentadas ferramentas para quantificação, visualização e análise de incertezas na modelagem geológica de reservatórios de petróleo.

Ao final é mostrado um pequeno exemplo do passo-a-passo da construção de um modelo fino de reservatório no Petrel e depois, uso do fluxo de trabalho *de* incertezas para gerar  $N$  diferentes realizações para na próxima etapa (não realizada) aplicar algum dos métodos do capítulo 4 para escolher  $n$  (um número muito menor que  $N$ ) modelos representativos.

# ABSTRACT

---

Abstract of Undergraduate Project presented to POLI/UFRJ as a partial fulfillment of the requirements for the degree of Engineer.

## UNCERTAINTY QUANTIFICATION AND VISUALIZATION IN PETROLEUM RESERVOIRS MODELING

Thiago Guimarães Saraiva

March/2016

Advisor: Virgílio José Martins Ferreira Filho

Course: Petroleum Engineering

In all areas, whether in the sciences or in day-to-day decision-making are required but never have a perfect and complete information about a particular physical phenomenon that is occurring in a given system. Sometimes excess of uncertainty, lack of knowledge about something, can lead to a bad decision, thereby causing losses of all possible types. Hence the need, if not obligation, to consider the modeling uncertainty in decision making.

This work aims primarily spend on major uncertainty definitions by various authors. After discussing some different conceptions of uncertainties, are some tools to quantify it and then use visualization techniques.

The end is shown a small example of step-by-step construction of a thin reservoir model in Petrel and then use the uncertainties workflow to generate  $N$  different achievements to the next step (unrealized) apply some of the methods chapter 4 to select  $n$  (far fewer than  $N$ ) representative models.

# ÍNDICE

---

Dedicatória .....	i
Agradecimentos .....	ii
Resumo .....	iii
Abstract .....	iv
Índice .....	v
Lista de figuras .....	viii
1. Introdução .....	2
Objetivos .....	4
2. Definições e modelagem de incertezas .....	5
2.1. O que é Incerteza? Algumas definições .....	5
2.2. Desafios e questões na modelagem de incertezas .....	6
2.3. Fontes de Incerteza .....	7
3. Quantificação de incertezas .....	11
3.1. Teoria da Probabilidade .....	11
3.1.1. Espaço amostral, Evento e Resultado .....	12
3.1.2. Probabilidade condicional .....	13
3.1.3. Lei de Bayes .....	13
3.2. Entropia da Informação .....	14
3.2.1. Entropia da Informação aplicada a um conceito Espacial .....	16
3.2.2. Entropia Total de um Modelo .....	18
3.2.3. Aplicação em Reservatórios .....	18
3.3. Análise de Sensibilidade .....	24
3.4. Simulação de Monte Carlo .....	26
4. Visualização de incertezas .....	29

4.1.	Visualização de Sumarização de Dados. ....	30
4.1.1.	Box Plot .....	30
4.2.	Outras Técnicas de Visualização .....	32
4.2.1.	Gráficos de Coordenadas paralelas.....	32
4.2.2.	Gráficos <i>Spagetti</i> e <i>Fan Charts</i> .....	33
4.2.3.	Mapas de Cores.....	36
4.2.4.	Justaposição e Superposição.....	38
4.2.5.	Cone HSV e mapas bivariados .....	41
4.2.6.	Mapas de deslocamento .....	46
4.2.7.	Renderizando volumes e Mapas vetoriais. (3D).....	46
5.	Análise de incertezas.....	50
5.1.	Definição de Cenários Geológicos e Realizações de um Cenário. ....	50
5.2.	Algumas Metodologias para escolha de modelos representativos.....	53
5.2.1.	Método do VME.....	54
5.2.2.	Escolha de Modelos em Espaços Paramétricos. ....	57
5.2.3.	Escolha de Modelos em Espaços Métricos.....	58
6.	Aplicação.....	63
6.1.	Descrição do Modelo Geológico Usado .....	63
6.2.	Fluxo de trabalho de incertezas do Petrel .....	67
6.2.1.	Funcionamento do Fluxo de trabalho e resultados .....	68
7.	Conclusões .....	77
8.	REferÊncias.....	80
	Apêndice.....	I
	A. Criação do Modelo Fino.....	I
	B. Criação do Modelo de Falha .....	II
	C. Definição do Pillar Malhading. ....	III

D. Definição dos Horizontes .....	VI
E. Definição das zonas do reservatório.....	IX
F. Definição do número de camadas (Layering).....	IX
G. Visualização de propriedades geométricas no Modelo Geológico .....	X
H. Utilizando as informações de <i>logs</i> dos poços. ....	XII
I. Transferindo dados sísmicos.....	XIV
J. Modelo Petrofísico de Porosidade.....	XVII
K. Modelo Petrofísico de NTG.....	XXII
L. Modelo Petrofísico de Permeabilidade.....	XXIV

# LISTA DE FIGURAS

---

Figura 2.1 - Uma tubulação esquemática que mostra as etapas em que incertezas são incorporadas no processo de visualização. Adaptado de (PANG, WITERNBRINK, & LODHA, 1996).....	9
Figura 3.1- Entropia de Shannon para um experimento de dois eventos possíveis em função do primeiro resultado.....	16
Figura 3.2 – Aplicação do conceito espacial da Entropia da Informação. Adaptado de (WELLMANN & REGENAUER-LIEB, 2012).....	17
Figura 3.3 – (a) Modelo geológico com as superfícies modeladas das unidades geológicas. (b) Visão do topo do modelo, pode-se ver as linhas de interseção do topo com as superfícies das unidades geológicas 1 e 2. (WELLMANN & REGENAUER-LIEB, 2012)....	21
Figura 3.4 – Mapas de Probabilidade para as unidades geológicas F1, F2 e F3. (WELLMANN & REGENAUER-LIEB, 2012).....	21
Figura 3.5 – Mapas de Entropia da Informação. Na figura foram gerados 3 mapas, um com o intervalo total de valores de H, e os outros com $H>0$ e com $H>1$ . (WELLMANN & REGENAUER-LIEB, 2012) .....	22
Figura 3.6 - Avaliação da unidade de imprecisão em um modelo com diferentes números de células. Adaptado de (WELLMANN & REGENAUER-LIEB, 2012).....	23
Figura 3.7 - Medida <i>fuzzy</i> de uma unidade geológica para conjuntos aleatoriamente escolhidos de 50 realizações simuladas. Adaptado de (WELLMANN & REGENAUER-LIEB, 2012).....	24
Figura 3.8 – Gráfico de Análise da Sensibilidade gerado pelo CMOST, simulador da CMG mostrando as variações do volume acumulado de água produzida durante as rodadas de experimentos para o ajuste de histórico.....	26
Figura 3.9 - Um histograma com 10.000 simulações sobre uma proxy, e o ajustamento de três Funções de Densidade de Probabilidade. (CAMARGO SILVA E FILHO, 2012).....	28

Figura 4.1 - Boxplots e suas modificações. A esquerda, a construção do boxplot. A direita, modificações visuais. (POTTER, K., 2010).....	31
Figura 4.2 - Modificações de densidade no boxplot. a) Histplot. b) Vaseplot. c) Box-percentile plot. d) Violin Plot (POTTER, K., 2010) .....	32
Figura 4.3 - Gráfico de Coordenadas Paralelas. Fonte: Stackoverflow (Hiperlink).....	33
Figura 4.4 – Gráfico Spagetti. Uma aplicação na área da meteorologia. ....	34
Figura 4.5 - Gráficos tipo espaguete (spaghetti plots), sobrepondo resultados de diversos membros do conjunto de simulações em uma mesma imagem, para uma aplicação na área de meteorologia. ....	35
Figura 4.6 – Exemplo de visualização da incerteza usando diagrama tipo <i>fan-chart</i> . Fonte: <a href="https://www.cnb.cz/en/monetary_policy/forecast/previous_forecasts/prognoza_1305.html">https://www.cnb.cz/en/monetary_policy/forecast/previous_forecasts/prognoza_1305.html</a> .....	36
Figura 4.7 – Corte latitudinal de um tronco humano. (POTTER, K., 2010). ....	37
Figura 4.8 - Média para variações de 50% (para mais ou para menos) da condutividade dos pulmões (mapa da esquerda). Desvio padrão para 50%(para mais ou para menos) da condutividade dos pulmões. (POTTER, K., 2010).....	37
Figura 4.9 - Diferentes escalas de cores para mapas de cores. (POTTER, K., 2010) ....	38
Figura 4.10 - Classificação esquemática de duas abordagens para visualização de conjuntos: (a) visualização condensada; e (b) visualização simultânea. ....	39
Figura 4.11 - Diversas formas de visualização de um mesmo conjunto de 20 séries temporais (LAMIGUEIRO, 2014). ....	40
Figura 4.12 – Aplicação em reservatórios. Realizações de um mesmo modelo geológico justapostas para visualização na sala UHD (Ultra High Definition) no CENPES da Petrobras (Um caso 5D). ....	41
Figura 4.13 - RGB, CMY e Azul. (POTTER, K., 2010). ....	42
Figura 4.14 - Cone HSV (Matiz, Saturação, Valor). (POTTER, K., 2010).....	43
Figura 4.15 - As 3 dimensões do HSV. (POTTER, K., 2010).....	43
Figura 4.16 – Média usando escala de cores. (POTTER, K., 2010).....	44

Figura 4.17 – Desvio padrão usando escala cinza ou valor (quantidade de cinza). (POTTER, K., 2010) .....	44
Figura 4.18 – Mapa bivariado da média com o desvio padrão. ....	45
Figura 4.19 – Uma aplicação, em modelos geológicos. Mapa bivariado na saturação ao invés do valor. ....	45
Figura 4.20 – Nessa figura o desvio padrão é representado pelo mapa de cores e a média pela altura do deslocamento da superfície. (POTTER, K., 2010). ....	46
Figura 4.21 – Renderização volumétrica da salinidade do oceano e sua incerteza. Cores representam valores de salinidade e regiões de maior opacidade representam maior incerteza. (DJURCILOV, KIM, LERMUSIAUX, & PANG, 2002). ....	47
Figura 4.22 - Função de Transferência 2D e valores de opacidade. (DJURCILOV, KIM, LERMUSIAUX, & PANG, 2002).....	48
Figura 4.23 – Uso de malha branca para realçar o grau de transparência do volume. (DJURCILOV, KIM, LERMUSIAUX, & PANG, 2002) .....	48
Figura 4.24 – Visualização do fluxo de correntes. As médias e desvios-padrão de vento e as correntes oceânicas são mostrados usando glifos. (WITTENBRINK, C. M.; PANG, A. T.; LODHA, S. K., 1996).....	49
Figura 5.1 – Cenários pessimista (a), provável (b) e otimista (c). (LIGERO, RISSO, & SCHIOZER, 2008). ....	51
Figura 5.2 – Análise de sensibilidade para o VPL (a) e para o Np (b). (LIGERO, RISSO, & SCHIOZER, 2008). ....	52
Figura 5.3 - Fluxo de trabalho genérico de múltiplas rodadas de simulação. (ROMEU, et al., 2013).....	54
Figura 5.4 - Espaços paramétricos de modelos e de respostas (para uma estratégia de produção u1), com diferentes números de parâmetros. (ROMEU, et al., 2013). ....	55
Figura 5.5 – Combinação das diversas estratégias de produção (em rosa) e modelos alternativos do mesmo reservatório (em azul), gerando uma matriz de resultados (em verde). Para cada estratégia de produção (cada linha da matriz), calcula-se VME e curva de risco. (ROMEU, et al., 2013) .....	56

Figura 5.6 - Resultados parciais de uma aplicação da metodologia de seleção de modelos representativos em espaços paramétricos. (MARQUES, 2012). .....	58
Figura 5.7 - Representação dos modelos no espaço métrico, por uma matriz de distâncias e por uma redução 2D dessa matriz pela técnica MDS. De (ROMEU, et al., 2013) e adaptado de (CAERS, 2011). .....	60
Figura 5.8 - Semelhante à Figura 4.7, mas usando uma métrica diferente para cálculo das distâncias, baseada em resultados de uma simulação por linhas de fluxo. (ROMEU, 2013 – adaptado de CAERS, 2011). .....	61
Figura 5.9 - Transformação da matriz da distância para um remapeamento vantajoso do espaço dos modelos (ROMEU, et al., 2013) e adaptado de (CAERS, 2011). .....	62
Figura 6.1 - Poços do reservatório em questão, presentes no projeto Petrel e as superfícies de topo e Base. ....	65
Figura 6.2- Visualização da porosidade estimada .....	66
Figura 6.3- Visualização do NTG estimado. ....	66
Figura 6.4 - Visualização da permeabilidade estimada. ....	67
Figura 6.5 – Fontes de informação e Níveis de Complexidade no estudo das incertezas . .....	68
Figura 6.6 - Configuração usada no <i>Fluxo de trabalho</i> de teste.....	69
Figura 6.7 - Outra opção de fluxo de trabalho.....	70
Figura 6.8 – Lista de Variáveis/Parâmetros incertos considerados no fluxo de trabalho, suas distribuições e seus <i>ranges</i> de ocorrência.....	71
Figura 6.9 – As variáveis de incertezas serão definidas nos campos “ <i>Range</i> ” e “ <i>Azimuth</i> ” do variograma. ....	72
Figura 6.10 – Escolha do Método de Monte Carlo de amostragem com habilitação de hipercubo latino. ....	72
Figura 6.11- Visualização justaposta dos mapas de porosidade de 4 realizações. ....	73
Figura 6.12 – Capturas de tela dos mapas de porosidade de todas as 50 realizações do modelo geológico base usado no fluxo de trabalho.....	74
Figura 6.13 – Funcionalidades do <i>Uncertatinty Viewer</i> plugin. ....	75

Figura 6.14 – Mapa de desvio padrão gerado a partir das realizações. ....	76
Figura 6.15 – Barra de cores e curva (amarela) de opacidade. ....	76
Figura A.1 - Criação do modelo fino.....	I
Figura B.1 - Criação do modelo de falha.....	II
Figura B.2 - Definição do <i>Fault modeling</i> .....	III
Figure C.1 - Definição do Pillar Malhading. ....	III
Figure C.2 - Definição do modelo geológico e do tamanho da malha do Pillar Malhading.....	IV
Figure C.3 - Definição dos limites do modelo.....	V
Figure C.4 - Definição dos limites do modelo – confirmação.....	V
Figure C.5 - Visualização da malha de referência.....	VI
Figure D.1 - Definição dos Horizontes.....	VII
Figure D.2 - Inclusão das superfícies de topo e base.....	VII
Figure D.3 - Associação dos marcadores dos poços.....	VIII
Figure D.4 - Final do processo de criação dos horizontes no modelo geológico. ....	VIII
Figure E.1 - Definição das zonas do reservatório.....	IX
Figure F.1 - Definição do número de camadas.....	X
Figure G.1 - Opção Geometrical Modeling.....	XI
Figure G.2 - Criação de propriedade em <i>Geometrical Modeling</i> . ....	XII
Figure G.3 - Visualização dos índices das camadas. ....	XII
Figure H.1 - Opção Scale up well logs.....	XIII
Figure H.2 - Seleção das informações de logs dos poços.....	XIII
Figure H.3 - Visualização dos dados de porosidade no modelo.....	XIV
Figure I.1 - Transformando dados da escala da sísmica.....	XV
Figure I.2 - Visualização da sísmica no modelo geológico.....	XVI

Figure I.3 - Transferência de um mapa de atributo sísmico ao modelo .....	XVI
Figure I.4 - Visualização do atributo sísmico no modelo geológico criado .....	XVII
Figure J.1 - Opção Petrophysical Modeling para porosidade .....	XVII
Figure J.2 - Detalhes de petrofísica com o cadeado aberto .....	XVIII
Figure J.3 - Método Sequential Gaussian Simulation .....	XIX
Figure J.4 - Parâmetros de variograma para a porosidade .....	XIX
Figure J.5 - Detalhes da aba <i>Distribution</i> para porosidade .....	XX
Figure J.6 - Detalhes da aba Co-kriging para porosidade .....	XXI
Figure J.7 - Visualização da porosidade estimada .....	XXI
Figure K.1 - Opção Petrophysical Modeling para NTG .....	XXII
Figure K.2 - Detalhes da aba <i>Distribution</i> para NTG .....	XXIII
Figure K.3 - Visualização do NTG estimado. ....	XXIII
Figure L.1 - Criação da propriedade Perm utilizando a calculadora. ....	XXIV
Figure L.2 - Opção Petrophysical Modeling para permeabilidade .....	XXV
Figure L.3 - Método Sequential Gaussian Simulation .....	XXV
Figure L.4 - Detalhes da aba <i>Distribution</i> para a permeabilidade. ....	XXVI
Figure L.5 - Detalhes da aba Co-kriging para permeabilidade .....	XXVII
Figure L.6 - Visualização da permeabilidade estimada. ....	XXVII

# 1. INTRODUÇÃO

---

Na modelagem de qualquer sistema físico há incertezas de várias naturezas. Com o passar dos anos, a consciência da importância do estudo das incertezas cresceu significativamente ao ponto de que considerá-las tanto na modelagem quanto na tomada de decisões se tornou crucial para a diminuição de riscos tais como riscos econômicos, sociais e ambientais. Tal aumento de importância dada à temática da incerteza é constatada através do número crescente de trabalhos científicos em diversas disciplinas ao longo dos últimos anos.

Todo o processo de modelagem de fenômenos que ocorrem em um sistema físico envolve incertezas e naturalmente, o resultado é um modelo que contém incertezas, o qual produz respostas que podem se ajustar de maneira razoável aos dados observados, mas que podem levar a previsões errôneas em virtude da representatividade incerta da realidade.

Discutindo-se o contexto da indústria do petróleo, pode-se facilmente vislumbrar o grande universo de incertezas que se faz presente nas múltiplas disciplinas que compõem as atividades de exploração desse recurso desde a aquisição de dados passando pela construção de modelos até a tomada de decisões sobre campos petrolíferos. Afinal, como ter garantia da veracidade e acurácia das características de um sistema, que é uma rocha saturada de fluidos complexos, soterrada a centenas (às vezes até milhares de metros) de profundidade se a maioria das propriedades é inferida, ou seja, medidas indiretamente através de correlações? Pode-se atentar também para a limitação da resolução, alcance e acurácia dos equipamentos usados para aquisição de dados sísmicos, geológicos, perfis de poços e medidores de pressões somada às diferentes interpretações que esses dados podem receber dos diferentes profissionais das disciplinas pertinentes e o custo para se obter mais informações.

Toda essa incerteza não pode ser ignorada, daí a necessidade de elaboração de métodos de visualização e quantificação de incertezas e de fluxos de trabalho que possibilitem a manipulação e análise de tais informações para que decisões errôneas possam ser evitadas.

Esta monografia está estruturada da seguinte maneira:

No capítulo 2 serão apresentadas algumas definições de incertezas, desafios e questões envolvidas na modelagem. Além disso, serão discutidas também fontes de incertezas.

No capítulo 3 será discutida a Quantificação de Incertezas. Também será feita uma revisão da Teoria da Probabilidade para, posteriormente, discutir-se o conceito de Entropia da Informação e como pode ser aplicada na modelagem geológica. Posteriormente, serão abordadas a Análise da Sensibilidade e a Simulação de Monte Carlo que não quantificam a incerteza de maneira direta como a Entropia, mas são de grande importância para a análise da incerteza em diversos parâmetros.

No capítulo 4 são discutidos métodos de visualização de incertezas. Primeiro apresentando técnicas de visualização da sumarização de dados através do *box plot*, suas primeiras versões e algumas de suas variações. Nas seções seguintes são mostrados métodos para visualização de vários tipos de dados de diferentes dimensionalidades, alguns já usuais nas disciplinas das Ciências da Terra e outras que já são usadas em outras áreas do conhecimento, mas que ainda não foram implementadas e podem ter potencial aplicabilidade.

O capítulo 5 trata de análise de incertezas. Primeiramente é justificada a análise da incerteza no processo decisório para depois definir cenários geológicos e realizações destes cenários. Os conceitos os quais são citados frequentemente nesse trabalho e no capítulo 6 serão de suma importância.

O capítulo 6 mostra uma aplicação. Primeiro é montado um modelo geológico no Petrel para depois usar o fluxo de trabalho de incertezas para gerar diversas realizações do modelo. Com isto estas informações são exportadas para o GOCAD, *software* de modelagem geológica da companhia Paradigm, a fim de gerar mapas de desvio padrão e média célula a célula levando em conta todas as realizações. Detalhes do processo de montagem do modelo geológico, passo a passo, podem ser conferidos no Apêndice A.

# Objetivos

Esse trabalho tem por objetivo:

- Apresentar métodos de quantificação de incertezas;
- Reunir as técnicas gráficas de visualização de dados já existentes na bibliografia que possam ser úteis para visualizar incertezas;
- Apresentar metodologias para escolha de modelos representativos de um reservatório;
- Mostrar uma aplicação de como realizações são geradas através de um fluxo de trabalho em *software* comercial usando Simulação de Monte Carlo.

## 2. DEFINIÇÕES E MODELAGEM DE INCERTEZAS

---

### 2.1. O que é Incerteza? Algumas definições.

Nas últimas décadas a produção acadêmica em torno da temática da visualização de incertezas aplicada a diferentes áreas do conhecimento ocasionou o surgimento de diversas definições de "incerteza".

A incerteza é a situação que envolve informação imperfeita ou desconhecida. É um termo usado de formas diferentes em vários campos, incluindo filosofia, física, estatística, economia, finanças, psicologia, sociologia, engenharia, meteorologia, ciência da informação e outras. Aplica-se a previsões de eventos futuros e para medições físicas realizadas. A incerteza surge em ambientes parcialmente observáveis e/ou estocásticos, bem como devido à ignorância.

Descendo a um nível mais específico, o das ciências exatas, incerteza pode ser definida como grau de falta de conhecimento sobre uma quantidade de erro (HUNTER & GOODCHILD, 1993). Em um trabalho posterior (PANG et al, 1996), a incerteza ganhou uma definição de caráter mais estatístico. Foram considerados três tipos de incerteza:

- **Estatística:** dada pela média estimada e pelo desvio padrão que podem ser usados para construir intervalos de confiança ou a distribuição dos dados;
- **Erro:** uma diferença entre as estimativas dos dados, ou entre um valor conhecido e uma estimativa;
- **Range:** que é um intervalo no qual o dado deve existir, mas não pode ser quantificado nas definições de estatísticas ou de erro.

Por muitos anos as incertezas têm sido quantificadas e traduzidas por meio de medidas estatísticas as quais nem sempre são suficientes para assumir tal função. Nas últimas décadas, fluxos de trabalho que consideram incertezas em diversos parâmetros, ferramentas integradas a

esses fluxos (análise da sensibilidade e simulação de Monte Carlo) e até novas correlações para quantificar diretamente as incertezas vêm sido criados e aprimorados afim de adquirir informações sobre o desconhecido.

No âmbito das disciplinas pertinentes à geologia, que é o nosso interesse aqui, um outro autor (CAERS, 2011) define que a incerteza é causada por um entendimento incompleto sobre o que se quer quantificar. Quantificar o que não se sabe é uma tarefa subjetiva. Qualquer avaliação de incerteza deverá ser feita através de algum tipo de modelo, isso porque não existe uma maneira absolutamente correta de se quantificar incerteza e sim modelos de quantificação de incerteza e que em cada problema em particular deve ser estudado qual o modelo mais apropriado a ser usado. Há incerteza até na escolha do modelo. Por essas dificuldades de entendimento da falta de informação sobre o que não se conhece, é dito que modelar incertezas em todos os níveis de estruturação de um problema, além de impossível, trabalhoso e dispendioso, pode não ser necessário para uma decisão relativamente simples.

Na próxima seção serão abordados alguns desafios e questões importantes sobre a modelagem de incertezas no campo das ciências da terra, que engloba a temática de reservatório de petróleo.

## **2.2. Desafios e questões na modelagem de incertezas.**

Algumas ferramentas de modelagem de fenômenos aleatórios através de modelos tradicionais de probabilidade são muito rígidas para manusear todas as suas complexidades (CAERS, 2011). Dentro desse cenário de complexidades, podem ser notadas algumas questões e desafios tais como:

- Aspecto do alto número de dimensões: não há dúvida do alto grau de complexidade dos sistemas físicos que ocorrem na geologia, para isso são necessárias muitas variáveis para descrevê-los. Quando se tem um modelo geológico, normalmente a malha contém um número de células que pode variar de algumas dezenas de milhares a milhões de células e cada célula contém algumas variáveis, então facilmente têm-se milhões de variáveis até em modelos relativamente pequenos. Por exemplo, a propriedade “porosidade” da malha de um modelo geológico, pode ter tantas dimensões quanto o número de células do modelo;
- A subjetividade da análise de incertezas: um sistema existe com toda a sua verdade, mas com propriedades desconhecidas. Qualquer análise deve ser feita com base em um

modelo, ou seja, cada caso de modelagem será um problema específico e particular, cada um com seu grau de detalhamento e exigência de informações;

- Modelar incertezas às vezes é uma tarefa que deve ser adaptada. Se a tarefa muda, então o tipo de modelagem e aproximação de incertezas também será diferente. Construir um modelo de incerteza que inclua todos os aspectos e fontes de incerteza é muito difícil e às vezes não necessário;
- Existem em geral poucas fontes de dados, além disso, estas fontes possuem diferentes resoluções. É um desafio lidar com esta variedade dos dados para restringir os modelos de incerteza. Sem dados, a modelagem se torna impossível. Tais dados podem ser obtidos através de poços (tais como testemunhos, ou testes de poços). Ou obtidos de forma mais indireta pela geofísica ou levantamentos remotos de sensoriamento a serem correlacionadas.

## 2.3. Fontes de Incerteza

Fontes de incertezas podem se diferenciar de acordo com a tarefa que se deseja realizar. Como esta monografia trata de visualização e quantificação de incertezas em reservatórios de petróleo, nesta seção serão abordadas as principais fontes de incertezas na modelagem geológica e no processo de visualização de dados. Nos dois casos as fontes podem ser análogas uma vez que quando se constrói um modelo de um sistema físico, posteriormente se deseja visualizar de algum modo os resultados de alguma simulação ou fluxo de trabalho.

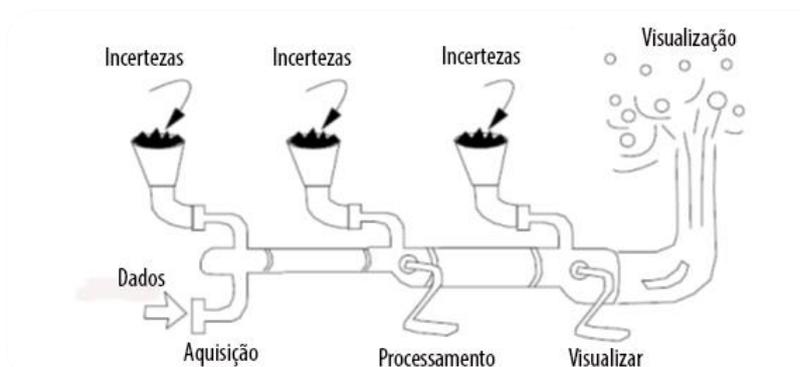
Fontes de incertezas já foram especificadas por vários autores e é possível observar que apesar de as agruparem em tipos e fontes aparentemente diferentes, consegue-se perceber aspectos e uma ideia central em comum a todas elas. Aqui nessa seção sintetizam-se tais fontes segundo basicamente dois autores ((CAERS, 2011) e (WITTENBRINK, C. M.; PANG, A. T.; LODHA, S. K., 1996)) sendo que o primeiro cita fontes de incertezas relacionadas ao processo de modelagem nas Ciências da Terra enquanto que o segundo discorre sobre fontes de incertezas no processo de visualização de dados em um contexto mais geral.

As fontes relacionadas por estes autores são as seguintes:

- a. Incerteza relacionada a medições de erros e processamento de medições diretas. Quanto mais uma medida for realizada, mais confiável essa será. Incertezas diferentes são associadas a uma medição dependendo se essa for feita por uma pessoa ou por uma máquina;

- b. Incerteza relacionada à interpretação de dados processados. Tais dados podem ser interpretados de diversas maneiras e cada interpretação e processamento necessita de um modelo próprio;
- c. Incerteza relacionada ao cenário geológico usado, o qual é interpretado através de dados ou baseados em modelos físicos os quais são incertos (na seção 5.1 será dada uma definição de cenário geológico);
- d. Incerteza Espacial: ainda que os dados fossem perfeitamente medidos, esses ainda são muito dispersos em relação à resolução que se deseja construir os modelos geológicos. Isso significa que vários modelos com diferentes distribuições espaciais de propriedades ou de estrutura de camadas podem ser gerados se encaixando perfeitamente com os dados medidos. É caso da estimativa das propriedades geológicas entre um poço e outro através dos dados de perfis de poço;
- e. Incerteza da Resposta: estão relacionadas aos processos do modelo como, fluxo, transporte, ondas, equações de calor ou ainda decisões feitas baseadas nesses modelos. Pode haver incerteza relacionada à física desses processos ou outros parâmetros que precisam ser especificados para especificar tais processos. Exemplo, resolver uma equação diferencial requer condições de contorno e iniciais, as quais podem ser incertas. Exemplo: pressões na fronteira de um reservatório, profundidade do contato óleo-água e outros.
- f. Aleatoriedade do Processo: provenientes da aleatoriedade e da natureza caótica dos fenômenos naturais. Exemplo: comportamento das nuvens, escoamento turbulento seja em um tubo ou sobre uma placa;
- g. Entendimento limitado: essa fonte diz respeito à falta de conhecimento daquele que realiza o estudo ou a modelagem de algum processo em questão;
- h. Ignorância da presença de fenômenos: essa incerteza está relacionada com fenômenos ou processos que nem se quer imagina-se que estão presentes ou que são plausíveis. Não se observa, nem se pensa que são teoricamente ou praticamente possíveis;
- i. Incerteza relacionada ao fato de que algum fenômeno não pode ser medido por estar ocorrendo em um lugar muito distante ou inacessível, por exemplo, propriedades no núcleo interno da Terra ou em outro planeta.

O segundo autor (WITTENBRINK, C. M.; PANG, A. T.; LODHA, S. K., 1996) expressa através de uma figura semelhante a Figura 2.1 as três fontes (ou etapas) principais de incertezas que são incorporados no processo de visualização de dados, começando pela aquisição de dados e outras etapas até gerar uma visualização repleta de incertezas.



**Figura 2.1 - Uma tubulação esquemática que mostra as etapas em que incertezas são incorporadas no processo de visualização. Adaptado de (PANG, WITERNBRINK, & LODHA, 1996).**

Para melhor explicar a Figura 2.1 a seguir são feitos alguns comentários em relação a partes específicas desta Figura.

**Incerteza na Aquisição de dados:** Assim como o primeiro autor (CAERS, 2011), este autor também cita as incertezas que são englobadas nesta fase inicial da visualização que na verdade, é uma etapa em comum com o processo de modelagem já que toda visualização de dados é resultado da simulação de algum modelo. A aquisição engloba dados oriundos da medição de instrumentos ou resultados de modelos numéricos os quais possuem variações estatísticas. No que se diz respeito à medição de instrumentos, a incerteza é proveniente da precisão ou calibração do equipamento, se essa medida é feita por uma máquina ou ser humano (aqui entra a incerteza referente às observações por diferentes pessoas, pois cada um possui percepções ou leituras ligeiramente diferentes do outro), tais componentes são causas das variações estatísticas dos dados. Referente às variações estatísticas dos resultados de saída dos modelos numéricos, esses são simplificações da realidade, então frequentemente têm-se parâmetros de entrada sendo inferidos pelos cientistas, pois é impossível saber ou medi-los com exatidão. Somado a isso, têm-se as correlações, equações diferenciais que tentam expressar os fenômenos que ocorrem em tal sistema. Tais equações por si só são aproximações da realidade, e muitas vezes são impossíveis de serem resolvidas de maneira analítica, e portanto, são usados

métodos numéricos para resolvê-las, de novo, aproximações. Aproximações e simplificações propagam incertezas.

**Incerteza no Processamento ou Transformação:** Antes de serem renderizados, muitas vezes os dados passam por transformações que podem acontecer mesmo sem o conhecimento da pessoa que está envolvida na tarefa de visualização. Essas transformações podem ser simples como uma conversão de unidades ou mudança de escala ou mais complexas como o processamento de tais dados por algum tipo de algoritmo ou correlação para originar um novo tipo de dado que seja desejável visualizar. A ideia principal dessa fonte é que tais transformações alteram a forma original dos dados e que correlações físicas e algoritmos inserem incertezas, começando pelo simples fato dos cálculos de um computador terem uma precisão limitada e pela falta de conhecimento por parte da pessoa que visualiza tais dados a quais transformações esses estão sendo submetidos. Tais transformações podem acontecer precocemente, logo após a aquisição de dados ou mais tardiamente em outros processos envolvidos na visualização;

**Incerteza na Visualização:** Incerteza introduzida nos processos relacionados à visualização de dados como diferentes métodos de renderização de volumes, de radiossidade (iluminação de objetos 3D na computação gráfica), de interpolação (1D, 2D ou 3D) em caso de dados dispersos, de contorno de isosuperfícies e de animação (evolução dos dados no tempo). Existem diversos métodos computacionais para realizar esses processos, cada um com suas limitações e por sua vez, produzem renderizações e disposições ligeiramente diferentes de um mesmo conjunto de dados. Tal dúvida sobre qual método produziria uma visualização mais “verdadeira” é a principal incerteza dessa categoria de fonte.

## 3. QUANTIFICAÇÃO DE INCERTEZAS

---

Vários autores, tais como (POTTER, ROSEN, & JOHNSON, 2012), mostram que ao lado da visualização da incerteza existem questões importantes relacionadas à quantificação da incerteza. Isto porque os recursos de visualização, segundo estes autores, são limitados, sendo necessárias formas de quantificação que simplifiquem e resumam informações de incerteza. Neste sentido este capítulo apresenta uma revisão da Teoria da Probabilidade que será necessária para definir os conceitos de entropia da informação. Embora não seja explorado nesse capítulo, é importante ressaltar que a quantificação muitas vezes foi associada a parâmetros estatísticos tais como a média e desvio-padrão, quantificando a incerteza entendida através da dispersão maior ou menor dos dados. Tais parâmetros são subsídios para construção de intervalos de confiança onde probabilidades são associadas.

Na seção 3.2 serão apresentados os conceitos da Entropia da Informação e suas aplicações para a quantificação de incertezas em reservatório de petróleo. Vale ressaltar que no estudo de caso que será apresentado no capítulo 6 não serão aplicadas nenhuma dessas definições de entropia. Aplicações que envolvem Entropia podem ser, portanto, o foco de trabalhos posteriores.

Na seção 3.3 será apresentado o conceito de Análise da Sensibilidade por ser de suma importância para o conhecimento de parâmetros que possuem impactos significativos sobre uma determinada função objetivo. Sua importância ficará mais clara no capítulo 5.

A simulação de Monte Carlo será abordada na seção 3.4 deste capítulo, ela é uma técnica que usa conceitos estatísticos e probabilísticos a fim de cobrir os intervalos de valores que parâmetros incertos podem assumir. Com esta técnica é possível gerar diferentes realizações de um reservatório dentro de um cenário.

### 3.1. Teoria da Probabilidade

A probabilidade é um conceito matemático, uma vez que não se pode observar a probabilidade na natureza. E o motivo de ser abordada aqui é que além de ser a forma mais

aceita, antiga e consolidada de se expressar incertezas e por seus conceitos serem amplamente usados na Simulação de Monte Carlo e comporem a definição de Entropia da Informação.

A probabilidade de um evento acontecer pode ser interpretada como a taxa de sucesso na realização de um experimento repetitivo, ainda que eventualmente, a repetição não seja explícita.

$$\text{Probabilidade (evento)} = \frac{(\# \text{ de eventos bem sucedidos})}{\# \text{ total de tentativas}} \quad (1)$$

### 3.1.1. Espaço amostral, Evento e Resultado

Para ilustrar os conceitos, seja a situação hipotética da exploração de diamantes. A ação de pegar uma pedra do depósito é considerada como experimento. O resultado do experimento não é conhecido com exatidão antes de sua realização. No entanto, assume-se que peso do diamante é menor do que um grande número, “GRANDE”, e maior que zero. O conjunto de todos os possíveis resultados é definido como Espaço Amostral (S). Por exemplo:

Tamanho dos diamantes:  $S = (0, \text{GRANDE})$

Qualquer subconjunto dessas amostras é considerado um evento:

Evento  $E_1 = \{\text{um diamante de tamanho superior a 5 ct}\}$

Evento  $E_2 = \{\text{um diamante de tamanho entre 2 e 4 ct}\}$

A probabilidade de um evento ocorrer é denotada como  $P(E)$ . Daí são definidos os axiomas da probabilidade que se seguem:

*Axioma 1:*  $0 \leq P(E) \leq 1$

*Axioma 2:*  $P(S) = 1$

*Axioma 3:*  $P(E_1 \cup E_2 \cup E_3 \dots) = P(E_1) + P(E_2) + P(E_3) + \dots$  se  $E_1, E_2, E_3, \dots$  são eventos mutuamente excludentes (o que indica que  $E_i$  e  $E_j$  não podem ocorrer ao mesmo tempo).

### 3.1.2. Probabilidade condicional

Todo dia faz-se perguntas do tipo:

Qual a probabilidade de se achar óleo na localização  $y$  dado que óleo foi achado na localização  $x$ ?

Qual a probabilidade do sítio  $x$  está contaminado dado que foi amostrado na localização  $z$  uma concentração de  $y$  ppm?

Geralmente, deseja-se calcular a probabilidade de um evento acontecer, dado que uma certa informação já está disponível previamente. A esta probabilidade é dado o nome de probabilidade condicional, a qual é representada da seguinte forma:

$$P(\text{Evento } E \text{ ocorrer} / \text{Evento } F \text{ ocorrer}) = P(E/F) \quad (2)$$

Ao avaliar probabilidades condicionais é necessário saber quais eventos estão relacionados ou não, caso não sejam,  $P(E/F)$  é simplesmente igual a  $P(E)$ .

Dado que uma moeda regular é jogada nove vezes foram observadas coroas em todas as nove vezes, qual é a probabilidade de ter coroa na décima jogada? A resposta é  $1/2$  pois os eventos são assumidos como independentes.

### 3.1.3. Lei de Bayes

A equação da Lei de Bayes é um dos conceitos chave na Teoria da Probabilidade e de enorme importância para as Ciências da Terra. Por exemplo, a probabilidade de se achar diamante na região é uma probabilidade incondicional. Quanto mais informação estiver disponível, a probabilidade de um evento “mudar”, crescendo ou diminuindo, isso é, se “aprende” através dos dados. A magnitude dessa mudança é governada pela Lei de Bayes.

A Regra de Bayes pode ser deduzida como se segue:

Sejam:

$$(I) P(E \text{ e } F) = P(E/F)P(F)$$

$$(II) P(E \text{ e } F) = P(F/E)P(E)$$

Dividindo se (I) por (II) obtém-se:

$$1 = \frac{P(E/F)P(F)}{P(F/E)P(E)} \quad (3)$$

Portanto:

$$P(E/F) = \frac{P(F/E)P(E)}{P(F)} \quad (4)$$

Essa equação oferece uma relação entre a probabilidade condicional e a incondicional. Se a probabilidade condicional é conhecida, então a outra probabilidade condicional pode ser obtida usando a regra de Bayes.  $P(E/F)$  é também referida como probabilidade “a posteriori” (depois de se aprender através dos dados), enquanto que  $P(E)$  é referida como probabilidade “a priori” (antes de se obter dados).

## 3.2. Entropia da Informação

A Entropia da Informação é uma forma de quantificar incerteza feita através da quantidade de informação e encontra aplicações em diversas áreas. Neste caso, entropia é incerteza. O objetivo desta seção é mostrar o significado da entropia da informação e como ela pode ser aplicada como parte da modelagem de quantificação de incertezas em reservatórios de petróleo.

O primeiro conceito de Entropia da Informação foi definido (SHANNON, 1948) para saber a quantidade de informação necessária para transmitir um texto em Inglês. A ideia era que dada a probabilidade das letras do alfabeto inglês ocorrerem em um texto de uma carta, é possível estabelecer uma medida descrevendo a informação que falta para determinar o texto inteiro de uma mensagem enviada parcialmente. Depois de mais algum embasamento teórico, a seguinte equação foi definida para medir a falta de informação:

$$H = - \sum_i^N p_i * \log(p_i) \quad (5)$$

onde  $H$  é a Entropia da Informação definida como a soma de todos os produtos de probabilidade  $p$  para cada resultado  $i$  possível de um total de  $N$  resultados possíveis pelo seu logaritmo. O valor mínimo é 0, porque  $\log 1 = 0$  e  $\lim_{x \rightarrow 0} (x \log x) = 0$ . O logaritmo pode ser tomado em qualquer base, dependendo da unidade da informação aplicada.

Para ilustrar o conceito básico da Entropia da Informação, considere o exemplo de um experimento qualquer com dois resultados possíveis. Se a probabilidade dos dois eventos ocorrerem for igual, então a entropia é máxima. Por outro lado, se um evento é mais provável que o outro, a entropia é menor, ou seja, intuitivamente se tem mais informação sobre a maior chance da ocorrência desse evento. Em um caso extremo onde a probabilidade de ocorrência de um evento é 1, na qual um resultado é sempre o mesmo, a entropia seria 0, ou seja, conhecimento máximo ou nenhuma incerteza sobre o resultado. Um experimento que se encaixa nesse caso é o de se arremessar uma moeda e observar o resultado (cara ou coroa). A Figura 3.1 é o gráfico gerado usando a equação de definição de entropia de Shannon em um experimento no qual há dois resultados possíveis. No gráfico é mostrado o comportamento da Entropia de Shannon em função da probabilidade do primeiro evento. Como mencionado antes, observa-se o máximo da entropia quando os eventos são equiprováveis ( $p_1 = p_2 = 0,5$ ).

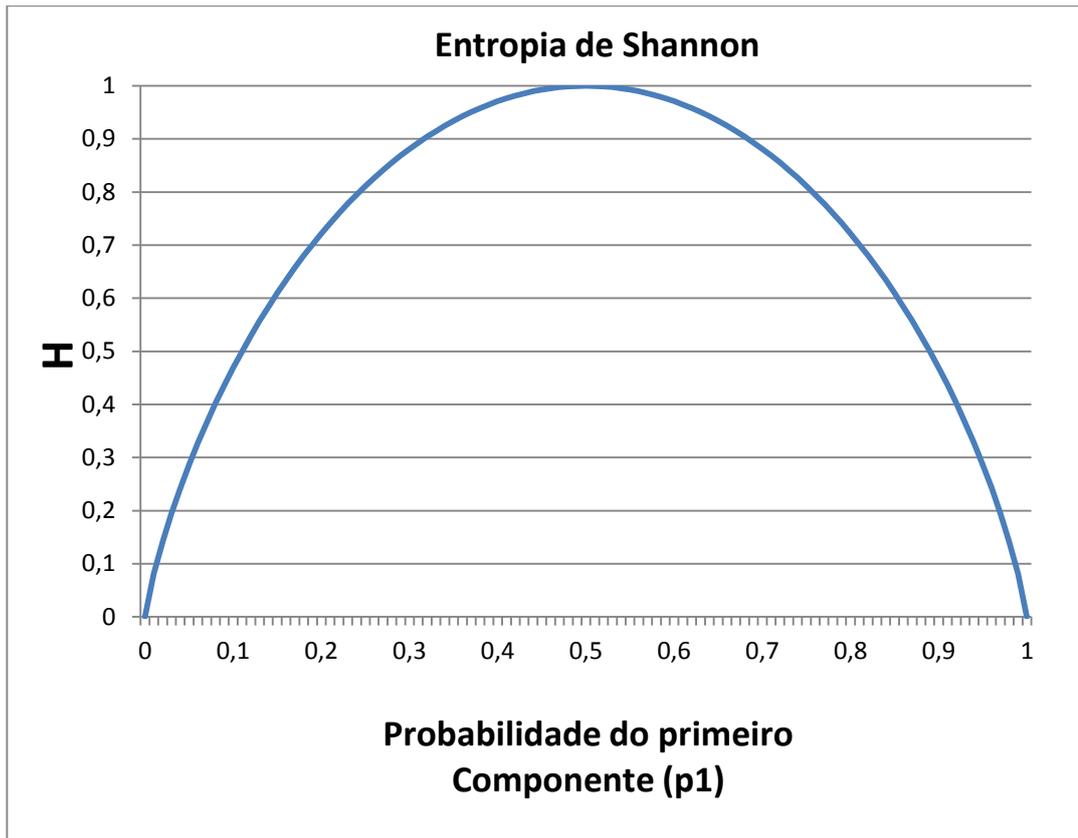


Figura 3.1- Entropia de Shannon para um experimento de dois eventos possíveis em função do primeiro resultado.

### 3.2.1. Entropia da Informação aplicada a um conceito Espacial

Em um contexto espacial, pode-se interpretar a entropia de uma sub-região (células de uma malha) como a quantidade de falta de informação com respeito a propriedades discretas das células. Isso é possível se cada unidade geológica, for considerado um evento exclusivo onde cada célula pode ter de 1 até M unidades geológicas.

Para cada sub-região discreta, pode-se descrever a entropia da informação como:

$$H(x, t) = - \sum_{m=1}^M p_m(x, t) \log p_m(x, t) \quad (6)$$

onde  $x$  é a localização da sub-região (célula ou elemento discreto da malha) e  $M$  o número de possíveis unidades geológicas (ou membros) que tal sub-região pode ter. A variável “ $t$ ” poderia ser o tempo físico ou qualquer outro parâmetro que descreva a evolução do modelo.

A Figura 3.2 ilustra o parágrafo anterior. Na parte (a) da figura há uma malha regular com regiões de cores diferentes indicando três tipos possíveis de unidades geológicas (vermelha, verde e azul), e fronteiras entre tais regiões que não estão claras (incertas), onde mais de uma unidade geológica pode ocorrer em cada célula. A ocorrência de cada unidade geológica em uma célula é tratada como um evento de probabilidade  $p_m$  e na parte (b) da figura são mostrados 3 mapas de probabilidade, um para cada região. Pode ser notado que na região das fronteiras incertas, as probabilidades de cada unidade geológica diminuem pois há ocorrência conjunta de uma ou mais unidades geológicas. Quanto mais unidades geológicas forem possíveis de ocorrer em uma sub-região, mais as probabilidades  $p_m$ 's são próximas de  $1/M$  e maior será a entropia, conseqüentemente, a incerteza de se saber em qual região nos encontramos. No mapa (c) observa-se a entropia de informação de cada célula. Quanto mais próximo da área central, onde há interseção de três fronteiras incertas, maior a entropia da informação pois não se sabe ao certo qual das três unidades geológicas estão ocorrendo ali.

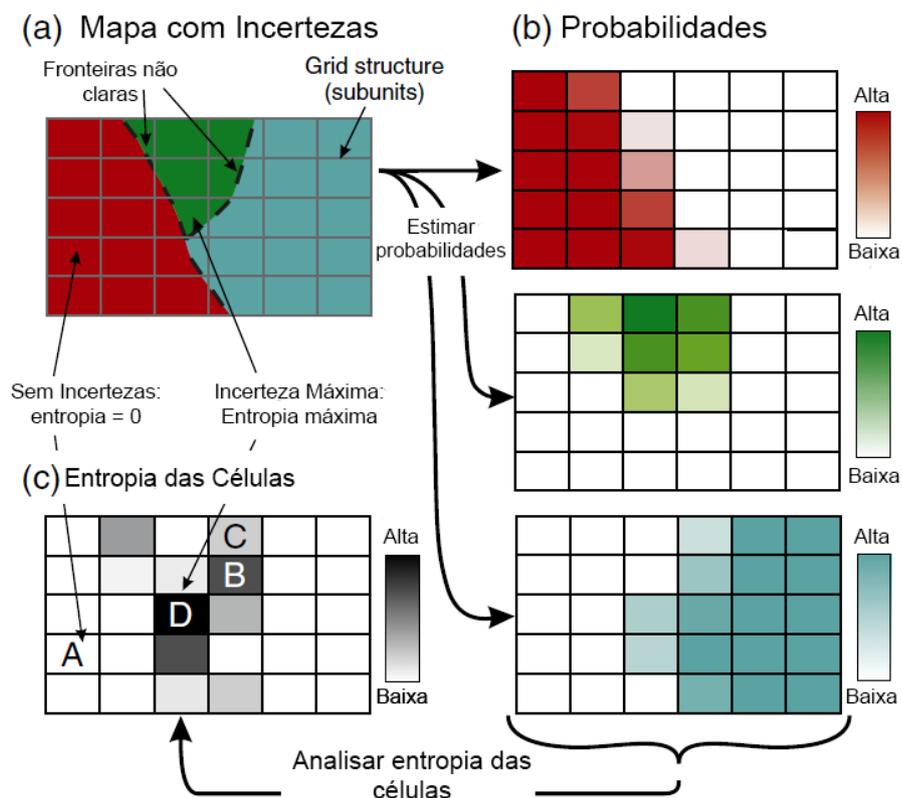


Figura 3.2 – Aplicação do conceito espacial da Entropia da Informação. Adaptado de (WELLMANN & REGENAUER-LIEB, 2012).

Há também uma outra interpretação da entropia da informação como medidas *fuzzy* (WELLMANN & REGENAUER-LIEB, 2012). O significado de tal interpretação se relaciona

à quão precisa é a ocorrência de uma unidade geológica específica em todo modelo geológico. Este cálculo de entropia será importante para as aplicações em modelos de reservatórios apresentadas nas próximas seções.

A associação entre a entropia e a teoria *Fuzzy* é realizada através de uma função (função de pertinência *fuzzy*)  $f$  que varia no intervalo  $[0,1]$  e que tem as seguintes propriedades:

- A medida deve ser 0, se e somente se,  $f$  é 0 ou 1 em todo modelo;
- O valor da medida atinge seu máximo quando a  $f$  assume valor 0,5 em todo modelo.

Tais propriedades são satisfeitas pela Entropia de Shannon onde a função  $f$  toma o lugar da probabilidade  $p_m$  de um resultado. Conforme apresentado em (WELLMANN & REGENAUER-LIEB, 2012). A medida *fuzzy* como entropia pode ser referida como  $H_m$ , normalizada pelo número total de células  $N$ :

$$H_m(t) = -\frac{1}{N} * \sum_{x=1}^N [p_m \log p_m + (1 - p_m(x, t)) \log(1 - p_m(x, t))] \quad (7)$$

### 3.2.2. Entropia Total de um Modelo

Esse é uma extensão dos conceitos apresentados até agora. A entropia total da informação  $H_t$  de um modelo espacial qualquer pode ser calculada do seguinte modo:

$$H_t(t) = -\frac{1}{N} * \sum_{x=1}^N \sum_{m=1}^M p_m(x, t) \log(p_m(x, t)) \quad (8)$$

onde o índice “ $m$ ” identifica o membro da sub-região de locação “ $x$ ”.

A entropia total do modelo é igual a zero quando todas as subdivisões  $x$  no modelo estão associadas a um membro, e é máxima quando em todas as subpartes do modelo a probabilidade de todos os membros é exatamente  $1/M$ .

### 3.2.3. Aplicação em Reservatórios

O conceito de entropia da informação foi aplicado em modelos geológicos gerando mapas espaciais (WELLMANN & REGENAUER-LIEB, 2012) como se a entropia da

informação fosse mais uma propriedade da malha. Conseguir observar espacialmente o quanto de incerteza se apresenta em cada localidade do modelo é de suma importância para decidir se serão necessárias novas aquisições de dados para alimentar o modelo de modo a diminuir tais incertezas.

Para o caso de  $t = 0$ , os conceitos mencionados nas seções anteriores deste capítulo podem ser diretamente aplicados em modelos geológicos. Se forem considerados os possíveis resultados, ou membros  $m \in M$  como unidades geológicas possíveis em um modelo, uma subdivisão do espaço modelo em um número discreto de células de igual tamanho com uma posição  $x$ , e a probabilidade que a unidade geológica existe em uma célula  $p_m(x)$ , pode-se aplicar a medida de entropia para :

- Calcular entropia das células para e usá-las para visualizar incertezas;
- Avaliar a incerteza de um modelo geológico inteiro através de sua medida *fuzzy* relacionada;
- Usar a entropia total do modelo como um meio de quantificar a incerteza de todo o modelo com um *único número* e comparar com outros modelos ou realizações.

A Entropia da Informação pode ser aplicada para quantificar incertezas em modelos geológicos desde que suas unidades geológicas sejam tidas como resultados exclusivos e discretos e que as probabilidades de cada célula possam ser estimadas.

Satisfazendo tal condição, há basicamente cinco passos a serem seguidos para o cálculo das incertezas em modelos geológicos oriundas da imprecisão de entrada de dados por meio de aproximação estocástica:

1. Construção de um modelo geológico que seja o mais representativo possível usando todas as informações disponíveis;
2. Atribuir distribuição de probabilidade aos dados de entrada (medidas de posição e orientação dos dados);
3. Baseado no modelo geológico inicial e definidas as distribuições de probabilidade,  $n$  conjuntos de dados de entradas diferentes são gerados por aproximação estocástica. Um exemplo de conjuntos de dados são as realizações (seção 5.1) gerada pelo fluxo de trabalho de análise de incertezas do *software* Petrel (seção 6.2), que utiliza Simulação de Monte Carlo como ferramenta de

amostragem das variáveis de incerteza a partir de uma distribuição de probabilidade pré-determinada (seção 3.4);

4. Os diferentes conjuntos de dados gerados são automaticamente recalculados para gerar  $n$  representações de modelos geológicos possíveis;
5. Os modelos geológicos são exportados em algum formato útil (universal) e são processados para analisar e visualizar as incertezas.

Primeiramente, determina-se uma função de indicação para cada unidade geológica  $F$ . Essa função é um subconjunto de todo o espaço modelado e definida como:

$$I_f(x) = \begin{cases} 1, & \text{se } x \in F \\ 0, & \text{se } x \notin F \end{cases} \quad (9)$$

Para os  $n$  modelos geológicos, obtém-se  $n$  campos indicadores para cada unidade geológica  $F$  e pode-se usar essas funções para estimar uma função de probabilidade indicadora:

$$P_f(x) = \sum_{k \in n} \frac{I_{Fk}(x)}{n} \quad (10)$$

Agora é possível usar a função de probabilidade indicadora para avaliar as medidas de entropia.

Para ilustrar o conceito, considere o modelo geológico na Figura 3.3 abaixo:

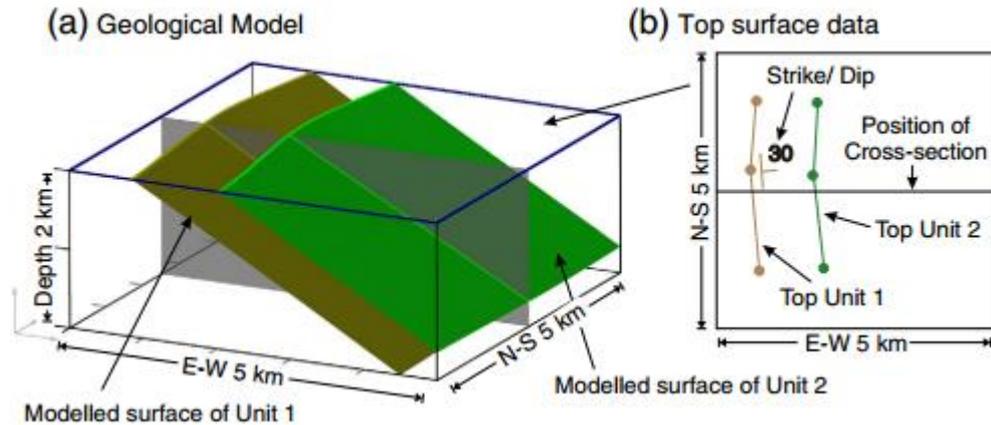


Figura 3.3 – (a) Modelo geológico com as superfícies modeladas das unidades geológicas. (b) Visão do topo do modelo, pode-se ver as linhas de interseção do topo com as superfícies das unidades geológicas 1 e 2. (WELLMANN & REGENAUER-LIEB, 2012).

Este modelo geológico mostra duas superfícies modeladas, subparalelas, que separam as unidades geológicas 1, 2 e 3. É perturbado 50 vezes, de maneira que são geradas 50 realizações desse modelo. Essas 50 realizações recebem uma malha idêntica. De maneira que a equação (9) possa ser usada. Ou seja, o número de ocorrência de cada unidade geológica F1, F2 e F3 é contado para cada posição x da malha. Suponha uma unidade geológica F1 em um ponto x da malha. Então são contadas quantas vezes F1 ocorreu naquele ponto x da malha nas 50 realizações. O mesmo para as outras unidades geológicas.

Usando a equação (10) é possível atribuir probabilidades. Se no ponto x, F1 ocorre 4 vezes nos 50 modelos então a Probabilidade de F1 naquele ponto x é de 4/50, se ocorre 50 vezes a probabilidade é 50/50, etc.... O mesmo para as outras unidades geológicas para o ponto x. Isto vai gerar mapas probabilidades tais como descritas nos cubos abaixo:

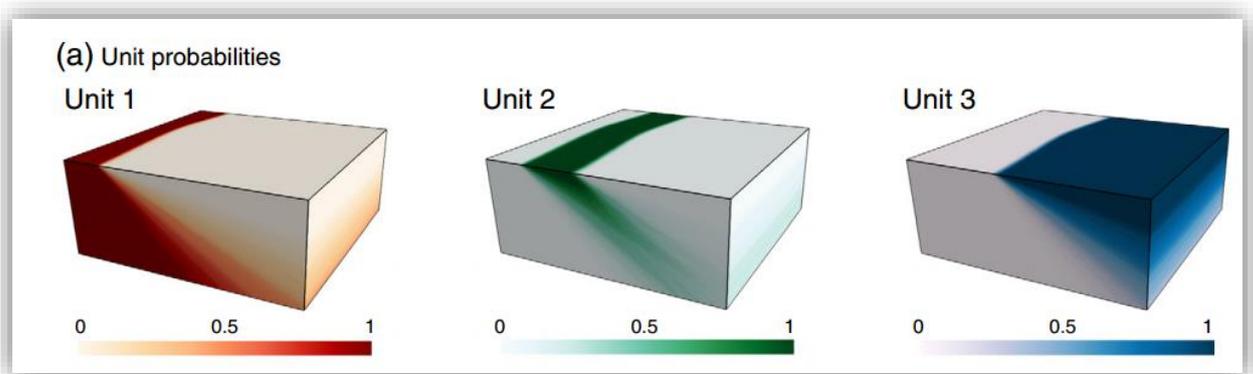


Figura 3.4 – Mapas de Probabilidade para as unidades geológicas F1, F2 e F3. (WELLMANN & REGENAUER-LIEB, 2012)

Com estas probabilidades acima é possível através da equação (6) gerar as Entropias mostradas na Figura 3.5:

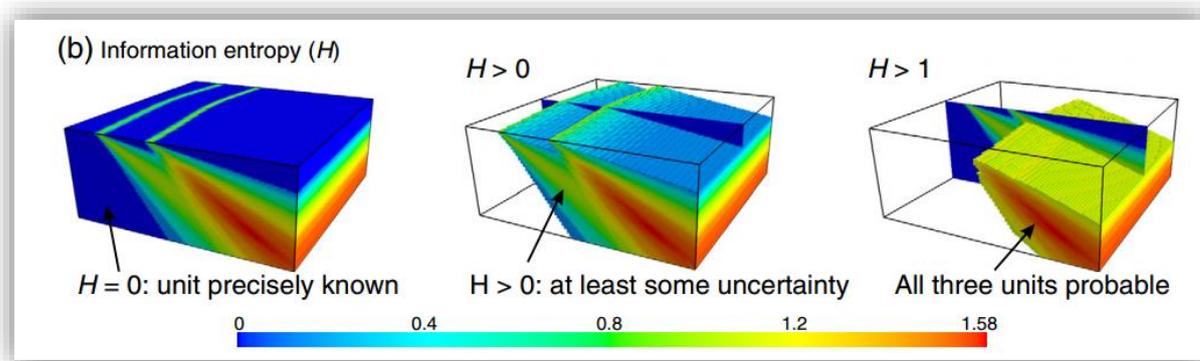
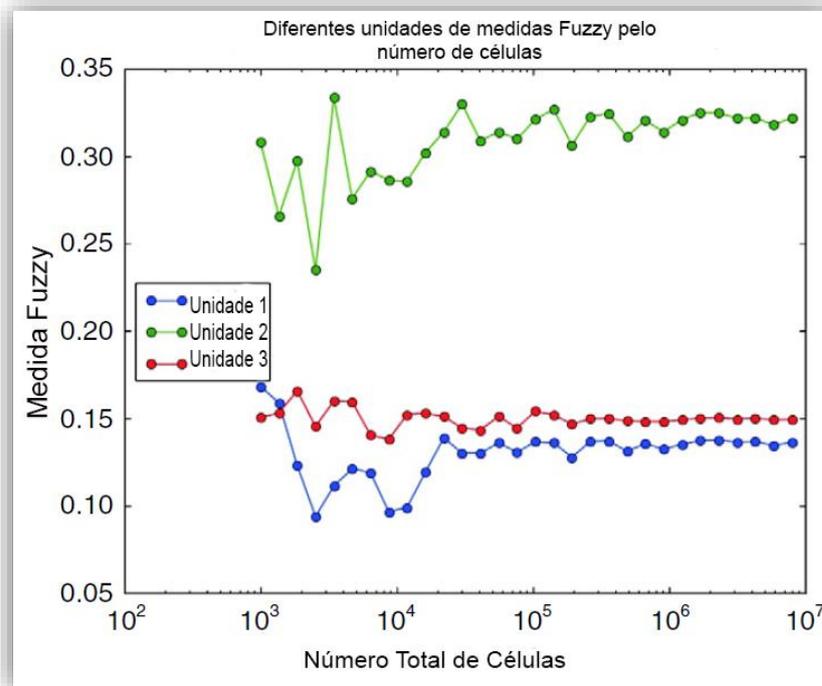


Figura 3.5 – Mapas de Entropia da Informação. Na figura foram gerados 3 mapas, um com o intervalo total de valores de  $H$ , e os outros com  $H > 0$  e com  $H > 1$ . (WELLMANN & REGENAUER-LIEB, 2012)

O conceito de entropia da informação ainda pode ser aplicado para estudar a diminuição de incertezas com a adição de novos dados (que na citação o exemplo foi da aquisição de novos dados provenientes da perfuração de um novo poço) nos modelos comparando as suas respectivas entropias médias (antes e depois do conhecimento dos novos dados) onde essas tendem a diminuir conforme se obtém mais informação.

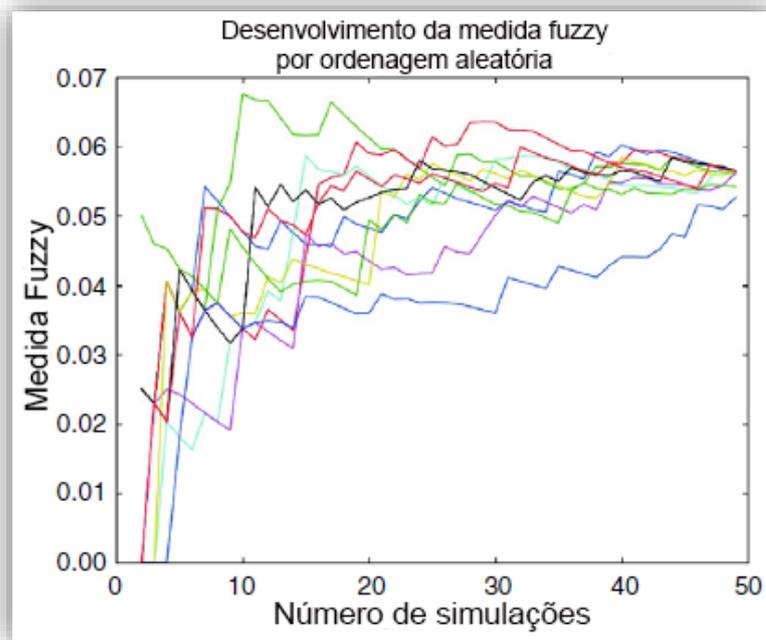
Outra aplicação é usar a entropia da informação como medida *fuzzy* para determinar a quantidade de células que um modelo deve ter afim de que se capture satisfatoriamente as incertezas. Uma malha com células grandes possui baixa resolução e representatividade do modelo. Perde-se informação, ou seja, a entropia da informação tende a ser maior em modelos com células largas já que esses tendem a ser pouco representativos. Tal prática seria feita medindo as entropias com unidade de medida *fuzzy* das células de um modelo e depois dividindo tais células em outras de menor tamanho, aumentando assim, o número total no modelo e posteriormente, medindo novamente a entropia desse modelo com células menores. De acordo com a Figura 3.6 observa-se como a oscilação da unidade da medida *fuzzy* diminui até chegar a aproximadamente um milhão de células, para todos os três esquemas (unidade 1, 2 e 3). Ou seja, a incerteza passa a ter oscilações cada vez menos significativas ao utilizar-se uma malha com número de células maior do que  $10^6$ , portanto, esse número de células seria o suficiente para capturar as incertezas do modelo. Por outro lado, é preciso conciliar o número

de células com o tempo computacional para simular as incertezas e posteriormente, fazer uma Transferência de Escala (*upscaling*) para simulação de fluxo, já que um modelo com muitas células se torna muito exigente em tempo computacional para simular qualquer processo.



**Figura 3.6 - Avaliação da unidade de imprecisão em um modelo com diferentes números de células. Adaptado de (WELLMANN & REGENAUER-LIEB, 2012).**

A teoria dos conjuntos *fuzzy* também pode ser utilizada como critério de convergência para auxiliar na decisão de quantas simulações da incerteza são necessárias afim de se obter uma estimativa estatística satisfatória sobre a simulação dessa. A ideia é simular um modelo geológico determinada quantidade de vezes, ordenar aleatoriamente tais simulações, e depois observar a oscilação da unidade de *fuzziness* associado a cada uma delas. Cada linha na Figura 3.7 representa uma ordem aleatória das mesmas simulações calculando a entropia da informação como uma medida *fuzzy*. É visto que por volta de 40-50 simulações as curvas tendem a convergir para valores entre 0,05 e 0,06, então, segundo esse critério, essa seria uma faixa aceitável de número de simulações para representar, de forma satisfatória, as incertezas do modelo.



**Figura 3.7 - Medida *fuzzy* de uma unidade geológica para conjuntos aleatoriamente escolhidos de 50 realizações simuladas. Adaptado de (WELLMANN & REGENAUER-LIEB, 2012).**

(WELLMANN & REGENAUER-LIEB, 2012) mostra aplicações úteis dos conceitos de entropia para analisar e visualizar incertezas em reservatórios de petróleo, no entanto, aplicações se mostram a nível básico, qualitativo, quando se sabe a probabilidade das unidades geológicas. Um problema de aplicabilidade já se mostra na tarefa de atribuir tais probabilidades célula a célula, principalmente tratando-se de modelos mais complexos de reservatórios e que critérios serão usados em tais atribuições, dando margem assim, ao desenvolvimento de novos trabalhos dessa matéria. Uma sugestão para a continuação das pesquisas a esta monografia, em nível mais avançado, seria integrar as aplicações de entropia nas convenções de análise de incerteza já existentes, tentando usar suas informações adicionais correlacionando-os com os resultados de procedimentos já existentes.

### **3.3. Análise de Sensibilidade**

A Análise da Sensibilidade não é um método direto de quantificação de incertezas como a entropia da informação da seção anterior, e sim uma ferramenta que considera tais incertezas carregadas através de inúmeros parâmetros do modelo para solucionar problemas que têm muitas ou infinitas soluções. Um dos seus principais objetivos é diminuir a dimensionalidade

de um determinado problema indicando os parâmetros de maior influência em alguma função objetivo. Os resultados da Análise de Sensibilidade são visualizados por meio do diagrama de tornado já bastante difundido em vários campos do conhecimento científico.

A Análise de Sensibilidade pode ser usada para determinar a variação dos resultados de modelos de simulação em relação a diferentes valores usados nos parâmetros de entrada do modelo, e assim identificar quais desses parâmetros tem um maior (ou menor) impacto sobre uma determinada função objetivo definida pelo usuário, como por exemplo, a produção acumulada de óleo. É estabelecido um número limitado de rodadas de simulação para determinar os parâmetros que podem ser alterados em estudos posteriores e a faixa de valores em que podem ser alterados. A informação obtida é usada no planejamento de processos como o ajuste de histórico da produção ou otimização, os quais requerem um número maior de rodadas de simulação.

Na Figura 3.8 observa-se um exemplo do diagrama de tornado usado para ilustrar a Análise da Sensibilidade do volume acumulado de água produzida de um determinado campo. De cima para baixo, as 3 primeiras barras mostram os valores máximos, alvo e mínimo (respectivamente) atingidos durante as rodadas simulações onde diversos parâmetros de entrada foram perturbados para saber a dimensão de s sobre o volume acumulado de água e assim, realizar um ajuste de histórico posterior.

Os parâmetros mais influentes na figura são **DWOC (Contato Óleo-água)**, **NTG (Net to Gross Ratio)** e **RadiusAq\*RadiusAq (Produto do Raio do Aquífero)** em ordem decrescente de influência. O Contato Óleo-Água teve uma influência de  $-8,265 \text{ E}+05$  bbl no volume acumulado de água produzida durante as rodadas do experimento.

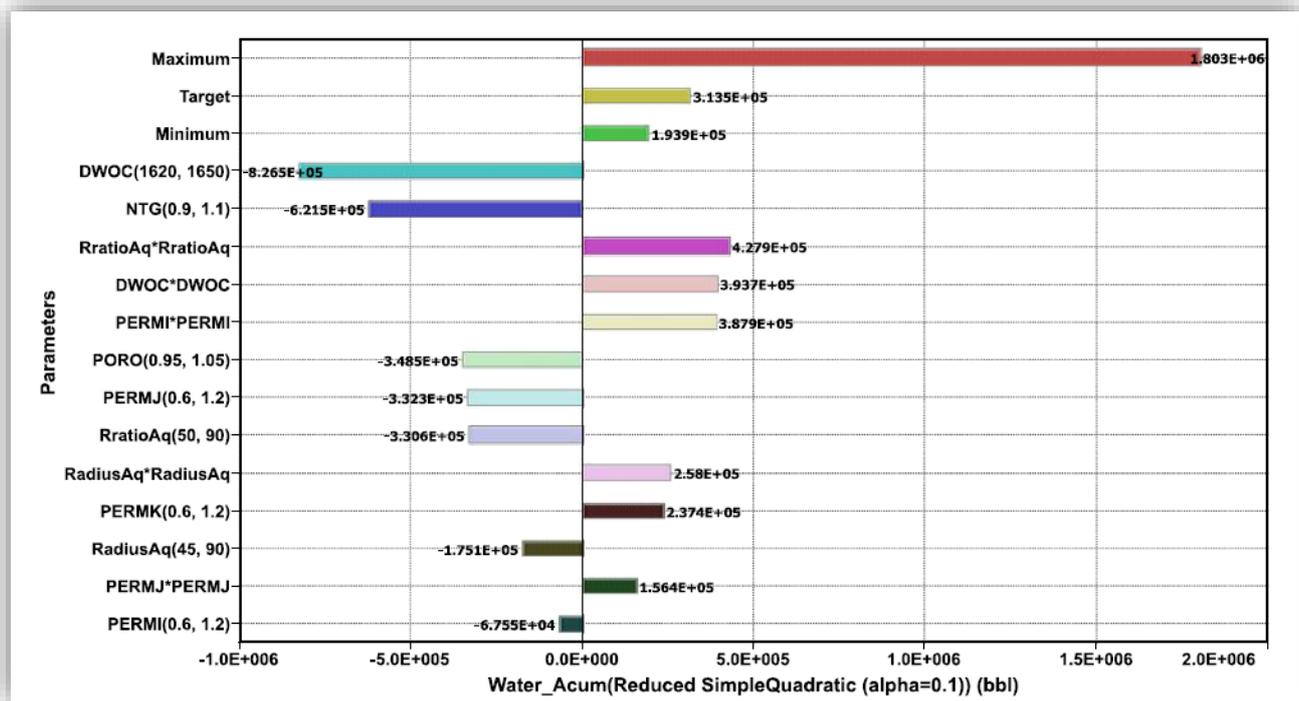


Figura 3.8 – Gráfico de Análise da Sensibilidade gerado pelo CMOST, simulador da CMG mostrando as variações do volume acumulado de água produzida durante as rodadas de experimentos para o ajuste de histórico.

### 3.4. Simulação de Monte Carlo

Como a Análise da Sensibilidade, a Simulação de Monte Carlo é outra técnica de quantificação indireta da incerteza usada em muitas áreas do conhecimento. Tenta considerar incertezas em parâmetros do modelo usando abordagens estatísticas e probabilísticas a fim de cobrir os intervalos de valores que tais parâmetros podem assumir. Cada conjunto de amostragens desses parâmetros que ocorrem em uma rodada de Simulação de Monte Carlo dará origem a uma realização do modelo, conceito que será apresentado no capítulo 5.

Simulação de Monte Carlo é uma técnica que visa emular um processo de amostragem de um dado fenômeno. Portanto, Monte Carlo muitas vezes é referido como “amostrar”, “retirar”, “extrair” de uma distribuição de probabilidade. Quando se está fazendo uma amostra real, e não uma amostragem de Monte Carlo, as amostras são obtidas diretamente do campo, por exemplo, a função de densidade da população é  $f_x(x)$ .

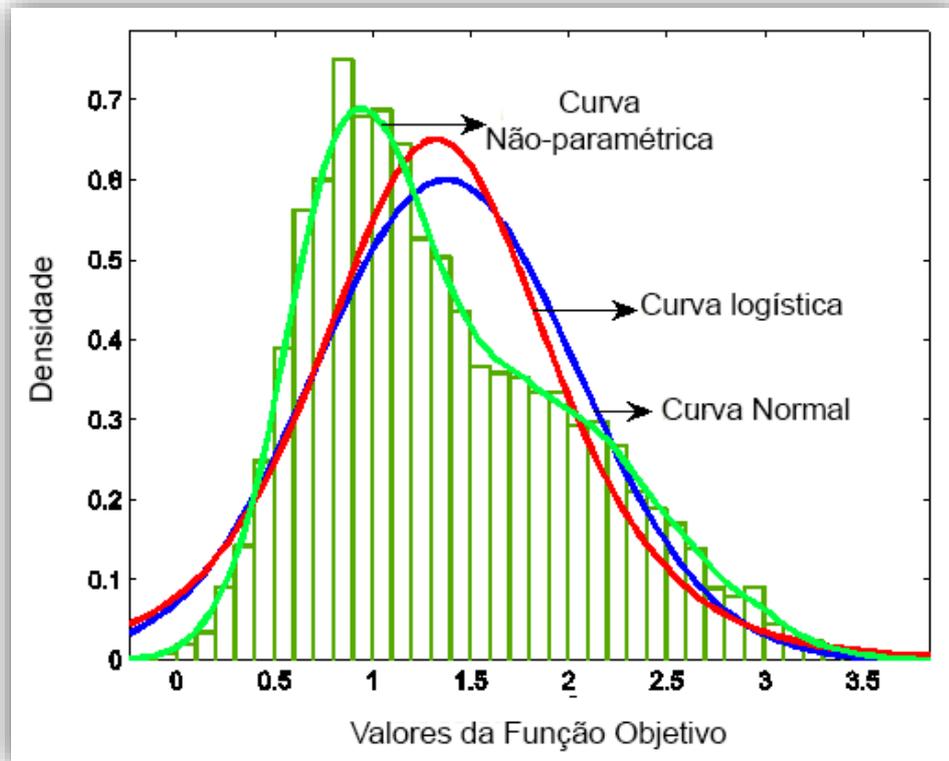
Na Simulação de Monte Carlo, assume-se que a distribuição de probabilidade  $f_x(x)$  é conhecida, e então, usa-se um programa de computador para extrair amostras dessa distribuição, já que no caso de um reservatório de petróleo, é inviável ir a campo e realizar tantas amostragens o quanto se deseja. Para construir um experimento de amostragem, deve-se de alguma maneira ter-se acesso a uma “entidade randômica”, pois deseja-se que a amostra seja justa, ou seja, nenhum valor em particular deva ocorrer mais frequentemente do que o descrito pela função de distribuição de probabilidade considerada. Por exemplo: Como deve-se simular a jogada de uma moeda em um computador, cujo eventos possíveis são perto de 50/50 cara e coroa quando um número grande de tentativas é realizado? Infelizmente, não existe uma máquina aleatória (um computador é uma máquina e ainda determinística) que possa fornecer uma entidade inteiramente aleatória. O que se tem disponível são geradores de números pseudoaleatórios. Um gerador de números pseudoaleatórios é um software que fornece como saída um número aleatório sob demanda. Tal número ou valor, em termos estatísticos, é simplesmente um evento de uma variável  $[0,1]$  aleatória uniforme. Portanto, é um número que sempre estará entre zero e um. Esses números são pseudoaleatórios, pois um gerador de números pseudoaleatórios sempre deve ser iniciado com o que é chamado de “semente aleatória”. Para uma dada semente aleatória, irá sempre se obter uma mesma sequência de números aleatórios. Por exemplo, usando-se MATLAB em um notebook específico, com a semente aleatória de 69071 e o gerador de números aleatórios do MATLAB, obtêm-se a seguinte lista de números aleatórios: 0,10135; 0,5382; 0,98182; 0,0534; 0,48321; 0,65310 e assim por diante. Usando a mesma máquina e o mesmo *software*, obter-se-á sempre a mesma sequência para uma dada semente aleatória.

Agora é possível gerar amostras a partir de uma distribuição especificada  $f_x(x)$ :

1. Sortear um número;
2. Usar a distribuição cumulativa para verificar o seu valor de amostra correspondente;
3. Repetir os passos 1 e 2 tantas vezes forem necessárias quanto o número de amostras.

A Figura 3.9 mostra um exemplo da Simulação de Monte Carlo na quantificação de incertezas em um problema de gerenciamento de reservatórios. Nesta Figura é mostrado um histograma de 10.000 simulações de reservatórios aplicando Monte Carlo sobre uma proxy como substituto de um simulador de reservatórios. Além disso, a Figura mostra o ajustamento

de três curvas de Funções de Densidade de Probabilidade. Através destas curvas é possível estudar o comportamento de uma função objetivo em termos de diferentes parâmetros de entrada do modelo de simulação.



**Figura 3.9 - Um histograma com 10.000 simulações sobre uma proxy, e o ajustamento de três Funções de Densidade de Probabilidade. (CAMARGO SILVA E FILHO, 2012).**

## 4. VISUALIZAÇÃO DE INCERTEZAS

---

Dados científicos não são considerados completos sem suas indicações de erro, acurácia ou níveis de confiança (POTTER, K., 2010). Usualmente, tais informações têm sido representadas em gráficos e tabelas anexas aos conjuntos de dados e não integrados à sua visualização. Dependendo do método visual usado para tentar incorporar incertezas à representação explícita dos dados, a visualização desses pode ficar muito complexa, com muita informação visual, podendo levar a falsas conclusões e previsões. No entanto, as incertezas são uma componente crucial para o entendimento do conjunto de dados e técnicas eficientes devem ser usadas para visualizar tais incertezas de maneira integrada a esse conjunto.

A maioria das técnicas de visualização incorpora incertezas como quantidade de informação desconhecida. Estes métodos criam essa impressão indicando áreas de menos confiança, maior erro, ou alta variação nos dados em sua visualização.

Por outro lado, se na área das disciplinas pertinentes a reservatórios de petróleo alguns fluxos de trabalho que consideram incertezas já foram convencionados, tanto na modelagem geológica quanto na simulação de fluxo, a quantificação de tais incertezas e a posterior visualização ainda estão pouco explícitas.

Compreender o nível de erro ou de confiança associado aos dados é um aspecto importante em sua análise e é frequentemente negligenciado nas visualizações, uma realidade inclusive no âmbito da modelagem e da engenharia de reservatórios. Há, portanto, demanda crescente de trabalho referentes à incorporação desta informação em visualizações (MACEACHREN, et al., 2005) e (PANG, WITERNBRINK, & Lodha, 1996), usando incerteza não só derivada dos dados, mas também presente ao longo de todo o fluxo de processamento da visualização.

Esse capítulo reúne algumas das principais técnicas de visualização de dados, além de artifícios gráficos e visuais para criar a impressão de incerteza nessa visualização. Inclui ainda outras técnicas de visualização de dados que embutem as estatísticas principais de conjunto de dados ou diferentes disposições desses para associar tais incertezas. Algumas das técnicas

apresentadas neste capítulo já foram amplamente usadas na indústria de petróleo e outras têm potencial aplicabilidade.

## 4.1. Visualização de Sumarização de Dados.

A disposição de informações estatísticas é fundamental em todos os campos da visualização uma vez que na literatura já foi referenciado (PANG, WITERNBRINK, & LODHA, 1996) que há incertezas de natureza estatística.

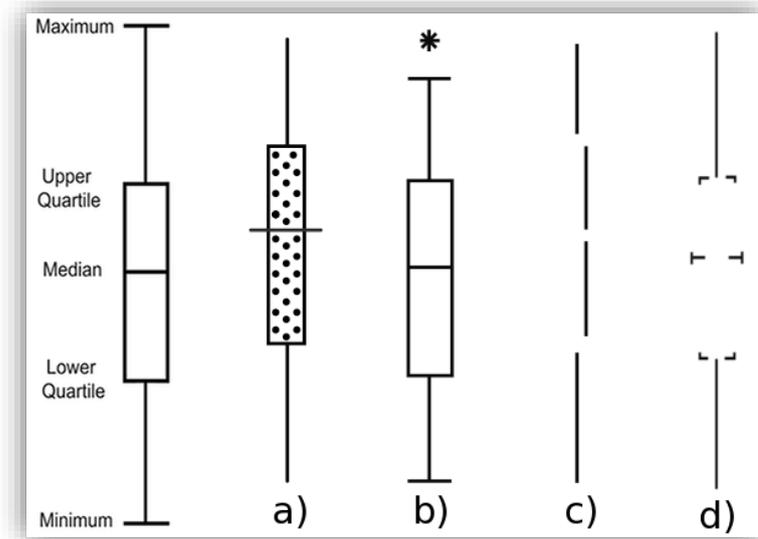
A medida que o conjunto de dados aumenta, mais necessário se torna um método que sumarie as características estatísticas fundamentais de tal conjunto. Tais técnicas devem mostrar informações de maneira simples e intuitiva de modo que se possa chegar a conclusões iniciais e de fundamental importância sobre um conjunto de dados.

Métodos para visualizar estatísticas básicas incluem tabelas, gráficos e plotagem. No entanto, achou-se válido abordar primeiramente o *boxplot* por reunir de forma de fácil visualização e entendimento tais números estatísticos mais importantes do conjunto de dados.

### 4.1.1. Box Plot

Uma das técnicas de visualização de dados estatísticos mais antigas é o *boxplot*. É prático e vantajoso, pois consegue a princípio resumir cinco parâmetros importantes do conjunto de dados em um só desenho. Esses são os valores mínimos e máximos do conjunto de dados, os quartis inferior e superior e a mediana como ilustrados na Figura 4.1. Ao passar do tempo essa técnica foi aprimorada para exibir mais informações dos conjuntos de dados.

Na Figura 4.1, à esquerda, observa-se o conjunto de dados sendo cortado pela mediana, representada por um traço. As bordas da caixa são usadas para demarcar os quartis inferior e superior assim como o intervalo interquartil. A caixa é interceptada por uma barra transversal na mediana do conjunto de dados. No extremo das linhas são observados os valores máximos e mínimos do conjunto de dados. À direita são outras versões de *boxplot*, algumas mais simplificadas, mas todas com interpretação equivalente.

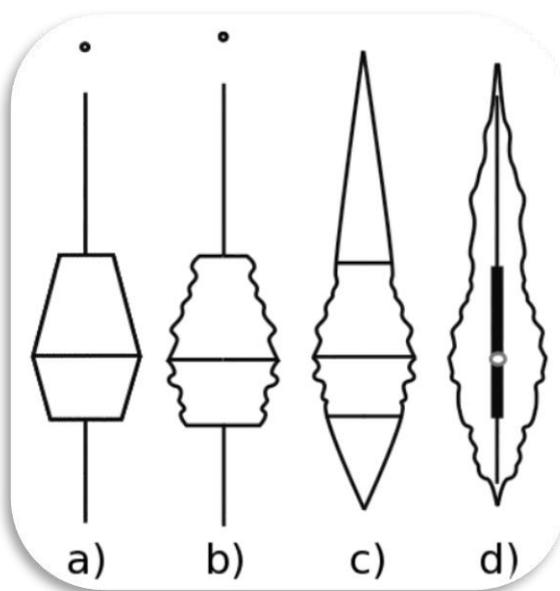


**Figura 4.1 - Boxplots e suas modificações. A esquerda, a construção do boxplot. A direita, modificações visuais. (POTTER, K., 2010)**

Versões posteriores de *boxplot* trazem a função de expressar aspectos da densidade da distribuição dos dados, que antes eram negligenciados.

O *Histplot*, mostrado na Figura 4.2(a), é uma aproximação simples para adicionar densidade a um *boxplot*. Aqui a densidade da distribuição é estimada apenas na região interquartil e mesmo assim, de forma linear, ainda dando margem para negligência de informações. Quanto mais largo for o *boxplot*, maior a densidade dos dados naquele determinado valor. O *vaseplot* da Figura 4.2(b), já representa de forma mais refinada a densidade, mas como o *histplot*, apenas na área central interquartil.

O *box-plot* de percentil, mostrada na Figura 4.2(c), é um outro método para adicionar o distribuição cumulativa empírica do conjunto de dados para o *boxplot*. Neste tipo de plotagem, assim como no *vaseplot*, a região interquartil possui a densidade da distribuição de dados bem refinada, o diferencial está nos percentis abaixo do 25° e acima do 75° a densidade da distribuição passa a ser aproximada de modo semelhante ao *histplot*.



**Figura 4.2 - Modificações de densidade no boxplot. a) Histplot. b) Vaseplot. c) Box-percentile plot. d) Violin Plot (POTTER, K., 2010)**

O *violin-plot* na Figura 4.2(d) é a modificação do *boxplot* que apresenta a maior riqueza de detalhes sobre a distribuição dos conjuntos de dados das quatro modificações apresentadas até agora. Não só a região interquartil central possui uma curva de densidade refinada, mas também a região periférica. Observa-se que agora a região central interquartil é representada por uma caixa preta e a mediana por um círculo branco em vez de um traço.

Outras formas de *boxplot* podem ser encontradas na bibliografia (POTTER, K., 2010) que são variações dessas noções básicas de *boxplot* cujas aplicações não são muito relevantes ao foco deste trabalho.

## 4.2. Outras Técnicas de Visualização

### 4.2.1. Gráficos de Coordenadas paralelas

O Gráfico de coordenadas paralelas é um método de visualização indireto de incertezas pois é uma ferramenta usada para dispor dados multidimensionais que, no entanto, são atemporais. Cada parâmetro (ou categoria de dados) é representado em um eixo vertical. O objetivo é analisar diversos objetos (no nosso caso, realizações de um modelo geológico) em relação a parâmetros de importância para a tomada de decisão em questão. Nesse gráfico cada objeto de análise é representado por uma linha que liga um valor de um eixo a outro valor de

um eixo ao lado. A linha “quebrada” ligando os valores de cada um dos eixos caracteriza tal objeto. É útil quando se quer observar determinados parâmetros e um conjunto de realizações de algum modelo de reservatório ou qualquer outro modelo estocástico quando se usa Simulação de Monte Carlo para gerar tais realizações. Na Figura 4.3 tem-se um gráfico programado em Matlab com quatro eixos paralelos cada um dispondo uma componente, ou variável do conjunto de resultados. Cada linha que liga os quatro eixos é uma realização ou resultado de simulação. Esse método não explicita números estatísticos sobre o conjunto de dados, apenas se visualiza graficamente a densidade da distribuição das curvas e o conjunto de valores das variáveis de interesse que cada realização assume.

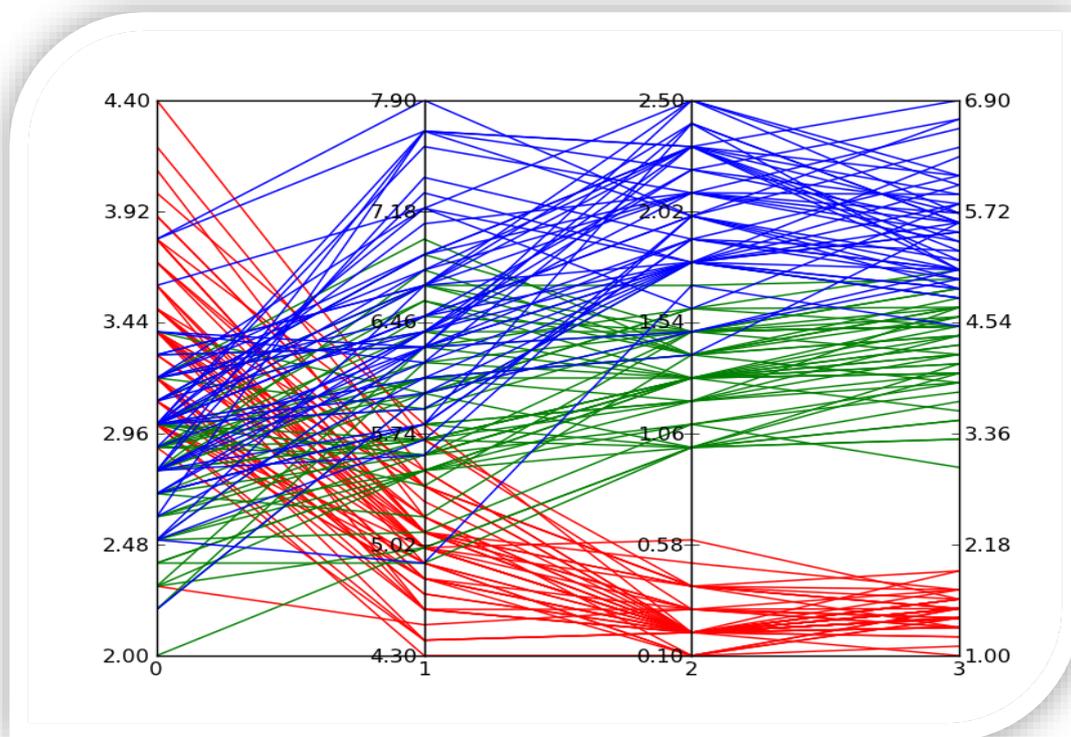


Figura 4.3 - Gráfico de Coordenadas Paralelas. Fonte: [Stackoverflow](#) (Hiperlink).

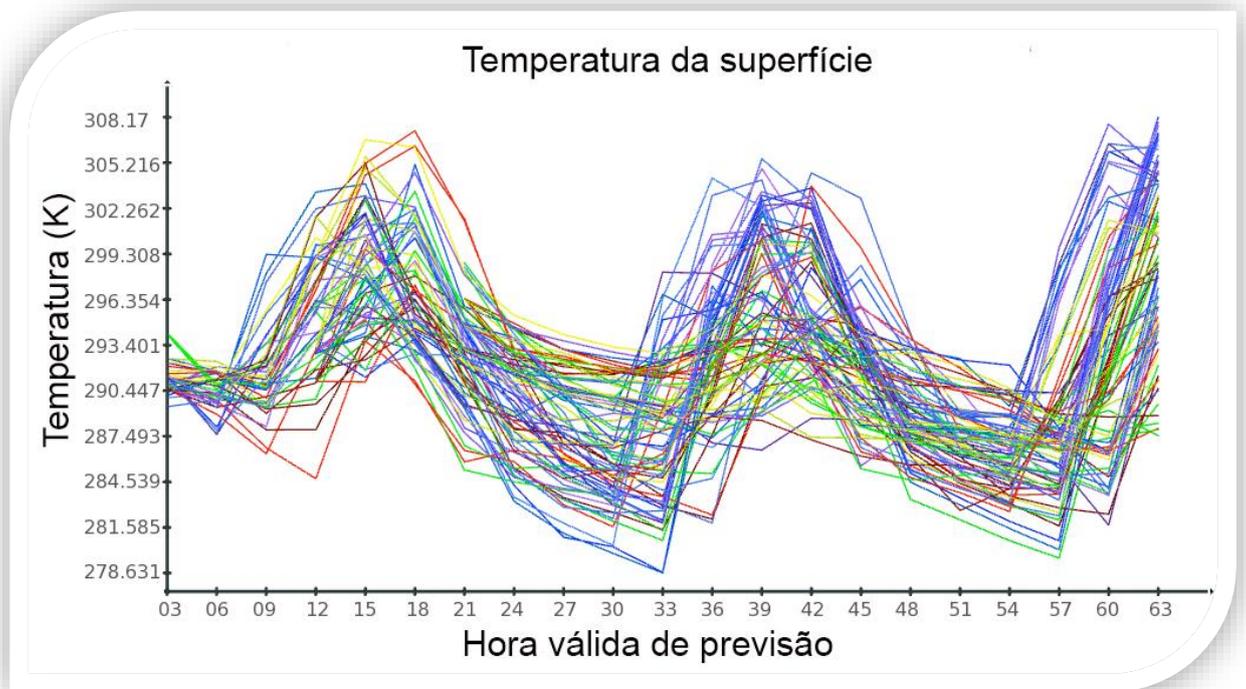
#### 4.2.2. Gráficos *Spagetti* e *Fan Charts*

O Gráfico *Spagetti* também é um método de visualização de conjunto de dados, que envolve incertezas, como os Gráficos de Coordenadas paralelas, mas ao contrário desse, o *spagetti* não exhibe dados multidimensionais e no caso da

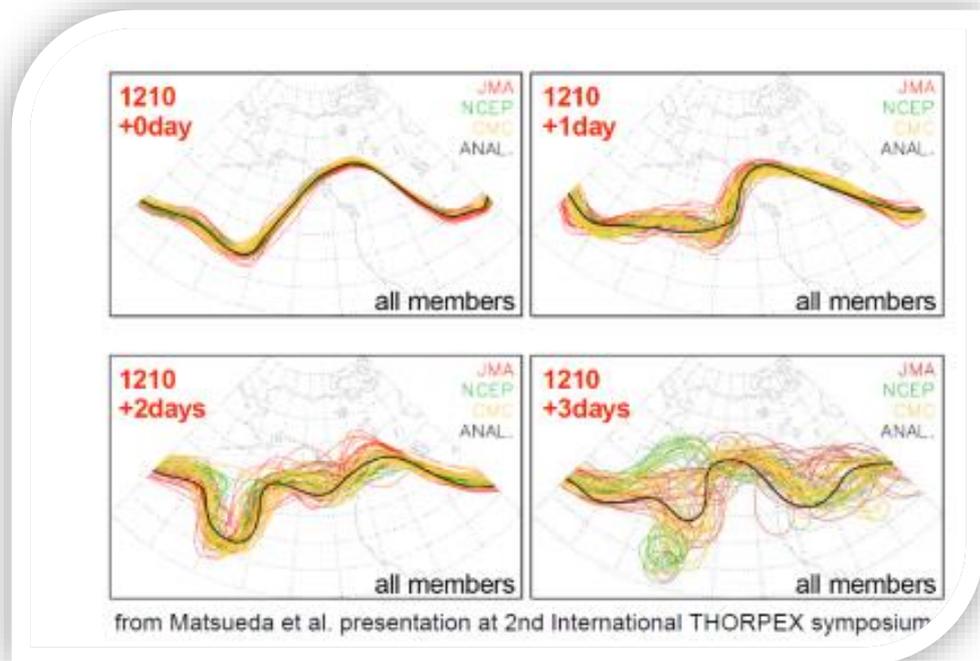
Figura 4.4, pode-se observar um exemplo de conjunto de dados que variam no tempo.

Nas

Figura 4.4 e Figura 4.5 podemos ver gráficos *spagetti*. Os dois são da área da meteorologia onde observa-se várias simulações da temperatura para diferentes horários de um determinado dia. Tais simulações são resultado da perturbação dos parâmetros de entrada do modelo através de Monte Carlo. Na Figura 4.5 tem-se outro gráfico *spaghetti* no qual são analisados resultados de uma simulação meteorológica, mas em um contexto espacial (2D) com a dimensão do tempo expressa na justaposição, lado a lado, das outras previsões dos dias consecutivos. A Figura 4.5 é, então, uma técnica de visualização 3D na qual os dados se dispõem de maneira bidimensional onde o tempo é a terceira dimensão expressa na posição dos quadros.

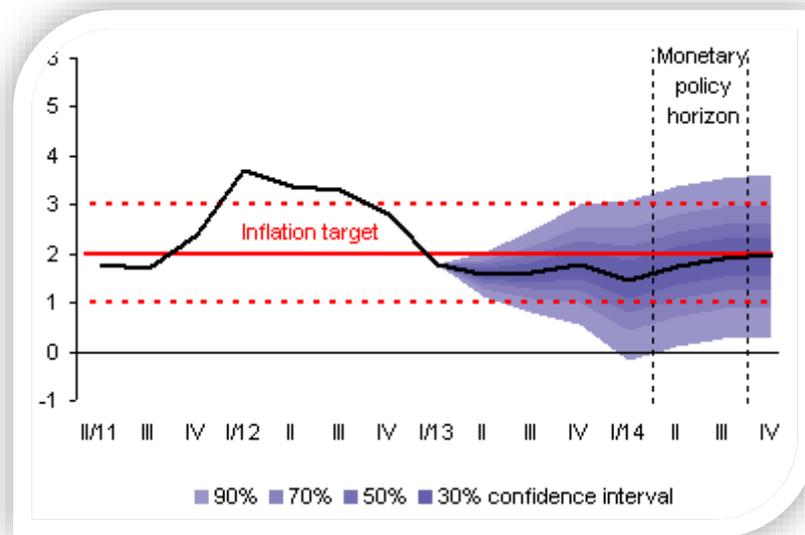


**Figura 4.4 – Gráfico Spagetti. Uma aplicação na área da meteorologia.**



**Figura 4.5 - Gráficos tipo espaguete (spaghetti plots), sobrepondo resultados de diversos membros do conjunto de simulações em uma mesma imagem, para uma aplicação na área de meteorologia.**

A Figura 4.6 apresenta um gráfico do tipo *Fan-Chart* onde são exibidas várias simulações da previsão do comportamento da inflação. O que caracteriza o *Fan-Chart* é a tentativa de captura das incertezas construindo intervalos de confiança por meio das bandas de cores diferentes geradas pela densidade das curvas de simulações das previsões da inflação a partir de dados observados previamente. A faixa mais escura é o intervalo de 30% de confiança. Conforme se considera faixas mais largas, mais chances se tem de o comportamento dos valores da inflação ocorrer ali.



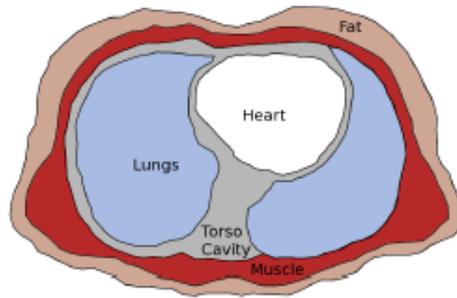
**Figura 4.6 – Exemplo de visualização da incerteza usando diagrama tipo *fan-chart*. Fonte: [https://www.cnb.cz/en/monetary\\_policy/forecast/previous\\_forecasts/prognoza\\_1305.html](https://www.cnb.cz/en/monetary_policy/forecast/previous_forecasts/prognoza_1305.html)**

### 4.2.3. Mapas de Cores

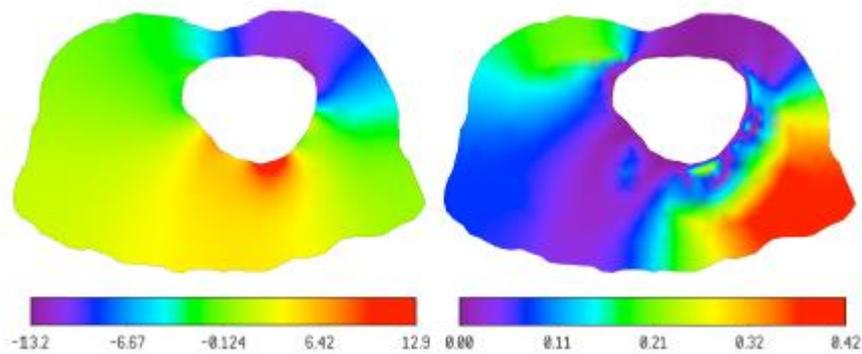
Mapas de cores é uma das técnicas mais usuais para representação de conjunto de dados 2D e 3D em diversas áreas (POTTER, K., 2010). Sua simplicidade possibilita a percepção e diferenciação imediata de baixos e altos valores, e no âmbito da incerteza, já foi usado para analisar tais conjuntos de dados ilustrando a média e observando o desvio-padrão, associando-o à uma medida de incerteza.

A Figura 4.7 mostra o corte latitudinal de um tronco humano que foi discretizado para se obter mapas da condutividade elétrica.

Foram realizadas 10.000 realizações perturbando a condutividade de entrada para observar a sensibilidade do mapa de tensões através do tronco. Essa grande quantidade de mapas 2D é impossível de ser analisada um a um a olho nu (ou por justaposição dos mapas, como será visto na seção 4.2.4), o ideal é ter algum método de visualização que sumarie as informações mais importantes pois cada célula terá um conjunto de 10.000 outros valores associados às realizações. Uma opção usual é usar tais amostras para contruir mapas de desvio padrão e média. (Figura 4.8).



**Figura 4.7 – Corte latitudinal de um tronco humano. (POTTER, K., 2010).**



**Figura 4.8 - Média para variações de 50% (para mais ou para menos) da condutividade dos pulmões (mapa da esquerda). Desvio padrão para 50%(para mais ou para menos) da condutividade dos pulmões. (POTTER, K., 2010).**

O esquema arco-íris de cores nem sempre é adequado. Alguns conjuntos de dados exibem comportamentos específicos, estão relacionados localmente, sofrem progressões de valores, crescem ou diminuem gradativamente em uma ou mais direções específicas e a escala arco-íris não dá tal ideia de progressão evidente. Para tais finalidades, são propostas outras escalas de cores na Figura 4.9.

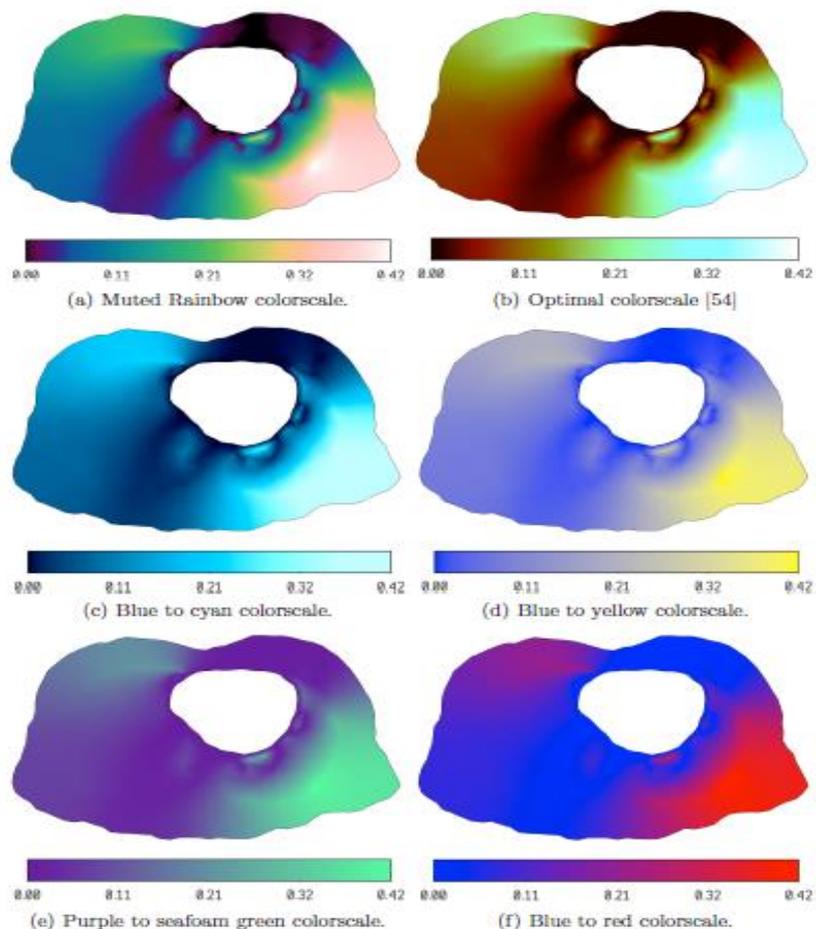
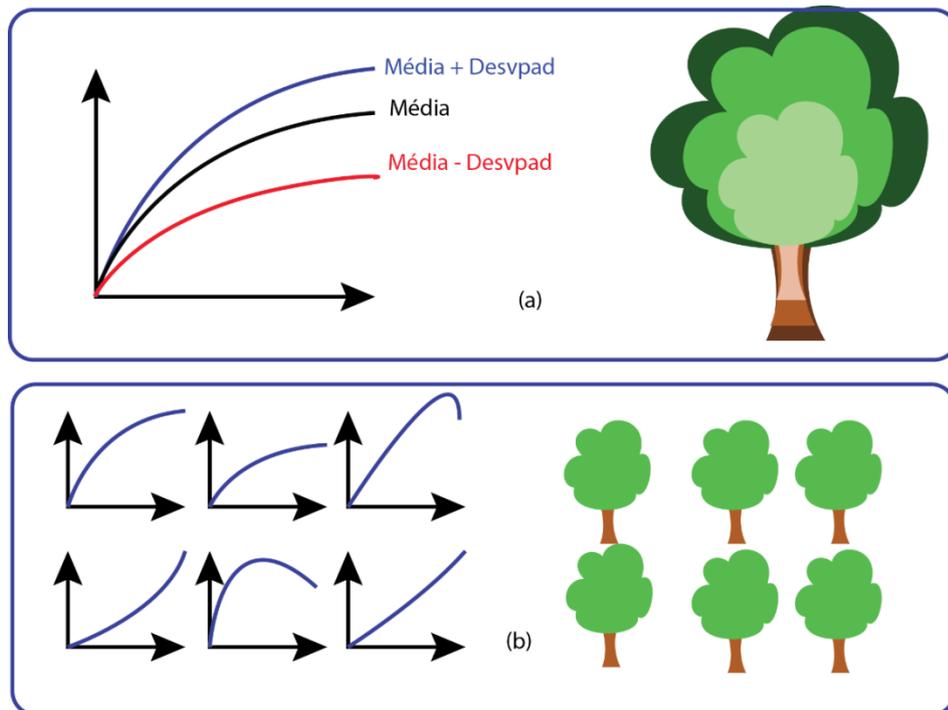


Figura 4.9 - Diferentes escalas de cores para mapas de cores. (POTTER, K., 2010)

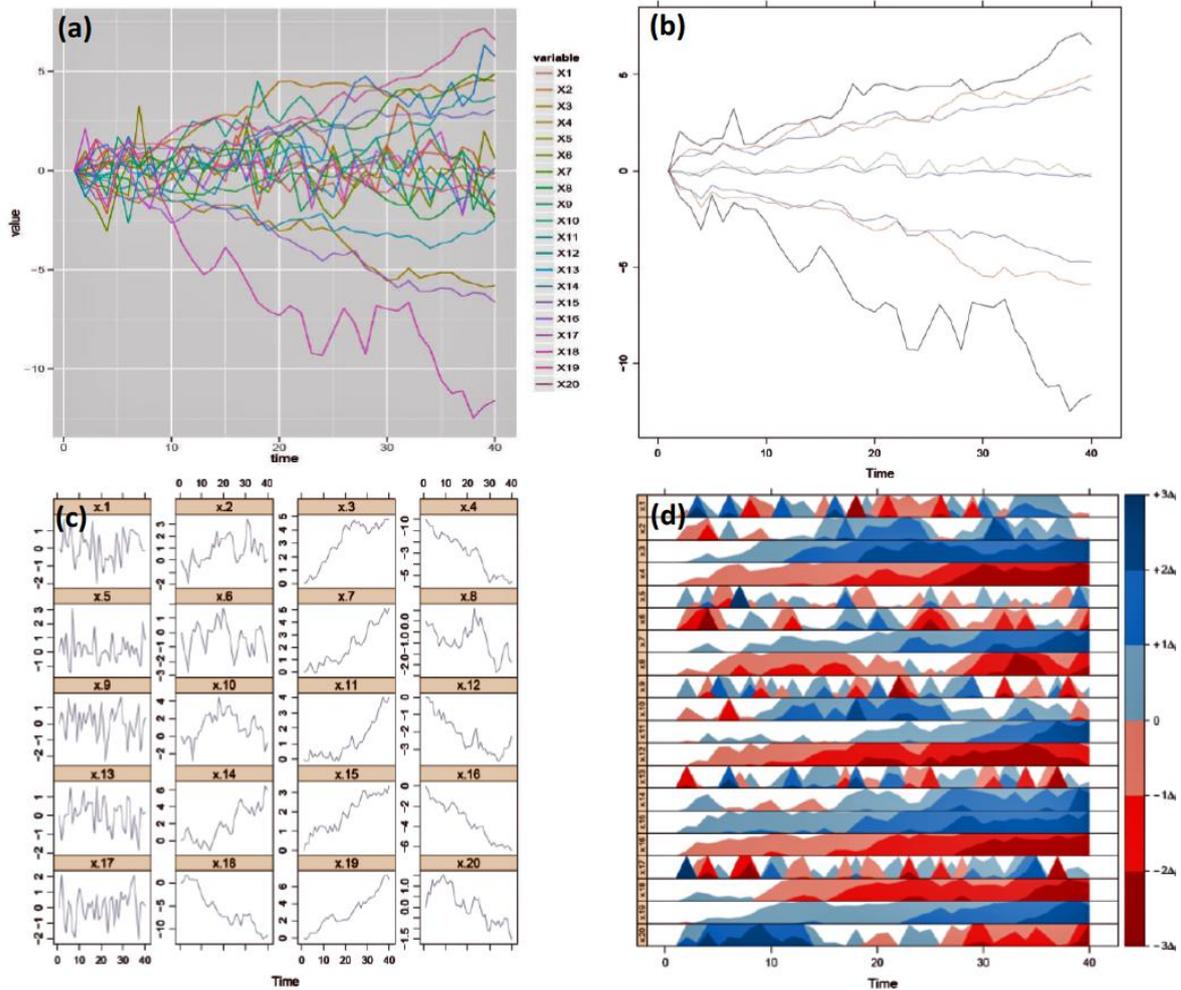
#### 4.2.4. Justaposição e Superposição.

Outros métodos de visualização de conjuntos são os de justaposição e superposição. Nas Figura 4.10 (a) e (b) estão exibidos tais métodos. A Figura (a) é o método de superposição ou visualização condensada. Nesse, várias informações – úteis e representativas do conjunto de dados – devem ser exibidas em um mesmo gráfico, de modo superposto. A superposição pode ser feita apresentando apenas curvas que representam as características estatísticas principais do conjunto de dados, ou pode-se condensar todas as curvas geradas conforme a conveniência. A Figura (b) trata do método justaposto ou visualização lado a lado.



**Figura 4.10 - Classificação esquemática de duas abordagens para visualização de conjuntos: (a) visualização condensada; e (b) visualização simultânea.**

Na Figura 4.11 é mostrada a aplicação de tais técnicas em um conjunto de 20 séries temporais: (a) e (b) são formas de visualização por superposição, mostrando, respectivamente, todas as curvas juntas ou somente estatísticas principais (máximo, mínimo, etc); em (c) e (d) são formas de visualização simultânea (justaposição), mostrando todas as curvas separadas, em espaços gráficos diferentes, dispostas, respectivamente, em um arranjo matricial ou empilhado. Essa última forma de visualização é conhecida como *horizon plot*, com as faixas azuis e vermelhas indicando respectivamente valores acima e abaixo da média, em tons progressivamente mais fortes quando os valores excedem uma ou duas vezes a escala vertical, o que permite compactar a informação de cada curva individual em uma faixa bastante estreita. Cada curva exibida em (c) está representada em (d). A justaposição vertical das curvas possibilita analisar quanto que o valor de cada uma dessas, em cada tempo, se distancia da média para aquele determinado tempo conforme a escala de cores exibida à direita da (d).

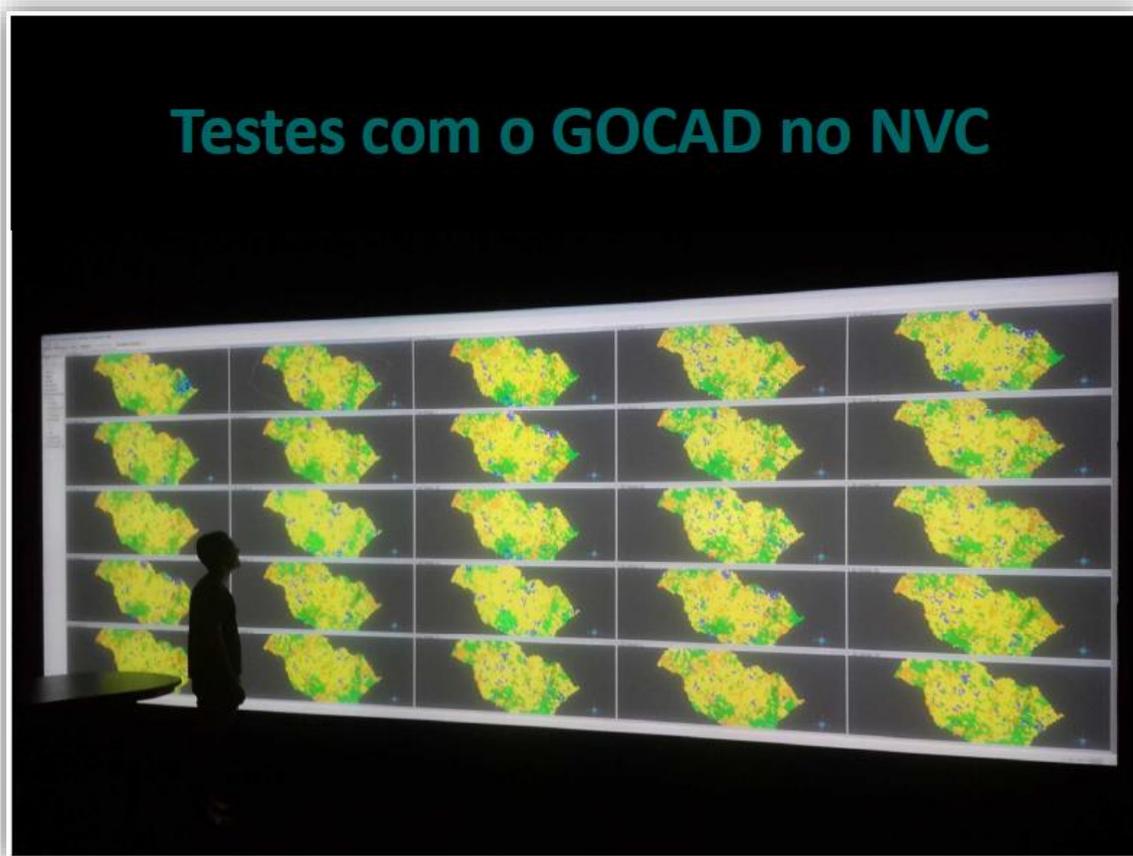


**Figura 4.11 - Diversas formas de visualização de um mesmo conjunto de 20 séries temporais (LAMIGUEIRO, 2014).**

A Figura 4.12 mostra a técnica de justaposição aplicada à visualização de múltiplas realizações de um modelo geológico base real na sala UHD do CENPES. A conclusão a que se chegou no experimento é que tais técnicas de justaposição e superposição devem ser usadas em conjunto ainda mais em casos de dados tridimensionais ou de dimensões maiores, que é o caso quando as mudanças de certas propriedades são analisadas inclusive no tempo. Há a análise 4D que é quando cada janela exibe um estado do mapa de determinada propriedade em um momento específico no tempo. Em uma análise 5D, cada janela representaria uma realização e a dimensão do tempo seria embutida através da animação das janelas mostrando a evolução dos mapas de determinada propriedade. No caso 4D têm-se as três dimensões espaciais e o tempo expresso pela posição de cada janela. No caso 5D, a dimensão extra é o número  $n$  de

realizações representado em cada janela, onde a outra dimensão, tempo, será embutida através das animações temporais dos mapas.

Há um limite da percepção visual humana para analisar conjunto de dados. Dependendo da quantidade de informação, detalhes importantes podem passar despercebidos ou serem vistos como redundantes, daí a necessidade do uso de métodos que integrem em uma mesma visualização as estatísticas principais de um conjunto de dados.

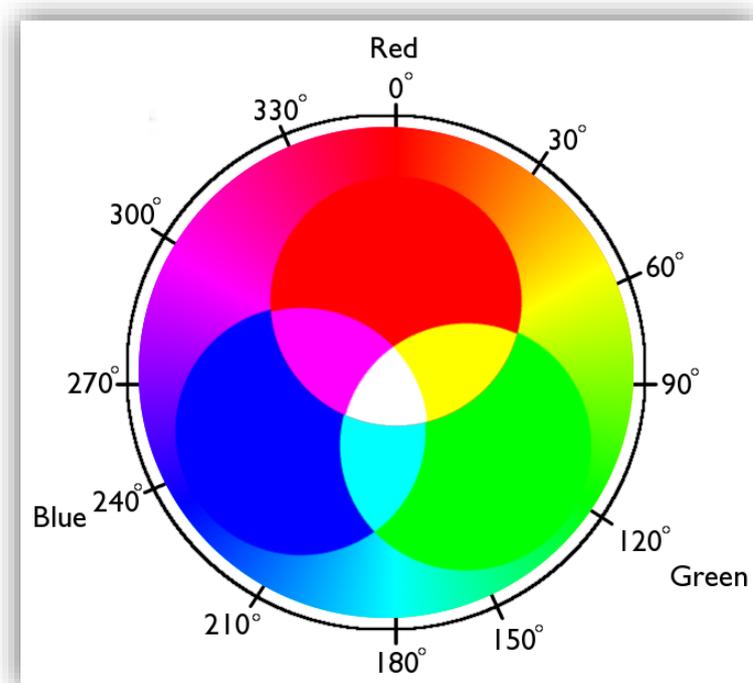


**Figura 4.12 – Aplicação em reservatórios. Realizações de um mesmo modelo geológico justapostas para visualização na sala UHD (Ultra High Definition) no CENPES da Petrobras (Um caso 5D).**

#### **4.2.5. Cone HSV e mapas bivariados**

Mapas de cores bivariadas servem tanto para visualização de dados 2D ou 3D, mas o espaço das cores é um espaço 3D. O sistema de cores RGB (Vermelho, verde e azul) embora seja um dos mais usuais e de mais fácil percepção das cores, se mostra de difícil navegação e

não práticos à associação intuitiva dessas (POTTER, 2010). Outra alternativa é o sistema CMY (ciano, magenta e amarelo) que também pode ser um espaço de difícil navegação. A relação entre os dois sistemas é mostrada na Figura 4.13. A interseção entre as cores azul e vermelho originam os tons de magenta, interseção entre azul e verde originam os tons de ciano e entre as cores verde e vermelho, os tons de amarelo.



**Figura 4.13 - RGB, CMY e Azul. (POTTER, K., 2010).**

O surgimento do espaço HSV (Matiz, saturação e valor) trouxe um sistema de cores mais intuitivo e de fácil navegação (POTTER, K., 2010). Atentando-se para a Figura 4.14 pode-se ver a relação entre essas 3 “dimensões” descritas em um cone, o cone HSV. A saturação é definida como nível de branco diluído em uma cor e que varia na direção radial do cone. A matiz varia de acordo com o deslocamento angular ao redor do eixo longitudinal (vertical) do cone, e o valor é o nível de cinza diluído na cor e que varia na direção do eixo vertical.

O uso de mapas de cores é extenso para visualização de mapas 2D e 3D de propriedades de modelos geológico ou de simulação, não trazendo assim, muita inovação. Já o uso de mapas bivariados não se mostra tão usual mas certamente possui grande potencial para visualização de dados conjuntamente com suas incertezas.

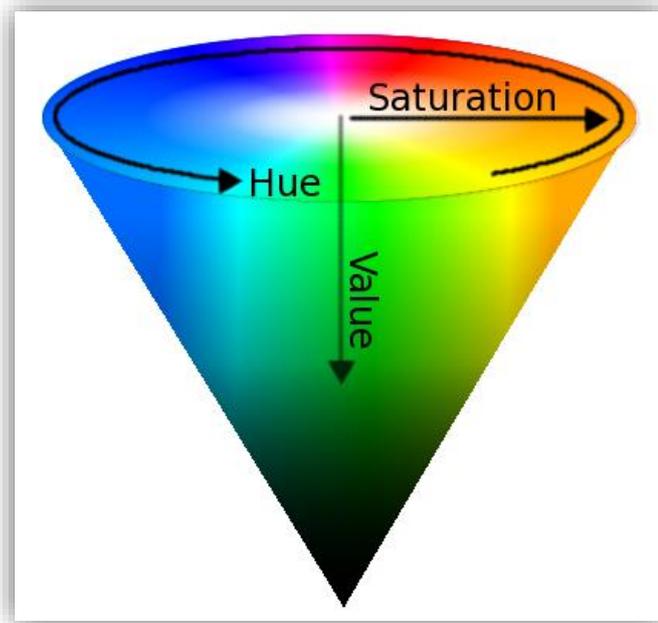


Figura 4.14 - Cone HSV (Matiz, Saturação, Valor). (POTTER, K., 2010).

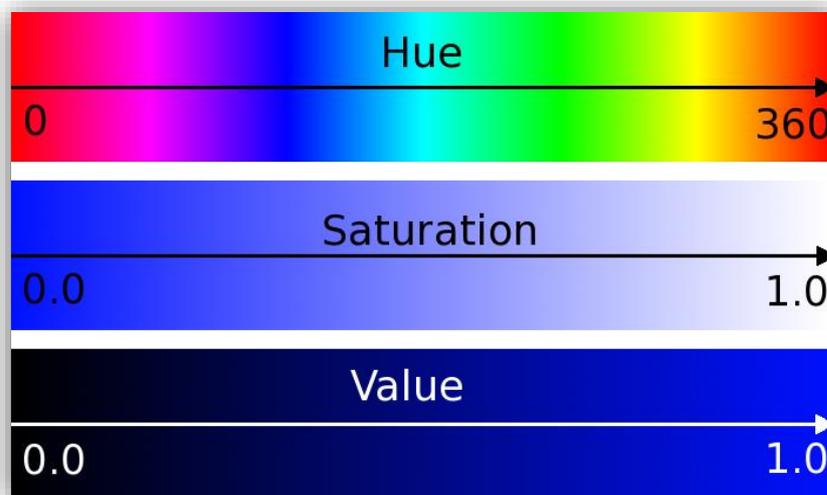
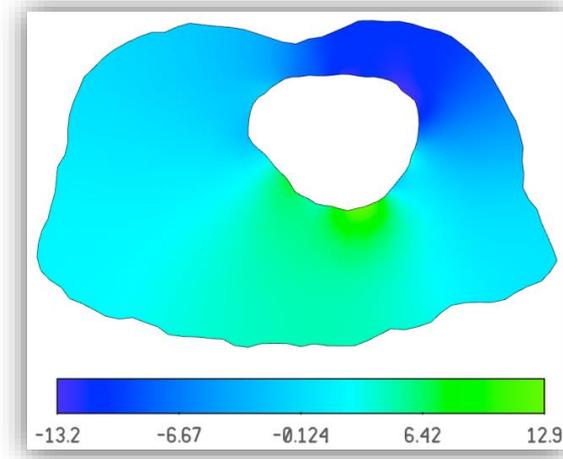


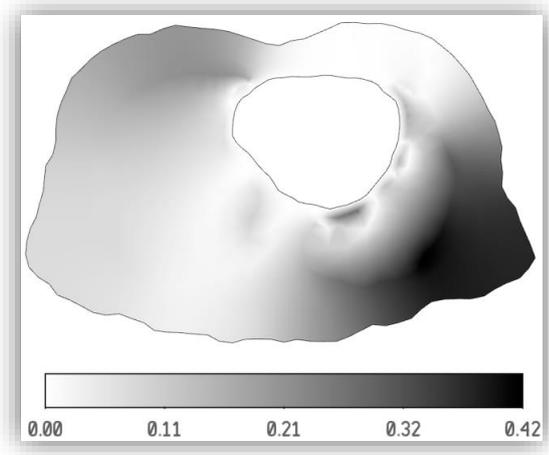
Figura 4.15 - As 3 dimensões do HSV. (POTTER, K., 2010).

Nas Figuras Figura 4.16 a Figura 4.18 observa-se uma aplicação das escalas do cone HSV. A média foi representada pelo mapa de cores na escala azul-verde e o desvio padrão foi representado pela escala de valor (*value*) onde valores baixos denotam desvios padrões baixos evoluindo para o preto, valores maiores.

ser a associada à uma medida de incerteza. Valores maiores de desvio padrão levam a valores de mais adição de cinza na escala da Figura 4.17, chegando ao preto. Associar a ideia de incerteza ao visual de áreas escuras se torna intuitiva e de fácil entendimento, uma vez que o desconhecido é facilmente remetido àquilo que não se vê com exatidão ou com nitidez.



**Figura 4.16 – Média usando escala de cores. (POTTER, K., 2010)**



**Figura 4.17 – Desvio padrão usando escala cinza ou valor (quantidade de cinza). (POTTER, K., 2010)**

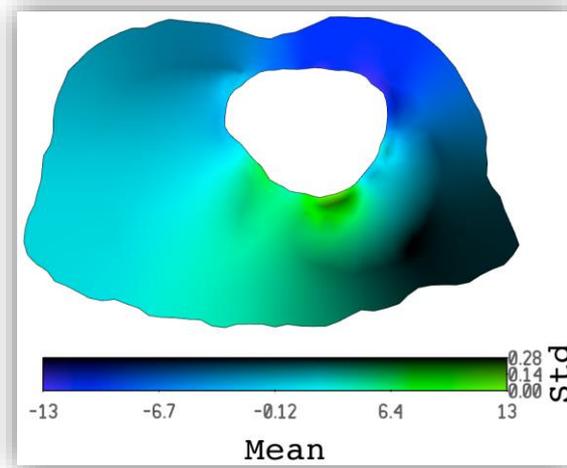


Figura 4.18 – Mapa bivariado da média com o desvio padrão.

A Figura 4.19 mostra outra aplicação de mapa bivariado, na área da geologia, mas dessa vez esse está bivariado na saturação das cores, ou seja, a saturação é usada para associar incerteza à visualização dos dados. A Figura 4.19(a) mostra a espessura em centímetros estimada do solo superficial de uma determinada região. No qual a Figura 4.19(b) mostra que a incerteza relativa é igual ou superior a 80 % nas áreas mais esbranquiçadas. E quanto menos saturada (ou mais esbranquiçada) uma área estiver, mais os dados naquela região serão incertos. A legenda mostra os valores da incerteza em faixas distintas de espessura do solo.

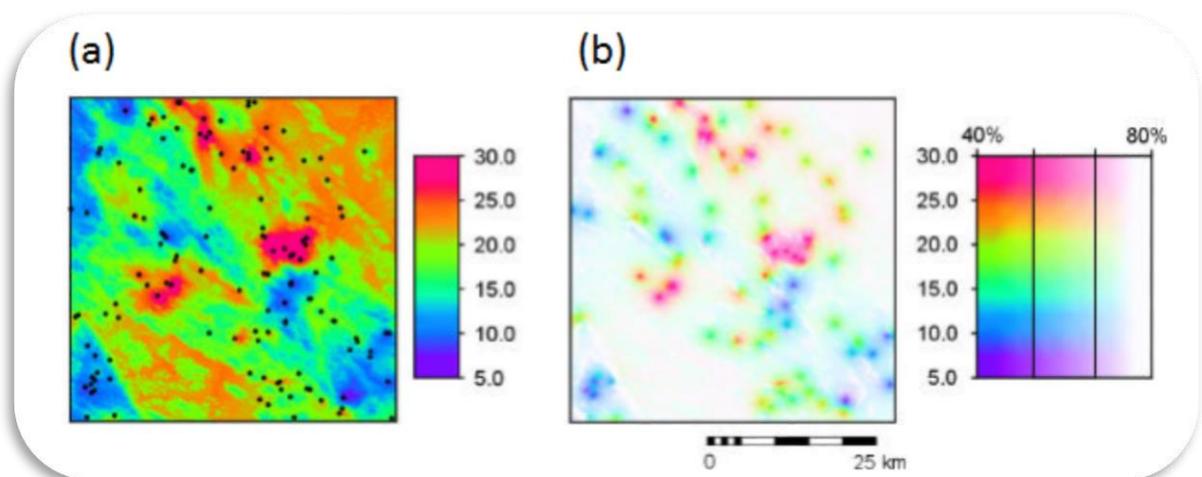


Figura 4.19 – Uma aplicação, em modelos geológicos. Mapa bivariado na saturação ao invés do valor.

#### 4.2.6. Mapas de deslocamento

Nessa aplicação ilustrada na Figura 4.20 o desvio padrão célula a célula das 10.000 realizações mencionadas na seção 4.2.3 foi representado pelo mapa de cores na direção do plano XY. A média do conjunto de dados está representada no eixo Z, diretamente proporcional à altura de um determinado ponto XY da superfície. Ou seja, a cada ponto da superfície tem-se uma cor associada a um valor de desvio padrão em uma escala de cores e, ao mesmo tempo, esse ponto estará deslocado ao longo do eixo Z, determinando, assim, o valor da média, naquele ponto do conjunto de dados.

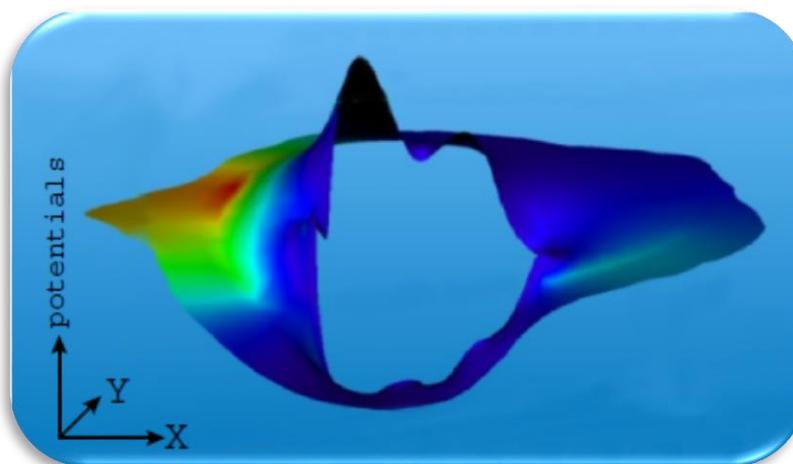
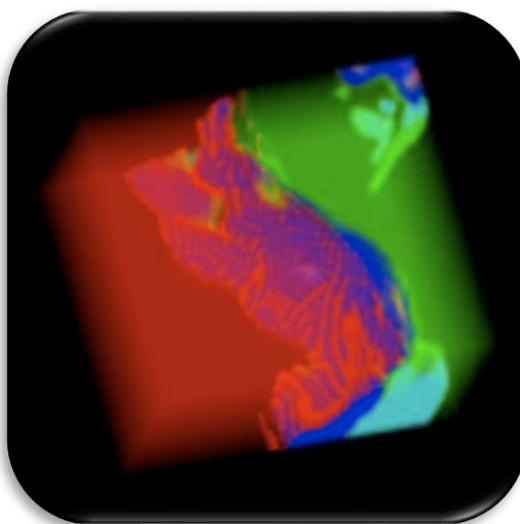


Figura 4.20 – Nessa figura o desvio padrão é representado pelo mapa de cores e a média pela altura do deslocamento da superfície. (POTTER, K., 2010).

#### 4.2.7. Renderizando volumes e Mapas vetoriais. (3D)

Essa é uma seção de grande importância nesse trabalho pois embora as técnicas que serão abordadas a seguir tenham sido usadas em outras áreas do conhecimento, não foram ainda implementadas em *softwares* de modelagem geológica e de visualização de resultados de simulação de fluxo como Petrel, IMEX e visualizadores da CMG, sendo incluído somente no *software* GOCAD o recurso de mapa bivariado na opacidade através da função de transferência.

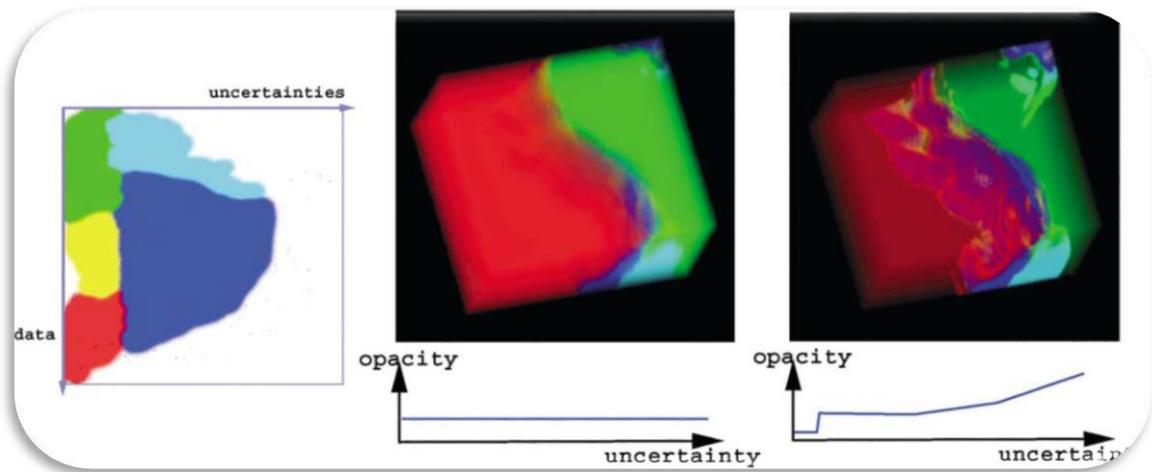
Na Figura 4.21 observa-se a renderização de um mapa 3D da salinidade cuja as áreas de opacas remetem a dados de maior valor de incerteza. Nesse caso, a associação da incerteza a uma visualização mais opaca dos dados é mais conveniente pois consegue-se enxergar as regiões de dados com mais incerteza através do volume e visualizá-las melhor. Em cada caso em específico, diferentes funções de transferência devem ser testadas, pois a qualidade da visualização é função das características do conjunto de dados assim como a definição de incerteza (desvio padrão, entropias, etc) a ser usada no modelo.



**Figura 4.21 – Renderização volumétrica da salinidade do oceano e sua incerteza. Cores representam valores de salinidade e regiões de maior opacidade representam maior incerteza. (DJURCILOV, KIM, LERMUSIAUX, & PANG, 2002).**

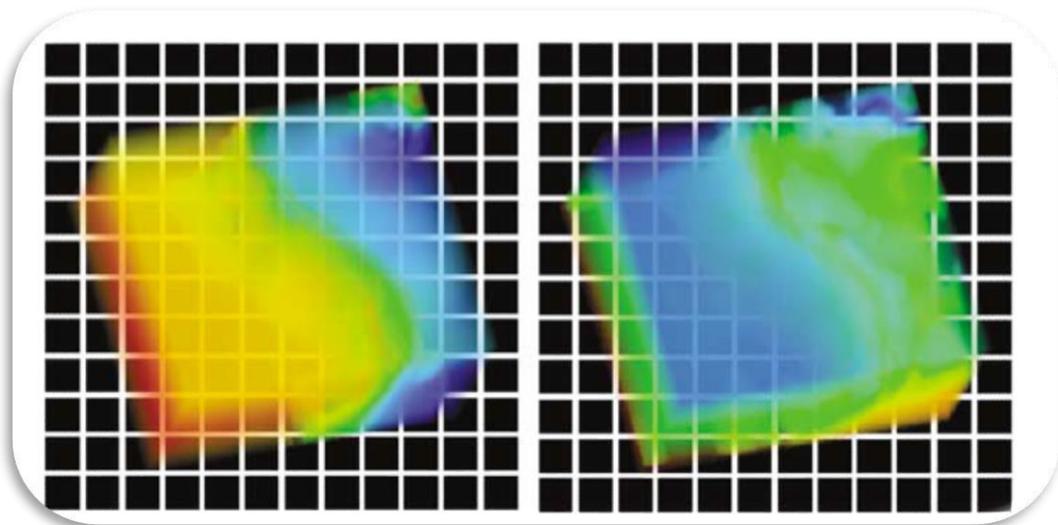
Na Figura 4.22 é exibido um exemplo de função de transferência 2D (DJURCILOV, KIM, LERMUSIAUX, & PANG, 2002). Pode-se observar pelo gráfico à esquerda da figura (gráfico de função de transferência 2D) dois eixos, um relacionado aos valores dos dados e outro a valores de incerteza. Os dados que se dispuserem mais a direita desse gráfico apresentam maiores valores de incerteza (cores azul escuro e azul claro). Se for usado um mapa bivariado de opacidade constante, o resultado da visualização é o mapa do meio da Figura 4.22 onde não se observa visualmente as diferenças de opacidade. Mas se, por outro lado, for usado um mapa bivariado que associe altos valores de opacidade a altos valores de incerteza e o contrário para baixos valores dessa, o resultado é o mapa de salinidade mais a direita na Figura 4.22, onde os dados em azul escuro, que possuem altos valores de incerteza, aparecem bem destacados na renderização do volume. Para mais exemplos de funções transferências assim

como outros métodos de renderização de incertezas em volumes, pode-se acessar o trabalho mencionado anteriormente. Todos os métodos podem ser válidos na visualização de mapas geológicos dependendo das características do conjunto de dados a ser estudado.



**Figura 4.22 - Função de Transferência 2D e valores de opacidade. (DJURCILOV, KIM, LERMUSIAUX, & PANG, 2002).**

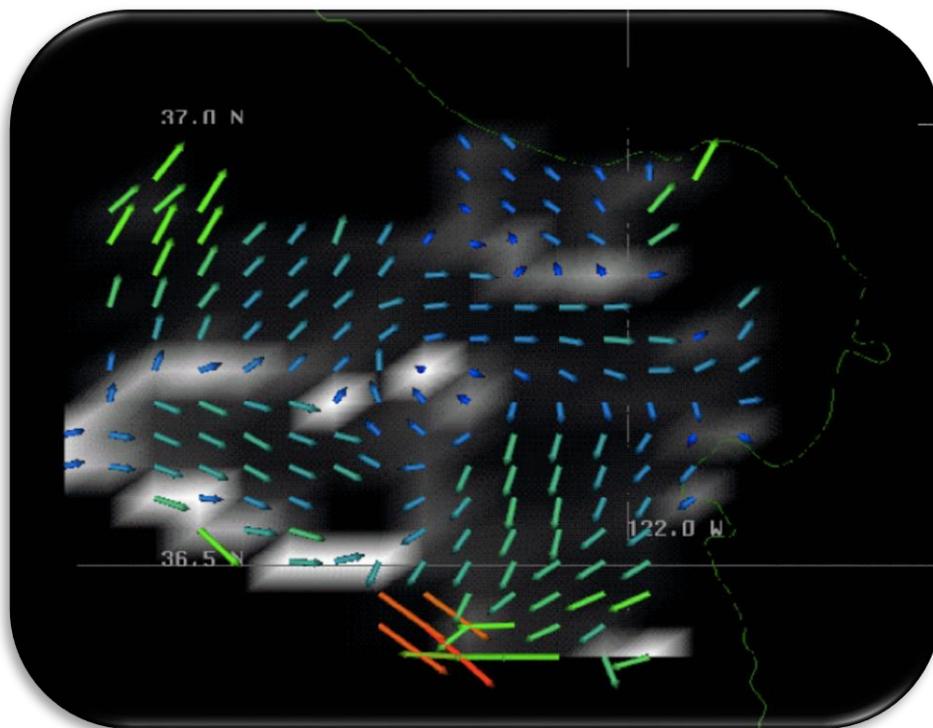
Na Figura 4.23 foi usado malhas brancas para realçar o grau de transparência seguindo a mesma relação com a incerteza da Figura 4.22.



**Figura 4.23 – Uso de malha branca para realçar o grau de transparência do volume. (DJURCILOV, KIM, LERMUSIAUX, & PANG, 2002)**

Os glifos (Figura 4.24) são muito úteis para visualizar dados vetoriais. Podem agregar diversas informações através dos parâmetros cor, orientação, localização e tamanho. A cor, orientação e localização podem ser usados para expressar as principais características do conjunto de dados enquanto que a distorção de sua forma é usada para representar as incertezas. Um meio de distorção é a nitidez que pode ser inversamente proporcional aos níveis de incerteza. Quanto maior a incerteza, menos nítido se torna o glifo, indicando assim, as regiões de alta incerteza por estarem pouco nítidas.

Por nem todos os dados serem visualizados de forma eficaz usando glifos, a adição de glifos para transmitir apenas informações de incerteza é muitas vezes preferível.



**Figura 4.24 – Visualização do fluxo de correntes. As médias e desvios-padrão de vento e as correntes oceânicas são mostrados usando glifos. (WITTENBRINK, C. M.; PANG, A. T.; LODHA, S. K., 1996).**

## 5. ANÁLISE DE INCERTEZAS

---

Após ver algumas das diversas técnicas que podem ser empregadas tanto para quantificar quanto para visualizar incertezas, é preciso saber como analisar tais informações e processar os dados de saída dos fluxos de trabalho, para assim, dispor de informações ainda mais qualificadas que irão auxiliar a tomada decisão relativa a algum problema a ser resolvido.

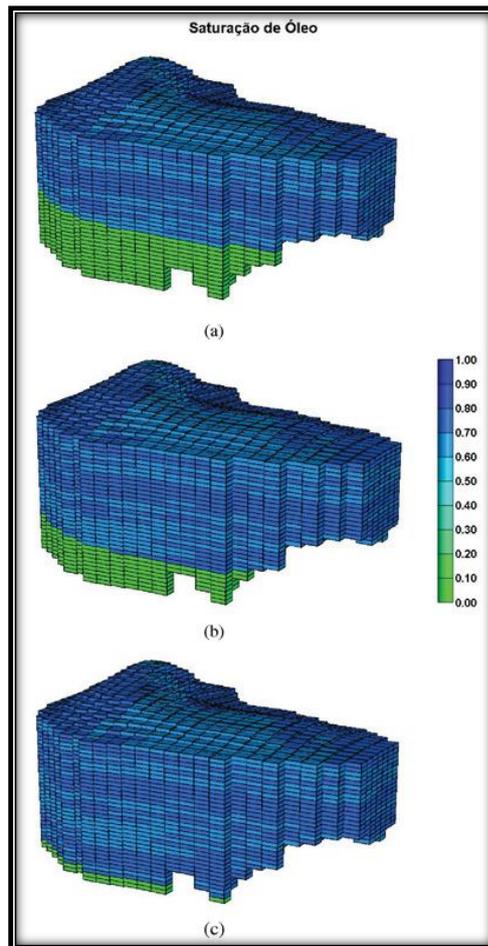
Nas próximas seções será apresentado brevemente o conceito de cenários e realizações de um modelo geológico, as razões que justificam o estabelecimento de tais cenários e algumas metodologias para a escolha de realizações representativas de um modelo geológico. A importância destas questões se deve ao fato de que não é economicamente viável e nem computacionalmente prático simular o fluxo de hidrocarbonetos em centenas de modelos em que muitos terão uma resposta muito parecida além de analisar e otimizar resultados de todas elas. Daí a necessidade escolher de um conjunto de  $N$  realizações um subconjunto de  $n$  realizações representativa no qual se preserve ao máximo as informações do conjunto maior.

### 5.1. Definição de Cenários Geológicos e Realizações de um Cenário.

Um cenário é uma hipótese, uma macro situação de um modelo geológico convencionada a priori. Tal hipótese pode ser originada de interpretações sísmicas alternativas de um determinado horizonte geológico, modelos geológicos conceituais alternativos, uma ou mais falhas serem selantes ou não, folhelhos serem selantes ou não, interpretações alternativas do zoneamento de um reservatório, etc. Um cenário alternativo envolverá a criação de um novo objeto determinístico (modelo).

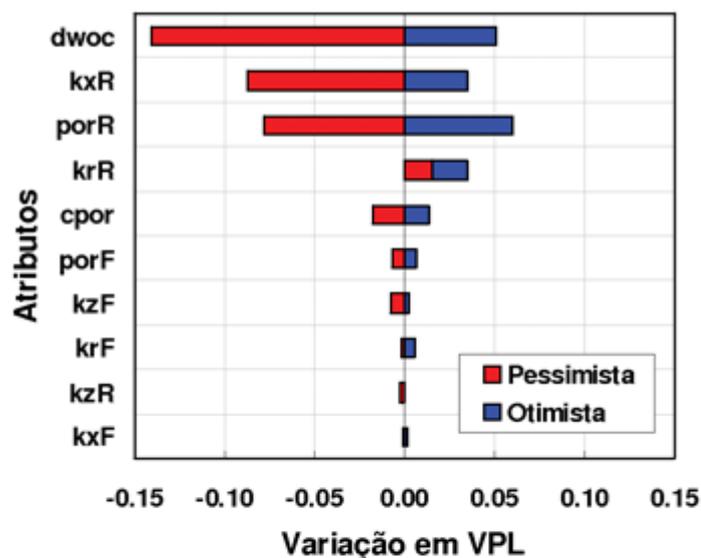
Na indústria do petróleo, na área da modelagem geológica, é usual definir cenários “otimista”, “intermediário” e “pessimista” para se definir um modelo geológico base. O que define os cenários em cada uma das categorias citadas anteriormente é o conjunto de hipóteses a serem consideradas na construção do modelo geológico, que supostamente, trariam respostas melhores ou piores de acordo com cada categoria que se encaixam.

Nas Figura 5.1 (a), (b) e (c) têm-se exemplos de mapas de saturação de óleo de cenários pessimista, intermediário e otimista, respectivamente. A hipótese a ser considerada nesse caso é a profundidade do contato óleo-água. Quanto mais profundo o contato-óleo água for, teoricamente, maiores serão os volumes ocupados pela parte de óleo no reservatório. Tais cenários alternativos da profundidade do contato óleo-água foram convencionados em (LIGERO, RISSO, & SCHIOZER, 2008) ao fazer análise econômica em reservatórios de petróleo considerando incertezas relativas, inclusive, à profundidade do contato óleo-água.

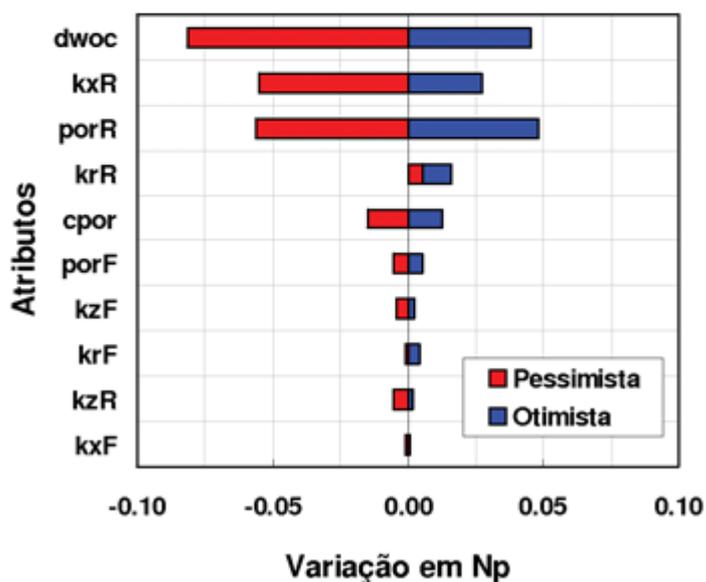


**Figura 5.1 – Cenários pessimista (a), provável (b) e otimista (c). (LIGERO, RISSO, & SCHIOZER, 2008).**

Tal influência sobre as funções objetivo VPL e  $N_p$  (volume de óleo acumulado produzido em um determinado período) foi medida através de uma análise da sensibilidade, exibida na Figura 5.2, onde o impacto sobre tais funções também foi observado a respeito de outras variáveis.



(a)



(b)

Figura 5.2 – Análise de sensibilidade para o VPL (a) e para o Np (b). (LIGERO, RISSO, & SCHIOZER, 2008).

O estabelecimento de cenários é uma tarefa multidisciplinar e que deve ser integrada com a participação de todos os profissionais que estão envolvidos no processo de modelagem uma vez que a necessidade de possíveis cenários vem da grande falta de informação (incerteza) do sistema petrolífero em questão já que provavelmente está em uma fase inicial de desenvolvimento. É subjetiva a escolha das variáveis decididas a priori que definirão os

cenários otimista, intermediário e pessimista, sendo apenas possível concluir qual é mais representativo e necessitará de menos ajustes aquele que produzir curvas de simulação semelhantes com as curvas de campo. Tal escolha pode envolver parâmetros mais complexos do que apenas atribuir incerteza ao contato óleo-água e irá variar muito intrinsecamente de acordo com o problema estudado e com as decisões assumida pela equipe envolvida no processo de modelagem. Exemplos de outros tipos de hipóteses é assumir se determinadas falhas são selantes ou não, o número de compartimentos do reservatório e se eles se comunicam entre si, métodos diferentes para cálculo de propriedades do reservatório, entre outros.

Um modelo é uma abstração de um sistema real e que pode ser utilizado para predição de algumas variáveis de interesse de tal sistema. Por ser uma abstração, necessita-se de diferentes modelos que representem as possíveis realidades (dentro de um mesmo cenário) de um mesmo reservatório para compensar tal falta de representatividade e conhecimento do sistema em questão. Tais modelos, gerados a partir da variação de parâmetros de um mesmo cenário, são chamados de realizações. Simulação de Monte Carlo é utilizada para produzir realizações a partir de sorteios de variáveis geológicas passíveis de abordagem probabilística (porosidade, saturação, contato óleo-água, Bo, NTG, Fácies, topo, isópacas. Uma prática que leva em consideração essa falta de conhecimento e incerteza é a de se gerar diversos modelos equiprováveis, submetê-los a uma função de transferência relativamente simples para depois escolher alguns modelos representativos para serem simulados. A questão que se abre é quais modelos escolher para a fase seguinte, a simulação.

## **5.2. Algumas Metodologias para escolha de modelos representativos.**

É frequente na área de reservatórios serem gerados de maneira estocástica centenas e até milhares de modelos a partir de cenários geológicos que também podem ser numerosos. A simulação de todos se torna muitas vezes dispendiosa ou impossível em termos de tempo computacional, mas mais limitante ainda é a capacidade de analisar, entender grande quantidade de dados e otimizar todos esses modelos. Tal questão levanta a necessidade de eleger, em meio a esse enorme conjunto de  $N$  modelos, um conjunto bem menor, de  $n$  modelos

sendo que o máximo possível de informações importantes sejam expressas e preservadas por esse conjunto  $n$ .

### 5.2.1. Método do VME.

Na Figura 5.3 é ilustrado de maneira genérica o fluxo de trabalho de múltiplas rodadas de simulação de reservatórios. Geralmente têm-se um simulador que pode ser alimentado com  $N$  modelos descritivos do reservatório diferentes, um conjunto  $U_k$  de estratégias de produção e na saída do simulador, para cada simulação de um modelo associado a uma estratégia de produção haverá um resultado de simulação diferente. Tais combinações podem ser verificadas na Figura 5.4.

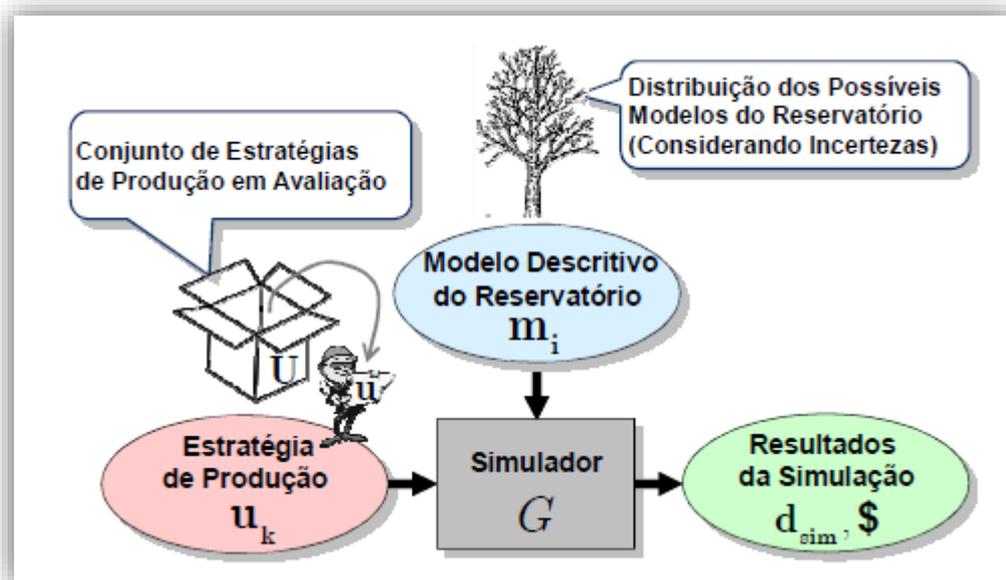
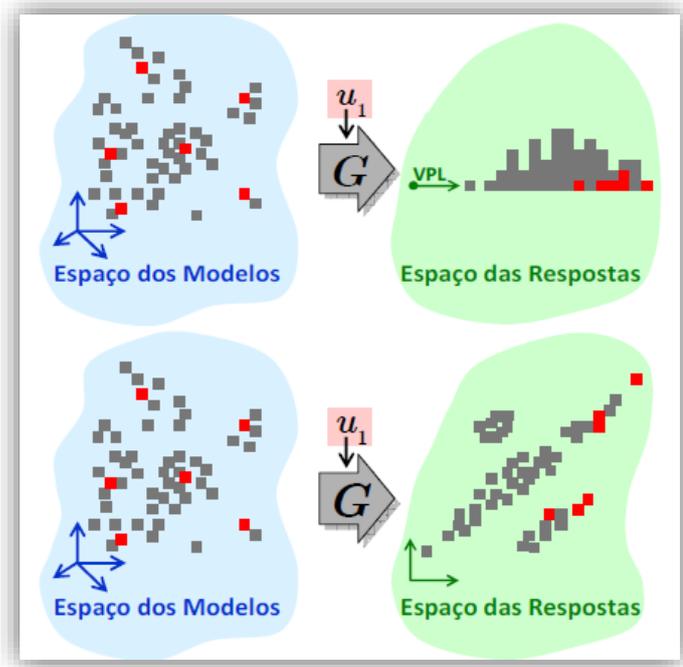


Figura 5.3 - Fluxo de trabalho genérico de múltiplas rodadas de simulação. (ROMEU, et al., 2013).

A principal motivação da seleção de modelos representativos é o critério de representatividade. O conjunto de modelos escolhidos deve ser muito menor que o conjunto inicial mas preservar as principais informações de modo que a decisão tomada a partir desse modelo menor seja a mesma que do que a tomada analisando o conjunto inicial por inteiro.

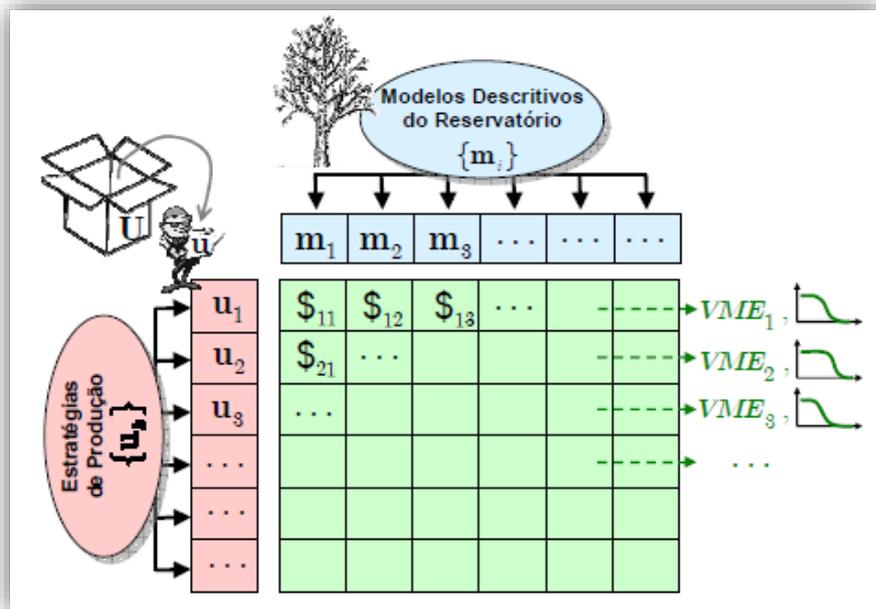
Tal questão é ilustrada na Figura 5.4 onde são amostrados determinados modelos, representados pelos pontos vermelhos, esses pontos parecem estar representando de maneira

eficiente o espaço dos modelos, e de fato estão, mas as respostas correspondentes a esses modelos escolhidos tendem a se concentrar em determinada região no espaço de respostas, à direita, ou seja, VPL médio desse conjunto escolhido para a estratégia  $u_1$  seria maior do que do conjunto inteiro. Mesmo se houvesse sido feita uma escolha em que o espaço de resposta tivesse sido bem representado, a boa representatividade não é garantida para as outras estratégias de produção.



**Figura 5.4 - Espaços paramétricos de modelos e de respostas (para uma estratégia de produção  $u_1$ ), com diferentes números de parâmetros. (ROMEU, et al., 2013).**

A Figura 5.5, como mencionado antes mostra combinações entre as diversas estratégias de produção e modelos descritivos gerando diversos valores de VPL para uma mesma estratégia de produção e um VME (valor monetário médio). Selecionar modelos representativos nessa situação significa trabalhar com um número muito menor de colunas, tentando preservar o máximo possível as informações do conjunto original. Dois modos de realizar tal escolha são considerados em (ROMEU, et al, 2013; ROMEU, 2013): um mais simples, que considera apenas o VME (valor monetário esperado) que será descrita a seguir e outro mais elaborado, que leva em conta todas as curvas de VPL, que não será apresentada aqui.



**Figura 5.5 – Combinação das diversas estratégias de produção (em rosa) e modelos alternativos do mesmo reservatório (em azul), gerando uma matriz de resultados (em verde). Para cada estratégia de produção (cada linha da matriz), calcula-se VME e curva de risco. (ROMEUE, et al., 2013)**

O critério de decisão baseado somente no VME visa fornecer um conjunto de modelos, levando em conta todas as estratégias de produção, que reproduza o melhor possível os VME que seriam calculados com todos os modelos do conjunto inicial. Tal método deverá apontar também, a partir dos modelos selecionados, a mesma estratégia de VME máximo que a escolhida através do conjunto inicial. Aqui escolher modelos representativos significaria escolher determinadas colunas da matriz da Figura 5.5.

O método sugerido em (ROMEUE, et al., 2013) consideram-se os VME, para cada estratégia de produção, calculados para o conjunto original de  $N$  modelos, com probabilidades  $p_i$ , e para o conjunto reduzido de  $n$  modelos, com probabilidades  $q_i$ :

$$VME_N(u_k) = \sum_{i=1}^N (p_i VPL(u_k, m_i)) \quad (11)$$

$$VME_n(u_k) = \sum_{i=1}^n (q_j VPL(u_k, m_i)) \quad (12)$$

O que se deseja é minimizar a diferença entre os resultados dos VME calculados para o conjunto inicial (equação 11) e o VME calculado para o conjunto reduzido (equação 12) para as estratégias de produção de interesse.

### 5.2.2. Escolha de Modelos em Espaços Paramétricos.

Essa metodologia consiste basicamente em amostrar visualmente modelos em um espaço paramétrico representado por quatro parâmetros: FR, Np, VPL e Wp. Tal método é uma aplicação recente citada por (ROMEU, et al., 2013) e desenvolvida por (MARQUES, 2012) cujo objetivo foi avaliar o valor da flexibilidade para, ou mitigar o risco, ou aumentar o valor do projeto. O estudo consistiu em gerar 292 os quais foram gerados estocasticamente variando parâmetros de entrada como permeabilidade horizontal, espessura porosa, entre outras.

Desses 292 modelos, 10 deveriam ser escolhidos de modo a representar o conjunto inteiro. A Figura 4.7 mostra os resultados parciais de (MARQUES, 2012) ao realizar a amostragem dos modelos representativos em dois tipos de espaços paramétricos 2D. Um composto pelo VPL e pelo NP (volume acumulado de líquidos produzidos) e Wp (volume acumulado de água produzida) e NP.

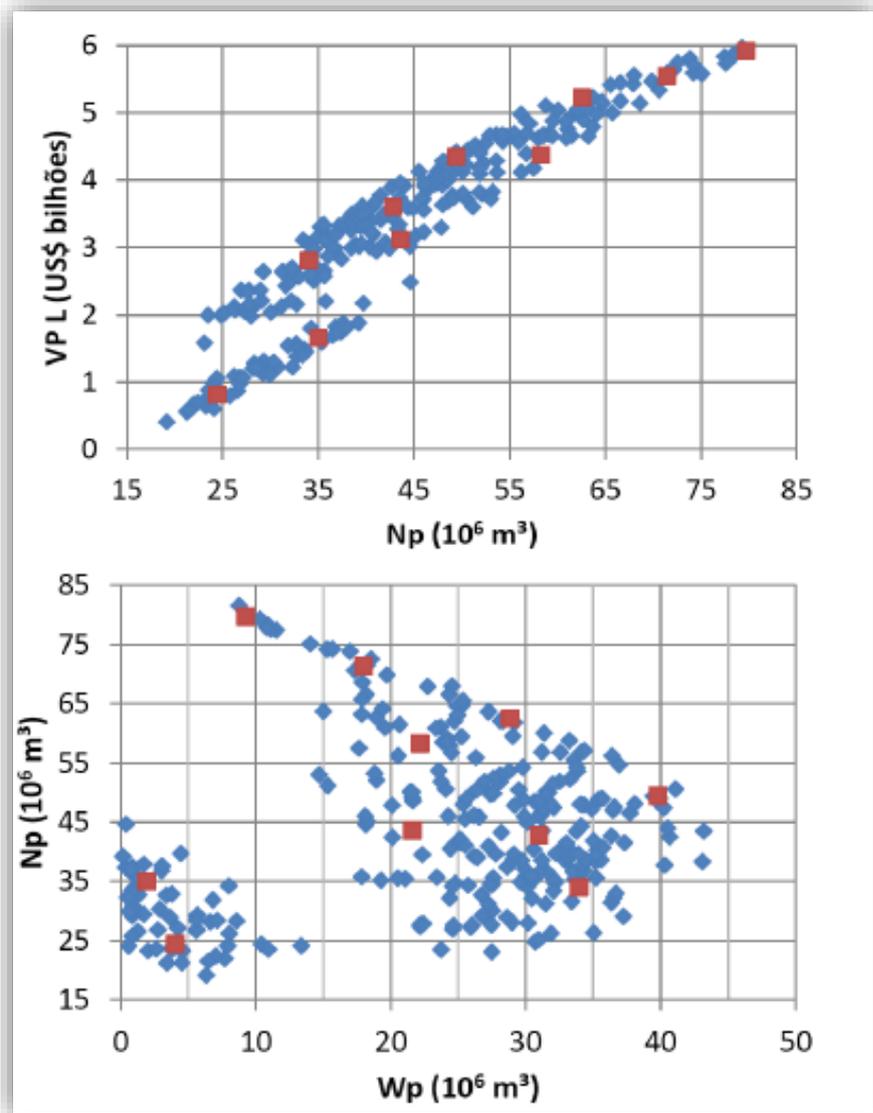


Figura 5.6 - Resultados parciais de uma aplicação da metodologia de seleção de modelos representativos em espaços paramétricos. (MARQUES, 2012).

### 5.2.3. Escolha de Modelos em Espaços Métricos.

Segundo (CAERS, 2011), a importante observação que se deve fazer para o entendimento das incertezas é que a complexidade e a dimensionalidade dos dados de entrada de um modelo geológico normalmente são muito maiores do que os da resposta alvo desejada. Exemplo de decisões: podem ser binárias como perfurar ou não, realizar mais testes ou não. Ao passo de que a complexidade do modelo de entrada pode ser enorme, contendo relações sofisticadas entre diferentes tipos de variáveis físicas essas ainda variam no espaço.

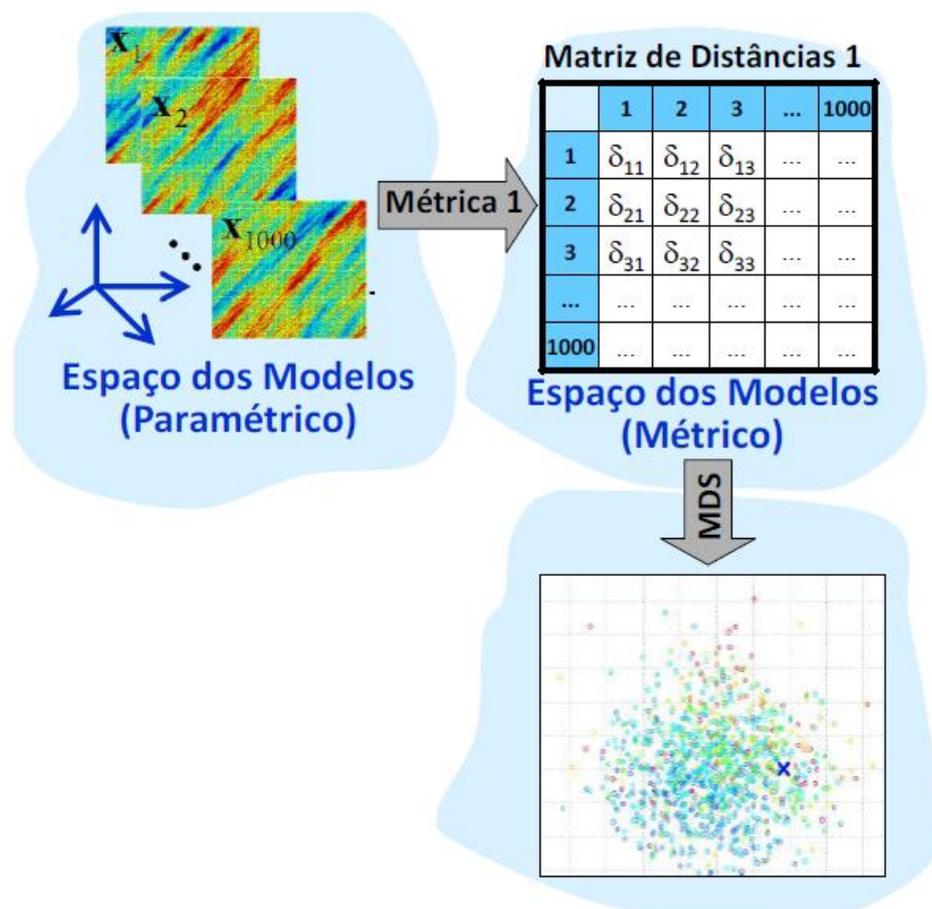
Essa observação sugere que a incerteza de um conjunto de modelos geológicos pode ser representada de um jeito mais simples. Muitos fatores podem afetar a resposta de um modelo, se a diferença em valor de uma única variável, como a porosidade leva a uma diferença considerável na resposta, então tal variável é crítica para o processo decisório. Mas como modelos geológicos têm um grande número de dimensões, às vezes tão grandes quanto o número de células, com propriedades que variam complexamente no espaço e no tempo, frequentemente não é simples discernir quais variáveis são críticas para o processo de tomada de decisão.

A distância é o valor simples e positivo que quantifica a diferença entre dois pontos em um determinado espaço. Nesse caso, dois modelos geológicos, ou realizações. Se existem  $N$  modelos geológicos, então uma matriz de distância  $N \times N$  pode ser especificada. Tratando-se de mapas 3D de propriedades (porosidades, permeabilidades, etc.), um dos métodos é a distância euclidiana dos valores célula a célula. Há outras alternativas de distâncias e cada uma irá fornecer uma configuração diferente no espaço métrico. Isso irá permitir “estruturar” a incerteza com uma resposta em particular em mente e criar um melhor esclarecimento de qual incerteza mais afeta a resposta dos modelos. Se os modelos (realizações) forem considerados peças de quebra-cabeças, duas peças bem similares podem ser agrupadas e representadas por uma peça “média” ou “intermediária” entre as duas.

Essa abordagem também é referida como “baseada em distância”, já que seu ponto de partida é o cálculo de distâncias entre cada modelo e todos os outros, expressando o quanto cada modelo é “diferente” do outro em aspectos de interesse.

Uma vez montada a matriz de distâncias, é usada uma forma visual de representar os aspectos essenciais dessa informação que é a técnica MDS (*Multidimensional Scaling* da álgebra linear). Essa técnica tenta capturar as informações da matriz de distâncias de modo a distribuir os  $N$  pontos em um espaço de duas (ou mais) dimensões, de tal forma que a distância euclidiana entre os  $N$  pontos nesse espaço de dimensão reduzida se aproxime ao máximo da distância verdadeira que consta na matriz de distâncias.

A principal diferença desse método para o anterior é que a escolha é feita no espaço métrico, considerando as distâncias entre os modelos, e não no espaço paramétrico. Uma vantagem é que o espaço métrico tem menos dimensões (talvez somente duas, dependendo da eficiência do MDS em capturar a informação da matriz de distâncias do caso em questão) do que o espaço paramétrico.



**Figura 5.7 - Representação dos modelos no espaço métrico, por uma matriz de distâncias e por uma redução 2D dessa matriz pela técnica MDS. De (ROMEU, et al., 2013) e adaptado de (CAERS, 2011).**

Posteriormente, a seleção de modelos representativos é feita nesse espaço métrico, por técnicas de clusterização, como *k-means* e *k-medoids*. Também é possível a escolha visual, podendo ser mais ou menos bem sucedida.

Na Figura 5.7 percebe-se uma disposição um tanto sem forma ou tendência aparente. (ROMEU, et al., 2013) atenta para a necessidade de haver um bom agrupamento dos modelos no espaço métrico para efetuar a escolha e as citadas foram duas.

A primeira é a busca de métricas mais significativas para o cálculo das distâncias, conforme a natureza do problema. Com a métrica apropriada, a distribuição dos pontos no espaço métrico resulta em um agrupamento mais estruturado, mais passível de interpretação segundo critérios relevantes para o problema de interesse (Figura 5.8).

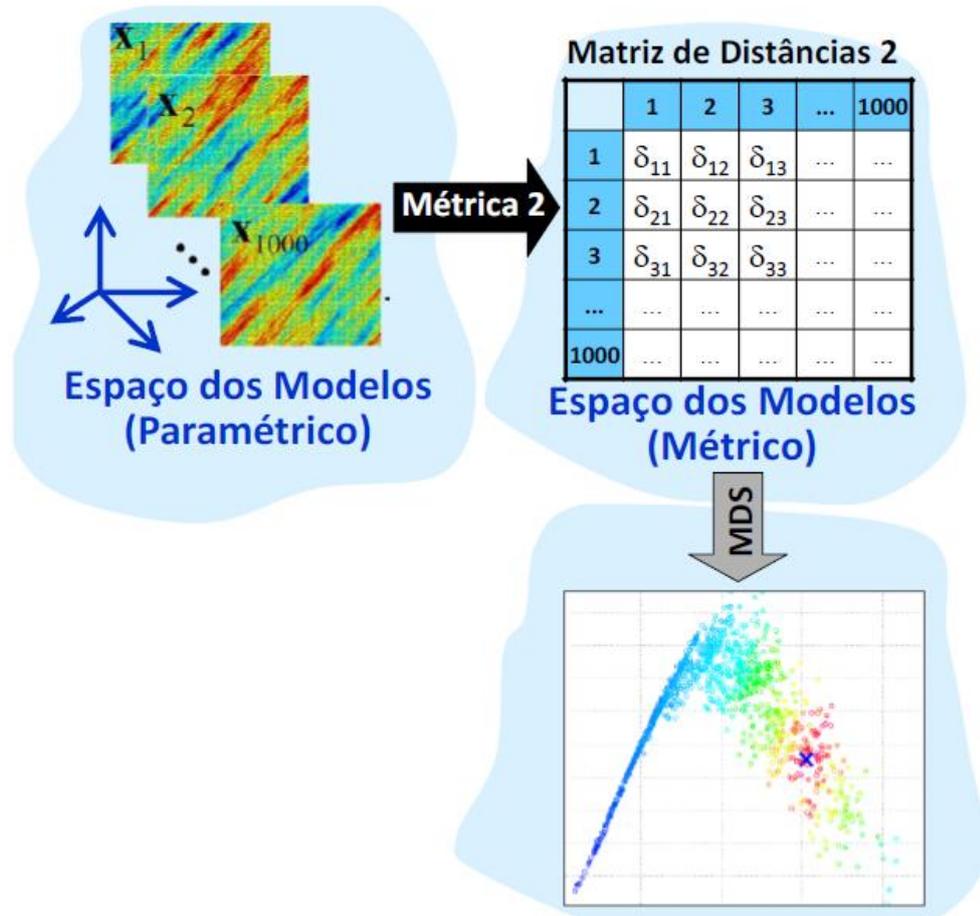


Figura 5.8 - Semelhante à Figura 4.7, mas usando uma métrica diferente para cálculo das distâncias, baseada em resultados de uma simulação por linhas de fluxo. (ROMEUE, 2013 – adaptado de CAERS, 2011).

A segunda seria usar a Transformada de Kernel na matriz de distâncias após a aplicação do MDS:

$$k(m_i, m_j) = \exp\left(-\frac{\delta(m_i, m_j)}{\sigma}\right)$$

Onde  $\delta$  é a distância original entre dois modelos, e  $\sigma$  é um fator de escala. Essa transformação é conhecida por agrupar de maneira ainda mais favorável à escolha de modelos, como exemplificado na Figura 5.9.

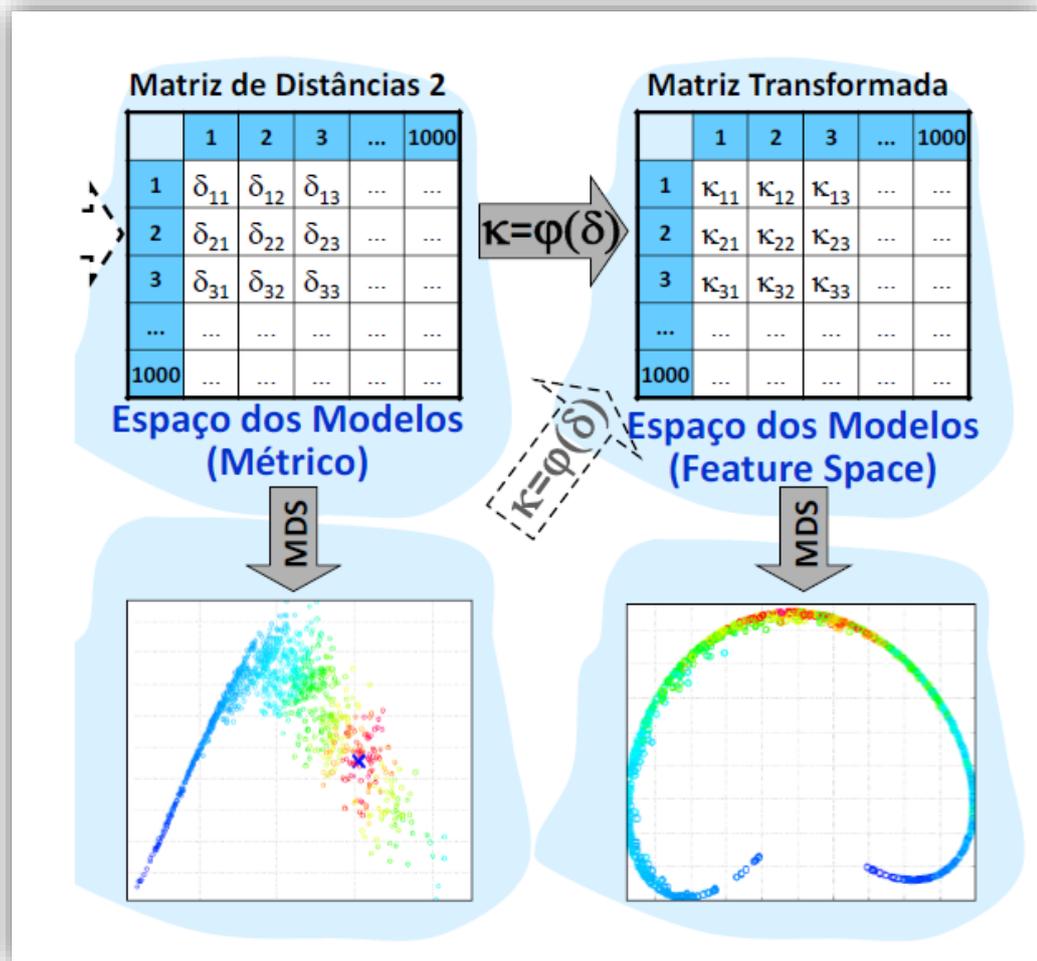


Figura 5.9 - Transformação da matriz da distância para um remapeamento vantajoso do espaço dos modelos (ROMEY, et al., 2013) e adaptado de (CAERS, 2011).

## 6. APLICAÇÃO

---

O Objetivo principal deste capítulo é apresentar uma visualização da incerteza na modelagem geológica, mostrando as realizações estocásticas desta modelagem. O capítulo mostra como são geradas diversas realizações de um modelo geológico através do fluxo de trabalho de incertezas do Petrel. As incertezas são consideradas ao se estabelecer variáveis como fontes de incertezas dentro de processos da construção do modelo geológico. A estas variáveis serão atribuídas funções de densidade de probabilidade para que dessas possam ser amostrados valores por meio da Simulação de Monte Carlo com hipercubo latino.

A seção 6.1 apresentará uma breve descrição do modelo geológico a ser usado como base. Nele serão apresentados os principais mapas de propriedades petrofísicas do modelo e outras características consideradas de importância. Os detalhes técnicos do uso do Petrel na construção do modelo podem ser conferidos no Apêndice.

A seção 6.2 trata do funcionamento do fluxo de trabalho de incertezas. Será determinado em que etapas serão atribuídas as incertezas, quais variáveis, os valores máximos e mínimos de ocorrência dessas variáveis. Por fim, serão mostrados as capturas de tela do mapa de porosidade estimada das 50 realizações geradas por meio desse fluxo de trabalho com posterior visualização de um mapa bivariado do desvio padrão célula a célula de todas essas realizações.

### 6.1. Descrição do Modelo Geológico Usado

O modelo usado representa um reservatório real que por motivo de sigilo de informação não pode ser identificado. Algumas características são válidas de serem abordadas:

- O reservatório possui dois poços produtores e um injetor;
- O poço injetor injeta aproximadamente 2000 m<sup>3</sup>/d de água na zona do aquífero. Embora a injeção não seja suficiente para manter o balanço de massa, o declínio da pressão é muito pequeno, indicando forte influência do aquífero;

- O principal poço produtor teve um pico de produção da ordem de 5200 m<sup>3</sup>/d. Este poço é o melhor do campo, tendo sido durante um período responsável por mais de 50% da produção total de óleo do campo;
- O segundo poço foi completado mais próximo ao aquífero e produz cerca de 5000 m<sup>3</sup>/d de líquido (1000 m<sup>3</sup>/d de óleo e 4000 m<sup>3</sup>/d de água).

O modelo geológico construído possui 122 x 86 x 47 células, de 50m x 50m x 1m. As propriedades estimadas no modelo, através da simulação sequencial gaussiana (SGS) foram Porosidade, Permeabilidade e NTG (Net to Gross), a partir das informações de três poços existentes na área.

A porosidade foi estimada utilizando o dado de amplitude sísmica como variável secundária.

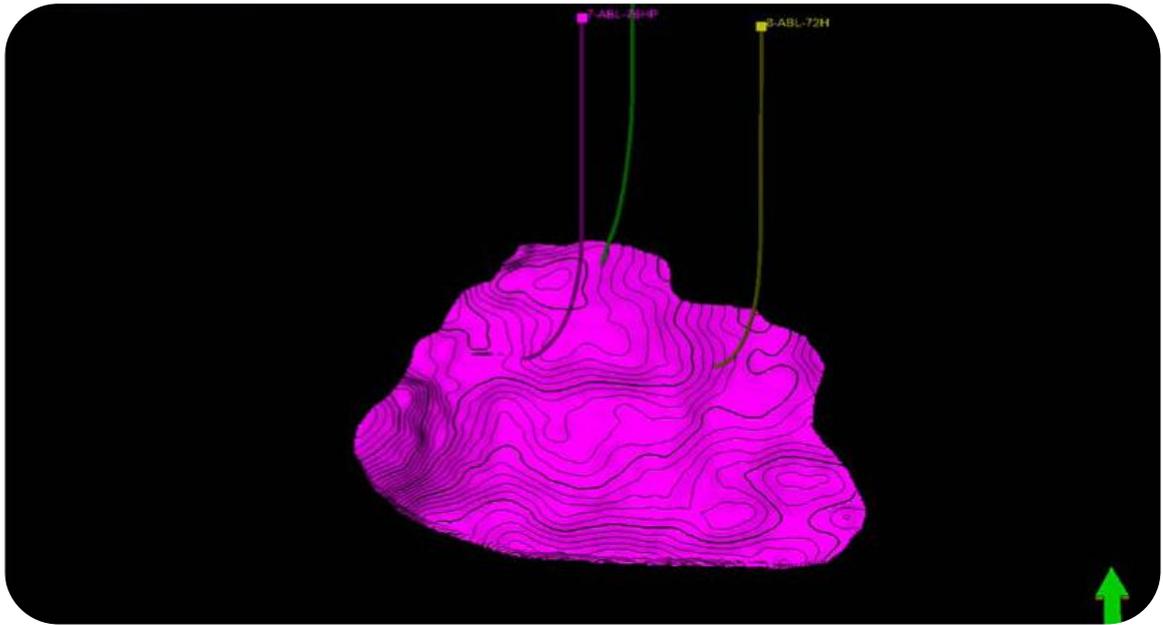
A propriedade permeabilidade foi calculada por co-krigagem, utilizando a porosidade como variável secundária e considerando uma distribuição normal com média de 5000 mD e desvio padrão de 1000 mD, dados estes baseados em teste de formação.

Por fim, a razão NTG foi estimada a partir da correlação com a amplitude sísmica da base do reservatório, que apresentou um R<sup>2</sup> igual a -0,7.

Não foi feito um modelo de simulação, pois não haviam pacotes que tivessem simultaneamente processos de modelagem geológica e de simulação para montar o fluxo de trabalho de incertezas.

Para a construção do modelo geológico, os dados de entrada utilizados foram as informações dos poços (perfis e marcadores estratigráficos) e de superfícies (topo e base do reservatório, previamente interpretados por um geofísico).

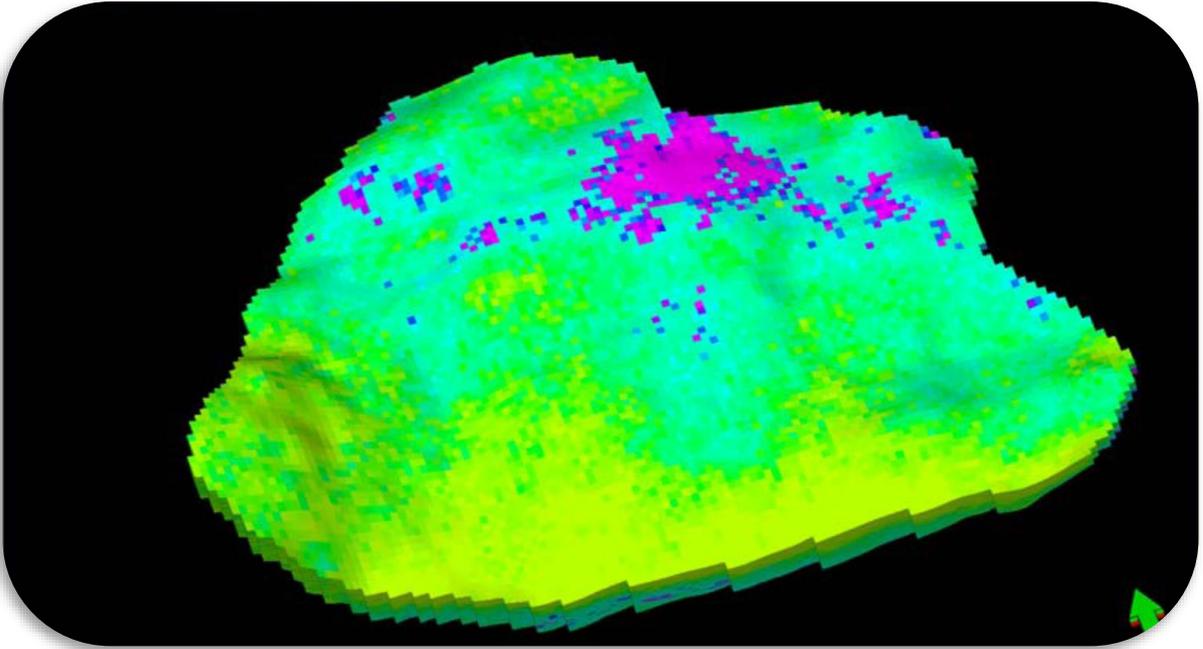
Além disso, outro dado utilizado para a modelagem foi o cubo de amplitude sísmica, que engloba toda a área do modelo. Na Figura 6.1 pode-se observar as superfícies de topo e base do modelo geológico em questão gerado no Petrel juntamente com as trajetórias dos poços.



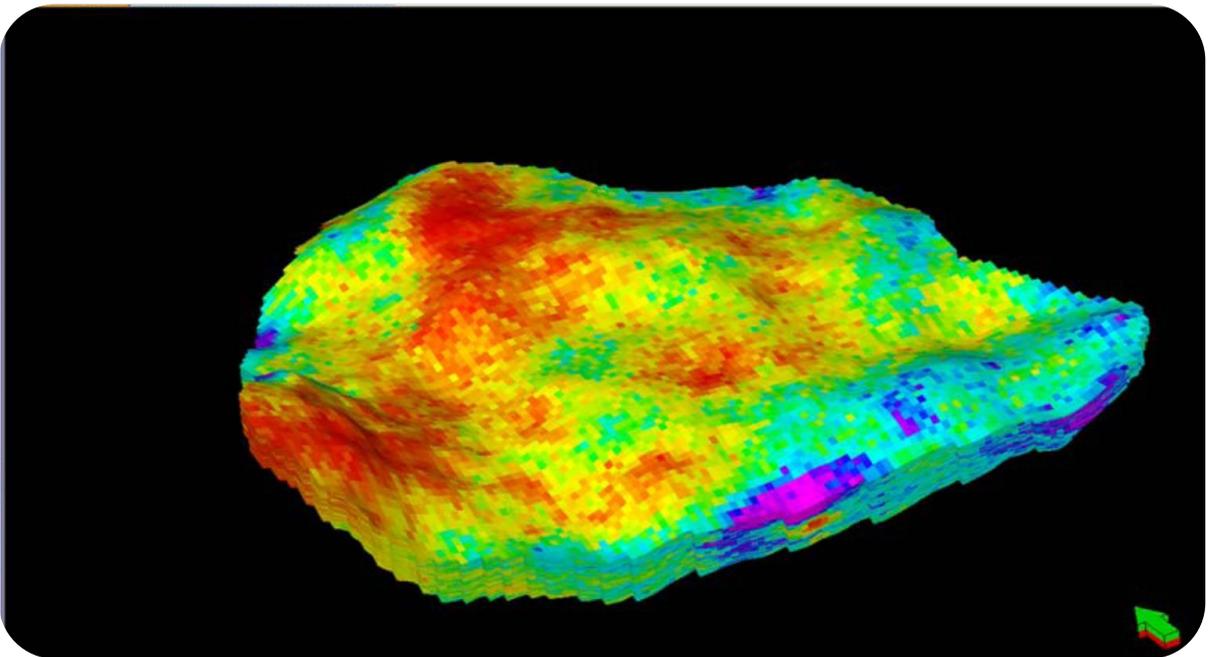
**Figura 6.1 - Poços do reservatório em questão, presentes no projeto Petrel e as superfícies de topo e Base.**

Nas Figura 6.2, Figura 6.3 e Figura 6.4 são mostrados os mapas da porosidade, NTG e permeabilidade estimados.

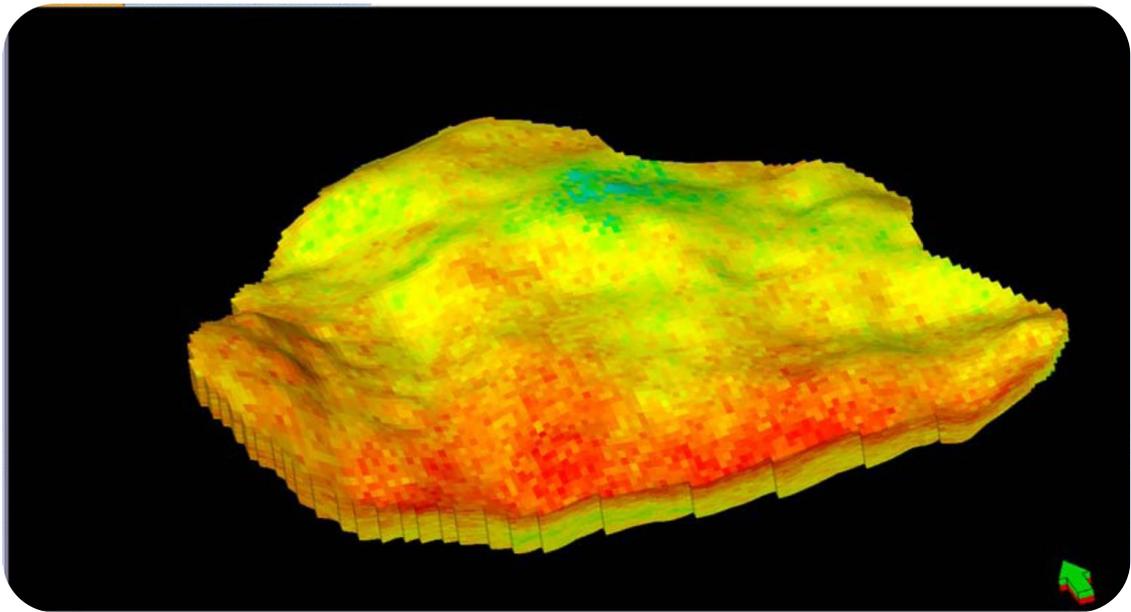
O Apêndice mostra o passo a passo no Petrel para a montagem do modelo geológico e obtenção de tais mapas.



**Figura 6.2- Visualização da porosidade estimada**



**Figura 6.3- Visualização do NTG estimado.**



**Figura 6.4 - Visualização da permeabilidade estimada.**

## **6.2. Fluxo de trabalho de incertezas do Petrel**

Realizada a criação do modelo geológico, pode-se iniciar o fluxo de trabalho de incertezas do Petrel para criação de realizações desse modelo para posterior comparação dessas realizações ou aplicação de técnicas tais como aquelas apresentadas no capítulo 5 para escolha de modelos representativos.

Na Figura 6.5 pode-se observar os diferentes níveis de complexidade em que se pode estudar as incertezas em um reservatório. No fluxo de trabalho o estudo é feito no nível da “semente aleatória” pois somente parâmetros estáticos que tenham uma abordagem geoestatística (que nesse caso serão NTG, porosidade e permeabilidade) serão abordados. Diferentes interpretações de dados sísmicos envolveria um estudo de incerteza que se situa em níveis superiores de complexidade (balão azul claro Modelos Deposicionais, Diagenéticos). Um nível ainda mais baixo de complexidade seria considerar apenas um único modelo como representativo do sistema.

Geradas as realizações, uma próxima etapa seria escolher entre elas os modelos representativos através de alguns dos métodos mencionados na seção 4.2. O Metrel, um *plugin* do Petrel, é capaz de fazer tais escolhas pela metodologia de escolha dos modelos no espaço métrico mas tal etapa não foi realizada.

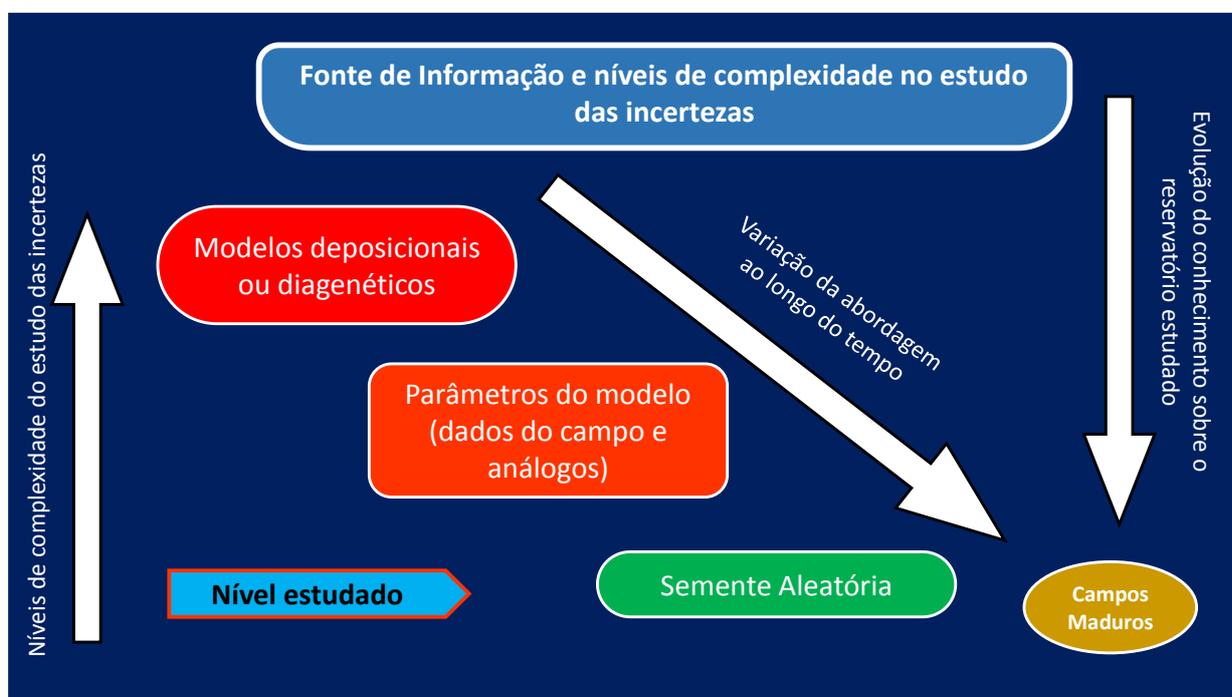


Figura 6.5 – Fontes de informação e Níveis de Complexidade no estudo das incertezas <sup>1</sup>.

### 6.2.1. Funcionamento do Fluxo de trabalho e resultados

Para dar início ao estudo de análise de incertezas, precisa-se primeiro criar um caso de simulação, ou um caso de cálculo de volumes (volumes porosos, inicial de água, inicial de gás, inicial de óleo). A Figura 5.6 descreve os processos do Petrel que são assumidos como tarefas no fluxo de trabalho para que sejam executadas de acordo com o número de rodadas for determinado.

No fluxo de trabalho colocam-se os processos, que nesse caso são de modelagem geológica, do programa nos quais deseja-se assimilar parâmetros considerando incertezas usando uma abordagem geoestatística.

<sup>1</sup> Figura adaptada de um slide selecionado de BLAUTH, M., ABREU, C. S., MONTEIRO, M., & SOUZA, O. G. (2013). Apresentação de slides da Petrobras (material completo não disponível, informação reservada)."

Na Figura 6.6 começa-se estabelecendo que o sistema forneça uma mensagem informando o número da interação do fluxo de trabalho (\$LOOP, que já é uma variável pré-estabelecida pelo sistema). Na terceira linha estabelece-se qual é a malha no qual os processos subsequentes serão rodados. Para efeitos de simplicidade, os processos das linhas 4 a 8 foram desconsiderados do fluxo de trabalho pois os critérios para se estabelecer variáveis de incerteza em tais processos é mais complexo e também pelo objetivo de o exemplo ser apenas ilustrar os fundamentos da geração de realizações de um modelo geológico base.

Nas linhas 9 e 10 foram alocados processos de *Petrophysical Modeling* para as propriedades NTG (*Net to gross*) e Porosidade nos quais serão estabelecidas variáveis de incertezas que a cada rodada do fluxo de trabalho assumirão valores amostrados através de alguma distribuição de probabilidade (nesse caso será a uniforme) pelo método da Simulação de Monte Carlo. Essas duas linhas são as linhas principais do fluxo de trabalho em questão pois é a etapa em que serão consideradas as incertezas através da perturbação de parâmetros do modelo base.

Na linha 11 é estabelecido o comando para o cálculo dos volumes. Na linha 12 o comando para exportar a realização gerada em um formato RESCUE (formato “padrão” que consegue ser lido por muitos *softwares* comerciais) e na linha 13 no formato EGRID (formato do Eclipse). Pode-se observar que as realizações levarão o número de sua interação em seu nome de arquivo. O formato RESCUE será importante para importar as realizações para o GoCad.

Por último, na linha 14, gerar-se-á um *print* da janela de visualização 3D em formato PNG também com o número da interação (\$LOOP).

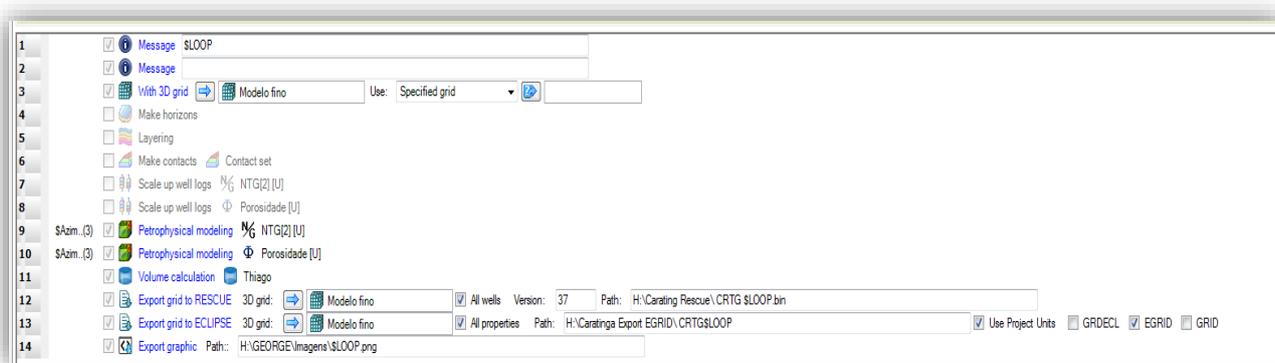
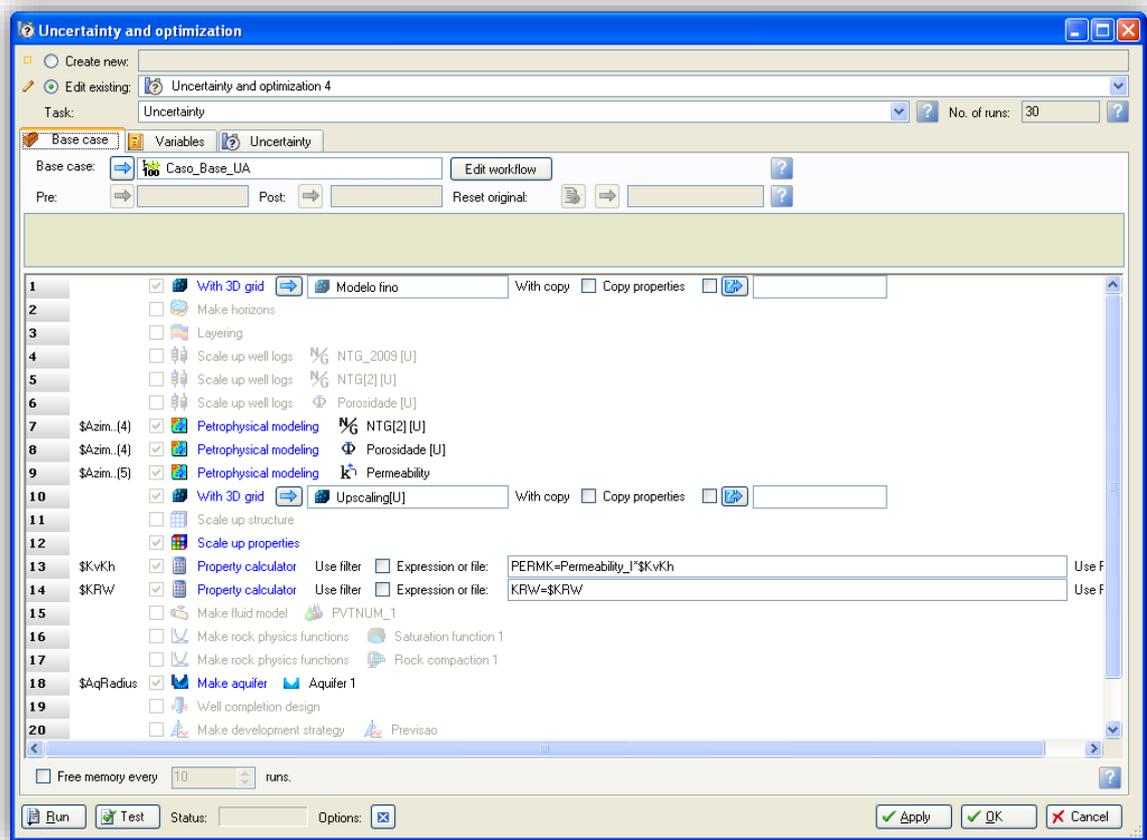


Figura 6.6 - Configuração usada no Fluxo de trabalho de teste.

A Figura 6.7 é um exemplo de fluxo de trabalho mais completo que foi usado em um trabalho anterior para apresentar o funcionamento do fluxo de trabalho para os geólogos e engenheiros de reservatório dos CENPES. Onde foi usado o *Property Calculator* para associar incerteza à razão kv/kh e o processo *Make Aquifer* para associar incerteza ao raio do aquífero atuante.



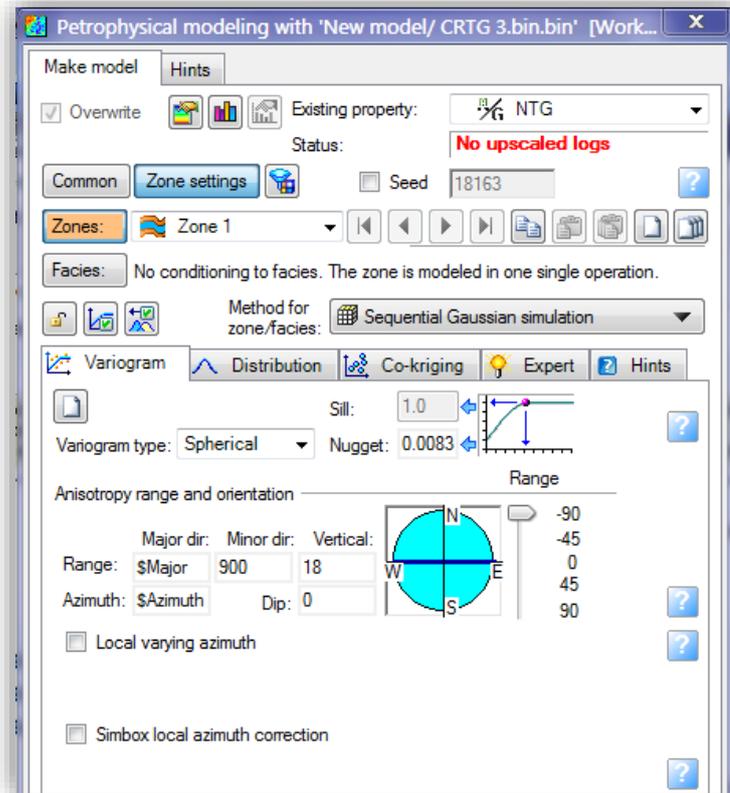
**Figura 6.7 - Outra opção de fluxo de trabalho.**

Na Figura 6.8 mostra-se a lista das variáveis de incerteza. As variáveis são aquelas na coluna “name” e que possuem um “\$” no seu início. A cada uma é associada valores máximos, mínimos e base que são parâmetros para a distribuição uniforme. Pode-se escolher dentre as distribuições Normal, Triangular e outras, dependendo da adequação dos dados com a distribuição em questão. Há também a opção de se importar uma distribuição.

	Type	Int	Name	Base value	Distribution	Arguments			
1	Uncertain	<input type="checkbox"/>	\$Azimuth	60	Uniform	Min	50	Max	70
2	Uncertain	<input type="checkbox"/>	\$Major	1100	Uniform	Min	1000	Max	1200
3	Uncertain	<input type="checkbox"/>	\$Minor	850	Uniform	Min	800	Max	900
4	Uncertain	<input type="checkbox"/>	\$azimuthP	60	Uniform	Min	40	Max	80
5	Uncertain	<input type="checkbox"/>	\$majorP	1200	Uniform	Min	1000	Max	1300
6	Uncertain	<input type="checkbox"/>	\$minorP	850	Uniform	Min	800	Max	900
7	Uncertain	<input type="checkbox"/>	\$majorPe	1100	Uniform	Min	1000	Max	1200
8	Uncertain	<input type="checkbox"/>	\$minorPe	900	Uniform	Min	850	Max	950

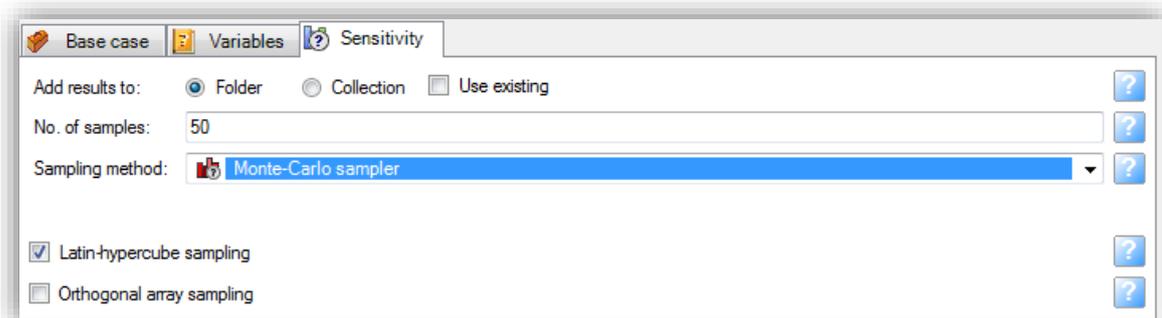
**Figura 6.8 – Lista de Variáveis/Parâmetros incertos considerados no fluxo de trabalho, suas distribuições e seus ranges de ocorrência.**

Na Figura 6.9 se pode ver onde tais variáveis serão definidas no processo *Petrophysical Modeling* para que sejam perturbadas pela Simulação de Monte Carlo. Serão definidas nos parâmetros do variograma do processo, onde assumiu-se que há incertezas. Os nomes das variáveis são sugestivos pois remete ao campo onde elas serão inseridas. “\$Major” será alocada no campo com coluna “Major” e linha “Range”, assim como “\$Azimuth” irá na linha “Azimuth” para NTG. Nos processos relativos a porosidade e permeabilidade, foram criadas outras variáveis com mesmo nome mas P (de porosidade) e Pe (de permeabilidade) no final de seus nomes, isso pode ser verificado na Figura 6.8. A cada rodada do fluxo de trabalho tais variáveis assumirão um valor amostrado das distribuições de probabilidade definidas na Figura 6.8. Tais valores serão determinantes para gerar mapas diferentes de NTG.



**Figura 6.9 – As variáveis de incertezas serão definidas nos campos “Range” e “Azimuth” do variograma.**

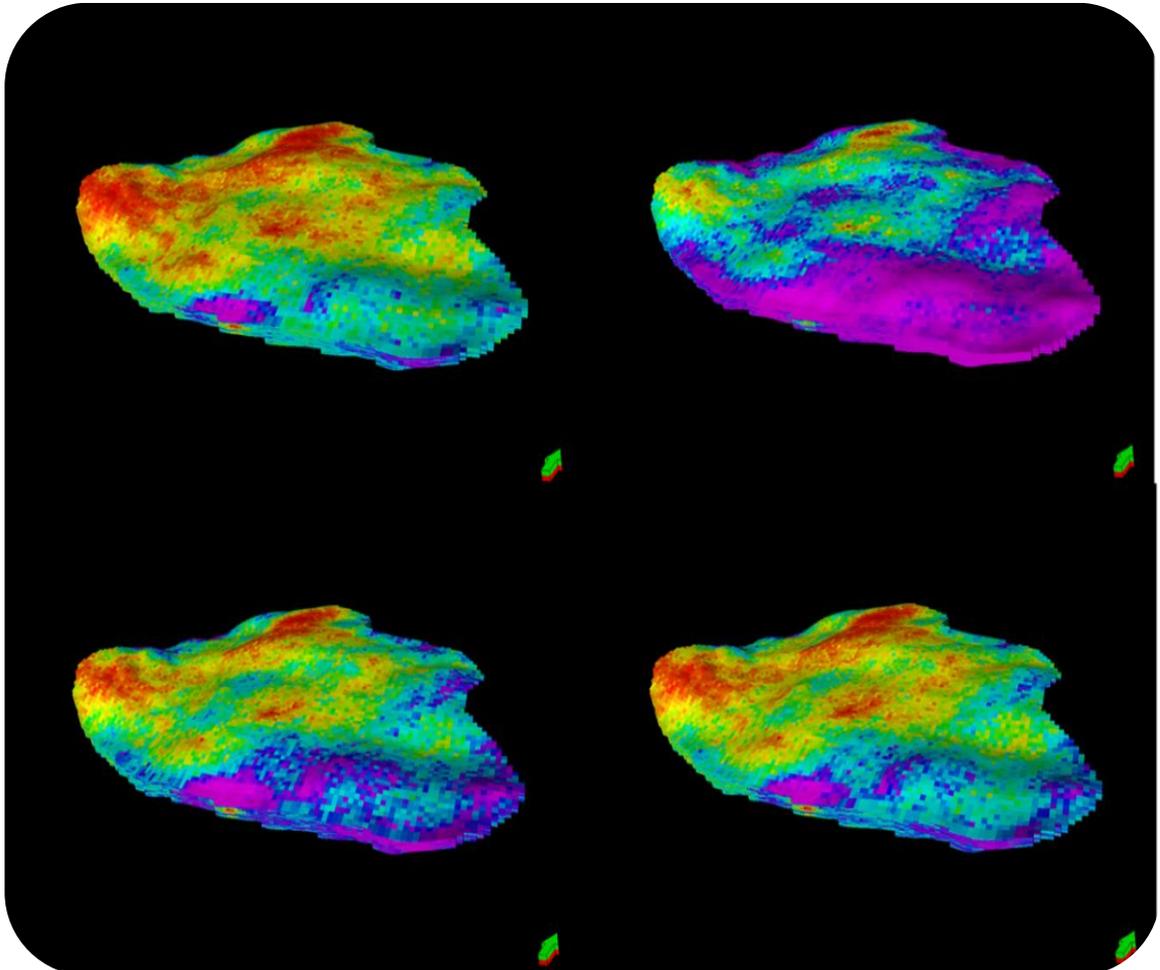
Na Figura 6.10 observa-se a escolha da Simulação de Monte Carlos como método de amostragem para as variáveis estabelecidas e a marcação da opção do Hipercubo Latino que possibilita a divisão da distribuição em quantis de modo que a amostragem seja representativa



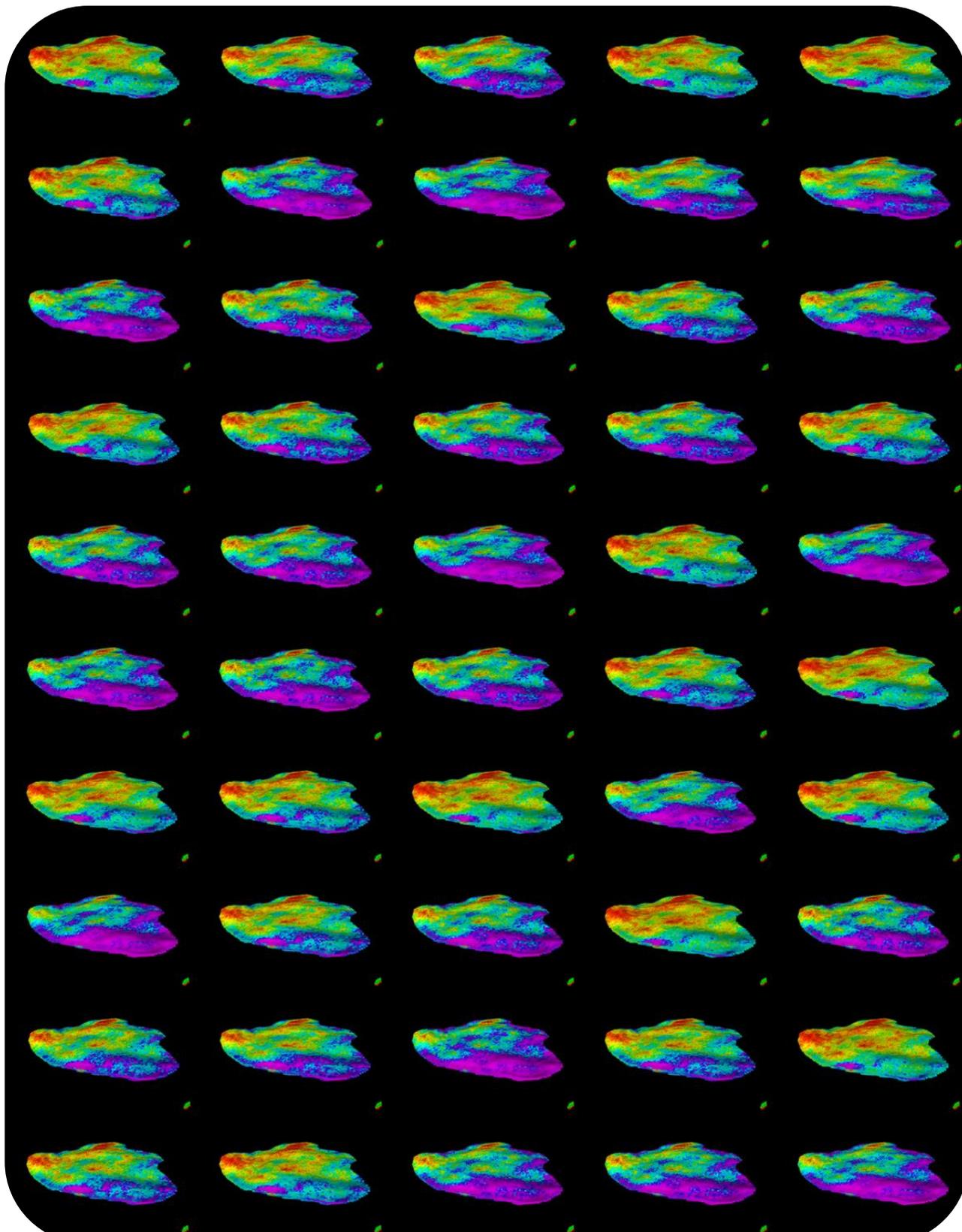
em toda a faixa de valores da distribuição.

**Figura 6.10 – Escolha do Método de Monte Carlo de amostragem com habilitação de hipercubo latino.**

A Figura 6.11 apresenta a visualização mais destacada de 4 realizações evidentemente diferentes. Na Figura 6.12 pode-se visualizar todos os mapas de porosidade das 50 realizações geradas pelo fluxo de trabalho aplicando a técnica de justaposição apresentada na seção 4.2.4.



**Figura 6.11- Visualização justaposta dos mapas de porosidade de 4 realizações.**



**Figura 6.12 – Capturas de tela dos mapas de porosidade de todas as 50 realizações do modelo geológico base usado no fluxo de trabalho.**

Geradas as realizações, usa-se as malhas em formato RESCUE para importá-los para o GOCAD, outro *software* de modelagem geológica (Paradigm). Há um *plugin* do GOCAD chamado *Uncertainty Viewer* que tem as funcionalidades mostradas na Figura 6.13. Tais funcionalidades usam informações fornecidas por diversos modelos (realizações, nesse caso) para gerar estatísticas, calcular percentis, ou intervalos interquantis. É possível calcular distância dos poços, consultar a qual percentil uma realização pertence ou até mesmo ordenar (*Rank*) as realizações segundo algum critério desejado.

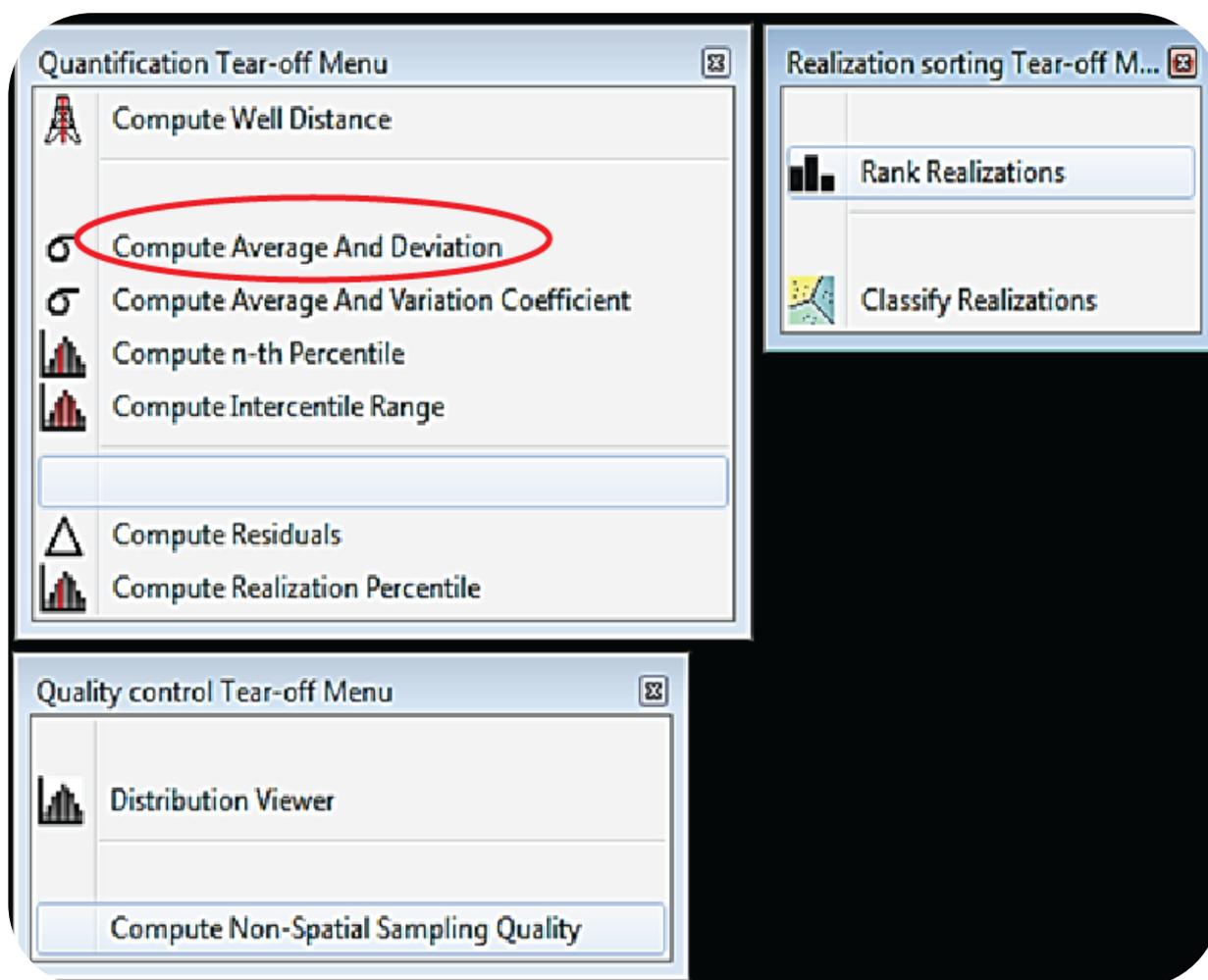


Figura 6.13 – Funcionalidades do *Uncertainty Viewer* plugin.

Depois de importadas as realizações usou-se a opção “*Compute Average and Deviation*” para gerar duas propriedades novas no modelo: Desvio Padrão e Média calculadas célula a célula de todas as realizações. Na Figura 6.14 o mapa de desvio padrão é mostrado. Na Figura 6.15 é possível observar a barra de cores com uma curva amarela de opacidade, se trata de um mapa bivariado (rever seção 3.2.4), que nesse caso, foi estabelecido um valor baixo de

opacidade para valores baixos de desvio padrão e um valor alto de opacidade para valores altos do desvio padrão. A opacidade varia de 0 a 1 mas nenhuma cor na escala recebeu opacidade zero. Os valores mais altos de opacidade foram escolhidos para enfatizar os valores maiores de desvio padrão, que está denotando a incerteza neste caso.

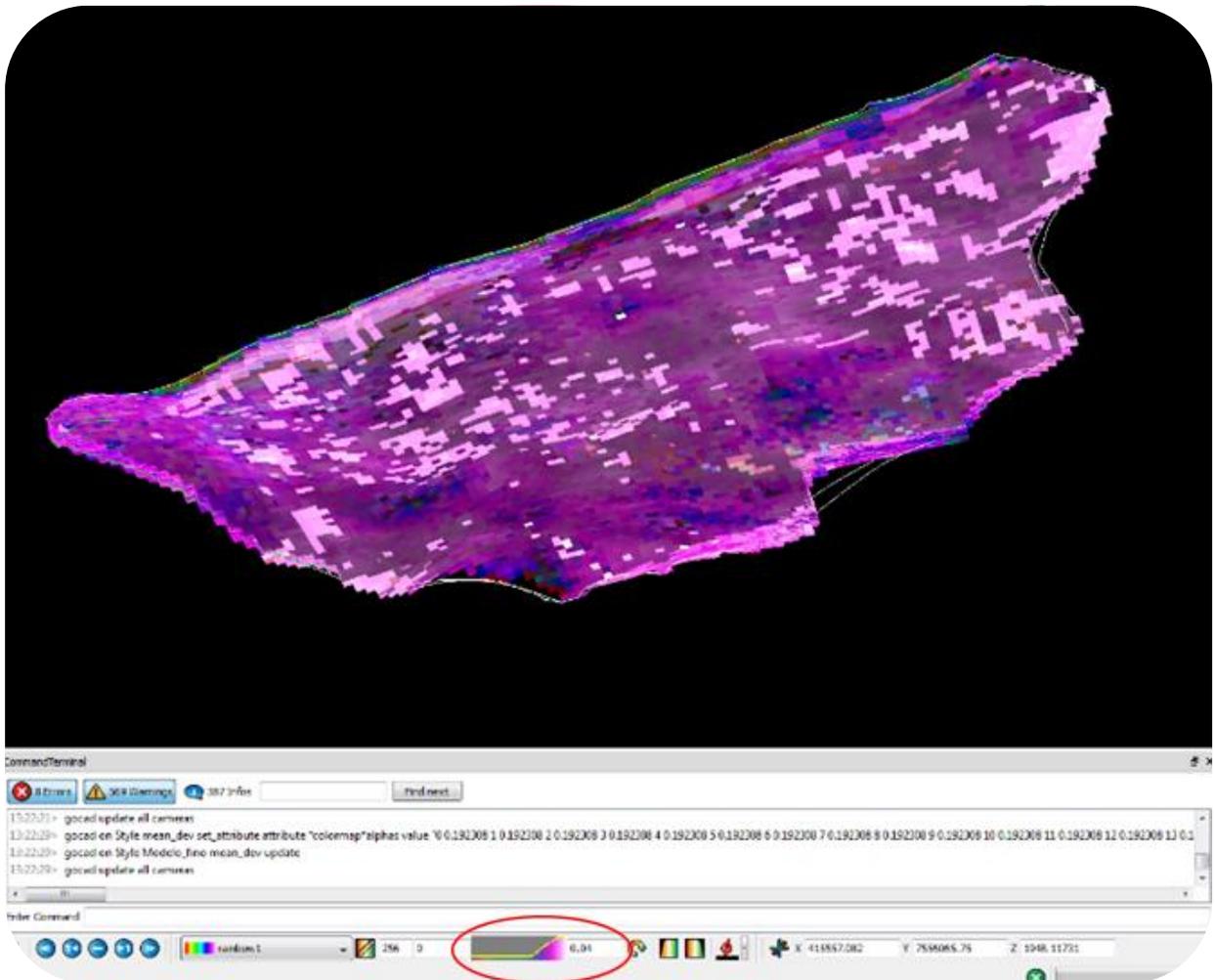


Figura 6.14 – Mapa de desvio padrão gerado a partir das realizações.



Figura 6.15 – Barra de cores e curva (amarela) de opacidade.

## 7. CONCLUSÕES

---

Com esse trabalho foi possível observar muitas abordagens que as incertezas podem ter tanto em termos de quantificação quanto no âmbito da visualização.

O primeiro objetivo, que era apresentar técnicas de quantificação de incertezas, foi atingido apresentando a abordagem probabilística clássica, que é a mais consolidada, bem aceita e a base de muitas outras ferramentas de estudo de incertezas, principalmente Simulação de Monte e a Entropia da Informação. Depois foram apresentadas a Simulação de Monte Carlo e Análise de Sensibilidade que são amplamente usadas tanto na indústria do petróleo quanto em outras áreas do conhecimento e se mostram essenciais para o estudo das incertezas. Viu-se que a Análise da Sensibilidade se mostra importante por diminuir a dimensionalidade dos problemas, explicitando as variáveis mais importantes e de maior impacto na função objetivo e servirá como subsídio para a Simulação de Monte Carlo no momento que forem definidos os parâmetros que serão perturbados para gerar as realizações dos modelos de reservatórios. A Entropia da Informação, método principal de quantificação de incertezas apresentado nesse trabalho, se mostrou uma abordagem adicional a subsidiar as tomadas de decisão podendo ser tema de futuros trabalhos.

Depois foram vistas as técnicas de visualização para representação de incertezas, começando por algumas variações de Box Plot que ao longo de sua evolução, incorporou de maneira intuitiva e de fácil entendimento recursos visuais de modo a representar os principais parâmetros estatísticos de um conjunto de dados. Em seguida foram analisadas outras técnicas como Gráfico de Coordenadas Paralelas, Gráficos Spagetti, Fan Charts, Mapas de Cores, Justaposição e Superposição, Cone HSV e mapas bivariados, Mapas de Deslocamento, métodos gráficos para renderizar volumes e mapas vetoriais 3D; cada uma com suas peculiaridades e utilidades. Foi observado que a escolha de cada técnica está relacionada à dimensionalidade, tipologia dos dados e o que se espera visualizar. O uso em conjuntos de várias técnicas simultaneamente é frequente.

Uma situação em que se tem em mãos dezenas (e as vezes até centenas) de realizações de modelos, justifica o terceiro objetivo que é apresentar métodos para a escolha de modelos representativos de um reservatório. Foram citados princípios básicos de dois métodos: escolha

de modelos em espaço paramétrico e escolha de modelos em espaço métrico. O primeiro o método consiste em escolher, a princípio, os modelos de maneira visual, o segundo envolve métodos de clusterização e ferramentas da álgebra linear que culminam que a escolha seja feita em um espaço métrico onde a dimensionalidade do problema e a representatividade do reservatório seja mais eficiente.

Por fim, o último objetivo foi cumprido ao descrever uma aplicação básica de como as incertezas podem ser incorporadas no fluxo de trabalho para gerar realizações de um modelo de reservatório. O estudo considerou incertezas em um nível baixo de complexidade, o da semente aleatória, no qual o modelo é perturbado somente em parâmetros que possuem abordagem geoestatística por meio da Simulação de Monte Carlo.

Com esse trabalho conclui-se que a temática do estudo da incerteza envolve a integração tanto no âmbito do uso das ferramentas a serem utilizadas quanto no âmbito do trabalho em equipe de profissionais de diversas disciplinas na indústria do petróleo. As subáreas para expansão se dividem tanto na implementação das técnicas de visualização, que são pouco exploradas aplicadas à incerteza, e em métodos quantitativos para medi-la, agregando mais informação à abordagem tradicional que é a probabilística. A necessidade de subsidiar tomadas de decisões com as incertezas cresce a cada dia com o surgimento de fontes não-convencionais de hidrocarbonetos ou em ambientes em que o risco de produção é muito grandes.



## 8. REFERÊNCIAS

---

- CAERS, J. (2011). *Modeling Uncertainty in Earth Sciences*. (J. Wiley, & Sons, Eds.) Hoboken, Nova Jersey, EUA: Wiley-Blackwell.
- DJURCILOV, S., KIM, K., LERMUSIAUX, P., & PANG, A. (2002). Visualizing scalar volumetric data with uncertainty. *Computers & Graphics* 26, pp. 239–248.
- HUNTER, G., & GOODCHILD, M. (1993). Managing Uncertainty in Spacial Databases: Putting Theory Into Praticce. *Journal of Urban and Regional Information Systems Association*, 5(2):55–62.
- LAMIGUEIRO, O. (2014). *Displaying time series, spatial, and space-time with R (The R Series)*. Boca Raton: CRC Press Chapman & Hall Book.
- LIGERO, E. L., RISSO, V. R., & SCHIOZER, D. J. (2008, Março). Análise econômica da redução de incertezas geológicas de reservatórios de Petróleo. *Revista Brasileira de Geociências*(38(1 - suplemento)), 106-116.
- LUCA, A., & TERMINI, S. (1975). A definition of nonprobabilistic entropy in the setting of fuzzy theory. *Information and Control* 20 (4), 301-312.
- MACEACHREN, A. M., ROBINSON, A., HOPPER, S., GARDNER, S., MURRAY, R., GAHEGAN, M., & HETZLER, E. (2005). Visualizing Geospatial Uncertainty: What We Know and What We Need To Know. *Cartography and Geographic Information Science*, 32, 139-160.
- MARQUES, M. (2012). *Metodologia para estimar valor de flexibilização no desenvolvimento de campos de petróleo*. Dissertação de mestrado, UNICAMP, Faculdade de Engenharia Mecânica e Instituto de Geociências, Campinas.
- PANG, A., WITERNBRINK, C., & LODHA, S. (1996). *Approaches to Uncertainty Visualization*. Paper, University of California, Computer Science Departament, California, USA.
- POTTER, K. (2010). *The Visualization Of Uncertainty*. Tese de Doutorado, University Of Utah, School Of Computing, Utah.

- POTTER, K., ROSEN, P., & JOHNSON, C. (2012). From quantification to visualization: a taxonomy of uncertainty visualization approaches. *In Uncertainty Quantification in Scientific Computing, IFIP Advances in Information and Communication Technology*, 377, pp. 226-249.
- RAIMOND, A. (2009). The communicative process of weather forecast issued in the probabilistic form. *Journal of Science Communication*, 8(1), 1-12.
- ROMEU, R. K., SCHIOZER, D. J., GHISI, I. T., GASPAR, A. S., SANTOS, A. A., GALLIN, A., & ARMSTRONG, M. (2013). Redução Ótima de Cenários e Outras Abordagens para Seleção de Modelos Representativos das Incertezas em Reservatórios de Petróleo. *Apresentação feita na UFRJ*, 2-5.
- SCHIOZER, D., LIGERO, E., SUSLICKA, S., COSTA, A., & SANTOS, J. (2004). Use of representative models in the integration of risk analysis and production strategy definition. *Journal of Petroleum Science and Engineering*(44), 131-141.
- SHANNON, E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423.
- TANJA, O. (2003). *Time-Lapse Seismic Within Reservoir Engineering*. Delft University of Technology, Department of Applied Earth Sciences. Drukkerij Augustijn, Enschede.
- WELLMANN, J. F., & REGENAUER-LIEB, K. (2012). Uncertainties have a meaning: Information entropy as a quality measure for 3-D geologicals models. *Tectonophysics*, 526-529, 207-216.
- WITTENBRINK, C. M.; PANG, A. T.; LODHA, S. K. (1996). Glyphs for visualizing uncertainty in vector fields. *IEEE Transactions on Visualization and Computer Graphics*, 2(3), pp. 266-279.
- ZADEH, A. (1965). Fuzzy Sets. *Information and control*, 8(3), 338-353.

# APÊNDICE

Neste Apêndice são apresentados os detalhes da construção dos modelos geológicos no *software* Petrel da Schlumberger que possui diversos recursos para montagem de modelos geológicos e possui interface de comunicação com o simulador desenvolvido por essa mesma companhia, o Eclipse.

## A. Criação do Modelo Fino

Primeiramente será mostrado, nas figuras a seguir, como é feita a criação de uma malha estrutural no Petrel. Essa malha é criada sem nenhuma propriedade, e será preenchida posteriormente.

Clica-se na opção *Define Model* e escolha um nome para o modelo, conforme na Figura A.1.

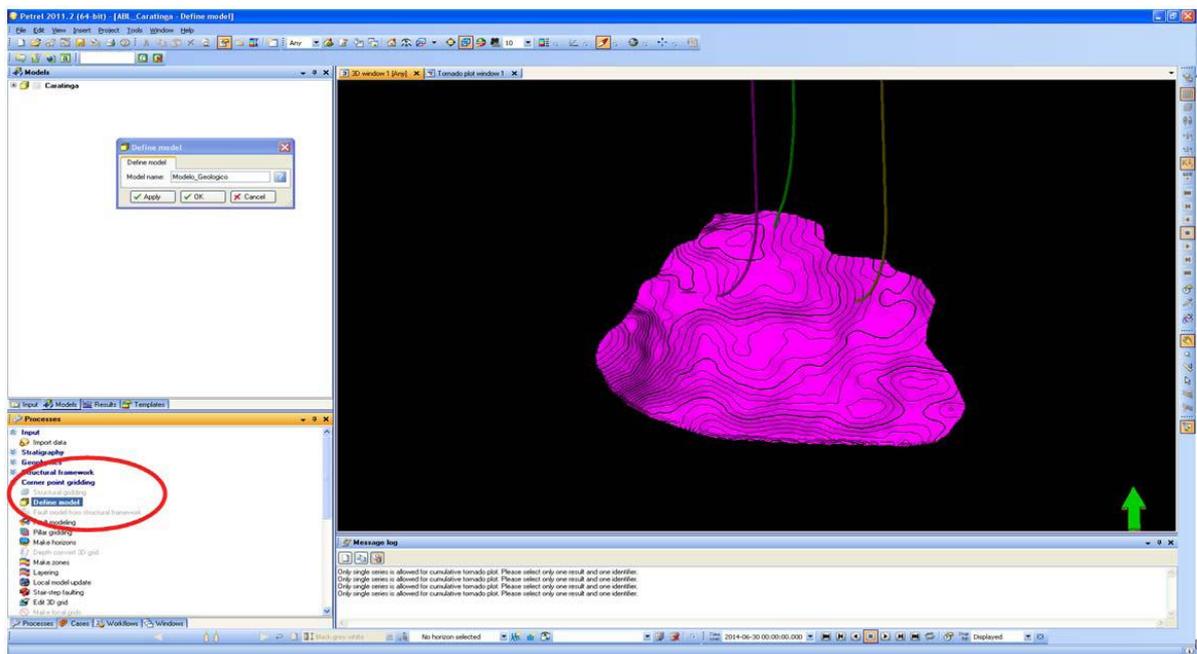


Figura A.1 - Criação do modelo fino

## B. Criação do Modelo de Falha

Para a definição do modelo de falha, clica-se na opção *Fault modeling*, conforme Figura B.1, e no modelo geológico criado no item anterior (no caso Modelo\_Geologico).

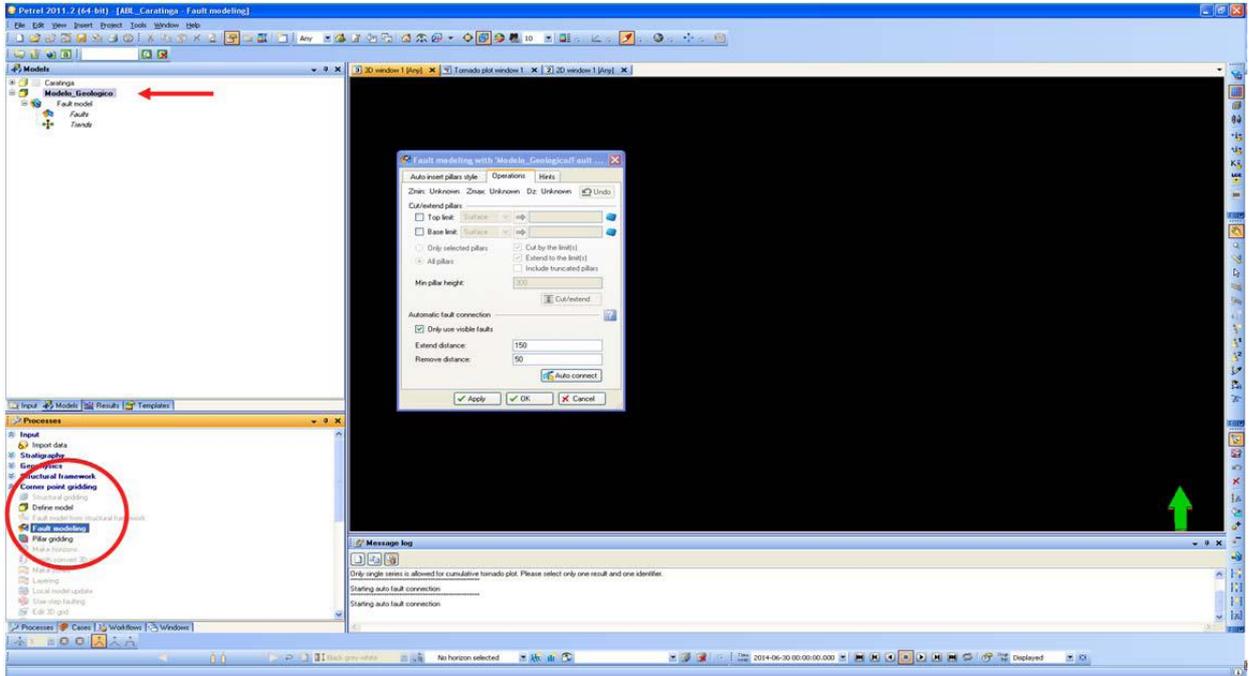


Figura B.1 - Criação do modelo de falha

A definição do *Fault modeling*, indicando os seus limites, é feito conforme a Figura B.2, através da indicação do Topo e da Base.

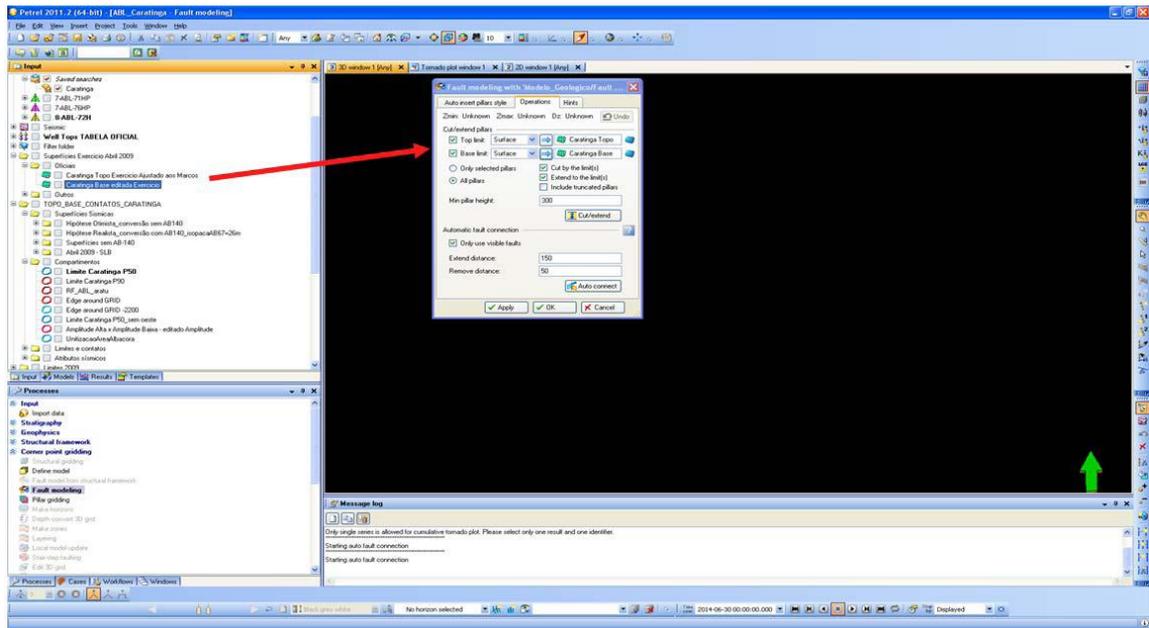


Figura B.2 - Definição do *Fault modeling*.

## C. Definição do *Pillar Malhading*.

A definição do *Pillar Malhading* é feita clicando-se na opção *Pillar Malhading*, conforme a Figure C.1. Estes pilares serão utilizados na definição dos vértices das células.

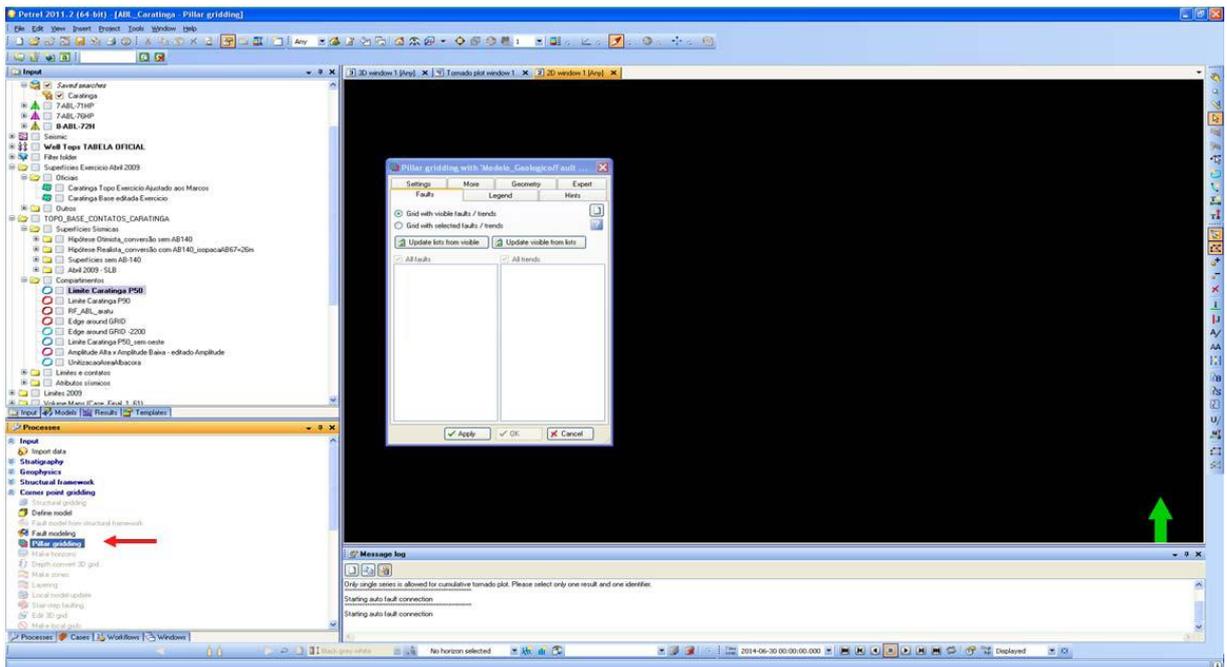


Figure C.1 - Definição do *Pillar Malhading*.

Na aba *Settings*, será dado o nome do modelo geológico e o tamanho da malha em I e J (50m x 50m) (Figure C.2).

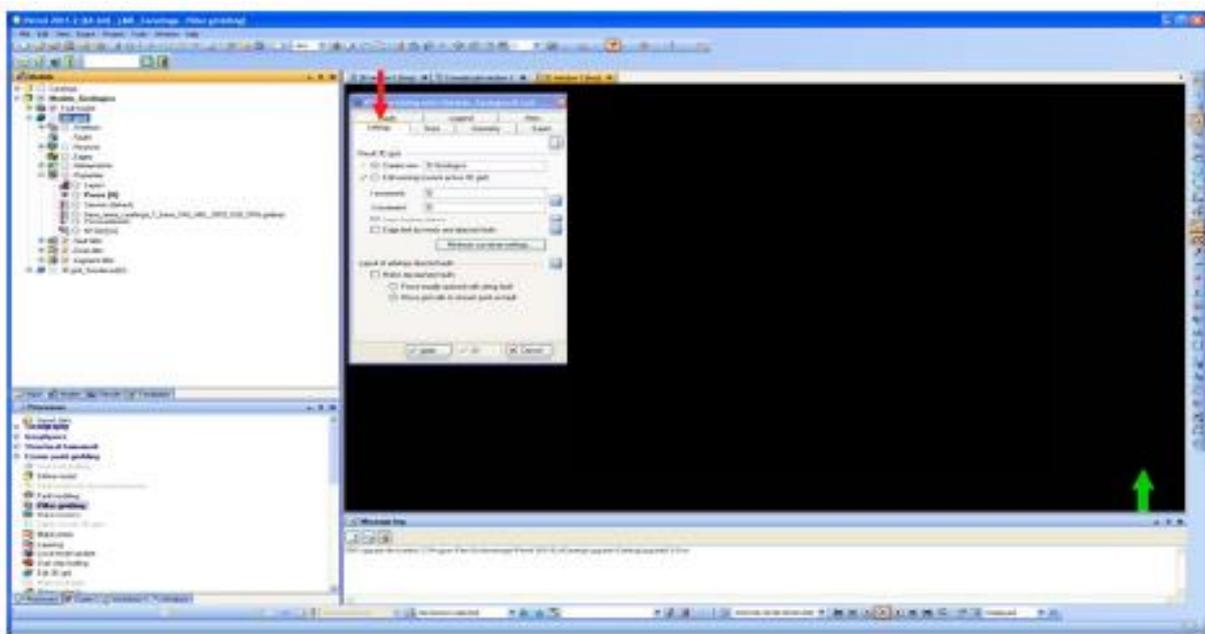
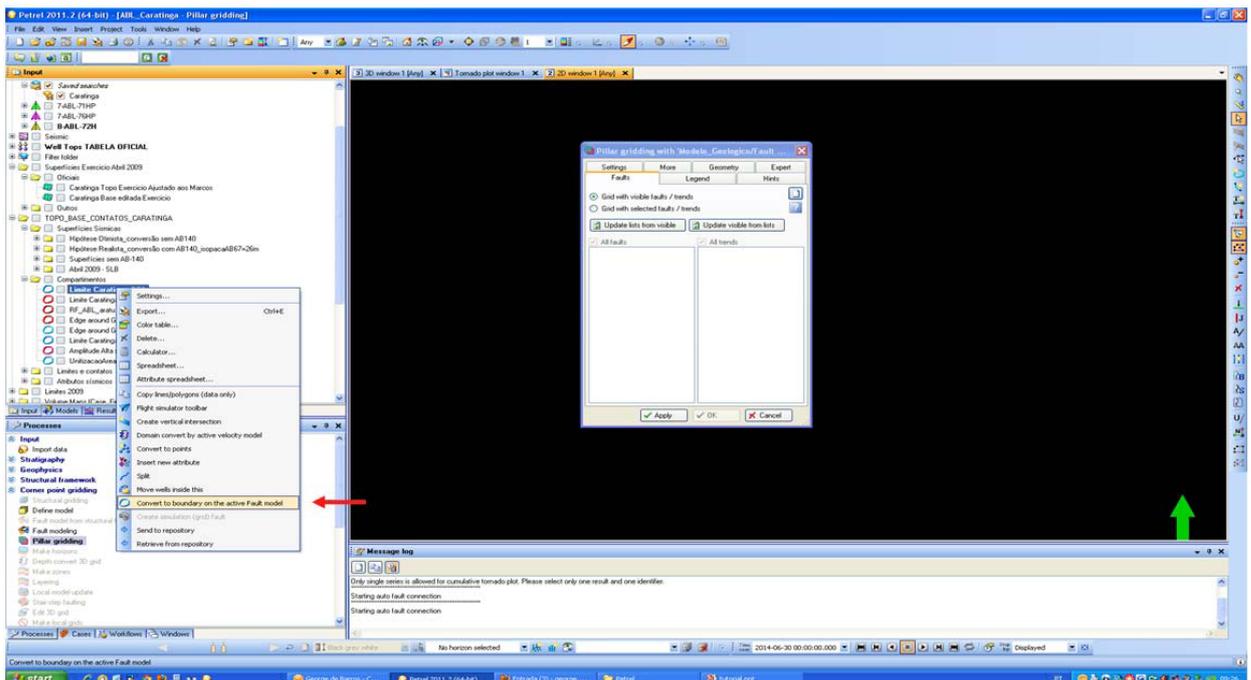


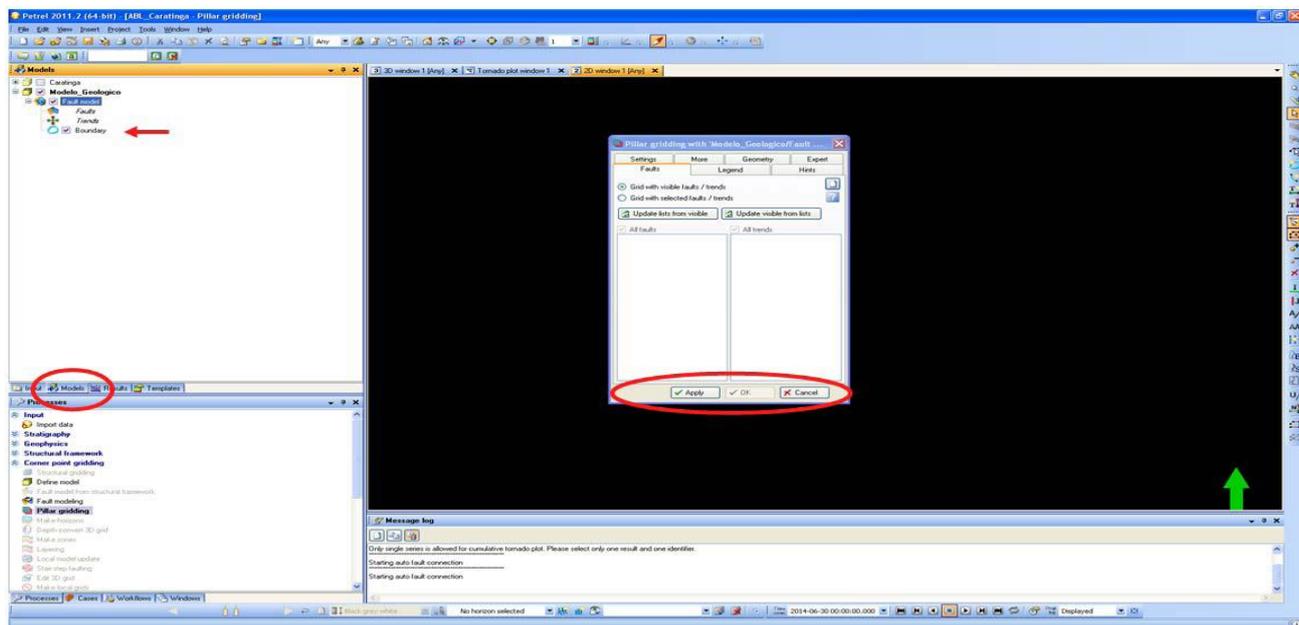
Figure C.2 - Definição do modelo geológico e do tamanho da malha do Pillar Malhading

Para definir o *Boundary* (fronteiras) do modelo, clique com o botão direito no objeto (Aba *Input*) e escolha a opção *Convert to boundary on the active Fault model* (Figura 8.6). Isso fará com que o *Boundary* seja criado na aba *Models* (Figure C.3).



**Figure C.3 - Definição dos limites do modelo**

Para finalizar, deve-se clicar primeiro em *Apply* e depois em *OK* (Figure C.4).



**Figure C.4 - Definição dos limites do modelo – confirmação.**

Desta maneira, pode-se visualizar a malha de referência criada, na qual se observa os pilares e o contorno do modelo (Figure C.5).

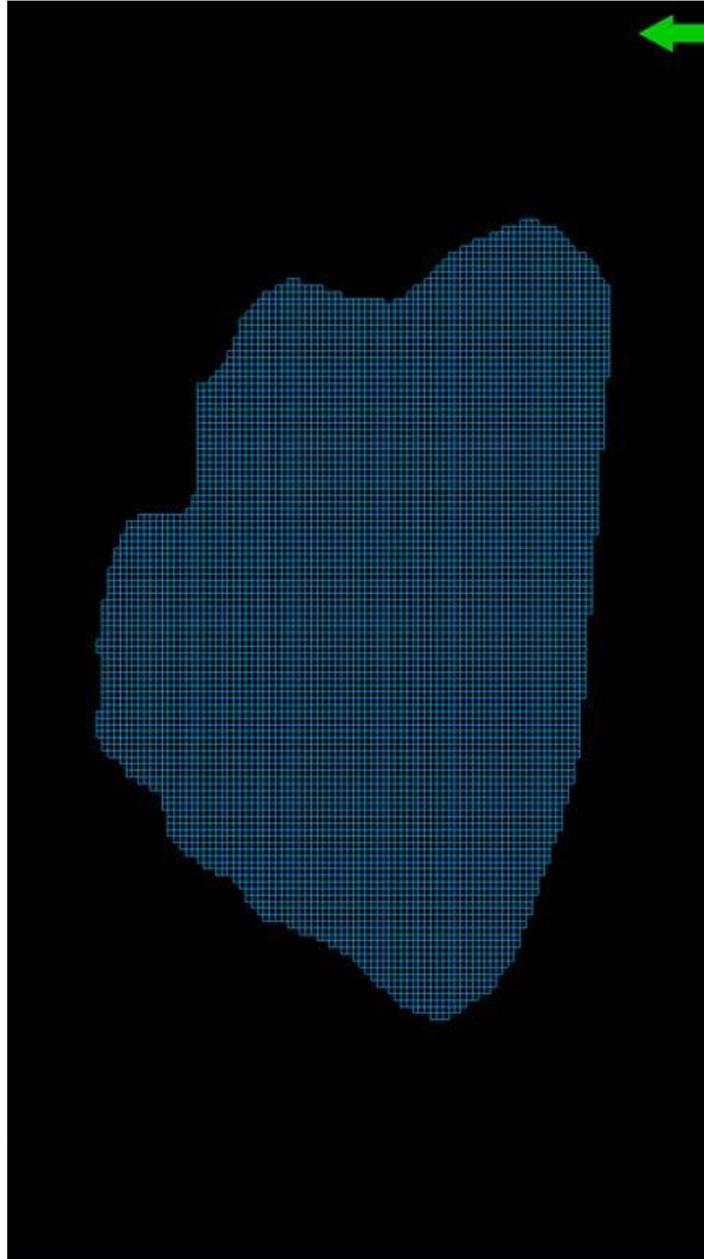


Figure C.5 - Visualização da malha de referência

## D. Definição dos Horizontes

Para a criação dos horizontes, que vão definir os limites superior e inferior do modelo geológico, deve-se informar as superfícies de topo e base do reservatório. Para isso, clique na opção *Make Horizons* (Figure D.1).

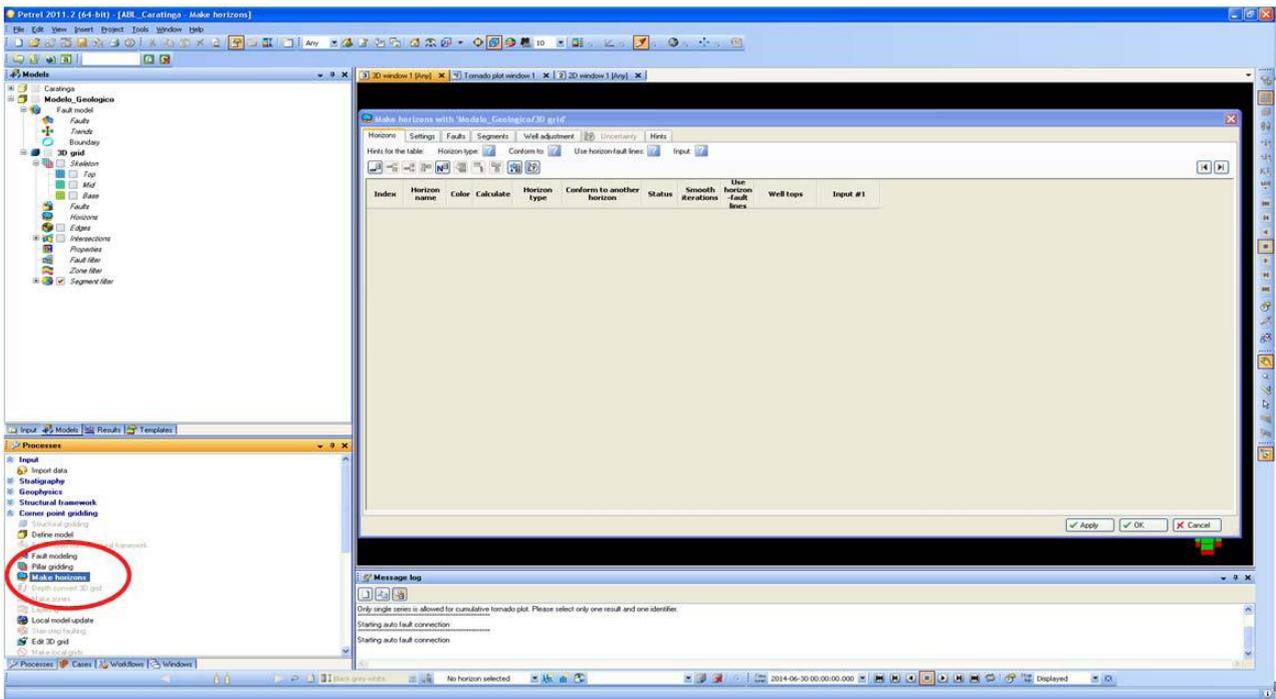


Figure D.1 - Definição dos Horizontes

Clique no botão assinalado (círculo menor) para incluir as superfícies de topo e base, que entram no item Input #1, indicado na seta (Figure D.2).

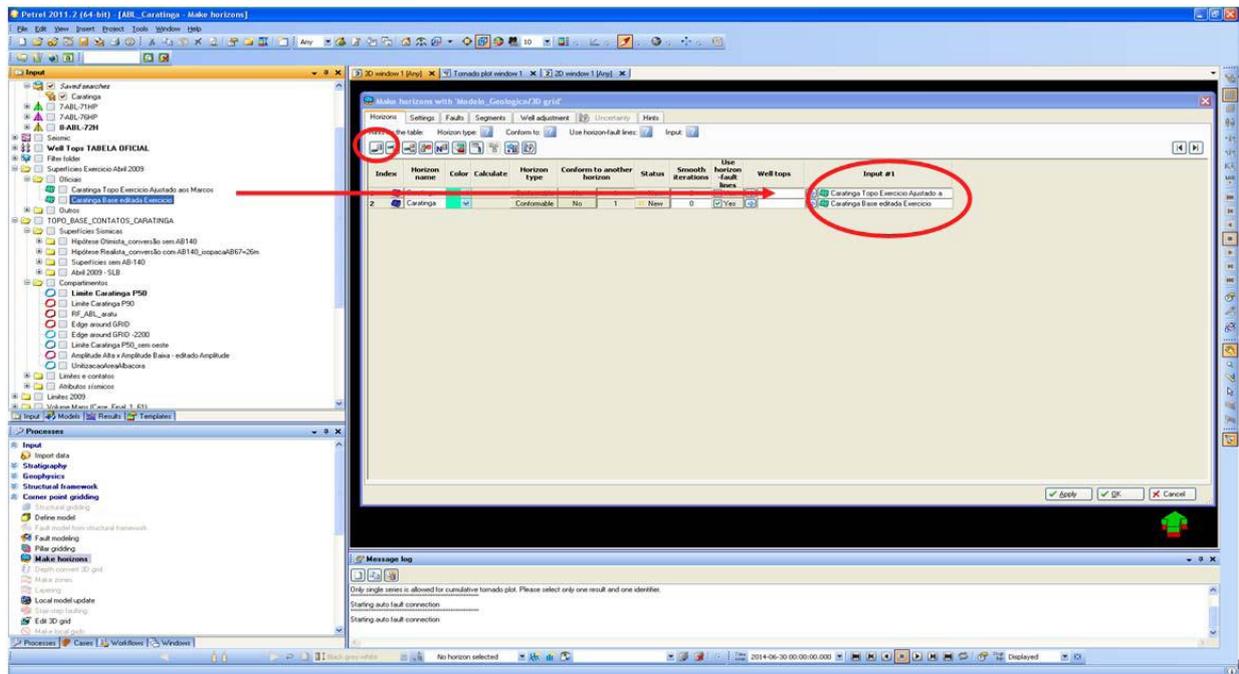


Figure D.2 - Inclusão das superfícies de topo e base

A seguir, associe os horizontes aos marcadores dos poços (que estão, neste exemplo, na aba *Input – Wells Tops TABELA OFICIAL*, pasta *Stratigraphy*). Depois clica-se no botão *OK* (Figure D.3).

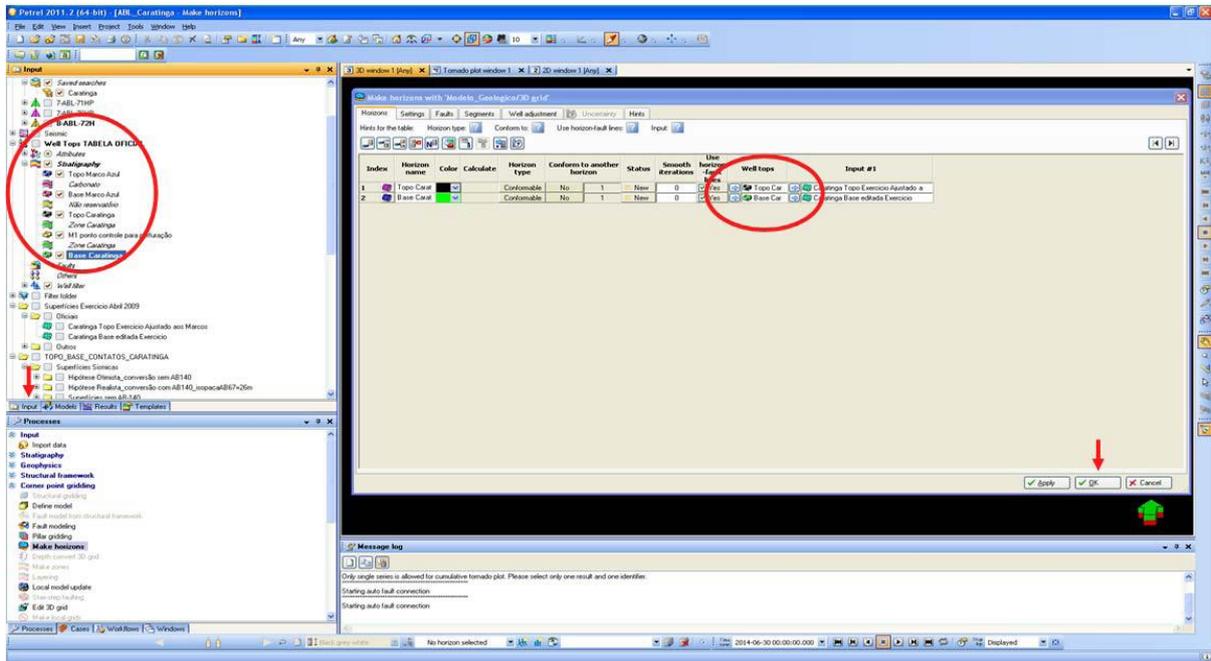


Figure D.3 - Associação dos marcadores dos poços

Finalmente os horizontes são criados no modelo geológico (Figure D.4).

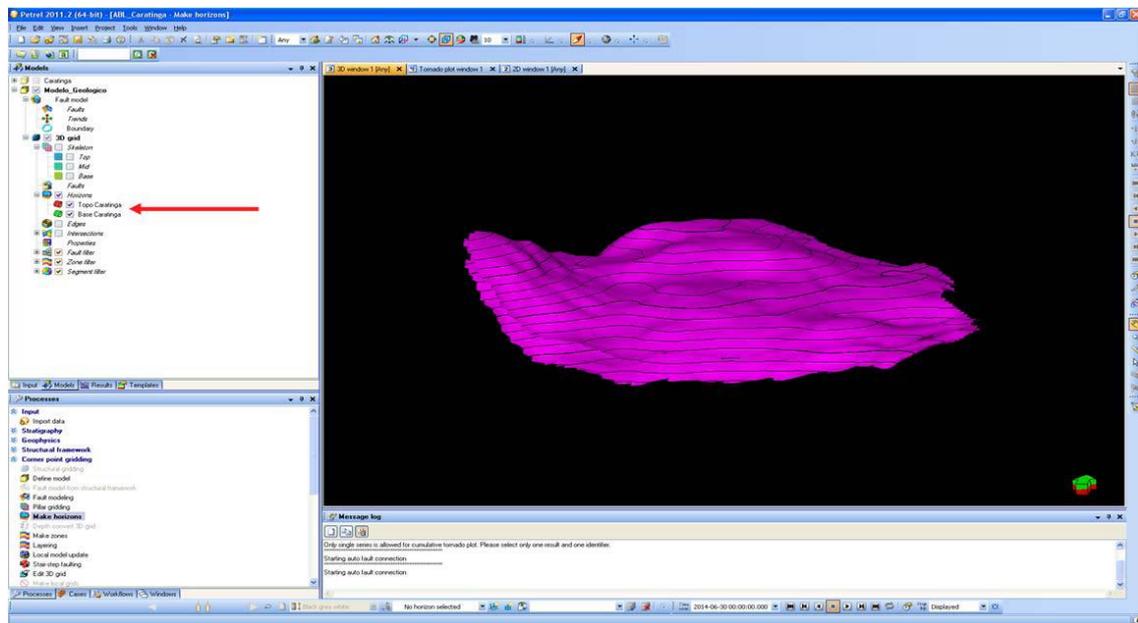


Figure D.4 - Final do processo de criação dos horizontes no modelo geológico.

## E. Definição das zonas do reservatório.

Neste modelo há somente uma zona, mas caso houvesse mais de uma, a opção *Make zones* deveria ser utilizada para se configurar diversas zonas (Figure E.1).

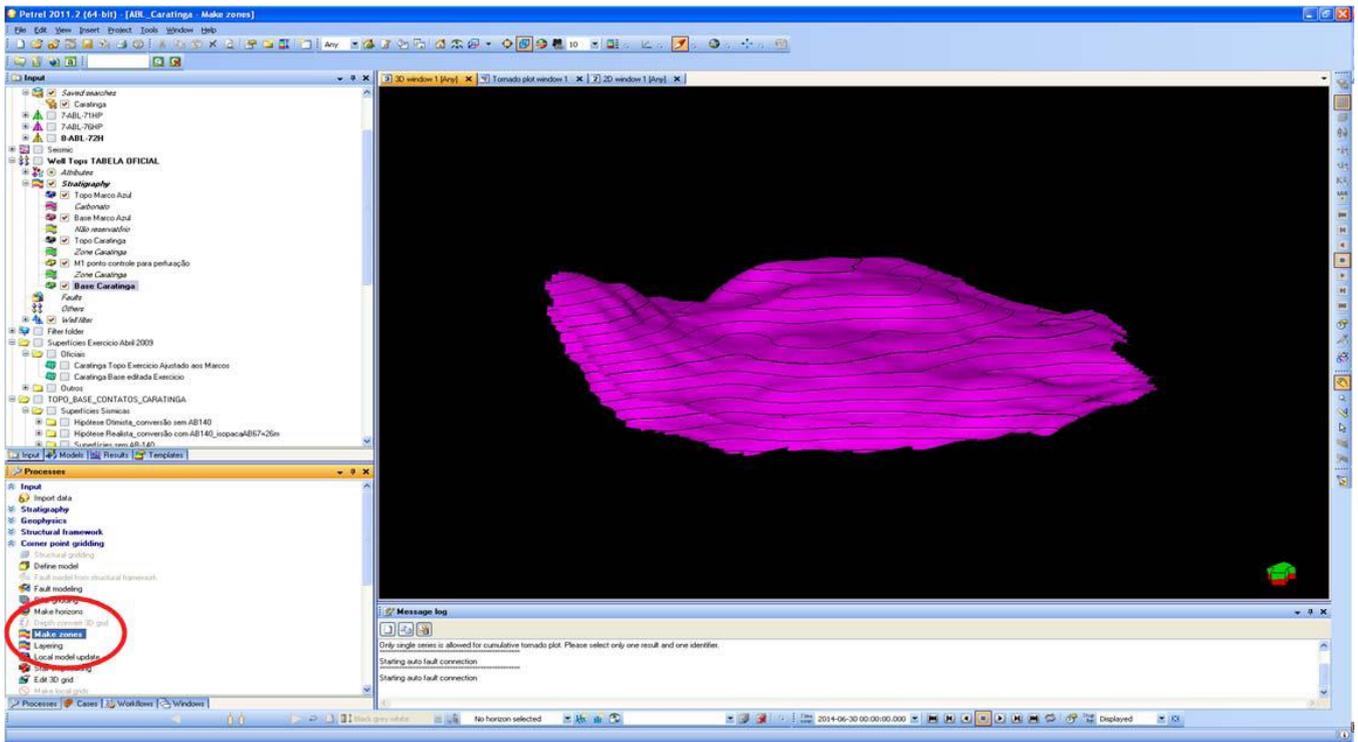


Figure E.1 - Definição das zonas do reservatório

## F. Definição do número de camadas (Layering)

A opção *Layering* é utilizada para a definição do número de camadas do modelo geológico (Figure F.1). Para este exemplo, foi utilizada a opção “Paralelo ao topo” (*Follow top*), com células de 1 metro de espessura.

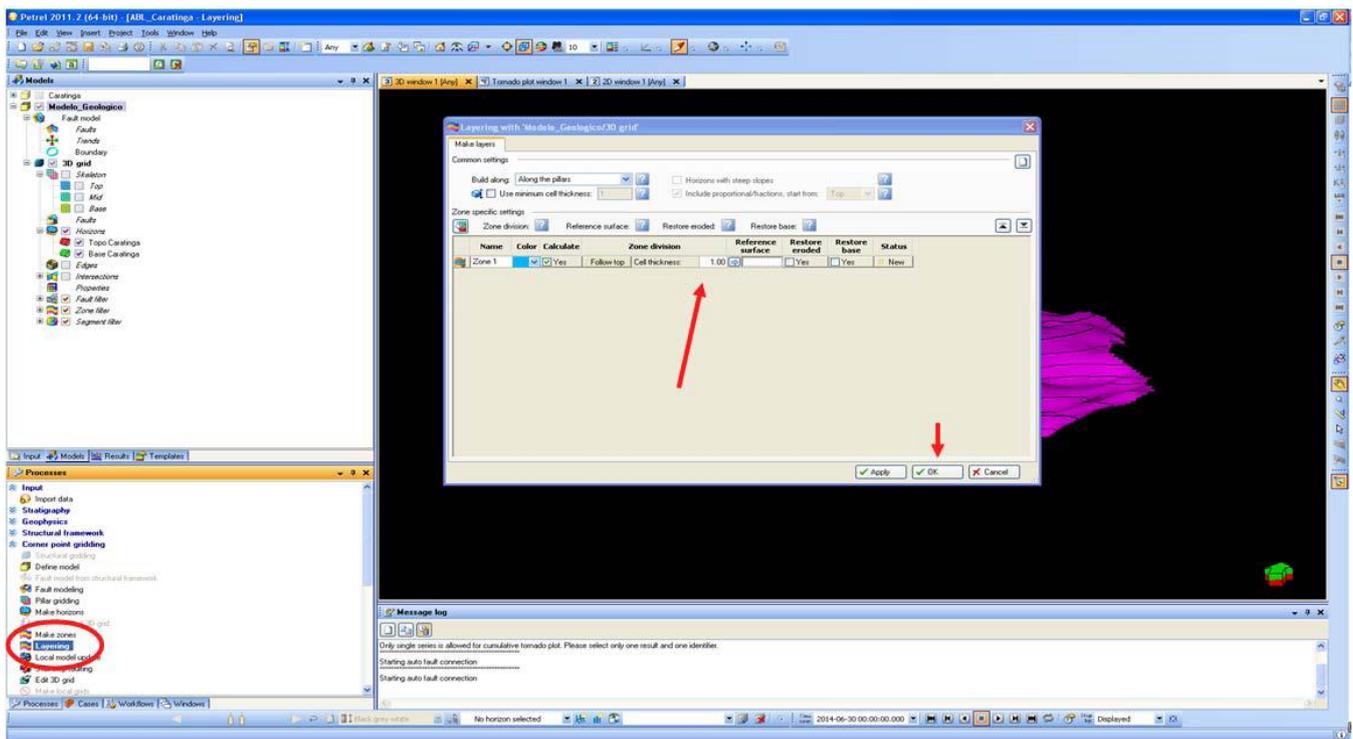
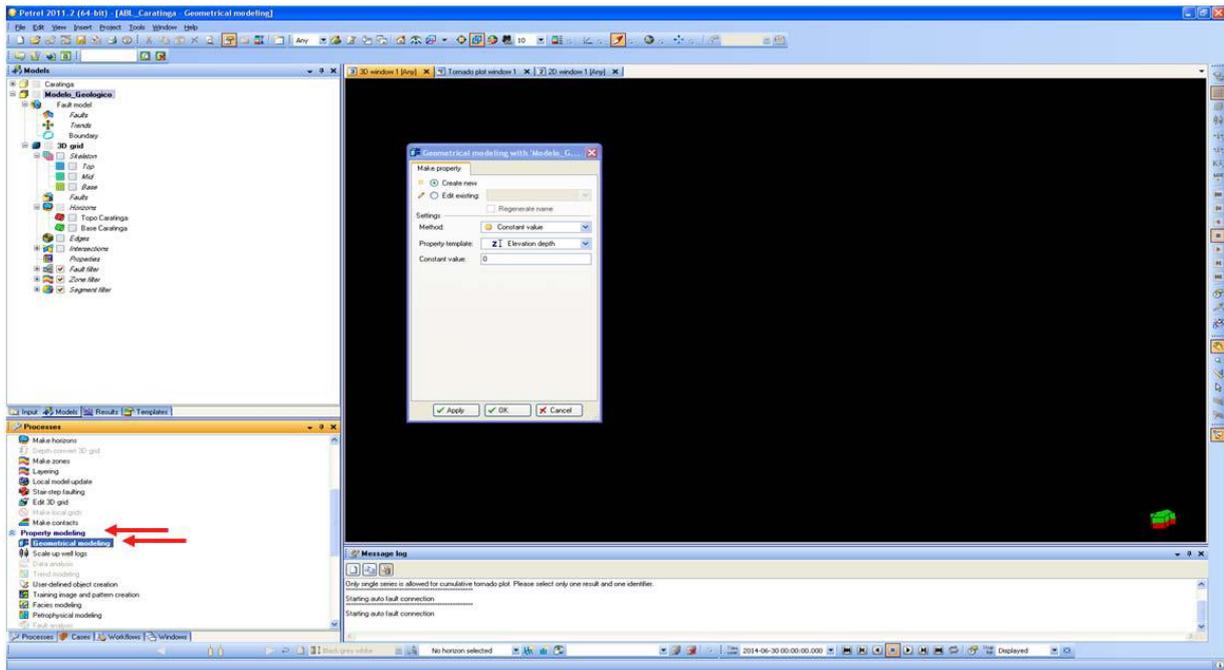


Figure F.1 - Definição do número de camadas

## G. Visualização de propriedades geométricas no Modelo Geológico

Para uma melhor visualização do modelo, é possível criar propriedades que representam características geométricas da malha, como por exemplo, os índices das camadas. Para isso, deve-se clicar em *Property modeling*, e depois em *Geometrical Modeling* (Figure G.1).



**Figure G.1 - Opção Geometrical Modeling**

Marque a opção *Zone Index* em *Method* (seta laranja na Figure G.2) e selecione a opção *From all Layers (K)* (seta vermelha). Será criada uma propriedade no modelo com o índice da camada (Figure G.3). Por convenção do Petrel, a primeira camada (K=1) é a superior, coincidente com a convenção utilizada na simulação numérica.

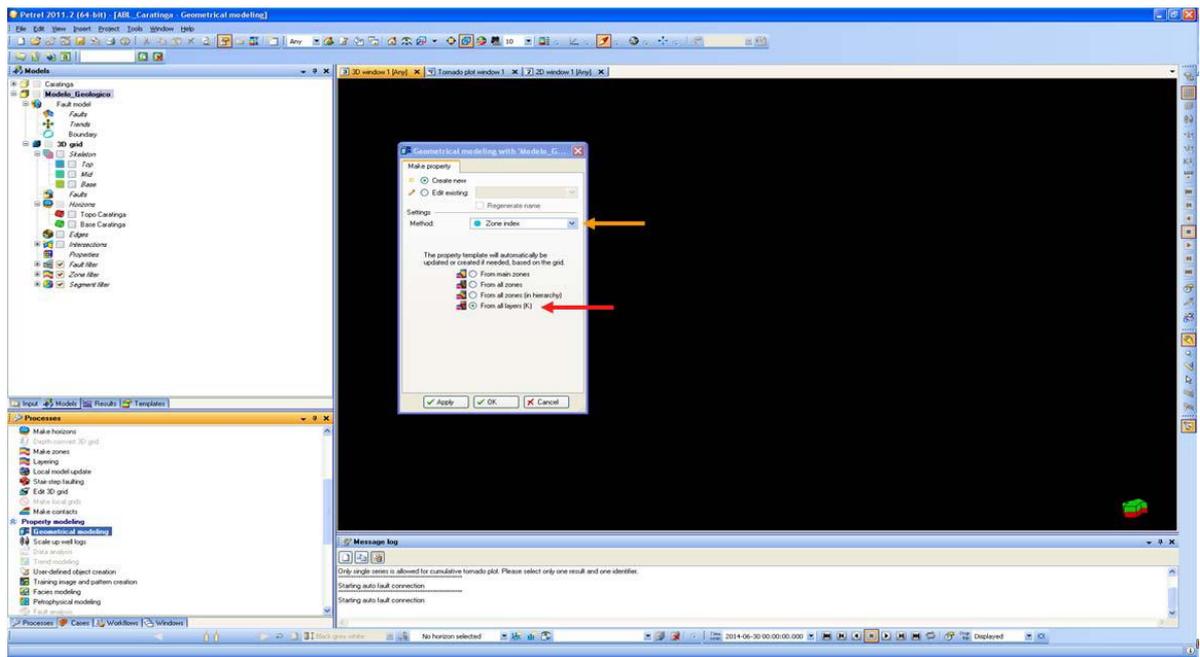


Figure G.2 - Criação de propriedade em *Geometrical Modeling*.

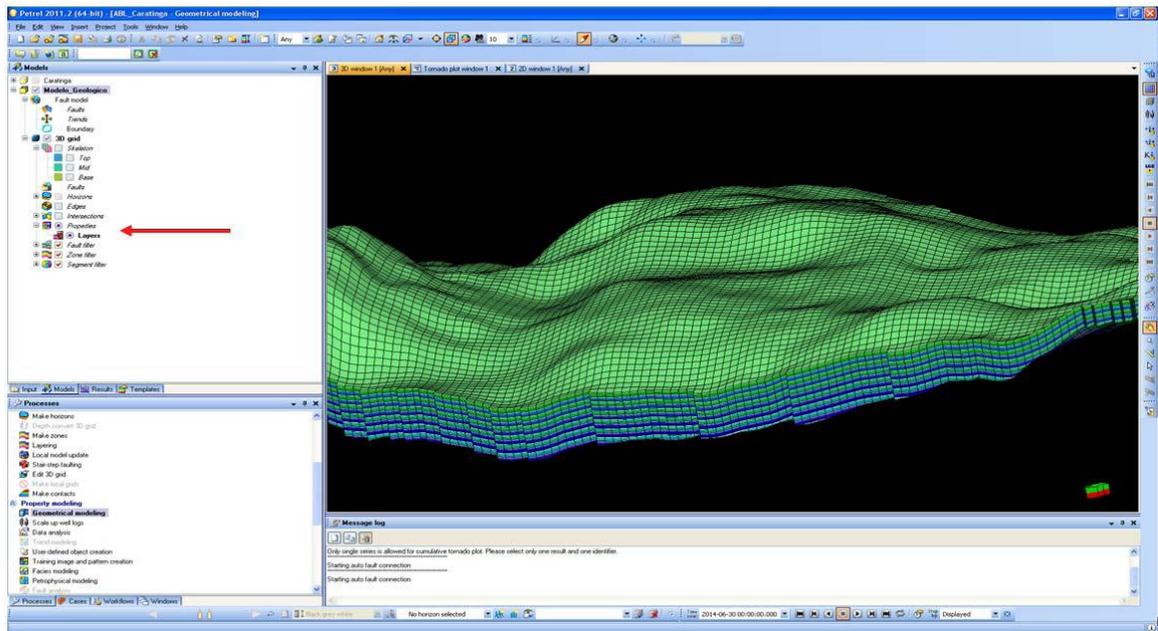


Figure G.3 - Visualização dos índices das camadas.

## H. Utilizando as informações de *logs* dos poços.

Para utilizar as informações dos perfis dos poços, deve-se transferi-las para o modelo geológico. Na aba *Process*, em *Property Modeling*, clica-se no processo *Scale up well logs* (Figure H.1).

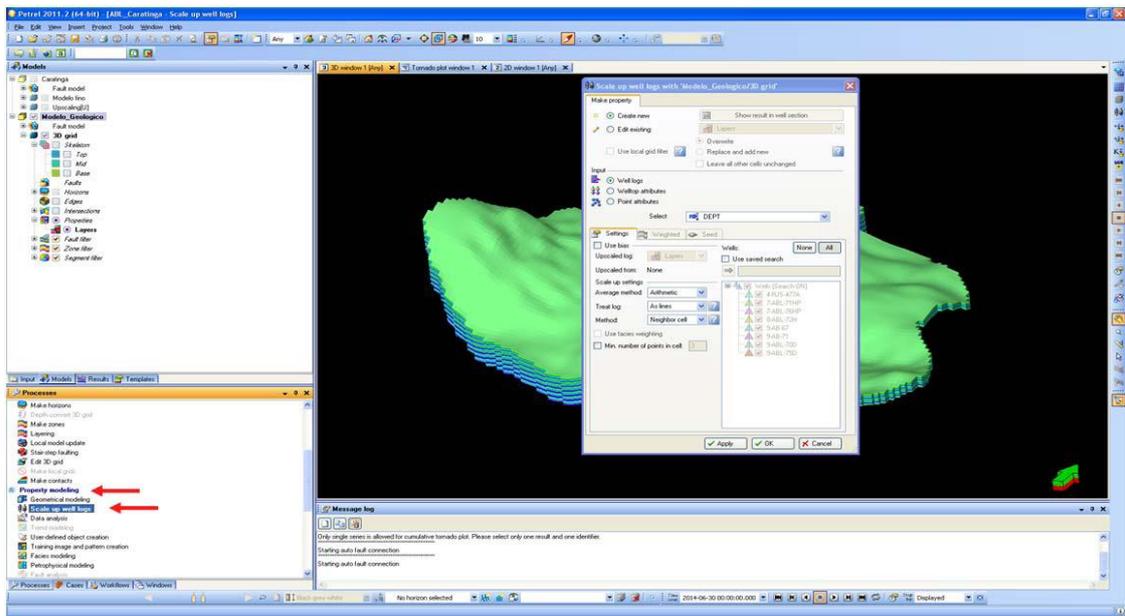


Figure H.1 - Opção Scale up well logs

No *Scale up* da Porosidade (Figure H.2), a informação na escala de poço (20 cm) será transferida para a escala do modelo (50m x 50m x 1m), **utilizando média aritmética.**

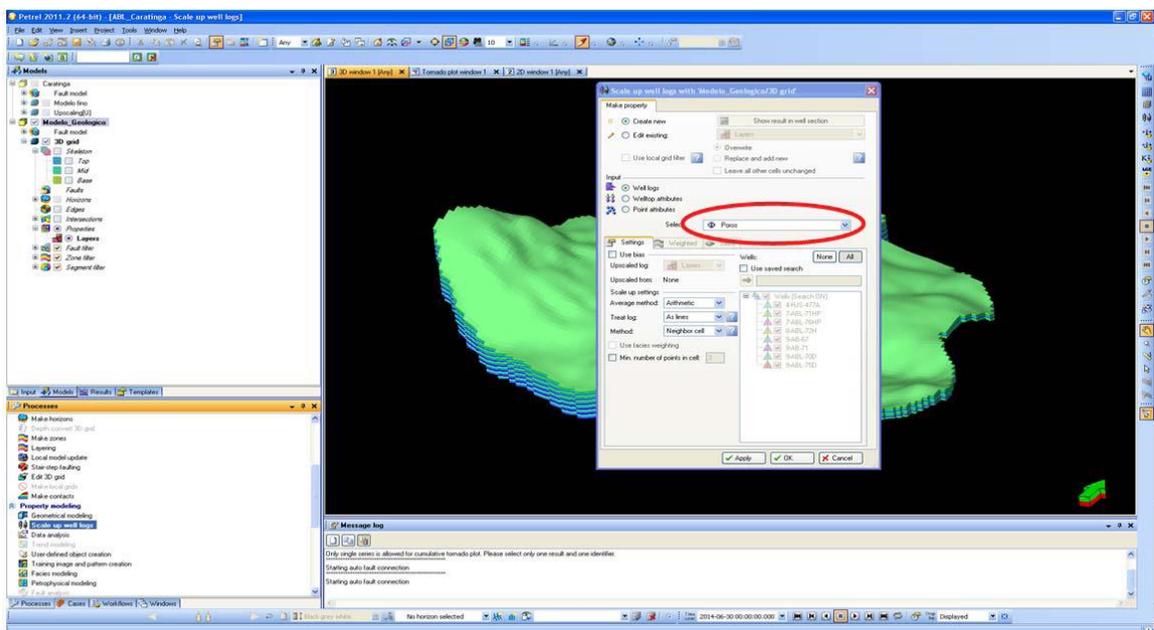


Figure H.2 - Seleção das informações de logs dos poços

Para a visualização dos dados de porosidade, clique na propriedade que foi criada no modelo (Figure H.3). Observe que essa propriedade só ocorre na área dos poços.

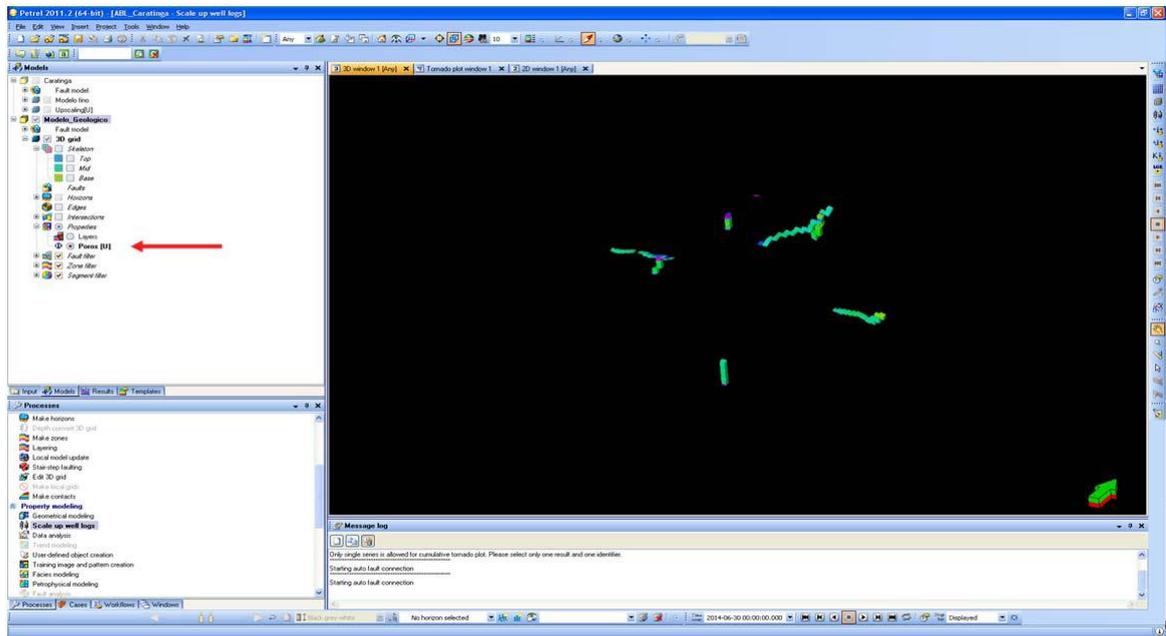


Figure H.3 - Visualização dos dados de porosidade no modelo

## I. Transferindo dados sísmicos

A informação do cubo de amplitude sísmica será utilizada como variável secundária no processo de modelagem da porosidade. Desta forma, é necessário que essa informação esteja na mesma escala do modelo geológico. Para transferir a amplitude da escala sísmica para a escala do modelo, em *Property modeling*, clique em *Geometrical Modeling*. Na janela aberta, escolha o método *Seismic Resampling* e selecione a sísmica na aba input, conforme a seta da Figure I.1.

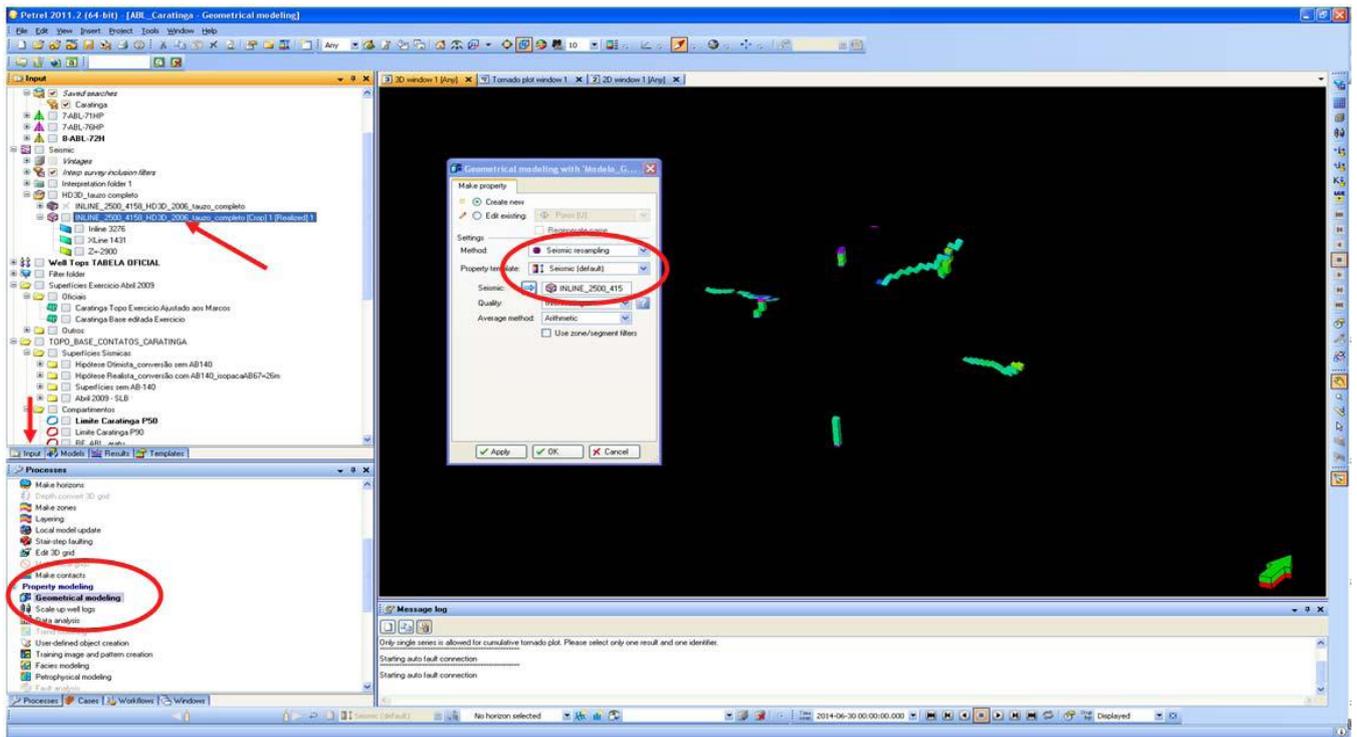


Figure I.1 - Transformando dados da escala da sísmica

Para visualização da sísmica no modelo geológico, selecione a nova propriedade criada no modelo (Figura 8.22). Para a modelagem da razão NTG, será utilizado um mapa 2D de amplitude da base do reservatório. Para a transferência deste mapa para a escala do modelo, deve-se clicar em *Geometrical Modeling* e selecionar o método *Constant or surface in segments and zones* e alterar o template para *Red white blue*. Em seguida, selecionar *All zones* e também o mapa do atributo escolhido para transferência na aba *Input* (*base\_areia\_caratinga\_1\_base...*), identificada pela seta em azul na Figure I.2. Dessa forma, o mapa sísmico foi replicado a todas as camadas do modelo geológico.

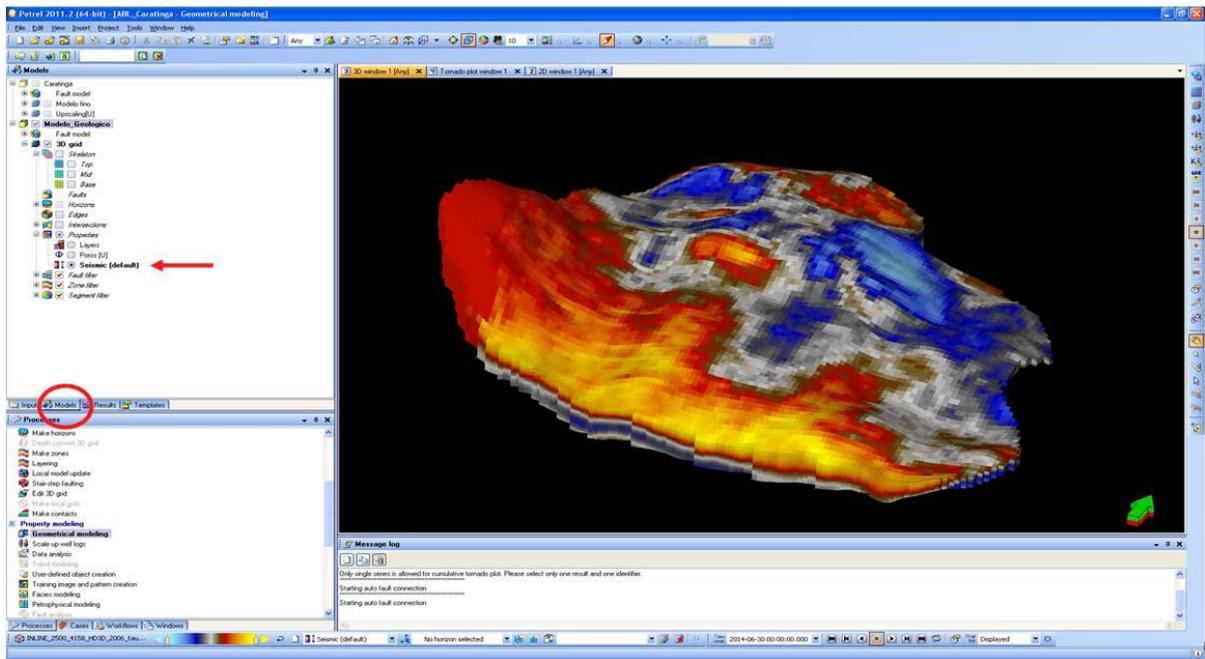


Figure I.2 - Visualização da sísmica no modelo geológico

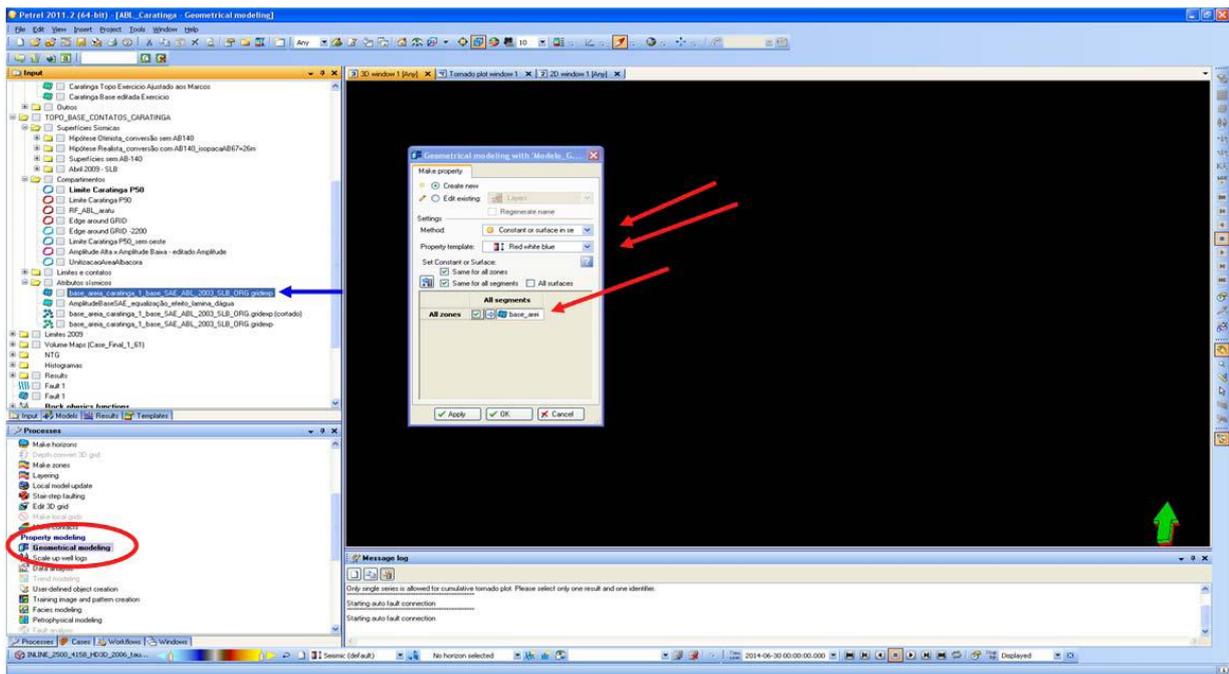


Figure I.3 - Transferência de um mapa de atributo sísmico ao modelo

O atributo criado no modelo é visualizado na Figure I.4.

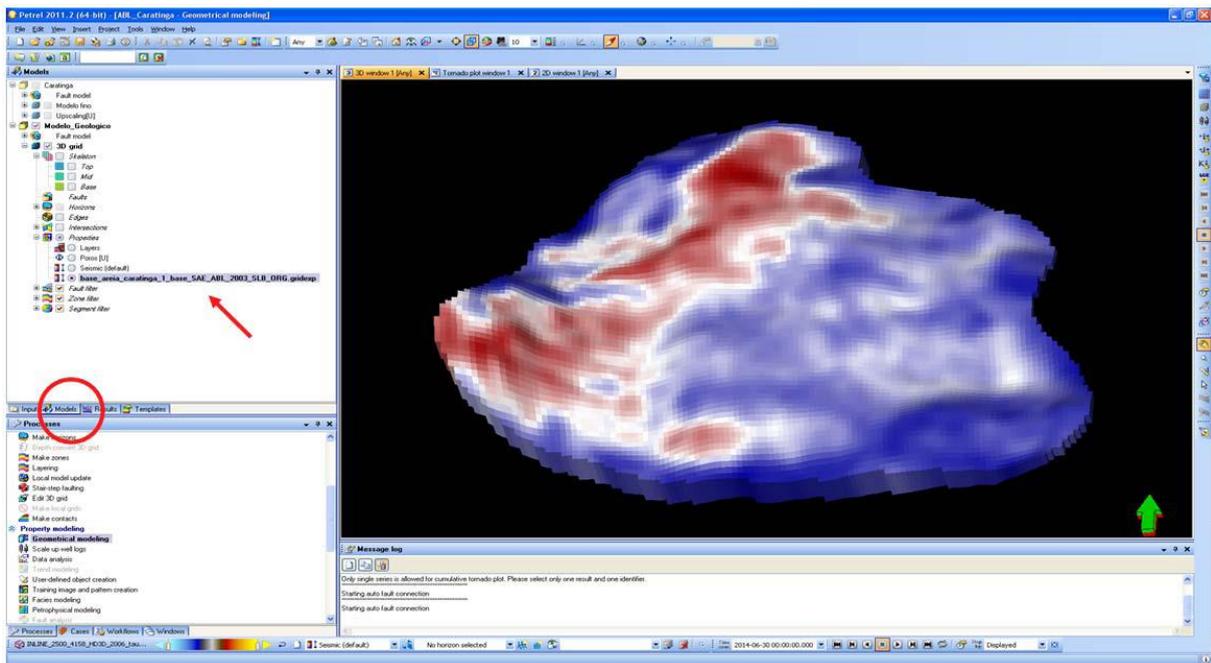


Figure I.4 - Visualização do atributo sísmico no modelo geológico criado

## J. Modelo Petrofísico de Porosidade

Para efetuar a distribuição da Porosidade no modelo, na opção *Property Modeling*, clique em *Petrophysical Modeling* (Figure J.1). Na janela aberta, selecione a Porosidade como propriedade e libere o cadeado (indicado pela seta em laranja) para permitir a edição dos parâmetros de simulação.

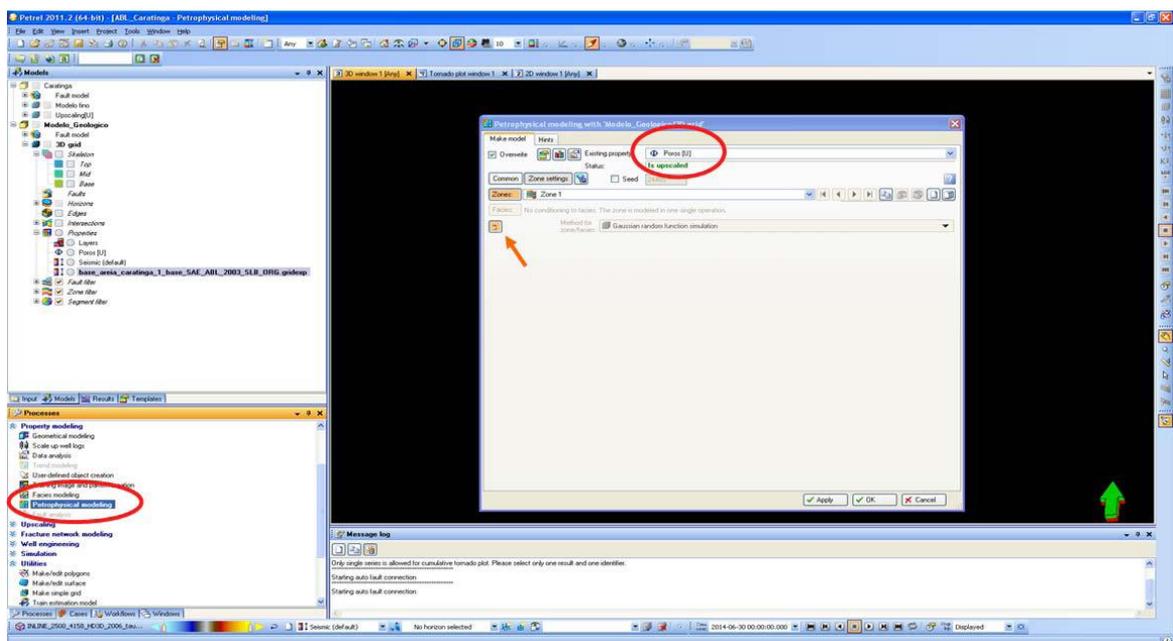


Figure J.1 - Opção Petrophysical Modeling para porosidade

Com o cadeado clicado, aparecem as opções em várias abas (Figure J.2) a serem configuradas.

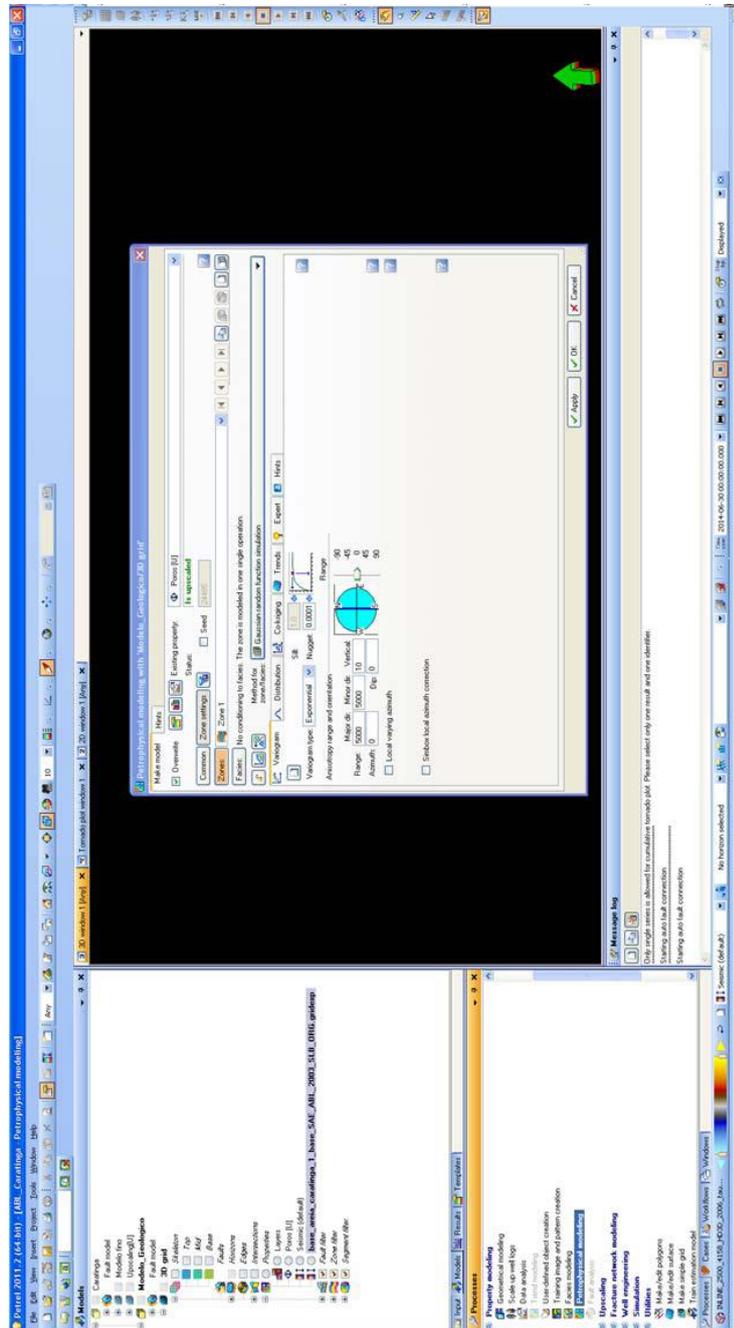


Figure J.2 - Detalhes de petrofísica com o cadeado aberto

Escolha o método *Sequential Gaussian Simulation* (Figure J.3).

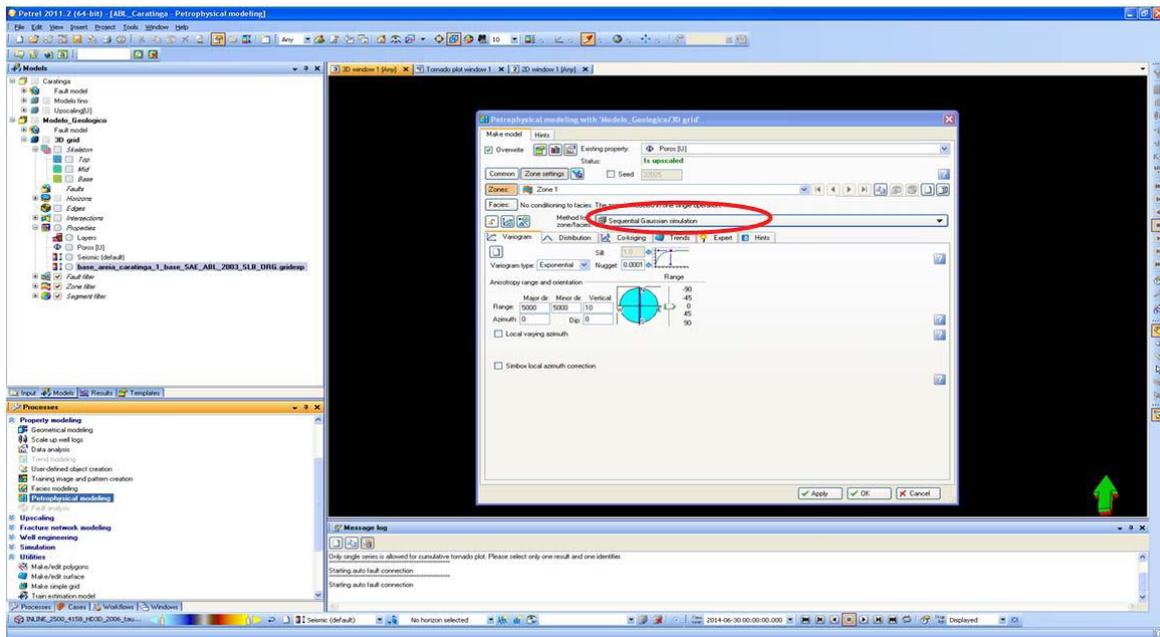


Figure J.3 - Método Sequential Gaussian Simulation

Depois, insere-se os demais parâmetros de variograma: *Nugget*, Range Maior, Range Menor, Range Vertical e Azimute, conforme.

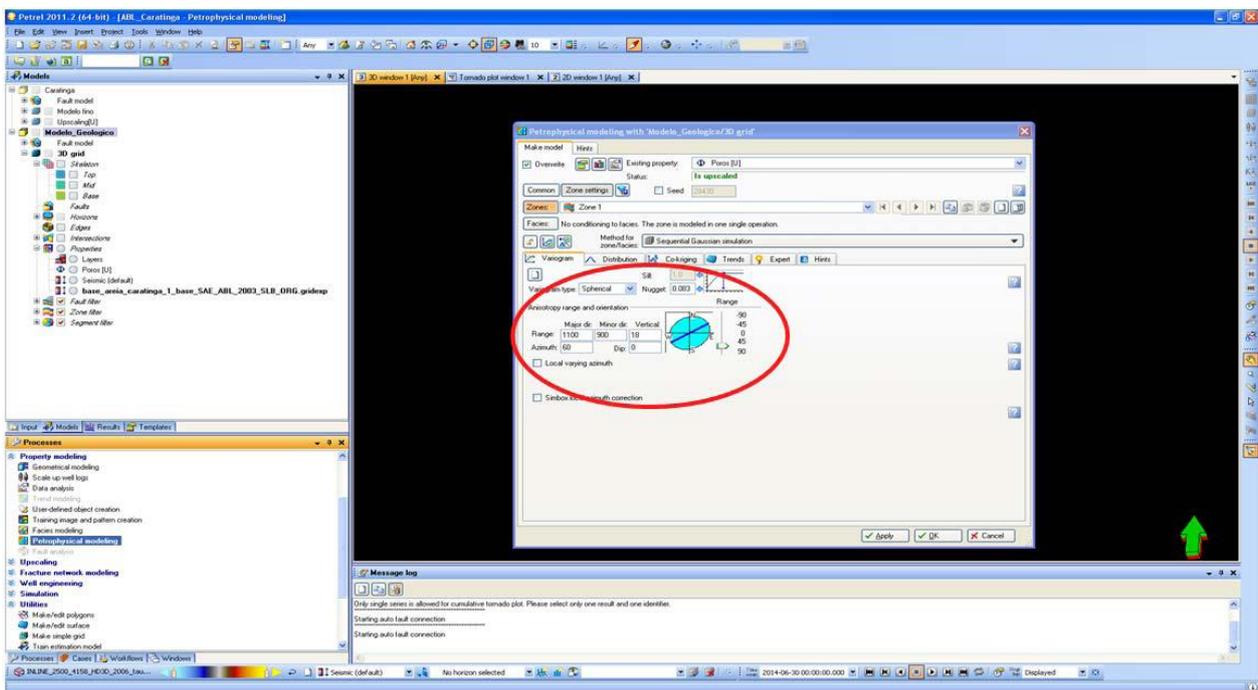


Figure J.4 - Parâmetros de variograma para a porosidade.

Na aba *Distribution* (Figure J.5), definem-se os valores mínimos e máximos da distribuição (pelo botão *Estimate*, em laranja) e o tipo de Distribuição: *Standard* e *From*

*upscaled logs* (setas em azul). Desta forma, a simulação vai reproduzir a distribuição de porosidade dos poços (min, max, média e desvio padrão).

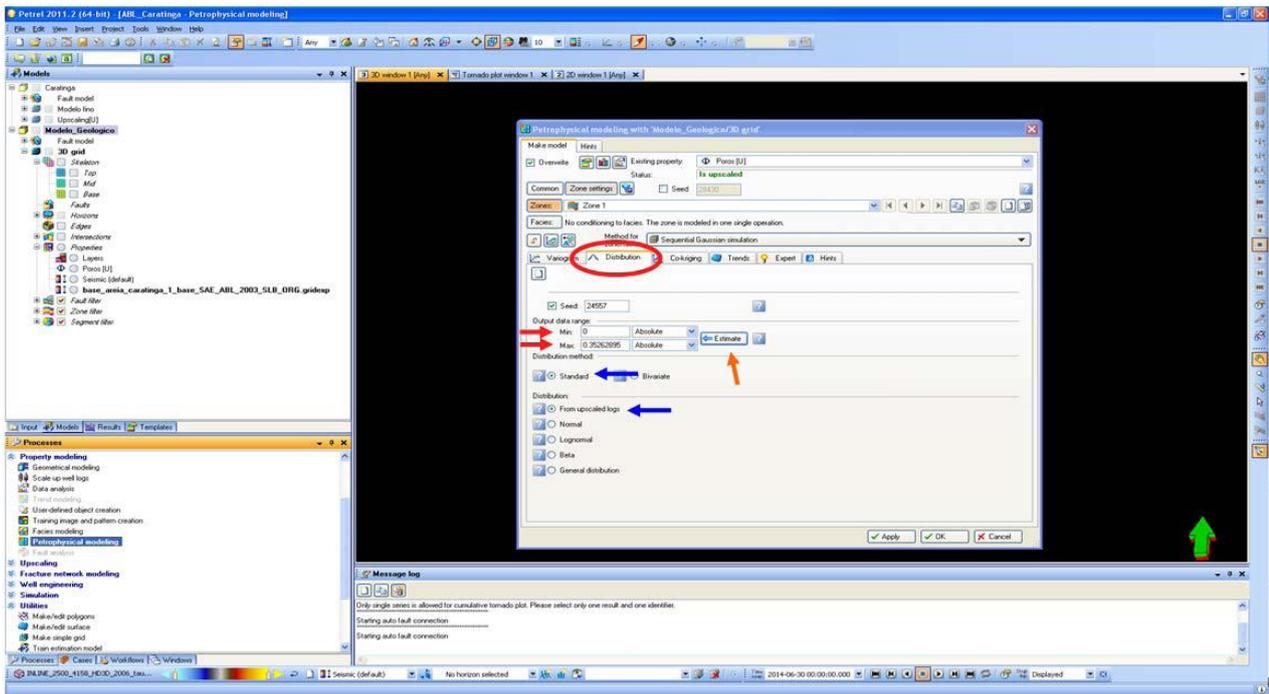


Figure J.5 - Detalhes da aba *Distribution* para porosidade.

Na aba *Co-kriging* (Figure J.6), define-se a propriedade que será a variável secundária, no caso a Sísmica (seta em azul). Em *Method*, marque a opção *Collocated-co-kriging*, e escolha a opção *Constant* para coeficiente (setas em vermelho). Depois, define-se um coeficiente de correlação ou o estime acionando o botão *Estimate* (setas em verde).

Finaliza-se clicando no botão *OK*.

As vantagens da co-krigagem são maiores quando a variável primária (NTG do poço) está sub-amostrada em relação à secundária (mapa da amplitude da base) e quando ambas apresentam alta correlação entre si. Mas uma elevada densidade amostral da variável secundária induz a que dados secundários mais próximos ao local onde se estima a variável primária filtrem a influência de dados mais afastados. Para evitar este efeito surgiu a variante do método de co-krigagem colocada, na qual é retido só o dado secundário colocado com o local onde se estima a variável primária.

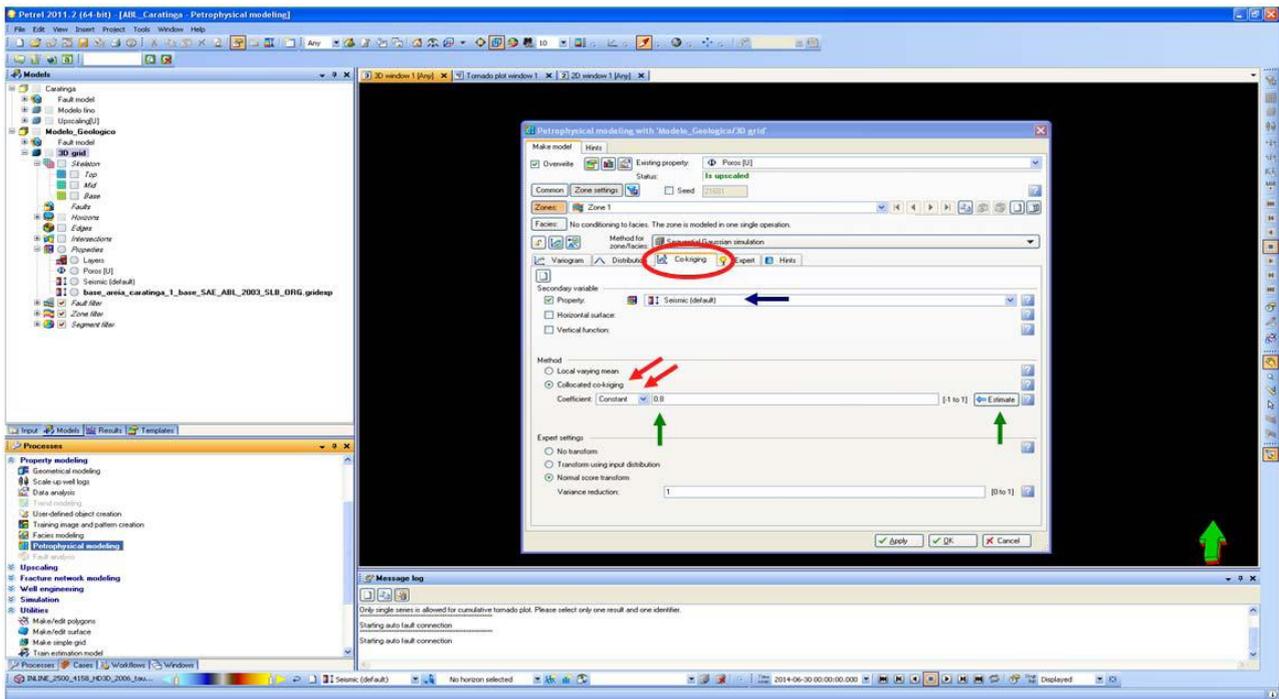


Figure J.6 - Detalhes da aba Co-kriging para porosidade.

Pode-se visualizar a porosidade estimada, conforme Figure J.7.

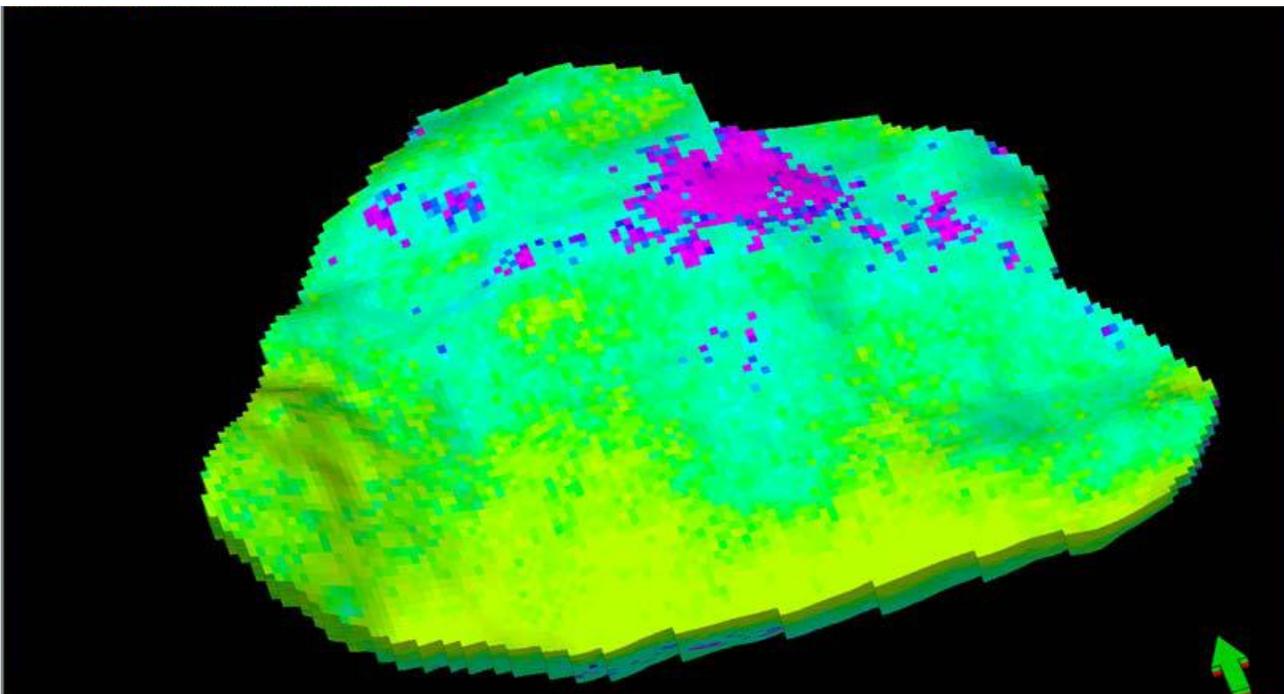


Figure J.7 - Visualização da porosidade estimada

## K. Modelo Petrofísico de NTG

A estimativa do NTG segue um fluxo similar. Selecione *Petrophysical modeling* e a propriedade NTG (Figure K.1). Escolha o método: *Sequential Gaussian Simulation*.

Para a definição dos parâmetros do variograma, utilize os mesmos da Porosidade. Na aba *Distribution* (Figure K.2), defina a distribuição *Sequential Gaussian simulation* e para valores Min: 0, Max: 1, Normal Mean: 0.89624 e Std: 0.24943.

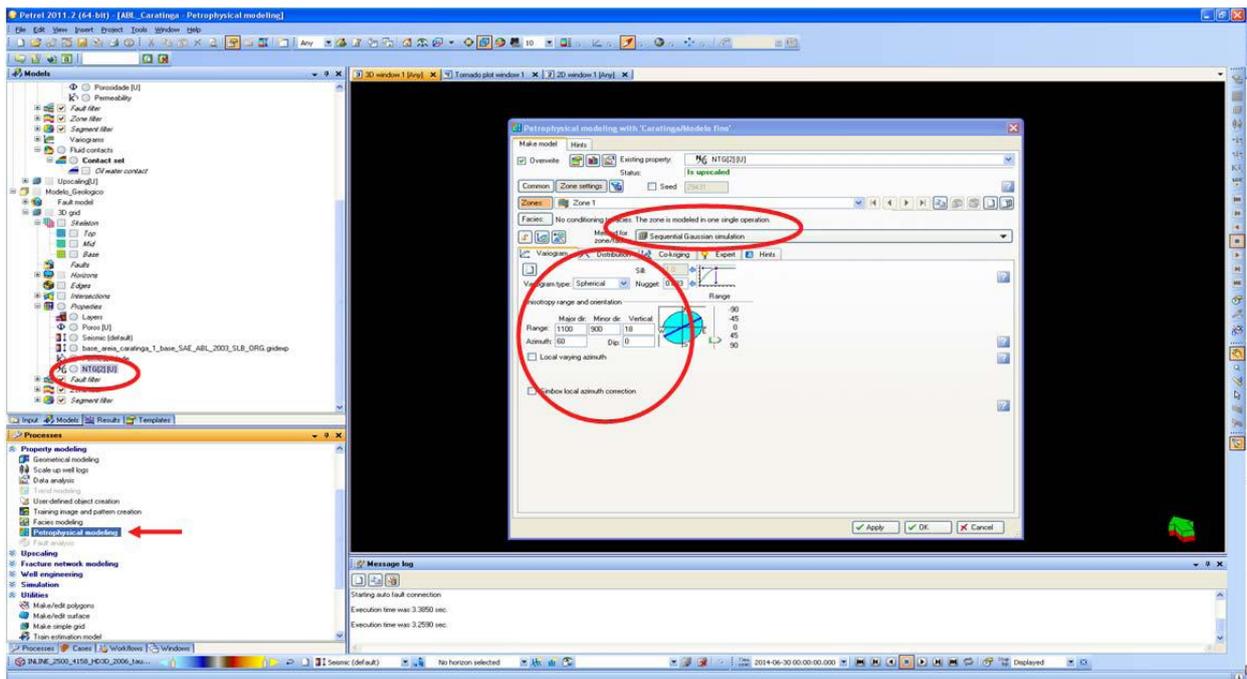


Figure K.1 - Opção Petrophysical Modeling para NTG

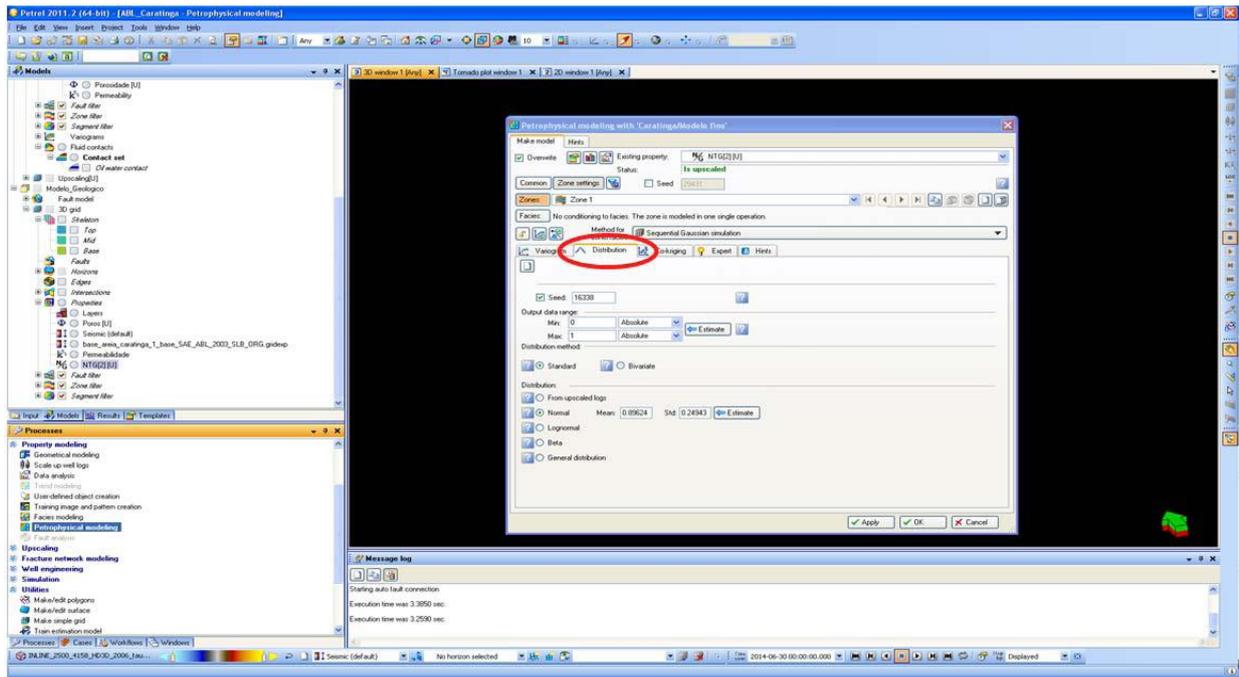


Figure K.2 - Detalhes da aba Distribution para NTG.

Na aba Co-kriging (Figure K.1), deve-se utilizar como variável secundária o mapa sísmico da base. É possível usar diretamente o próprio mapa da aba Input (base\_areia....malhaexp), através da opção *Horizontal surface*. Também se pode usar a propriedade já criada com a transferência do mapa ao modelo (Figure K.2), através da opção Property. Como coeficiente de correlação, foi utilizado o valor de -0.7.

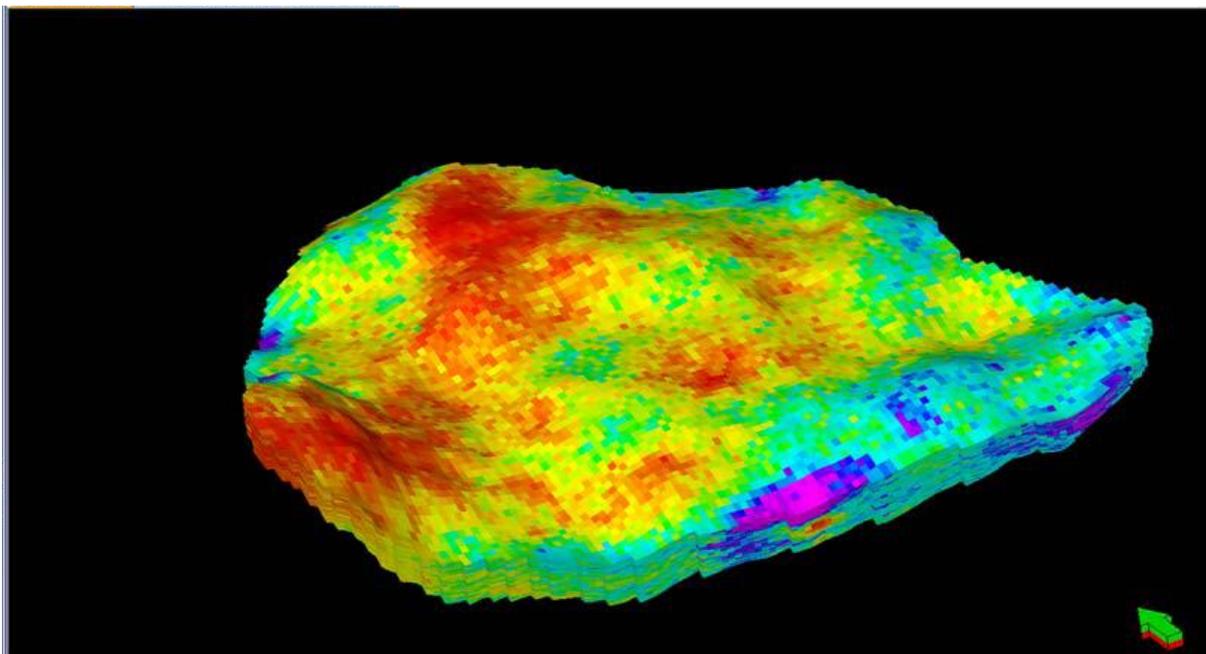


Figure K.3 - Visualização do NTG estimado.

## L. Modelo Petrofísico de Permeabilidade

Uma vez que, para este caso, a Permeabilidade não foi calculada nos poços, deve-se primeiro criá-la no modelo, utilizando a calculadora. Na aba *Models*, clique com o botão direito em *Properties*, e selecione *Calculator*. Na janela aberta, insira Permeabilidade = U (este U significa *undefined*) e altere o *template* para *Permeability*,

conforme a figura. Ao se clicar em *ENTER*, será criada uma propriedade no modelo sem nenhum valor (Figure L.1).

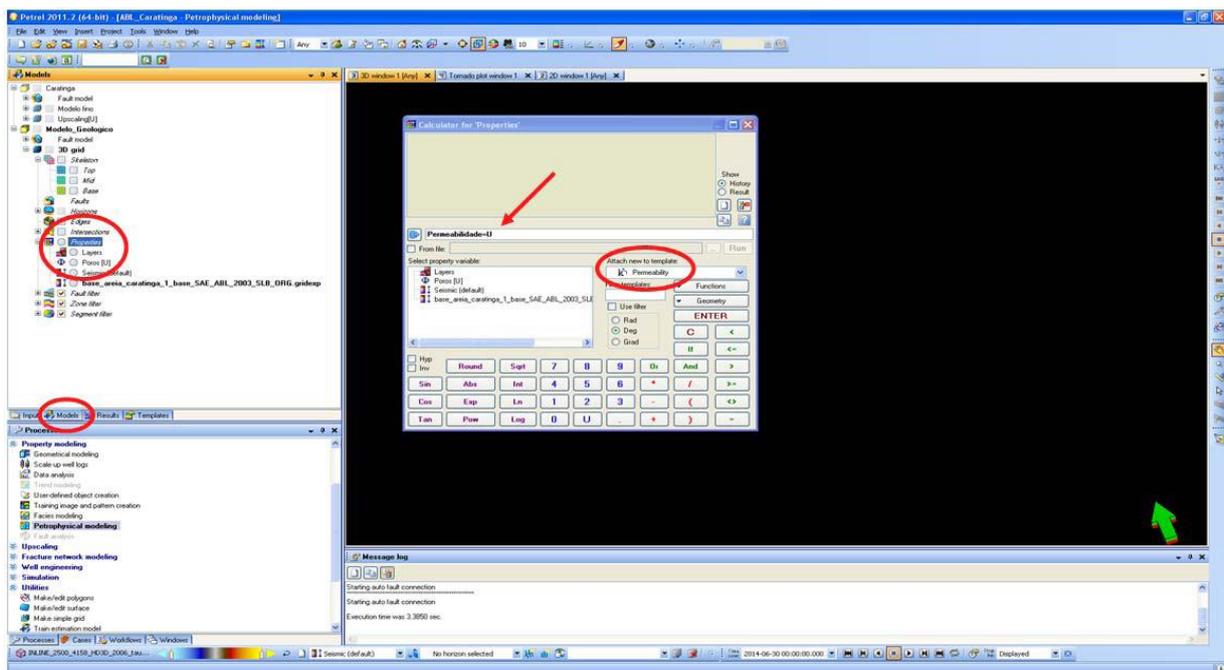


Figure L.1 - Criação da propriedade Perm utilizando a calculadora.

Com a propriedade criada, agora se pode fazer uma estimativa da permeabilidade pela opção *Petrophysical Modeling*, de forma similar à feita para a porosidade. Escolha a propriedade Permeabilidade e depois abra o cadeado (Figure L.2).

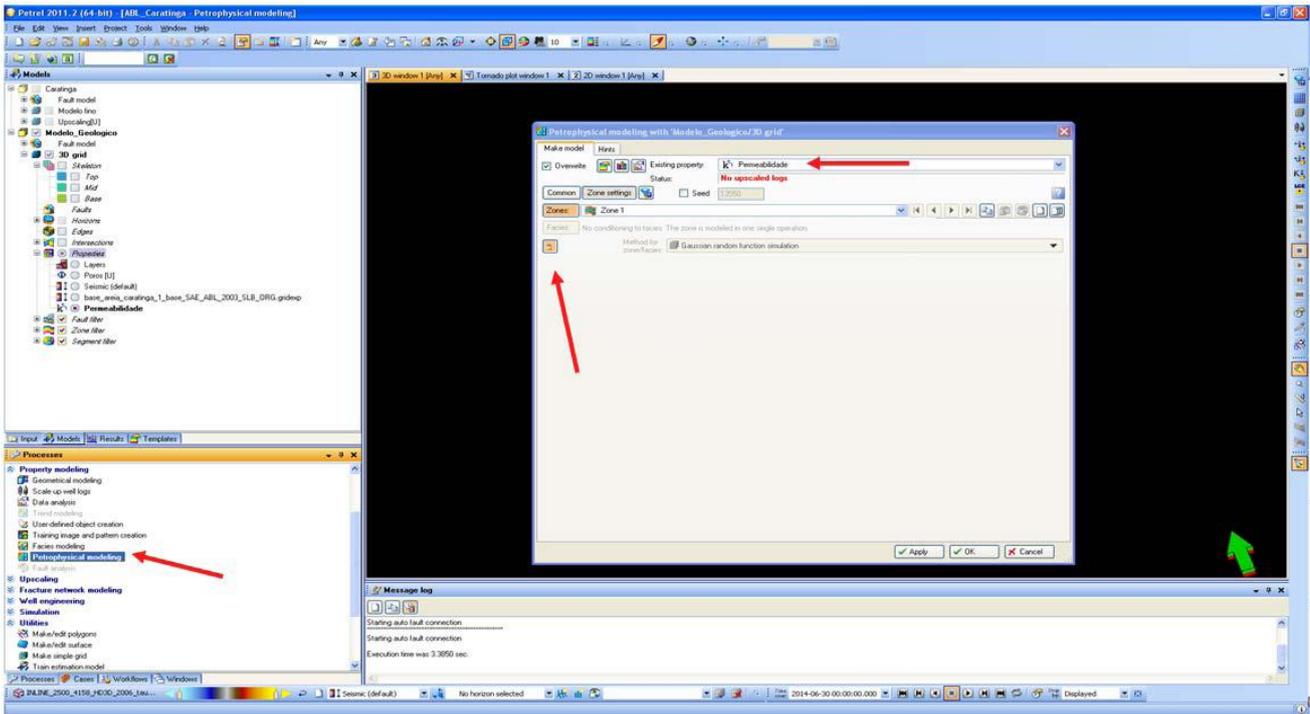


Figure L.2 - Opção Petrophysical Modeling para permeabilidade.

Altere o método para *Sequential Gaussian simulation* e utilize o mesmo variograma da Porosidade (Figure L.3).

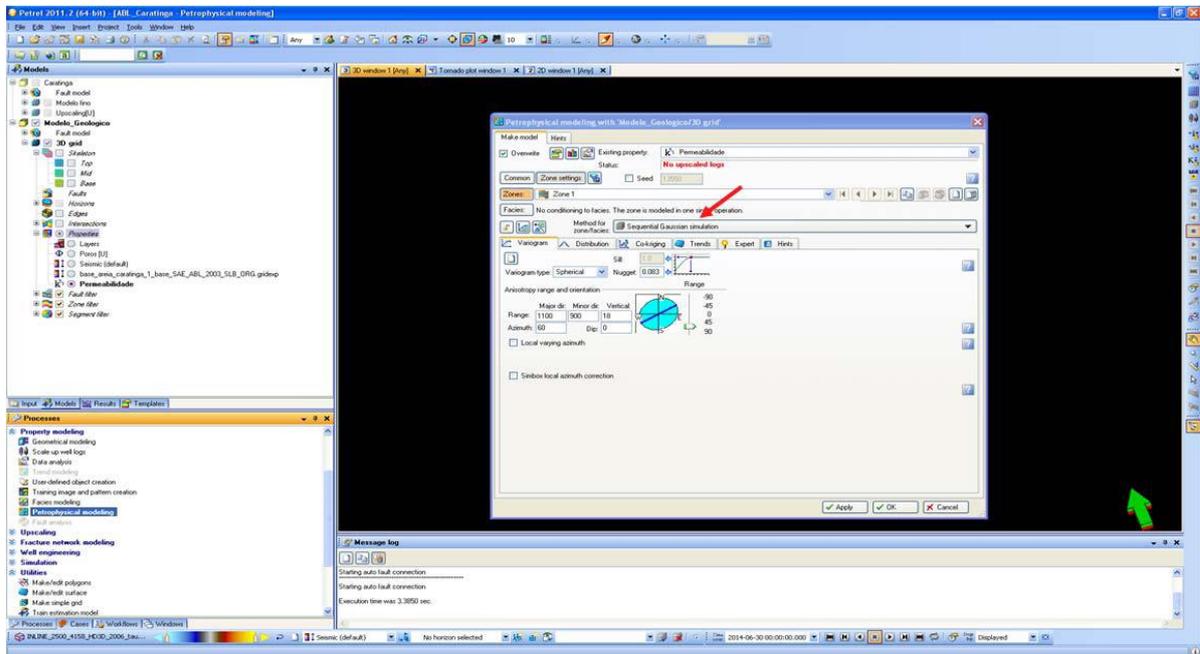
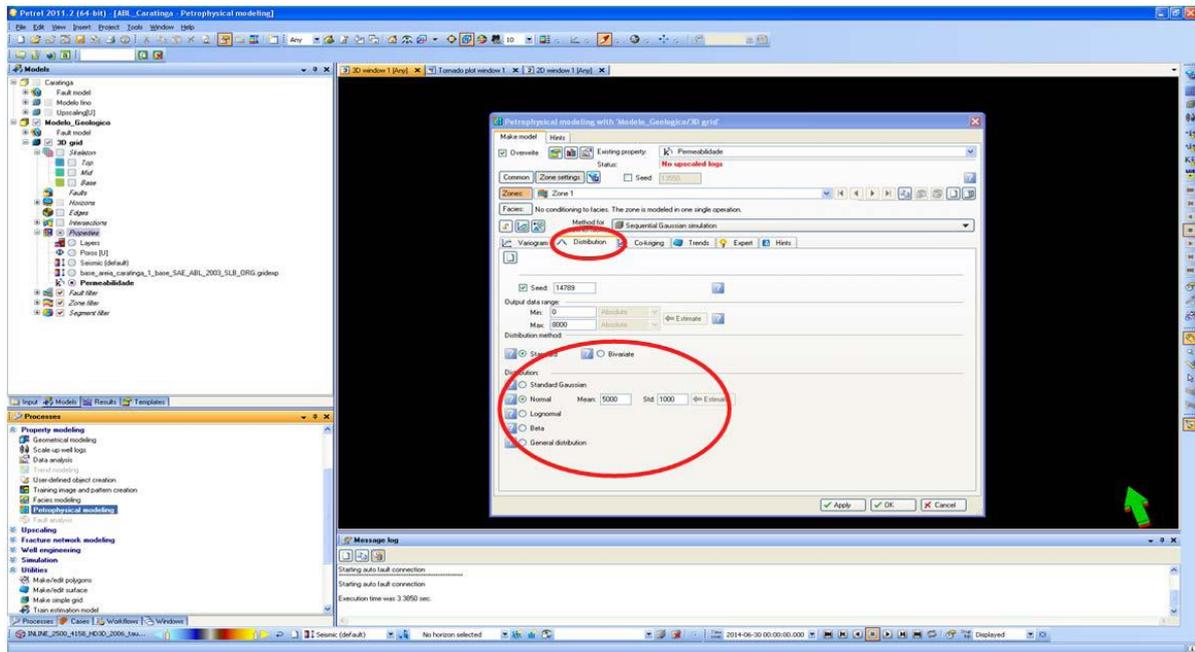


Figure L.3 - Método Sequential Gaussian Simulation.

Na aba *Distribution* (Figure L.4), complete os parâmetros da distribuição: Min: 0, Max: 8000. Selecione distribuição Normal com Mean: 5000, Std: 1000. A permeabilidade média foi obtida a partir dos resultados de teste de formação. Os demais parâmetros foram estimados.



**Figure L.4 - Detalhes da aba *Distribution* para a permeabilidade.**

Na aba *Co-kriging*, defina a porosidade como variável secundária, o método *collocated co-kriging* e a correlação entre as propriedades igual a 0.8 (Figure L.5).

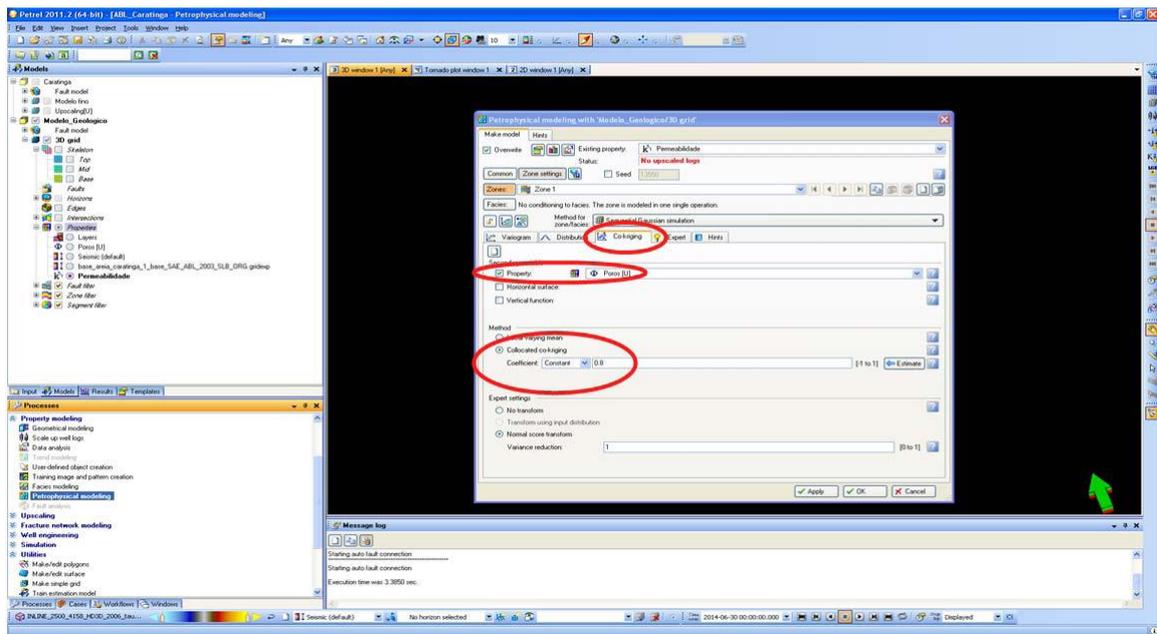


Figure L.5 - Detalhes da aba Co-kriging para permeabilidade.

Selecione a propriedade Permeabilidade para visualizá-la (Figure L.6).

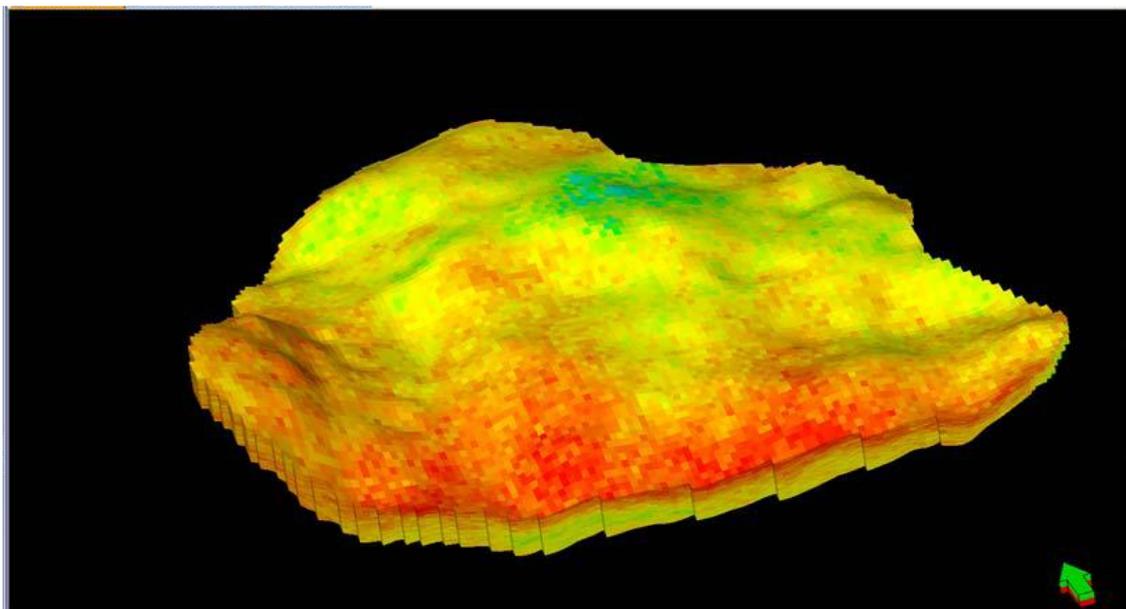


Figure L.6 - Visualização da permeabilidade estimada.