

MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
Escola de Engenharia  
Programa de Pós-Graduação em Engenharia de Minas,  
Metalúrgica e de Materiais (PPGE3M)

**GEOESTATÍSTICA DE MÚLTIPLOS PONTOS APLICADA À  
SIMULAÇÃO DE MODELOS GEOLÓGICOS EM GRIDS  
ESTRATIGRÁFICOS**

**Luiz Gustavo Rasera**

Dissertação para obtenção do título de  
Mestre em Engenharia

Porto Alegre, RS  
2014

MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
Escola de Engenharia  
Programa de Pós-Graduação em Engenharia de Minas,  
Metalúrgica e de Materiais (PPGE3M)

**GEOESTATÍSTICA DE MÚLTIPLOS PONTOS APLICADA À  
SIMULAÇÃO DE MODELOS GEOLÓGICOS EM GRIDS  
ESTRATIGRÁFICOS**

**Luiz Gustavo Rasera**

Dissertação realizada no Laboratório de Pesquisa Mineral e Planejamento Mineiro da Escola de Engenharia da UFRGS, dentro do Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais (PPGE3M), como parte dos requisitos para a obtenção do Título de Mestre em Engenharia.

Área de Concentração: Tecnologia Mineral, Ambiental e Metalurgia Extrativa

Porto Alegre, RS

2014

Esta dissertação foi julgada adequada para a obtenção do Título de Mestre em Engenharia e aprovada em sua forma final, pelo Orientador e pela Banca Examinadora do Curso de Pós-Graduação.

Orientador:

Prof. Dr. João Felipe Coimbra Leite Costa

Banca Examinadora:

Dr. Luiz Eduardo Varella

Prof. Dr. Paulo Salvadoretti

Prof.<sup>a</sup> Dr.<sup>a</sup> Vanessa Cerqueira Koppe

Prof. Dr. Telmo Roberto Strohaecker  
Coordenador do PPGE3M

## Agradecimentos

Ao Prof. Dr. João Felipe C. L. Costa pela orientação na elaboração deste trabalho e pela oportunidade de fazer parte do Laboratório de Pesquisa Mineral e Planejamento Mineiro (LPM). Gostaria também de agradecer por todos os ensinamentos e conselhos recebidos, que certamente contribuíram para o meu crescimento profissional e pessoal.

Ao Prof. Dr. Alexandre Boucher por toda atenção e apoio, e pela oportunidade de participar de um dos seus projetos de pesquisa.

Ao Mestrando Geól. Eduardo Motta pela sugestão na escolha do depósito mineral para ilustrar o estudo de caso, construção do *wireframe* do modelo geológico e exportação do banco de dados.

A todos os colegas do LPM pela amizade e pelos bons momentos de convivência.

À Companhia Vale pela disponibilização do banco de dados.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela bolsa concedida.

# Sumário

<b>Lista de figuras</b> .....	<b>vi</b>
<b>Lista de tabelas</b> .....	<b>viii</b>
<b>Resumo</b> .....	<b>ix</b>
<b>Abstract</b> .....	<b>x</b>
<b>1 Introdução</b> .....	<b>1</b>
1.1 Meta .....	7
1.2 Objetivos específicos .....	7
1.3 Metodologia .....	8
1.4 Organização da dissertação .....	9
<b>2 Representação de modelos geológicos em <i>grids</i> estratigráficos</b> .....	<b>10</b>
2.1 Introdução .....	10
2.2 <i>Grids</i> estratigráficos .....	14
2.3 Interpolação de superfícies e volumes por krigagem.....	16
2.3.1 <i>Dual simple kriging</i> .....	17
2.3.2 <i>Dual ordinary kriging</i> .....	18
2.4 Construção de <i>grids</i> estratigráficos através do mapeamento topológico de <i>wireframes</i> .....	21
2.5 Estudo de caso.....	28
<b>3 Geoestatística de múltiplos pontos aplicada à simulação de modelos geológicos</b> .....	<b>32</b>
3.1 Introdução .....	32
3.2 Simulação sequencial.....	35
3.3 O algoritmo <i>SNESIM</i> .....	38
3.3.1 Construção da árvore de busca .....	40
3.3.2 Recuperando probabilidades da árvore de busca.....	42
3.3.3 Simulação com múltiplos <i>grids</i> .....	43

3.3.4	Reprodução da distribuição marginal .....	47
3.3.5	Particionamento da árvore de busca .....	48
3.4	Simulação de contatos litológicos com o algoritmo <i>SNESIM</i> .....	50
3.5	Estudo de caso.....	56
3.5.1	Geologia local.....	57
3.5.2	Banco de dados.....	58
3.5.3	Construção da imagem de treinamento.....	59
3.5.4	Definição da zona de incerteza e classificação dos contatos .....	61
3.5.5	Simulação dos contatos litológicos.....	63
3.5.6	Validação e pós-processamento das realizações.....	66
<b>4</b>	<b>Conclusões e recomendações .....</b>	<b>69</b>
4.1	Conclusões .....	69
4.2	Recomendações.....	71
	<b>Referências.....</b>	<b>73</b>

## Lista de figuras

Figura 1.1 – sequencia metodológica adotada.....	8
Figura 2.1 – seções verticais dos <i>grids</i> : (a) regular, (b) regular com subblocagem, (c) estratigráfico com células alinhadas verticalmente e (d) estratigráfico com células alinhadas com relação a linhas perpendiculares às superfícies do sólido geológico .....	13
Figura 2.2 – posicionamento dos vértices de referência das células unitárias dos <i>grids</i> : (a) regular (1 vértice) e (b) estratigráfico (8 vértices).....	15
Figura 2.3 – distorção no cálculo da distância horizontal causada pelo arranjo vertical das células ao longo do eixo $Z$ .....	16
Figura 2.4 – mapeamento do <i>grid</i> Cartesiano $G_B$ sobre o <i>grid</i> estratigráfico $G_\Omega$ .....	22
Figura 2.5 – mapeamento das superfícies de referência do domínio físico $\Omega$ sobre o domínio $B$ .....	24
Figura 2.6 – condições de contorno aplicadas no mapeamento das superfícies de referência $\Omega_i$ .....	24
Figura 2.7 – variáveis $\xi$ (a), $\eta$ (b), $\zeta$ (c) definidas no sistema de coordenadas $x, y, z$ .....	28
Figura 2.8 – variáveis $x$ (a), $y$ (b), $z$ (c) definidas no sistema de coordenadas $\xi, \eta, \zeta$ .....	29
Figura 2.9 – interpolação das variáveis $x$ (a), $y$ (b), $z$ (c) por <i>DuOK</i> no <i>grid</i> Cartesiano $G_B$ .....	29
Figura 2.10 – diagramas de dispersão entre os valores reais das coordenadas $x, y, z$ dos pontos do <i>wireframe</i> e das estimativas por <i>DuOK</i> .....	30
Figura 2.11 – <i>grid</i> estratigráfico gerado com base no <i>wireframe</i> da envoltória do corpo geológico .....	30
Figura 2.12 – detalhe da geometria irregular dos blocos que compõem o <i>grid</i> estratigráfico.....	31
Figura 3.1 – exemplo de um <i>template</i> de busca composto por 4 vetores.....	40
Figura 3.2 – <i>TI</i> contendo duas categorias e <i>template</i> de busca composto por 5 nós em um arranjo em forma de cruz .....	40
Figura 3.3 – árvore de busca construída com base na <i>TI</i> e no <i>template</i> de busca da Figura 3.2 .....	41
Figura 3.4 – três níveis de múltiplos <i>grids</i> : grosseiro, médio e fino .....	45
Figura 3.5 – <i>template</i> de busca com geometria $3 \times 3$ em três níveis de múltiplos <i>grids</i> .....	46
Figura 3.6 – decomposição de uma <i>TI</i> em um repositório de padrões .....	52
Figura 3.7 – seção vertical típica do depósito estudado .....	57
Figura 3.8 – modelo de blocos interpretado (a) e furos de sondagem (b) .....	59

Figura 3.9 – imagem de treinamento utilizada.....	61
Figura 3.10 – classes de partição geradas nas zonas de incerteza do modelo de blocos (a) e (c); e da <i>TI</i> (b) e (d).....	62
Figura 3.11 – conjunto de realizações geradas pelo algoritmo <i>SNESIM</i> .....	64
Figura 3.12 – conjunto de realizações geradas pelo algoritmo <i>SNESIM</i> .....	65
Figura 3.13 – conjunto de realizações geradas pelo algoritmo <i>SNESIM</i> .....	66
Figura 3.14 – proporções das categorias no modelo de referência (MR) e nas 25 realizações.....	67
Figura 3.15 – <i>E-type</i> das 4 categorias simuladas na zona de incerteza: (a) estéril, (b) itabirito duro, (c) itabirito friável e (d) hematitito.....	68
Figura 3.16 – <i>E-type</i> das 4 categorias simuladas na zona de incerteza: (a) estéril, (b) itabirito duro, (c) itabirito friável e (d) hematitito.....	68



## Lista de tabelas

Tabela 3.1 – proporções das categorias no modelo de blocos e na amostragem.....	59
Tabela 3.2 – parâmetros utilizados na construção da <i>TI</i> .....	60
Tabela 3.3 – volume da zona de incerteza no modelo de blocos e na <i>TI</i> .....	61
Tabela 3.4 – parâmetros utilizados na simulação do modelo geológico .....	63

## Resumo

Apesar da sua ampla aplicação na modelagem de formações geológicas, os algoritmos de simulação de múltiplos pontos (*MPS*) não foram desenvolvidos para simular sólidos geológicos de grandes dimensões, cujas características não são repetitivas, frequentemente encontrados na mineração. A maioria dos métodos de *MPS* baseia-se na repetição de padrões para a simulação, mas a localização espacial específica destes padrões não é um fator crítico. O método proposto trata-se de uma adaptação do algoritmo *SNESIM* para a simulação de contatos litológicos a partir de modelos de blocos interpretados, utilizando a abordagem de partição da árvore de busca. A técnica perturba o modelo geológico em um *grid* de simulação, após a definição de uma zona de incerteza ao redor dos contatos interpretados. O modelo de blocos e a imagem de treinamento (*TI*) são pré-processados por um algoritmo de agrupamento que determina a zona de incerteza e classifica os diferentes contatos em classes de partição. No modelo de blocos, a zona de incerteza corresponde ao domínio de simulação, e na *TI*, ela serve como um repositório de padrões. O tamanho da zona de incerteza controla o grau de perturbação do modelo de blocos, e as classes de partição permitem que o algoritmo lide com tendências e padrões locais da *TI*. Outro fator importante no processo de modelagem geológica é a construção de modelos numéricos que possam representar estruturas geológicas complexas e que possuam alta aderência geométrica com relação ao modelo geológico de referência. Devido a isto, a dissertação propõe uma metodologia geoestatística para a construção de *grids* estratigráficos através do mapeamento topológico das superfícies do sólido geológico. As duas metodologias são ilustradas por um estudo de caso em um depósito de minério de ferro. Os resultados obtidos mostraram que as metodologias permitem a construção de *grids* estratigráficos que se ajustam à geometria dos corpos geológicos, bem como, a geração de modelos simulados consistentes com o modelo de referência.

## Abstract

Despite its wide application in the modeling of geological formations, multiple-point simulation (*MPS*) algorithms were not designed to simulate large non-repetitive geological objects, often found in mining. Most *MPS* methods rely on pattern repetition for simulation, but the specific locations of these patterns are not critical. The proposed method is an adaptation of *SNESIM* algorithm for the simulation of lithological contacts from interpreted block models, using the search tree partition approach. The technique perturbs the geological model on a simulation grid, after defining a zone of uncertainty around the interpreted contacts. The block model and the training image (*TI*) are pre-processed by a clustering algorithm that determines the uncertainty zone and classifies the different contacts in partition classes. In the block model, the zone of uncertainty corresponds to the simulation domain, and in the *TI*, it serves as a repository of patterns. The size of the uncertainty zone controls the amount of perturbation of the block model, and the partition classes enable the algorithm to handle trends and location-specific patterns in the *TI*. Another important factor in the geological modeling process is the construction of numerical models that can represent complex geological structures and have high geometrical adherence in respect to the reference geological model. Due to this, the dissertation proposes a geostatistical methodology for the construction of stratigraphic grids through the topological mapping of the geological solid surfaces. Both methodologies are illustrated by a case study in an iron ore deposit. The results showed that the methodologies allow the construction of stratigraphic grids that fit the geometry of the geological bodies, as well as the generation of simulated models consistent with the reference model.

# Capítulo 1

## Introdução

A construção de modelos geológicos numéricos geralmente possui propósitos definidos. Na indústria mineral, esses modelos normalmente são a base de referência para o planejamento e a execução de projetos de engenharia, incluindo a sua avaliação econômica por meio de diversos tipos de funções de transferência. A modelagem geológica do corpo de minério consiste na base para todos os processos de estimativa de teores *in situ*, recursos e reservas minerais, assim como, para o projeto da mina e o planejamento da produção de longo prazo. No entanto, na prática, a maioria dos estudos de avaliação de recursos, planejamento de lavra e processamento mineral negligencia a incerteza associada à geometria e posição dos limites desses corpos geológicos.

Na avaliação de recursos minerais, os modelos geológicos são frequentemente utilizados como guias para a discriminação de teores pertencentes a diferentes populações estatísticas. Desta forma, os teores contidos em cada domínio geológico são caracterizados através de modelos de distribuição e semivariogramas específicos, possibilitando que a estimativa ou simulação dos teores seja realizada de forma diferenciada para cada população. Todavia, enquanto há uma grande preocupação com a avaliação da incerteza dos teores, a análise da incerteza com relação ao modelo geológico é frequentemente omitida.

Os métodos tradicionais de modelagem geológica de depósitos minerais fundamentam-se na definição explícita dos volumes dos corpos de minério através de processos de digitalização manual. Esses métodos requerem a construção de complexos polígonos em um conjunto de planos representando as projeções das seções dos corpos geológicos. Os polígonos criados são então ligados uns aos outros com o auxílio de *tie lines*,

linhas-guia utilizadas para definir as relações de conectividade espacial entre os objetos. Por fim, as seções são interpoladas por triangulação para gerar os sólidos tridimensionais. Esse processo é extremamente laborioso e demorado, contudo, fundamental para definir as posições espaciais e os limites entre os diferentes domínios geológicos.

Atualmente, sabe-se que o uso de métodos tradicionais de modelagem geológica não é recomendado em situações onde há um número limitado de informações e um elevado grau de incerteza associado ao modelo geológico. Tal prática é, também, uma das principais fontes responsáveis por problemas de reconciliação entre as estimativas e a produção, em termos de volumes e massas de minério (e estéril). O modelo obtido por meio dessas técnicas consiste apenas em uma representação de uma possível realidade. Tal fato impossibilita a avaliação da incerteza associada a geometria do corpo de minério, ao menos que outra equipe de especialistas gere outros modelos geológicos, o que seria tecnicamente inviável. Em alguns casos, técnicas automatizadas, tais como, a modelagem implícita através de funções de base radial (Cowan *et al.*, 2002; Cowan *et al.*, 2003), o método dos campos potenciais (Aug *et al.*, 2005; Calcagno *et al.*, 2007; Renard *et al.*, 2013) e a interpolação suavizada discreta (Mallet, 2002), podem ser utilizadas para construir diretamente o sólido geológico (*wireframe*) a partir dos dados de sondagem e afloramentos, fornecendo também, medidas de incerteza.

Os fenômenos naturais responsáveis pela formação de depósitos minerais são governados por complexos processos físicos, químicos, termodinâmicos e biológicos. Intrinsecamente, esses fenômenos são determinísticos, no entanto, as suas condições iniciais e os processos de formação e transformação envolvidos, não são conhecidos com o grau de confiança exigido para serem modelados deterministicamente a partir de um pequeno número de amostras. O nosso conhecimento limitado sobre os processos geológicos é o que torna a geoestatística extremamente útil na modelagem da distribuição espacial de variáveis regionalizadas (Matheron, 1976; David, 1977; Journel & Huijbregts, 1978; David, 1988; Christakos, 1992; Chilès & Delfiner, 1999). A geoestatística consiste em uma abordagem estatística para modelagem, em oposição aos modelos determinísticos que se preocupam com a descrição genética dos processos físicos. Os métodos geoestatísticos baseiam-se no conceito de modelos de funções aleatórias, e descrevem os fenômenos espaciais como conjuntos de variáveis aleatórias governadas por leis probabilísticas.

A abordagem geoestatística tradicional para fins de modelagem geológica é realizada, geralmente, através da simulação sequencial de variáveis categóricas, das quais, destacam-se

a simulação gaussiana truncada (Journel & Isaaks, 1984; Matheron *et al.*, 1987; Xu & Journel, 1993) e a simulação sequencial dos indicadores (Alabert, 1987; Journel, 1989; Alabert & Massonnat, 1990). O objetivo da simulação sequencial, da forma que foi originalmente proposta, é a reprodução do histograma e do modelo de covariância das propriedades a serem simuladas através da tiragem sequencial de distribuições condicionais (*ccdfs*). Um caminho aleatório visita sequencialmente cada nó do *grid* e valores simulados são retirados da distribuição condicional do valor naquele nó, com base na vizinhança de dados condicionantes e nós previamente simulados. Outros exemplos de técnicas de simulação utilizadas na modelagem geológica são os algoritmos baseados em objetos (*object-based*) ou booleanos (Haldorsen & Lake, 1984; Haldorsen & Chang, 1986; Stoyan *et al.*, 1987), algoritmos baseados em processos (*process-based*) (Bridge & Leeder, 1979; Lopez, 2003), métodos de modelagem de superfícies (*surface-based*) (Deutsch & Xie, 2001; Pyrcz & Deutsch, 2005) e, mais recentemente, algoritmos de simulação de múltiplos pontos baseados em *pixels* (*pixel-based*) (Guardiano & Srivastava, 1993; Strebelle, 2002).

No entanto, os algoritmos tradicionais de simulação sequencial, limitados à reprodução de estatísticas de dois pontos, tal como o modelo de semivariograma, não são capazes de reproduzir estruturas geológicas complexas. Essas estruturas, tais como, canais meandantes, por exemplo, possuem geometria curvilínea e apresentam complexas interdependências. Desta forma, a sua modelagem requer as correlações entre múltiplos pontos e diferentes escalas, as quais estão muito além da capacidade de caracterização oferecida pelos modelos de covariância (Arpat, 2005).

Dentre as diversas técnicas de simulação geoestatística recentemente desenvolvidas, encontra-se a simulação geoestatística de múltiplos pontos (*multiple-point statistics – MPS*). O conceito de *MPS* foi proposto inicialmente por Journel (1992), tendo o seu primeiro algoritmo elaborado por Guardiano & Srivastava (1993) baseado em uma extensão do processo de simulação sequencial. O método visava, inicialmente, combinar a capacidade de condicionamento oferecida pelos algoritmos baseados em *pixels*, com o potencial dos algoritmos baseados em objetos em reproduzir formas geológicas complexas. Esta vertente das técnicas de simulação estocástica se trata de uma alternativa para a simulação tradicional baseada em semivariogramas. O método substitui o processo de descrição da continuidade espacial através de modelos analíticos, com o uso de imagens de treinamento (*training image – TI*), objetos que consistem em representações explícitas dos padrões geológicos.

As imagens de treinamento refletem um conceito geológico prévio, podendo ser consideradas um grupo de modelos de funções aleatórias que fornecem uma quantificação das heterogeneidades existentes nos padrões geológicos. O processo de construção de *TIs* tridimensionais é geralmente realizado por meio de algoritmos de simulação baseados em objetos, ou por algoritmos de simulação de processos físicos. Esses métodos possibilitam a geração de complexas feições geológicas, no entanto, apresentam capacidade de condicionamento limitada. Outra fonte comum para a obtenção de imagens de treinamento é através da digitalização manual de sólidos geológicos em *softwares* de modelagem 3D.

Apesar das vantagens oferecidas em teoria, o algoritmo originalmente proposto por Guardiano & Srivastava (1993) exigia uma elevada demanda computacional (CPU e RAM), fator que acabou inviabilizando a sua aplicação prática durante os anos 90. A utilização da técnica só foi possível após o desenvolvimento do algoritmo *SNESIM* (Strebelle, 2002). O método introduziu a utilização de uma estrutura de dados especial chamada *search tree* (árvore de busca) para armazenar e recuperar de forma rápida todos os eventos relevantes presentes na imagem de treinamento. Posteriormente, diversos trabalhos ilustraram aplicações práticas do algoritmo por meio de estudos de caso reais em reservatórios de petróleo (Caers *et al.*, 2001; Strebelle *et al.*, 2002; Caers *et al.*, 2003; Liu *et al.*, 2004; Harding *et al.*, 2005).

Todavia, a aplicação dos algoritmos de *MPS* se torna problemática em situações onde o ambiente geológico é constituído por sólidos maciços com grandes dimensões e que não se repetem. Tais casos são comuns, sendo frequentemente encontrados na mineração, como, por exemplo, depósitos de filões auríferos, pórfiros cupríferos e formações ferríferas (Boucher *et al.*, 2014). Nessas situações, geralmente, a localização dos corpos de minério é conhecida, contudo, a sua geometria e extensão precisa são incertas. Os algoritmos tradicionais de *MPS* não foram desenvolvidos para tais aplicações. Geralmente, as imagens de treinamento contêm todos os padrões geológicos relevantes, contudo, a localização espacial específica destes não é um fator crítico. Essa abordagem é adotada por diversos algoritmos, tais como: *SNESIM* (Strebelle, 2002), *FILTERSIM* (Zhang *et al.*, 2006; Wu *et al.*, 2008), *SIMPAT* (Arpat & Caers, 2007), *IMPALA* (Straubhaar *et al.*, 2010), *Direct Sampling* (Mariethoz *et al.*, 2010), assim como, nos trabalhos apresentados por Boucher (2009), Peredo & Ortiz (2011), Parra & Ortiz (2011) e Tahmasebi *et al.* (2012).

Uma exceção a essa regra é o algoritmo *DISPAT* (Honarkhah & Caers, 2010) que possibilita perturbar uma imagem de treinamento definida no próprio *grid* de simulação.

Neste caso, o grau de perturbação é controlado pela distância máxima entre o nó a ser simulado e os padrões considerados, o que promove uma restrição dos padrões encaixantes àqueles próximos ao local a ser simulado. Outros algoritmos caracterizam-se por impor as posições espaciais das unidades geológicas através do uso de campos de probabilidade para controlar a localização das categorias a serem simuladas (Strebelle, 2002; Harding *et al.*, 2005). A abordagem baseada em campos de probabilidade força a simulação dos corpos pertencentes a determinada categoria através do fornecimento de valores elevados de probabilidade de ocorrência em locais predefinidos. Entretanto, segundo Boucher (2013), esta metodologia é inadequada pois leva somente em consideração as tendências das proporções entre categorias, e não as tendências com relação aos tipos de padrões. Determinados algoritmos possibilitam a reprodução de tendências referentes aos tipos de padrões, como o *IMPALA* (Straubhaar *et al.*, 2010), por exemplo. Tal característica, possibilitaria a simulação de corpos geológicos com grandes dimensões por meio do uso de funções de tendência apropriadas. Cabe salientar, no entanto, que para alguns ambientes geológicos, tais como, depósitos de veios mineralizados (Boisvert *et al.*, 2008), os processos tradicionais de construção de imagens de treinamento e simulação por *MPS* podem ser mais facilmente adaptados.

Outra metodologia proposta para a simulação de modelos geológicos consiste na inferência direta de estatísticas de múltiplos pontos a partir de dados de sondagem. Tal abordagem é apresentada por Dimitrakopoulos *et al.* (2010). Nesse trabalho, os autores inferiram e modelaram estatísticas de múltiplos pontos usando cumulantes, que foram subsequentemente utilizadas para simulação através do algoritmo *HOSIM* (Mustapha & Dimitrakopoulos, 2010). Todavia, o uso dessa metodologia é somente adequado em situações onde há uma elevada densidade de dados de sondagem.

Recentemente, diversos trabalhos propuseram diferentes abordagens para a aplicação dos algoritmos de *MPS* em problemas reais de mineração. Osterholt & Dimitrakopoulos (2007) apresentaram um estudo de caso ilustrando a aplicação do algoritmo *SNESIM* para a modelagem geológica de um depósito de minério de ferro. Goodfellow *et al.* (2012) utilizaram o mesmo algoritmo e uma imagem de treinamento definida no próprio *grid* de simulação, composto por um único corpo maciço mineralizado, para simular um depósito polimetálico. Um exemplo de adaptação do algoritmo *SNESIM* para a simulação de um depósito de minério de ferro foi apresentado por Pasti *et al.* (2012). Aplicações do mesmo algoritmo utilizando a abordagem por campos de probabilidade foram realizadas por Silva Jr.



*et al.* (2013) e Müller *et al.* (2013). Os autores ilustraram os trabalhos com estudos de caso em depósitos de minério de ferro e cobre, respectivamente. Deraisme & Assibey-Bonsu (2013) apresentaram a aplicação do algoritmo *IMPALA* na modelagem de zonas mineralizadas em um depósito hidrotermal. Recentemente, Boucher *et al.* (2014) propuseram uma modificação do algoritmo *SNESIM* para a simulação de contatos geológicos, utilizando o próprio modelo de blocos interpretado como imagem de treinamento. O método foi ilustrado através de um estudo de caso em um depósito de minério de ferro.

O presente estudo trata-se de uma extensão do trabalho de Boucher *et al.* (2014). No entanto, esta dissertação propõe o uso de uma imagem de treinamento gerada exclusivamente por um processo de simulação de objetos. A ideia é fazer uso da simulação geoestatística de múltiplos pontos como instrumento para perturbar o modelo geológico interpretado em um *grid* de simulação. O resultado dessa abordagem consistiria em um conjunto de cenários equiprováveis que poderia ser utilizado posteriormente para caracterizar a incerteza com relação à geometria e a localização dos contatos entre os diferentes domínios geológicos, otimizar a locação de furos de sondagem, e como parâmetro adicional em estudos de sensibilidade e análise de risco.

O método proposto fundamenta-se na hipótese básica de que a incerteza com relação aos contatos entre diferentes litotipos reduz a medida que nos distanciamos dos contatos interpretados em um modelo geológico. Desta forma, assume-se que a posição espacial de um determinado conjunto de corpos geológicos é conhecida, contudo, a disposição dos contatos interpretados é classificada como incerta. Este pressuposto permite a simplificação do problema da modelagem geológica por meio da simulação geoestatística, visto que as únicas porções a serem simuladas do *grid* consistem nos nós localizados próximos aos contatos interpretados (zonas de incerteza). Neste caso, a imagem de treinamento necessita conter somente os padrões geométricos encontrados nos contatos interpretados pelo especialista.

A construção de modelos geológicos numéricos de alta resolução é essencial para o aumento da aderência geométrica de modelos simulados com relação ao modelo geológico de referência. Tal prática contribui para a obtenção de estimativas mais precisas de volume/massa, reduzindo problemas de diluição ocasionados pelo ajuste geométrico inadequado dos modelos de blocos tradicionais às estruturas curvilíneas dos corpos geológicos. Entretanto, na indústria mineral, este fator é frequentemente negligenciado. Devido a esta deficiência, essa dissertação propõe o desenvolvimento de uma metodologia

para a construção de *grids* estratigráficos, através do mapeamento topológico das superfícies do *wireframe* da envoltória do corpo mineralizado.

## 1.1 Meta

A meta desta dissertação consiste em avaliar o uso da geoestatística de múltiplos pontos como ferramenta auxiliar no processo de modelagem geológica de depósitos minerais. O presente estudo apresenta uma adaptação do algoritmo *SNESIM* para a simulação de contatos litológicos baseada na abordagem de partição da árvore de busca. O método é ilustrado por meio de um estudo de caso em um depósito de minério de ferro localizado no estado de Minas Gerais. Adicionalmente, propõe-se uma metodologia para a construção de *grids* estratigráficos, possibilitando o aumento da aderência geométrica dos modelos simulados com relação ao modelo de referência, assim como, a reprodução de modelos geológicos mais realistas.

## 1.2 Objetivos específicos

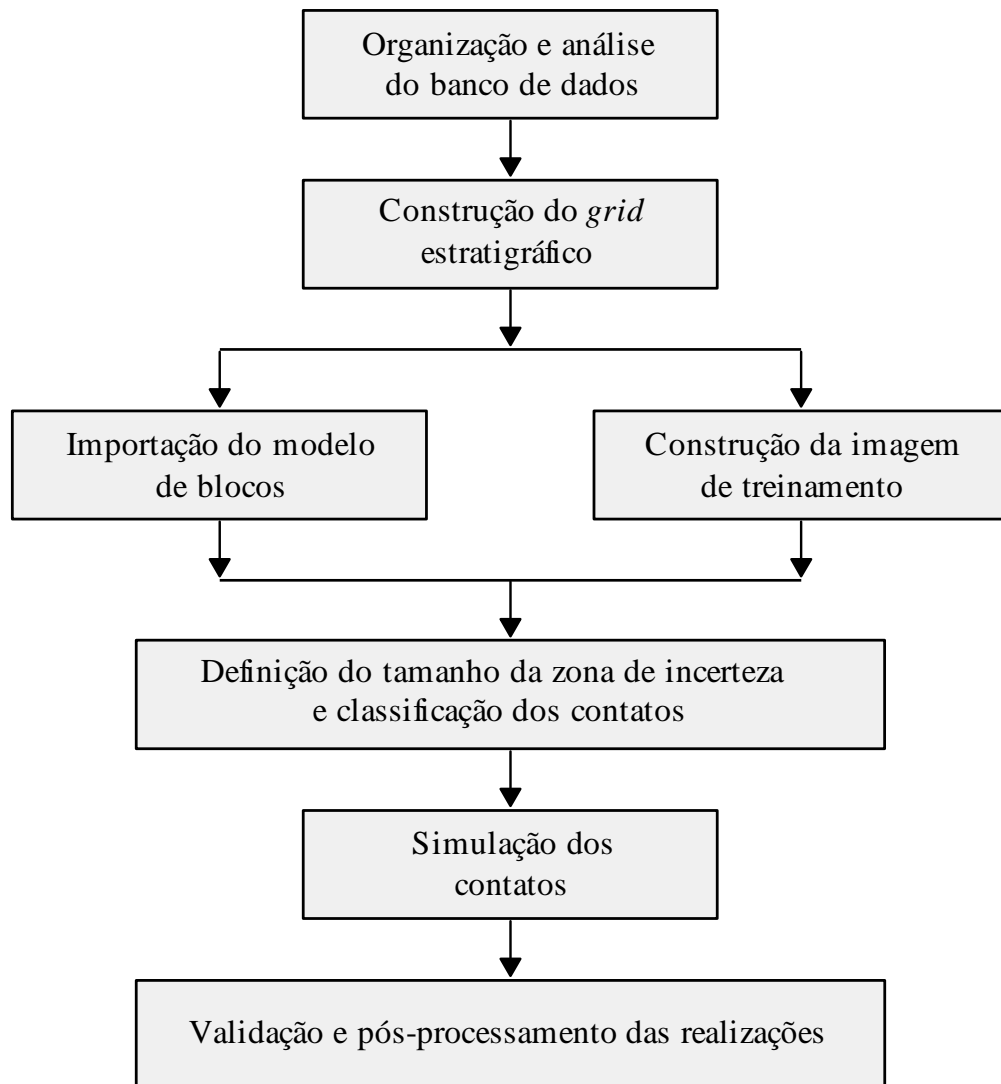
Esta dissertação propõe atingir a sua meta cumprindo os seguintes objetivos específicos:

- Desenvolver uma metodologia para a construção de um *grid* estratigráfico com base no modelo geológico-estrutural do depósito mineral, definido a partir das superfícies de referência do *wireframe* da envoltória do corpo de minério;
- Gerar uma imagem de treinamento 3D contendo os padrões geométricos dos contatos litológicos interpretados nas seções verticais por meio do algoritmo *TetrisTiGen*;
- Desenvolver um algoritmo para detectar e classificar os contatos litológicos presentes no modelo de blocos interpretado e na imagem de treinamento em diferentes classes, com base em diversos tamanhos da zona de incerteza. O resultado desta classificação definirá quais blocos devem ser simulados;

- Simular os contatos litológicos do modelo de blocos através de uma adaptação do algoritmo *SNESIM* utilizando a metodologia de partição da árvore de busca;
- Validar os resultados e pós-processar as realizações.

### 1.3 Metodologia

O processo de simulação do modelo geológico apresentado no estudo de caso exigiu a definição e a execução de uma sequência lógica de etapas. Desta forma, para atingir os objetivos propostos nesta dissertação foi adotada a seguinte metodologia de trabalho, apresentada na Figura 1.1.



**Figura 1.1** – sequência metodológica adotada.

## 1.4 Organização da dissertação

Os capítulos desta dissertação estão organizados da seguinte maneira:

O Capítulo 1 realiza uma introdução ao tema referente à aplicação da simulação geoestatística de múltiplos pontos na modelagem geológica de depósitos minerais, assim como, a exposição do problema e a justificativa para o trabalho. Em seguida, são apresentadas a meta da dissertação, os objetivos específicos e a sequência metodológica adotada.

O Capítulo 2 apresenta uma abordagem geoestatística para a representação de modelos geológicos numéricos em *grids* estratigráficos através da interpolação por krigagem. A metodologia proposta é ilustrada na forma de um estudo de caso.

O Capítulo 3 faz uma breve revisão da simulação sequencial e dos fundamentos teóricos da geoestatística de múltiplos pontos e a sua aplicação na simulação de modelos geológicos. Posteriormente, é feita a apresentação do algoritmo *SNESIM* e propõe-se uma adaptação do mesmo para a simulação de contatos litológicos. O método é ilustrado através de um estudo de caso em um depósito de minério de ferro.

O Capítulo 4 inclui as conclusões do presente estudo e recomendações para trabalhos futuros.

## Capítulo 2

# Representação de modelos geológicos em *grids* estratigráficos

Este capítulo introduz os principais fundamentos e características das técnicas de representação numérica de modelos geológicos. Em seguida, é apresentada a metodologia desenvolvida para a construção de *grids* estratigráficos. A Seção 2.1 realiza um breve histórico da evolução das técnicas informatizadas de representação de modelos geológicos utilizadas nos setores de exploração de recursos minerais e energéticos. A Seção 2.2 introduz os conceitos básicos e as propriedades da estrutura de dados dos *grids* estratigráficos. A Seção 2.3 apresenta os fundamentos teóricos de uma das variações das técnicas de krigagem, chamada *dual kriging*, desenvolvida especificamente para a interpolação de superfícies e volumes. A Seção 2.4 propõe uma metodologia geoestatística para a construção de *grids* estratigráficos através do mapeamento topológico das superfícies de referência do *wireframe* do modelo geológico. Por fim, a Seção 2.5 ilustra a aplicação da metodologia proposta na forma de um estudo de caso em um depósito de minério de ferro.

### 2.1 Introdução

Fatores como o esgotamento de reservas minerais e energéticas, problemas relacionados a contaminações ambientais, e a elevação constante dos custos em todos os setores econômicos ao longo dos anos, tornaram necessário o aumento da precisão e eficiência das técnicas de modelagem geológica e avaliação de recursos (Houlding, 1994). No entanto, a evolução da informatização desse processo tomou caminhos distintos, influenciada

principalmente pelas necessidades específicas de cada setor geocientífico, e devido às limitações tecnológicas impostas pela época. Nos últimos anos, com o advento de computadores de alto desempenho, estes diferentes processos de caracterização convergiram para uma forma mais genérica.

No setor mineral, o principal objetivo das técnicas computacionais de avaliação de recursos sempre foi a estimativa da variabilidade espacial dos teores, com aplicações iniciais em grandes depósitos de metais base do tipo pórfiro disseminado. A construção de modelos geológicos numéricos envolve a representação do subsolo através da discretização dos *wireframes* em conjuntos de blocos retangulares uniformes. Esses blocos constituem *grids* regulares tridimensionais (3D), os quais são chamados modelos de blocos. Esta estrutura de dados permite que os sólidos geológicos sejam subdivididos em matrizes de blocos, onde as dimensões dos blocos são definidas de acordo com o método de lavra adotado, o espaçamento da malha de amostragem e as condições geológicas do depósito. Posteriormente, com base em informações amostrais, diversas técnicas de interpolação são utilizadas para estimar (ou simular) o teor médio de cada bloco. Esta metodologia é análoga a abordagem *raster* na computação, uma vez que cada bloco representa uma medida contínua de variabilidade em uma densidade espacial fixa de informação. A representação por meio de modelos de blocos permite que valores monetários possam ser informados para cada bloco, com base no teor de metal, custos de lavra e beneficiamento. Desta forma, em operações de lavra a céu aberto, por exemplo, algoritmos de otimização podem ser aplicados sobre o próprio *grid* para a determinação dos limites da cava ótima (*pit optimization*).

Com o passar do anos, os modelos de blocos foram aprimorados através da adição de ferramentas que possibilitaram a imposição do controle geológico no processo de estimativa (ou simulação) de teores. Tal prática faz uso de indicadores para representar um determinado litotipo para cada bloco, possibilitando assim, a discriminação de diferentes mineralizações, por exemplo. Contudo, a abordagem envolve aproximações significativas, uma vez que a precisão da representação das estruturas geológicas é controlada pelo tamanho dos blocos (Figura 2.1a). Uma solução frequentemente aplicada em casos onde uma maior resolução é necessária (ou para remediar, até certo ponto este problema), consiste na utilização de algoritmos que permitem a subdivisão dos blocos (subblocagem) (Figura 2.1b). No entanto, para determinadas situações, tal prática torna-se computacionalmente inviável, até mesmo para computadores de alto desempenho.

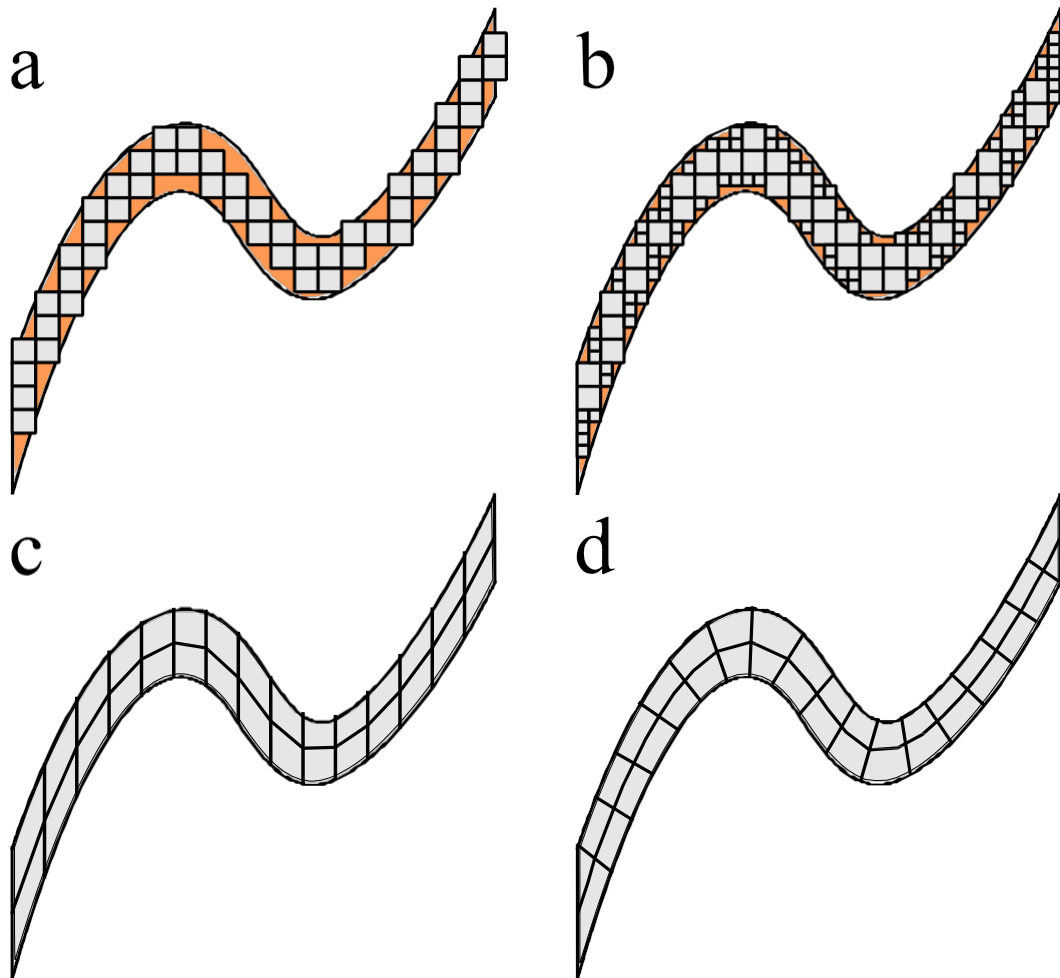
Os *grids* regulares 3D consistem em uma estrutura de dados apropriada para a representação da variabilidade espacial de teores. No entanto, são totalmente inadequados para representar a geometria curvilinear dos corpos geológicos, que são raramente encontrados na forma de blocos regulares. Além disso, na prática, o tamanho de célula ótimo para a representação da variabilidade espacial dos teores é raramente adequado para reproduzir a geologia com um nível de precisão aceitável.

Por outro lado, no setor de recursos energéticos (petróleo, gás natural e carvão), sempre houve uma maior ênfase na representação da estratigrafia e geometria das estruturas geológicas (Houlding, 1994). Em aplicações ligadas à exploração de petróleo e gás, existe uma grande preocupação na identificação e representação das formações geológicas que podem consistir em potenciais reservatórios. As etapas iniciais do processo de modelagem de reservatórios consistem no estabelecimento da arquitetura, ou geometria, da formação que contem os hidrocarbonetos, e de outras formações geológicas adjacentes (Deutsch, 2002). Na mineração de carvão, por exemplo, são realizados diversos estudos para estimar as espessuras e extensões das camadas de carvão e estéril.

Inicialmente, essas aplicações levaram ao desenvolvimento de metodologias de modelagem geológica baseadas em superfícies (*surface-based*). Tais métodos se caracterizam por serem derivados de observações de características estruturais das formações geológicas, ao invés de propriedades físicas e químicas das variáveis. Estas representações envolvem a construção de *grids* 2D horizontais regulares (*surface models*). A abordagem permite que cada estrato, ou superfície de interesse, pertencente a determinada formação, seja definida com base em valores de elevação e profundidade (ou espessura) em cada nó do *grid*, através de uma grande variedade de técnicas de interpolação. Esta metodologia é análoga a abordagem computacional para a construção de superfícies de contorno (mapas de isópacas e isócoras, por exemplo), com a exceção de que um único modelo pode conter diversas superfícies ou camadas. Um modelo de superfície consiste, essencialmente, em uma estrutura de dados do tipo *raster*, pois assume uma variação contínua no plano horizontal, e a sua resolução espacial é controlada pelo tamanho das células.

Com o passar dos anos, a representação por malhas superficiais (*gridded surfaces*) tem sido refinada através da implementação de *grids* 3D deformados, também conhecidos como *grids* curvilineares ou estratigráficos. Esse tipo de estrutura de dados permite que a malha seja regular no plano horizontal, mas irregular na terceira dimensão (vertical) (Figura 2.1c); ou, até

mesmo, irregular em todas as três dimensões (Figura 2.1d). Basicamente, o produto resultante consiste em uma combinação das abordagens *raster* e vetorial. A utilização de malhas 3D promove uma melhor representação da geologia se comparada ao uso de malhas superficiais, contudo, cria também, um problema adicional na integração de técnicas de interpolação. As implementações tradicionais de determinados algoritmos de estimativa e simulação geoestatística não podem ser aplicadas diretamente em *grids* curvilíneas. No caso de *grids* estratigráficos não estruturados, por exemplo, estes algoritmos requerem modificações que os permitam lidar com vizinhanças de busca irregulares e células de diferentes suportes (Deutsch *et al.*, 2002). Mais recentemente, metodologias similares têm sido desenvolvidas utilizando superfícies triangulares, as quais fornecem grande flexibilidade, e um elevado grau de detalhe, consistindo em uma abordagem verdadeiramente vetorial para o problema.



**Figura 2.1** – seções verticais dos *grids*: (a) regular, (b) regular com subblocagem, (c) estratigráfico com células alinhadas verticalmente e (d) estratigráfico com células alinhadas com relação a linhas perpendiculares às superfícies do sólido geológico.



## 2.2 *Grids* estratigráficos

*Grids* (ou malhas) consistem no produto da discretização de um domínio geométrico 2D, ou 3D, em um conjunto de formas geométricas menores. Eles fornecem um meio eficiente para a representação da variabilidade espacial de propriedades pertencentes a um determinado meio físico. Esta estrutura de dados permite a representação de uma medida espacial contínua da variação de uma propriedade, atribuindo valores representativos de uma pequena porção finita de volume, na forma de nós dispostos em intervalos regulares. No caso de modelos de blocos 3D, tradicionalmente, os nós localizam-se nos centróides dos blocos (Figura 2.2a). O valor estimado em cada bloco, geralmente, consiste no teor médio obtido a partir da discretização do bloco em um conjunto de pontos. Assim, as estimativas são realizadas nos pontos, e têm a sua média armazenada no centróide do bloco.

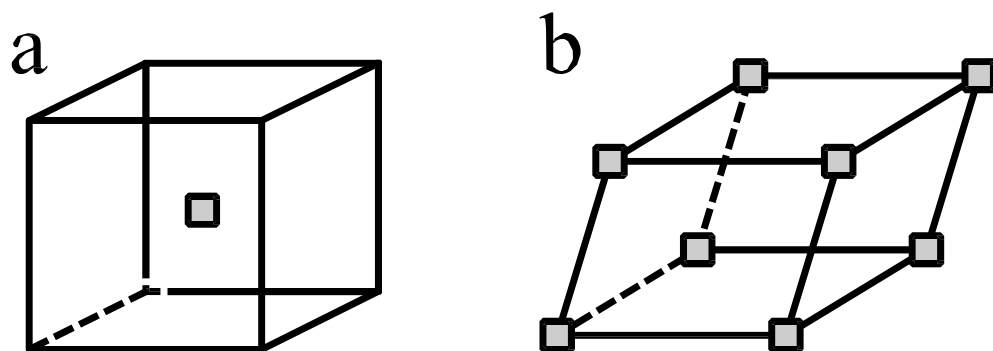
*Grids* estruturados 3D (*structured grids*) consistem, essencialmente, no produto da discretização de um espaço euclidiano 3D através de um conjunto de hexaedros, onde cada célula que compõe o *grid* é identificada por um índice. Estes modelos consistem em uma estrutura de dados que apresenta conectividade regular, o que os torna altamente eficientes para o armazenamento de informações, uma vez que as relações de vizinhança entre os dados são definidas pelo arranjo de armazenamento (Castillo, 1991; George, 1991).

Em contrapartida, *grids* 3D não estruturados (*unstructured grids*) tratam-se do produto da discretização de um espaço 3D em formas geométricas (usualmente tetraédros) dispostas em um arranjo irregular. Estes modelos caracterizam-se por apresentar conectividade irregular, o que impede que sejam facilmente representados por meio de vetores na memória do computador. Devido à esta característica, eles requerem uma lista de conectividade que especifique a maneira em que um conjunto de vértices de referência deve formar os elementos estruturais. Esta peculiaridade os torna relativamente ineficientes em processos de armazenamento de dados se comparados aos *grids* estruturados, pois os mesmos requerem a armazenagem explícita das relações de vizinhança dos dados (Mavriplis, 1996). No entanto, os *grids* não estruturados oferecem um grau muito maior de flexibilidade na representação de formas geométricas complexas. Esse tipo de estrutura de dados não é abordada no presente trabalho.

As principais variedades de *grids* estruturados, são: o *grid* Cartesiano, os *grids* retilineares, e os *grids* curvilineares ou estratigráficos. O *grid* Cartesiano 3D trata-se de um

caso especial dos *grids* regulares, onde as células consistem em cubos unitários e seus vértices são pontos localizados em coordenadas inteiras. Os *grids* reticulados 3D consistem na tesselação por paralelepípedos (retos ou oblíquos) que podem ser, ou não, congruentes uns aos outros. No caso de *grids* 3D curvilíneos ou estratigráficos, os modelos de blocos apresentam a mesma estrutura de dados de *grids* regulares, contudo, as células consistem em hexaedros irregulares. Esta característica os torna extremamente úteis para a modelagem de formas geométricas complexas, como as encontradas em formações geológicas.

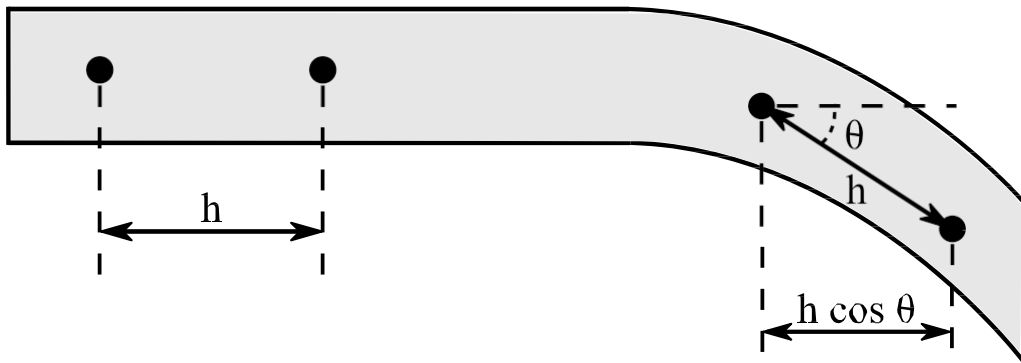
*Grids* estratigráficos 3D são estruturas de dados que permitem um elevado grau de flexibilidade na representação de corpos geológicos, visto que podem se conformar com facilidade no processo de discretização de *wireframes*. No entanto, eles exigem um número maior de informações para construir e posicionar as suas células, se comparados aos modelos de blocos regulares. Para definir cada célula unitária, as quais consistem em hexaedros irregulares, esses modelos necessitam das coordenadas dos oito vértices de referência que as compõem (Figura 2.2b). Por outro lado, esta propriedade permite que os vértices dos blocos se adaptem de acordo com os limites do corpo geológico, o que possibilita uma redução significativa do número total de blocos. Portanto, mesmo que os modelos sejam limitados a um número pequeno de blocos, eles frequentemente permitem uma representação muito mais econômica e acurada dos sólidos geológicos.



**Figura 2.2** – posicionamento dos vértices de referência das células unitárias dos *grids*:  
(a) regular (1 vértice) e (b) estratigráfico (8 vértices).

Geralmente, os *grids* estratigráficos são construídos adotando um empilhamento das células ao longo do eixo vertical Z (Figura 2.1c). No entanto, os modelos também podem ser construídos com as células alinhadas às linhas perpendiculares às superfícies de referência do sólido geológico (Figura 2.1d). A última alternativa possibilita uma representação mais correta das distâncias horizontais entre os nós da malha, pois leva em consideração a

correlação cronoestratigráfica entre os estratos, contudo, ao preço de um custo computacional adicional (Deutsch, 2002). O presente estudo faz uso específico desta última estrutura de dados. A seguir, a Figura 2.3 ilustra a distorção no cálculo da distância horizontal ao se adotar um alinhamento vertical das células ao longo do eixo  $Z$ . Ao analisarmos a situação com base nas coordenadas estratigráficas, ambos os pares de pontos encontram-se separados pela mesma distância  $h$ ; contudo, no par localizado à direita, os pontos parecem estar mais próximos.



**Figura 2.3** – distorção no cálculo da distância horizontal causada pelo arranjo vertical das células ao longo do eixo  $Z$  (modificado de Deutsch, 2002).

### 2.3 Interpolação de superfícies e volumes por krigagem

A técnica de interpolação geoestatística conhecida como *dual kriging* (*DuK*) consiste em uma das diversas variações dos algoritmos de krigagem. Neste método, as estimativas são expressas como combinações lineares dos valores de covariância, ao invés dos valores dos dados (Dubrule, 1983; Journel, 1989). O formalismo da *DuK* explora as propriedades de filtragem da krigagem, além de reduzir o custo computacional da krigagem quando aplicada usando uma vizinhança de busca global (Goovaerts, 1997). A interpolação por *DuK* combina as qualidades de acurácia da formulação tradicional da krigagem, com a capacidade de geração de linhas de contorno suaves características de interpoladores baseados em *splines* (Dubrule, 1983).

A krigagem ordinária (*OK*), em sua formulação tradicional, pode ser classificada como um estimador discreto, no sentido em que exige a solução de um sistema de krigagem para cada ponto no espaço no qual se realiza uma estimativa. A formulação dupla da krigagem ordinária, conhecida como *dual ordinary kriging* (*DuOK*), fornece um estimador contínuo uma vez que, para um dado conjunto de dados, requer a solução de apenas um único sistema

de krigagem, fazendo com que a estimativa resultante assuma a forma de uma função definida continuamente no espaço. No entanto, os benefícios da técnica só podem ser capitalizados se uma vizinhança de busca global é utilizada (Auñón & Gómez-Hernández, 2000).

A *DuOK* (Matheron, 1975; Matheron, 1980; Dubrule, 1983; Galli & Murillo, 1984; Royer & Vieira, 1984; Journel, 1989; Zhu, 1992) trata-se de uma formulação alternativa da *OK*, onde o valor estimado é obtido por meio da combinação linear de funções de covariância. Tal característica faz com que o método seja particularmente útil em problemas que necessitam de estimativas para um grande número de locais, ou em situações onde a disponibilidade de uma expressão analítica é vantajosa, tais como: a renderização de superfícies, a determinação de perfis longitudinais, ou o cálculo de volumes (Auñón & Gómez-Hernández, 2000). O tipo de superfície a ser gerada pelo interpolador depende do tipo de função de covariância utilizada. Se o objetivo do estudo é gerar de superfícies suavizadas, por exemplo, funções de covariância como o modelo gaussiano são preferidas em relação aos modelos exponencial e esférico. A seguir, são apresentados os fundamentos teóricos de duas variações da técnica: *dual simple kriging (DuSK)* e *dual ordinary kriging (DuOK)*.

### 2.3.1 Dual simple kriging

Considere, inicialmente, o estimador de krigagem simples (*SK*) de uma determinada variável  $z$  no local  $\mathbf{u}$  (Equação 2.1):

$$\hat{z}(\mathbf{u}) = \sum_{\alpha} \lambda_{\alpha}^{SK}(\mathbf{u}) z(\mathbf{u}_{\alpha}) \quad (2.1)$$

Os pesos de *SK*  $\lambda_{\alpha}^{SK}(\mathbf{u})$  são obtidos através da resolução do sistema de  $n(\mathbf{u})$  equações lineares, conforme a Equação 2.2:

$$\sum_{\alpha} \lambda_{\alpha}^{SK}(\mathbf{u}) C(\mathbf{u}_{\alpha} - \mathbf{u}) = C(\mathbf{u}_{\alpha} - \mathbf{u}) \quad (2.2)$$

onde, temos que os pesos de krigagem aparecem como funções lineares dos valores de covariância  $C(\mathbf{u}_{\alpha} - \mathbf{u})$ . Conseqüentemente, o estimador de *SK* apresentado pela Equação 2.1

pode ser expresso como sendo a combinação linear destes valores de covariância mais uma média estacionária  $m$  (Equação 2.3):

$$Z(\mathbf{u}) = \sum_{\alpha} d_{\alpha}^{SK}(\mathbf{u}) Z(\mathbf{u}_{\alpha}) + m \quad (2.3)$$

onde,  $d_{\alpha}^{SK}(\mathbf{u})$  é o peso duplo associado a covariância  $C(\mathbf{u}_{\alpha} - \mathbf{u})$ . A Equação 2.3 é conhecida como a forma dupla do estimador tradicional de  $SK$  (Equação 2.1). Os pesos duplos  $d_{\alpha}^{SK}(\mathbf{u})$  são obtidos por meio da identificação dos valores dos dados através da seguinte expressão de krigagem (Equação 2.4):

$$\sum_{\alpha} d_{\alpha}^{SK}(\mathbf{u}) Z(\mathbf{u}_{\alpha}) = Z(\mathbf{u}) \quad (2.4)$$

Em contraste com o sistema tradicional de  $SK$  (Equação 2.2), o sistema duplo expresso pela Equação 2.4 não resulta da redução da variância do erro, em vez disso, o mesmo é estabelecido a partir da propriedade de exatidão do estimador de krigagem (Goovaerts, 1997).

Da mesma forma que a notação  $\lambda_{\beta}^{SK}(\mathbf{u})$  utilizada para o pesos originais no sistema apresentado pela Equação 2.2, a dependência em  $\mathbf{u}$  dos pesos duplos  $d_{\beta}^{SK}(\mathbf{u})$  refere-se ao fato de que os  $n(\mathbf{u})$  dados contidos podem variar de um local  $\mathbf{u}$  para outro. Para uma determinada configuração de dados, os pesos duplos  $d_{\beta}^{SK}(\mathbf{u})$  e  $d_{\beta}^{SK}(\mathbf{u}')$  são os mesmos para todos  $\beta$ .

### 2.3.2 Dual ordinary kriging

A  $OK$ , em sua formulação tradicional, fornece uma estimativa local por meio da combinação linear dos dados amostrais na forma de um conjunto de pesos, os quais necessitam ser determinados para que cada nova coordenada espacial possa ser estimada. A sua expressão é dada pela Equação 2.5:

$$Z(\mathbf{u}) = \sum_{\alpha} \lambda_{\alpha}(\mathbf{u}) Z(\mathbf{u}_{\alpha}) \quad (2.5)$$

onde,  $z^*(\mathbf{u})$  é o valor estimado,  $z(\mathbf{u}_\alpha)$  são os dados amostrais,  $n(\mathbf{u})$  o número de dados, e  $\lambda^{OK}_\alpha(\mathbf{u})$  são os coeficientes de ponderação, os quais não dependem somente dos dados amostrais, mas também do local  $\mathbf{u}$  a ser estimado.

O sistema de *OK* inclui  $(n(\mathbf{u}) + 1)$  equações lineares com  $(n(\mathbf{u}) + 1)$  incógnitas: os  $n(\mathbf{u})$  pesos  $\lambda^{OK}_\alpha(\mathbf{u})$  e o parâmetro de Lagrange  $\mu_{OK}(\mathbf{u})$ , responsável pela condição de soma dos pesos (Equação 2.6):

$$\begin{cases} \sum_{\alpha} (\dots) (\dots) (\dots) (\dots) (\dots) & (\dots) \\ \sum_{\alpha} (\dots) & (\dots) \end{cases} \quad (2.6)$$

De forma análoga à *DuSK*, a estimativa por krigagem ordinária é expressa na sua forma dupla como uma combinação linear das covariâncias  $C(\mathbf{u}_\alpha - \mathbf{u})$  mais a estimativa da tendência  $m^*_{OK}(\mathbf{u})$ , conforme a Equação 2.7:

$$(\dots) \sum_{\alpha} (\dots) (\dots) (\dots) (\dots) \quad (2.7)$$

O sistema de *DuOK* inclui  $(n(\mathbf{u}) + 1)$  equações lineares com  $(n(\mathbf{u}) + 1)$  incógnitas: os  $n(\mathbf{u})$  pesos duplos  $d^{OK}_\alpha(\mathbf{u})$  e a estimativa da tendência  $m^*_{OK}(\mathbf{u})$  (Equação 2.8):

$$\begin{cases} \sum_{\alpha} (\dots) (\dots) (\dots) (\dots) (\dots) & (\dots) \\ \sum_{\alpha} (\dots) & (\dots) \end{cases} \quad (2.8)$$

Da mesma forma que o sistema de *DuSK* (Equação 2.4), as primeiras  $n(\mathbf{u})$  equações no sistema de *OK* (Equação 2.8) podem ser vistas como condições para a identificação dos dados. Lembre-se que para um dado conjunto de  $n(\mathbf{u})$  dados, os pesos  $d^{OK}_\beta(\mathbf{u})$  e estimativa média  $m^*_{OK}(\mathbf{u})$  não dependem de  $\mathbf{u}$ .

A Equação 2.8 é deduzida a partir da propriedade de que a estimativa por OK  $z^*_{OK}(\mathbf{u})$  identifica a estimativa da tendência  $m^*_{OK}(\mathbf{u})$  sempre que todos os  $n(\mathbf{u})$  dados são igualmente correlacionados com o valor desconhecido (Zhu, 1992). Nestes casos, todas as covariâncias  $C(\mathbf{u}_\alpha - \mathbf{u})$  são iguais a uma constante  $q$ , e o estimador duplo expresso pela Equação 2.7 é reescrito conforme (Equação 2.9):

$$\left( \begin{array}{c} \left( \right) \\ \left( \right) \end{array} \right) \sum \left( \begin{array}{c} \left( \right) \\ \left( \right) \end{array} \right) \left( \begin{array}{c} \left( \right) \\ \left( \right) \end{array} \right) \quad (2.9)$$

o que leva à restrição de que a soma dos pesos de krigagem  $d^{OK}_\alpha(\mathbf{u})$  deve ser igual a zero. O formalismo da krigagem dupla também pode ser facilmente estendido para o caso em que o componente de tendência é modelado como uma combinação linear de funções de tendência (*dual kriging with trend*) (Journel & Rossi, 1989).

Uma vez que o mesmo conjunto de dados é utilizado para estimar  $z$  em dois locais diferentes  $\mathbf{u}$  e  $\mathbf{u}'$ , o sistema duplo permanece inalterado. Os dois conjuntos de pesos duplos e as duas estimativas de tendência são idênticos:  $d^{OK}_\alpha(\mathbf{u}) = d^{OK}_\alpha(\mathbf{u}')$ ,  $\alpha$  e  $m^*_{OK}(\mathbf{u}) = m^*_{OK}(\mathbf{u}')$ . Caso os mesmos  $n$  dados sejam utilizados para estimar o atributo  $z$  em todos os locais, ou seja,  $n(\mathbf{u}) = n$ ,  $\mathbf{u}$ , o estimador duplo apresentado pela Equação 2.7, torna-se igual a (Equação 2.10):

$$\left( \begin{array}{c} \left( \right) \\ \left( \right) \end{array} \right) \sum \left( \begin{array}{c} \left( \right) \\ \left( \right) \end{array} \right) \quad (2.10)$$

onde, os  $n$  pesos  $d^{OK}_\alpha$  e a estimativa da tendência  $m^*_{OK}$  são independentes da localização, portanto, precisam ser calculados somente uma única vez.

O estimador apresentado pela Equação 2.10 é considerado uma função determinística de  $\mathbf{u}$ . O valor estimado  $z^*_{OK}(\mathbf{u})$ , em qualquer ponto  $\mathbf{u}$ , é facilmente obtido através do cálculo dos  $n$  vetores  $\mathbf{h} = \mathbf{u}_\alpha - \mathbf{u}$ , e da inserção dos mesmos na Equação 2.10. Contudo, em casos de estimativas globais, deve-se ter cuidado, pois o modelo de covariância  $C(\mathbf{h})$  é raramente conhecido sobre grandes distâncias, e a solução de um único, mas grande sistema duplo (Equação 2.8) pode ser instável (Goovaerts, 1997).

## 2.4 Construção de *grids* estratigráficos através do mapeamento topológico de *wireframes*

Os processos de construção de *grids* estruturados, segundo Soni (2000), podem ser baseados em métodos de algébricos de interpolação (Gordon & Thiel, 1982; Soni, 1992) ou em equações diferenciais parciais envolvendo sistemas elípticos (Steger & Chausee, 1980; Thompson, 1987) e sistemas hiperbólicos (Brackbill, 1993). O presente estudo faz uso da primeira vertente metodológica, aplicando o formalismo da *DuK* para a geração de um *grid* estratigráfico, a partir do *wireframe* de um modelo geológico. A ideia básica do método consiste em estimar as coordenadas espaciais  $x, y, z$  dos vértices de referência que definem os blocos do *grid* estratigráfico, com base nas coordenadas dos pontos que compõem as superfícies do *wireframe*.

Segundo Sánchez *et al.* (2013), *grids* estruturados podem ser definidos através do mapeamento de um cubo unitário  $B$  sobre um domínio físico representado por uma região conectada  $\Omega \subset \mathbb{R}^3$ . Em termos topológicos, a projeção do domínio  $B$  sobre  $\Omega$  pode ser classificada como um homeomorfismo. Por definição, homeomorfismo consiste em uma função contínua entre espaços topológicos que possui uma função inversa também contínua. Na categoria dos espaços topológicos, os homeomorfismos consistem em isomorfismos, ou seja, eles são mapeamentos que preservam todas as propriedades topológicas de um determinado espaço. Em termos mais simples, um homeomorfismo pode ser definido como um alongamento contínuo de um objeto (espaço topológico), em uma nova forma contínua.

Desta forma, o processo de construção de um *grid* estratigráfico 3D  $G_\Omega$  pode ser tratado como a definição das funções contínuas  $x(\xi, \eta, \zeta)$ ,  $y(\xi, \eta, \zeta)$  e  $z(\xi, \eta, \zeta)$ , que referem-se às coordenadas  $x, y, z$  do sistema Cartesiano, através do homeomorfismo  $\mathbf{X}$  (Equação 2.11) (Sánchez *et al.*, 2009):

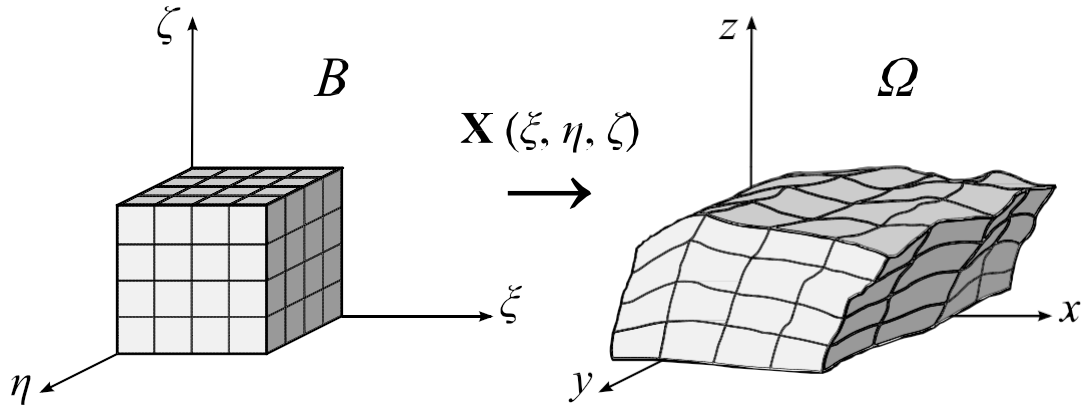
$$\text{com } \left( \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{X} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} \right) \quad (2.11)$$

a partir de um domínio contínuo  $B$ , representado por um *grid* Cartesiano  $G_B$ , com dimensões em número de nós ( $n_{cp}$ ) iguais a  $n_{cp} \times n_{cp} \times n_{cp}$ , origem em  $[0, 0, 0]$ , definido em um sistema de coordenadas curvilineares  $\xi, \eta, \zeta$  (Equação 2.12):



$$\{( \quad ) | \quad \} \quad (2.12)$$

sobre um domínio físico  $\Omega$  (*wireframe* do modelo geológico) (Figura 2.4).



**Figura 2.4** – mapeamento do *grid* Cartesiano  $G_B$  sobre o *grid* estratigráfico  $G_\Omega$ .

Podemos definir as fronteiras, ou limites, dos domínios  $B$  e  $\Omega$ , como a união dos conjuntos das superfícies  $B_i$  e  $\Omega_i$ ,  $i = 1, \dots, 6$ , respectivamente (Equação 2.13):

$$\cup \quad \cup \quad (2.13)$$

Este processo de mapeamento induz a uma decomposição natural de  $\partial\Omega$  em seis superfícies  $\Omega_i$ , uma vez que cada face de  $B$  é mapeada sobre uma superfície de referência do domínio físico  $\Omega$ . Assim, temos que (Equação 2.14):

$$( \quad ) \quad (2.14)$$

onde,

$$( \quad ) \quad (2.15)$$

logo, o mapeamento de  $B_i$  induz a construção de uma malha sobre a superfície  $\Omega_i$ , conforme a Equação 2.16:

$$| \quad (2.16)$$

Na realidade, os sistemas de coordenadas curvilineares, assim como o sistema Cartesiano, são todos ortogonais, e podem ser utilizados para mapear regiões com geometria curvilinear, da mesma maneira que um *grid* Cartesiano é utilizado para mapear um hexaedro regular. A principal característica que distingue esses sistemas é o fato de que os eixos das coordenadas curvilineares, normalmente, coincidem com a geometria de  $\partial\Omega$ . Portanto, os sistemas curvilineares podem ser considerados logicamente retangulares, e do ponto de vista computacional, não diferem conceitualmente do sistema Cartesiano (Thompson, Soni & Weatherill, 1999).

Contudo, para que o *grid* estratigráfico  $G_\Omega$  possa ser construído, é necessário, inicialmente, que exista um domínio  $B$  definido no sistema de coordenadas  $\xi, \eta, \zeta$ . Desta forma, torna-se possível aplicar o homeomorfismo expresso pela Equação 2.11, para determinar as coordenadas  $x, y, z$  dos vértices de referência dos blocos irregulares que compõem  $G_\Omega$ . Isto implica em um mapeamento através de um homeomorfismo inverso  $\mathbf{X}^{-1}$ , realizado a partir das superfícies de referência da envoltória do corpo geológico  $\Omega$ , sobre a superfície do domínio topológico  $B$ .

Desta forma, o processo de mapeamento das superfícies do *wireframe* se resume na definição de funções contínuas  $\xi(x, y, z)$ ,  $\eta(x, y, z)$  e  $\zeta(x, y, z)$ , que representam o conjunto de coordenadas curvilineares, através de uma função de mapeamento inversa  $\mathbf{X}^{-1}$  (Equação 2.17):

$$\text{com } \left( \begin{matrix} \xi \\ \eta \\ \zeta \end{matrix} \right) = \left( \begin{matrix} \xi \\ \eta \\ \zeta \end{matrix} \right) \left( \begin{matrix} x \\ y \\ z \end{matrix} \right) \quad (2.17)$$

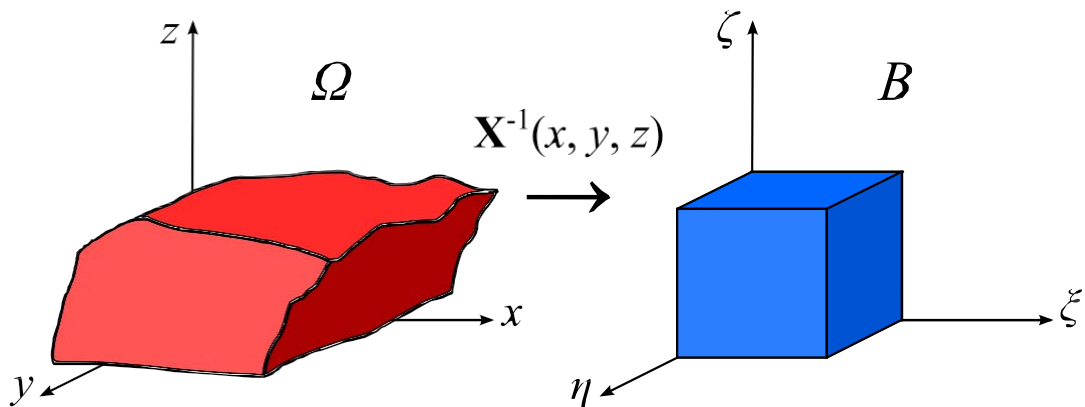
a partir de um domínio contínuo  $\Omega$ , definido em um sistema de coordenadas  $x, y, z$ , que neste caso, é constituído pelos pontos extraídos das superfícies do sólido geológico (Equação 2.18):

$$\left\{ \left( \begin{matrix} x \\ y \\ z \end{matrix} \right) \mid \left( \begin{matrix} \xi \\ \eta \\ \zeta \end{matrix} \right) \in \Omega \right\} \quad (2.18)$$

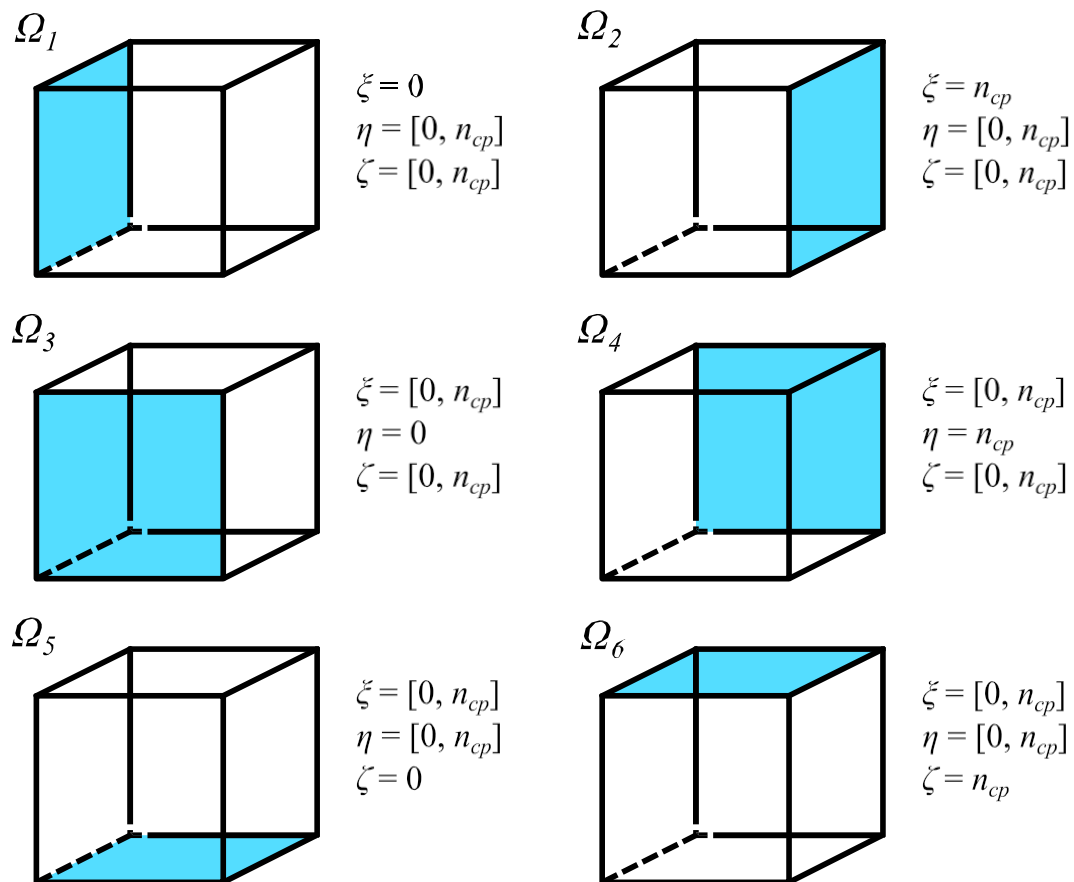
sobre um domínio contínuo  $B$  (Figura 2.5).

A partir do exposto acima, a construção de um *grid* estratigráfico conformado a uma região  $\Omega$ , torna-se um problema de valores limite. Neste caso, os valores internos das coordenadas curvilineares são gerados com base em valores específicos localizados nas fronteiras, ou superfícies de referência, do domínio  $\Omega$  (Thompson, Soni & Weatherill, 1999). Note que, temos como problema, a estimativa dos valores internos das coordenadas  $\xi, \eta, \zeta$  a

partir de valores constantes específicos (condições de contorno), impostos pelas superfícies opostas  $\Omega_i$  (Figura 2.6). Conforme a Figura 2.6, os valores de  $\xi$ ,  $\eta$ ,  $\zeta$  devem variar monotonicamente e dentro de um mesmo intervalo, sobre as superfícies nas quais não estão especificados. Assim, entre os pares de superfícies  $\Omega_1$  e  $\Omega_2$ ,  $\Omega_3$  e  $\Omega_4$ , e  $\Omega_5$  e  $\Omega_6$ , as variáveis  $\xi$ ,  $\eta$ ,  $\zeta$  podem assumir valores entre 0 e  $n_{cp}$ , respectivamente.



**Figura 2.5** – mapeamento das superfícies de referência do domínio físico  $\Omega$  sobre o domínio  $B$ .



**Figura 2.6** – condições de contorno aplicadas no mapeamento das superfícies de referência  $\Omega_i$ .

Este problema também pode ser visto de uma outra perspectiva, ou seja, na forma de um problema de valores limite no sistema de coordenadas  $\xi, \eta, \zeta$ , onde as variáveis a serem estimadas são as coordenadas  $x, y, z$ . Esta abordagem é mais simples, pois ela aproveita o fato de  $B$  ter a geometria de um hexaedro regular, o que possibilita a discretização do domínio por um *grid* Cartesiano. Desta forma, as coordenadas  $\xi, \eta, \zeta$  dos pontos que encontram-se na porção interna de  $B$  são automaticamente determinadas pelo arranjo dos blocos que compõem o *grid* Cartesiano  $G_B$ . As coordenadas curvilineares  $\xi, \eta, \zeta$  passam a informar a localização espacial da projeção dos pontos  $\mathbf{u}_{j}^{\Omega_i}, j = 1, \dots, J_{\Omega_i}$ , extraídos do *wireframe*, sobre a face correspondente  $B_i$  no hexaedro regular  $B$ , onde  $J_{\Omega_i}$  refere-se ao número total de pontos encontrados na superfície  $\Omega_i$ . Esse processo permite a transformação das coordenadas espaciais dos pontos pertencentes a  $\partial\Omega$  para o sistema de coordenadas  $\xi, \eta, \zeta$ , no qual o domínio  $B$  encontra-se definido.

A metodologia utilizada para realizar esse procedimento é apresentada pelo Algoritmo 2.1. O primeiro passo do método consiste na definição das superfícies de referência do domínio físico  $\Omega$ , as quais são representadas por  $\Omega_i, i = 1, \dots, 6$ . Para isso, é necessário que todas as superfícies adjacentes  $\Omega_i$  compartilhem os pontos  $\mathbf{u}^{\Omega_i}$  localizados em suas extremidades, de tal forma que definam um domínio topológico contínuo. Outro parâmetro exigido, é o valor da constante  $n_{cp}$ , que corresponde à resolução de  $G_B$  com relação aos eixos  $\xi, \eta, \zeta$ . Após a identificação de cada superfície, o algoritmo realiza a coleta das coordenadas  $x, y, z$  dos pontos pertencentes as mesmas. Em seguida, é aplicada a transformação expressa pela Equação 2.17, que resulta na definição das coordenadas estratigráficas  $\xi, \eta, \zeta$ . Este processo representa a projeção dos pontos  $\mathbf{u}^{\Omega_i}$  sobre as faces correspondentes do domínio  $B$ . Isso possibilita o mapeamento topológico de cada superfície, através da identificação da localização do conjunto  $\mathbf{u}^{\Omega_i}$  neste novo domínio. Por fim, todas as coordenadas dos pontos são armazenadas em um novo banco de dados definido no sistema de coordenadas  $\xi, \eta, \zeta$ .

Após a obtenção do conjunto de dados  $\mathbf{u}^{\hat{B}}$  no sistema  $\xi, \eta, \zeta$ , realiza-se a interpolação dos valores das coordenadas  $x, y, z$  na porção interna do domínio  $B$ . Para isso,  $B$  é discretizado na forma de um *grid* Cartesiano  $G_B$ , onde os valores das funções  $x(\xi, \eta, \zeta), y(\xi, \eta, \zeta)$  e  $z(\xi, \eta, \zeta)$  são interpolados por *dual kriging* com base nos valores de variáveis  $x, y, z$  pertencentes ao conjunto de dados  $\mathbf{u}_{k}^{\hat{B}}, k = 1, \dots, \Sigma$ .

---

**Algoritmo 2.1 – mapeamento topológico das superfícies de referência do *wireframe***


---

- 01:** Divida o *wireframe* da envoltória do corpo geológico  $\Omega$  em 6 superfícies de referência  $\Omega_i, i = 1, \dots, 6$
- 02:** Defina o valor do parâmetro de resolução  $n_{cp}$
- 03:** **para** cada superfície de referência  $\Omega_i$  **faça**
- 04:**     **para** cada ponto  $\mathbf{u}^{\Omega_i}_j, j = 1, \dots, J_{\Omega_i}$  que compõe  $\Omega_i$  **faça**
- 05:**         Extraia os valores das coordenadas  $x(\mathbf{u}^{\Omega_i}_j), y(\mathbf{u}^{\Omega_i}_j), z(\mathbf{u}^{\Omega_i}_j)$
- 06:**         Calcule os valores das coordenadas  $\zeta(\mathbf{u}^{B_i}_j), \eta(\mathbf{u}^{B_i}_j), \zeta(\mathbf{u}^{B_i}_j)$ , que informam a localização espacial relativa de  $\mathbf{u}^{\Omega_i}_j$  no hexaedro regular  $B$ , através da sua projeção sobre a face  $B_i$ , definidos em um intervalo entre  $[0, n_{cp}]$
- 07:**         Adicione os valores de  $\zeta(\mathbf{u}^{B_i}_j), \eta(\mathbf{u}^{B_i}_j), \zeta(\mathbf{u}^{B_i}_j)$  ao conjunto de dados  $\zeta(\mathbf{u}^{\hat{B}_k}), \eta(\mathbf{u}^{\hat{B}_k}), \zeta(\mathbf{u}^{\hat{B}_k}), k = 1, \dots, \sum$ , respectivamente
- 08:**     **fim para**
- 09:** **fim para**
- 

O método desenvolvido para a construção do *grid* estratigráfico é apresentado pelo Algoritmo 2.2. A primeira parte do algoritmo realiza, inicialmente, a construção de um *grid* Cartesiano  $G_B$  com dimensões iguais a  $n_{cp} \times n_{cp} \times n_{cp}$ , tamanho de célula de  $1 \times 1 \times 1$  m, e origem em  $[0, 0, 0]$ , o qual é definido a partir das coordenadas  $\zeta, \eta, \zeta$  do conjunto de pontos  $\mathbf{u}^{\hat{B}_k}, k = 1, \dots, \sum$ . Em seguida, são realizadas as interpolações por *DuOK* dos nós  $g^{GB}_l, l = 1, \dots, n_{cp}^3$  de  $G_B$ . As estimativas são feitas condicionais ao conjunto de dados  $\mathbf{u}^{\hat{B}_k}, k = 1, \dots, \sum$ , fazendo uso de uma vizinhança de busca global. O resultado desse processo é a geração das variáveis  $x^*(g^{GB}_l), y^*(g^{GB}_l)$  e  $z^*(g^{GB}_l)$ , as quais, futuramente, consistirão nas coordenadas dos vértices estruturais das células do *grid* estratigráfico  $G_\Omega$ .

A segunda parte do algoritmo dedica-se, exclusivamente, à construção de  $G_\Omega$ . Inicialmente, a partir das dimensões informadas para  $G_B$ , o algoritmo organiza a disposição e define o número total de blocos de  $G_\Omega$  no sistema de coordenadas  $x, y, z$ . Para cada nó  $g^{G_\Omega}_m, m = 1, \dots, (n_{cp} - 1)^3$  pertencente a  $G_\Omega$ , as coordenadas dos vértices estruturais (*corner points*) de cada bloco irregular são definidas com base nas estimativas de  $x^*(g^{GB}_l), y^*(g^{GB}_l), z^*(g^{GB}_l)$ . Desta forma, assim que os oito vértices de referência pertencentes a cada nó  $g^{G_\Omega}_m$  são informados, realiza-se a construção de um bloco irregular  $v^{G_\Omega}_m$  que compõe  $G_\Omega$ .

Cabe salientar que o modelo resultante  $G_\Omega$  sempre irá possuir um número total de blocos inferior ao *grid*  $G_B$ . Em contraste aos blocos regulares, cada bloco de  $G_\Omega$  requer oito vértices de referência para ser construído. Desta forma, cada nó interpolado em  $G_B$  irá constituir um dos vértices de referência (*corner points*) dos hexaedros irregulares de  $G_\Omega$ .

Portanto, o número total de blocos em um *grid* estruturado cúbico, definido no domínio  $\Omega$ , é expresso por  $N_{G\Omega}$  (Equação 2.19):

$$( \quad ) \quad (2.19)$$

onde,  $n_{cp}$  é o número de blocos com relação aos eixos  $\xi, \eta, \zeta$  de  $G_B$ . No entanto, caso  $G_B$  não seja um cubo, e possua dimensões distintas com relação aos eixos  $\xi, \eta, \zeta$ , temos que o número total de células  $N_{G\Omega}$  é dado por (Equação 2.20):

$$( \quad )( \quad )( \quad ) \quad (2.20)$$

onde,  $\xi_{cp}, \eta_{cp}, \zeta_{cp}$  são iguais às dimensões, em número de nós, com relação aos eixos  $\xi, \eta, \zeta$  de  $G_B$ .

A configuração da distribuição do número de nós é um aspecto fundamental no processo de geração de qualquer malha, pois define a resolução do modelo geológico final. Em algumas situações, por exemplo, a representação da variabilidade vertical de determinada propriedade é crucial, logo, a definição de um número maior de nós com relação ao eixo vertical é obrigatória.

---

### Algoritmo 2.2 – construção do *grid* estratigráfico

---

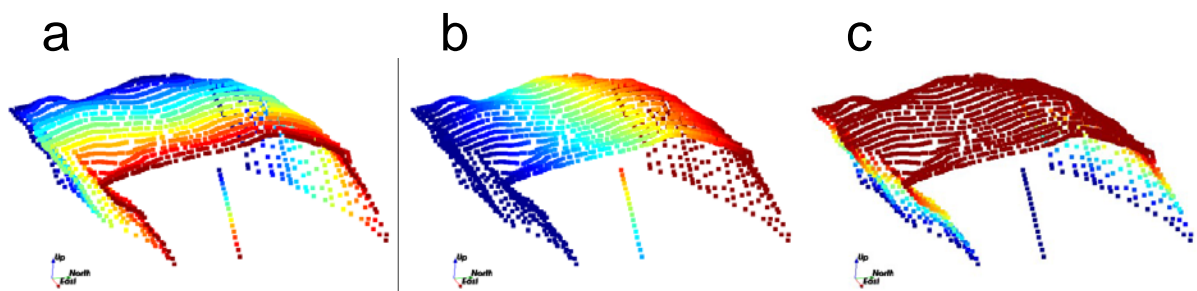
- 01:** Construa um *grid* Cartesiano  $G_B$  com dimensões iguais a  $[n_{cp} \times n_{cp} \times n_{cp}]$ , tamanho de célula igual a  $[1 \times 1 \times 1 \text{ m}]$  e origem em  $[0, 0, 0]$  definido no sistema de coordenadas  $\xi, \eta, \zeta$
  - 02:** **para** cada nó  $g^{GB}_l, l = 1, \dots, n_{cp}^3$  pertencente a  $G_B$  **faça**
  - 03:** Realize as interpolações dos valores de  $x^*(g^{GB}_l), y^*(g^{GB}_l), z^*(g^{GB}_l)$  por *DuOK* condicionais ao conjunto de dados  $x(\mathbf{u}^{\partial B}_k), y(\mathbf{u}^{\partial B}_k), z(\mathbf{u}^{\partial B}_k), k = 1, \dots, \Sigma$ , respectivamente
  - 04:** **fim para**
  - 05:** Baseando-se em um *grid* estratigráfico  $G_\Omega$  com dimensões iguais a  $[n_{cp} - 1 \times n_{cp} - 1 \times n_{cp} - 1]$  definido no sistema de coordenadas  $x, y, z$
  - 06:** **para** cada nó  $g^{G\Omega}_m, m = 1, \dots, (n_{cp} - 1)^3$  pertencente a  $G_\Omega$  **faça**
  - 07:** **para** cada vértice  $v_p(g^{G\Omega}_m), p = 1, \dots, 8$  de  $g^{G\Omega}_m$  **faça**
  - 08:** Defina como coordenadas  $x(v_p(g^{G\Omega}_m)), y(v_p(g^{G\Omega}_m)), z(v_p(g^{G\Omega}_m))$  os valores de  $x^*(g^{GB}_l), y^*(g^{GB}_l), z^*(g^{GB}_l)$ , respectivamente
  - 09:** **fim para**
  - 10:** Construa o bloco irregular  $v^{G\Omega}_m$  unindo o conjunto de vértices  $v_1(g^{G\Omega}_m), \dots, v_8(g^{G\Omega}_m)$
  - 11:** **fim para**
-

## 2.5 Estudo de caso

A metodologia exposta na Seção 2.4 foi utilizada para construir um *grid* estratigráfico com base no *wireframe* da envoltória do corpo de geológico de um depósito de minério de ferro, localizado no estado de Minas Gerais. O banco de dados é apresentado em detalhe no estudo de caso do Capítulo 3.

O *wireframe* da envoltória do corpo de minério foi construído em um software de modelagem geológica a partir de 17 seções verticais com espaçamento de 50 m, compostas por um total de 2.865 pontos. O modelo tridimensional foi separado em seis superfícies de referência, consistindo nas superfícies de base, topo, *footwall*, *hangingwall*, e nas duas seções verticais localizadas nas extremidades do sólido geológico. Posteriormente, os dados foram importados para o *software SGeMS*, e o Algoritmo 2.1 foi aplicado sobre os conjuntos de pontos que compõem cada superfície, gerando três novas variáveis, as coordenadas curvilineares  $\xi$ ,  $\eta$ ,  $\zeta$ . O número de *corner points* escolhido para os três eixos de coordenadas foi de  $n_{cp} = 201$ .

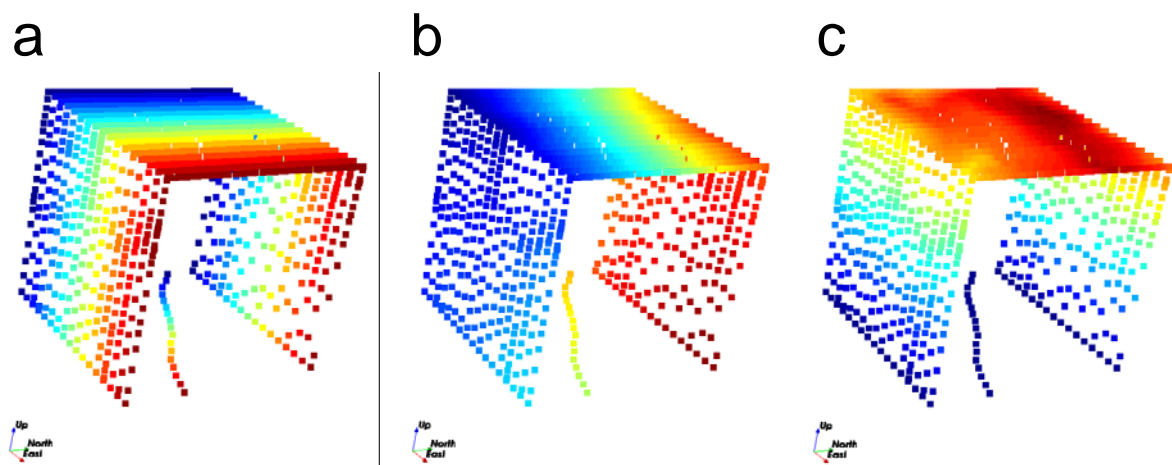
A Figura 2.7 ilustra o banco de dados composto pelos pontos extraídos do *wireframe* definido no sistema de coordenadas  $x$ ,  $y$ ,  $z$ , e as novas propriedades  $\xi$ ,  $\eta$ ,  $\zeta$ . Devido ao fato das duas seções verticais utilizadas para delimitar o sólido geológico consistirem em planos, elas apresentam pontos somente em seu contorno (Figura 2.7). O mesmo se aplica à superfície de base do *wireframe*, contudo, esta possui um conjunto adicional de pontos dispostos na forma de uma linha diagonal, localizada na porção central da superfície.



**Figura 2.7** – variáveis  $\xi$  (a),  $\eta$  (b),  $\zeta$  (c) definidas no sistema de coordenadas  $x$ ,  $y$ ,  $z$ . escala de cores: azul ( $\xi$ ,  $\eta$ ,  $\zeta = 0$ ) e vermelho ( $\xi$ ,  $\eta$ ,  $\zeta = 201$ ).

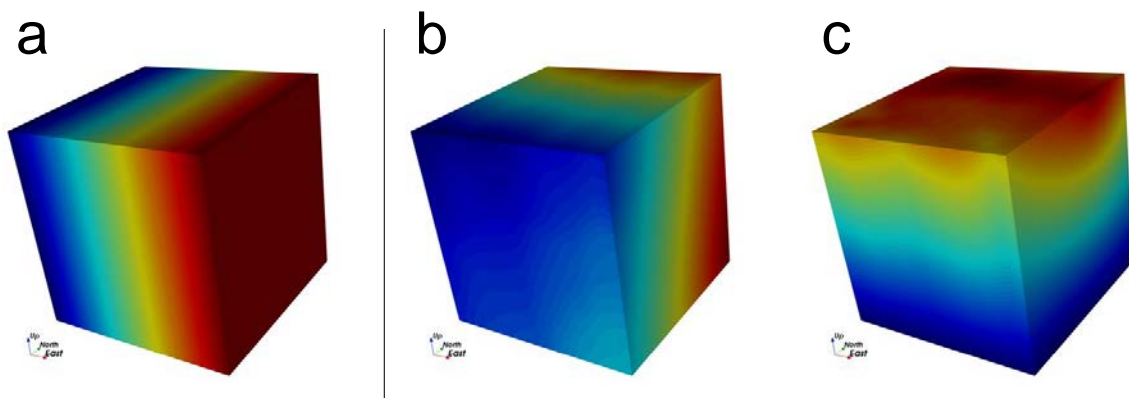
Em seguida, o banco de dados foi carregado novamente, mas no sistema de coordenadas  $\xi$ ,  $\eta$ ,  $\zeta$ . Note que nesse caso, há uma inversão de variáveis:  $\xi$ ,  $\eta$ ,  $\zeta$  tornam-se as coordenadas espaciais, e as variáveis  $x$ ,  $y$ ,  $z$ , as propriedades que devem ser interpoladas

(Figura 2.8). O Algoritmo 2.2 foi utilizado para a construção de um *grid* Cartesiano com dimensões iguais a  $201 \times 201 \times 201$  m, composto por um total de 8.120.601 blocos, com tamanho de  $1 \times 1 \times 1$  m, e origem em  $[0, 0, 0]$ , definido no sistema  $\zeta, \eta, \zeta$ . As variáveis  $y$  e  $z$  foram estimadas por *DuOK* com base nos conjuntos de pontos apresentados pela Figura 2.8b e Figura 2.8c, respectivamente. As estimativas adotaram uma vizinhança de busca global e modelos de semivariograma isotrópicos, compostos por apenas uma estrutura esférica. Neste caso, visto que todas as seções verticais são ortogonais ao eixo  $\zeta$ , e apresentam espaçamento regular (Figura 2.8a), optou-se por interpolar a variável  $x$  através de um modelo de regressão linear.



**Figura 2.8** – variáveis  $x$  (a),  $y$  (b),  $z$  (c) definidas no sistema de coordenadas  $\zeta, \eta, \zeta$ .

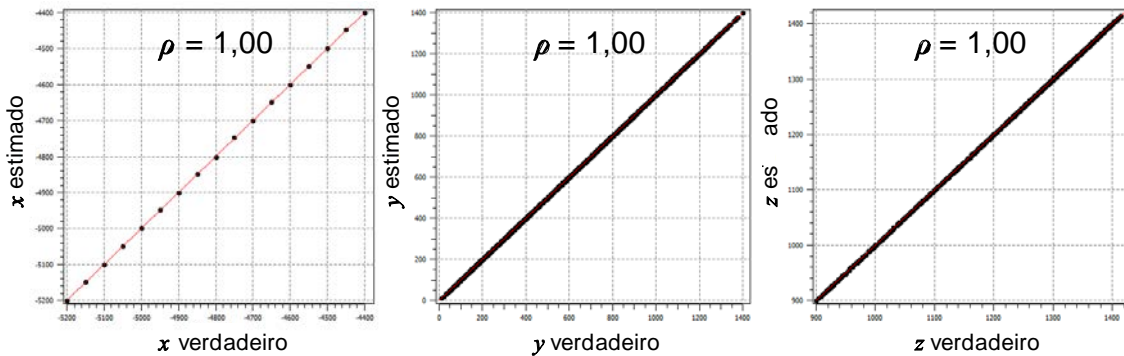
Os resultados das estimativas das variáveis  $x, y, z$  são apresentados na Figura 2.9. Verifica-se que as interpolações por *DuOK* não criaram nenhum artefato indesejado. Os modelos gerados apresentam mudanças bastante graduais entre os valores estimados, o que está de acordo com as variações das coordenadas espaciais dos pontos das superfícies do *wireframe* (Figura 2.8).



**Figura 2.9** – interpolação das variáveis  $x$  (a),  $y$  (b),  $z$  (c) por *DuOK* no *grid* Cartesiano  $G_B$ .

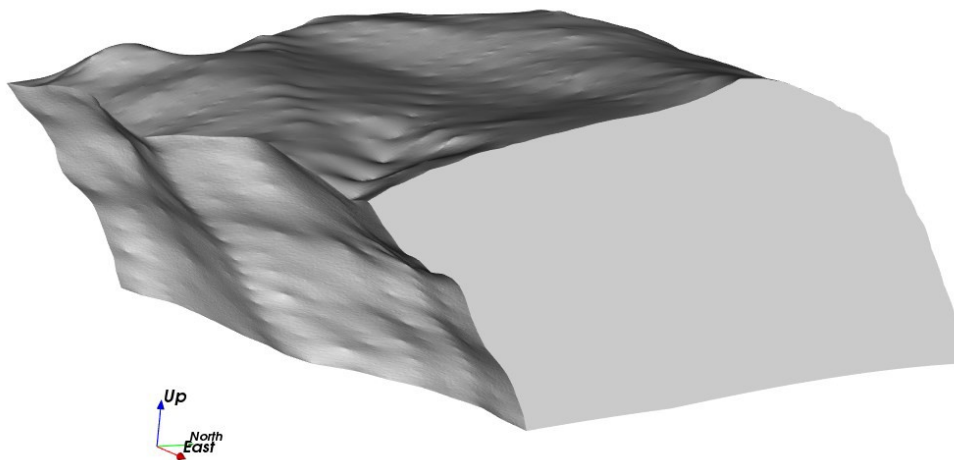


Os diagramas de dispersão apresentados na Figura 2.10 mostram que as interpolações por *DuOK* honraram exatamente os valores das coordenadas dos pontos extraídos do *wireframe*. Todas as retas de regressão possuem coeficientes de correlação linear  $\rho = 1$ .



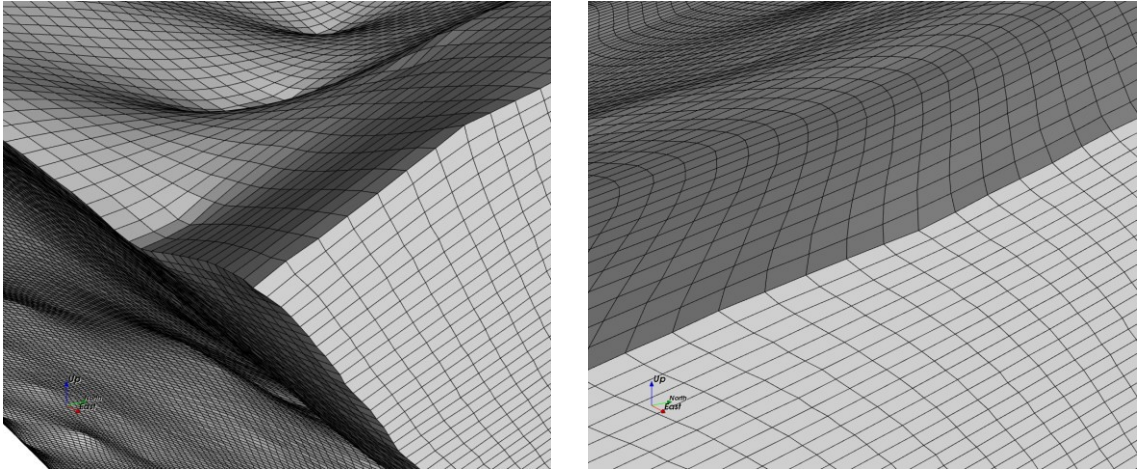
**Figura 2.10** – diagramas de dispersão entre os valores reais das coordenadas  $x$ ,  $y$ ,  $z$  dos pontos do *wireframe* e das estimativas por *DuOK*.

Por fim, o Algoritmo 2.2 realizou a importação das estimativas das propriedades  $x$ ,  $y$ ,  $z$  para o *SGeMS* utilizando a estrutura de dados chamada *structured grid*. Isto permite que o *software* identifique os vetores dessas variáveis, e os leia como as coordenadas dos vértices de referência que compõem aos blocos do *grid* estratigráfico. O *grid* estratigráfico resultante possui um total de 8 milhões de blocos, e dimensões com relação aos eixos  $x$ ,  $y$ ,  $z$  de  $200 \times 200 \times 200$  blocos, os quais apresentam geometrias irregulares e volumes distintos. O tamanho médio das células é de aproximadamente  $4 \times 4 \times 2$  m. A Figura 2.11 apresenta o modelo de blocos gerado, note que o mesmo se conforma de acordo com os limites da envoltória do corpo de minério, apresentada anteriormente na Figura 2.7.



**Figura 2.11** – *grid* estratigráfico gerado com base no *wireframe* da envoltória do corpo geológico.

A Figura 2.12 apresenta em detalhe a geometria dos blocos que compõem o *grid*. Verifica-se que, apesar do arranjo curvilíneo do modelo de blocos, as células encontram-se dispostas de forma regular, e estão alinhadas de acordo com as superfícies da envoltória definida na modelagem do *wireframe*.



**Figura 2.12** – detalhe da geometria irregular dos blocos que compõem o *grid* estratigráfico.

## Capítulo 3

# Geoestatística de múltiplos pontos aplicada à simulação de modelos geológicos

Este capítulo apresenta os principais fundamentos da geoestatística de múltiplos pontos, e a sua aplicação na simulação de modelos geológicos. Em seguida, propõe-se uma metodologia para a simulação de contatos litológicos a partir de modelos geológicos interpretados. A Seção 3.1 realiza uma breve introdução aos métodos de simulação de múltiplos pontos. A Seção 3.2 introduz os fundamentos e as propriedades básicas da simulação sequencial aplicada à variáveis categóricas. A Seção 3.3 apresenta, em detalhe, o algoritmo *SNESIM* e o conceito de partição da árvore de busca. A Seção 3.4 propõe um método de simulação de contatos litológicos através de uma adaptação do algoritmo *SNESIM*, utilizando o particionamento da árvore de busca. Por fim, a Seção 3.5 ilustra a aplicação do método na forma de um estudo de caso em um depósito de minério de ferro.

### 3.1 Introdução

Tradicionalmente, as duas principais alternativas existentes para a modelagem geoestatística de variáveis categóricas consistiam nos algoritmos de simulação baseados em *pixels* (*pixel-based*) e os baseados em objetos (*object-based*) (Liu, 2006). Os algoritmos baseados em *pixels* caracterizam-se por construir as realizações um *pixel* por vez, o que oferece grande capacidade de condicionamento, no entanto, compromete a sua velocidade de execução. Outra limitação imposta especificamente por algoritmos tradicionais baseados estatísticas bipontuais reside no fato de que os mesmos conseguem reproduzir somente o

modelo de semivariograma, ou covariância, falhando na reprodução de geometrias e padrões complexos (Remy *et al.*, 2009). Por outro lado, os algoritmos baseados em objetos constroem as realizações inserindo, no *grid* de simulação, um objeto ou uma determinada forma geométrica por vez (Chilès & Delfiner, 1999; Stoyan *et al.*, 1987; Haldorsen & Damsleth, 1990; Lantuéjoul, 2002; Mallet, 2002). Esta característica possibilita a reprodução de formas geológicas complexas, e os torna relativamente rápidos se comparados aos algoritmos baseados em *pixels*. Contudo, dificulta também a sua capacidade de condicionamento, particularmente quando os dados são numerosos, possuem pequeno suporte e são de diferentes tipos (Remy *et al.*, 2009).

Um aspecto fundamental dos métodos geoestatísticos é que não há como existir qualquer processo probabilístico de simulação, ou estimativa, sem que existam as estatísticas de múltiplos pontos necessárias conectando simultaneamente o conjunto de dados com os lugares desconhecidos (Journel, 1994). As estatísticas de múltiplos pontos podem ser derivadas explicitamente a partir de um modelo analítico multivariado, uma imagem de treinamento, ou fornecidas de forma implícita pelo próprio algoritmo de simulação.

Os algoritmos tradicionais, baseados em semivariogramas, adotam implicitamente as estatísticas de alta ordem embutidas no próprio método de simulação, as quais possuem, frequentemente, caráter de alta entropia. Modelos de máxima entropia caracterizam-se por maximizar o grau de desordem além do alcance dos modelos de semivariograma (Goovaerts, 1997; Journel & Alabert, 1989; Journel & Deutsch, 1993). Em certos casos, o uso de modelos de alta entropia em aplicações geocientíficas é questionável. Isto porque, sabe-se que na natureza, é comum a existência de estruturas complexas com geometria curvilínea, geralmente, com níveis mais baixos de entropia, envolvendo mais que dois pontos simultaneamente (Remy *et al.* 2009).

Considere, inicialmente, uma variável aleatória  $S$ , a qual pode assumir  $K$  classes  $\{s_k, k = 1, \dots, K\}$ , onde  $S$  representa diferentes categorias ou litotipos. Uma imagem de treinamento (*training image - TI*) consiste em uma representação de como os valores de uma determinada propriedade  $s_k$  estão conjuntamente distribuídos no espaço (Farmer, 1992; Strebelle, 2002; Journel, 2002; Zhang, 2006). Uma *TI* é, essencialmente, uma realização não condicional de um modelo de função aleatória  $S(\mathbf{u})$ , que consiste em uma representação conceitual prévia da distribuição espacial dos valores da variável  $s_k$ , e que não precisa necessariamente honrar a localização do conjunto de  $n(\mathbf{u})$  amostras de  $s_k$ ,  $k = 1, \dots, K$ . Assume-se que a distribuição

conjunta em um espaço  $A$  dos valores desconhecidos  $\{s_k(\mathbf{u}), s_k \in A\}$  é semelhante a da  $TI$ , no entanto, a mesma honra o grupo de  $n$  amostras.

De forma análoga aos algoritmos de simulação baseados em estatísticas de dois pontos, que têm como objetivo gerar realizações que honrem o conjunto de dados e um modelo de semivariograma, a simulação de múltiplos pontos visa gerar diferentes realizações adaptando a  $TI$  de tal forma que ela respeite todas as amostras condicionantes e as estruturas de múltiplos pontos presentes na mesma.

Segundo Jones *et al.* (2013), a utilização de imagens de treinamento para inferência estatística apresenta algumas vantagens sobre os métodos alternativos, dentre elas:

- As estatísticas de múltiplos pontos oferecerem substancialmente mais informação se comparadas às estatísticas espaciais de dois pontos;
- Não há um agrupamento ou a tomada de valores médios dos dados dispostos em conjuntos de *lags*;
- Não existe a necessidade de supor a existência de uma estrutura espacial Gaussiana, ou de um modelo de campo aleatório;
- Permitem que complexas relações geométricas entre unidades geológicas possam ser identificadas (Lyster, 2009).

Strebelle (2002) e Arpat (2005) classificam os métodos que fazem uso de informações derivadas de múltiplos pontos em quatro grupos principais:

- A simulação *annealing* (*simulated annealing*) tem sido utilizada para reproduzir algumas estatísticas de múltiplos pontos específicas previamente modeladas a partir de uma  $TI$  (Farmer, 1990; Deutsch, 1992). Nesta abordagem, estatísticas de múltiplos pontos de alta ordem são usadas como restrições explícitas que cada realização deve honrar através de uma função objetivo. Além do fato de que um pequeno número de estatísticas pode ser considerado simultaneamente, o método sofre de problemas de convergência relacionados à dificuldade de escolha de um conjunto adequado de parâmetros;
- Os campos aleatórios de Markov exigem uma especificação prévia sobre a distribuição de probabilidade de múltiplos pontos da variável simulada, ou alguma razão entre os valores de probabilidade condicional, dada a ocorrência de um evento de dados (*data event*) de múltiplos pontos (Tjelmeland, 1996). Apesar do grande apelo teórico da

técnica, o método é iterativo, demanda extremamente da CPU, e pode não convergir de forma satisfatória;

- Caers & Journel (1998) e Caers *et al.* (1999) propuseram um método que utiliza redes neurais artificiais (*artificial neural networks - ANN*) para modelar estatísticas de múltiplos pontos inferidas de uma *TI* e, posteriormente, simular valores. O algoritmo desenvolvido produz resultados relativamente bons, no entanto, possui natureza iterativa, ou seja, exige da CPU, e é sujeito a ter problemas de convergência;
- Outra abordagem geostatística de múltiplos pontos foi apresentada por Srivastava (1992) e Guardiano & Srivastava (1993). O método baseia-se em uma extensão do conceito da simulação sequencial. A ideia se fundamenta na inferência de probabilidades condicionais locais em cada nó não amostrado do *grid* de simulação, através da varredura de uma *TI* em busca de réplicas de um determinado *data event*. O nó do *grid* é então simulado utilizando esta *cpmf* (*conditional probability mass function*), e é posteriormente, tratado como dado condicionante. Como é permitido que haja variações na configuração dos dados, o processo de simulação não sofre de problemas de convergência, característicos dos algoritmos iterativos.

O conceito de simulação de múltiplos pontos proposto por Guardiano & Srivastava (1993) visava combinar a capacidade de condicionamento oferecida pelos algoritmos baseados em *pixels*, com o potencial dos algoritmos baseados em objetos em reproduzir formas geológicas complexas. O método caracterizava-se por operar de maneira sequencial, onde a inferência das probabilidades condicionais dos valores simulados era realizada através de proporções condicionais obtidas a partir de uma *TI*.

No entanto, a implementação original do método proposto por Guardiano & Srivastava (1993) demandava extremamente da CPU, pois a varredura da *TI* precisava ser feita para cada nó não amostrado. A aplicação prática do método só foi possível depois do trabalho de Strebelle (2000), onde o autor propôs o uso de uma estrutura de dados especial chamada *search tree* (árvore de busca).

### 3.2 Simulação sequencial

O amplo grupo de algoritmos de simulação conhecidos pelo nome genérico de algoritmos de simulação sequencial (*sequential simulation algorithms*), ao invés de modelarem a *cpmf* de  $N$  pontos, estes métodos realizam a modelagem de uma distribuição condicional em apenas um ponto, e em seguida, extraem um valor simulado, visitando todos os  $N$  nós do *grid* aleatoriamente. Para garantir a reprodução do modelo de covariância de determinada variável categórica  $s_k$ ,  $k = 1, \dots, K$ , cada *cpmf* referente a determinado ponto é feita condicional não somente aos  $n$  dados originais, mas também, a todos os valores simulados em locais previamente visitados.

A implementação da simulação sequencial consiste na reprodução de variáveis regionalizadas por meio da utilização sequencial de distribuições condicionais. Considere um conjunto de  $N$  variáveis aleatórias  $S(\mathbf{u}_\alpha)$ ,  $\alpha = 1, \dots, N$  definido em  $N$  locais  $\mathbf{u}_\alpha$ . O objetivo do método é gerar  $L$  realizações conjuntas  $\{s_k^{(l)}(\mathbf{u}_\alpha), \alpha = 1, \dots, N\}$ , com  $l = 1, \dots, L$ , de  $N$  variáveis aleatórias, condicionais a um conjunto de  $n$  dados, e que reproduzem as propriedades de uma dada distribuição multivariada.

Considere uma propriedade  $S$ , a qual pode assumir  $K$  classes  $\{s_k, k = 1, \dots, K\}$ , onde  $S$  representa diferentes categorias ou litotipos, o valor desconhecido no local  $\mathbf{u}_\alpha$  é associado a um conjunto de  $K$  indicadores definidos por  $I_k(\mathbf{u}) = 1$  se  $S(\mathbf{u}) = s_k$ , do contrário,  $I_k(\mathbf{u}) = 0$ . Um mapa dos  $N$  indicadores da  $k$ -ésima categoria pode ser representado por um vetor  $\mathbf{i}_k = [i_k(\mathbf{u}_\alpha), \alpha = 1, \dots, N]$ . Este vetor de indicadores é considerado uma realização conjunta de  $N$  variáveis aleatórias  $I_k(\mathbf{u}_1), \dots, I_k(\mathbf{u}_N)$ . Em teoria, o processo de simulação sequencial consiste na geração de categorias a partir de uma *pmf* (*probability mass function*)  $N$ -variada dos  $N$  indicadores das variáveis aleatórias, condicionais a um conjunto de  $n$  amostras, conforme a Equação 3.1:

$$P(\mathbf{I} | \mathbf{D}) = \prod_{\alpha=1}^N P(I_\alpha | \mathbf{D}, \mathbf{I}_{-\alpha}) \quad (3.1)$$

No entanto, a *cpmf*  $N$ -variada expressa pela Equação 3.1 é analiticamente intratável, o que a torna difícil de ser amostrada. Uma solução prática é gerar amostras desta *pmf* através da aplicação recursiva da regra de Bayes (Johnson, 1987). Na geoestatística, a aplicação recursiva desta regra é conhecida como simulação sequencial dos indicadores (Alabert, 1987; Journel, 1989; Alabert & Massonnat, 1990). Desta forma, a distribuição multivariada

composta de  $N$  pontos é decomposta em um conjunto de  $N$  distribuições condicionais univariadas (*cpmfs*), conforme a Equação 3.2:

$$\begin{aligned}
 (i_k | \mathbf{i}_{k-1}) &= \{ (i_k | i_{k-1}) | (i_k | \mathbf{i}_{k-1}) \} \\
 &= \{ (i_k | i_{k-1}) | (i_k | i_{k-2}, i_{k-1}) \} \\
 &= \{ (i_k | i_{k-1}, i_{k-2}, \dots, i_1) | (i_k | \mathbf{i}_{k-1}) \} \\
 &= \{ (i_k | \mathbf{i}_{k-1}) | (i_k | \mathbf{i}_{k-1}) \} \\
 &= (i_k | \mathbf{i}_{k-1})
 \end{aligned} \tag{3.2}$$

A simulação com base na *cpmf*  $f(\mathbf{i}_k | n)$  torna-se significativamente mais simples, visto que agora basta gerar sequencialmente  $N$  valores categóricos a partir das  $N$  *cpmfs* locais, ao invés de ter que gerar todos os  $N$  valores de uma só vez a partir de  $f(\mathbf{i}_k | n)$ . No processo de simulação sequencial, não existe nenhuma restrição com relação à sequência de decomposição da Equação 3.2. Esta sequência geralmente consiste no caminho aleatório da simulação. Desta forma, as amostras e todos os nós previamente simulados são utilizados para condicionar a determinação da *cpmf* local em qualquer nó ainda não visitado ao longo desse caminho aleatório (Deutsch & Journal, 1998).

A medida que a simulação avança, o processo de condicionamento envolve um número cada vez maior de nós previamente simulados, o que o torna mais complexo. Este problema é contornado através da definição de uma vizinhança de busca  $W(\mathbf{u}_\alpha)$  centrada em  $\mathbf{u}_\alpha$  utilizada para selecionar somente os dados mais importantes para a inferência da *cpmf* local. Desta forma, temos que  $A_{(\alpha-1)}$  denota o conjunto de nós previamente simulados antes do  $n$ -ésimo passo da decomposição da simulação sequencial expressa por Equação 3.2, e  $A_{(0)}$  corresponde ao conjunto inicial de dados condicionantes, ou seja, o grupo de  $n$  dados amostrais. Ao considerarmos que  $A^{W(\mathbf{u}_\alpha)}_{(\alpha-1)} \in A_{(\alpha-1)}$  é o subconjunto de nós previamente retidos na vizinhança  $W(\mathbf{u}_\alpha)$ , temos que a Equação 3.2 é reescrita conforme (Equação 3.3):

$$(i_k | \mathbf{i}_{k-1}) = \prod_{(\mathbf{i}_{k-1}) \in A^{W(\mathbf{u}_\alpha)}_{(\alpha-1)}} \{ (i_k | \mathbf{i}_{k-1}) | (i_k | \mathbf{i}_{k-1}) \} \tag{3.3}$$

Geralmente, a vizinhança de busca  $W(\mathbf{u}_\alpha)$  consiste em um *template*, ou um elipsoide, definido por um conjunto de  $N_W$  nós, onde, usualmente,  $N_W \ll N$ . O valor de  $N_W$  corresponde



ao número máximo de valores previamente simulados que podem ser retidos em  $W(\mathbf{u}_\alpha)$  para calcular a *cpmf* local  $Prob\{I_k(\mathbf{u}_\alpha) = 1 | A^{W(\mathbf{u}_\alpha)}_{(\alpha-1)}\}$  de cada categoria no local  $\mathbf{u}_\alpha$ . Em seguida, um valor simulado é gerado através das  $K$  *cpmfs* locais em  $\mathbf{u}_\alpha$ , e é posteriormente utilizado para o condicionamento dos próximos nós a serem visitados ao longo do caminho aleatório da simulação. Outra realização pode ser gerada se alterarmos o caminho aleatório, e utilizarmos um outro grupo de números aleatórios para fazer a tiragem das *cpmfs* locais.

Assim, de uma forma generalizada, o processo de simulação sequencial aplicado a atributos categóricos segue a seguinte sequência de etapas:

- i. Defina o conjunto inicial de  $n$  dados condicionantes  $A_{(0)} = s_k(\mathbf{u})$ ,  $k = 1, \dots, K$ ;
- ii. Defina um caminho aleatório visitando os  $N$  nós não informados;
- iii. Para cada nó  $\mathbf{u}_\alpha$ ,  $\alpha = 1, \dots, N$  faça;
  - a. Obtenha a vizinhança de dados condicionantes  $A^{W(\mathbf{u}_\alpha)}_{(\alpha-1)}$ ;
  - b. Estime a  $Prob\{I_k(\mathbf{u}_\alpha) = 1 | A^{W(\mathbf{u}_\alpha)}_{(\alpha-1)}\}$ , ou seja, a *cpmf* local para cada categoria;
  - c. Tire um valor simulado  $s_k^{(l)}(\mathbf{u}_\alpha)$  de  $Prob\{I_k(\mathbf{u}_\alpha) = 1 | A^{W(\mathbf{u}_\alpha)}_{(\alpha-1)}\}$ ;
  - d. Adicione  $s_k^{(l)}(\mathbf{u}_\alpha)$  ao conjunto de dados condicionantes  $A_{(\alpha)} = A_{(\alpha-1)} \cup s_k^{(l)}(\mathbf{u}_\alpha)$ ;
- iv. Continue até todos os nós do *grid* terem sido simulados;

O principal aspecto que diferencia os algoritmos de simulação sequencial existentes é a maneira em que se obtém, e se realiza, a tiragem das distribuições condicionais. O tipo e forma da *cpmf* depende diretamente dos dados condicionantes e das propriedades estatísticas desejáveis para uma determinada realização (Arpat, 2005).

### 3.3 O algoritmo *SNESIM*

O algoritmo *SNESIM* (*Single Normal Equation Simulation*) (Strebelle, 2000; Strebelle, 2002) baseia-se no paradigma da simulação sequencial, onde cada valor simulado se torna um dado condicionante para as simulações dos nós a serem visitados posteriormente (Goovaerts, 1997). O algoritmo recebeu esse nome devido ao fato dele utilizar apenas uma única equação normal para modelar a probabilidade de determinada categoria em um nó particular do *grid* de simulação. Esta equação normal se refere a relação de Bayes que define uma probabilidade condicional. Journel (1992) demonstrou a conexão entre a geoestatística de múltiplos pontos e a avaliação estendida de probabilidades através de um sistema estendido de equações normais de krigagem. Ao invés de modelar as estatísticas de múltiplos pontos a partir de algumas

estatísticas de ordem inferior, a probabilidade de múltiplos pontos é identificada às proporções experimentais correspondentes, lidas em uma imagem de treinamento. Assim, o método elimina a necessidade de resolver um sistema completo de krigagem, em vez disso, ele obtém as probabilidades diretamente por meio de uma única equação normal (*single normal equation*) que equivale à identificação de uma proporção.

O algoritmo realiza a varredura da imagem de treinamento por meio de um *template* de busca pré-definido, utilizado para extrair os eventos de dados (*data events*) existentes na *TI*. Para cada evento de dados, o algoritmo procura por réplicas daquele evento e então recupera o histograma referente ao valor central do evento. Uma vez que os eventos de dados e as probabilidades associadas aos seus valores centrais são coletados a partir da *TI*, o algoritmo os armazena em uma estrutura de dados dinâmica, chamada *search tree* (árvore de busca) (Roberts, 1998). Esta abordagem vai de encontro a proposta original de Srivastava (1992), que exigia uma nova varredura da *TI* para cada novo evento de dados encontrado.

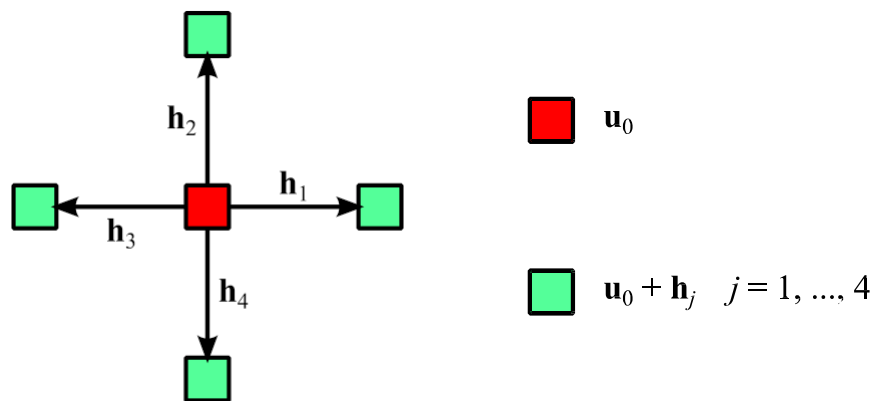
Devido ao fato que os eventos de dados de condicionamento local incluem nós previamente simulados, e os nós a serem visitados ao longo de um caminho aleatório, a geometria deste evento de dados é modificada em cada nó visitado. Originalmente, Guardiano & Srivastava (1993) haviam proposto a varredura completa da *TI* em cada nó não amostrado para inferir a distribuição de probabilidade condicional específica aos dados informando aquele nó. Esta varredura repetitiva demandava em excesso a CPU, especialmente quando se consideravam *TIs* com grandes dimensões, ou quando se gerava um número elevado de realizações em *grids* com um grande número de nós (Strebelle, 2002).

O algoritmo desenvolvido por Strebelle (2000) pode ser dividido em duas partes principais: i) a construção da árvore de busca, onde as proporções de todos os padrões encontrados na *TI* são armazenadas. Mais precisamente, o algoritmo armazena na árvore de busca somente os números de ocorrência dos eventos de dados e os valores centrais encontrados na *TI*, dos quais as proporções são calculadas. A construção da árvore de busca requer a varredura da *TI* somente uma única vez, antes de iniciar o processo de simulação; ii) a segunda parte refere-se ao processo de simulação em si, que segue a estrutura dos demais algoritmos de simulação sequencial, onde é realizada a leitura dessas proporções, para que, posteriormente, as mesmas sejam utilizadas na tiragem de valores simulados.

As seções seguintes apresentam uma descrição detalhada do conjunto de etapas e propriedades do algoritmo *SNESIM*.

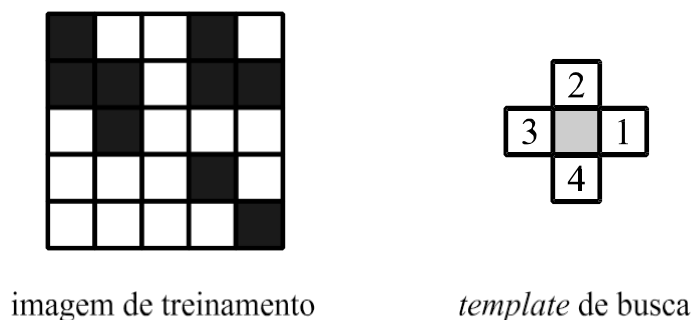
### 3.3.1 Construção da árvore de busca

A construção de uma árvore de busca é realizada após a definição de um *template* de busca  $\tau_J$ , o qual é composto por um conjunto de  $J$  vetores  $\{\mathbf{h}_j, j = 1, \dots, J\}$  que se irradiam a partir de um nó central  $\mathbf{u}_0$ , e um conjunto de  $J$  nós  $(\mathbf{u}_0 + \mathbf{h}_j, j = 1, \dots, J)$ . Este *template* é utilizado para fazer a varredura da *TI* e armazenar todos os padrões de treinamento  $pat(\mathbf{u}'_0) = \{t(\mathbf{u}'_0); t(\mathbf{u}_0 + \mathbf{h}_j), j = 1, \dots, J\}$ , onde  $\mathbf{u}'_0$  é qualquer nó central da *TI*, e  $t(\mathbf{u}_0 + \mathbf{h}_j)$  é o valor da *TI* em determinado nó  $\mathbf{u}'_0 + \mathbf{h}_j$  do *grid*. A Figura 3.1 apresenta um exemplo de *template* de busca 2D composto por de 4 vetores  $\mathbf{h}_j, j = 1, \dots, 4$ .



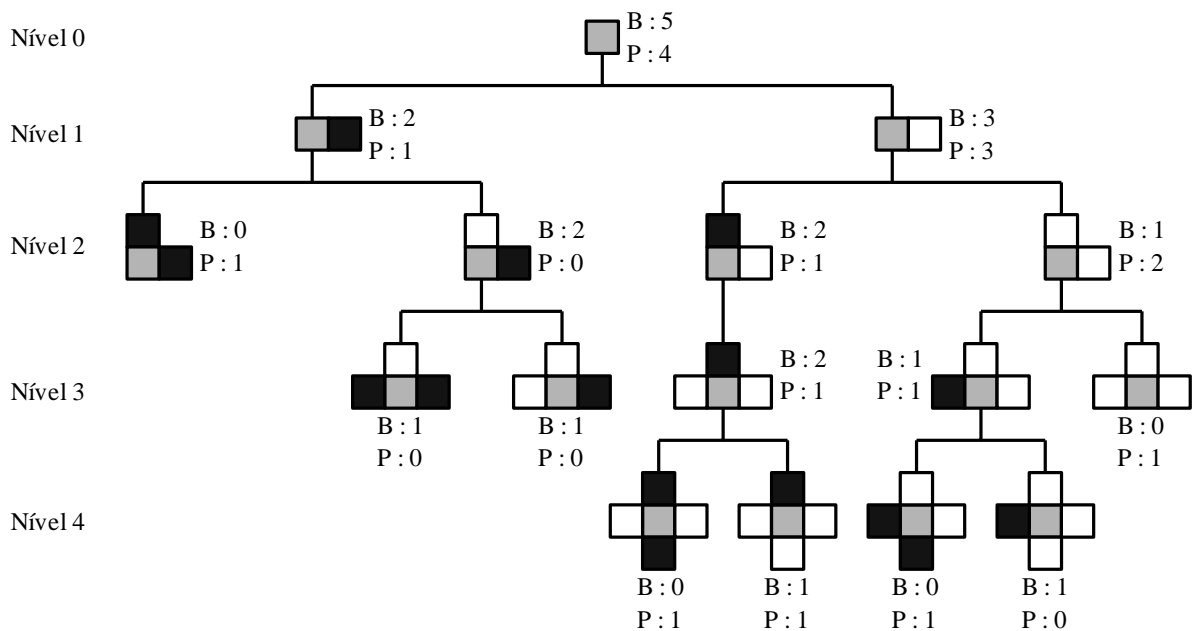
**Figura 3.1** – exemplo de um *template* de busca composto por 4 vetores.

Considere agora a *TI* e o *template* de busca apresentados na Figura 3.2. A árvore de busca correspondente é apresentada na Figura 3.3. A *TI* ilustrada na Figura 3.2 possui duas categorias (branco e preto). O *template* de busca é composto por cinco *pixels* dispostos em um arranjo em forma de cruz. O *pixel* central de cor cinza corresponde ao *pixel* com valor de categoria desconhecido, os outros quatro *pixels* adjacentes enumerados de 1 a 4 são onde as categorias podem ser ou não conhecidas, definindo assim, um evento de dados.



**Figura 3.2** – *TI* contendo duas categorias e *template* de busca composto por 5 nós em um arranjo em forma de cruz (modificado de Boucher, 2007).

A árvore de busca armazena todos os padrões encontrados na *TI* e as suas frequências de ocorrência em função do tamanho e da geometria do *template* de busca escolhido. Cada nível da árvore de busca corresponde a um dos *pixels* que compõe o *template* de busca apresentado na Figura 3.2. A construção de uma árvore de busca binária se inicia com a definição da sua raiz (*root*), que corresponde a um evento de dados vazio (nível 0 da árvore de busca). A frequência associada a raiz da árvore consiste na proporção global de *pixels* pretos e brancos encontrada em uma *TI* erodida (*eroded TI*). Uma *TI* erodida consiste no maior subconjunto da imagem de treinamento original, dentro do qual, o *template* de busca é totalmente contido, eliminando qualquer efeito de borda. Na Figura 3.2, a imagem erodida corresponde ao subconjunto de  $3 \times 3$  *pixels* localizado na porção central da *TI*. Note que neste subconjunto são encontrados 5 *pixels* brancos ( $B = 5$ ) e 4 *pixels* pretos ( $P = 4$ ), o que fornece a proporção global das duas categorias no nível 0 da *search tree*, conforme é ilustrado na Figura 3.3. A frequência de cada padrão é computada a partir da *TI* erodida. As variáveis  $B$  e  $P$  indicam o número de ocorrência das categorias branco e preto, respectivamente, no nó central do *template* de busca dado um determinado evento.



**Figura 3.3** – árvore de busca construída com base na *TI* e no *template* de busca da Figura 3.2 (modificado de Boucher, 2007).

O nível 1 da árvore de busca apresentada na Figura 3.3 armazena o número de ocorrências levando em consideração as categorias encontradas no nó central e no *pixel* #1 do *template* de busca da Figura 3.2. Caso o *pixel* #1 seja preto, os valores de  $B$  e  $P$  correspondem ao número de vezes que o nó central do *template* é branco e preto, respectivamente. Neste

caso, existem três ocorrências do *pixel* #1 do *template* ser preto dentro da *TI* erodida, em duas delas o nó central é branco ( $B = 2$ ), e em uma ele é preto ( $P = 1$ ) (Figura 3.3). De forma análoga, o lado direito da árvore de busca no nível 1 informa o número de vezes que o nó central do *template* de busca é branco ou preto, caso o *pixel* #1 seja branco.

Neste exemplo, o processo de construção da árvore de busca é realizado de tal forma que o *pixel* preto governa as ramificações do lado esquerdo da árvore, e o *pixel* branco as ramificações do lado direito. Posteriormente, esse procedimento é aplicado aos *pixels* #3 e #4 para completar os outros níveis da árvore de busca. Cabe salientar que as árvores de busca não são limitadas a apenas duas categorias, no entanto, a enumeração das combinações de padrões contendo diversas classes pode levar à construção de árvores muito grandes, acarretando em um aumento significativo da demanda computacional. Segundo Mahajara (2004), o tamanho da árvore de busca é controlado por três fatores principais: o tamanho da imagem de treinamento, o número de nós que compõem o *template* de busca, e o número de categorias presentes na *TI*. Note também que a árvore de busca é específica ao tamanho e geometria de determinado *template* de busca.

### 3.3.2 Recuperando probabilidades da árvore de busca

Além de possibilitar o armazenamento dos padrões existentes na *TI*, a construção da árvore de busca permite a recuperação rápida das suas frequências de ocorrência. A *cpmf* local em qualquer nó ao longo do caminho aleatório pode ser lida diretamente a partir da árvore de busca, sem a necessidade de se realizar a varredura da *TI* novamente. O armazenamento de todos os padrões encontrados na *TI* nesta estrutura de dados, permite:

- i. a recuperação do número total  $c$  de padrões com exatamente os mesmos  $J$  valores de dados  $D_J = \{d_j, j = 1, \dots, J\}$ , onde tal padrão é representado por  $\{t(\mathbf{u}'_0 + \mathbf{h}_j) = d_j, j = 1, \dots, J\}$ ;
- ii. para esse mesmo conjunto de padrões, a obtenção do número  $c_k$ , o qual possui um valor específico  $t(\mathbf{u}'_0) = s_k$ ,  $k = 1, \dots, K$  no local central  $t(\mathbf{u}'_0)$ , onde  $K$  é o número total de categorias presentes na *TI*.

A razão entre os termos  $c_k$  e  $c$  fornece a proporção de padrões de treinamento que apresenta o valor central  $t(\mathbf{u}'_0) = s_k$ , dentre todos aqueles que identificaram os  $J$  valores  $t(\mathbf{u}'_0 + \mathbf{h}_j) = d_j$ , conforme a Equação 3.4:

$$(\dots) \mid (\dots) - \quad (3.4)$$

A recuperação da probabilidade condicional associada a um determinado padrão por meio de um *template* de busca completamente informado é relativamente simples. O valor pode ser recuperado através da varredura ao longo das ramificações da árvore de busca até o nível final ser atingido. No entanto, esse processo é mais complexo quando o evento de dados é incompleto. Quando há um *pixel* não informado no *template* adotado, a passagem para os próximos níveis da árvore somente é permitida se aquele *pixel* considerar a possibilidade de assumir todas as classes possíveis, para que, posteriormente, novas ramificações da árvore sejam construídas. Caso um evento de dados não for encontrado na árvore de busca (pelo fato da *TI* não contê-lo), o *pixel* mais distante do nó central do *template* de busca é ignorado até o evento de dados reduzido ser encontrado na árvore, e a *cpmf* possa ser recuperada (Boucher, 2007).

### 3.3.3 Simulação com múltiplos *grids*

O processo de simulação realizado pelo algoritmo *SNESIM* segue o formalismo da simulação sequencial, onde todos os nós que compõe um *grid* de simulação  $G$  são visitados seguindo um caminho aleatório. Os dados originais são realocados para os nós mais próximos, e em todos os nós não informados a simulação sequencial é realizada. Em cada nó  $\mathbf{u}$  pertencente a  $G$ , o *template* de busca  $\tau_J$  é utilizado para recuperar o evento de dados condicional  $dev(\mathbf{u})$ , o qual é definido através da Equação 3.5:

$$(\dots) \{ {}^{(l)}(\dots) {}^{(l)}(\dots) \} \quad (3.5)$$

onde,  $s_k^{(l)}(\mathbf{u} + \mathbf{h}_j)$  é um valor informado para um determinado nó da  $l$ -ésima realização. Tal valor pode ser um dos dados originais, assim como, um valor previamente simulado. Note que pode existir qualquer número de nós com valores não informados entre as  $J$  possíveis localizações do *template*  $\tau_J$  centrado em  $\mathbf{u}$ .

Em seguida, são recuperados todos os  $c$  padrões na *TI* que possuam os mesmos valores de  $dev_J(\mathbf{u})$ . Caso o valor de  $c$  seja inferior a constante  $c_{min}$  (número mínimo de réplicas), define-se um evento de dados menor  $dev_{J-1}(\mathbf{u})$ , descartando o nó mais distante de  $dev_J(\mathbf{u})$ ,

para que, posteriormente, a busca seja reiniciada. Este passo é repetido até  $c \geq c_{min}$ . Neste caso, se assume que  $J'$  ( $J' \leq J$ ) é o tamanho do evento de dados no qual  $c \geq c_{min}$ . A probabilidade condicional, da qual o valor simulado  $s_k^{(l)}(\mathbf{u})$  é obtido, é definida igual a proporção correspondente da *TI*, conforme a Equação 3.6:

$$P(c \geq c_{min} | J' \leq J) = \frac{P(c \geq c_{min}, J' \leq J)}{P(J' \leq J)} \quad (3.6)$$

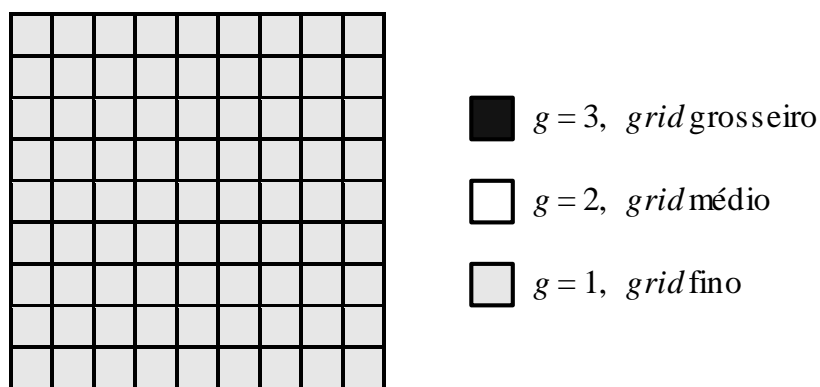
Todavia, quando uma *TI* possui grandes dimensões e um número elevado de padrões, o tamanho da árvore de busca rapidamente torna-se proibitivo. Para contornar este problema, uma abordagem por múltiplos *grids*, ou simulação em cascata, é tipicamente implementada (Boucher, 2007). Em teoria, conforme a Equação 3.1, cada *cpmf* obtida deve ser condicional ao conjunto de dados originais e a todos os nós previamente simulados. Contudo, em situações onde se trabalha em *grids* de simulação com grandes dimensões, esta condição torna-se problemática, especialmente para os algoritmos de simulação sequencial. O condicionamento das realizações torna-se um desafio devido ao aumento significativo da quantidade de nós previamente simulados, o que implica que a *cpmf* do último nó simulado deve ser condicional a todos  $N - 1$  valores previamente simulados. Como resultado, o cálculo de todas estas *cpmfs* exige uma elevada demanda da CPU e RAM.

Para reduzir o consumo dos recursos computacionais, o *template* de busca é utilizado para filtrar o número de dados condicionantes, facilitando assim, a inferência das *cpmfs*. Na prática, o *template* de busca escolhido não deve ser muito pequeno, pois impossibilita a reprodução de padrões de grande escala existentes na *TI*. Por outro lado, um *template* que contém um número muito elevado de nós implica em um grande número de *cpmfs* na árvore de busca, o que aumenta o custo das operações da CPU e a demanda de memória (Strebelle, 2002). Uma solução criada para reproduzir as correlações de longo alcance é fornecida através do conceito de múltiplos *grids*, proposto inicialmente por Gómez-Hernández (1991), e posteriormente desenvolvido por Tran (1994).

A simulação com o uso de múltiplos *grids* (Tran, 1994) é utilizada para capturar estruturas geológicas de grande escala através da adoção de um *template* de busca  $\tau_j$  de grandes dimensões, no entanto, com um número relativamente pequeno de nós. O método consiste na expansão do *template* de busca e na construção de uma nova árvore de busca para simular somente os nós que pertencem a este *template* expandido.

A abordagem por múltiplos *grids* é, essencialmente, um conjunto de simulações em cascata de  $N_G$  *grids* gradativamente menores. Assim, o  $g$ -ésimo ( $1 \leq g \leq N_G$ ) *grid* é constituído por cada  $2^{g-1}$ -ésimo nó do *grid* de simulação final ( $g = 1$ ). O *template* de busca adotado para esse conjunto de *grids* não precisa ter necessariamente a mesma configuração geométrica. O método permite que *templates* de busca maiores sejam utilizados em *grids* grosseiros (*coarse grids*) para capturar estruturas de grande escala encontradas na *TI*, para que, posteriormente, *templates* de busca menores extraíam as informações de pequena escala em *grids* mais finos (*fine grids*). No entanto, note que para cada múltiplo *grid*  $G^g$  criado, uma nova árvore de busca  $T^g$  precisa ser construída, podendo inclusive, utilizar uma imagem de treinamento diferente, que reflita as heterogeneidades específicas daquela escala. Quando o  $g$ -ésimo *grid* é completamente simulado, os valores simulados são retidos e adicionados ao conjunto de dados originais, para posteriormente, serem utilizados no condicionamento das próximas simulações realizadas nos *grids* mais finos. Durante o processo de simulação, todos os nós previamente simulados nos *grids* grosseiros não são revisitados.

Considere, por exemplo, um *grid* Cartesiano 3D  $G$  no qual a simulação é realizada,  $G^g$  pode ser definido como o  $g$ -ésimo subconjunto de  $G$ , tal que  $G^1 = G$ . O *grid*  $G^g$  é obtido através da subamostragem de  $G^{g-1}$  por um fator igual a 2 ao longo dos três eixos Cartesianos. Desta forma,  $G^g$  consiste em um subconjunto de  $G^{g-1}$  obtido pela retenção de todos os outros nós de  $G^{g-1}$ .  $G^g$  é chamado de  $g$ -ésimo nível de múltiplo *grid*. A seguir, a Figura 3.4 ilustra um *grid* de simulação 2D que contém 3 níveis de múltiplos *grids*: grosseiro, médio e fino.



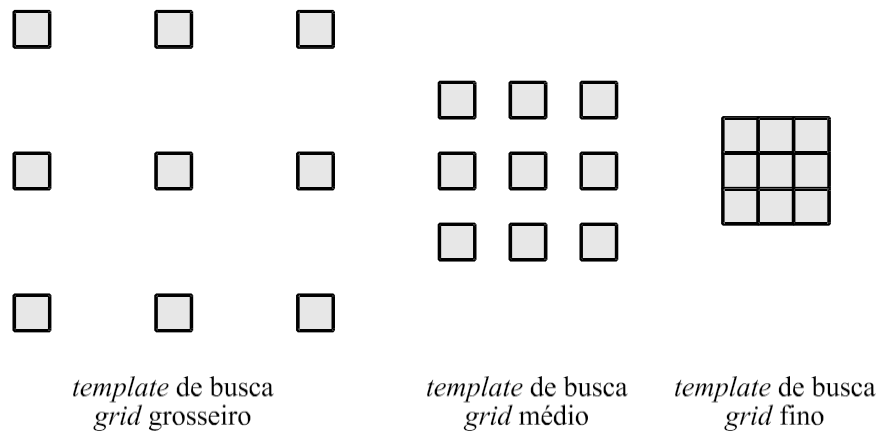
**Figura 3.4** – três níveis de múltiplos *grids*: grosseiro, médio e fino.

No  $g$ -ésimo múltiplo *grid*  $G^g$ , o *template* de busca correspondente  $\tau_j$  é reescalado por um fator igual a  $2^{g-1}$ . Desta forma, temos que o *template* de busca reescalado  $\tau_j^g$  pode ser expresso conforme a Equação 3.7:



$$\{ \quad \quad \quad \} \quad (3.7)$$

Note que  $\tau_J^g$  possui o mesmo número de nós de  $\tau_J$ , contudo, apresenta uma maior extensão espacial, permitindo que ele capture estruturas de grande escala sem a necessidade de aumentar o tamanho da árvore de busca. A Figura 3.5 ilustra a expansão de um *template* de busca 2D com geometria  $3 \times 3$  nós, em três níveis de múltiplos *grids*: grosseiro, médio e fino. A seguir, o Algoritmo 3.1 descreve o *SNESIM* com o uso do conceito de múltiplos *grids*.



**Figura 3.5** – *template* de busca com geometria  $3 \times 3$  em três níveis de múltiplos *grids*.

---

**Algoritmo 3.1** – *SNESIM* com múltiplos *grids*

---

- 01:** Baseando-se em um modelo de blocos e uma *TI* que contém a variável categórica  $s_k$ ,  $k = 1, \dots, K$  definidos em um *grid*  $G$
  - 02:** Defina o número de múltiplos *grids*  $N_G$
  - 03:** **para** cada realização  $l$ ,  $l = 1, \dots, L$  **faça**
  - 04:** **para** cada múltiplo *grid*  $G^g$ ,  $g = N_G, \dots, 1$  **faça**
  - 05:** Realoque o conjunto de dados  $s_k(\mathbf{u}_\alpha)$ ,  $k = 1, \dots, K$ ,  $\alpha = 1, \dots, n(\mathbf{u})$  para os nós  $\mathbf{u}^g$  mais próximos que compõem  $G^g$
  - 06:** Construa o *template* de busca  $\tau_J^g$  através do reescalonamento do *template* de busca  $\tau_J$
  - 07:** Construa a árvore de busca  $T^g$  com base na *TI* e no *template*  $\tau_J^g$
  - 08:** Defina um caminho aleatório visitando todos os nós  $\mathbf{u}^g$  a serem simulados
  - 09:** **para** cada nó  $\mathbf{u}^g_\beta$ ,  $\beta = 1, \dots, N^g$  ao longo do caminho aleatório **faça**
  - 10:** Encontre o evento de dados  $dev_J(\mathbf{u}^g_\beta)$  definido por  $\tau_J^g$
  - 11:** Obtenha a *cpmf*  $Prob(S(\mathbf{u}^g_\beta) = s_k \mid dev_J(\mathbf{u}^g_\beta))$  a partir de  $T^g$
  - 12:** Simule o valor  $s_k^{(l)}(\mathbf{u}^g_\beta)$  a partir da distribuição condicional e o adicione ao conjunto de dados  $s_k(\mathbf{u}_\alpha)$ ,  $\alpha = 1, \dots, n(\mathbf{u})$
  - 13:** **fim para**
  - 14:** **se**  $g > 1$  **faça**
  - 15:** Remova o conjunto de dados  $s_k(\mathbf{u}_\alpha)$ ,  $\alpha = 1, \dots, n(\mathbf{u})$  realocado em  $G^g$
  - 16:** **fim para**
  - 17:** **fim para**
-

### 3.3.4 Reprodução da distribuição marginal

Em algumas situações é necessário que o histograma das realizações reproduza, ou seja próximo, ao histograma de uma determinada distribuição, como, por exemplo, o histograma desagrupado de um conjunto de amostras, ou o de um modelo de blocos. Na prática, recomenda-se a escolha de uma *TI* cujo histograma seja razoavelmente próximo ao da distribuição marginal que se deseja reproduzir (Remy *et al.*, 2009). No entanto, o algoritmo *SNESIM* fornece um parâmetro chamado fator *servo-system*, o qual permite a correção da *cpmf* obtida em cada nó da árvore de busca, aproximando gradualmente o histograma dos valores previamente simulados à distribuição marginal informada pelo usuário.

Suponha que  $p_k^c$ ,  $k = 1, \dots, K$  denote as proporções dos valores da  $k$ -ésima classe simulada até o momento, e  $p_k^l$ ,  $k = 1, \dots, K$  denote as proporções marginais, as quais se deseja reproduzir. O método realiza uma modificação na linha 11 do Algoritmo 3.1, adotando os seguintes passos:

- i. Calcule o valor da *cpmf* da forma originalmente descrita na linha 11 do Algoritmo 3.1;
- ii. Corrija os valores das probabilidades  $Prob(S(\mathbf{u}) = s_k | dev_J(\mathbf{u}))$ , utilizando a Equação 3.8:

$$\left( \left( \right) \mid \left( \right) \right) \left( \left( \right) \mid \left( \right) \right) \left( \right) \quad (3.8)$$

onde,  $\omega \in [0, 1)$  é o fator de intensidade do *servo-system*. Caso  $\omega = 0$ , nenhuma correção é aplicada. Se  $\omega \rightarrow 1$ , a reprodução da distribuição marginal controla totalmente o processo de simulação, com o risco de falhar na reprodução dos padrões encontrados na *TI*.

Se o valor de  $Prob^*(S(\mathbf{u}) = s_k | dev_J(\mathbf{u})) \notin [0, 1]$ , ele é redefinido para o valor limite mais próximo. Todos os valores de probabilidade atualizados são reescalados de tal forma que a sua soma seja igual a 1 (Equação 3.9):

$$\left( \left( \right) \mid \left( \right) \right) \frac{\left( \left( \right) \mid \left( \right) \right)}{\sum \left( \left( \right) \mid \left( \right) \right)} \quad (3.9)$$

### 3.3.5 Particionamento da árvore de busca

Aplicações práticas dos algoritmos baseados em imagens de treinamento mostraram que a *TI* não deve ser um análogo do fenômeno natural que se deseja modelar, mas sim, um repositório de padrões e das suas respectivas propabilidades de ocorrência nesse fenômeno (Boucher, 2009). Para ser adequadamente reproduzida, a *TI* precisa ter padrões repetitivos, e não pode conter tendências e informações locais, uma vez que as posições absolutas dos eventos de dados são perdidas durante a construção da árvore de busca. A qualidade de uma *TI* é medida pelo tamanho da sua árvore de busca, e não pela sua atual dimensão. Desta forma, uma *TI* que contém muitos padrões terá uma árvore de busca maior do que uma que possui pouca variedade de padrões.

O tamanho da árvore de busca é controlado diretamente pelo número de padrões encontrados na *TI*. Ele é uma função do tamanho da *TI*, da dimensão do *template* de busca e do número de categorias presentes na imagem de treinamento (Mahajara, 2004). Conseqüentemente, a velocidade do algoritmo *SNESIM* é função do tamanho da árvore de busca. Devido ao fato de que grande parte dos eventos de dados são incompletos, uma vez que alguns nós dos *templates* de busca podem estar não informados, a recuperação da probabilidade condicional associada a determinado padrão envolve visitas recursivas à árvore de busca. Assim, quanto maior for a *search tree*, mais lento será o processo de busca.

A combinação entre *TIs* contendo um número elevado de padrões, e o uso de *templates* de busca com grandes dimensões, frequentemente, resulta em árvores de busca muito grandes, que não podem ser armazenadas na memória. O tamanho da árvore de busca pode ser reduzido utilizando-se *TIs* e *templates* de busca menores, no entanto, ao custo de sacrificar a qualidade dos resultados. Dentre as principais abordagens propostas para solucionar este problema, destaca-se a utilização de métodos hierárquicos (Maharaja, 2004), que são especialmente recomendados para a simulação de categorias aninhadas.

Outra solução desenvolvida baseia-se na divisão do *grid* de simulação em regiões (Liu, 2006; Wu *et al.*, 2008), e na utilização de uma *TI* específica para cada região. Contudo, esta abordagem não garante compatibilidade entre as *TIs*, o que pode gerar eventuais discontinuidades nas fronteiras entre regiões, visto que não há um modelo informando os padrões de transição entre as mesmas. A aplicação de campos de probabilidade é outra metodologia utilizada para controlar a localização das categorias a serem simuladas

(Strebelle, 2002; Harding *et al.*, 2004). O problema desta abordagem é que a integração de informação secundária através dos campos de probabilidade modifica as probabilidades condicionais recuperadas da árvore de busca, podendo afetar a reprodução dos padrões existentes na *TI*.

Boucher (2009) apresentou um método que fornece grande flexibilidade no processo de simulação com o algoritmo *SNESIM* através do particionamento da *TI* em diferentes classes. O particionamento deve ser realizado de tal forma que os padrões associados com cada classe sejam homogêneos, e possam ser armazenados de forma eficiente em uma árvore de busca. As classes de partição são definidas por um conjunto de filtros que possibilitam a discriminação entre os diferentes padrões encontrados na *TI*. Assim, o método realiza a construção de uma árvore de busca específica para cada classe de partição. Adicionalmente, a definição das classes de partição sobre a *TI* garante a reprodução dos padrões que conectam uma determinada partição com as demais, eliminando assim, a principal desvantagem da abordagem por regiões, que consiste na modelagem das transições entre as mesmas.

As classes de partição podem ser criadas a partir de filtros aplicados sobre as propriedades de determinada *TI*. Estes filtros podem ser simples, como o método das médias móveis (*moving average*), ou mais complexos, tais como os utilizados pelo algoritmo *FILTERSIM* (Zhang *et al.*, 2006), por exemplo. O processo resulta na definição de novas variáveis na *TI*, as quais podem ser armazenadas em um vetor  $\mathbf{a}(\mathbf{u})$  e agrupadas por meio de um algoritmo  $\psi(\mathbf{a}(\mathbf{u}))$  em um conjunto de classes de partição  $\Theta_d$ ,  $d = 1, \dots, D$ . Assim, temos que a classe de partição  $\Theta_d(\mathbf{u})$  em determinado nó  $\mathbf{u}$  da *TI* pode expressa pela Equação 3.10:

$$\left( \right) \quad \left( \left( \right) \right) \quad (3.10)$$

onde,  $\psi()$  é o algoritmo de agrupamento que realiza a partição dos filtros em um conjunto de  $D$  classes.

Em seguida, os padrões encontrados em cada classe de partição são armazenados em uma árvore de busca específica  $T_d$  correspondente a classe de partição  $\Theta_d$ . Desta forma, cria-se uma árvore de busca  $T_d$  para cada de classe de partição  $\Theta_d$ ,  $d = 1, \dots, D$ , para que, posteriormente, cada evento de dados centrado em um local  $\Theta_d(\mathbf{u})$  seja gravado em sua árvore de busca correspondente. A árvore de busca original  $T$  é substituída por um vetor composto por um conjunto de  $D$  árvores menores  $T_d$ ,  $d = 1, \dots, D$ .

Durante recuperação das probabilidades através das *cpmfs*, é essencial permitir que os eventos de dados se estendam sobre as classes de partição vizinhas. Esta sobreposição garante que todos os eventos encontrados na *TI* sejam gravados. Como nenhuma das árvores de busca contém todos os padrões existentes na *TI*, a variedade de padrões encontrados na *TI* somente pode ser restituída por meio da conexão espacial conjunta das árvores.

O particionamento da árvore de busca não modifica o processo de simulação em si, e pode ser facilmente estendido para a simulação em múltiplos *grids*. No entanto, em qualquer nó  $\mathbf{u}$  ao longo do caminho aleatório, a classe partição  $\Theta_d(\mathbf{u})$  daquele local deve ser recuperada para selecionar a árvore de busca apropriada  $T_d$  para obter a *cpmf* correspondente. Após a categoria ser obtida através da *cpmf*, o próximo nó ao longo do caminho é simulado com a sua respectiva árvore de busca, até que todos os nós tenham sido visitados.

Além de proporcionar uma maior capacidade de condicionamento dos padrões simulados através do uso das classes de partição, a abordagem de particionamento da árvore de busca oferece duas melhorias significativas à implementação clássica do algoritmo *SNESIM*: i) fornece maior robustez com respeito a divergência da condição de estacionariedade, e; ii) aumenta consideravelmente a velocidade do processo de simulação (Boucher, 2007).

### **3.4 Simulação de contatos litológicos com o algoritmo *SNESIM***

Apesar da sua ampla aplicação na modelagem de formações geológicas, a grande maioria dos algoritmos tradicionais de *MPS* não foi desenvolvida para a simulação de corpos geológicos maciços com grandes dimensões, e que não possuem um padrão de repetição. Na mineração, estes objetos são frequentemente encontrados em depósitos minerais do tipo filões auríferos, pórfiros cupríferos e formações ferríferas (Boucher *et al.*, 2014). Nestes casos, a localização dos corpos de minério é geralmente conhecida, contudo, a sua geometria e extensão precisa são incertas.

A metodologia proposta fundamenta-se na hipótese básica de que a incerteza geológica reduz a medida que nos distanciamos dos contatos interpretados em um modelo geológico. Pressupõe-se que a posição espacial dos corpos geológicos é conhecida, contudo, a disposição dos contatos interpretados é dita incerta. Esta hipótese permite uma simplificação

do problema da modelagem geológica via simulação geoestatística, uma vez que as únicas porções a serem simuladas do *grid* consistem nos blocos próximos aos contatos interpretados.

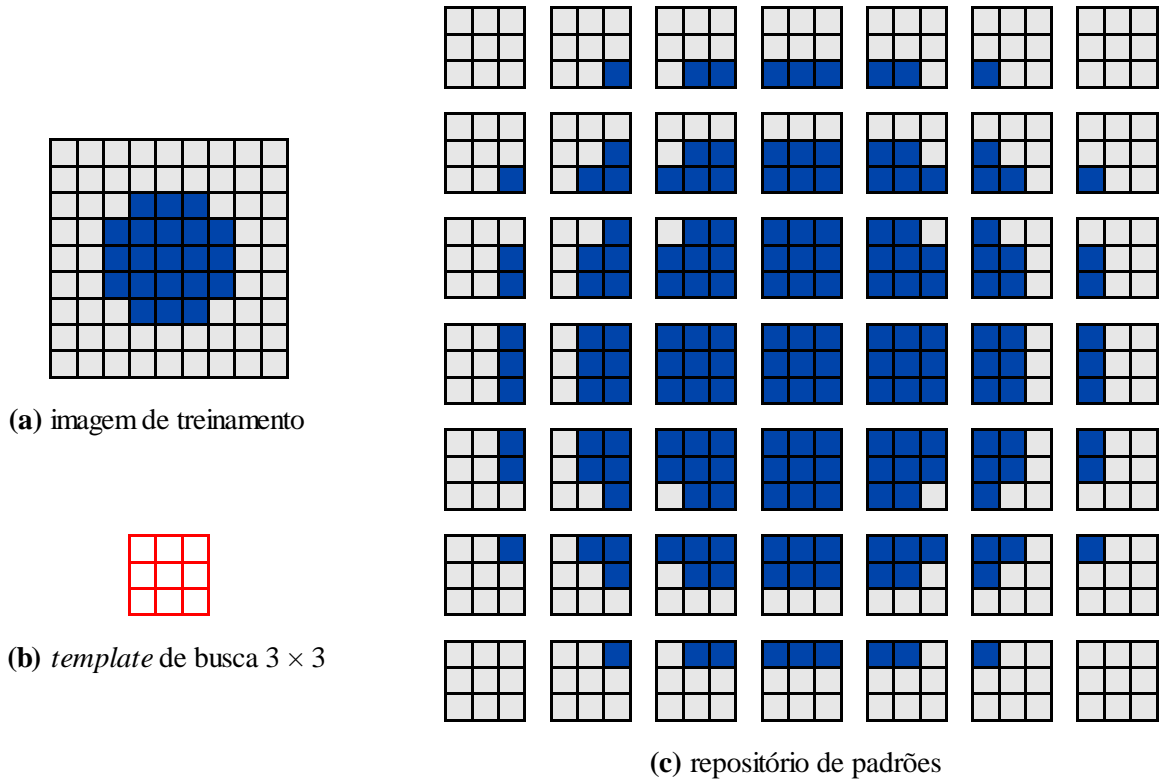
A técnica apresentada baseia-se na metodologia proposta por Pasti *et al.* (2012), e consiste em uma variação do algoritmo desenvolvido por Boucher *et al.* (2014) para a simulação de contatos litológicos a partir de modelos de blocos interpretados, utilizando o conceito de partição da árvore de busca. O método realiza, inicialmente, a definição da zona de incerteza e o agrupamento dos contatos litológicos do modelo de blocos e da *TI* em um conjunto de classes de partição. Em seguida, o processo de simulação decorre somente na zona de incerteza, com base nas classes de partição definidas anteriormente. A metodologia pode ser descrita em quatro etapas:

- i. Defina o tamanho do raio da zona de incerteza ao redor dos contatos litológicos do modelo de blocos e da *TI*;
- ii. Realize o particionamento dos blocos localizados nas zonas de incerteza do modelo de blocos e da *TI* em classes referentes aos diferentes tipos de contatos litológicos;
- iii. Extraia e armazene os padrões da zona de incerteza da *TI* em uma árvore de busca para cada classe de partição criada;
- iv. Remova os valores dos blocos localizados na zona de incerteza do modelo de blocos e os simule com o algoritmo *SNESIM* condicionais aos valores dos blocos localizados fora da zona de incerteza e aos dados de sondagem.

Neste caso, os padrões de treinamento relevantes são aqueles localizados nas zonas de transição entre os diferentes litotipos. Assim, o algoritmo precisa armazenar somente os padrões dos contatos interpretados, ao invés da geometria do sólido geológico como um todo. Isto é realizado por meio da definição de zonas de incerteza ao redor dos contatos interpretados no modelo de blocos e na *TI*. Na imagem de treinamento, esta zona serve como um repositório de padrões, e no modelo de blocos, como um domínio de simulação.

Considere a *TI* 2D binária com dimensão de  $9 \times 9$  *pixels* composta por um fundo de cor cinza e um círculo azul, apresentada pela Figura 3.6a. Ao realizarmos a varredura desta *TI* utilizando um *template* de busca completo, com arranjo de  $3 \times 3$  *pixels* (Figura 3.6b), decompomos a imagem original em um repositório de padrões, formado por um total de 49 padrões de treinamento (Figura 3.6c). No entanto, note que a informação de que tínhamos um

círculo azul no centro da imagem é perdida. O que resta é um conjunto de padrões ilustrando a transição do círculo azul para o fundo cinza (contatos), e alguns padrões contendo somente informação da categoria azul (parte interna do círculo), ou da categoria cinza (porção externa ao redor do círculo).



**Figura 3.6** – decomposição de uma *TI* em um repositório de padrões.

Desta forma, temos que após a definição da zona de incerteza, a *TI* é reduzida aos padrões encontrados ao redor dos contatos. Estes padrões são então agrupados em classes de acordo com a combinação de litotipos, ou a orientação dos padrões, por exemplo. Filtros automatizados também podem ser aplicados, tais como os utilizados pelos algoritmos *FILTERSIM* (Zhang *et al.*, 2006) e *DISPAT* (Honarkhah & Caers, 2010). Em seguida, os padrões de cada classe de partição são armazenados em uma *search tree*, através da abordagem de partição da árvore de busca (Boucher, 2009). Durante o processo de simulação, as classes de partição são definidas no *grid* de simulação, e servem para informar qual árvore de busca deve ser utilizada para recuperar as probabilidades condicionais.

No presente estudo, as classes de partição são geradas através de um algoritmo de agrupamento  $\psi()$  aplicado diretamente sobre a *TI* e o modelo de blocos. O algoritmo  $\psi()$  é utilizado para definir a classe de partição  $\Theta_d$ ,  $d = 1, \dots, D$  de cada nó  $\mathbf{u}$  do *grid*, através da

análise combinatória das categorias dos blocos  $s_k(\mathbf{u})$ ,  $k = 1, \dots, K$  localizados em uma vizinhança de busca  $W(\mathbf{u})$  centrada em  $\mathbf{u}$ , conforme a Equação 3.11:

$$\binom{K}{\psi(\cdot)} = \binom{K}{\psi(\cdot)} \quad (3.11)$$

onde,  $\psi(\cdot)$  é o algoritmo de agrupamento que define a classe de partição de  $\Theta_d(\mathbf{u})$ ,  $d = 1, \dots, D$  com base nos valores dos blocos  $s_k(\mathbf{u})$ ,  $k = 1, \dots, K$  retidos em cada vizinhança de busca  $W(\mathbf{u})$ .

O número de combinações possíveis entre o conjunto de dados armazenados no vetor  $W(\mathbf{u})$  pode ser determinado através da análise combinatória simples dos valores da variável categórica  $s_k(\mathbf{u})$ ,  $k = 1, \dots, K$ . Neste caso, a ordem da combinação dos elementos não importa, e os mesmos precisam ser contados uma única vez. O número de possibilidades de combinações entre os elementos é dado pelo coeficiente binomial  $C_{\varepsilon}^K$ , conforme a Equação 3.12:

$$\binom{K}{\varepsilon} = \frac{K!}{\varepsilon!(K-\varepsilon)!} \quad (3.12)$$

onde,  $K$  é o número total de elementos, e  $\varepsilon = [2, K]$  o intervalo referente ao número de elementos escolhidos. Note que, quando  $\varepsilon = 2$ , estamos considerando todas as possibilidades de contatos binários entre todas as  $K$  categorias. Em contraste, se  $\varepsilon = K$ , há apenas uma combinação possível, a que envolve as  $K$  categorias simultaneamente.

Logo, temos que o número total de combinações possíveis entre as  $K$  categorias, o qual corresponde ao número total de classes de partição  $D$ , é obtido através da soma do conjunto de coeficientes binomiais  $C_{\varepsilon}^K$ ,  $\varepsilon = 2, \dots, K$ , expressa pela Equação 3.13:

$$\sum_{\varepsilon=2}^K \binom{K}{\varepsilon} = D \quad (3.13)$$

onde,  $D$  é o número total de classes de partição  $\Theta_d$ ,  $K$  corresponde ao número total de categorias, e  $\varepsilon = [2, K]$  é o intervalo referente ao número de elementos escolhidos.

O algoritmo desenvolvido para identificar e classificar os contatos litológicos existentes na *TI* e no modelo de blocos, é apresentado em detalhe pelo Algoritmo 3.2. Inicialmente, deve se garantir que o modelo de blocos e a *TI* contêm o mesmo conjunto de



categorias  $s_k$ ,  $k = 1, \dots, K$ . Os únicos parâmetros que precisam ser definidos pelo usuário são: o tamanho do raio  $r$  da zona de incerteza, e a proporção marginal mínima  $p_{min}$ .

O parâmetro  $r$  corresponde ao raio da vizinhança de busca  $W(\mathbf{u})$ , e controla diretamente o grau de perturbação do modelo de blocos. Quanto mais alto for o valor de  $r$ , maior será a quantidade de blocos a serem simulados e, conseqüentemente, maior será o grau de variabilidade das realizações. A proporção marginal  $p_{min}$  serve como parâmetro de controle do nível de sensibilidade do processo de classificação dos contatos. Ela corresponde ao valor percentual mínimo que uma determinada categoria  $s_k$  deve representar dentro da população que compõe a vizinhança de busca  $W(\mathbf{u})$  para ser detectada.

Antes de definir a classe de partição  $\Theta_d(\mathbf{u}_\beta)$ ,  $d = 1, \dots, D$  de determinado nó  $\mathbf{u}_\beta$ ,  $\beta = 1, \dots, N$ , de um *grid*  $G$ , o algoritmo faz a contagem do número total de categorias  $K_{W(\mathbf{u}_\beta)}$  retidas em  $W(\mathbf{u}_\beta)$  para verificar se  $\mathbf{u}_\beta$  encontra-se, ou não, dentro da zona de incerteza  $R$ . Se  $K_{W(\mathbf{u}_\beta)} \geq 2$ , o bloco  $\mathbf{u}_\beta$  é dito pertencente a um contato, e é identificado por uma classe. Caso contrário,  $\mathbf{u}_\beta$  não recebe um valor de classe de partição. Desta forma, o seu valor não é simulado, mas é utilizado como dado condicionante durante o processo de simulação.

---

**Algoritmo 3.2 – definição da zona de incerteza e classificação dos contatos**

---

- 01:** Baseando-se em um modelo de blocos e uma *TI* que contém a variável categórica  $s_k$ ,  $k = 1, \dots, K$  definidos em um *grid*  $G$
  - 02:** Defina o tamanho do raio da zona de incerteza  $r$  e a proporção marginal  $p_{min}$
  - 03:** **para** cada bloco  $\mathbf{u}_\beta$ ,  $\beta = 1, \dots, N$  que compõe  $G$  **faça**
  - 04:**     Armazene todos os valores de blocos  $s_k(\mathbf{u})$  que estão a uma distância  $\leq r$  de  $\mathbf{u}_\beta$  em uma vizinhança de busca  $W(\mathbf{u}_\beta)$
  - 05:**     **para** cada categoria  $s_k$ ,  $k = 1, \dots, K$  encontrada em  $W(\mathbf{u}_\beta)$  **faça**
  - 06:**         se a proporção  $p_k(W(\mathbf{u}_\beta)) < p_{min}$  **faça**
  - 07:**             Remova todos os valores dos blocos  $s_k(\mathbf{u})$  de  $W(\mathbf{u}_\beta)$
  - 08:**     **fim para**
  - 09:**     se o número total de categorias  $K_{W(\mathbf{u}_\beta)} \geq 2$  em  $W(\mathbf{u}_\beta)$  **faça**
  - 10:**         Defina a classe de partição  $\Theta_d(\mathbf{u}_\beta)$ ,  $d = 1, \dots, D$  de  $\mathbf{u}_\beta$  através da análise combinatória simples de todos os valores dos blocos  $s_k(\mathbf{u})$  retidos em  $W(\mathbf{u}_\beta)$
  - 11:**     **do contrário faça**
  - 12:**         Informe que  $\mathbf{u}_\beta$  não pertence à zona de incerteza
  - 13:**     **fim para**
  - 14:** Defina como zona de incerteza de  $G$  a região  $R$  que contém os valores de blocos  $\Theta_d(\mathbf{u})$ ,  $d = 1, \dots, D$
- 

O Algoritmo 3.3 apresenta em detalhe a adaptação do algoritmo *SNESIM* para a simulação dos contatos litológicos de um modelo de blocos interpretado, utilizando

simultaneamente as abordagens de múltiplos *grids* (Tran, 1994) e de partição da árvore de busca (Boucher, 2009). Note que, inicialmente, é necessário que o modelo de blocos e a *TI* tenham sido previamente pré-processados e recebido o mesmo conjunto de classes de partição  $\Theta_d, d = 1, \dots, D$  por meio do Algoritmo 3.2.

Inicialmente, o algoritmo realiza a remoção de todos os valores de blocos  $s_k(\mathbf{u})$  que encontram-se na zona de incerteza do *grid* de simulação (região  $R$ ), a qual foi delimitada anteriormente através do Algoritmo 3.2. Em seguida, os padrões de treinamento encontrados em cada classe da *TI* são armazenados em uma árvore de busca específica  $T_d$  correspondente a uma determinada classe de partição  $\Theta_d, d = 1, \dots, D$ . Para cada múltiplo *grid*  $G^g$ , o conjunto de dados  $s_k(\mathbf{u}_\alpha), \alpha = 1, \dots, n(\mathbf{u})$  é realocado, e o *template* de busca original  $\tau_j$  é reescalado para ajustar-se ao nível de múltiplo *grid* utilizado naquele momento. Adicionalmente, para cada classe de partição, constrói-se uma nova árvore de busca  $T_{\Theta_d}^g$  com base no *template* reescalado  $\tau_j^g$ .

Durante o processo de simulação, em qualquer nó do modelo de blocos visitado pelo caminho aleatório, se recolhe o valor da classe de partição daquele nó  $\Theta_d(\mathbf{u}^g)$  para selecionar qual árvore de busca  $T_{\Theta_d}^g$  deve ser utilizada para recuperar a *cpmf* local. Um valor  $s_k^{(l)}(\mathbf{u}^g)$  é obtido daquela *cpmf* para que, posteriormente, o próximo nó ao longo do caminho aleatório seja simulado de acordo com a sua árvore de busca correspondente, até que todos os nós localizados na zona de incerteza tenham sido visitados.

O algoritmo utiliza as classes de partição  $\Theta_d, d = 1, \dots, D$  como um guia para informar quais padrões de treinamento devem ser recuperados por determinada árvore de busca. Os padrões fornecidos pela árvore de busca correspondente são aqueles encontrados na *TI* justamente nos nós que possuem a mesma classe de partição do nó a ser simulado no modelo de blocos. Isto garante que todos os eventos de dados encontrados na *TI* sejam armazenados.

A qualidade das realizações depende, inicialmente, do uso de uma *TI* que possua padrões de treinamento semelhantes aos padrões encontrados nos contatos litológicos interpretados no modelo geológico. Além disso, é essencial que o tamanho do *template* de busca adotado seja levemente superior ao diâmetro da zona de incerteza utilizada. O particionamento da árvore de busca faz com que nenhuma das árvores contenha todos os padrões presentes na *TI*. Estes padrões somente poderão ser reproduzidos se houver uma conexão espacial entre o conjunto de árvores de busca.

---

**Algoritmo 3.3 – adaptação do SNESIM para a simulação de contatos litológicos**


---

- 01: Baseando-se em um modelo de blocos e uma *TI* que contém a variável categórica  $s_k$ ,  $k = 1, \dots, K$  e as classes de partição  $\Theta_d$ ,  $d = 1, \dots, D$  definidos em um *grid*  $G$
  - 02: Remova todos os valores dos blocos  $s_k(\mathbf{u})$  que encontram-se na zona de incerteza definida pela região  $R$  em  $G$
  - 03: Defina o número de múltiplos *grids*  $N_G$
  - 04: **para** cada realização  $l$ ,  $l = 1, \dots, L$  **faça**
  - 05:     **para** cada múltiplo *grid*  $G^g$ ,  $g = N_G, \dots, 1$  **faça**
  - 06:         Realoque o conjunto de dados  $s_k(\mathbf{u}_\alpha)$ ,  $\alpha = 1, \dots, n(\mathbf{u})$  para os nós  $\mathbf{u}^g$  mais próximos que compõem  $G^g$
  - 07:         Construa o *template* de busca  $\tau_J^g$  através do reescalonamento do *template* de busca  $\tau_J$
  - 08:         **para** cada classe de partição  $\Theta_d$ ,  $d = 1, \dots, D$  **faça**
  - 09:             Construa a árvore de busca  $T_{\Theta_d}^g$  com base na *TI* e no *template*  $\tau_J^g$
  - 10:         **fim para**
  - 11:         Defina um caminho aleatório visitando todos os nós  $\mathbf{u}^g$  a serem simulados
  - 12:         **para** cada nó  $\mathbf{u}^g_\beta$ ,  $\beta = 1, \dots, N^g$  ao longo do caminho aleatório **faça**
  - 13:             Encontre o evento de dados  $dev_J(\mathbf{u}^g_\beta)$  definido por  $\tau_J^g$
  - 14:             Obtenha a classe de partição  $\Theta_d(\mathbf{u}^g_\beta)$  de  $\mathbf{u}^g_\beta$
  - 15:             Obtenha a *cpmf*  $Prob(S(\mathbf{u}^g_\beta) = s_k \mid dev_J(\mathbf{u}^g_\beta))$  a partir da árvore de busca correspondente  $T_{\Theta_d}^g$
  - 16:             Simule o valor  $s_k^{(l)}(\mathbf{u}^g_\beta)$  condicional ao conjunto de dados  $s_k(\mathbf{u}_\alpha)$ ,  $\alpha = 1, \dots, n(\mathbf{u})$  e aos valores dos blocos  $s_k(\mathbf{u}^g)$  previamente simulados e aos que encontram-se fora da zona de incerteza  $R$
  - 17:             Adicione  $s_k^{(l)}(\mathbf{u}^g_\beta)$  ao conjunto de dados  $s_k(\mathbf{u}_\alpha)$ ,  $\alpha = 1, \dots, n(\mathbf{u})$
  - 18:         **fim para**
  - 19:         **se**  $g > 1$  **faça**
  - 20:             Remova o conjunto de dados  $s_k(\mathbf{u}_\alpha)$ ,  $\alpha = 1, \dots, n(\mathbf{u})$  realocado em  $G^g$
  - 21:         **fim para**
  - 22: **fim para**
- 

### 3.5 Estudo de caso

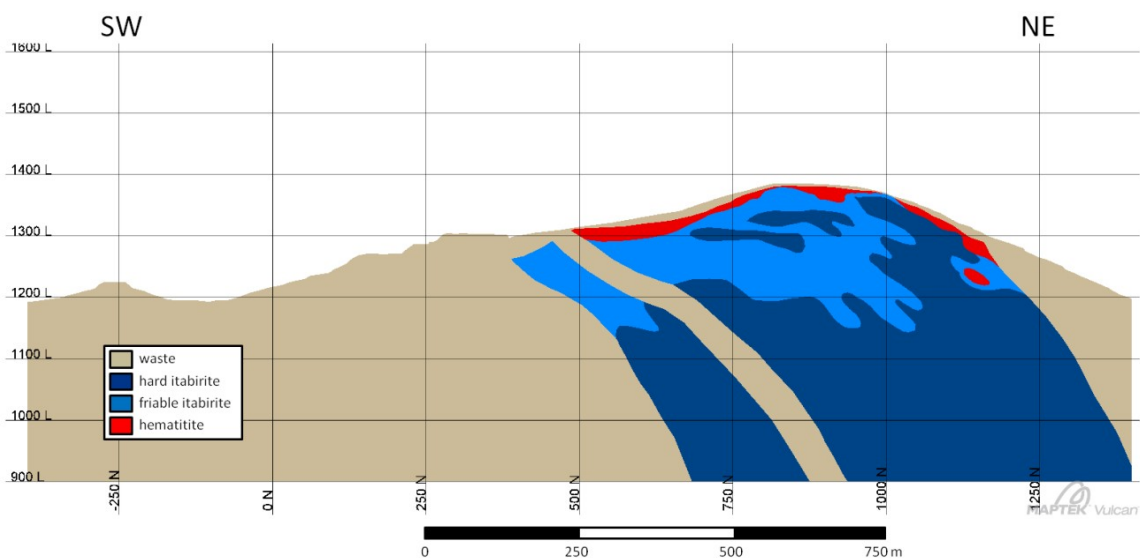
A metodologia proposta na Seção 3.4 é ilustrada na forma de um estudo de caso em um depósito de minério de ferro. O depósito em questão está situado na região do Quadrilátero Ferrífero, no estado de Minas Gerais, na fronteira sudoeste do Cráton São Francisco. A região do Quadrilátero Ferrífero é, atualmente, a principal área produtora de minério de ferro no Brasil, e um dos mais importantes distritos de produção de minério de ferro no mundo.

### 3.5.1 Geologia local

Os minérios de ferro em todo o Quadrilátero Ferrífero foram formados pelo enriquecimento supergênico de itabiritos. Eles são classificados de acordo com o teor de ferro, como minérios de alta qualidade e de grau intermediário. Minérios de alto teor apresentam conteúdo de Fe superior a 64%, e teores baixos de contaminantes, como  $\text{SiO}_2$ ,  $\text{Al}_2\text{O}_3$ , CaO e MgO. Os teores de Fe dos minérios intermediários de baixa qualidade, chamado minério itabirítico, pode variar de 32 a 64%, apresentando quantidades variáveis de contaminantes (Rosière & Chemale, 2000).

O depósito mineral, objeto deste estudo de caso, apresenta um corpo de minério composto por lentes descontínuas de hematitito rico em ferro intercaladas por itabiritos macios ricos em ferro. O sólido geológico possui aproximadamente 800 m de comprimento, 800 m de largura e mais de 420 m de profundidade (Figura 3.7). O corpo de minério está localizado na porção central do depósito, e é composto predominantemente por minério itabirítico friável e macio. As rochas encaixantes consistem em materiais impermeáveis, tais como, filitos, rochas vulcânicas extrusivas básicas e diques, que acabaram favorecendo a atuação dos processos de enriquecimento supergênico.

A morfologia do depósito é parcialmente controlada por dobramentos, e os corpos de minério estão orientados de acordo com lineamentos definidos por diferentes eventos tectônicos. É frequente a ocorrência de intercalações de lentes descontínuas de minério com geometria irregular.



**Figura 3.7** – seção vertical típica do depósito estudado.

A porção superior do depósito é marcada pela presença de complexas intercalações entre lentes de itabirito duro, itabirito friável e hematitito. Na superfície, há uma extensa área de contato entre os corpos de hematitito e estéril (Figura 3.7). A porção inferior tem como característica mais proeminente a existência de um dique (classificado como estéril), que atravessa o corpo de itabirito duro ao longo de toda a sua extensão (Figura 3.7). O mesmo se encontra, aproximadamente, paralelo às superfícies de *footwall* e *hangingwall*.

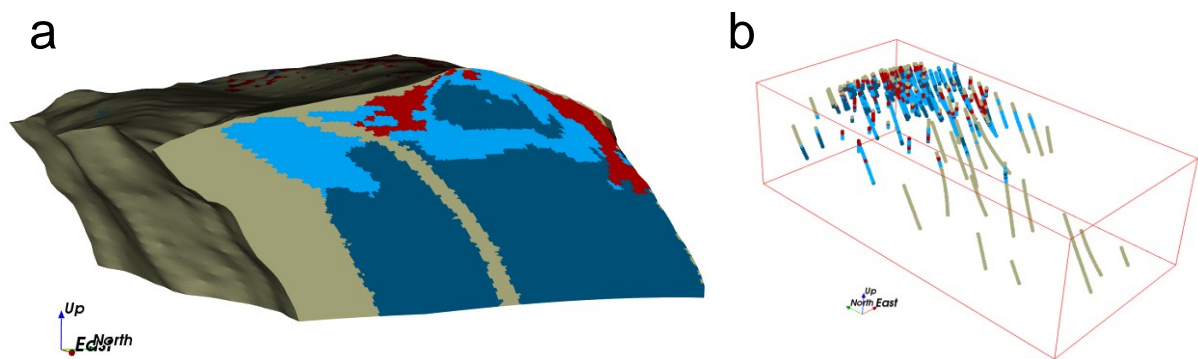
### 3.5.2 Banco de dados

O banco de dados é composto por 1.720 amostras de furos de sondagem com descrição litológica, 17 seções verticais interpretadas, um *wireframe* da envoltória do corpo mineralizado, e um modelo de blocos interpretado, em um suporte de  $10 \times 10 \times 10$  m. A malha de sondagem apresenta arranjo aproximadamente regular com espaçamento médio de 50 m, estando localizada preferencialmente na porção superior do depósito (Figura 3.8b). As seções verticais possuem orientação paralela ao eixo X, e encontram-se regularmente espaçadas a cada 50 m. No presente estudo, os litotipos foram simplificados e divididos em 4 categorias: estéril, itabirito duro, itabirito friável e hematitito.

O *wireframe* da envoltória do corpo geológico foi construído através da extrusão das seções verticais por triangulação. O modelo de blocos fornecido foi gerado através da interpolação da variável categórica por krigagem dos indicadores, utilizando informações dos furos de sondagem e das seções verticais. Posteriormente, este modelo de blocos foi importado para um *grid* estratigráfico de alta resolução, composto por um total de 8 milhões de células, com tamanho médio de  $4 \times 4 \times 2$  m, construído através do mapeamento topológico do *wireframe* da envoltória do corpo geológico. A metodologia utilizada para a construção do *grid* estratigráfico é descrita e ilustrada em detalhe nas Seções 2.4 e 2.5, respectivamente.

No entanto, quando os valores do modelo de blocos original foram importados para o *grid* estratigráfico, foi necessário realizar um pós-processamento dos mesmos. Isto se deve ao fato do *grid* estratigráfico possuir células com suportes muito menores do que os blocos do modelo de referência. Desta forma, as células não informadas do *grid* estratigráfico foram estimadas por meio do interpolador *nearest neighbor* (vizinho mais próximo), com base nos valores centrais dos blocos do modelo original (Figura 3.8a).

A Tabela 3.1 ilustra as discrepâncias existentes entre as proporções das quatro categorias no modelo de blocos, e no conjunto de amostras obtidas por sondagem. Observe que os percentuais das classes de itabirito friável e hematitito são, aproximadamente, 3 e 6 vezes maiores na amostragem, respectivamente. Isto ocorre pelo fato dos furos de sondagem estarem localizados preferencialmente na porção superior do depósito, onde os corpos de itabirito friável e hematitito estão localizados (Figura 3.8). No entanto, neste caso, note que as proporções a serem reproduzidas nos modelos simulados, devem ser próximas às proporções do modelo de blocos interpretado.



**Figura 3.8** – modelo de blocos interpretado (a) e furos de sondagem (b). legenda: cinza: estéril, azul escuro: itabirito duro, azul claro: itabirito friável, vermelho: hematitito.

**Tabela 3.1** – proporções das categorias no modelo de blocos e na amostragem.

Categoria	Proporção (%)	
	Modelo de blocos	Amostragem
Estéril	34,40	30,70
Itabirito duro	52,59	20,99
Itabirito friável	10,47	32,44
Hematitito	2,54	15,87
<b>Total</b>	100,00	100,00

### 3.5.3 Construção da imagem de treinamento

A *TI* utilizada pelo processo de simulação foi gerada por meio de um algoritmo de simulação de objetos implementado no *SGeMS*, chamado *TetrisTiGen*. Ela foi construída com base nos padrões interpretados das seções verticais do depósito em estudo (Figura 3.7). Como os contatos litológicos apresentam padrões curvilineares, o objeto básico utilizado para a simulação foi uma esfera. Neste caso, a principal informação estrutural que pretende-se

extrair de um objeto esférico, não é a sua forma geométrica como um todo, mas sim, os padrões curvilíneos encontrados nas suas extremidades. Cabe salientar, que neste estudo de caso, o modelo de blocos original não pode ser utilizado como *TI*, pois foi construído em um suporte maior do que a resolução do *grid* estratigráfico. Logo, ele não apresenta os padrões na escala exigida para a construção de um modelo de alta resolução.

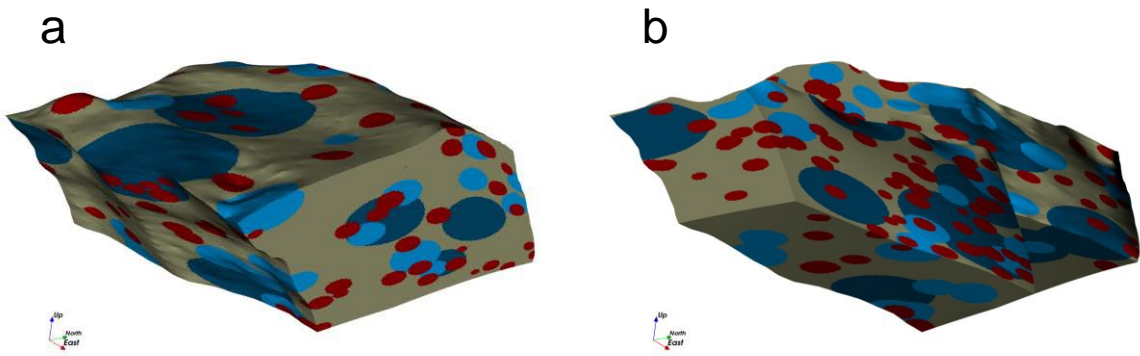
O processo de simulação foi realizado com base em três esferas de raios iguais a 40, 20 e 10 blocos, representando os litotipos itabirito duro, itabirito friável e hematítico, respectivamente. Os raios das esferas foram definidos de acordo com os tamanhos dos corpos geológicos e a geometria dos padrões dos contatos entre os litotipos.

A simulação dos objetos foi realizada em três fases, uma para cada objeto simulado, adotando um posicionamento aleatório dos elementos (Tabela 3.2). O *grid* de simulação utilizado foi o próprio *grid* estratigráfico que comporta o modelo geológico. A sequência de critérios de parada adotada pelas simulações (parâmetro referente ao percentual limite de blocos a serem classificados com determinada categoria) foi selecionada para gerar uma distribuição aproximadamente uniforme entre as três categorias, garantindo que haja um número suficiente de ocorrências dos diferentes padrões de contatos. Adicionalmente, todos os nós não informados com valores simulados foram classificados como estéril.

**Tabela 3.2** – parâmetros utilizados na construção da *TI*.

Parâmetro	Simulação #1	Simulação #2	Simulação #3
Categoria	Itabirito duro	Itabirito friável	Hematítico
Objeto	Esfera	Esfera	Esfera
Raio (num. de blocos)	40	20	10
Posicionamento	Aleatório	Aleatório	Aleatório
Critério de parada (%)	30	25	20

A Figura 3.9 apresenta a *TI* gerada em diferentes pontos de vista: em perspectiva (Figura 3.9a), e através de um conjunto de seções transversais e longitudinais (Figura 3.9b). Note que, neste caso, o objetivo do processo de simulação por *MPS* não é simular um conjunto de esferas. Após o agrupamento das categorias presentes na *TI* em classes de partição, o posicionamento espacial e a forma dos objetos é perdida. A única propriedade de interesse são os padrões de contato entre os elementos.



**Figura 3.9** – imagem de treinamento utilizada em perspectiva (a) e através de seções transversais e longitudinais (b). legenda: cinza: estéril, azul escuro: itabirito duro, azul claro: itabirito friável, vermelho: hematítico.

### 3.5.4 Definição do tamanho da zona de incerteza e classificação dos contatos

A definição do tamanho da zona de incerteza, e o agrupamento dos contatos litológicos do modelo de blocos e da *TI*, utilizados no presente estudo de caso, foi realizada através do Algoritmo 3.2. O tamanho do raio ( $r$ ) escolhido para a zona de incerteza foi de 5 blocos, tanto para o modelo de blocos, como para a *TI*. Em unidades métricas, este raio corresponde a, aproximadamente, 20 m no plano XY, e 10 m com relação ao eixo Z. O parâmetro referente à proporção marginal mínima  $p_{min}$  foi definido igual a zero.

O tamanho do raio da zona de incerteza controla diretamente a quantidade de células a serem simuladas no modelo de blocos. Neste caso, no modelo de blocos interpretado, a zona de incerteza corresponde a um total de 1.786.169 blocos, ou seja, cerca de 22% dos 8 milhões de blocos que compõem o *grid* de simulação. Na *TI*, ela totaliza 3.623.093 blocos, e representa 45% do volume do *grid* (Tabela 3.3). A zona de incerteza adotada permite que qualquer contato litológico possa variar em até 10 blocos, aproximadamente 40 m, entre as realizações.

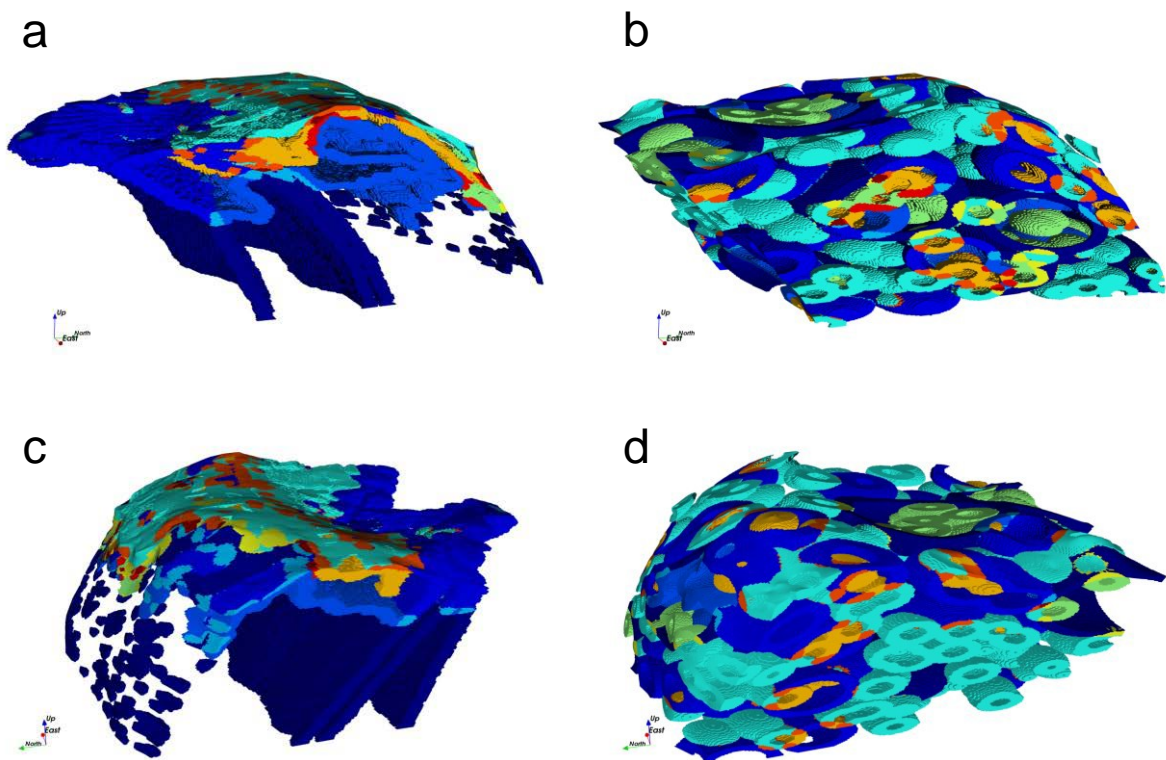
**Tabela 3.3** – volume da zona de incerteza no modelo de blocos e na *TI*.

Zona de incerteza	Volume (blocos)	Volume (%)
Modelo de blocos	1.786.169	22,33
Imagem de treinamento	3.623.093	45,29



O número de classes de partição foi determinado por meio da Equação 3.12. Com base nas 4 categorias, foi gerado um número total de 11 classes, as quais correspondem aos diferentes tipos de contatos existentes no modelo de blocos e na *TI*. A Figura 3.10 apresenta as zonas de incerteza e a disposição espacial do conjunto de classes de partição no modelo de blocos (Figura 3.10a e 3.10c) e na *TI* (Figura 3.10b e 3.10d). Cada cor representa um grupo de padrões de contato. A classe de cor azul escuro, por exemplo, refere-se ao contato entre os litotipos itabirito duro e estéril (Figura 3.10).

Devido à complexidade das intercalações entre o grande número de corpos de hematítico, itabirito friável e estéril, note que, praticamente toda a porção superior do modelo geológico é simulada (Figura 3.10a e 3.10c). Em contraste, as únicas estruturas a serem simuladas em profundidade são o dique, o *footwall*, e algumas porções do *hangingwall* (Figura 3.10a e 3.10c). Na *TI*, todos os contatos entre as esferas servirão como um repositório de padrões (Figura 3.10b e 3.10d).



**Figura 3.10** – classes de partição geradas nas zonas de incerteza do modelo de blocos (a) e (c); e da *TI* (b) e (d).

### 3.5.5 Simulação dos contatos litológicos

Os modelos geológicos simulados foram gerados através do Algoritmo 3.3, fazendo uso da versão do algoritmo *SNESIM* com particionamento da árvore de busca, implementada no *SGeMS*. O *grid* de simulação definido foi o próprio modelo de blocos interpretado, e a *TI* utilizada foi o modelo apresentado pela Figura 3.9. A única região do *grid* a ser simulada foi a zona de incerteza ilustrada anteriormente nas Figuras 3.10a e 3.10c. As simulações foram condicionadas com base em 1.720 amostras de furos de sondagem e blocos localizados fora da zona de incerteza.

Foram geradas um total de 25 realizações, utilizando três níveis de múltiplos *grids* e um *template* de busca isotrópico, composto por um total de 60 nós. O *template* foi definido com um raio igual a 6 blocos, um pouco maior do que aquele usado na zona de incerteza. A distribuição marginal das categorias foi definida de acordo com as proporções dos litotipos encontrados no modelo de blocos interpretado (Tabela 3.1). O restante dos parâmetros adotados no processo de simulação são apresentados em detalhe na Tabela 3.4.

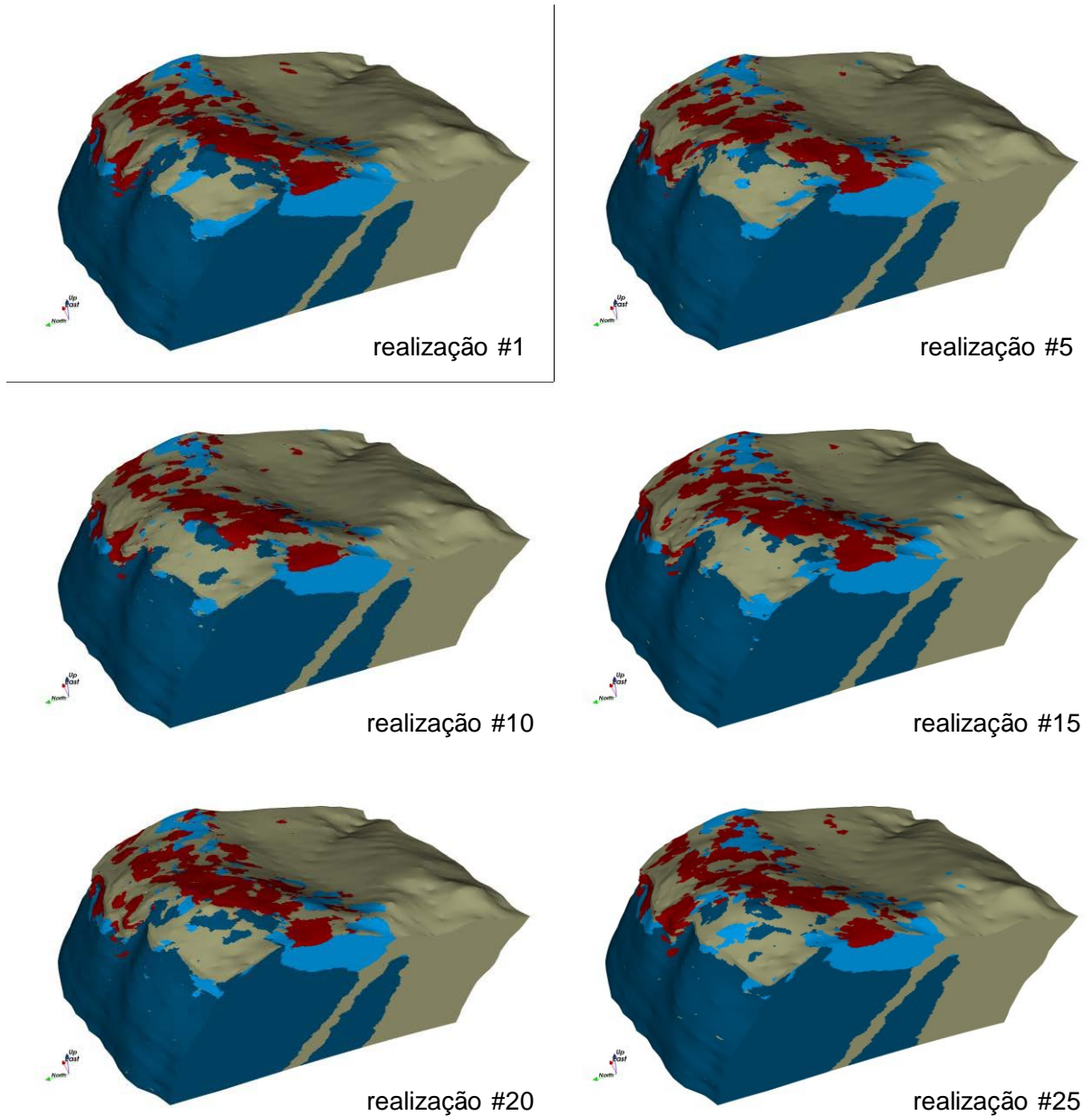
**Tabela 3.4** – parâmetros utilizados na simulação do modelo geológico.

número de realizações	25
semente	7287559
número de categorias	4
distribuição marginal	0.34/0.53/0.10/0.03
número de nós do <i>template</i> de busca	60
tamanho dos raios do <i>template</i> de busca (nós)	6, 6, 6
ângulos de rotação do <i>template</i> de busca	0, 0, 0
número mínimo de réplicas	1
número de múltiplos <i>grids</i>	3

As simulações foram realizadas em um computador com processador Intel Core i7-3930K 3.20 GHz, com 32.0 GB de memória RAM, e sistema operacional Windows 7 64-bit. Nestas condições, o tempo de execução necessário para gerar as 25 realizações foi de 2h 41min.

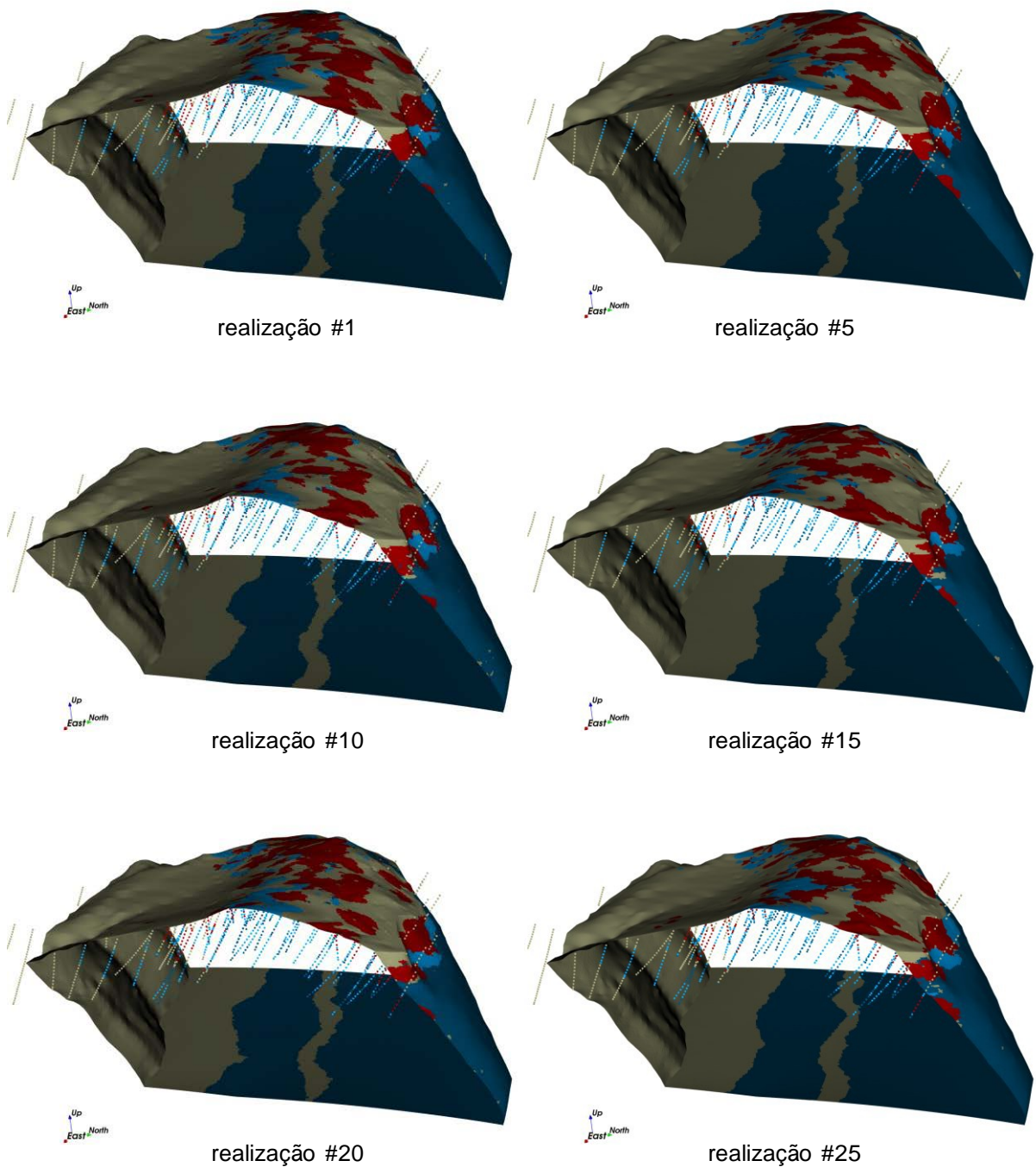
Os modelos simulados não apresentaram nenhuma incoerência com relação ao modelo de blocos original. Isto se deve, em parte, ao fato de que somente os contatos entre os litotipos foram simulados, o que manteve o posicionamento das unidades geológicas de acordo com o modelo interpretado. A Figura 3.11 ilustra um conjunto de 6 realizações obtidas através do

algoritmo adaptado. Na porção superficial, é possível visualizar a complexidade dos contatos entre o hematítico e os corpos de itabirito friável e estável. Em profundidade, se observa a reprodução do dique e do *footwall* (Figura 3.11).



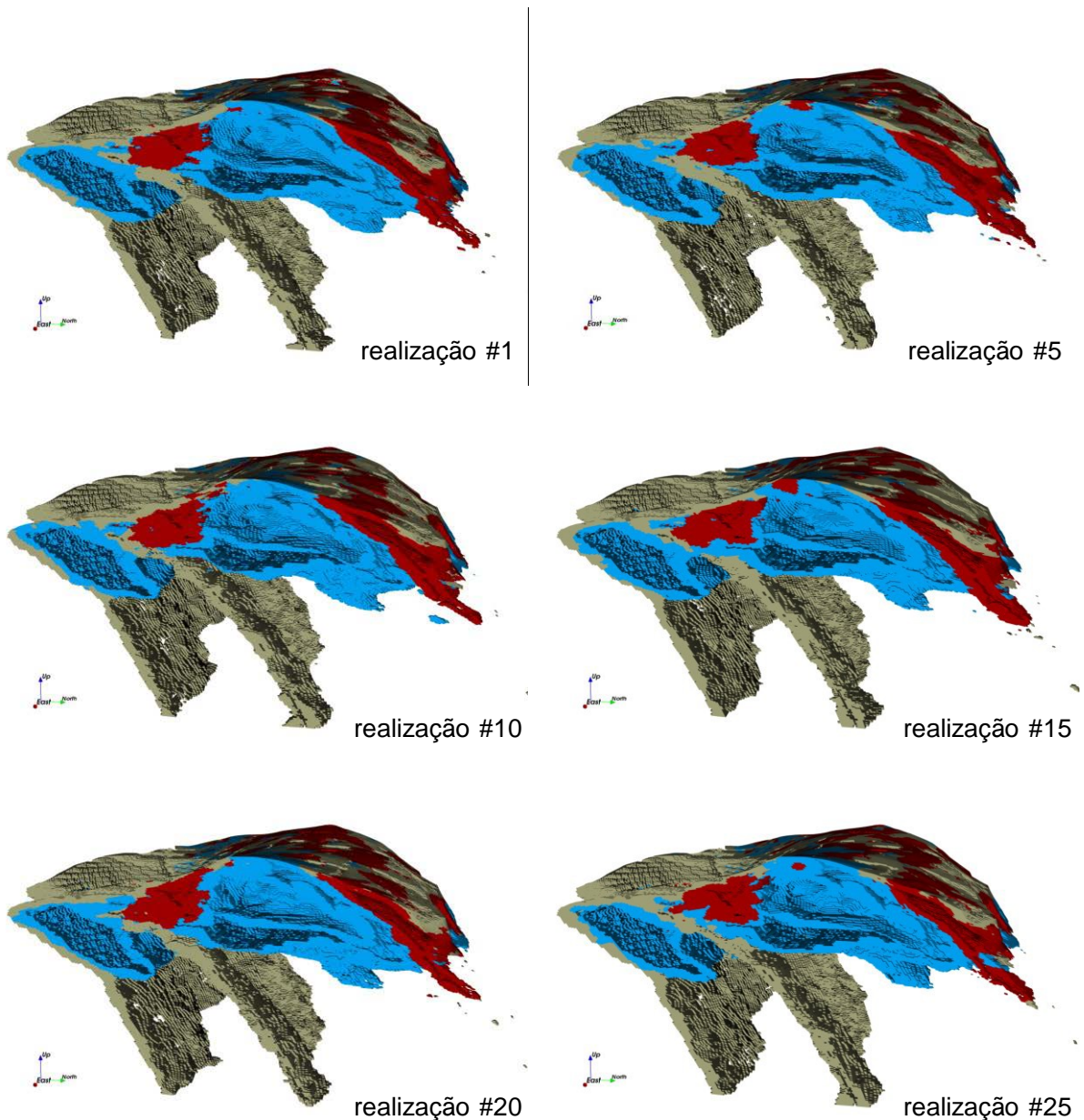
**Figura 3.11** – conjunto de realizações geradas pelo algoritmo *SNESIM*. legenda: cinza: estável, azul escuro: itabirito duro, azul claro: itabirito friável, vermelho: hematítico.

A Figura 3.12 apresenta o mesmo conjunto de realizações através de outro ponto de vista. Nela estão representadas as superfícies da envoltória dos modelos simulados e os furos de sondagem. Na seção da base, pode-se perceber a continuidade dos contatos do dique e do *footwall*, ao longo de toda a extensão do depósito. Observe que os perfis dos contatos seguem um padrão curvilíneo.



**Figura 3.12** – conjunto de realizações geradas pelo algoritmo *SNESIM*. legenda: cinza: estéril, azul escuro: itabirito duro, azul claro: itabirito friável, vermelho: hematítico.

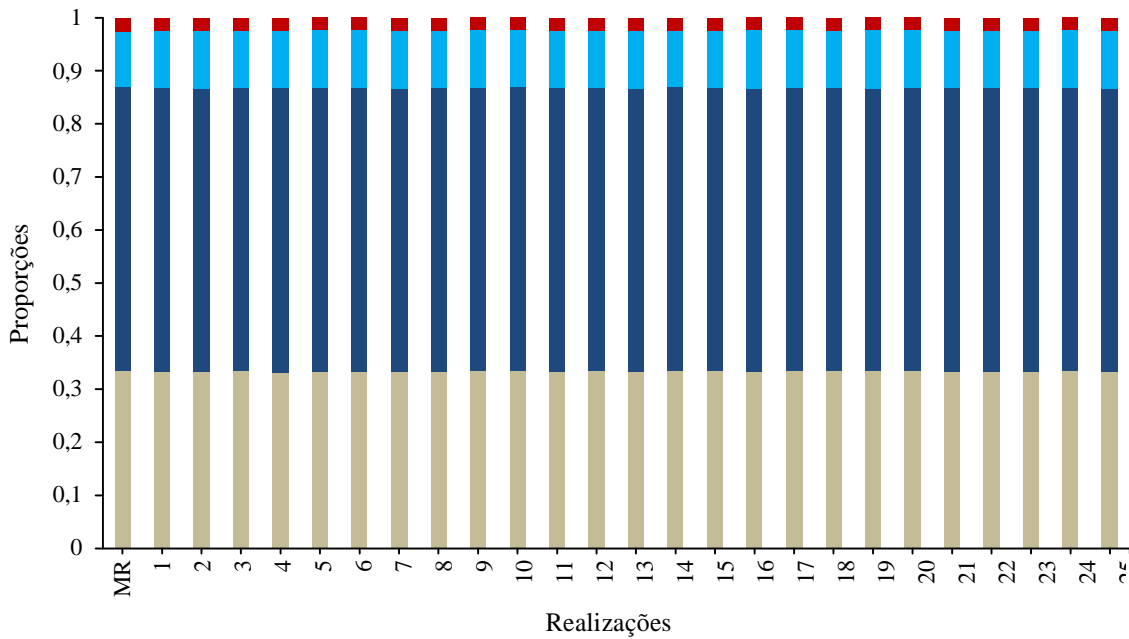
Na Figura 3.13, o corpo de itabirito duro foi filtrado para facilitar a visualização das simulações do dique e da superfície do *footwall*. Pode se observar que ambas as estruturas foram reproduzidas ao longo de toda a profundidade do depósito. Em superfície, nota-se que o hematítico aparece sobre os corpos de itabirito friável, comprovando que as realizações são consistentes com o modelo de referência.



**Figura 3.13** – conjunto de realizações geradas pelo algoritmo *SNESIM*. legenda: cinza: estéril, azul claro: itabirito friável, vermelho: hematítico.

### 3.5.6 Validação e pós-processamento das realizações

O conjunto de 25 realizações foi validado através da comparação entre as proporções das categorias simuladas e a distribuição do modelo de blocos interpretado (modelo de referência). O gráfico apresentado pela Figura 3.14 ilustra as proporções das quatro categorias no modelo de referência (MR) e nas 25 realizações. Pode-se observar que as simulações reproduziram de forma satisfatória as proporções dos litotipos do modelo de blocos interpretado.

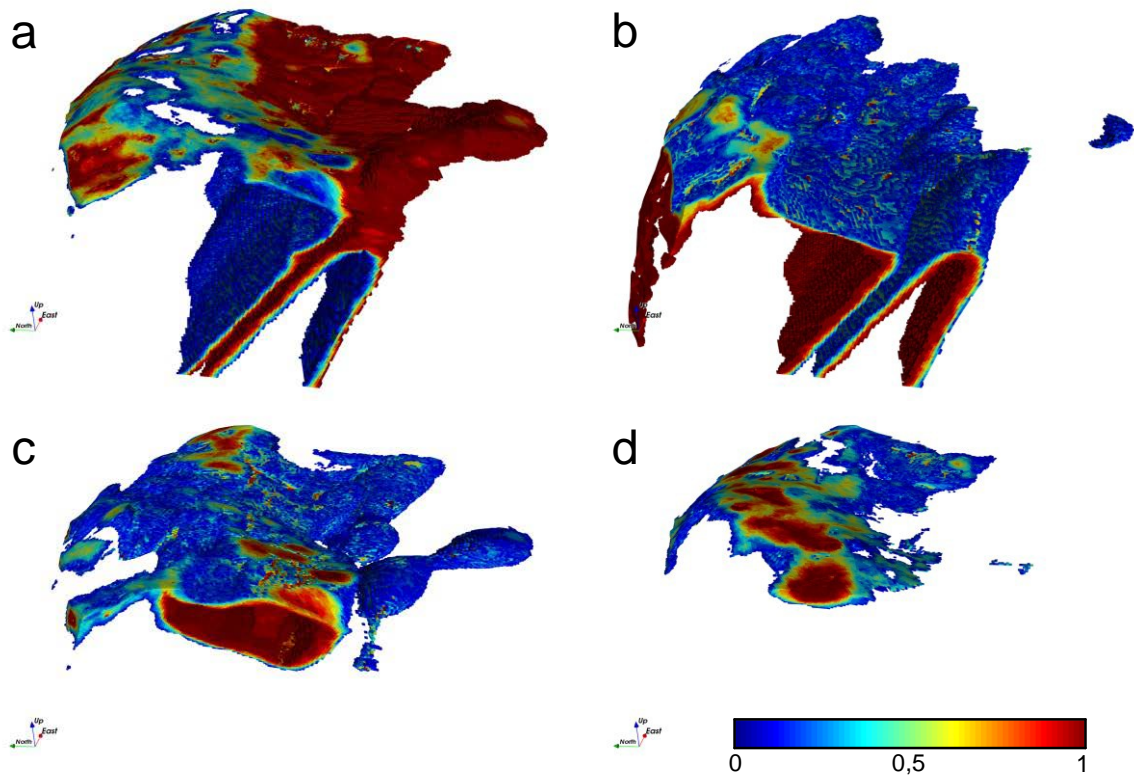


**Figura 3.14** – proporções das categorias no modelo de referência (MR) e nas 25 realizações. legenda: cinza: estéril, azul escuro: itabirito duro, azul claro: itabirito friável, vermelho: hematítico.

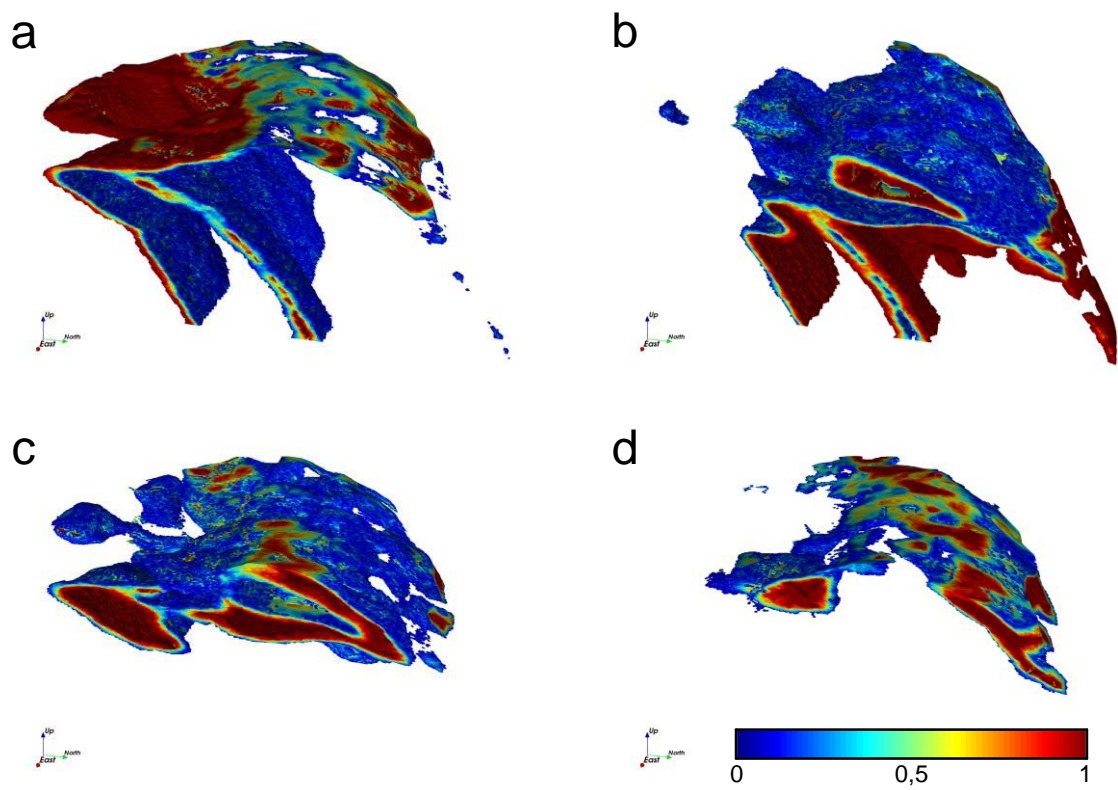
Adicionalmente, foi gerado um mapa *E-type* para cada categoria simulada, através do cálculo da média do conjunto de 25 realizações. A seguir, as Figuras 3.15 e 3.16 apresentam os mapas *E-type* dentro da zona de incerteza definida para o processo de simulação. As porções de cor azul nos mapas indicam baixa probabilidade do bloco pertencer a determinada categoria, as regiões em vermelho, indicam alta probabilidade. Conforme o esperado, em todos os casos há uma mudança gradual das probabilidades ao longo da zona de incerteza.

Verifica-se, inicialmente, a existência de um maior grau de incerteza na superfície do depósito, principalmente devido à complexidade das intercalações entre os corpos de estéril e hematítico e, em menor escala, itabirito friável. Nas Figuras 3.15a e 3.15b, observa-se uma boa definição dos contatos do dique e do *footwall*.

No entanto, nas Figuras 3.16a e 3.16b é possível notar um grau maior de incerteza na porção superior do dique, que provoca, inclusive, a sua segmentação em determinadas partes. Isso se deve ao fato da região ter sido completamente definida dentro da zona de incerteza, devido à pequena espessura do dique nesta área. A medida que se aumenta o tamanho da zona de incerteza, maior é o número de inconsistências com relação ao modelo interpretado. Isto ocorre porque o *SNESIM*, assim como a maioria dos métodos de *MPS*, é um algoritmo baseado em *pixels*, logo, não possui nenhum conhecimento específico das condições geológicas, além dos padrões registrados na árvore de busca.



**Figura 3.15** – *E-type* das 4 categorias simuladas na zona de incerteza: (a) estéril, (b) itabirito duro, (c) itabirito friável e (d) hematitito.



**Figura 3.16** – *E-type* das 4 categorias simuladas na zona de incerteza: (a) estéril, (b) itabirito duro, (c) itabirito friável e (d) hematitito.

## Capítulo 4

### Conclusões e recomendações

Neste capítulo, são apresentadas as conclusões do presente estudo e da metodologia proposta, assim como, recomendações para trabalhos futuros.

#### 4.1 Conclusões

Na indústria mineral, os modelos geológicos são a base de referência para todos os processos de estimativa de teores, classificação de recursos e reservas minerais, projeto da mina, e planejamento da produção de longo prazo. Atualmente, a geoestatística de múltiplos pontos oferece uma alternativa viável para a modelagem e a avaliação da incerteza destes modelos. Entretanto, os algoritmos tradicionais de *MPS* não foram originalmente desenvolvidos para simular corpos geológicos maciços, tipicamente encontrados na mineração. Nestes casos, é comum, as posições dos corpos de minério serem conhecidas, contudo, a sua geometria e extensão serem incertas. Outro aspecto importante a ser considerado no processo de modelagem geológica, é a necessidade de construir modelos numéricos de alta resolução, que possibilitem também, uma elevada aderência geométrica às estruturas curvilineares, frequentemente encontradas em formações geológicas.

No Capítulo 2 foi apresentada uma metodologia para a construção de *grids* estratigráficos a partir do mapeamento topológico das superfícies do *wireframe* do corpo geológico. O método fundamenta-se na projeção do conjunto de superfícies da envoltória mineralizada em um outro domínio topológico, definido em um novo sistema de coordenadas espaciais. Neste novo sistema, as variáveis a serem interpoladas são as coordenadas



Cartesianas das superfícies e da parte interna do sólido geológico. O processo de interpolação é realizado por *DuOK* em um *grid* regular, e os valores estimados determinam as coordenadas dos vértices que formam os blocos do *grid* estratigráfico. A *DuOK* proporciona uma interpolação rápida e acurada dos dados, possibilitando a geração de superfícies e sólidos suavizados. Além de permitir uma melhor representação das formações geológicas, os *grids* estratigráficos contribuem para a qualidade das estimativas de volume/massa, uma vez que as células destes *grids* se conformam de forma adequada às estruturas dos corpos geológicos, o que não ocorre nos modelos de blocos tradicionais.

O método proposto no Capítulo 3 consiste em uma adaptação do algoritmo *SNESIM* para a simulação de contatos litológicos, através do uso do conceito de partição da árvore de busca. A técnica permite a geração de diferentes cenários através da perturbação do modelo geológico interpretado, sem alterar as posições espaciais das unidades geológicas principais. O domínio de simulação consiste na zona de incerteza definida ao redor dos contatos interpretados. Todos os blocos retidos neste zona são agrupados em diferentes classes de partição de acordo com o tipo de contato litológico, e são simulados com base em uma árvore de busca específica para aquela classe de partição. Esta abordagem exige uma alteração da concepção e do processo de construção da *TI*, visto que ela necessita conter somente os padrões dos contatos litológicos. Entretanto, note que, caso haja a disponibilidade de um modelo geológico de referência com resolução compatível à do *grid* estratigráfico, o mesmo pode ser pré-processado e utilizado como *TI*.

O nível de perturbação do modelo de blocos é controlado pelo tamanho da zona de incerteza. Quanto maior for o diâmetro desta zona, maior será o grau de variabilidade das realizações. Note que, o aumento do tamanho da zona de incerteza acarreta também, em um número mais elevado de inconsistências com relação ao modelo interpretado. Os padrões dos contatos são agrupados em diferentes classes de partição que subdividem a *TI* em um conjunto de repositórios menores, de padrões mais homogêneos. A partição dos contatos em classes facilita o condicionamento das realizações, possibilitando que o algoritmo lide com tendências e padrões locais da *TI*. O processo de classificação dos contatos tem papel fundamental na qualidade dos resultados das simulações. O objetivo do algoritmo de agrupamento é minimizar o grau de variabilidade dos padrões armazenados em cada classe de partição.

Apesar do grande apelo teórico e prático, a técnica tem também as suas limitações. O método não é apropriado em duas ocasiões: i) quando a localização dos corpos geológicos é desconhecida, ou; ii) quando os corpos de minério apresentarem pequenas dimensões. A metodologia depende da disponibilidade de um modelo geológico já existente, e que o mesmo esteja definido em um *grid* de simulação. Por isso, a aplicação da técnica não é recomendada em situações onde há um baixo nível de conhecimento geológico do depósito mineral, como em fases iniciais de exploração, por exemplo. Em aplicações mineiras, os modelos de referência devem ser gerados por meio de informações obtidas a partir de sondagens, trincheiras e afloramentos. No setor de recursos energéticos, estes modelos podem ser construídos através da interpretação de perfis sísmicos.

Por fim, o método proposto possibilita a geração de um conjunto de cenários equiprováveis a partir de informações existentes, sem comprometer a interpretação do modelo geológico e sem gerar custos adicionais. Os modelos simulados podem ser utilizados para quantificar a incerteza em relação à geometria e a posição dos contatos entre as diferentes unidades geológicas. Posteriormente, estas informações podem auxiliar na locação de furos de sondagem, assim como, servir como parâmetros adicionais em estudos de sensibilidade dos processos de estimativa de teores, classificação de recursos e planejamento de produção.

## 4.2 Recomendações

Para trabalhos futuros, recomenda-se que a metodologia proposta seja aplicada em outros tipos de depósitos minerais. Adicionalmente, aconselha-se que seja realizada uma análise de sensibilidade com relação ao tamanho da zona de incerteza, e às configurações do próprio processo de simulação, para avaliar o grau de impacto de cada parâmetro no resultado final das realizações.

Dentre as diversas técnicas de krigagem existentes, a *DuK* é a mais indicada para a interpolação de superfícies e volumes, visto que permite a geração de superfícies suaves, com alto grau de acurácia e eficiência computacional. No entanto, outras técnicas não geostatísticas também podem ser utilizadas para a construção de *grids* estratigráficos, incluindo interpoladores iterativos, tais como, o método de mínima curvatura, ou técnicas baseadas em interpolação polinomial segmentada, como *splines*, por exemplo. O critério de

qualidade dos resultados é subjetivo, contudo, sugere-se a escolha de funções que reproduzam a variabilidade espacial desejada, e honrem os valores do conjunto de amostras.

No presente estudo, a determinação do tamanho da zona de incerteza depende do grau de conhecimento geológico do depósito. O ideal é que sempre haja a intervenção de um especialista para defini-la. Em algumas situações, critérios como a densidade de dados, podem ser utilizados para inferir o tamanho deste zona, como a variância de krigagem, por exemplo. Desta forma, áreas densamente amostradas teriam zonas de incerteza menores do que áreas com informações esparsas. A vizinhança de busca adotada para a definição da zona de incerteza pode ter outras geometrias, fazendo uso de diferentes anisotropias de acordo com o tipo de contato litológico detectado. Sugere-se também, a aplicação de outros filtros para classificar os contatos. Estes filtros podem levar em consideração não somente o número de categorias retidas na vizinhança de busca, mas também, informações relacionadas à geometria, ou a orientação dos contatos, por exemplo.

## Referências

- Alabert, F. G. 1987. *Stochastic imaging of spatial distributions using hard and soft information*. M.Sc. Thesis, Stanford University, Stanford, USA.
- Alabert F. G. & Massonnat, G. J. 1990. Heterogeneity in a complex turbiditic reservoir: Stochastic modelling of facies and petrophysical variability. In: *65th Annual Technical Conference and Exhibition*, number 20604, pp. 775-790. Society of Petroleum Engineers.
- Arpat, G. B. 2005. *Sequential simulation with patterns*. Ph.D. Thesis, Stanford University, Stanford, USA.
- Arpat, G. B. & Caers, J. 2007. Conditional simulations with patterns, *Mathematical Geosciences*, 39, 177-203.
- Aug C., Chilès, J. P., Courrioux, G. & Lajaunie, C. 2005. 3D Geological modeling an uncertainty: the potential-field method. In: Leuangthong, O., Deutsch, C. V. (eds.), *Geostatistics Banff 2004*. Springer, Dordrecht, pp. 145-154.
- Auñón, J. & Gómez-Hernández, J. J. 2000. Dual Kriging with Local Neighborhoods: Application to the Representation of Surfaces, *Mathematical Geology*, 32(1), 69-85.
- Boisvert, J. B., Leuangthong O., Ortiz, J. M. & Deutsch, C. V. 2008. A methodology to construct training images for vein-type deposits, *Computers & Geosciences*, 34, 491-502.
- Boucher, A. 2007. *Downscaling of satellite remote sensing data: Application to land cover mapping*. Ph.D. Thesis, Stanford University, Stanford, USA.
- Boucher, A. 2009. Considering complex training images with search tree partitioning, *Computers & Geosciences*, 35, 1151-1158.
- Boucher, A. 2013. Strategies for modeling with multiple-point simulation algorithms. In: Garner, D., Thenin, D. & Deutsch, C. V. (eds.), *Closing the gap: Advances in applied geomodeling for hydrocarbon reservoirs*. Canadian Society of Petroleum Geologists, Calgary, pp. 67-73.
- Boucher, A., Costa, J. F. C. L., Rasera, L. G. & Motta, E. 2014. Simulation of geological contacts from interpreted geological model using multiple-point statistics, *Mathematical Geosciences*, Special Issue.

- Brackbill, J. U. 1993. An adaptive grid with directional control, *Journal of Computational Physics*, 108(1), 38-50.
- Bridge, J. & Leeder, M. 1979. A simulation model of alluvial stratigraphy, *Sedimentology*, 26(5), 617-644.
- Caers, J. & Journel, A. G. 1998. Stochastic reservoir simulation using neural networks trained on outcrop data. In: *SPE ATCE Proceedings*, number SPE 49026. Society of Petroleum Engineers.
- Caers, J., Srinivasan, S. & Journel, A. G. 1999. Geostatistical quantification of geological information for a fluvial-type north sea reservoir. In: *SPE ATCE Proceedings*, number SPE 56655. Society of Petroleum Engineers.
- Caers, J., Avseth, P. & Mukerji, T. 2001. Geostatistical integration of rock physics, seismic amplitudes and geological models in north-sea turbidite systems. In: *SPE ATCE Proceedings*, number SPE 71321. Society of Petroleum Engineers.
- Caers, J., Strebelle, S. & Payrazyan, K. 2003. Stochastic integration of seismic and geological scenarios: A submarine channel saga. *The Leading Edge*, pp. 192-196.
- Calcagno, P., Chilès, J. P., Courrioux, G. & Guillen, A. 2008. Geological modelling from field data and geological knowledge, Part I - Modelling method coupling 3D potential-field interpolation and geological rules, *Physics of the Earth and Planetary Interiors*, 171(1-4), 147-157.
- Castillo, J. E. 1991. *Mathematical Aspects of Grid Generation*, Society for Industrial and Applied Mathematics, Philadelphia, 157 p.
- Chilès, J. P. & Delfiner, P. 1999. *Geostatistics: Modeling Spatial Uncertainty*, John Wiley & Sons, Inc., New York, 695 p.
- Christakos, G. 1992. *Random Fields Models in Earth Sciences*, Academic Press, San Diego, 474 p.
- Chugunova, T. L. & Hu, L. Y. 2008. Multiple-point simulations constrained by continuous auxiliary data, *Mathematical Geosciences*, 40, 133-146.
- Cowan, E. J., Beatson, R. K., Fright, W. R., McLennan, T. J. & Mitchell, T. J. 2002. Rapid Geological Modelling. In: Vearncombe, S. (Ed.), *Applied Structural Geology for Mineral Exploration and Mining International Symposium Abstract Volume*, Australian Institute of Geoscientists Bulletin, 36, 39-41 (AIG: West Perth).
- Cowan, E. J., Beatson, R. K., Ross, H. J., Fright, W. R., McLennan, T. J., Evans, T. R., Carr, J. C., Lane, R. G., Bright, D. V., Gillman, A. J., Oshust, P. A. & Titley, M. 2003. Practical Implicit Geological Modelling. In: Dominy, S. (Ed.), *5th International Mining Geology Conference*, AusIMM, pp. 89-99.
- David, M. 1977. *Geostatistical Ore Reserve Estimation*, Elsevier Scientific Publishing Company, 364 p.

- David, M. 1988. *Handbook of Applied Advanced Geostatistical Ore Reserve Estimation*, Elsevier, Amsterdam, 216 p.
- de Vries, L. M., Carrera, J., Falivene, O., Gratacs, O. & Slooten, L. J. 2009. Application of multiple point geostatistics to non-stationary images, *Mathematical Geosciences*, 41, 29-42.
- Deraisme, J. & Assibey-Bonsu, W. 2013. A mining application of multiple point simulations: modeling of mineralized zones in a hydrothermal deposit. In: Costa, J. F., Koppe, J., Peroni, R. (eds.), *36th APCOM – Application of Computers and Operations Research in the Mineral Industry*, Fundação Luiz Englert, pp. 199-208.
- Deutsch, C. V. 2002. *Geostatistical Reservoir Modeling*, Oxford University Press, New York, 376 p.
- Deutsch, C. V., Pyrcz, M. J. & Tran, T. T. 2002. Geostatistical Assignment of Reservoir Properties on Unstructured Grids, SPE 77427, *SPE Annual Technical Conference and Exhibition*, San Antonio, Texas.
- Deutsch, C. V. & Journel, A. G. 1998. *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York, 369 p.
- Deutsch, C. V., Xie, Y. & Cullick, A. S. 2001. Surface geometry and trend modeling for integration of stratigraphic data in reservoir models, paper SPE 68817 at *SPE Western region meeting*, Bakersfield, California, USA.
- Dimitrakopoulos, R., Mustapha, H. & Gloaguen, E. 2010. High-order statistics of spatial random fields: Exploring spatial cumulants for modeling complex non-Gaussian and non-linear phenomena, *Mathematical Geosciences*, 42, 65-99.
- Dubrule, O. 1983. Two methods with different objectives: Splines and kriging, *Mathematical Geology*, 15(2), 245-257.
- Farmer, C. 1992. Numerical rocks. In: King, P. (Ed.), *The Mathematical Generation of Reservoir Geology*, Clarendon Press, Oxford, pp. 22-33.
- Galli, A. & Murillo, E. 1984. Dual kriging – Its properties and its uses in direct contouring. In G. Verly and others (eds.), *Geostatistics for Natural Resources Characterization: D*. Reidel Publishing Company, Vol. 2, pp. 621-634.
- George, P. L. 1991. *Automatic Mesh Generation: Application to Finite Element Methods*, Wiley, New York, 333 p.
- me - ern nde , . . 2 1. *A stochastic approach to the simulation of block conductivity fields conditioned upon data measured at a smaller scale*. Ph.D. Thesis, Stanford University, Stanford, USA.
- Goodfellow, R., Consuegra, F. A., Dimitrakopoulos, R. & Lloyd, T. 2012. Quantifying multi-element and volumetric uncertainty, Coleman McCreedy deposit, Ontario, Canada, *Computers & Geosciences*, 42, 71-78.

- Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation*, Oxford University Press, New York, 483 p.
- Gordon, W. J. & Thiel, L. C. 1982. Transfinite mappings and their application to grid generation. In: Thompson, J. F. (Ed.), *Numerical Grid Generation*, North-Holland, Amsterdam.
- Guardiano, F. & Srivastava, R. M. 1993. Multivariate geostatistics: beyond bivariate moments. In: Soares, A. (Ed.), *Geostatistics-Troia*, Vol. 1, Dordrecht, Kluwer Academic Publishers, pp. 133-144.
- Haldorsen, H. H. & Lake, L. W. 1984: A new approach to shale management in field-scale models. *Society of Petroleum Engineers Journal*, Vol. 24, no. 8, 447-452.
- Haldorsen, H. H. & Chang, D. W. 1986. Notes on stochastic shales: from outcrop to simulation model. In: Lake, L. & Carroll, H. B. (eds.), *Reservoir characterization*: London, Academic Press, pp. 445-485.
- Haldorsen, H. H. & Damsleth, E. 1990. Stochastic modeling, *Journal of Petroleum Technology*, pp. 404-412.
- Harding, A., Strebelle, S., Levy, M., Thorne, J., Xie, D., Leigh, S. & Preece, R. 2005. Reservoir facies modeling: New advances in MPS. In: Leuangthong, O., Deutsch, C. V. (eds.), *Geostatistics Banff 2004*, Springer, Dordrecht, pp. 559-568.
- Honarkhah, M. & Caers, J. 2010. Stochastic simulation of patterns using distance-based pattern modeling, *Mathematical Geosciences*, 42, 487-517.
- Houlding, S. W. 1994. *3D Geoscience Modeling: Computer Techniques for Geological Characterization*, Springer-Verlag, New York, Berlin, Heidelberg, 309 p.
- Johnson, M. E. 1987. *Multivariate Statistical Simulation*, John Wiley & Sons Publ., New York, 230 p.
- Jones, P., Douglas, I. & Jewbali, A. 2013. Modeling combined geological and grade uncertainty: application of multiple-point simulation at the Apensu gold deposit, Ghana, *Mathematical Geosciences*, 45, 949-965.
- Journel, A. G. 1989. *Fundamentals of Geostatistics in Five Lessons*, Vol. 8, Short Course in Geology, American Geophysical Union, Washington, D.C.
- Journel, A. G. 1994. Modeling uncertainty: Some conceptual thoughts. In: Dimitrakopoulos, R. (Ed.), *Geostatistics for the Next Century*, Kluwer, pp. 30-43.
- Journel, A. G. 2002. Combining knowledge from diverse sources: An alternative to traditional data independence hypotheses, *Mathematical Geology*, 34(5), 573-596.
- Journel, A. G. & Alabert, F. G. 1989. Non-Gaussian data expansion in the earth sciences, *Terra Nova*, 1, 123-134.

- Journel, A. G. & Isaaks, E. H., 1984. Conditional indicator simulation: Application to a Saskatchewan uranium deposit, *Mathematical Geology*, 16(7), 685-718.
- Journel, A. G. & Huijbregts, C. J. 1978. *Mining Geostatistics*, Academic Press, London, 600 p.
- Journel, A. G. & Rossi, M. E. 1989. When do we need a trend model in kriging? *Mathematical Geology*, 21(7), 715-739.
- Journel, A. G. & Deutsch, C. V. 1993. Entropy and spatial disorder, *Mathematical Geology*, 25(3), 329-355.
- Lantuéjoul, C. 2002. *Geostatistical Simulation: Models and Algorithms*, Springer-Verlag, Berlin, Germany, 256 p.
- Liu, Y., Harding, A. & Gilbert, R. 2004. Multiple-point geostatistical simulation. In: *Geostatistics Banff 2004, 7th International Geostatistics Congress*.
- Liu, Y. 2006. Using the Snesim program for multiple-point statistical simulation, *Computers & Geosciences*, 32, 1544-1563.
- Lopez, S. 2003. *Modelisation de reservoir chenalisés meandriforme: Approche génétique et stochastique*, Ph.D. Thesis, École des Mines de Paris, France.
- Lyster, S. J. 2009. *Simulation of geologic phenomena using multiple-point statistics in a Gibbs sampler algorithm*. Ph.D. Thesis, University of Alberta, Edmonton, Canada.
- Mahajara, A. 2004. *Hierarchical simulation of multiple-facies reservoirs using multiple-point geostatistics*. M.Sc. Thesis, Stanford University, Stanford, USA.
- Mallet, J. L. 2002. *Geomodeling*, Oxford University Press, New York, 624 p.
- Mariethoz, G., Renard, P. & Straubhaar, J. 2010. Direct sampling method to perform multiple-point geostatistical simulations, *Water Resources Research*, 46, W11536.
- Matheron, G. 1975. *The theory of regionalized variables and its applications*, English translation of Le Cahiers du Centre de Morphologie Mathématique, Fascicule 5: Ecole des Mines de Paris, Fontainebleau, 175 p.
- Matheron, G. 1976. Forecasting block grade distributions: The transfer functions. In: Guarascio, M., David, M. & Huijbregts, C. J. (eds.), *Advanced Geostatistics in the Mining Industry*, Roma, Italy, D. Reidel Publishing Company, pp. 237-251.
- Matheron, G. 1980. Splines and kriging – their formal equivalence. In: Merriam, D. F. (Ed.), *Down-to-earth statistics – Solutions looking for geological problems*: Syracuse University Geology Contributions, p. 77-95.
- Matheron, G., Beucher, H., Fouquet, C., Galli, A., Guerillot, D. & Ravenne, C., 1987. Conditional simulation of the geometry of fluvio-deltaic reservoirs. *SPE*.



- Mavriplis, D. J. 1996. *Mesh Generation and adaptivity for complex geometries and flows*, Handbook of Computational Fluid Mechanics, Peyret, R. (Ed.), Academic Press, London, 467 p.
- Müller, E. M., Diedrich, C. & Costa, J. F. C. L. 2013. Multipoint simulation applied to the Sossego copper deposit, Brazil. In: Costa, J. F., Koppe, J., Peroni, R. (eds.), *36th APCOM – Application of Computers and Operations Research in the Mineral Industry*, Fundação Luiz Englert, pp. 189-198.
- Mustapha, H. & Dimitrakopoulos, R. 2010. HOSIM: A high-order stochastic simulation algorithm for generating three-dimensional complex geological patterns, *Computers & Geosciences*, 37, 1242-1253.
- Osterholt, V. & Dimitrakopoulos, R. 2007. Simulation of wireframes and geometric features with multiple-point techniques: Application at Yandi iron ore deposit. In: *Orebody Modelling and Strategic Mine Planning*, AusIMM, Spectrum Series, 2nd Ed., 14, pp. 95-124.
- Parra, A. & Ortiz, J. M. 2011. Adapting a texture synthesis algorithm for conditional multiple point geostatistical simulation, *Stochastic Environmental Research and Risk Assessment*, 25, 1101-1111.
- Pasti, H. A., Costa, J. F. C. L. & Boucher, A. 2012. Multiple-point geostatistics for modelling lithological domains at a Brazilian iron ore deposit using the single normal equations simulation algorithm. In: Abrahamsen, P., Hauge, R., Kolbjornsen, O. (eds), *Geostatistics Oslo*, Springer, Dordrecht, pp. 397-408.
- Peredo, O. & Ortiz, J. M. 2011. Parallel implementation of simulated annealing to reproduce multiple-point statistics, *Computers & Geosciences*, 37, 1110-1121.
- Pyrz, M. J., Catuneanu, O. & Deutsch, C. V. 2005. Stochastic surface-based modeling of turbidite lobes, *AAPG Bulletin*, 89(2), 177-191.
- Remy, N., Boucher, A. & Wu, J. 2009. *Applied Geostatistics with SGeMS*, Cambridge University Press, New York, 264 p.
- Renard, D., Wagner, L., Chilès, J. P., Deraisme, J., Jahoda, R. & Vann, J. 2013. Modelling the geometry of a mineral deposit domain with a potential field. In: Costa, J. F., Koppe, J., Peroni, R. (eds.), *36th APCOM – Application of Computers and Operations Research in the Mineral Industry*, Fundação Luiz Englert, pp. 145-155.
- Roberts, E. S. 1998. *Programming Abstractions in C: A Second Course in Computer Science*, Addison-Wesley, Reading, MA, 819 p.
- Rosière, C. A. & Chemale, F. Jr. 2000. Itabiritos e minérios de ferro de alto teor do Quadrilátero Ferrífero – uma visão geral e discussão, *Geonomos*, 8(2), 27-43.
- Royer, J. J. & Vieira, P. C. 1984. Dual formalism of kriging, In: G. Verly and others (eds.), *Geostatistics for Natural Resources Characterization*, D. Reidel Publishing Company, Vol. 2, pp. 691-702.

- Sánchez, P. B., Noda, L. C., Mota, F. J. D., Flores, G. F. G. & Domínguez, A. P. 2009. Adaptive discrete harmonic grid generation. *Mathematics and Computers in Simulation*, 79, 1792-1809.
- Sánchez, P. B., Cortés, J. J. & Flores, G. G. 2013. Harmonic hexahedral structured grid generation. *Mathematical and Computer Modelling*, 57, 2289-2301.
- Silva Jr., A. A., Costa, J. F. C. L., & Rasera, L. G. 2013. Quantifying uncertainty in lithotype modeling of a Brazilian iron ore deposit with multiple-point geostatistics. In: Costa, J. F., Koppe, J., Peroni, R. (eds.), *36th APCOM – Application of Computers and Operations Research in the Mineral Industry*, Fundação Luiz Englert, pp. 123-132.
- Soni, B. K. 1992. Grid generation for internal flow configurations, *Computers & Mathematics with Applications*, 24(5-6), 191-201.
- Soni, B. K. 2000. Grid generation: Past, present, and future. *Applied Numerical Mathematics*, 32, 361-369.
- Srivastava, R. M. 1992. Iterative methods for spatial simulation. In: *SCRF Annual Meeting Report 5, Stanford Center for Reservoir Forecasting*, Stanford, CA, USA, 94305-2220.
- Steger, J. L. & Chausee, D. S. 1980. Generation of body-fitted coordinates using hyperbolic partial differential equations, *SIAM Journal on Scientific and Statistical Computing*, 1(4), 431-437.
- Stoyan, D., Kendall, W. S. & Mecke, J. 1987. *Stochastic geometry and its applications*, John Wiley & Sons, New York, 345 p.
- Straubhaar, J., Renard, P., Mariethoz, G., Froidevaux, R. & Besson, O. 2010. An Improved Parallel Multiple-point Algorithm Using a List Approach. *Mathematical Geosciences*, 43, 305-328.
- Strebelle, S. 2000. *Sequential simulation drawing structures from training images*. Ph.D. Thesis, Stanford University, Stanford, USA.
- Strebelle, S. 2002. Conditional simulation of complex geological structures using multiple-point statistics, *Mathematical Geology*, 34(1), 1-21.
- Strebelle, S., Payrazyan, K. & Caers, J. 2002. Modeling of a deepwater turbidite reservoir conditional to seismic data using multiple-point geostatistics. In: *SPE ATCE Proceedings*, number SPE 77425. Society of Petroleum Engineers.
- Tahmasebi, P., Hezarkhani, A. & Sahimi, M. 2012. Multiple-point geostatistical modeling based on the cross-correlation functions. *Computational Geosciences*, 16, 779-797.
- Thompson, J. F. 1987. A general three-dimensional elliptic grid generation system on a composite block structure, *Computer Methods in Applied Mechanics and Engineering*, 64, 377-411.

- Thompson, J. F., Soni, B. K. & Weatherill, N. P. 1999. *Handbook of Grid Generation*, CRC Press, Boca Raton, 1136 p.
- Tjelmeland, H. 1996. *Stochastic models in reservoir characterization and markov random fields for compact objects*. Ph.D. Thesis, Norwegian University of Science and Technology, Trondheim, Norway.
- Tran, T. T., 1994. Improving variogram reproduction on dense simulation grids, *Computers & Geosciences*, 20(7), 1161-1168.
- Wu, J., Boucher, A. & Zhang, T. 2008. A SGeMS code for pattern simulation of continuous and categorical variables: FILTERSIM, *Computers & Geosciences*, 34, 1863–1876.
- Xu, W. & Journel, A. G. 1993. GTSIM: Gaussian truncated simulation of lithofacies. In: *Report 6, Stanford Center of Reservoir Forecasting*, Stanford, CA, USA.
- Zhang, T. 2006. *Filter-based training pattern classification for spatial pattern simulation*. Ph.D. Thesis, Stanford University, Stanford, USA.
- Zhang, T., Switzer, P. & Journel, A. G. 2006. Filter-based classification of training image patterns for spatial simulation, *Mathematical Geology*, 38, 63-80.
- Zhu, H. 1992. Dual kriging. *NACOG Geostats Newsletter*, Vol. 4, p. 4-5.