

Roberto Mentzingen Roló

# **Aprimoramentos em modelagem geológica implícita com funções distância assinaladas**

**Proposta de tese para o exame de qualificação**

Porto Alegre

2018



# Listas de ilustrações

Figura 1 – Ilustração esquematizando a modelagem geológica explícita usando polilinhas e linhas guias. . . . .	9
Figura 2 – Ilustração esquemática destacando a incerteza na localização do limite entre os domínios azul e vermelho, entre os dados amostrais. . . . .	10
Figura 3 – Busto reconstruído a partir de pontos esparsos. . . . .	12
Figura 4 – Ilustração esquemática da modelagem geológica implícita. . . . .	12
Figura 5 – Histograma do banco de dados. . . . .	15
Figura 6 – Vista das amostras em perspectiva. . . . .	16
Figura 7 – Amostras codificadas em indicadores para cada uma das três categorias do banco de dados. . . . .	17
Figura 8 – Ilustração esquemática mostrando o cálculo das distâncias assinaladas. . . . .	17
Figura 9 – Distâncias assinaladas calculadas para cada uma das categorias do banco de dados. . . . .	18
Figura 10 – Variogramas experimentais das distâncias assinaladas e modelos para cada uma das categorias do banco de dados. . . . .	19
Figura 11 – Variogramas experimentais dos indicadores e modelos para cada uma das categorias do banco de dados. . . . .	20
Figura 12 – Fluxograma da modelagem geológica implícita usando tabelas de covariância. . . . .	21
Figura 13 – Efeito da resolução do <i>grid</i> na reprodução de estruturas geológicas. . . . .	22
Figura 14 – Modelos geológicos em diferentes resoluções . . . . .	24
Figura 15 – Interpolação das distâncias calculadas por diferentes métodos. . . . .	26
Figura 16 – Modelo não realista gerado por krigagem global para a categoria 2. . . . .	27
Figura 17 – Particionamento. . . . .	29
Figura 18 – Iso superfícies para a categoria 1 extraída dos diferentes modelos implícitos. . . . .	31
Figura 19 – Iso superfície extraída do modelo implícito interpolado por RBF para a categoria 1. . . . .	32
Figura 20 – Iso superfície extraída do modelo implícito interpolado por RBF para a categoria 2. . . . .	32
Figura 21 – Esquema para criação de um modelo implícito multi categórico. . . . .	34
Figura 22 – Modelo geológico multi categórico. . . . .	35
Figura 23 – Esquema mostrando os vetores de anisotropia local para cada nó do <i>grid</i> . . . . .	36
Figura 24 – Iso superfície para a categoria 1 extraída de um modelo implícita gerado por krigagem com anisotropia local variável mostrando os vetores. . . . .	37
Figura 25 – Esquema mostrando os vetores de anisotropia local para cada centro de partição. . . . .	38

Figura 26 – Iso superfície para a categoria 1 extraída de um modelo implícita gerado por krigagem com anisotropia local variável mostrando os vetores. . . . .	38
Figura 27 – Esquema mostrando etapas usando o refinamento iterativo para funções de bases radiais. . . . .	39
Figura 28 – Modelo geológico híbrido criado a partir de furos de sondagem e seções interpretadas. . . . .	41
Figura 29 – Distâncias estimadas em (a) e transformadas em probabilidades em (b) para um mesmo bloco, com cinco categorias. . . . .	43
Figura 30 – Seções em XZ do mapa de probabilidades para a categoria 1 com diferentes valores para $\gamma$ . . . . .	44
Figura 31 – Histograma do bootstrap espacial da média das distâncias assinaladas para a categoria 1. . . . .	45
Figura 32 – Coeficiente U calculado para todos os nós do <i>grid</i> . . . . .	46
Figura 33 – Distâncias simuladas na zona de incerteza para as categorias do banco de dados. . . . .	47
Figura 34 – Diferentes realizações do modelo geológico. . . . .	48
Figura 35 – Diferentes realizações do modelo geológico. . . . .	50
Figura 36 – Calibração do parâmetro C para a categoria 1. . . . .	51
Figura 37 – Classificação dos locais comparando valores estimados e simulados. . . . .	52
Figura 38 – Interpolação, simulação e classificação na zona de incerteza. . . . .	54
Figura 39 – Seções verticais de uma realização para a categoria 1. . . . .	55
Figura 40 – Esquema do método mostrando os passos necessários. . . . .	56
Figura 41 – Comparação de modelos gerados por diferentes métodos de estimativas. . . . .	60
Figura 42 – Figura esquemática de um neurônio. . . . .	61
Figura 43 – Rede neural. . . . .	62
Figura 44 – Zonas de incerteza e contatos. . . . .	63
Figura 45 – Probabilidades de cada categoria em dois diferentes blocos. . . . .	64
Figura 46 – Entropias calculadas para diferentes valores de $\gamma$ . . . . .	65
Figura 47 – Calibração do $\gamma$ . . . . .	66
Figura 48 – Comparação multi categórica. . . . .	67
Figura 49 – Esquema da SGT. Note que as litologias sandstone e shale não aparecerão juntas nas realizações. . . . .	68
Figura 50 – Esquema da SPT. Note que as litologias 1 e 3 não aparecerão juntas nas realizações. . . . .	69
Figura 51 – Diferentes templates, para o caso de duas variáveis latentes e quatro categorias. . . . .	69
Figura 52 – Esquema do uso de MDS para definição da regra de truncagem. . . . .	70
Figura 53 – Diferentes regras de truncagem definidas para cada bloco a partir das probabilidades calibradas e calculadas. . . . .	71

Figura 54 – Modelo implícito ruim, com a presença de estruturas e formas indesejadas ( " <i>blobs</i> " ). . . . .	72
Figura 55 – Cronograma de atividades. . . . .	74



# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>7</b>
<b>2</b>	<b>ESTADO DA ARTE</b>	<b>15</b>
<b>2.1</b>	<b>O banco de dados</b>	<b>15</b>
<b>2.2</b>	<b>A função distância assinalada</b>	<b>16</b>
2.2.1	Calculando a função distância assinalada	16
<b>2.3</b>	<b>Variografia das distâncias calculadas</b>	<b>18</b>
<b>2.4</b>	<b>Interpolação das distâncias calculadas</b>	<b>21</b>
2.4.1	Resolução do grid	22
2.4.2	Métodos de interpolação	24
2.4.3	Funções de bases radiais - RBF	27
2.4.3.1	O <i>kernel</i>	28
2.4.3.2	Decomposição do domínio	28
<b>2.5</b>	<b>Visualização do modelo geológico</b>	<b>29</b>
2.5.1	Adaptação do método para múltiplas categorias simultâneas	32
<b>2.6</b>	<b>Incorporação da não estacionariedade de segunda ordem</b>	<b>35</b>
2.6.1	Krigagem com anisotropia local variável	36
2.6.2	Funções de bases radiais com anisotropia local variável	37
2.6.3	Refinamento iterativo	38
<b>2.7</b>	<b>Incorporação de informação secundária</b>	<b>40</b>
<b>2.8</b>	<b>Avaliação da incerteza</b>	<b>42</b>
2.8.1	Avaliação heurística da incerteza	42
2.8.2	<i>Boundsim</i>	44
2.8.3	Simulação direta das distâncias assinaladas	46
2.8.4	Simulação multi ponto	49
2.8.5	<i>Boundary simulation</i>	50
2.8.6	Sumário dos métodos de avaliação de incerteza	56
<b>3</b>	<b>PROPOSTA DE TESE</b>	<b>57</b>
<b>3.1</b>	<b>Problemas e relevância da tese</b>	<b>57</b>
<b>3.2</b>	<b>Meta</b>	<b>58</b>
<b>3.3</b>	<b>Objetivos</b>	<b>58</b>
3.3.1	Interpolador	58
3.3.2	Zona de incerteza	62
3.3.3	Avaliação de incerteza	66
3.3.3.1	<i>Boundary simulation</i> multi categórico	67

3.3.3.2	Simulação <i>P-field</i>	67
3.3.3.3	Simulação plurigaussiana truncada	68
3.3.4	Validação	72
3.3.5	Sumário	72
<b>3.4</b>	<b>Cronograma e atividades</b>	<b>73</b>
	<b>REFERÊNCIAS</b>	<b>75</b>

# 1 Introdução

Estimativas de recursos e reservas minerais demandam a construção de modelos numéricos de teores de longo prazo, que abrangem toda a extensão do depósito mineral e compreendem todo o tempo de vida da mina. Esses modelos são atualizados em intervalos de um a três anos de operação. Modelos de médio prazo podem ser construídos para planejar de um a seis meses no futuro do empreendimento. Enquanto modelos de curto prazo, são construídos com o objetivo de balizar, semanalmente ou diariamente, as decisões relativas a controle de teores e planejamento detalhado da mina.

Construir modelos numéricos de longo, médio e curto prazo para avaliação de recursos/reservas e planejamento de mina exige quatro grandes atividades ([ROSSI; DEUTSCH, 2013](#)):

1. Coleta e gerenciamento de dados;
2. Interpretação e modelagem geológica;
3. Atribuição de teores;
4. Avaliação e gerenciamento da incerteza geológica e de teores.

Após a coleta, gerenciamento e checagem dos dados, é necessário identificar diferentes domínios. A determinação dos domínios deve ser baseada no conhecimento geológico, i.e. zonas de oxidação, litologias, alteração ou limites estruturais e deve ser suportada por extensiva análise estatística (análise exploratória dos dados), incluindo análises variográficas. É importante que a definição dos domínios faça sentido espacial e geológico. Os domínios são, algumas vezes, baseados na combinação de duas ou mais variáveis para as quais a correlação entre os teores pode ser demonstrada. O procedimento pode ser bastante demorado, principalmente em depósitos multivariados, entretanto, as estimativas são melhoradas significativamente quando cuidadosamente limitadas por variáveis geológicas ([ROSSI; DEUTSCH, 2013](#)).

A definição de diferentes domínios de estimativas, ou zonas estacionárias no interior do depósito, é necessária porque a inferência geoestatística, geralmente, o método de escolha para a construção dos modelos numéricos de teores, exige decisão de estacionariedade ([MCLENNAN, 2007](#)). Isto é, os teores em cada domínio estacionário pertencem a populações estatísticas distintas, são caracterizados através de modelos de distribuição e de covariância específicos e devem ser modelados como funções aleatórias estacionárias diferentes. ([JOURNEL; HUIJBREGTS, 1978](#)).

Uma função aleatória estacionária é uma representação probabilística de uma propriedade petrofísica com valor esperado e covariância constantes independente da localização (MCLENNAN, 2007). Estacionariedade, definida por Matheron (1962), é uma propriedade do modelo de função aleatória, não é uma característica intrínseca da variável. É uma decisão tomada pelo geomodelador e necessária para a inferência.

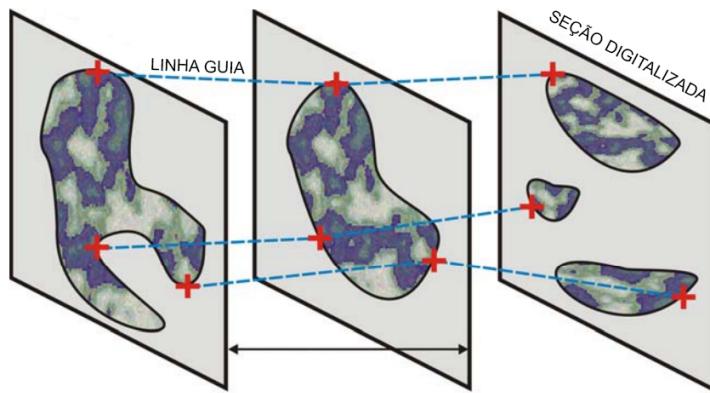
Os domínios estacionários não devem ser muito pequenos, dessa forma não existirão dados suficientes para descrição e inferência estatística confiável, e nem muito grandes, já que os dados, provavelmente, poderiam ser subdivididos em grupos geologicamente mais homogêneos. A definição adequada dos domínios estacionários é uma etapa importante na avaliação de recursos/reservas, a mistura de populações produz estimativas sub ótimas que superestimam ou subestimam teores e tonelagens. Além disso, nenhuma técnica geoestatística pode compensar uma definição de estacionariedade inadequada (ROSSI; DEUTSCH, 2013).

Uma vez que os domínios estacionários tenham sido definidos, um modelo tridimensional que defina os limites de cada função aleatória estacionária deve ser construído: o modelo geológico. É necessário identificar a "jurisdição espacial" de cada função aleatória estacionária. Inevitavelmente, existe incerteza associada à localização dos limites, e essa incerteza geológica deve ser quantificada (MCLENNAN, 2007). A etapa de Interpretação e modelagem geológica compreende a definição dos domínios estacionários e a criação do modelo geológico tridimensional que forma o alicerce para todo o trabalho de estimativa subsequente. Muitas vezes, o modelo geológico é o fator de maior importância na estimativas das tonelagens mineralizadas (ROSSI; DEUTSCH, 2013).

A abordagem tradicional para a criação de modelos geológicos tridimensionais é através da triangulação de polilinhas<sup>1</sup> (*strings*) gerando sólidos que representam os corpos minerais. As polilinhas são digitalizadas manualmente em seções deslocadas, a partir dos dados de sondagem e conhecimento geológico prévio, por um geomodelador. As seções são unidas por linhas guias, que demarcam a conectividade entre as seções digitalizadas para triangulação adequada dos sólidos. Esse procedimento é referido como modelagem geológica explícita, já que os limites entre os diferentes domínios são definidos explicitamente por coordenadas (x,y,z) que posicionam a "colcha de retalhos" de triângulos (MCLENNAN; DEUTSCH, 2006; COWAN et al., 2003). A Figura 1 ilustra a metodologia explícita para um típico depósito de veio.

<sup>1</sup> Em computação gráfica: uma linha composta por vários segmentos, usada para compor imagens na tela (OXFORD, 2018).

Figura 1 – Ilustração esquematizando a modelagem geológica explícita usando polilinhas e linhas guias.



Modificado de [McLennan e Deutsch \(2006\)](#)

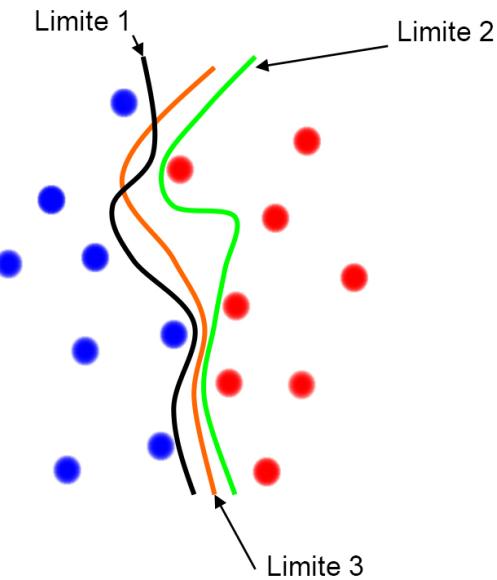
De acordo com [McLennan e Deutsch \(2006\)](#), embora a metodologia seja direta e simples, principais motivos da sua ampla implementação na prática, e que *softwares* modernos de mineração forneçam ferramentas computais para visualizar os dados de sondagem e agilizar o processo de digitalização ([SILVA, 2015](#)), existem uma série de limitações e desvantagens. O processo é tedioso e demorado, digitalizar manualmente as polilinhas e uni-las por linhas guia, exige muito tempo de um profissional experiente. Em depósitos de alta complexidade, não é raro o geomodelador trabalhar até três meses no modelo geológico explícito. A geometria dos corpos geológicos muitas vezes precisa ser simplificadas para que o modelo seja concebido em tempo hábil. Por esse motivo, para a maioria das minas, apenas um modelo é construído e mantido, raramente há oportunidade de modelar interpretações geológicas alternativas e comparar estimativas de recursos baseadas em diferentes modelos ([COWAN et al., 2003](#)).

Os modelos criados explicitamente são subjetivos e não replicáveis. O volume mineralizado é essencialmente composto por uma série de pequenas decisões subjetivas, já que cada ponto da polilinha em cada seção é escolhido por um geomodelador. Inevitavelmente, a "assinatura" do profissional é impressa nos limites dos domínios geológicos. Geomodeladores diferentes criam modelos diferentes a partir do mesmo banco de dados. Ademais, modelos explícitos são inflexíveis, é laborioso e demorado atualizar modelos explícitos a medida que novos dados são obtidos já que uma nova digitalização manual é necessária.

Em virtude da onerosidade em construir múltiplos modelos geológicos explícitos, é custoso avaliar a incerteza na localização dos limites dos diferentes domínios entre os dados amostrais, como mostrado na ilustração esquemática na [Figura 2](#). Além da tediosa redigitalização manual de polilinhas e retriangulação, não existe forma direta de incorporar

múltiplas realizações possíveis para a localização dos limites que representam a incerteza associada.

Figura 2 – Ilustração esquemática destacando a incerteza na localização do limite entre os domínios azul e vermelho, entre os dados amostrais.



Modificado de Cáceres et al. (2011)

Em muitos casos, a incerteza do modelo geológico pode ser uma fonte de incerteza crucial. Em depósitos de veio de ouro, por exemplo, o volume mineralizado é um indicador econômico vital para o gerenciamento do projeto. Ignorar a incerteza volumétrica, considerando apenas um único modelo geológico explícito, pode ser devastador para o empreendimento. A incerteza associada ao modelo geológico deve ser avaliada.

Embora a metodologia explícita seja demorada, laboriosa, subjetiva, não replicável, inflexível e incapaz de avaliar incertezas, os limites entre domínios resultantes, geralmente, são realistas. Realismo geológico é um dos principais objetivos da modelagem e pode ser diretamente controlado durante o processo de digitalização.

Dadas as desvantagens do método explícito de modelagem geológica, pesquisas vêm sendo realizadas com o objetivo de automatizar, ou pelo menos, semi automatizar a modelagem. Métodos matemáticos menos sofisticados, como o vizinho mais próximo podem ser utilizados na criação automática de modelos geológicos determinísticos. Métodos geoestatísticos também têm aplicação, a krigagem dos indicadores, (JOURNEL, 1982) consiste na geração de uma distribuição acumulada de probabilidades a partir da transformação não linear dos dados. Para variáveis categóricas, é possível estimar diretamente a probabilidade de cada local não amostrado pertencer à cada uma das diferentes litologias

do depósito mineral.

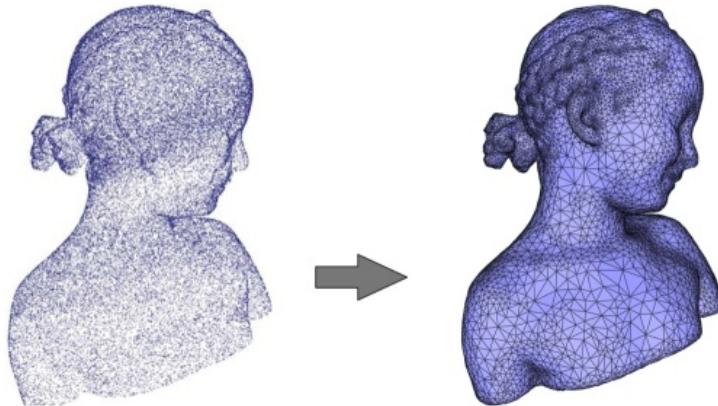
A necessidade de avaliação da incerteza associada ao modelo geológico impulsionou o desenvolvimento de métodos estocásticos baseados em múltiplas realizações equiprováveis de um atributo. A partir da análise conjunta das realizações é possível avaliar a incerteza.

São metodologias estabelecidas da geoestatística clássica, baseada no variograma: a simulação sequencial dos indicadores ([GOOVAERTS, 1997](#)), que compartilha do mesmo princípio da krigagem dos indicadores. A simulação gaussiana truncada ([MATHERON et al., 1987](#)), que tem como ideia central gerar múltiplas realizações de uma variável aleatória gaussiana contínua e truncá-las em uma série de limites pré definidos, indicando as diferentes litologias. A simulação plurigaussiana truncada, ([GALLI et al., 1994](#)) que é uma extensão da simulação gaussiana truncada, baseada na truncagem de duas funções aleatórias gaussianas que podem ou não ser correlacionadas.

Outras metodologias não baseadas em variograma foram desenvolvidas: a simulação geoestatística multiponto ([STREBELLE, 2002](#)), tem o mesmo formalismo das simulações sequenciais baseadas em variograma, com a diferença que a distribuição local de probabilidade é encontrada a partir da amostragem de uma imagem de treinamento com um modelo multi ponto ([PYRCZ; DEUTSCH, 2014](#)). Do mesmo modo, a simulação baseada em objetos ([ALABERT; MODOT et al., 1992; BRATVOLD et al., 1995; HOLDEN et al., 1998](#)), metodologia que produz modelos atrativos (visualmente limpos) que honram a geometria idealizada interpretada a partir de afloramentos ou dados de sísmica. Quando geometrias específicas do depósito ou reservatório são identificadas de forma correta, parametrizadas e têm significância para a função de transferência, podem ser integradas diretamente no modelo pelos métodos baseados em objetos ([PYRCZ; CATUNEANU; DEUTSCH, 2005](#)).

Uma outra família de métodos, ditos métodos implícitos, compreendem técnicas importadas da computação gráfica. O uso da modelagem implícita foi introduzida no campo da computação gráfica para a geração de objetos de diferentes geometrias e complexidades por [Bloomenthal e Bajaj \(1997\)](#). A ideia geral é usar uma função implícita para demarcar regiões de diferentes formas e extensões no espaço. [Cowan et al. \(2002\)](#), [Cowan et al. \(2003\)](#) popularizou a modelagem implícita no contexto geológico, baseado no trabalho de [Savchenko et al. \(1995\)](#) que modela objetos tridimensionais a partir de dados esparsos interpolando uma função volume ([Figura 3](#)).

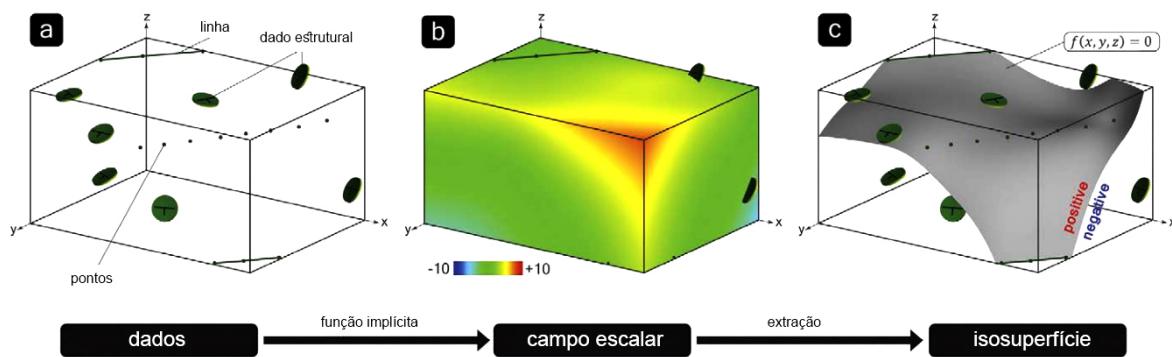
Figura 3 – Busto reconstruído a partir de pontos esparsos.



Fonte: [CGAL \(2015\)](#)

Modelos geológicos implícitos são criados a partir de dados esparsos, geralmente, dados categóricos ([Figura 4 \(a\)](#)). Os pontos amostrais podem ser usados para derivar uma função implícita que fornece uma representação matemática contínua de um atributo através de um volume, um campo escalar ou modelo implícito ([Figura 4 \(b\)](#)). Modelos implícitos contém um número infinito de isosuperfícies, para visualizar o modelo geológico uma isosuperfície específica deve ser extraída do modelo e mostrada no espaço tridimensional (geralmente a isosuperfície zero),  $f(x, y, z) = s$ , onde  $s$  é o valor de interesse no campo escalar ([Figura 4 \(c\)](#)). A localização dessa superfície é conhecida implicitamente como função da localização no domínio ([MARTIN; BOISVERT, 2017](#)).

Figura 4 – Ilustração esquemática da modelagem geológica implícita.



Modificado de ([AILLERES et al., 2015](#))

Diferentes tipos de dados amostras podem ser usados para derivar diferentes funções

---

volume na modelagem implícita. Mallet (2004) propõe uma função volumétrica cronológica, levando em consideração a posição estratigráfica das diferentes unidades geológicas, enquanto Lajaunie, Courrioux e Manuel (1997) usam co-krigagem de incrementos em um campo potencial, omitindo a função volume. A função volume mais comumente utilizada é a função distância assinalada (OSHER; FEDKIW, 2003), a aplicação dessa metodologia é encontrada por toda a literatura de interpolação de dados esparsos, uma das aplicações é a reconstrução de superfícies a partir de escaneamento (Figura 3). Na modelagem geológica a metodologia é competente em capturar a geometria e extensão de corpos geológicos e tem sido aplicada com sucesso há mais de 10 anos na exploração mineral tendo ganhado espaço em softwares comerciais como o *Leapfrog* ®. Para bancos de dados pontuais, que representam a posição espacial de diferentes litologias no espaço, o uso das funções distância assinaladas torna os processos de interpretação, cálculo, modelagem, modificação do modelo de acordo com a interpretação dos geomodeladores e avaliação da incerteza diretos e rápidos. Além disso, informações a respeito da orientação das unidades geológicas (dados estruturais) pode ser imputada como restrição na criação do modelo geológico (MARTIN; BOISVERT, 2017).



## 2 Estado da arte

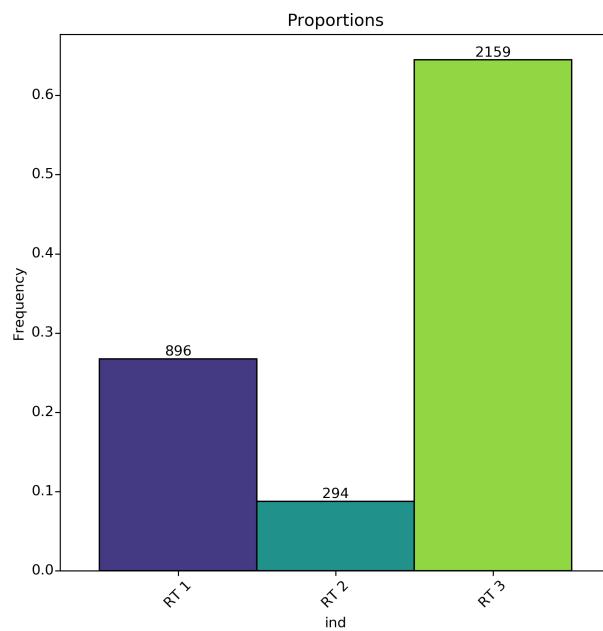
Esse capítulo detalha a modelagem geológica implícita com funções distância assinaladas, desde o cálculo das distâncias até a avaliação de incerteza dos modelos geológicos.

O fluxo de trabalho e os diferentes métodos disponíveis na literatura serão apresentados com exemplos práticos em um banco de dados tridimensional sintético que emula um depósito de cobre pôrfiro. Pode ser acessado na íntegra, com toda a memória de cálculo no repositório: [implicit modeling and uncert final.html](#)

### 2.1 O banco de dados

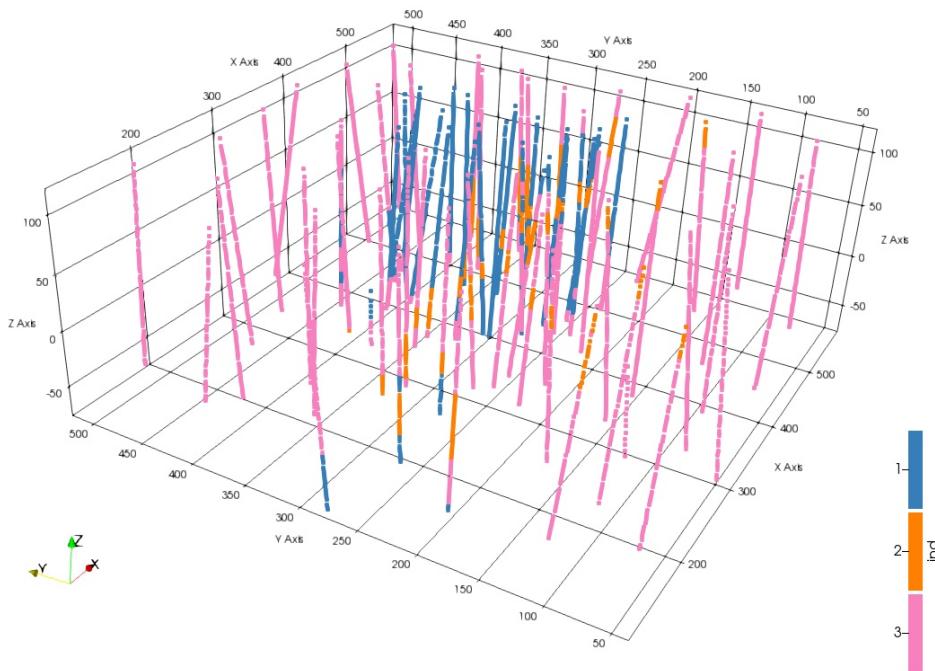
O banco de dados categóricos contém 72 furos totalizando 3349 amostras distribuídas entre 3 diferentes categorias conforme o histograma da [Figura 5](#).

Figura 5 – Histograma do banco de dados.



A [Figura 6](#) mostra a posição espacial das amostras do banco de dados tridimensional.

Figura 6 – Vista das amostras em perspectiva.



## 2.2 A função distância assinalada

A modelagem implícita é geralmente aplicada à variáveis categóricas que representam diferentes litologias do depósito mineral para definir a localização, extensão e geometria dos corpos minerais. A variável de indicadores deve ser codificada em uma função volumétrica e interpolada exaustivamente para todos os locais de interesse. A função distância assinalada ([OSHER; FEDKIW, 2003](#)) é a função volume mais utilizada na modelagem implícita de variáveis categóricas.

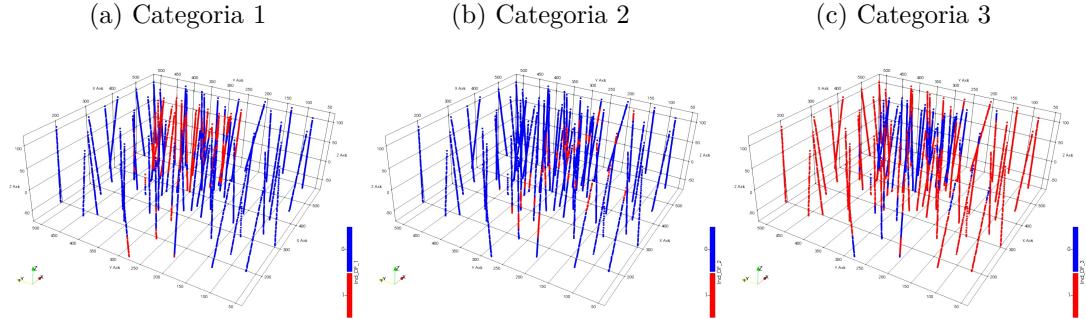
### 2.2.1 Calculando a função distância assinalada

O primeiro passo no cálculo das distâncias assinaladas é transformar as amostras  $z(u_\alpha)$ ,  $\alpha = 1, \dots, n$  em indicadores binários de acordo com a [Equação 2.1](#), especificando se a amostra pertence ou não ao domínio  $k$  que está sendo modelado.

$$i_k(u_\alpha) = \begin{cases} 1, & \text{se } z(u_\alpha) \text{ se pertence ao domínio } k \\ 0, & \text{caso contrário} \end{cases} \quad (2.1)$$

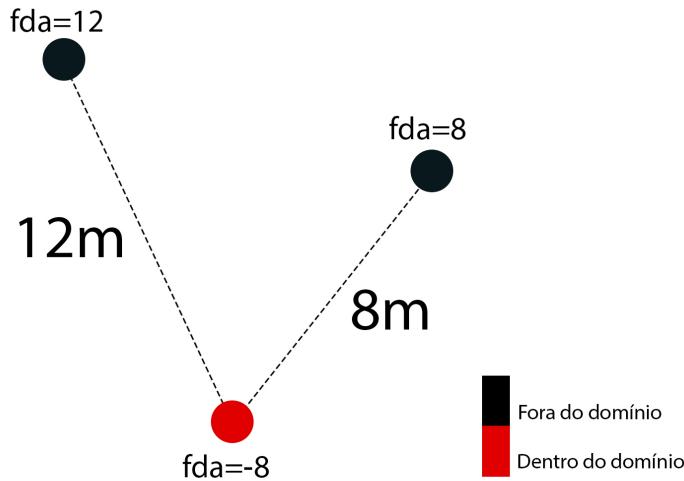
A [Figura 7](#) mostra as amostras codificadas em indicadores para cada uma das três categorias do banco de dados.

Figura 7 – Amostras codificadas em indicadores para cada uma das três categorias do banco de dados.



O segundo passo é o cálculo das distâncias assinaladas. Para cada ponto amostral, a menor distância euclideana até um outro ponto amostral que pertence à um indicador oposto é computada, e esse valor atribuído àquele ponto, com o sinal negativo caso pertença ao domínio modelado e com o sinal positivo, caso contrário como ilustrado na [Figura 8](#).

Figura 8 – Ilustração esquemática mostrando o cálculo das distâncias assinaladas.



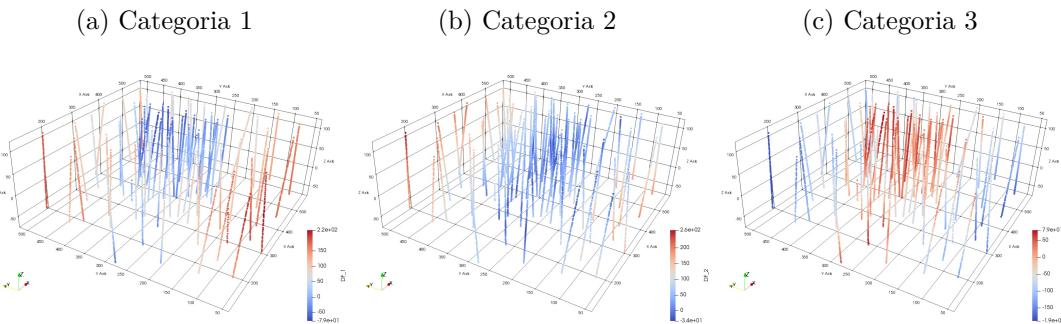
Desse modo, o valor da função distância assinalada  $d_k$  para a o domínio  $k$  modelado no local  $u_\alpha$  é:

$$d_k(u_\alpha) = \begin{cases} - \| u_\alpha - u_\beta \|, & \text{se } u_\alpha \text{ pertence ao domínio} \\ + \| u_\alpha - u_\beta \|, & \text{se } u_\alpha \text{ não pertence ao domínio} \end{cases} \quad (2.2)$$

O local  $u_\beta$  corresponde à amostra mais próxima codificada com um indicador diferente de  $u_\alpha$ .

A função distância assinalada calculada para cada uma das categorias do banco de dados é vista na [Figura 9](#).

Figura 9 – Distâncias assinaladas calculadas para cada uma das categorias do banco de dados.



Para corpos geológicos que se estendem muito mais em uma direção em relação às demais, como os corpos tabulares por exemplo, as distâncias calculadas podem ser anisotrópicas. Sendo assim, as coordenadas originais  $x$ ,  $y$  e  $z$  dos dados devem ser rotacionadas e/ou contraídas/dilatadas, a partir da transformação da [Equação 2.3](#). Então, as distâncias euclidianas ([Equação 2.12](#)) são calculadas normalmente para as novas coordenadas  $x''$ ,  $y''$ ,  $z''$ .

$$\begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = \begin{bmatrix} \frac{1}{a_{max}} & 0 & 0 \\ 0 & \frac{1}{a_{min}} & 0 \\ 0 & 0 & \frac{1}{a_{vert}} \end{bmatrix} \begin{bmatrix} \cos \alpha \cos \phi - \sin \alpha \sin \beta \sin \phi & -\sin \alpha \cos \phi - \cos \alpha \sin \beta \sin \phi & \cos \beta \sin \phi \\ \sin \alpha \cos \beta & \cos \alpha \cos \beta & \sin \beta \\ -\cos \alpha \sin \phi - \sin \alpha \sin \beta \cos \phi & \sin \alpha \sin \phi - \cos \alpha \sin \beta \cos \phi & \cos \beta \cos \phi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (2.3)$$

Onde  $a_{max}$ ,  $a_{min}$  e  $a_{vert}$  são as dimensões dos eixos máximo, médio e mínimo de anisotropia, e  $\alpha$ ,  $\beta$  e  $\phi$ , os ângulos de azimute, mergulho e *rake*, respectivamente.

## 2.3 Variografia das distâncias calculadas

No caso das distâncias assinaladas, obter um modelo de covariância a partir dos variogramas experimentais não é trivial. Distâncias assinaladas não são estacionárias, o variograma não se estabiliza em um patamar. Além disso, o caráter extremamente contínuo das distâncias torna a identificação analítica das direções principais um processo embaralhado.

[Manchuck e Deutsch \(2015\)](#) sugerem alternativas para modelagem do variograma das distâncias:

- Treinar o variograma usando validação cruzada;
- Tentar modelar interativamente os variogramas experimentais;

- Calcular e modelar os variogramas para as propriedades de indicadores e transformá-los em um equivalente gaussiano para as distâncias assinaladas;
- inferir um modelo de covariância plausível visualmente a partir das amostras ou de mapas delineados a mão.

O alcance do variograma, juntamente com a anisotropia, controlam a extensão e forma dos domínios modelados. Os variogramas experimentais das distâncias assinaladas nunca exibem efeito pepita. Porém, pode ser adicionado arbitrariamente pelo usuário para controlar a interconectividade dos domínios, adicionar ruído ao modelo quando desejável e controlar a extração que pode ser excessiva em alguns casos devido a alta continuidade dos modelos Gaussianos.

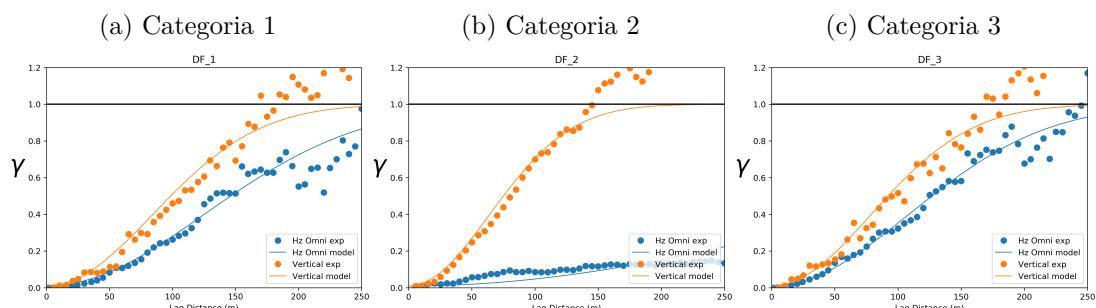
Os variogramas experimentais das distâncias assinaladas e dos indicadores foram calculados para todas as categorias do banco de dados. A categoria 1 é uma intrusão vertical. Então foi calculado o variograma omnidirecional no plano horizontal e na direção vertical. A categoria 2 é um intrusão tabular (veio) com direção N351 e mergulho -48. Sendo assim, foi calculado o variograma experimental omnidirecional no plano de altitude e na direção perpendicular ao plano. A categoria 3 é a rocha encaixante, as direções foram as mesmas da categoria 1. A Tabela 1 mostra os parâmetros usados no cálculo dos variogramas experimentais das distâncias assinaladas e dos indicadores.

Número de lags	Distância do lag	Tolerância linear	Tolerância do azimute	Tolerância do mergulho	Banda horizontal
50	5m	2.5m	90	22.5	20m

Tabela 1 – Parâmetros de cálculo dos variogramas experimentais.

O modelo mais indicado para modelar o variograma das distâncias assinaladas é o modelo Gaussiano. A Figura 10 mostra os variogramas experimentais e modelos ajustados das distâncias assinaladas para cada uma das categorias do banco de dados. Os modelos foram ajustados de forma automática pelo software PyGeostat ((CCG), 2015–2019), com efeito pepita de 0.01 (para evitar instabilidades nas matrizes de krigagem), uma estrutura Gaussiana, patamar estandardizado e com um máximo de 2500 iterações.

Figura 10 – Variogramas experimentais das distâncias assinaladas e modelos para cada uma das categorias do banco de dados.



Os modelos variográficos para cada categoria são expressos pelas Equação 2.4, Equação 2.5 e Equação 2.6.

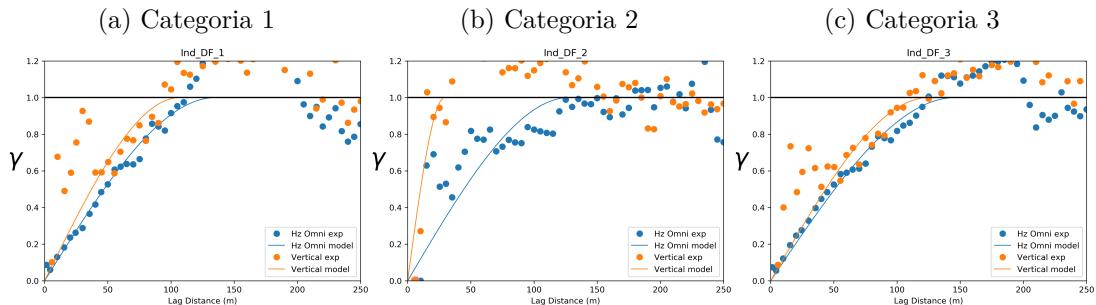
$$\gamma_{SD1}(h) = 0.01 + 0.99\text{gauss}\left(\frac{h_{omni}}{306m}, \frac{h_{normal}}{206m}\right) \quad (2.4)$$

$$\gamma_{SD2}(h) = 0.01 + 0.99\text{gauss}\left(\frac{h_{omni}}{877m}, \frac{h_{normal}}{157m}\right) \quad (2.5)$$

$$\gamma_{SD3}(h) = 0.01 + 0.99\text{gauss}\left(\frac{h_{omni}}{265m}, \frac{h_{normal}}{193m}\right) \quad (2.6)$$

O modelo mais indicado para modelar o variograma dos indicadores é o modelo esférico. A Figura 11 mostra os variogramas experimentais e modelos ajustados para os indicadores de cada uma das categorias do banco de dados. Os modelos foram ajustados de forma automática pelo software PyGeostat ((CCG), 2015–2019), com efeito pepita de 0, uma estrutura esférica, patamar estandardizado e com um máximo de 2500 iterações.

Figura 11 – Variogramas experimentais dos indicadores e modelos para cada uma das categorias do banco de dados.



Os modelos variográficos para cada categoria são expressos pelas Equação 2.7, Equação 2.8 e Equação 2.9.

$$\gamma_{IND1}(h) = 1.0\text{sph}\left(\frac{h_{omni}}{135m}, \frac{h_{normal}}{106m}\right) \quad (2.7)$$

$$\gamma_{IND2}(h) = 1.0\text{sph}\left(\frac{h_{omni}}{128m}, \frac{h_{normal}}{29m}\right) \quad (2.8)$$

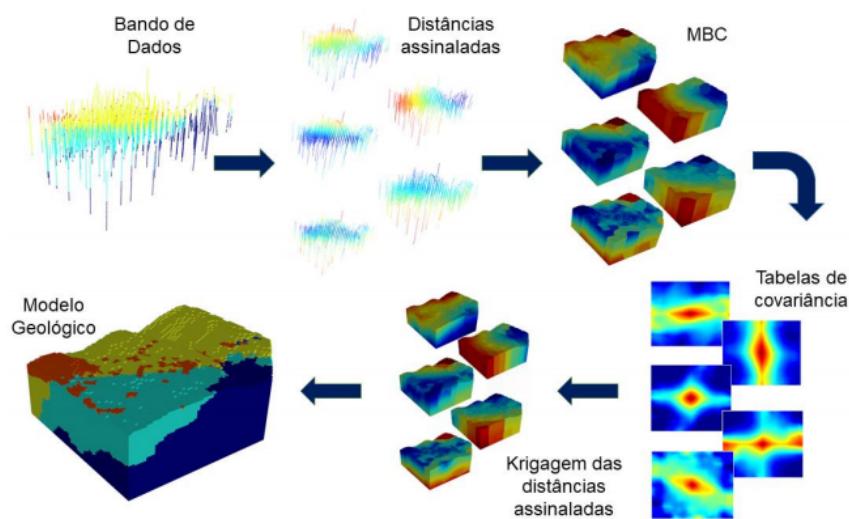
$$\gamma_{IND3}(h) = 1.0\text{sph}\left(\frac{h_{omni}}{146m}, \frac{h_{normal}}{125m}\right) \quad (2.9)$$

Em teoria, o ajuste dos variogramas dos indicadores é mais fácil se comparado ao variograma das distâncias assinaladas, além de atingirem um patamar, a menor continuidade espacial em relação às distâncias torna a identificação analítica das direções

principais um processo mais simples. No entanto, muitas vezes, na prática se observa o contrário.

Uma outra alternativa ao cálculo e modelagem de variogramas é o uso de tabelas de covariância proposto por Kloeckner et al. (2018), onde propõem a substituição da modelagem tradicional dos variogramas por tabelas de covariância, tornando o processo de modelagem geológica completamente automático e livre da subjetividade do geomodelador. A Figura 12 mostra o fluxograma da metodologia proposta.

Figura 12 – Fluxograma da modelagem geológica implícita usando tabelas de covariância.



Kloeckner et al. (2018)

O trabalho propõe uma técnica para obter uma tabela de covariância, a partir de um modelo base, que reproduz a continuidade espacial de um dado atributo usando o teorema de convolução via transformada rápida de Fourier. Embora a metodologia tenha apresentado resultados satisfatórios, a construção automática das tabelas de covariância ainda apresenta alguns problemas operacionais que tornam sua aplicação inviável para todos os tipos de depósito.

## 2.4 Interpolação das distâncias calculadas

Uma vez que as distâncias tenham sido calculadas a partir da localização das amostras, o próximo passo é interpolar o novo banco de dados para o *grid* definido para a modelagem geológica. Embora métodos implícitos não demandem um *grid* - são métodos conhecidos como *gridless* - o objetivo final é preencher o *grid* de estimativas com as categorias correspondentes ou visualizar o modelo através de iso superfícies. Técnicas de

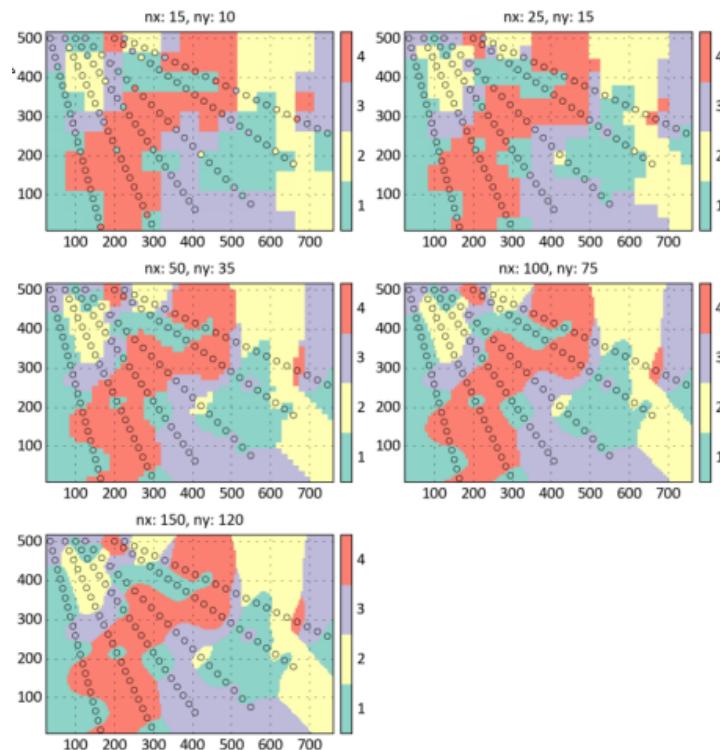
extração de iso superfícies exigem a função volumétrica definida exaustivamente em um *grid* retilíneo ([MARTIN; BOISVERT, 2017](#)).

### 2.4.1 Resolução do grid

Para um número constante de amostras, o tempo necessário para interpolação da função volume para todos os nós do *grid* aumenta linearmente com o número de nós. Os parâmetros do *grid* de estimativas, muitas vezes, são determinado por requisitos de engenharia do projeto. Todavia, o *grid* para definição do modelo geológico pode ser diferente do *grid* definido para os modelos numéricos de teores. A consideração mais importante é a capacidade de reproduzir a menor estrutura geológica de interesse.

Considere os modelos implícitos da [Figura 13](#) gerados em *grids* de diferentes resoluções.

Figura 13 – Efeito da resolução do *grid* na reprodução de estruturas geológicas.



Fonte: [Martin e Boisvert \(2017\)](#)

A resolução adequada para esse modelo é entre  $(nx, ny) = (50, 35)$  e  $(nx, ny) = (100, 75)$ . Para os dois primeiros *grids*, há classificação errônea dos nós colocados com amostras, além de  $(nx, ny) = (100, 75)$  há pouco benefício na reprodução das estruturas levando em consideração o aumento do tempo de interpolação.

É preciso encontrar um balanço entre resolução e tempo de interpolação. A resolução do *grid* pode ser aumentada a medida que o projeto avança e mais informação é adquirida (MARTIN; BOISVERT, 2017).

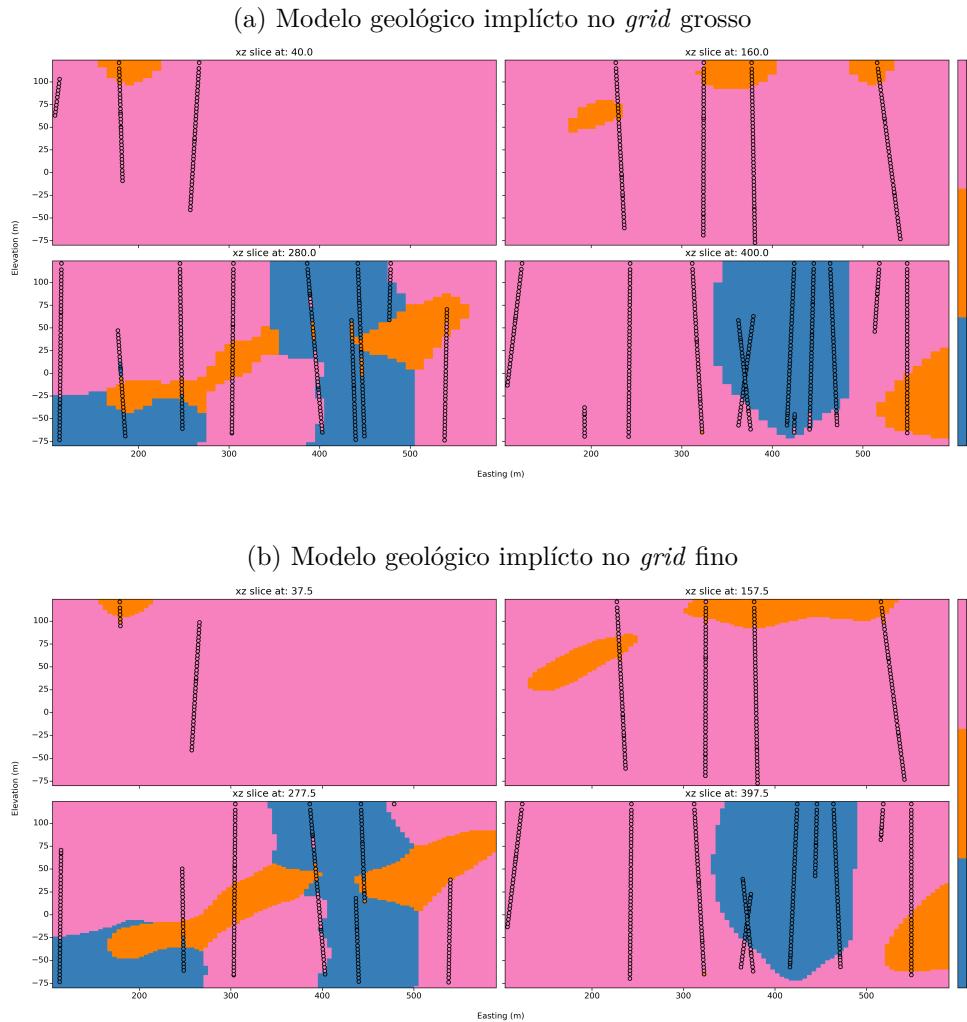
Para o estudo de caso, foram gerados dois *grids* com resoluções diferentes, seus parâmetros são mostrados na [Tabela 2](#).

	Grosso	Fino
nx	49	97
ny	49	98
nz	51	102
sx	10m	5m
sy	10m	5m
sz	4m	2m
num	122451	969612

Tabela 2 – Parâmetros dos *grids* de definição dos modelos geológicos implícitos.

A [Figura 14](#) mostra seções em XZ de dois modelos geológicos implícitos gerados em *grids* com resoluções diferentes para o banco de dados do estudo de caso. No *grid* fino ([46a](#)), as formas reproduzidas são suaves, apresentam realismo geológico, e o número de classificações errôneas é menor.

Figura 14 – Modelos geológicos em diferentes resoluções



## 2.4.2 Métodos de interpolação

Qualquer método de interpolação pode ser utilizado para as distâncias assinaladas, mesmo os métodos do inverso da distância produzem resultados realistas. A krigagem e as funções de bases radiais permitem incorporar informação adicional através dos modelos de covariância parametrizados a partir das amostras e minimizam a variância do erro. A literatura recomenda uso de métodos globais, que usam todas as amostras para cada estimativa, para evitar o surgimento de artefatos nos modelos.

[Hosseini e Deutsch \(2007\)](#) utilizaram inverso da distância, [Silva \(2015\)](#) utilizou krigagem ordinária global, [Rolo \(2017\)](#) utilizou krigagem ordinária, [Silva e Deutsch \(2015\)](#) aplicaram *dual kriging*, [Boisvert \(2013\)](#) gerou modelos implícitos através de distâncias assinaladas com anisotropia variável local (*Locally varying anisotropy kriging - LVA*) e [Cowan et al. \(2002\)](#) propôs a utilização de interpoladores baseados em funções de bases

radiais (*radial basis functions - RBF*). [Manchuck e Deutsch \(2018\)](#) ainda propuseram a utilização de mínimos quadrados móveis para incorporar interpretação manual e avaliar incerteza. A escolha do interpolador deve ser baseada em suas propriedades, na informação exigida para parametrizá-lo (o variograma, por exemplo) e nas sua capacidade de se ajustar a formas geológicas.

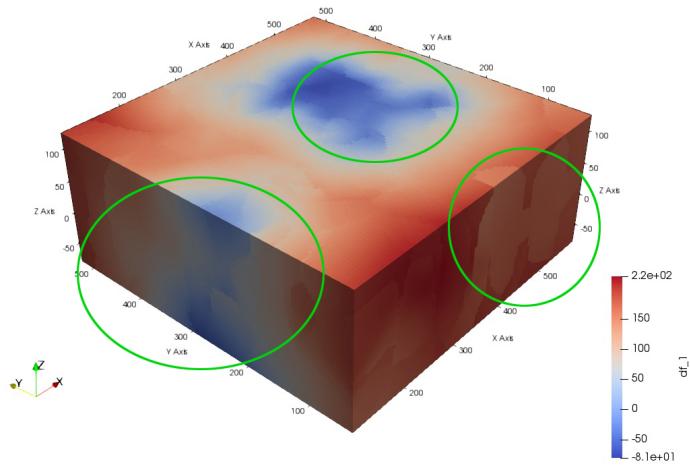
Além das características de cada interpolador, para as funções distância assinaladas existe um problema adicional relacionado à estacionariedade, de primeira e segunda ordem. Existe uma forte tendência (*trend*) na variável distância que restringe os estimadores baseados em krigagem à krigagem ordinária para que a estacionariedade seja mantida. Além disso, formas geológicas são curvilíneas e não são bem descritas por funções de covariância lineares. Por esses motivos, muitos autores preferem as funções de bases radiais já que ela não se baseia em estacionariedade de primeira ordem e honra localmente formas geológicas sem a necessidade de parametrização especial ([MARTIN; BOISVERT, 2017](#)).

Independente do método de interpolação, os modelos gerados devem ser livres de artefatos para que os artefatos do modelo implícito não sejam transferidos para os modelos geológicos. Funções de bases radiais e krigagem global garantem modelos livres de artefato por utilizarem todas as amostras em cada estimativa, já para os métodos baseados em inverso da distância e krigagem (não global) os parâmetros de busca são uma escolha crítica ([MARTIN; BOISVERT, 2017](#)). No entanto, métodos globais possuem limitações em bancos de dados volumosos por dependerem de muito processamento e memória RAM.

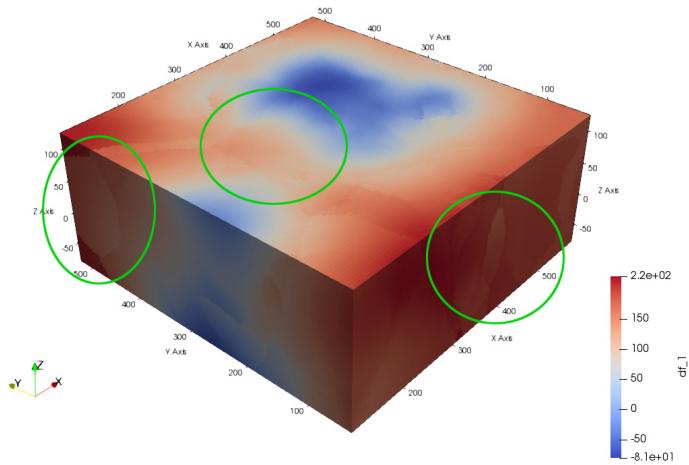
A [Figura 15](#) mostra três modelos implícitos, para a categoria 1, interpolados por krigagem ordinária usando 40 amostras, krigagem ordinária usando 100 amostras e funções de bases radiais, que é um método global. A krigagem ordinária global gerou modelos implícitos extremantes contínuos, mais do que o desejado, implicando em modelos geológicos não realistas, como mostrado na [Figura 16](#).

Figura 15 – Interpolação das distâncias calculadas por diferentes métodos.

(a) OK com 40 amostras



(b) OK com 100 amostras



(c) RBF global

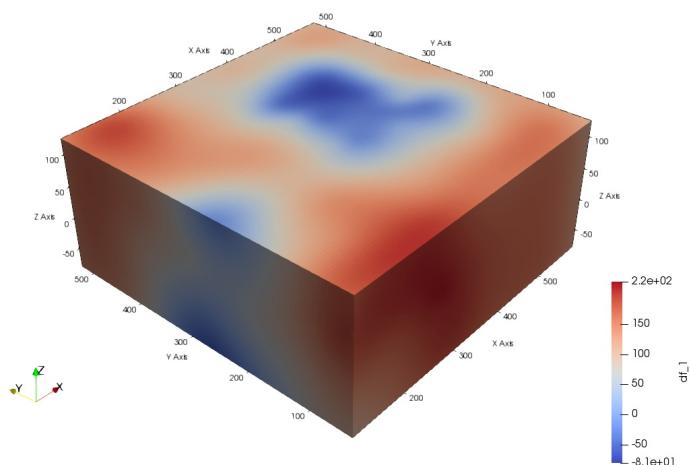
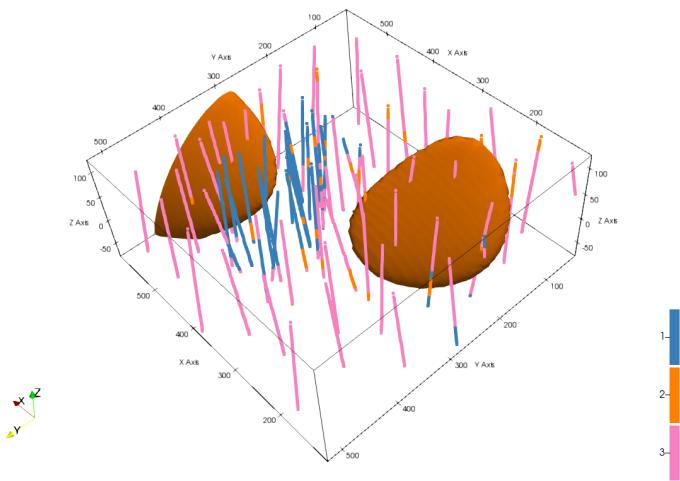


Figura 16 – Modelo não realista gerado por krigagem global para a categoria 2.



É evidente a presença de artefatos, realçados em verde, nos dois primeiros modelos. As funções de bases radiais criam um modelo suave e livre de artefatos.

A Tabela 3 mostra um *Benchmark* dos diferentes métodos de interpolação. Todos os algoritmos utilizados são da biblioteca GSLib e foram executados em um core i7 7700HQ @ 2.8 GHz com 16 Gb de RAM.

O algoritmo de krigagem global não terminou de estimar os quase um milhão de blocos do *grid* fino sem terminar com uma mensagem de erro. A krigagem ordinária com 100 amostras leva 45 minutos contra pouco mais de um minuto do RBF, subdividindo o problema, é possível reduzir o tempo de interpolação para 16 segundos. Não é difícil perceber porque o RBF é o interpolador preferido na modelagem geológica implícita. O software comercial Leapfrog® tem um algoritmo patenteado de RBF ainda mais rápido.

Método	Tempo grid grosso	Tempo grid fino	Classificação errônea grosso	Classificação errônea fino
krigagem global isotrópica	28min		121	
krigagem global anisotrópica	30min 34s		282	
krigagem ordinária anisotrópica (40)	1min 3s	38min	137	135
krigagem ordinária anisotrópica (100)		45min 51s		181
RBF isotrópico	21.5s		57	
RBF anisotrópico		1min 22s		38
Particionado RBF anisotrópico		1min 2s		39
Particionado RBF artefatos	16.5s			29
LVA OK			8	
LVA RBF				8
Krigagem dos indicadores		33min 27s		2

Tabela 3 – *Benchmark* dos diferentes métodos de interpolação.

### 2.4.3 Funções de bases radiais - RBF

A demonstração detalhada das equações e considerações pode ser encontrada em Martin e Boisvert (2015). Os interpoladores baseados em funções de bases radiais são

globais e não necessitam de um *grid*. Os pesos são treinados baseados nos valores da função distância assinalada calculados para cada ponto amostral. Com os pesos treinados o valor interpolado da função distância assinalada pode ser obtido rapidamente para qualquer localização do depósito. Esse procedimento permite treinar o sistema linear de equações antes de escolher o *grid* de interpolação (MARTIN; BOISVERT, 2017).

#### 2.4.3.1 O *kernel*

O *kernel* radial estabelece as relações espaciais entre todos os pontos do domínio. O *kernel* é sinônimo da função covariância usada na krigagem, e pode/deve ser parametrizado a partir do variograma, calculado e modelado a partir das amostras. Para modelagem geológica o *kernel* Gaussiano, representado pela Equação 2.10, é o mais indicado. Na prática tanto o variograma das funções distância assinaladas: que são modelados com um pequeno efeito pepita (para evitar instabilidades matemáticas e/ou controlar o ruído do modelo) e estrutura Gaussiana, quanto o variograma dos indicadores: que é modelado com uma estrutura esférica (MARTIN; BOISVERT, 2017), são adequados para a parametrização do *kernel*. Os programas da biblioteca GSLib trabalham apenas com uma única estrutura.

$$\phi(r) = \exp^{-\epsilon^{-2}r^2} \quad (2.10)$$

onde  $\epsilon$  é o parâmetro de suporte, que define o "range" do *kernel*.

Para parametrizar as funções de bases radiais com o modelo de variograma (das distâncias ou indicadores), a distância de suporte do *kernel* pode ser estimada a partir do alcance na direção de maior continuidade do variograma análogo. Os ângulos e relações de anisotropia são parametrizados a partir da orientação e alcances dos variogramas análogos (azimuth, dip, rake,  $r1 = \frac{a_{hmin}}{a_{hmax}}$ ,  $r2 = \frac{a_{hvert}}{a_{hmax}}$ ).

Fasshauer (2007) advoga que o *kernel* pode ser parametrizado pela distância que representa a maior esfera que pode ser colocada entre amostras no domínio. Para modelagem geológica essa técnica tende a super estimar o suporte (MARTIN; BOISVERT, 2017).

#### 2.4.3.2 Decomposição do domínio

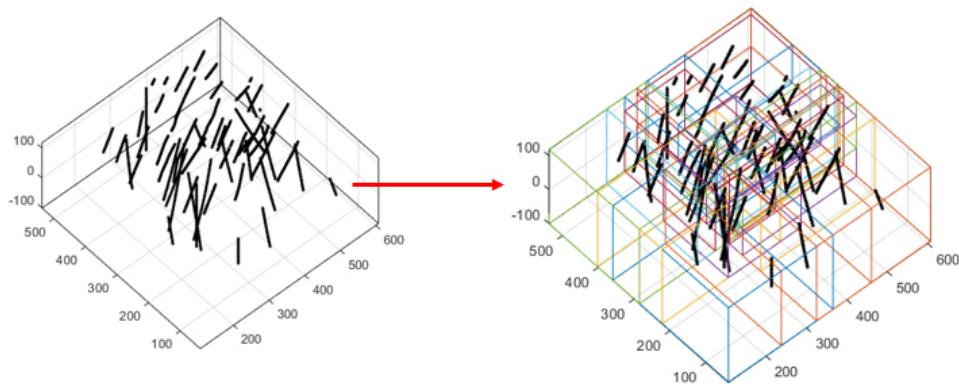
Como abordado na subseção 2.4.2, a limitação dos interpoladores globais são bancos de dados volumosos. Existem algumas técnicas para superar essa limitação como a solução iterativa (BEATSON; CHERRIE; MOUAT, 1999) por exemplo. Porém, um método que apresenta resultados satisfatórios e está presente na biblioteca GSLib é a decomposição do domínio (*Partition of unity - POU*) (WENDLAND, 2004), que transforma um problema volumoso e que demanda muito esforço computacional em diversos problemas menores e eficientes que são, ao final, unidos.

A demonstração das equações pode ser encontrada em [Wendland \(2004\)](#) e [Martin e Boisvert \(2015\)](#). Um domínio  $A$  é subdividido em uma série de partições sobrepostas  $\beta$  de modo que a união de todas as  $k$  partições em  $A$ ,  $\{\beta\}_{j=1}^K$  compreende o domínio. Para cada partição, os dados correspondentes são utilizados para interpolar a função escalar  $s_j(x)$ .

A função peso para cada partição deve ser igual a um no centro e 0 na fronteira, a contribuição de cada partição nos locais de sobreposição é dado pela função peso.

As partições são definidas encontrando uma série de coordenadas para os centros. Diversos métodos podem ser utilizados: um *grid* regular, k-means, árvore binária, *oct tree*.

Figura 17 – Particionamento.



Fonte: [Martin e Boisvert \(2015\)](#)

A [Figura 17](#) mostra um domínio particionado: o particionamento começa considerando o domínio completo como sendo uma partição, então a partição "total" é subdividida recursivamente em duas mantendo um nível de sobreposição. O particionamento termina quando cada partição individual tem menos amostras que o máximo especificado pelo usuário ([MARTIN; BOISVERT, 2017](#)). Dessa forma áreas com alta densidade amostral ficam em partições pequenas enquanto áreas esparsas ocupam partições maiores.

Deve haver um número de amostras em cada partição e sobreposição das partições suficientes para evitar o surgimento de artefatos quando as partições independentes forem unidas.

## 2.5 Visualização do modelo geológico

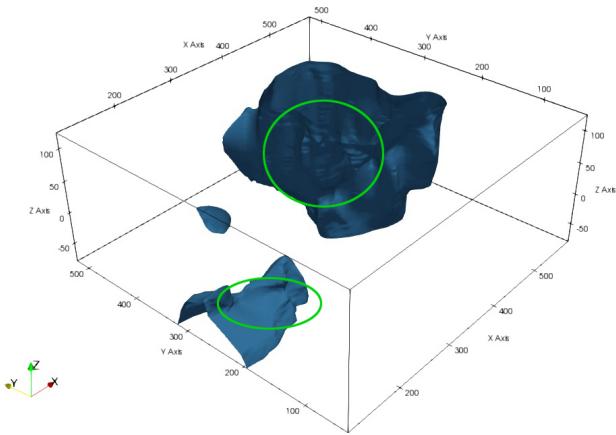
Como as funções distância são negativas no interior do domínio e positivas no exterior, um bom palpite inicial para a interface que separa os domínios no espaço, seria a linha (em duas dimensões) ou superfície (em três dimensões) que corresponda ao valor zero da função distância assinalada ([WILDE; DEUTSCH, 2011a](#)).

A [Figura 18](#) mostra a iso superfície que representa a categoria 1, extraída no software Paraview ([AHRENS; GEVECI; LAW, 2005](#)), dos diferentes modelos implícitos da [Figura 15](#). O algoritmo de extração de iso superfícies mais comumente utilizado é o dos cubos marchantes ([LORENSEN; CLINE, 1987](#)).

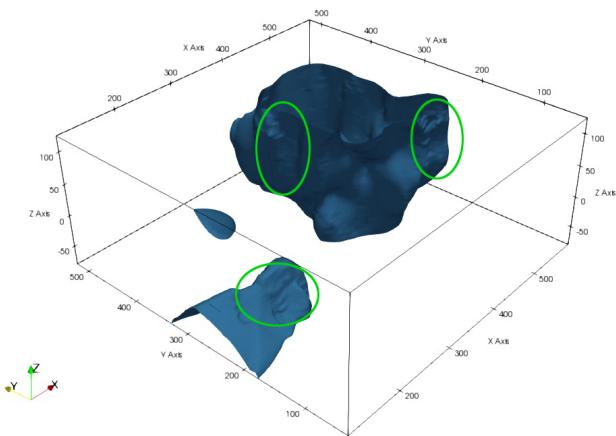
Note que os artefatos gerados nos modelos implícitos quando as distâncias foram interpoladas por krigagem são transferidos para as respectivas iso superfícies da [Figura 18](#) e realçados em verde. Daí a importância de gerar modelos implícitos livres de artefatos.

Figura 18 – Iso superfícies para a categoria 1 extraída dos diferentes modelos implícitos.

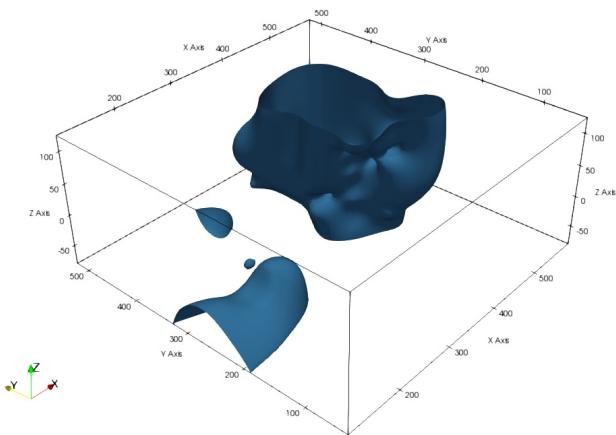
(a) OK com 40 amostras



(b) OK com 100 amostras



(c) RBF global



Na presença de múltiplos domínios, as iso superfícies devem ser extraídas uma a uma e algum tipo de regra hierárquica baseada na sequência estratigráfica ou ordem dos eventos geológicos deve ser aplicada para a criação dos modelos multi categóricos.

As Figura 19 e Figura 20 mostram iso superfícies extraídas dos modelos implícitos interpolados por RBF e parametrizados a partir do variograma dos indicadores juntamente com as amostras categóricas no *grid* fino.

Figura 19 – Iso superfície extraída do modelo implícito interpolado por RBF para a categoria 1.

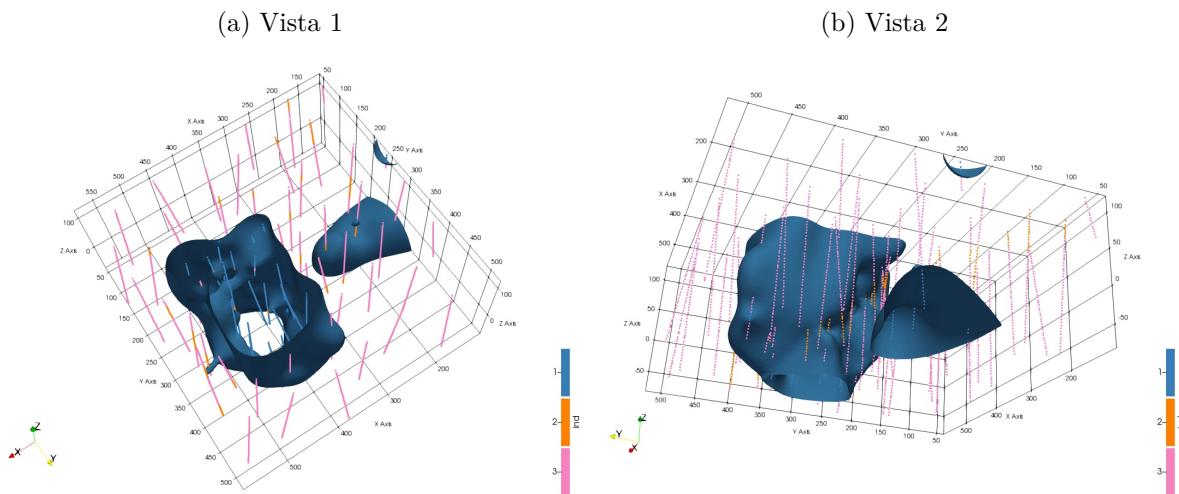
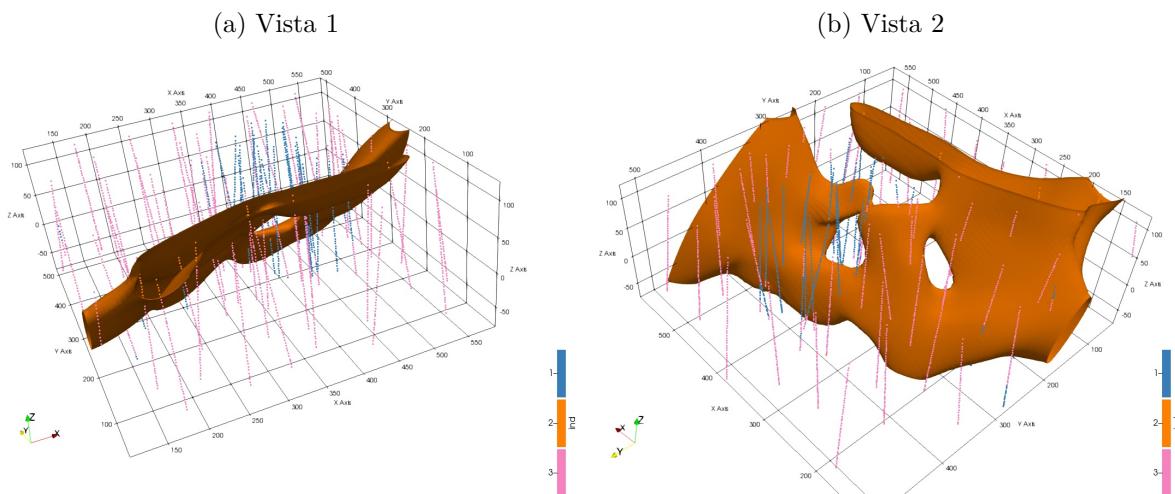


Figura 20 – Iso superfície extraída do modelo implícito interpolado por RBF para a categoria 2.



### 2.5.1 Adaptação do método para múltiplas categorias simultâneas

Silva e Deutsch (2012) propuseram, uma adaptação para o método com o objetivo de modelar múltiplos domínios geológicos simultaneamente, de forma similar ao caso binário.

Se existirem  $K$  múltiplos domínios no depósito mineral, para cada ponto amostral  $z(u_\alpha)$ ,  $\alpha = 1, \dots, n$ , um vetor de indicadores de  $K$  elementos é codificado de acordo com a [Equação 2.11](#).

$$i_k(u_\alpha) = \begin{cases} 1, & \text{se } z(u_\alpha) = k \\ 0, & \text{se } z(u_\alpha) \text{ caso contrário} \end{cases} \quad k = 1, \dots, K \quad (2.11)$$

A função distância é calculada, individualmente para cada elemento  $k$  do vetor, de acordo com a [Equação 2.12](#).

$$d_k(u_\alpha) = \begin{cases} - \| u_\alpha - u_\beta \|, & \text{se } i_k(u_\alpha) = 1 \\ + \| u_\alpha - u_\beta \|, & \text{se } i_k(u_\alpha) = 0 \end{cases} \quad k = 1, \dots, K \quad (2.12)$$

As distâncias calculadas são então interpoladas pelo método escolhido, individualmente para todos os nós do *grid* de acordo com a [Equação 2.13](#).

$$d_k^*(u) = \sum_{\alpha=1}^n \lambda_\alpha(u) d_k(u_\alpha) \quad k = 1, \dots, K \quad (2.13)$$

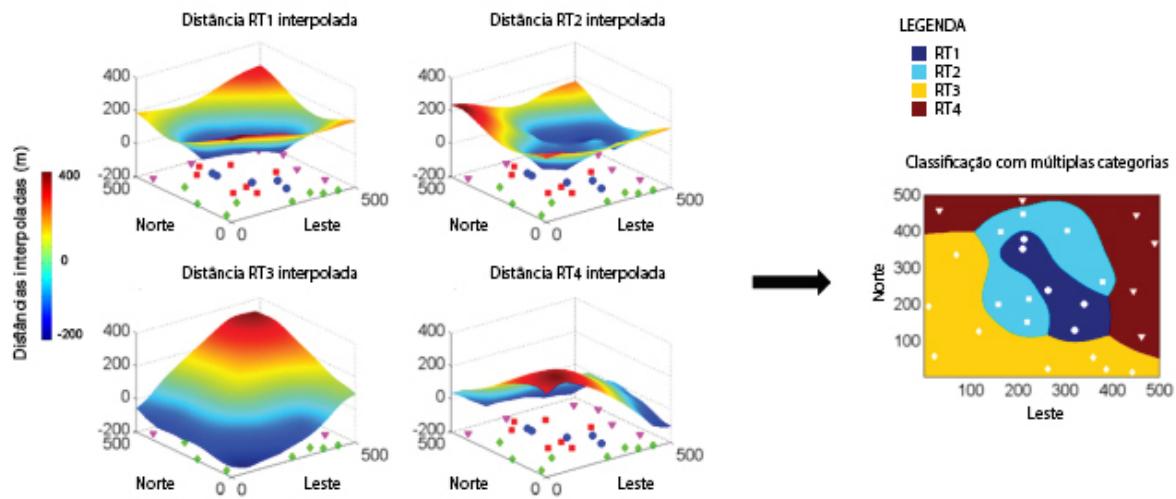
Por fim, cada bloco é classificado pela [Equação 2.14](#)

$$i^*(u) = k' \text{ de modo que } d_{k'}^* = \min \{d_k^*(u)\}_{k=1}^K \quad (2.14)$$

As distâncias estimadas fornecem uma medida de proximidade ao domínio oposto mais próximo. Sendo assim, a mínima distância assinalada estimada é tida como o domínio mais provável de ser encontrado numa região não amostrada. A [Equação 2.14](#) sumariza essa ideia ([SILVA, 2015](#)). A categoria associada com a menor distância estimada é retida em cada bloco.

A [Figura 21](#) mostra um exemplo de um modelo geológico criado a partir da mínima distância assinalada estimada. A esquerda, a figura mostra a projeção das distâncias assinaladas para cada uma das quatro categorias, e à direita a classificação final.

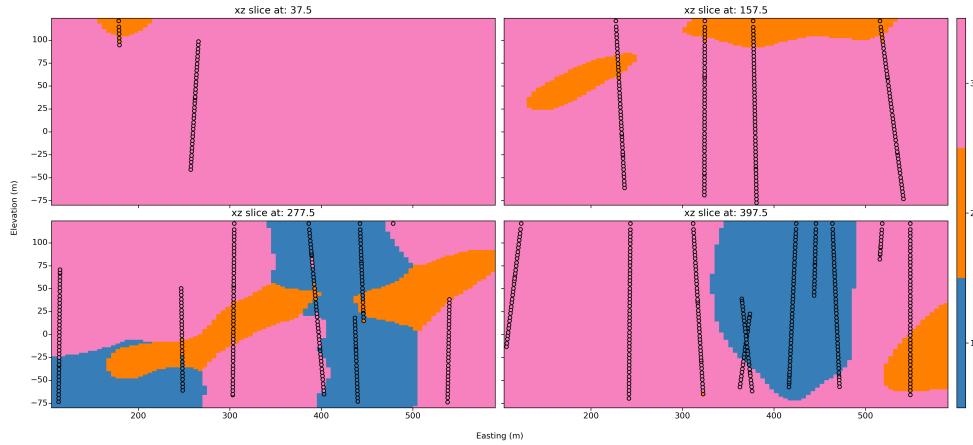
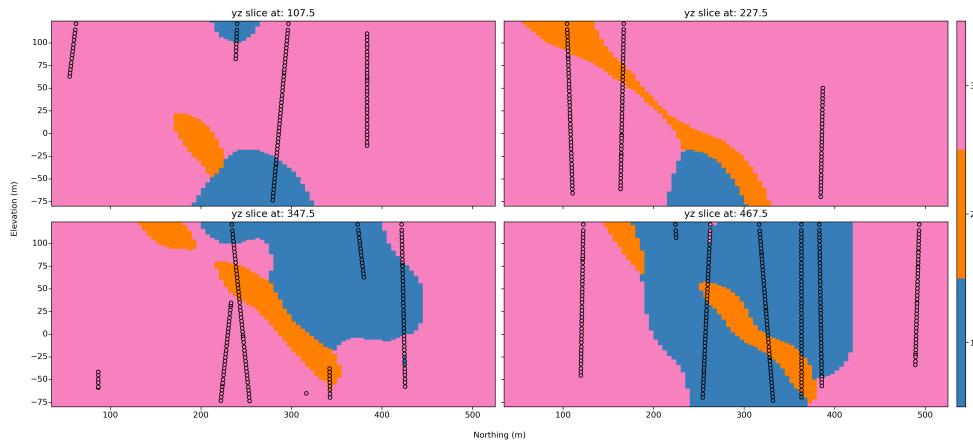
Figura 21 – Esquema para criação de um modelo implícito multi categórico.



Fonte: Modificado de [Silva \(2015\)](#)

A [Figura 22](#) mostra seções verticais em XZ e YZ de um modelo geológico implícito multi categórico criado por RBF parametrizado pelo variograma dos indicadores para o banco de dados do estudo de caso no *grid* fino. Nas seções também são mostradas amostras a 3 metros do centroide dos blocos na direção X e Y.

Figura 22 – Modelo geológico multi categórico.

(a) Seções em XZ do modelo implícito gerado por RBF no *grid* fino.(b) Seções em YZ do modelo implícito gerado por RBF no *grid* fino.

A variação do método para múltiplos domínios ou categorias simultâneos deve ser utilizado com cautela, em muitos casos, em especial em depósitos com geologia complexa que apresentam intrusões, falhas e dobras é indicado a modelagem e análise das litologias individualmente, ou até mesmo com algum nível de interpretação geológica manual via digitalização de seções.

## 2.6 Incorporação da não estacionariedade de segunda ordem

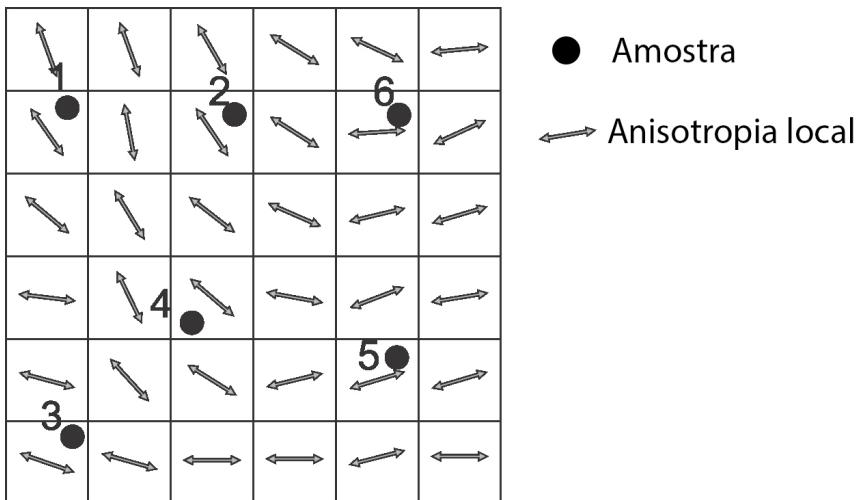
Corpos geológicos são complexos da escala macro à escala microscópica. Essa característica pode tornar difícil a captura de suas feições com funções de covariância. Segundo Martin e Boisvert (2017) a não estacionariedade de segunda ordem deve ser incorporada quando estruturas complexas estão sendo modeladas.

Anisotropia é modelada na geostatística como um conjunto de rotações e relações anisotrópicas: *azimuth*, *dip*, *rake*,  $r1 = \frac{a_{hmin}}{a_{hmax}}$ ,  $r2 = \frac{a_{vert}}{a_{hmax}}$ .

### 2.6.1 Krigagem com anisotropia local variável

A krigagem com anisotropia variável exige que os parâmetros locais de anisotropia sejam definidos em todos os nós do *grid* (Figura 23), e variem de forma suave pelo domínio. Isso permite que estruturas curvilineares em escala menor que o espaçamento entre as amostras sejam capturadas (MARTIN; BOISVERT, 2017). Porém, torna o método computacionalmente exigente.

Figura 23 – Esquema mostrando os vetores de anisotropia local para cada nó do *grid*.



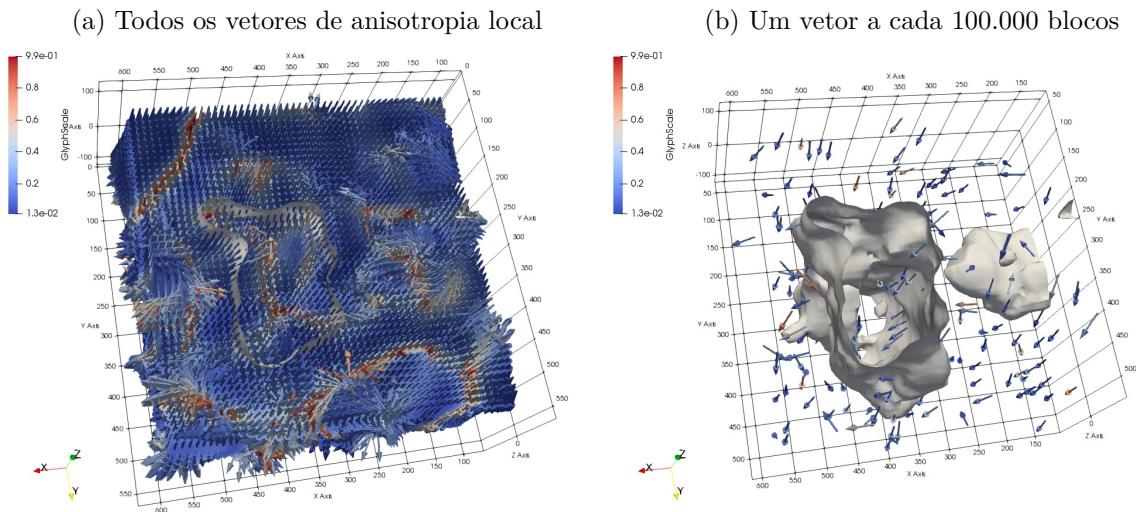
Fonte: Martin e Boisvert (2017)

Existem diversos métodos para definição da anisotropia local (LILLAHI; BOISVERT, 2015): pode ser inferida por um geomodelador, estimada diretamente a partir dos dados usando técnicas automáticas, interpolada de uma série de pontos onde a orientação foi medida ou pode ser extraída de dados em um *grid* que representam a variabilidade local (teores krigagdos, por exemplo). A biblioteca GSLib tem diversos softwares para definição de anisotropia local.

A Figura 24 mostra a iso superfície para a categoria 1 extraída de um modelo implícito gerado por krigagem com anisotropia local variável. Os vetores de anisotropia são plotados e foram gerados pelo software `imorient` da biblioteca GSLib a partir de um modelo implícito para a categoria 1 previamente interpolado. A interpolação foi feita no *grid* grosso, a condição da anisotropia local para cada nó do *grid* torna o algoritmo computacionalmente exigente.

Esse programa analisa uma janela centrada em cada nó e estima a orientação dominante nessa janela, uma janela pequena captura variações em menor escala e uma janela grande variação de larga escala. É preciso encontrar um balanço (LILLAH; BOISVERT, 2015).

Figura 24 – Iso superfície para a categoria 1 extraída de um modelo implícita gerado por krigagem com anisotropia local variável mostrando os vetores.

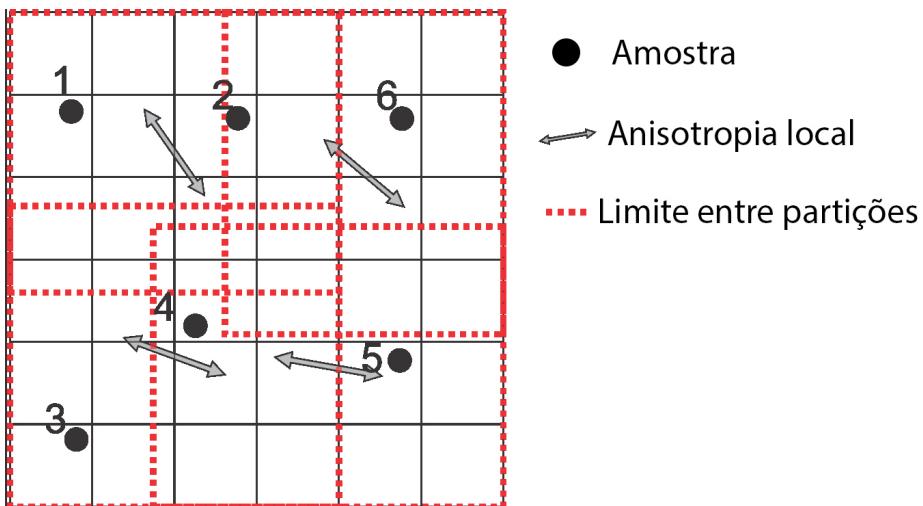


## 2.6.2 Funções de bases radiais com anisotropia local variável

RBF com anisotropia local variável é baseado na decomposição de domínios (subseção 2.4.3.2), um kernel anisotrópico pode ser usado para melhor suportar os dados em cada partição, já que as partições são independentes uma matriz de rotação é usada para definir *kernels* anisotrópicos locais que melhor se ajustam às propriedades espaciais de cada partição. O interpolador de bases radiais é obtido independentemente para cada partição e a solução final é obtida ponderando as soluções independentes nos locais de sobreposição (MARTIN; BOISVERT, 2017).

Ao contrário da krigagem com anisotropia local variável, que requer os vetores de anisotropia em todas as nós do *grid*, a implementação para funções de bases radiais requer os vetores somente no centro de cada partição (Figura 23) e que variem suavemente pelo domínio. Na figura é possível observar quatro partições sobrepostas cada uma com sua definição de anisotropia local.

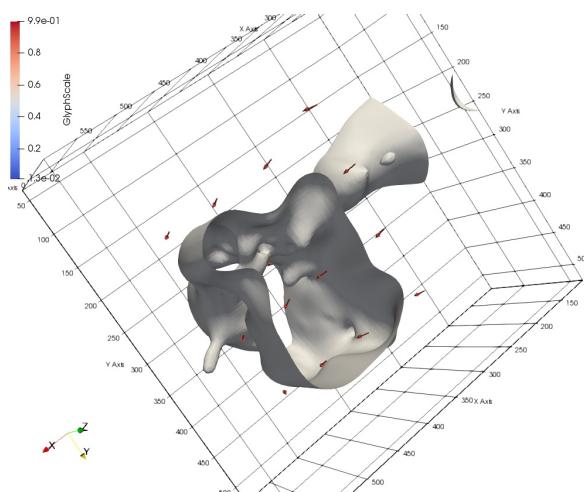
Figura 25 – Esquema mostrando os vetores de anisotropia local para cada centro de partição.



Fonte: [Martin e Boisvert \(2017\)](#)

A [Figura 26](#) mostra a iso superfície para a categoria 1 extraída de um modelo implícito gerado por funções de bases radiais com anisotropia local variável. Os vetores de anisotropia são plotados e foram gerados pelo software `rbfiterref` ([MARTIN; BOISVERT, 2015](#)) da biblioteca GSLib. A interpolação foi feita no *grid* fino.

Figura 26 – Iso superfície para a categoria 1 extraída de um modelo implícita gerado por krigagem com anisotropia local variável mostrando os vetores.



### 2.6.3 Refinamento iterativo

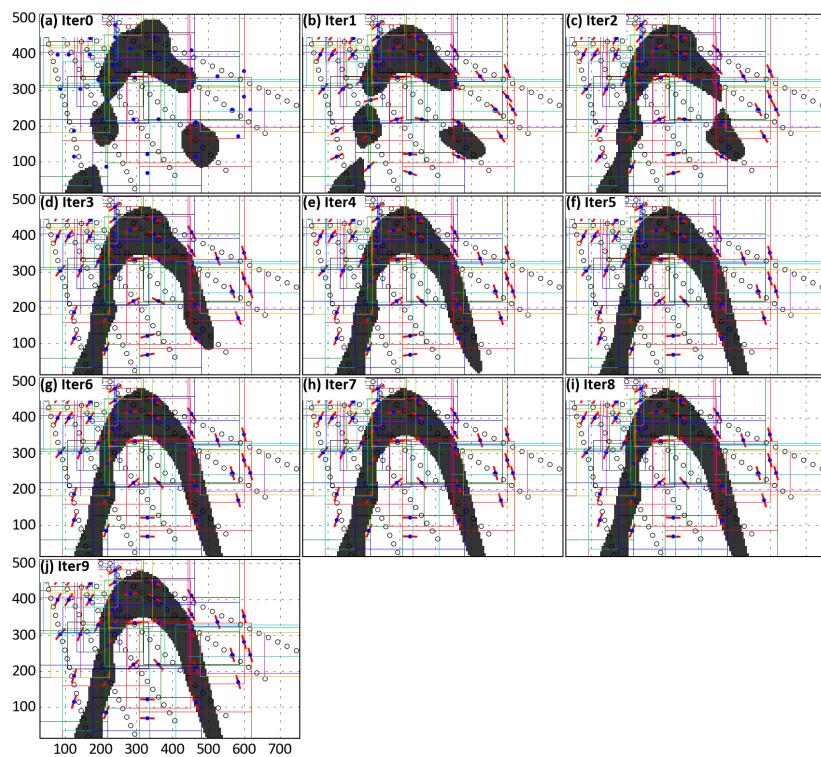
Extrair orientações locais de um modelo criado com anisotropia global e utilizar essas orientações em uma nova interpolação não estacionária resulta em um modelo

geológico mais refinado considerando informação local. Esse processo pode ser repetido, continuamente refinando o campo de anisotropias locais.

O refinamento iterativo com funções de bases radiais tem um benefício extra (em relação à krigagem com variação local de anisotropia), a possibilidade de incluir algum tipo de critério de parada, quando o modelo for considerado refinado. Como as partições são independentes, se o refinamento de uma partição em particular não resultará em mudanças significativas, pode ser omitido na solução final. Essa característica torna o processo significativamente mais rápido (MARTIN; BOISVERT, 2017).

A Figura 27 mostra um processo de refinamento iterativo onde a semente é um modelo isotrópico. O programa `rbfiterref` começa sua execução inferindo orientações anisotrópicas em cada flanco da dobra, que são rasos demais para que sejam definidos adequadamente. Porém, nas próximas iterações do refinamento, o mergulho se torna mais profundo e os flancos são definidos implicitamente.

Figura 27 – Esquema mostrando etapas usando o refinamento iterativo para funções de bases radiais.



Fonte: Martin e Boisvert (2017)

## 2.7 Incorporação de informação secundária

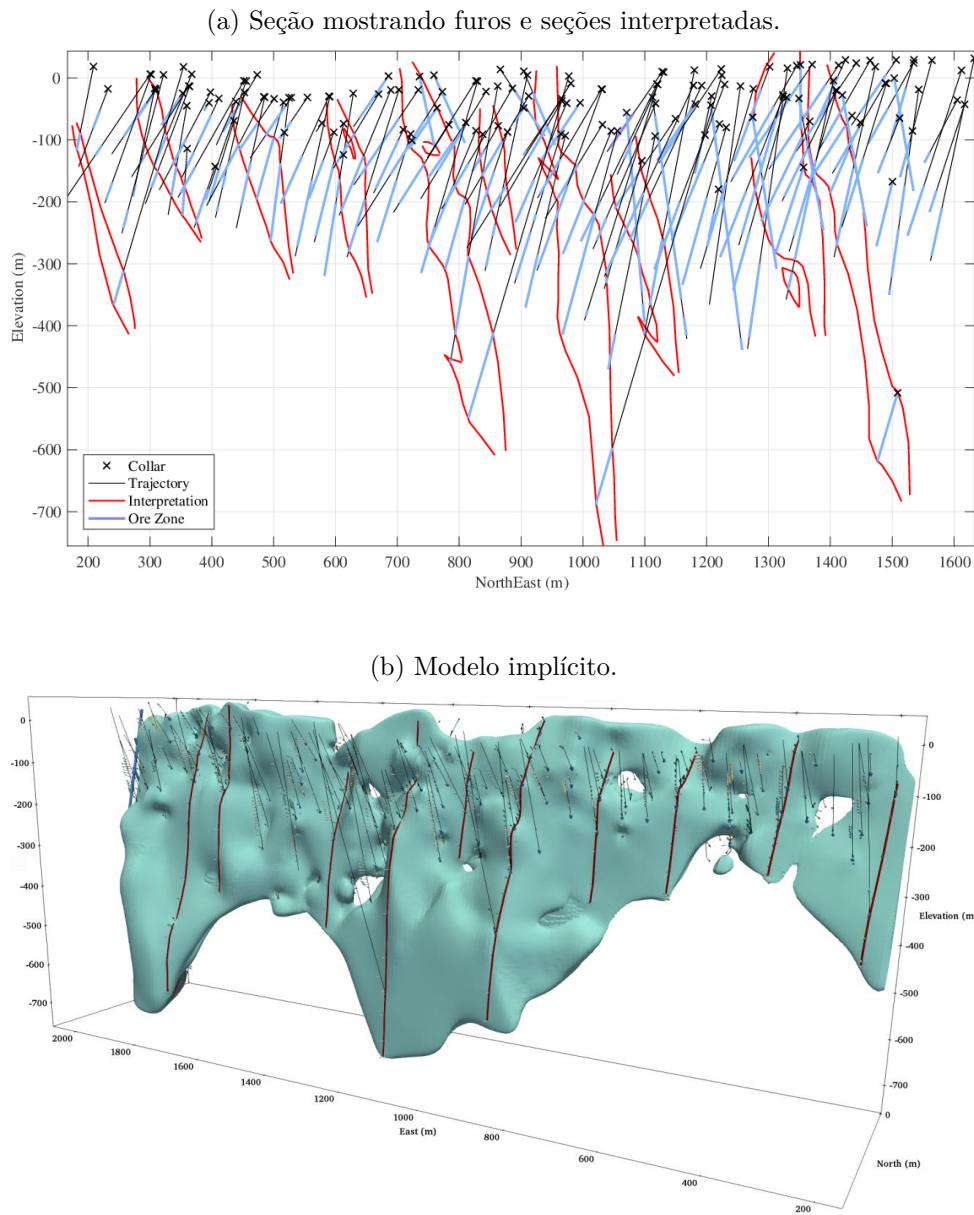
Modelos explícitos englobam julgamento profissional que é difícil de superar por qualquer algoritmo implícito devido à complexidade de representar objetos naturais com funções numéricas, principalmente quando os dados são esparsos.

Uma desvantagem da modelagem implícita é que muitas vezes o modelo resultante não está de acordo com a interpretação e expectativa do geomodelador. Assistência humana usualmente é necessária para construir modelos geológicos semi automáticos utilizando técnicas que utilizem informação explícita e implícita como a interpolação suave discreta (*Discrete smooth interpolation - DSI*) ([MALLET J., 2002](#)). Nas primeiras fases da mineração a proporção explícita da modelagem é maior, a medida que mais dados são obtidos, a proporção de intervenção humana é menor ([MANCHUCK; DEUTSCH, 2018](#)).

O uso da krigagem ou funções de bases radiais limita os dados ao suporte puntual. [Manchuck e Deutsch \(2018\)](#) propõe o uso de uma regressão linear local chamada mínimos quadrados móveis (*Moving least squares - MLS*) que ajusta uma função aos dados para integrar modelagem geológica implícita e explícita. MLS é um algoritmo de interpolação geral que pode ser usado para ajustar uma função em pontos, linhas e polígonos. MLS foi escolhido por sua capacidade de interpolar dados em linha de forma exata. Representar as linhas como pontos discretos ou utilização de integração numérica não é necessário.

[Manchuck e Deutsch \(2018\)](#) conduziram um estudo de caso em um banco de dados real de ouro, o banco de dados e o resultado final podem ser vistos na [Figura 28](#).

Figura 28 – Modelo geológico híbrido criado a partir de furos de sondagem e seções interpretadas.



Fonte: [Manchuck e Deutsch \(2018\)](#)

A incerteza pode ser avaliada pelo método por meio do parâmetro de incerteza C e simulações não condicionais.

A integração da modelagem implícita a partir das amostras e da interpretação geológica feita pelo geomodelador, a partir dos dados e do seu conhecimento prévio, é muito promissora. Alia a rapidez e simplicidade da metodologia implícita com a experiência do geomodelador, eliminando a necessidade de interpretar um número elevado de seções. O método é recente e ainda apresenta alguns desafios operacionais, principalmente em

relação a exigência de estruturas planares e avaliação de incerteza.

## 2.8 Avaliação da incerteza

Sempre existirá algum grau de incerteza nos volumes geológicos modelados, que varia de acordo com o número de amostras e a informação geológica disponível. A incerteza do modelo geológico, muitas vezes, pode ser a maior fonte de incerteza do projeto mineiro. Por esse motivo, diversos métodos foram desenvolvidos com o objetivo de avaliar a incerteza dos modelos geológicos a partir das funções distâncias assinaladas.

Essa seção abordará os métodos disponíveis na literatura com exemplos práticos no banco de dados apresentado.

### 2.8.1 Avaliação heurística da incerteza

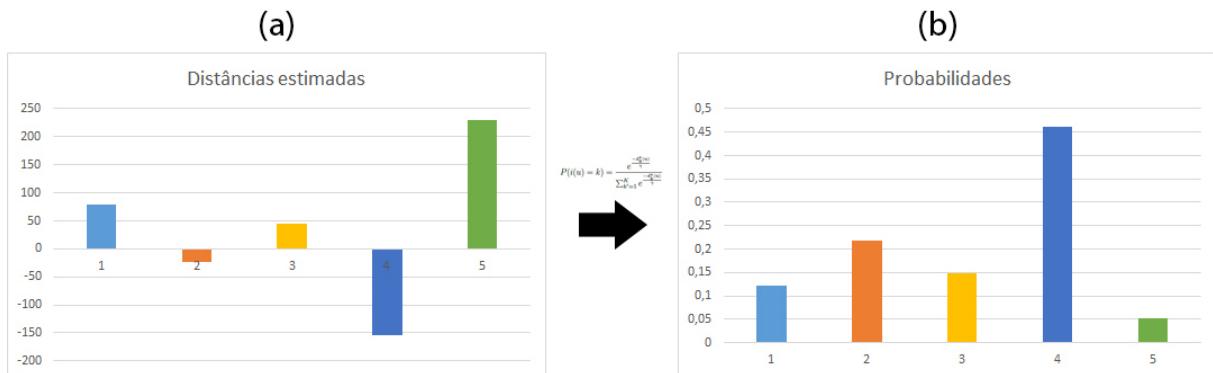
O método mais simples de avaliação de incerteza em modelos geológicos implícitos, proposto por [Silva e Deutsch \(2012\)](#), consiste em transformar distâncias estimadas em probabilidades via *softmax transformation* essa técnica é amplamente usada em métodos de classificação para múltiplas classes ([MCCULLAGH; NELDER, 1989](#)). A ideia é transformar as distâncias estimadas em probabilidades a partir da [Equação 2.15](#). Os valores transformados encontram-se entre zero e um e sua soma deve ser igual a um, para cada bloco estimado.

$$P(i(u) = k) = \frac{e^{\frac{-d_k^*(u)}{\gamma}}}{\sum_{k'=1}^K e^{\frac{-d_{k'}^*(u)}{\gamma}}} \quad (2.15)$$

$P(i(u) = k)$  representa a probabilidade de um local  $u$  pertencer à categoria  $k$ ,  $d_k^*(u)$  é a distância estimada para a categoria  $k$  e  $\gamma$  é um parâmetro que regula a inter-relação entre as probabilidades das  $K$  diferentes categorias. Quanto maior  $\gamma$ , maior as diferenças entre as probabilidades (maior a banda de incerteza), como pode ser observado na [Figura 30](#). [Silva e Deutsch \(2013\)](#) advogam que o parâmetro pode ser a maior distância estimada entre todas as categorias.

Para exemplificar a transformação, a [Figura 29](#) mostra à esquerda as distâncias estimadas para cada uma de cinco categorias em um mesmo bloco em particular. À direita as distâncias transformadas pela [Equação 2.15](#) em probabilidades daquele bloco pertencer a cada uma das cinco categorias. Observe que distâncias menores (mais negativas) geram probabilidades maiores, e vice-versa.

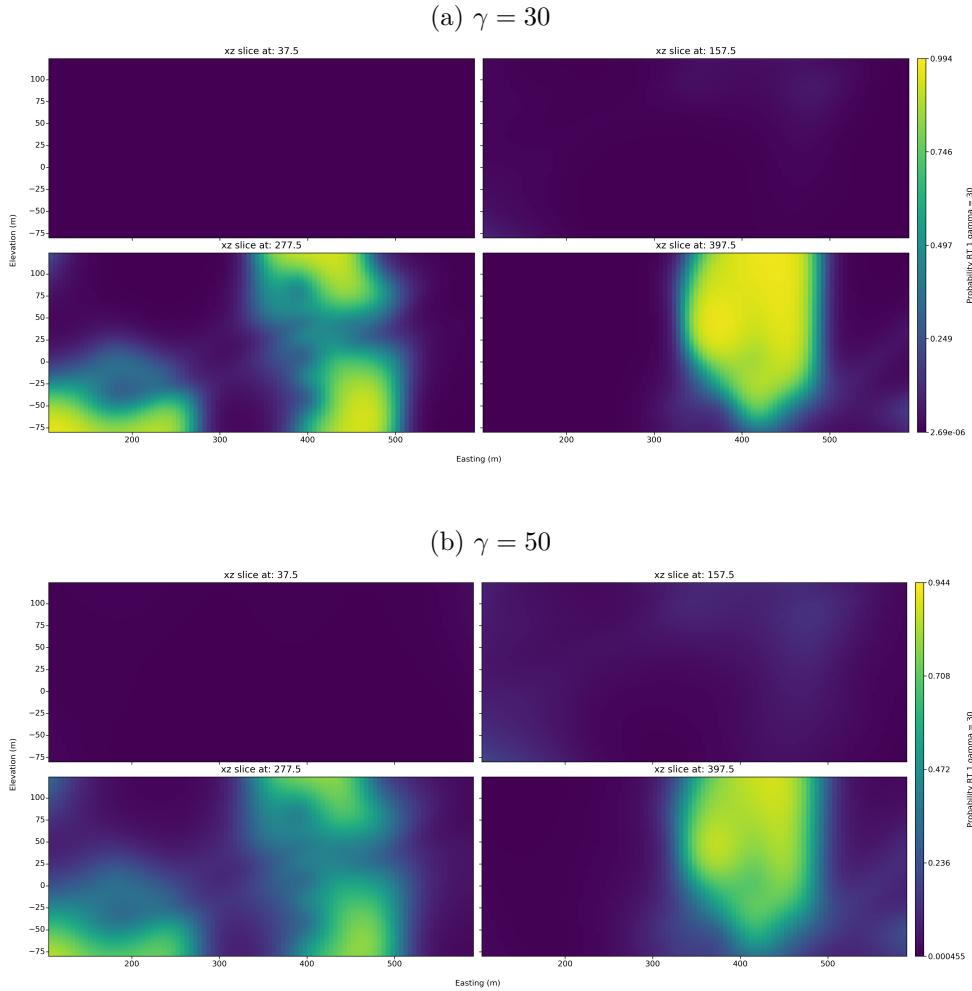
Figura 29 – Distâncias estimadas em (a) e transformadas em probabilidades em (b) para um mesmo bloco, com cinco categorias.



Fonte: [Rolo \(2017\)](#)

A [Figura 30](#) mostra seções em XZ do mapa de probabilidades para a categoria 1, com diferentes valores para  $\gamma$ .

Figura 30 – Seções em XZ do mapa de probabilidades para a categoria 1 com diferentes valores para  $\gamma$ .



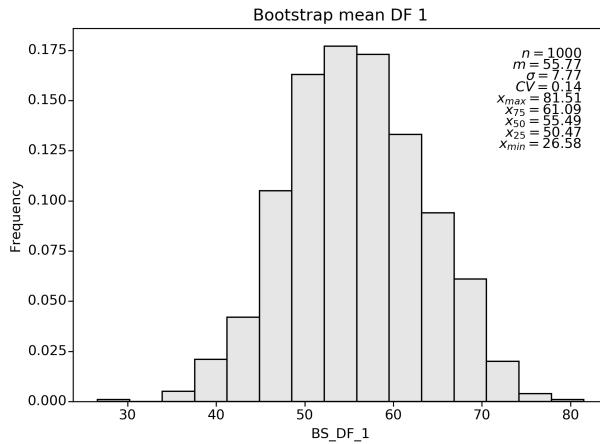
Apesar do método ser simples e direto, suportar múltiplas categorias simultaneamente, seu resultado não pode ser utilizado em etapas subsequentes do processo de avaliação, por não gerar múltiplas realizações do modelo geológico. Além disso, A escolha do parâmetro  $\gamma$  é subjetiva e depende da distribuição espacial das amostras, dependendo da escolha para o parâmetro nem mesmo blocos colocados com amostras recebem a probabilidade 100%. Alguns autores recomendam que o parâmetro  $\gamma$  seja igual a maior distância estimada entre todas as categorias.

### 2.8.2 *Boundsim*

Uma outra metodologia proposta por [McLennan e Deutsch \(2006\)](#) consiste em realizar um *bootstrap* espacial ([DEUTSCH, 2004](#)) da média da distância assinada calculada para uma dada categoria. O procedimento de *bootstrap* é amplamente utilizado para quantificar incerteza em parâmetros estatísticos. As duas mais importantes suposições são:

os dados são representativos e os dados são independentes. Por esse motivo, a técnica de amostragem deve ser modificada para variáveis que apresentem continuidade espacial. A Figura 31 mostra o histograma da média para as distâncias assinaladas da categoria 1 do banco de dados.

Figura 31 – Histograma do bootstrap espacial da média das distâncias assinaladas para a categoria 1.



Valores para a média são tomados do histograma da média, gerado pelo *bootstrap* espacial. Esses valores podem ser correspondentes ao p10, p50 e p90 por exemplo.

Então, as distâncias assinaladas são interpolados para os nós do *grid* por krigagem simples, que tem como hipótese a estacionariedade da média por todo o domínio. A média estacionária para a krigagem simples são os valores tomados do histograma.

O último passo consiste em extrair a iso superfície zero dos modelos implícitos gerados por krigagem simples com os diferentes valores da média estacionária tomadas do histograma da média. Em tese, quanto maior o percentil menor o risco associado, o volume do sólido modelado infla ou contrai de acordo com o valor escolhido para a média estacionária.

A Tabela 4 mostra os blocos classificados como pertencentes à categoria 1 usando a metodologia apresentada, com valores de média estacionária correspondentes à p10, p50 e p90.

Percentil	Blocos dentro
10	109886
50	110446
90	111069

Tabela 4 – Blocos classificados como pertencentes à categoria 1.

A metodologia é simples e direta, porém, deve ser aplicada à uma categoria por vez e na presença de múltiplos domínios alguma abordagem hierárquica deve ser estabelecida. A sensibilidade da krigagem simples à média depende da configuração espacial das amostras, muitas vezes a diferença entre os volumes obtidos é irrelevante, como no caso do banco de dados em estudo: a diferença no volume do sólido do caso mais conservador para o de maior risco é de pouco mais de 1%. As iso superfícies são praticamente coincidentes.

O produto final do método pode ser utilizado nas etapas subsequentes do processo de avaliação, todavia, o método não avalia incerteza de forma adequada. Os autores do estudo original aplicaram a metodologia apenas em bancos de dados sintéticos e de geometria regular.

### 2.8.3 Simulação direta das distâncias assinaladas

[Cáceres et al. \(2011\)](#) e [Radtke \(2017\)](#) propuseram que as distâncias assinaladas sejam simuladas de forma direta: a partir de um modelo geológico implícito multi categórico ([subseção 2.5.1](#)), um coeficiente U de incerteza deve ser calculado de acordo com [Equação 2.16](#):

$$U(u) = \frac{\max\{D_{min}\} - \min\{d_k^*(u)\}_{k=1}^K}{\max\{D_{min}\} - \min\{D_{min}\}} \quad (2.16)$$

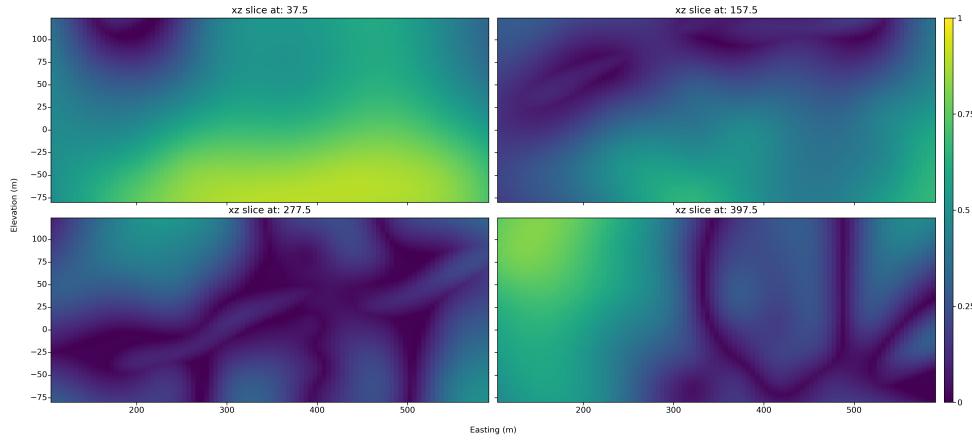
Onde:

$$D_{min} = \{\min\{d_k^*(u_1)\}, \dots, \{\min\{d_k^*(u_n)\}\}_{k=1}^K\} \quad (2.17)$$

E  $d_k^*(u)$  é a distância estimada no local u para a categoria k.

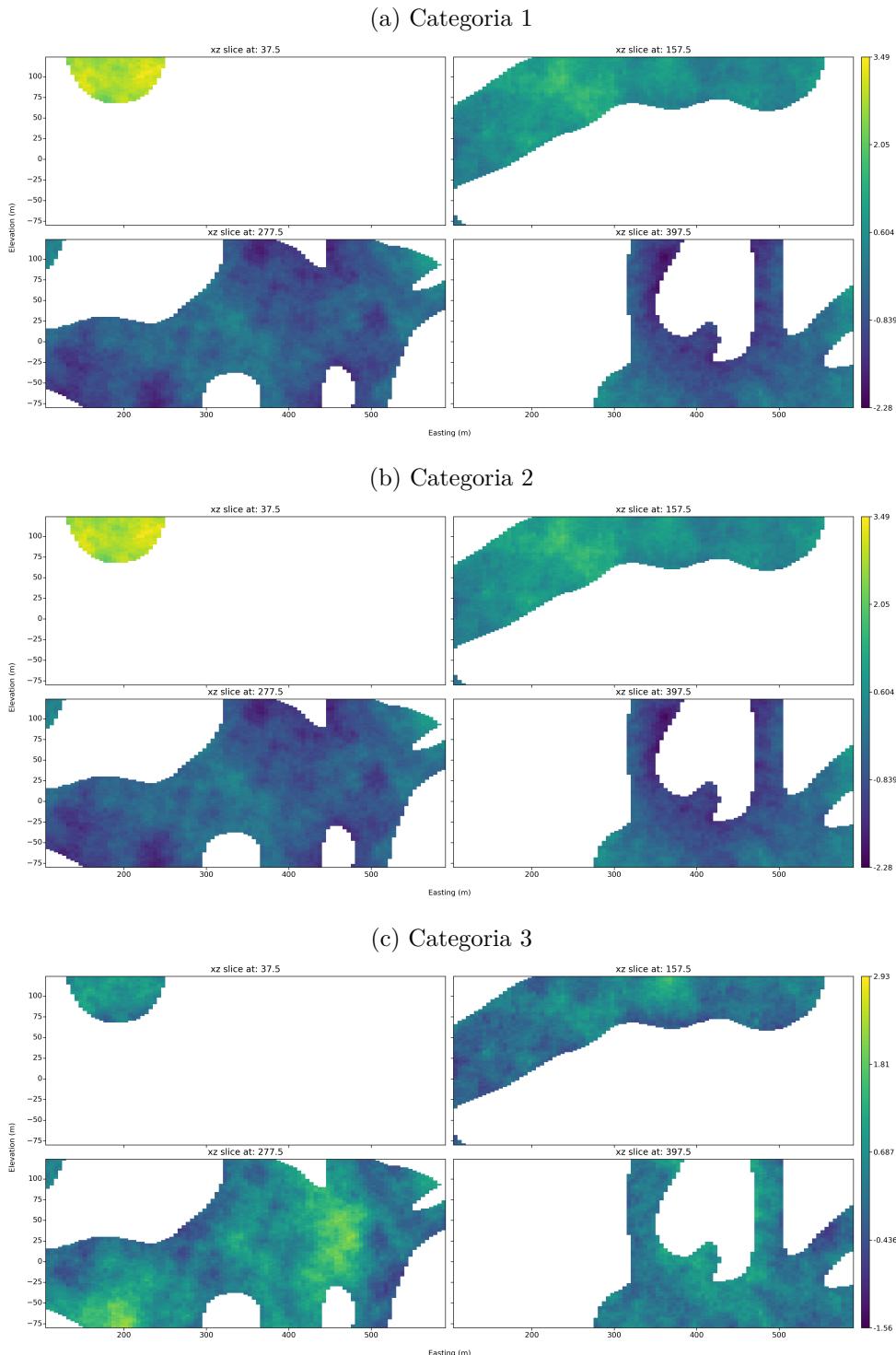
Os valores de U variam entre 0 e 1, um *cutoff* em U determina a extensão da zona de incerteza ao redor dos contatos: quanto mais próximo de 1, mais estreita a zona. A [Figura 32](#) mostra o coeficiente U calculado para todos os nós do *grid* a partir do modelo da [Figura 22](#).

Figura 32 – Coeficiente U calculado para todos os nós do *grid*.



Um *cutoff* em 0.8 foi aplicado em U e as distâncias assinaladas foram simuladas por simulação sequencial Gaussiana na região de incerteza. A [Figura 33](#) mostra as distâncias simuladas para as três categorias do banco de dados na zona de incerteza.

Figura 33 – Distâncias simuladas na zona de incerteza para as categorias do banco de dados.

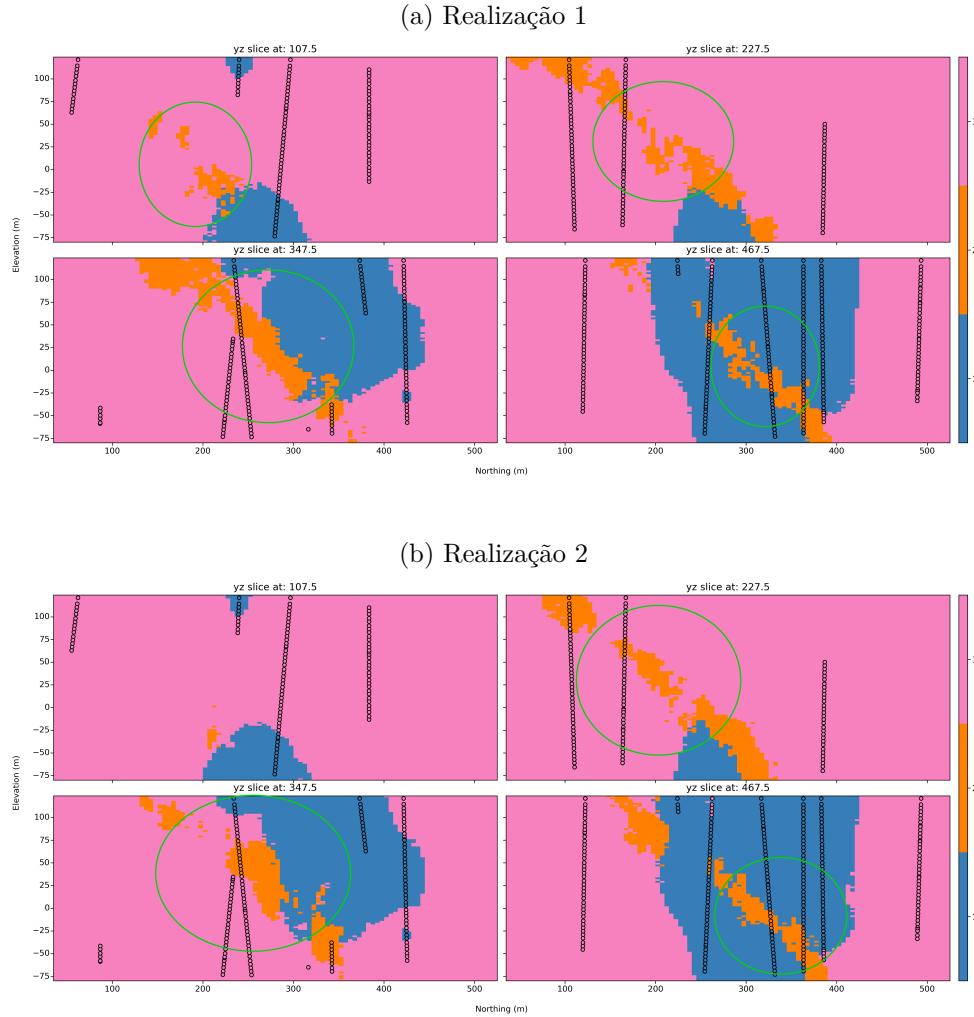


A categoria correspondente ao menor valor simulado é retida em cada bloco e os blocos simulados são concatenados com os blocos estáticos (congelados), atribuídos pela interpolação.

A Figura 34 mostra duas diferentes realizações do modelo geológico estocástico

criado a partir da simulação direta das distâncias.

Figura 34 – Diferentes realizações do modelo geológico.



O método é baseado em múltiplas realizações do modelo geológico, então, seu produto final pode ser usado nas etapas posteriores do processo de avaliação. Entretanto, a metodologia tem muitos passos ([RADTKE, 2017](#)):

1. Cálculo das distâncias assinaladas para todas as amostras e categorias;
2. Variografia das distâncias no espaço original para todas as categorias;
3. Interpolação das distâncias;
4. criação do modelo com base na menor distância interpolada;
5. Criação da zona de incerteza;
6. Transformação Gaussiana das distâncias;

7. Variografia das distâncias no espaço gaussiano para todas as categorias;

8. Geração de múltiplos modelos baseados na menor distância simulada;

9. Validação e pós processamento das realizações.

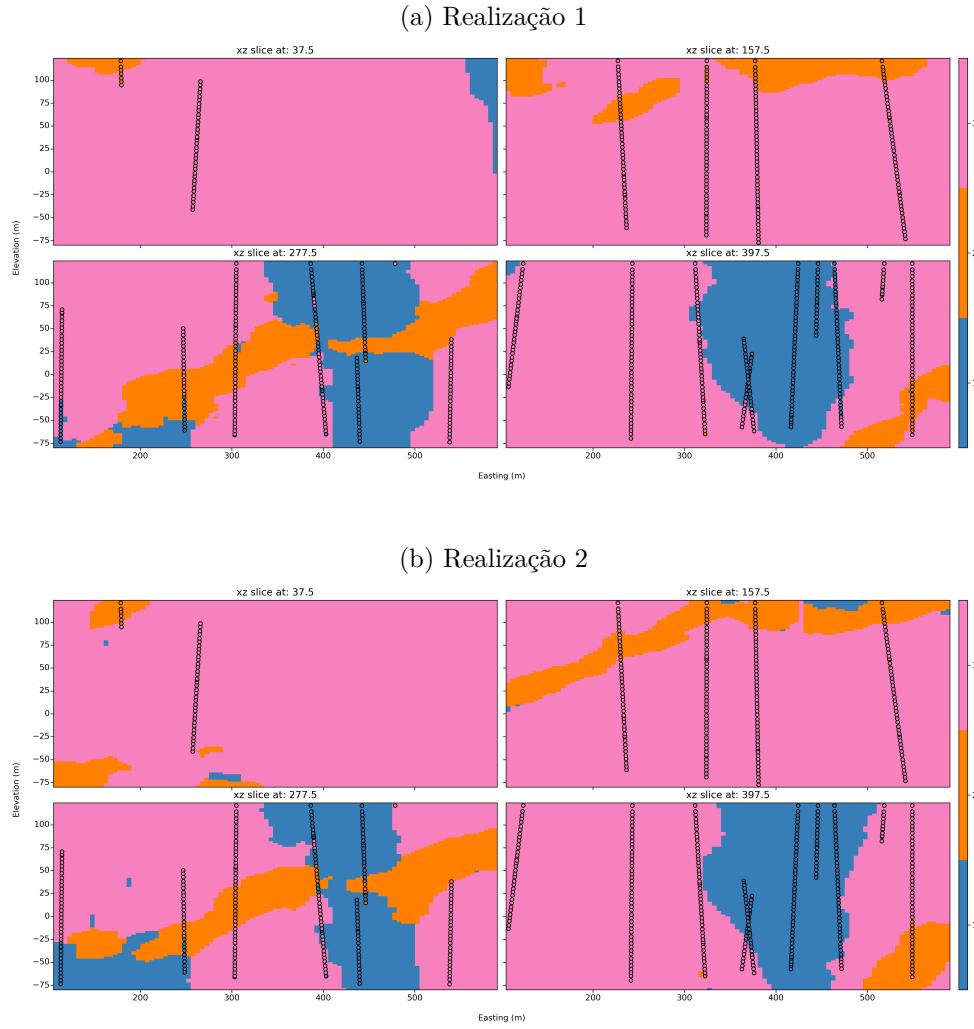
Além disso, a simulação é muito sensível aos parâmetros, o que muitas vezes, gera modelos com muito ruído, realçados em verde, que não são geologicamente realistas, e não honram as amostras, como os da Figura 34. O resultado final não justifica o número excessivo de passos e parâmetros. O coeficiente U não controla a magnitude da incerteza, apenas diminui o tempo das simulações, o magnitude da incerteza é controlada pelo parâmetros da simulação.

#### 2.8.4 Simulação multi ponto

Silva (2015) propôs uma metodologia que integra modelagem geológica implícita com funções distâncias assinaladas e simulação geostatística multi ponto. A geostatística multi ponto utiliza imagens de treinamento para extrair e reproduzir estruturas geológicas complexas. Geralmente, imagens de treinamento são baseadas em modelos conceituais do fenômeno geoético. Em seu trabalho, Silva (2015) cria imagens de treinamento a partir de modelos geológicos implícitos determinísticos criados com funções distância assinaladas e modelos estocásticos criados por simulação sequencial dos indicadores. No trabalho, é introduzida uma metodologia para integrar múltiplas imagens de treinamento, uma metodologia para calibrar a contribuição de cada imagem de treinamento e uma medida de entropia multi ponto ao longo dos furos de sondagem.

Silva (2015) concluiu que imagens de treinamento geradas a partir das amostras produzem modelos geológicos estocásticos por MPS melhores. A Figura 35 mostra duas realizações do modelo geológico criado a partir da aplicação do algoritmo **snesim** (STRÉBELLE; JOURNEL, 2000) usando a Figura 22 como imagem de treinamento.

Figura 35 – Diferentes realizações do modelo geológico.



Os modelos mostram realismo geológico, não apresentam ruido excessivo. Além disso, o método é rápido e não depende de um grande número de parâmetros. Em contrapartida, não é possível controlar a magnitude da incerteza, que depende apenas da imagem de treinamento. Não é possível definir uma zona de incerteza para que os limites variem em seu interior e também não é possível controlar a natureza dos contatos, suaves ou rugosos.

### 2.8.5 Boundary simulation

A metodologia proposta por [Wilde e Deutsch \(2011c\)](#) consiste em gerar realizações das fronteiras geológicas comparando distâncias modificadas interpoladas com simulações não condicionais.

O primeiro passo é calibrar os parâmetros de incerteza. [Munroe e Deutsch \(2008\)](#) propõem a calibração completa dos parâmetros  $C$  e  $\beta$ , onde  $C$  controla a largura da

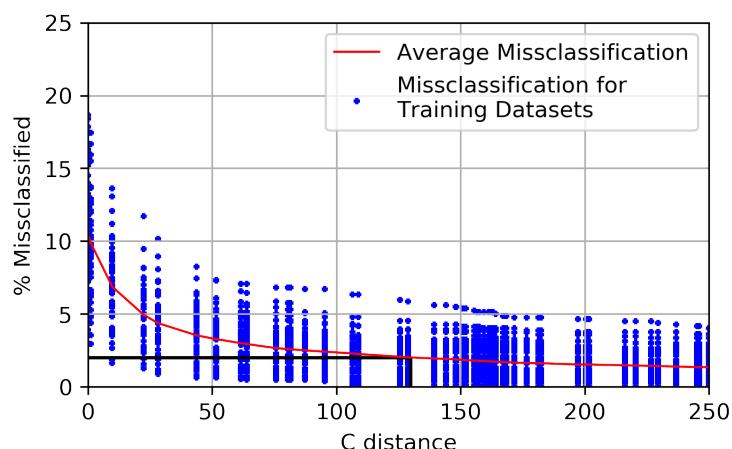
banda de incerteza e  $\beta$  controla o viés. Os parâmetros devem ser otimizados para que a incerteza apropriada seja avaliada. Otimizar esses parâmetros é uma operação computacionalmente cara que requer múltiplos modelos de referência e duas funções objetivo (WILDE; DEUTSCH, 2012).

Kaperov e Deutsch (2016) propuseram uma metodologia de calibração para o parâmetro C simplificada, de modo empírico, considerando a malha amostral e a experiência e conhecimento do geomodelador.

Wilde e Deutsch (2011b) propuseram uma outra metodologia de calibração. Nessa metodologia, apenas o parâmetro C é calibrado usando apenas as amostras disponíveis. Um subconjunto das amostras é removido antes do cálculo das distâncias assinaladas. Isso cria dois diferentes bancos dados: o banco de dados das distâncias assinaladas e o banco de dados *jackknife*. O banco de dados das distâncias assinaladas é usado para condicionar a estimativa das distâncias nos locais do banco de dados *jackknife*. Algumas amostras do banco de dados *jackknife* que foram codificadas como fora do domínio terão distâncias estimadas negativas (dentro do domínio) e vice-versa. O parâmetro C é ajustado até que uma proporção pré definida de classificação errônea seja atingida.

A Figura 36 mostra a calibração do parâmetro C para a categoria 1 do banco de dados. Um quarto das amostras foi removido de forma aleatória 50 vezes, gerando 50 bancos de dados de distância e 50 bancos de dados *jackknife*. As distâncias foram estimados nos locais das amostras *jackknife* e a taxa de classificação errônea foi registrada no eixo y. o parâmetro C foi incrementado, e somado ou subtraídos das distâncias de acordo com a Equação 2.18. O processo é repetido enquanto C varia de 0 até 250m. Quanto maior for o valor de C, menor o índice de classificação errônea.

Figura 36 – Calibração do parâmetro C para a categoria 1.

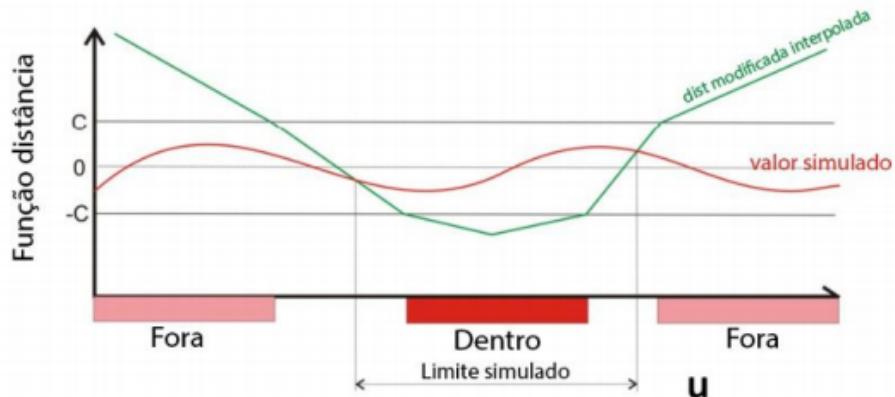


$$d_k(u_\alpha) = \begin{cases} -\|u_\alpha - u_\beta\| - C, & \text{se } u_\alpha \text{ pertence ao domínio} \\ +\|u_\alpha - u_\beta\| + C, & \text{se } u_\alpha \text{ não pertence ao domínio} \end{cases} \quad (2.18)$$

O valor do parâmetro C deve ser determinado a partir da calibração ([Figura 36](#)). [Manchuck e Deutsch \(2015\)](#) sugerem 2,5% de classificação errônea em nível aceitável para depósitos tabulares, uma outra forma de escolha é o método do cotovelo, o valor tomado deve ser o ponto de inflexão da curva. Para a categoria 1, foi tomado o valor de C igual a 130 que corresponde a um índice de classificação errônea de 2%. As distâncias modificadas pelo parâmetro C a partir da [Equação 2.18](#), devem ser interpoladas para todos os nós do grid, e a truncagem entre -C e C define a zona de incerteza.

Valores para a função distância são então simulados uniformemente entre -C e C. Caso o valor interpolado seja menor que o valor simulado para a função distância, o local é considerado pertencente ao domínio, caso o valor seja maior, o local é considerado externo ao domínio. Nos locais onde a interpolação apresenta o mesmo valor da simulação, são estabelecidos os contatos dos domínios como esquematizado na [Figura 37](#).

[Figura 37 – Classificação dos locais comparando valores estimados e simulados.](#)



Fonte: Modificado de [Wilde e Deutsch \(2011c\)](#)

Para que a simulação seja realizada de forma uniforme entre -C e +C, o desvio padrão  $y'(u)$ , deve ser simulado e transformado pela relação:

$$df'(u) = 2 * C * G^{-1}(y'(u)) - C \quad (2.19)$$

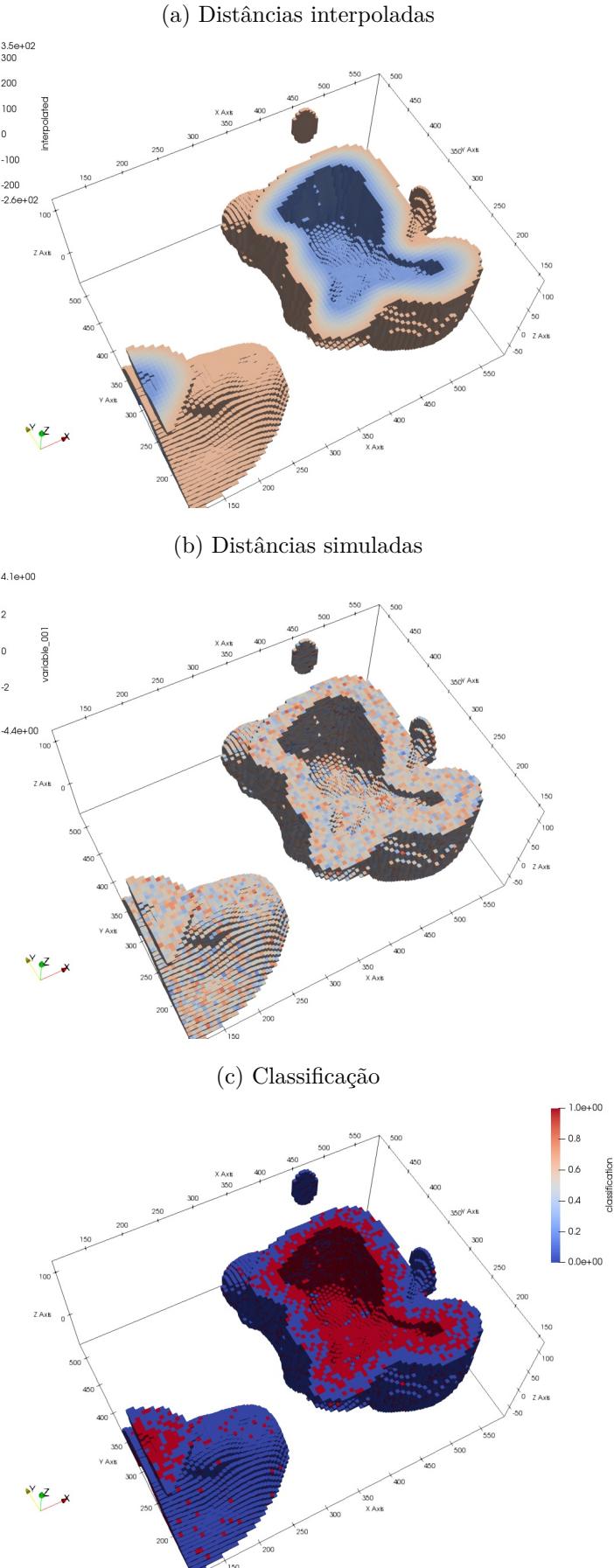
onde:  $df'(u)$  é o valor da função distância simulada,  $y'(u)$  o valor normal padrão da simulação não condicional, e  $G^{-1}$  representa a determinação do valor da distribuição

acumulada padrão normal correspondente a  $y'(u)$ . Para garantir que os valores pertençam a região estabelecida, os valores são multiplicados por  $2C$  e subtraídos de  $C$ .

O variograma utilizado na simulação não condicional pode ser o mesmo das distâncias assinaladas. O alcance do variograma determina a natureza do contato entre os domínios, menores alcances geram contatos mais rugosos enquanto maiores alcances, contatos suaves.

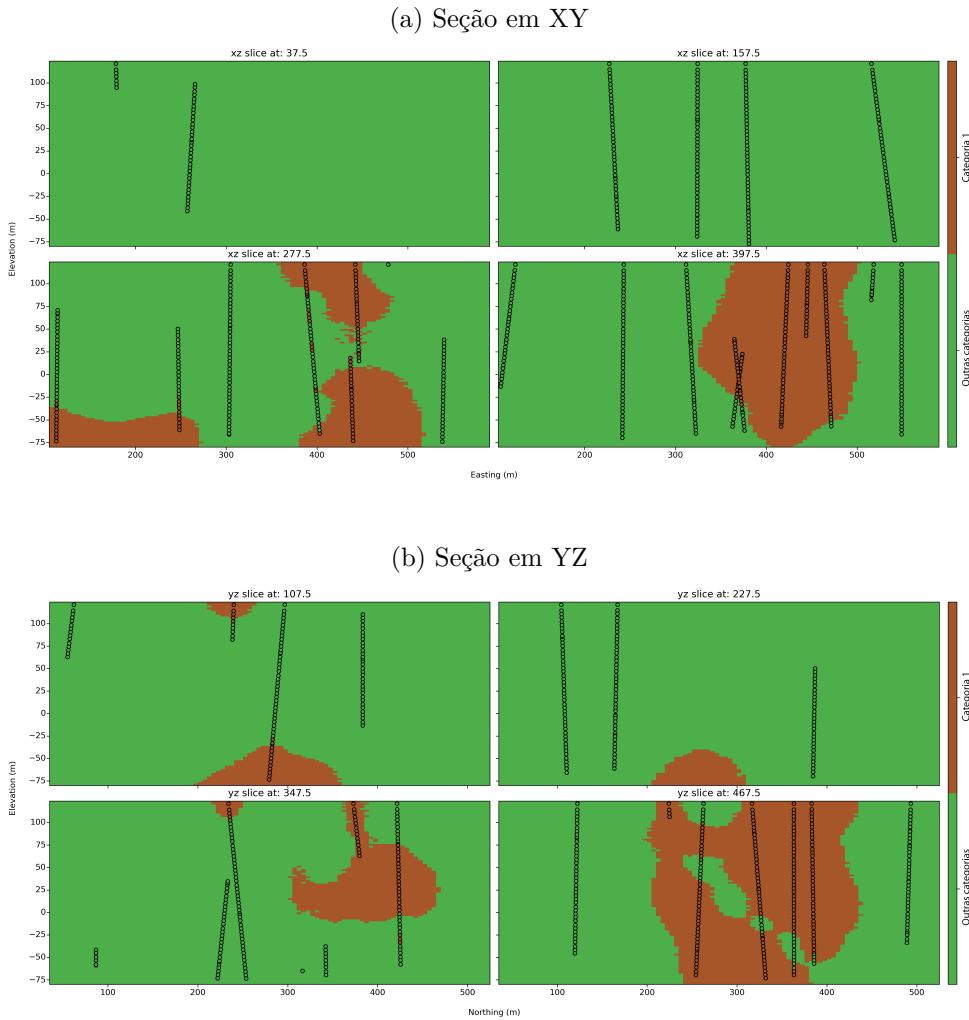
A [Figura 38](#) mostra as distâncias interpoladas e simuladas na zona de incerteza para a categoria 1, definida entre -120m e 120m e os blocos classificados em pertencentes à categoria ou não. As probabilidades de cada bloco pertencer à categoria 1 pode ser calculada tomando a média dos blocos classificados com 1 ou 0 em todas as realizações.

Figura 38 – Interpolação, simulação e classificação na zona de incerteza.



A Figura 39 mostra seções em XY e YZ de uma das realizações para a categoria 1.

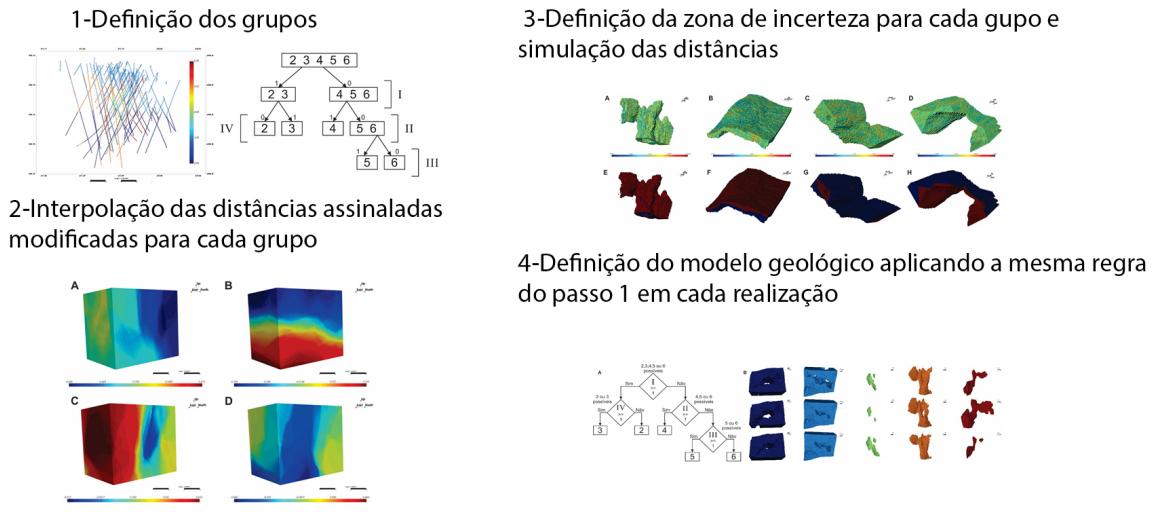
Figura 39 – Seções verticais de uma realização para a categoria 1.



O método é rápido, já que é baseado em simulações não condicionais, gera modelos realistas sem ruído e com fronteiras contínuas. Tanto a magnitude da incerteza quanto a natureza da fronteira podem ser controlados. Entretanto, o método funciona apenas para uma categoria por vez e mesmo que simplificações tenham sido desenvolvidas, a calibração do parâmetro de incerteza ainda pode ser laboriosa e subjetiva.

[Amarante, Rolo e Costa \(2018\)](#) propuseram uma adaptação do método para múltiplas categorias simultâneas, uma série de regras hierárquicas baseadas na cronologia dos eventos geológicos devem ser definidas. A Figura 40 mostra um esquema dos passos do método. Embora seja uma alternativa para ambientes multi categóricos, a definição das regras de hierarquização dependem de conhecimento geológico além das amostras, e muitas vezes indisponível.

Figura 40 – Esquema do método mostrando os passos necessários.



Fonte: Modificado de Amarante, Rolo e Costa (2018)

## 2.8.6 Sumário dos métodos de avaliação de incerteza

A Tabela 5 mostra um sumário dos métodos de incerteza que avalia: a simplicidade, o método deve, preferencialmente, ser simples e direto, não deve envolver um número excessivo de passos ou matemática complicada. A velocidade, o método deve ser de rápida execução. A capacidade de trabalhar em ambientes multi categóricos de forma direta, sem a necessidade de abordagens hierárquicas subjetivas. O realismo geológico, as formas geradas devem ser suaves e não apresentar ruído excessivo. O controle da incerteza, a incerteza deve ser controlada a partir de bandas de incerteza ou a partir dos parâmetros de simulação. Finalmente, a natureza do contato, ruidoso ou suave, deve ser controlada.

Método	Simplicidade	Velocidade	Multi categórico	Realismo geológico	Controle da incerteza	Controle do tipo de contato
Heurístico	simples	rápido	sim	não	sim	não
Boundsim	simples	rápido	não	sim	não	não
Simulação direta	complexo	demorado	sim	não	sim	não
MPS	simples	rápido	sim	sim	não	não
Boundary simulation	simples	rápido	não	sim	sim	sim

Tabela 5 – Sumário dos métodos de avaliação de incerteza de modelos geológicos.



# 3 Proposta de tese

## 3.1 Problemas e relevância da tese

A modelagem geológica implícita com funções distância assinaladas é uma metodologia bem estabelecida que permite ao gemodelador gerar de forma rápida modelos geológicos reproduutíveis, consistentes e checáveis a partir de uma função volume derivada a partir dos dados amostrais.

Porém, existem alguns pontos problemáticos e limitantes na metodologia que merecem atenção. A não estacionariedade da função distância assinalada torna a modelagem dos variogramas arbitrária e questionável.

Apesar dos métodos implícitos serem totalmente independentes de um *grid*, o objetivo final da modelagem geológica implícita é preencher o *grid* de estimativas com as litologias correspondentes. Para um número constante de amostras, o tempo requerido para interpolar a função volume para todos os nós do *grid* aumenta linearmente em relação ao número de nós. Os parâmetros do *grid* de estimativa podem ser pré-determinados pelos parâmetros de engenharia do projeto, porém, o *grid* de definição do modelo geológico não necessariamente deve ser o mesmo do *grid* de estimativa. A consideração mais importante na definição do *grid* é que a resolução seja suficiente para captar as estruturas geológicas de interesse (MARTIN; BOISVERT, 2017). É preciso encontrar um balanço entre número de nós e resolução necessária. A resolução do *grid* influencia diretamente a avaliação de incertezas.

A função volume, i.e. função distância assinalada, calculada em cada ponto amostral é interpolada. Diferentes métodos podem ser usados, desde métodos que não apresentam muita complexidade como os do inverso da distância, quanto métodos mais complexo como krigagem global ou funções de bases radiais com anisotropia global ou local. A escolha do interpolador muitas vezes é subjetiva e confusa e depende dos objetivos e da complexidade da geologia do depósito. Existem problemas relacionados à estacionariedade quando a krigagem é o método utilizado. A estratégia de busca é ponto crítico para nos interpoladores não globais.

Na presença de múltiplos domínios, principalmente em ambientes geológicos complexos, é necessário a aplicação de uma lógica de precedência de estruturas ao invés de simplesmente tomar a menor distância assinalada para a criação de modelos realistas. Mesmo em ambientes sem complexidade, categorias com poucas amostras disponíveis muitas vezes desaparecem dos modelos gerados tomando a menor distância estimada (se não for usada uma sequência hierárquica.).

Alguns dos métodos de avaliação de incerteza, encontrados na literatura, exigem a calibração de uma zona de incerteza entre diferentes domínios geológicos. Enquanto para alguns métodos, a definição da zona é subjetiva e não segue nenhuma regra matemática ou geológica, em outros a definição da zona de incerteza é extremamente laboriosa e complicada. Existem diferentes métodos para avaliação de incerteza em modelos geológicos implícitos. Em alguns deles, é possível avaliar a incerteza de todas as diferentes litologias do depósito mineral simultaneamente, em outros a incerteza associada à cada litologia deve ser avaliada separadamente. Além disso, a magnitude da incerteza e a natureza dos contatos não podem ser controladas na maior parte dos métodos encontrados na literatura.

Estruturas geológicas específicas, como lentes ou diques, podem desaparecer, ou não serem bem reproduzidas nos modelos implícitos, principalmente, quando sub-amostradas em relação às demais estruturas. Além disso, a metodologia, por si só, não produz bons resultados ao tentar reproduzir falhas e dobras.

Segundo o estatístico [Box \(1979\)](#): "*Todos os modelos estão errados, mas alguns são úteis.*" A modelagem implícita é conhecida por criar modelos de forma rápida e simples, porém, modelos ruins de forma rápida e simples. É necessário checar se os modelos implícitos honram a geologia do depósito e serão úteis para o processo de avaliação de recursos/reservas.

Essa tese encontra relevância ao buscar solução para os problemas abordados e propor melhorias para a técnica, que vem sendo aplicada com sucesso na indústria mineral há mais de uma década.

## 3.2 Meta

A meta proposta para a tese é investigar e desenvolver possíveis melhorias na modelagem geológica implícita com funções distância assinaladas através da solução dos diversos problemas pontuais apresentados na [seção 3.1](#). Para isso, os seguintes objetivos foram delineados:

## 3.3 Objetivos

### 3.3.1 Interpolador

O primeiro objetivo da tese é investigar a aplicabilidade de um interpolador baseado em aprendizado de máquina: [Samson e Deutsch \(2018b\)](#) desenvolveram um algoritmo, usando [TensorFlow](#)<sup>1</sup> para implementar a arquitetura da rede neural. O método é baseado

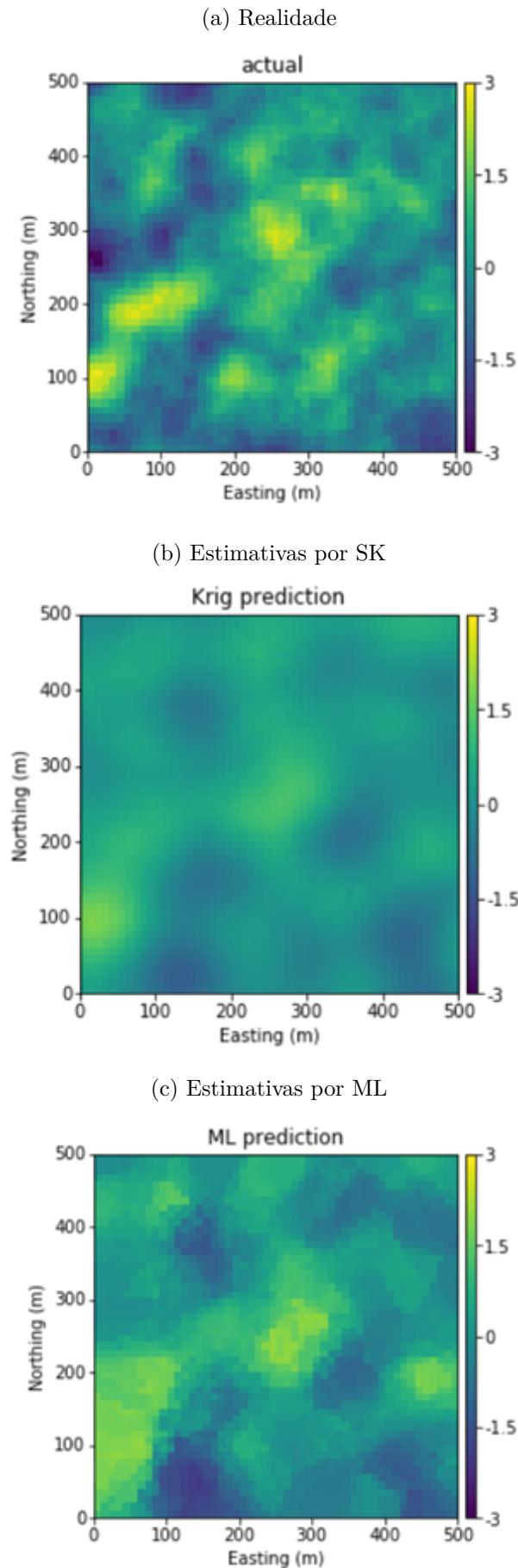
---

<sup>1</sup> TensorFlow é uma biblioteca de software de código aberto para computação numérica usando grafos computacionais. Foi originalmente desenvolvido pela Google Brain Team na organização de pesquisa Machine Intelligence do Google para aprendizado de máquina e pesquisa de redes neurais profundas

em aprendizado não supervisionado, *K-means* e supervisionado, *radial basis neural network - RBFN* em combinação com um algoritmo de otimização de parâmetro para produzir estimativas de teores.

Em seu trabalho, [Samson e Deutsch \(2018b\)](#) dividiram os dados disponíveis em treinamento (66%) e teste (33%). O algoritmo *K-means* é usado para determinar o número de unidades de bases radiais usadas na rede neural. A função de ativação alimentou o algoritmo com as coordenadas x, y e z e teores das 5 amostras mais próximas do local a ser estimado. A [Figura 41](#) mostra os resultados obtidos pela metodologia descrita e por krigagem simples a partir do mesmo banco de dados. Os resultados mostram uma boa reprodução das características visuais da imagem de referência, bem como do variograma e histograma dos dados.

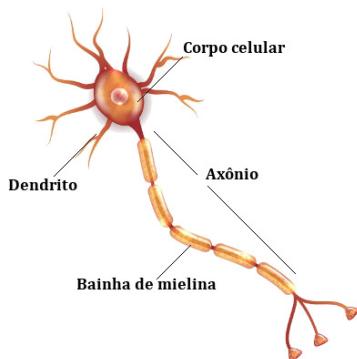
Figura 41 – Comparação de modelos gerados por diferentes métodos de estimativas.



Fonte: [Samson e Deutsch \(2018b\)](#)

Redes neurais são baseadas em neurônios ([Figura 42](#)). Os dendritos representam o *input* e os axônios representam o *output*. Redes neurais são tipicamente usadas em problemas de identificação ([SAMSON; DEUTSCH, 2018a](#)).

[Figura 42 – Figura esquemática de um neurônio.](#)



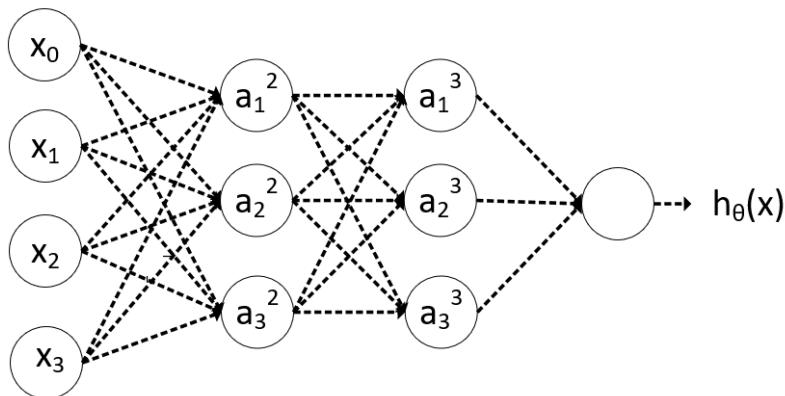
Fonte: <<https://brasilescola.uol.com.br/o-que-e/biologia/o-que-e-neuronio.htm>> acesso em: março de 2019

A [Figura 43](#) mostra um exemplo simples de rede neural. Os nós "x" na camada um da rede representam os *inputs*, os nós "a" na camada dois e três representam as camadas escondidas, e a camada quatro é o nó de *output* e representa a hipótese  $H_{\theta}(x)$ .

Redes neurais trabalham em um sistema binário de zeros e uns e por isso podem se basear nas propriedades da função sigmóide <sup>2</sup>, que tem seu valor em y entre zero e um para qualquer valor de x. Em uma rede neural, uma hipótese é gerada para cada nó e é determinado se seu valor será zero ou um, o valor binário é então passado para a próxima camada. Na rede neural da [Figura 43](#), uma hipótese para cada nó deve ser definida, a hipótese em um nó é o peso do *input* multiplicado pelo valor binário (0,1). O valor da hipótese é então imputado na função sigmóide que converte o valor para zero ou um ([SAMSON; DEUTSCH, 2018a](#)).

<sup>2</sup> A função sigmóide é uma função matemática de amplo uso em campos como a economia e a computação. O nome "sigmóide" vem da forma em S do seu gráfico. ([Estimation lemma, 2010](#))

Figura 43 – Rede neural.



Fonte: Samson e Deutsch (2018a)

Redes neurais de bases radiais apresentam uma arquitetura semelhante às redes neurais artificiais (Figura 43), que consiste em uma camada de *input*, camadas escondidas e uma camada de *output*. A principal diferença reside no número de camadas escondidas, que nas redes neurais de bases radiais se reduz a apenas uma. Além disso, redes neurais de bases radiais se baseiam em funções de bases radiais como função de ativação

A metodologia se mostrou competente em fazer estimativas a partir de amostras esparsas sem a necessidade do cálculo e modelagem do variograma. Em seu trabalho, Samson e Deutsch (2018b) não treinaram o parâmetro  $\beta$ , que controla o espalhamento da função de base radial e sugere que isso seja feito em trabalhos futuros para otimização do método. A aplicação de redes neurais de bases radiais na modelagem geológica implícita é promissora porque evita a tarefa de modelagem e cálculo do variograma. Além disso, a função radial gaussiana já é utilizada com sucesso na criação dos modelos implícitos.

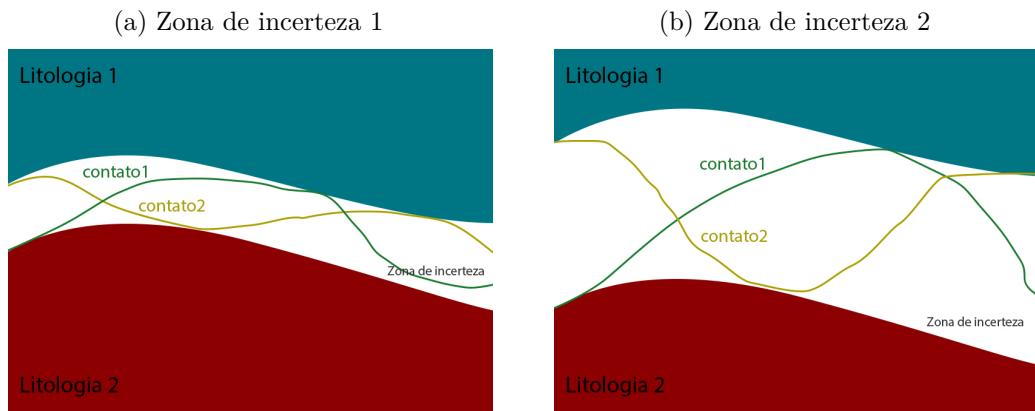
### 3.3.2 Zona de incerteza

Um outro objetivo é desenvolver uma metodologia objetiva e direta para definição da zona de incerteza: tanto o método baseado em múltiplas categorias simultâneas (coeficiente U da subseção 2.8.3) quanto o método da zona de incerteza calibrada (parâmetro C da subseção 2.8.5), para cada litologia, de forma independente podem gerar zonas de incertezas em interseção com amostras. É necessário definir uma zona de incerteza entre os contatos multi categóricos, que seja de fato "incerta", justa e livre de viés.

Gerar múltiplas realizações com a finalidade de avaliar a incerteza de modelos geológicos em todos os nós do *grid* é desperdício de tempo e poder computacional, já que no interior dos domínios não há incerteza.

A calibração correta do tamanho da zona de incerteza é importante porque no método de avaliação de incerteza proposto nessa tese, os diferentes contatos, para cada realização, variam dentro da zona de incerteza e a variação de volume de cada litologia depende diretamente da zona de incerteza como mostrado na [Figura 44](#).

Figura 44 – Zonas de incerteza e contatos.



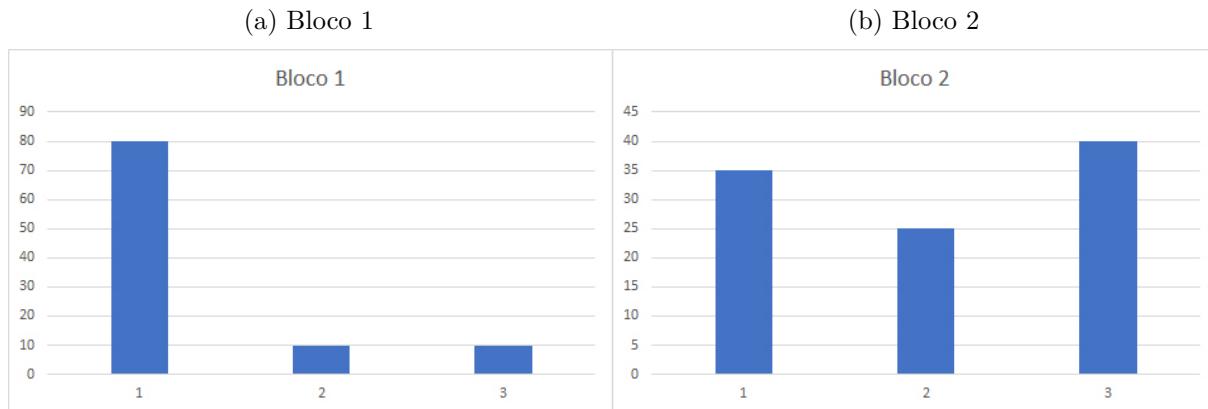
Em física, a entropia é uma medida do grau de desordem em um sistema. Um sistema ordenado, como um diamante, por exemplo, tem uma baixa entropia. Enquanto um sistema desordenado, como uma mistura de gases em alta temperatura, tem uma alta entropia. O conceito de entropia também surge na matemática. Em teoria da probabilidade, é uma forma de medir o grau médio de incerteza a respeito de fontes de informação.

Desse modo, a entropia é uma medida de incerteza. A entropia da informação ou entropia de Shannon ([SHANNON, 1948](#)) para um evento X com n resultados possíveis com probabilidades ( $p_1, \dots, p_n$ ) é dada por:

$$H(X) = H(p_1, \dots, p_n) = - \sum_{i=1}^n p_i \log_2 p_i \quad (3.1)$$

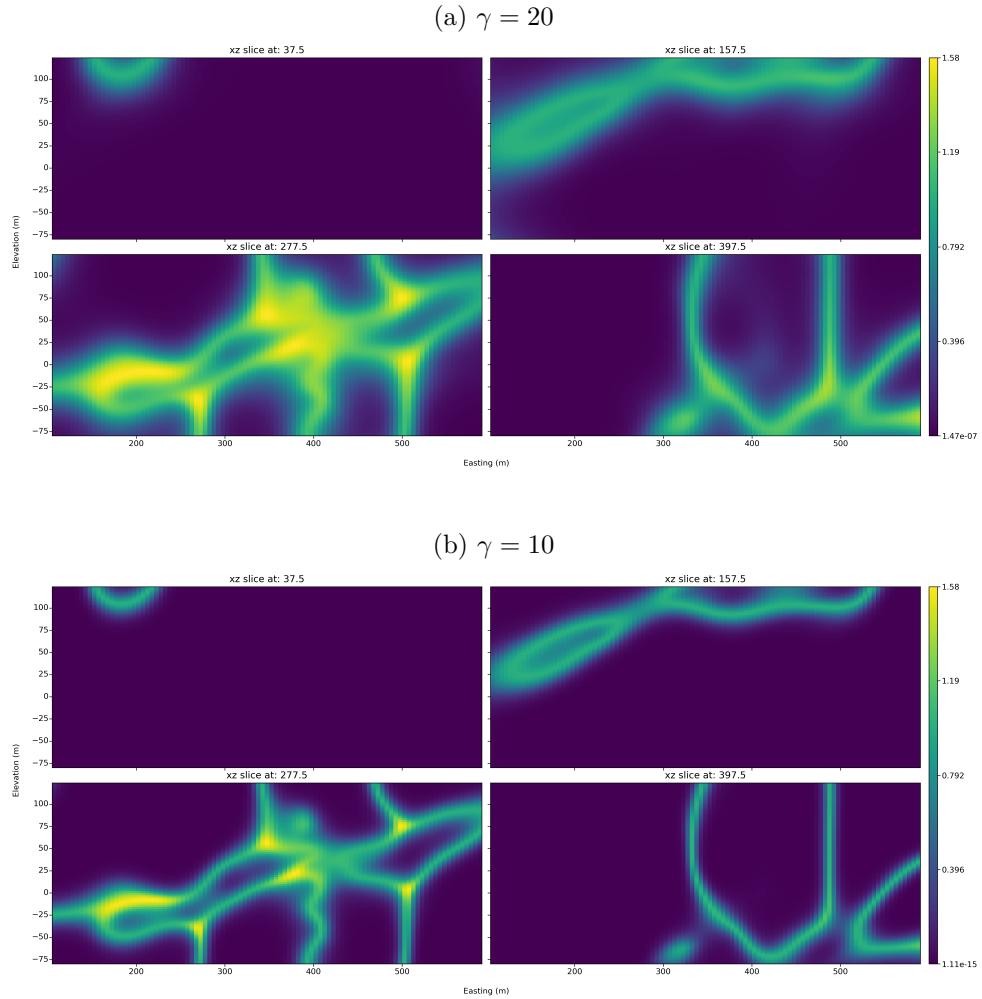
Calcular as probabilidades de cada litologia em um determinado local é trivial, como mostrado na [subseção 2.8.1](#), o que a torna a entropia da informação uma poderosa ferramenta tanto na definição da zona de incerteza como para de fato avaliar a incerteza do modelo geológico. Considere dois diferentes blocos onde as probabilidades de cada litologia foram calculadas e são mostradas na [Figura 45](#), o bloco 1 apresenta uma entropia menor que o bloco 2. Conhecimento e entropia são opostos: quando o conhecimento a respeito da mais provável litologia para um bloco é alto a entropia é baixa e vice-versa.

Figura 45 – Probabilidades de cada categoria em dois diferentes blocos.

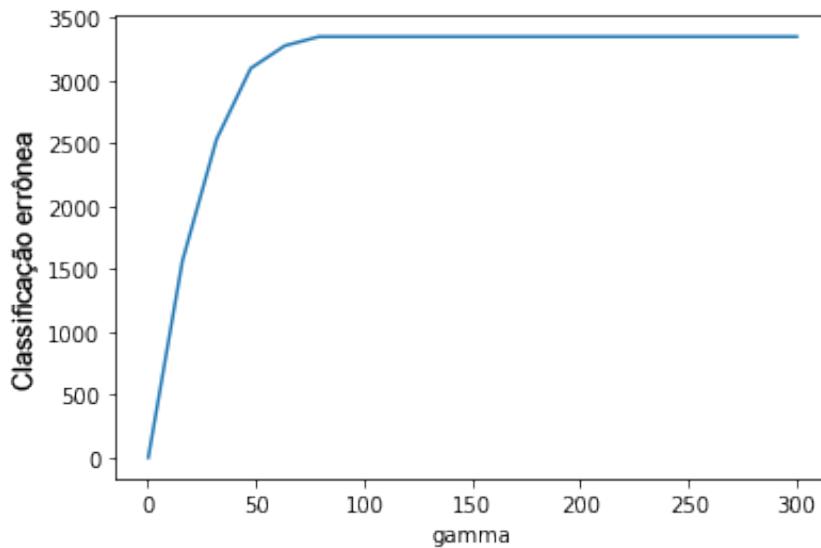


Algumas das propriedades notáveis e desejáveis da entropia são: a entropia é máxima para distribuições uniformes; é aditiva para eventos independentes; adicionar um novo evento com probabilidade zero não altera a entropia; é contínua e não negativa; a entropia para eventos como o mesmo resultado é zero e invariante à permutação.

A [Figura 46](#) mostra 4 seções verticais da entropia calculada para o banco de dados em estudo a partir de probabilidades calculadas com diferentes valores de  $\gamma$  ([subseção 2.8.1](#)). A entropia é um parâmetro semelhante ao coeficiente U ([subseção 2.8.3](#)); porém, a variação da entropia pode ser controlada pela magnitude das proporções calculadas pela [Equação 2.15](#). O parâmetro  $\gamma$  deve ser calibrado para gerar proporções que resultem em um campo de entropia plausível, a partir do qual será definida a zona de incerteza.

Figura 46 – Entropias calculadas para diferentes valores de  $\gamma$ .

A probabilidade de uma dada categoria em um ponto amostral deve ser um, se a amostra pertencer à essa mesma categoria, e zero caso contrário, somando um para todas as litologias e variando de suavemente, entre 0 e 1, no espaço. O parâmetro  $\gamma$  poderia ser calibrado de forma, que na validação, as duas imposições sejam atendidas (probabilidade igual a 1 se amostra pertence a categoria e zero caso contrário), uma série de valores permissíveis para  $\gamma$  poderia ser definido dessa forma. O  $\gamma$  escolhido deve ser o que contempla maior incerteza possível (maior valor). Dessa forma, o campo de entropia pode fornecer uma zona de incerteza de fato incerta, justa e livre de viés. A Figura 47 mostra a calibração do parâmetro para o banco de dados em estudo. Muitas vezes, a imposição é muito restritiva e gera zonas de incerteza muito pequenas ou em alguns casos ausentes. A partir de 0.4, um valor muito pequeno para o parâmetro, já existe classificação errônea, e a partir de 110 todas as amostras são classificadas de forma errada.

Figura 47 – Calibração do  $\gamma$ .

Conforme o valor de  $\gamma$  aumenta, o número de amostras classificadas erroneamente - diferente de 0 ou diferente de 1 - aumenta, da mesma forma para todas as litologias (por isso apenas um gráfico de calibração). É preciso desenvolver uma metodologia diferente para a calibração do parâmetro  $\gamma$ . Alguns autores recomendam a maior distância estimada entre todas as categorias.

Resta então definir um *cut off* no campo de entropia para sinalizar quais blocos variam e quais ficam congelados na geração dos modelos estocásticos. É preciso definir uma regra matemática, que pode ser baseada na entropia de formação do depósito mineral, depósitos de mais baixa entropia, como os de bauxita, teriam seu *cut off* definido em um valor maior do que depósitos de alta entropia, como um depósito de ouro primário.

### 3.3.3 Avaliação de incerteza

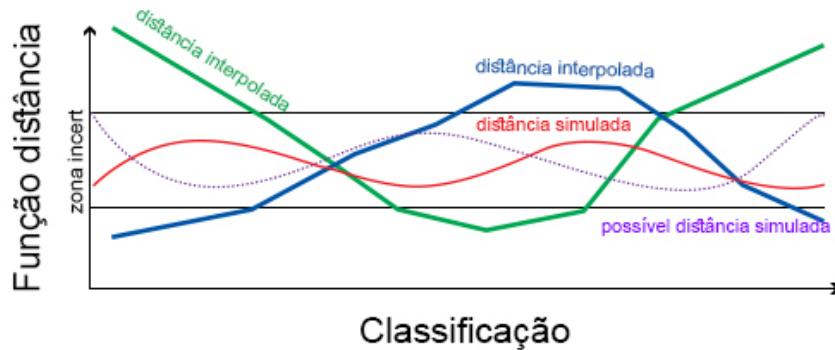
Essa tese ainda propõe investigar diferentes técnicas para a geração de modelos geológicos estocásticos multi categóricos baseados em múltiplas realizações. As técnicas que simulam diretamente as distâncias assinaladas geram modelos ruidosos, que não respeitam as zonas de incertezas definidas. Além disso, os resultados são muito sensíveis aos parâmetros e são necessários muitos passos nos *workflows*. Embora para geologias complexas seja mais indicado a modelagem e avaliação de incerteza estrutura à estrutura, em muitos casos, que apresentam geologia simples ou para definição de *grade shells* uma metodologia, multi categórica torna o processo de avaliação de incerteza mais simples e rápido.

Por esses motivos, o uso de uma variável contínua auxiliar, para derivar os modelos categóricos é uma alternativa melhor. Já que gera modelos menos ruidosos que apresentam contatos contínuos. A variável auxiliar é simulada de forma não condicional.

### 3.3.3.1 *Boundary simulation* multi categórico

Atualmente, o método de avaliação de incerteza que produz os melhores resultados é apresentado na [subseção 2.8.5](#). Esse método permite controlar o volume da zona de incerteza e a natureza dos contatos. Nesse método, o contato entre litologias varia dentro da zona de incerteza calibrada, sem a geração de ruído. Além disso, a natureza do contato pode ser controlada pelo variograma da simulação não condicional. Porém, o método trabalha apenas com uma litologia por vez e a simulação não condicional ainda depende de um variograma. Essa tese propõe uma adaptação desse método para múltiplas categorias simultâneas. As distâncias interpoladas para as diferentes litologias serão comparadas, apenas na nova zona de incerteza, com apenas um campo simulado. Dessa forma, as distâncias devem ser, de alguma forma, estandardizadas, ou com n campos simulados, um para cada distância, como mostra a [Figura 48](#).

Figura 48 – Comparaçāo multi categórica.



### 3.3.3.2 Simulação *P-field*

Uma alternativa à simulação não condicional das distâncias na zona de incerteza são os campos de probabilidade (*P-field*). A ideia central desse método de simulação é dissociar a tarefa de estimar distribuições de probabilidades locais para geração de múltiplas realizações equiprováveis ([FROIDÉVAUX, 1993](#)). Uma premissa é de que as distribuições locais são conhecidas, o que é uma premissa razoável, já que as distribuições locais podem ser calculadas pela [Equação 2.15](#). As simulações condicionais são obtidas tomando valores dessas CDFs locais. O valor de probabilidade usados para amostrar as distribuições locais constituem o campo de probabilidades (*P-field*) e são realizações

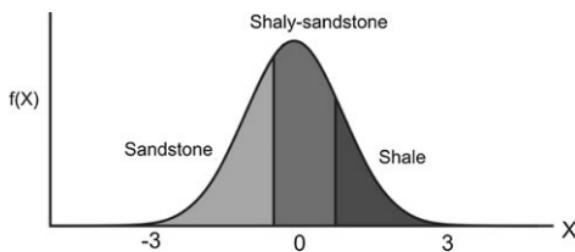
de uma função aleatória caracterizada por uma distribuição uniforme e um modelo de covariância.

N realizações do campo de probabilidades podem ser geradas na zona de incerteza, para cada bloco e em cada realização, a litologia correspondente é tomada da CDF local a partir da probabilidade simulada naquele bloco. O variograma utilizado para geração do campo de probabilidades controla a natureza dos contatos: suave ou rugoso. É preciso definir um modelo de covariância que se adeque aos contatos de todas as litologias de forma simultânea.

### 3.3.3.3 Simulação plurigaussiana truncada

Simulação gaussiana truncada (SGT) foi proposta por [Matheron et al. \(1987\)](#) e é uma alternativa para simular variáveis categóricas que usem uma única variável aleatória gaussiana. Realizações da variável gaussiana são tratadas como uma variável latente subjacente ([SILVA; DEUTSCH, 2016](#)) e então são truncadas para gerar realizações categóricas. Todavia, o uso de apenas uma variável gaussiana faz com que a técnica seja limitada apenas a simulação de variáveis categóricas ordenadas, como ambientes sedimentares estratigráficos, por exemplo. A [Figura 49](#) mostra a variável gaussiana truncada para representar 3 diferentes categorias. A aplicação de uma regra local de truncagem na SGT gera resultados similares à simulação *P-field*.

Figura 49 – Esquema da SGT. Note que as litologias sandstone e shale não aparecerão juntas nas realizações.

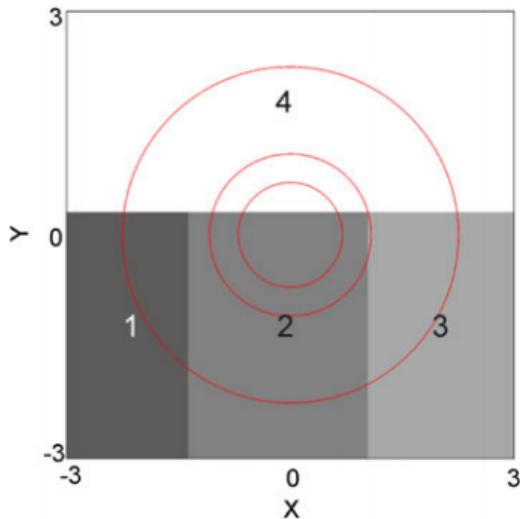


Fonte: [Deutsch e Deutsch \(2014\)](#)

A simulação plurigaussiana truncada (SPT), proposta por [Galli et al. \(1994\)](#), é uma extensão da SGT para permitir o uso de múltiplas variáveis aleatórias gaussianas (variáveis latentes). Assim, permite a simulação de estruturas mais complexas. A SPT pode reproduzir proporções, continuidade espacial e probabilidades de transição. O poder da SPT vem do uso de múltiplas variáveis gaussianas para representar as estruturas complexas observadas nos dados. Todavia, seu uso tem sido limitado a duas funções gaussianas,  $Y_1$  e  $Y_2$ , correlacionadas ou não, pelo fato de que a definição e interpretação

de regras de truncagem é complicada em mais de duas dimensões (SILVA; DEUTSCH, 2016). A Figura 49 mostra duas variáveis gaussianas Y e X truncadas segundo uma regra definida para definir quatro diferentes litologias.

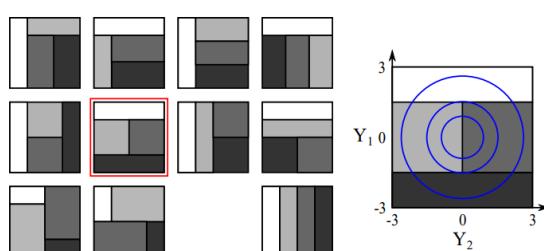
Figura 50 – Esquema da SPT. Note que as litologias 1 e 3 não aparecerão juntas nas realizações.



Fonte: [Deutsch e Deutsch \(2014\)](#)

A regra de truncagem ou regra de litologia é uma parte importante da metodologia SPT já que ela controla os contatos entre as categorias, sua transição e proporções. A definição das regras de truncagem é geralmente baseada nas probabilidades de transição observadas nos dados e nos modelos conceituais (MARIETHOZ et al., 2009). A partir dos dados, é possível calcular a probabilidade de transição de uma categoria para outra e expressá-las em forma de matriz (DEUTSCH; DEUTSCH, 2013).

Figura 51 – Diferentes templates, para o caso de duas variáveis latentes e quatro categorias.



Fonte: Modificado de [Armstrong et al. \(2011\)](#)

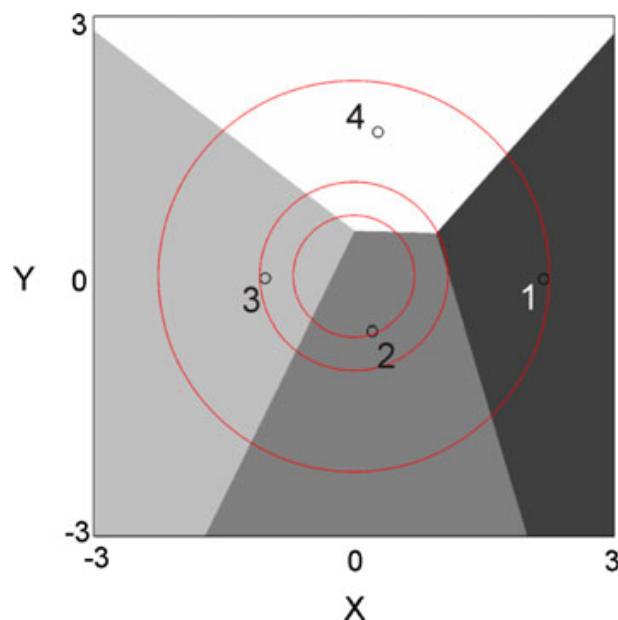
Diversas técnicas para definir a regra de truncagem foram propostas: Armstrong et al. (2011) propõem o particionamento de espaço gaussiano multivariado em paralelepípedos (ou retângulos no caso bivariado) através de limiares (Figura 51). Essa técnica é adequada ao caso bivariado e com poucas categorias.

Allard et al. (2012) introduziram o *assignment diagram*, que automaticamente constrói a regra de truncagem para o caso bivariado usando regressão baseada em kernel em variáveis auxiliares.

Sadeghi e Boisvert (2012) usaram *simulated annealing* para otimizar a regra de truncagem. A função objetivo é a minimização da classificação errônea entre as probabilidades de transição calculadas de realizações e das probabilidades de transição calculadas a partir dos dados.

Deutsch e Deutsch (2014) usaram *multidimensional scaling* (MDS) para definir regras de truncagem complexas com o foco em reproduzir as probabilidades de transição. Essa metodologia pode ser aplicada para qualquer número de variáveis gaussianas. A metodologia é uma ótima alternativa para ambientes geológicos complexos que apresentam contatos múltiplos entre as diferentes litologias, como mostrado na Figura 52. Além disso, a geração da regra de truncagem é automática.

Figura 52 – Esquema do uso de MDS para definição da regra de truncagem.



Fonte: Deutsch e Deutsch (2014)

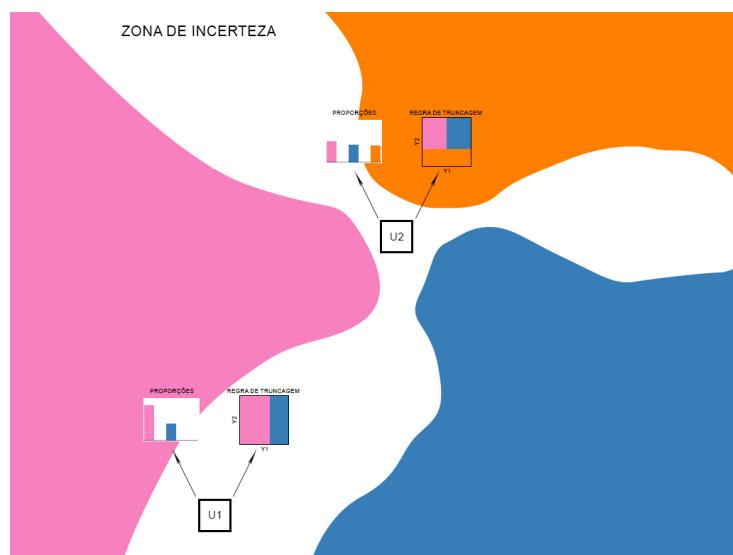
Astrakova, Oliver e Lantuéjoul (2015) propuseram uma metodologia semelhante a de Deutsch e Deutsch (2014) usando entropia bayesiana máxima em conjunto com *simulated annealing* para otimizar a regra de truncagem bigaussiana.

Madani e Emery (2015) e Silva e Deutsch (2016) propuseram uma abordagem hierárquica para definir a regra de truncagem em espaços multi dimensionais.

Em muitos casos, as proporções não são estacionárias. As curvas de proporção verticais, propostas por Matheron et al. (1987), mostram esse comportamento a partir do cálculo das proporções de cada litologia baseados na profundidade. Nesses casos, a regra de truncagem não deve ser global e deve variar localmente (SADEGHI; BOISVERT, 2012).

Essa tese propõe a definição dinâmica da regra de truncagem bivariada para cada bloco, dentro da zona de incerteza, a partir das probabilidades calibradas e calculadas pela Equação 2.15. A Figura 53 mostra um desenho esquemático do método proposto.

Figura 53 – Diferentes regras de truncagem definidas para cada bloco a partir das probabilidades calibradas e calculadas.



O próximo passo é simular as variáveis latentes gaussianas, cada uma com uma função de auto covariância. Segundo Sadeghi e Boisvert (2016), os variogramas devem ser modelados de forma que variável categórica original tenha a estrutura espacial correta após a truncagem. Diversas metodologias foram desenvolvidas para a definição dos variogramas. Na prática, se existem duas importantes litologias, minério mais rico por exemplo, o variograma dessas litologias devem ser atribuídos às variáveis latentes  $Y_1$  e  $Y_2$ . O variograma dos indicadores deve ser invertido numericamente para corresponder ao seu variograma gaussiano (JOURNEL; KYRIAKIDIS, 2004) (o variograma gaussiano tem alcance e anisotropias similares; porém, a forma é mais suave, é mais parabólica ou gaussiana em pequenas distâncias). Na prática, o variograma dos indicadores é alterado para o modelo gaussiano (PYRCZ; DEUTSCH, 2014).

Na metodologia original (ARMSTRONG et al., 2011) a simulação das variáveis

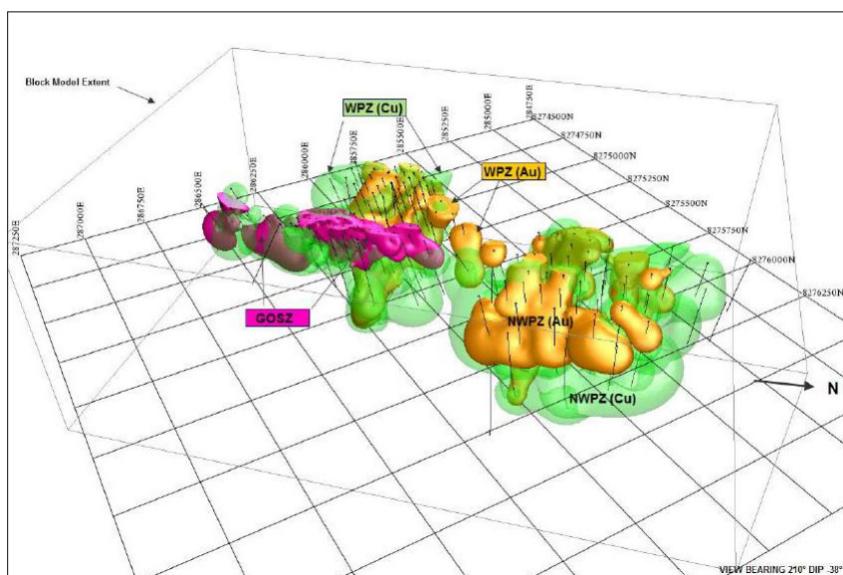
latentes é condicionada a valores gaussianos gerados a partir dos dados por um amostrador de Gibbs. Como não existem amostras na nova zona de incerteza proposta, a simulação deve ser não condicional.

O último passo é aplicar a regra de truncagem em cada bloco e para cada realização das variáveis latentes simuladas para classificar os blocos da zona de incerteza.

### 3.3.4 Validação

Finalmente, é preciso desenvolver metodologias para validação e checagem dos modelos. A modelagem implícita tende a gerar estruturas em formatos circulares ("blobs") ao redor das amostras, como o modelo apresentado na [Figura 54](#), ou em forma de salsicha em torno de furos de sondagem. Um índice de quanto as estruturas geradas se parecem com essas estruturas indesejadas deve ser desenvolvido. Também, a checagem de locais classificados que são co-locados com amostras (e deveriam ter sido classificados com a mesma litologia da amostra) e checagem "visual", que pode ser realizada por um algoritmo de aprendizado de máquinas.

Figura 54 – Modelo implícito ruim, com a presença de estruturas e formas indesejadas ("blobs").



Fonte: <<https://www.linkedin.com/pulse/implicit-modelling-disasters-making-part-1-ron-reid/>>, acesso em junho de 2019

### 3.3.5 Sumário

Em suma, essa tese propõe desenvolver e investigar:

- Um interpolador baseado em redes neurais independente de variograma para a variável distância assinalada;
- Uma metodologia para calibração e definição de uma zona de incerteza multi categórica estatisticamente justa, livre de viés e que seja de fato incerta;
  - Calibração do parâmetro  $\gamma$ ;
  - Truncagem no campo de entropia.
- Uma metodologia para avaliação da incerteza que gere múltiplas realizações de modelos geológicos multi categóricos dentro da zona de incerteza, sem ruído e com a possibilidade de controlar a natureza dos contatos.
  - Definição de uma regra de truncagem local para cada bloco.
  - Definição de um modelo de covariância para cada variável latente, que seja representativo.
- Uma metodologia para validação de modelos geológicos implícitos.
  - Identificação automática de estruturas indesejadas.

A aplicabilidade das novas metodologias propostas será avaliada a partir de estudos de casos em bancos de dados reais. As características e limitações das técnicas serão exaustivamente discutidas.

## 3.4 Cronograma e atividades

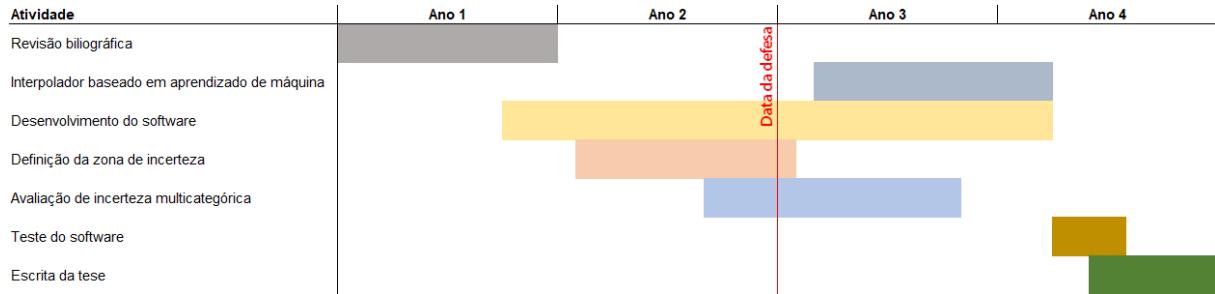
As atividades pretendidas para essa tese incluem: uma revisão bibliográfica da modelagem geológica implícita e todos os métodos de avaliação de incerteza disponíveis na literatura ([Capítulo 2](#)), desenvolvimento de um conjunto de programas de modelagem geológica implícita em python que inclua ferramentas: para o cálculo das distâncias, interpolação via redes neurais de bases radiais, calibração das probabilidades para cada litologia, definição da zona de incerteza, avaliação de incerteza via geração de múltiplas realizações equiprováveis do modelo geológico multi categórico e ferramentas de validação do modelo geológico. Os programas se encontram em desenvolvimentos e podem ser acessados no repositório: [<https://github.com/robertorolo/enhanced\\_dfmod>](https://github.com/robertorolo/enhanced_dfmod).

- Progresso de desenvolvimento de software:
  - ✓ Cálculo das distâncias assinaladas
  - ✓ Validação cruzada

- Interpolação tradicional por RBF
- Interpolação baseada em redes neurais
- Calibração do parâmetro  $\gamma$  a partir das amostras
- Calibração do parâmetro  $\gamma$  (nova abordagem)
- Cálculo das probabilidades
- Cálculo da entropia
- Definição do *cutt-off* no campo de entropia
- Simulação da variável contínua auxiliar
- Definição da regra de truncagem local
- Algoritmo de validação visual

A [Figura 55](#) mostra o cronograma de atividades para o período da tese.

Figura 55 – Cronograma de atividades.



# Referências

- AHRENS, J.; GEVECI, B.; LAW, C. Paraview: An end-user tool for large data visualization. *The visualization handbook*, Elsevier München, v. 717, 2005. Citado na página 30.
- AILLERES, L. et al. *Uncertainty in 3D models and implicit modeling of poly-deformaed terranes*. 2015. Presentation. Disponível em: <[http://www.tectonique.net/tectask/3DIGI/presentations/T1.2\\_ailleres\\_Goodbye2DEarthPresentation.pdf](http://www.tectonique.net/tectask/3DIGI/presentations/T1.2_ailleres_Goodbye2DEarthPresentation.pdf)>. Acesso em: 23-07-2018. Citado na página 12.
- ALABERT, F.; MODOT, V. et al. Stochastic models of reservoir heterogeneity: Impact on connectivity and average permeabilities. In: SOCIETY OF PETROLEUM ENGINEERS. *SPE Annual Technical Conference and Exhibition*. [S.l.], 1992. Citado na página 11.
- ALLARD, D. et al. Non-parametric diagrams for plurigaussian simulations of lithologies. In: *9th international geostatistical congress, Oslo, Norway*. [S.l.: s.n.], 2012. v. 1115. Citado na página 70.
- AMARANTE, F.; ROLO, R.; COSTA, J. Incerteza associada ao modelamento implícito e hierárquico. In: *CBMINA 2019*. [S.l.]: CBMINA, 2018. p. xx–xx. Citado 2 vezes nas páginas 55 e 56.
- ARMSTRONG, M. et al. *Plurigaussian simulations in geosciences*. [S.l.]: Springer Science & Business Media, 2011. Citado 2 vezes nas páginas 69 e 71.
- ASTRAKOVA, A.; OLIVER, D. S.; LANTUÉJOUL, C. Truncation map estimation based on bivariate probabilities and validation for the truncated plurigaussian model. *arXiv preprint arXiv:1508.01090*, 2015. Citado na página 70.
- BEATSON, R. K.; CHERRIE, J. B.; MOUAT, C. T. Fast fitting of radial basis functions: Methods based on preconditioned gmres iteration. *Advances in Computational Mathematics*, Springer, v. 11, n. 2-3, p. 253–270, 1999. Citado na página 28.
- BLOOMENTHAL, J.; BAJAJ, C. *Introduction to implicit surfaces*. [S.l.]: Morgan Kaufmann, 1997. Citado na página 11.
- BOISVERT, J. Automatic geological modeling with iterative refinement. *CCG Annual Report*, 2013. Citado na página 24.
- BOX, G. E. Robustness in the strategy of scientific model building. In: *Robustness in statistics*. [S.l.]: Elsevier, 1979. p. 201–236. Citado na página 58.
- BRATVOLD, R. B. et al. Storm: Integrated 3d stochastic reservoir modeling tool for geologists and reservoir engineers. *SPE Computer Applications*, Society of Petroleum Engineers, v. 7, n. 03, p. 58–67, 1995. Citado na página 11.
- CÁCERES, A. et al. Stochastic geological modeling using implicit boundary simulation. In: *Proceedings of the 2nd International Seminar on Geology for the Mining Industry GEOMIN*. [S.l.: s.n.], 2011. p. 21. Citado 2 vezes nas páginas 10 e 46.

- (CCG), C. for C. G. *PyGeostat: Python Geostatistical Package*. 2015–2019. <<http://www.ccgalberta.com/pygeostat/welcome.html>>. Citado 2 vezes nas páginas 19 e 20.
- CGAL. *Poisson Surface Reconstruction*. 2015. Manual do usuário. Disponível em: <[https://doc.cgal.org/latest/Poisson\\_surface\\_reconstruction\\_3/index.html](https://doc.cgal.org/latest/Poisson_surface_reconstruction_3/index.html)>. Acesso em: 25-07-2018. Citado na página 12.
- COWAN, E. et al. Rapid geological modelling. *Applied Structural Geology for Mineral Exploration and Mining, International Symposium*, 2002. Citado 2 vezes nas páginas 11 e 24.
- COWAN, E. et al. Practical implicit geological modelling. In: *Fifth International Mining Geology Conference*. [S.l.: s.n.], 2003. p. 17–19. Citado 3 vezes nas páginas 8, 9 e 11.
- DEUTSCH, C. A statistical resampling program for correlated data. *CCG Annual Report*, 2004. Citado na página 44.
- DEUTSCH, J.; DEUTSCH, C. Advances in truncated plurigaussian simulation for reproduction of transotion probabilities. *CCG Annual Report*, 2013. Citado na página 69.
- DEUTSCH, J. L.; DEUTSCH, C. V. A multidimensional scaling approach to enforce reproduction of transition probabilities in truncated plurigaussian simulation. *Stochastic environmental research and risk assessment*, Springer, v. 28, n. 3, p. 707–716, 2014. Citado 3 vezes nas páginas 68, 69 e 70.
- Estimation lemma. *Estimation lemma — Wikipedia, The Free Encyclopedia*. 2010. [Online; acesso 22-março-2019]. Disponível em: <[https://pt.wikipedia.org/wiki/Fun%C3%A7%C3%A3o\\_sigm%C3%B3ide](https://pt.wikipedia.org/wiki/Fun%C3%A7%C3%A3o_sigm%C3%B3ide)>. Citado na página 61.
- FASSHAUER, G. E. *Meshfree approximation methods with MATLAB*. [S.l.]: World Scientific, 2007. v. 6. Citado na página 28.
- FROIDEVAUX, R. Probability field simulation. In: *Geostatistics tróia'92*. [S.l.]: Springer, 1993. p. 73–83. Citado na página 67.
- GALLI, A. et al. The pros and cons of the truncated gaussian method. In: *Geostatistical simulations*. [S.l.]: Springer, 1994. p. 217–233. Citado 2 vezes nas páginas 11 e 68.
- GOOVAERTS, P. *Geostatistics for natural resources evaluation*. [S.l.]: Oxford University Press on Demand, 1997. Citado na página 11.
- HOLDEN, L. et al. Modeling of fluvial reservoirs with object models. *Mathematical Geology*, Springer, v. 30, n. 5, p. 473–496, 1998. Citado na página 11.
- HOSSEINI, A.; DEUTSCH, C. V. A distance function based algorithm to quantify uncertainty in areal limits. *CCG Annual Report*, 2007. Citado na página 24.
- JOURNEL, A. The indicator approach to estimation of spatial distributions. In: *Proceedings of the 17th APCOM International Symposium, New York*. [S.l.: s.n.], 1982. p. 793–806. Citado na página 10.
- JOURNEL, A. G.; HUIJBREGTS, C. J. *Mining geostatistics*. [S.l.]: Academic press, 1978. Citado na página 7.

- JOURNEL, A. G.; KYRIAKIDIS, P. C. *Evaluation of mineral reserves: a simulation approach*. [S.l.]: Oxford University Press, 2004. Citado na página 71.
- KAPEROV, T.; DEUTSCH, C. A review of boundary modeling and one way of selecting the c parameter to model boundaries with uncertainty. *CCG Annual Report*, 2016. Citado na página 51.
- KLOECKNER, J. et al. Modelagem geológica implícita automática via tabela de covariância com funções distância assinaladas. In: *Anais do IX Congresso Brasileiro de Mineração Belo Horizonte*. [S.l.: s.n.], 2018. p. xx–xx. Citado na página 21.
- LAJAUNIE, C.; COURRIOUX, G.; MANUEL, L. Foliation fields and 3d cartography in geology: principles of a method based on potential interpolation. *Mathematical Geology*, Springer, v. 29, n. 4, p. 571–584, 1997. Citado na página 13.
- LILLAHI, M.; BOISVERT, J. B. Inference of locally varying anisotropy fields from diverse data sources. *Computers & Geosciences*, Elsevier, v. 82, p. 170–182, 2015. Citado na página 36.
- LORENSEN, W. E.; CLINE, H. E. Marching cubes: A high resolution 3d surface construction algorithm. In: ACM. *ACM siggraph computer graphics*. [S.l.], 1987. v. 21, n. 4, p. 163–169. Citado na página 30.
- MADANI, N.; EMERY, X. Simulation of geo-domains accounting for chronology and contact relationships: application to the río blanco copper deposit. *Stochastic environmental research and risk assessment*, Springer, v. 29, n. 8, p. 2173–2191, 2015. Citado na página 70.
- MALLET J., L. *Geomodeling*. 1. ed. [S.l.]: Oxford, 2002. (Applied Geostatistics Series). ISBN 0195144600, 9780195144604. Citado na página 40.
- MALLET, J.-L. Space-time mathematical framework for sedimentary geology. *Mathematical geology*, Springer, v. 36, n. 1, p. 1–32, 2004. Citado na página 13.
- MANCHUCK, J.; DEUTSCH, C. Mls for boundary interpretations with uncertainty. *CCG Annual Report*, 2018. Citado 3 vezes nas páginas 24, 40 e 41.
- MANCHUCK, J. A.; DEUTSCH, C. V. Geometric modeling of irregular tabular deposits. *CCG Annual Report*, 2015. Citado 2 vezes nas páginas 18 e 52.
- MARIETHOZ, G. et al. Truncated plurigaussian simulations to characterize aquifer heterogeneity. *Groundwater*, Wiley Online Library, v. 47, n. 1, p. 13–24, 2009. Citado na página 69.
- MARTIN, R.; BOISVERT, J. Review of radial basis functions and domain decomposition for implicit geological modelling. *CCG Annual Report*, 2015. Citado 4 vezes nas páginas 27, 28, 29 e 38.
- MARTIN, R.; BOISVERT, J. Implicit modeling. *CCG Guidebook Series*, v. 22, 2017. Citado 13 vezes nas páginas 12, 13, 22, 23, 25, 28, 29, 35, 36, 37, 38, 39 e 57.
- MATHERON, G. *Traité de géostatistique appliquée*, tome i: Mémoires du bureau de recherches géologiques et minières. *Editions Technip, Paris*, v. 14, 1962. Citado na página 8.

- MATHERON, G. et al. Conditional simulation of the geometry of fluvio-deltaic reservoirs. In: SOCIETY OF PETROLEUM ENGINEERS. *SPE Annual Technical Conference and Exhibition*. [S.l.], 1987. Citado 3 vezes nas páginas 11, 68 e 71.
- MCCULLAGH, P.; NELDER, J. A. *Generalized linear models*. [S.l.]: Chapman and Hall, 1989. Citado na página 42.
- MCLENNAN, J. A. *The decision of stationarity*. Tese (Doutorado) — University of Alberta, 2007. Citado 2 vezes nas páginas 7 e 8.
- MCLENNAN, J. A.; DEUTSCH, C. V. Implicit boundary modeling (boundsim). *CCG Annual Report*, 2006. Citado 3 vezes nas páginas 8, 9 e 44.
- MUNROE, M.; DEUTSCH, C. Full calibration of c and in the framework of vein-type deposit tonnage uncertainty. *CCG Annual Report*, 2008. Citado na página 50.
- OSHER, S.; FEDKIW, R. *Level set methods and dynamic implicit surfaces*. [S.l.]: Springer-Verlag New York, Inc., 2003. Citado 2 vezes nas páginas 13 e 16.
- OXFORD, U. *Oxford online dictionary*. 2018. Definição. Disponível em: <<http://www.oed.com>>. Acesso em: 15-06-2018. Citado na página 8.
- PYRCZ, M. J.; CATUNEANU, O.; DEUTSCH, C. V. Stochastic surface-based modeling of turbidite lobes. *AAPG bulletin*, American Association of Petroleum Geologists, v. 89, n. 2, p. 177–191, 2005. Citado na página 11.
- PYRCZ, M. J.; DEUTSCH, C. V. *Geostatistical reservoir modeling*. [S.l.]: Oxford university press, 2014. Citado 2 vezes nas páginas 11 e 71.
- RADTKE, R. *Modelagem geológica implícita através de simulações de funções distância assinaladas*. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Sul, 2017. Citado 2 vezes nas páginas 46 e 48.
- ROLO, R. *Modelagem geológica implícita com funções distância assinaladas*. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Sul, 2017. Citado 2 vezes nas páginas 24 e 43.
- ROSSI, M. E.; DEUTSCH, C. V. *Mineral resource estimation*. [S.l.]: Springer Science & Business Media, 2013. Citado 2 vezes nas páginas 7 e 8.
- SADEGHI, S.; BOISVERT, J. Optimizing thresholds in truncated pluri-gaussian simulation. *CCG Annual Report*, 2012. Citado 2 vezes nas páginas 70 e 71.
- SADEGHI, S.; BOISVERT, J. Review of truncated plurigaussian simulation (case study). *CCG Annual Report*, 2016. Citado na página 71.
- SAMSON, M.; DEUTSCH, C. A brief introduction into machine learning. *CCG Annual Report*, 2018. Citado 2 vezes nas páginas 61 e 62.
- SAMSON, M.; DEUTSCH, C. Estimation with machine learning. *CCG Annual Report*, 2018. Citado 4 vezes nas páginas 58, 59, 60 e 62.
- SAVCHENKO, V. V. et al. Function representation of solids reconstructed from scattered surface points and contours. In: WILEY ONLINE LIBRARY. *Computer Graphics Forum*. [S.l.], 1995. v. 14, n. 4, p. 181–188. Citado na página 11.

- SHANNON, C. E. A mathematical theory of communication. *Bell system technical journal*, Wiley Online Library, v. 27, n. 3, p. 379–423, 1948. Citado na página 63.
- SILVA, D. *Signed distance function modeling with multiple categories*. 2015. Artigo publicado no site geostatisticslessons.com. Disponível em: <<http://www.geostatisticslessons.com/lessons/signeddistancefunctions>>. Acesso em: 28-12-2016. Citado na página 34.
- SILVA, D.; DEUTSCH, C. Modeling multiple rock types with distance functions: methodology and software. In: *CCG Annual Report*. [S.l.: s.n.], 2012. Citado 2 vezes nas páginas 32 e 42.
- SILVA, D.; DEUTSCH, C. Correcting distance function models to correct proportions. In: *CCG Annual Report*. [S.l.: s.n.], 2013. Citado na página 42.
- SILVA, D.; DEUTSCH, C. Speeding up the signed distance implicit geologic modeling method for multiple rock types. *CCG Annual Report*, 2015. Citado na página 24.
- SILVA, D.; DEUTSCH, C. Hierarchical approach to truncated plurigaussian simulation. *CCG Annual Report*, 2016. Citado 3 vezes nas páginas 68, 69 e 70.
- SILVA, D. M. A. *Enhanced geologic modeling with data-driven training images for improved resources and recoverable reserves*. Tese (Doutorado) — University of Alberta, 2015. Citado 4 vezes nas páginas 9, 24, 33 e 49.
- STREBELLE, S. Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 2002. Citado na página 11.
- STRÉBELLE, S.; JOURNEL, A. Sequential simulation drawing structures from training images. Stanford University California, 2000. Citado na página 49.
- WENDLAND, H. *Scattered data approximation*. [S.l.]: Cambridge university press, 2004. v. 17. Citado na página 28.
- WILDE, B.; DEUTSCH, C. A new way to calibrate distance function uncertainty. In: *CCG Annual Report*. [S.l.: s.n.], 2011. Citado na página 29.
- WILDE, B.; DEUTSCH, C. A new way to calibrate to calibrate distance functions uncertainty. *CCG Annual Report*, 2011. Citado na página 51.
- WILDE, B.; DEUTSCH, C. Simulating boundary realizations. *CCG Annual Report*, 2011. Citado 2 vezes nas páginas 50 e 52.
- WILDE, B. J.; DEUTSCH, C. V. Kriging and simulation in presence of stationary domains: developments in boundary modeling. In: *Geostatistics Oslo 2012*. [S.l.]: Springer, 2012. p. 289–300. Citado na página 51.