

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

PPGEM

METODOLOGIA PARA COMPARAR A EFICIÊNCIA DE ALTERNATIVAS PARA
DISPOSIÇÃO DE AMOSTRAS

Vanessa Cerqueira Koppe

Tese para obtenção do título de

Doutora em Engenharia

Porto Alegre, RS

2009

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

PPGEM

METODOLOGIA PARA COMPARAR A EFICIÊNCIA DE ALTERNATIVAS PARA
DISPOSIÇÃO DE AMOSTRAS

Vanessa Cerqueira Koppe

Engenheira de Minas

Tese apresentada ao programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais – PPGEM, como parte dos requisitos para a obtenção do título de Doutora em Engenharia.

Área de concentração: Metalurgia Extrativa e Tecnologia Mineral

Porto Alegre, RS

2009

Essa Tese foi julgada adequada para a obtenção do Título de Doutora em Engenharia, área de concentração Metalurgia Extrativa e Tecnologia Mineral 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

Co-Orientador: Prof. Dr. Rodrigo de Lemos Peroni

Banca Examinadora:

Prof. Dr. Giorgio Francesco Cesare de Tomi

Prof. Dr. Luis Eduardo de Souza

Dr. Marcelo Godoy

Prof. Dr. Carlos Pérez Bergmann

Coordenador do PPGEM

Aos meus pais,
pelo apoio e dedicação.

AGRADECIMENTOS

Ao Prof. Dr. João Felipe Coimbra Leite Costa pelo tema e orientação.

Ao Prof. Dr. Jair Carlos Koppe pelo apoio.

Ao Prof. Dr. Rodrigo de Lemos Peroni pela co-orientação.

À UFRGS e ao Laboratório de Pesquisa Mineral e Planejamento Mineiro pelas condições oferecidas para realização deste trabalho.

À Datamine Latin América pelo apoio.

Ao CNPq pela bolsa concedida.

Resumo

A avaliação de depósitos minerais é baseada na estimativa de valores de atributos de interesse ao longo do depósito. A amostragem do atributo, em cada local, é, comumente, realizada por sondagem, canaletas ou trincheiras. Contudo, devido aos elevados custos, em especial no que se refere à sondagem, o número de amostras disponíveis é limitado. Em geral, as informações amostradas são utilizadas na obtenção de estimativas ou simulações nos locais não amostrados, criando-se os ditos modelos de blocos. Após criados, os modelos de blocos podem, então, alimentar funções de transferência como, por exemplo, o VPL, que necessitam dos valores do atributo ao longo do depósito. Uma incerteza é associada às estimativas ou simulações, e conseqüentemente, ao valor calculado a partir destas para a função. Para reduzir a incerteza sobre a função, novas amostras devem ser coletadas. Como o número de sondagens é limitado, essas devem ser locadas a fim de trazer o maior benefício em termos de redução de incerteza. A locação de amostras adicionais seguindo uma malha regular e a locação de amostras nas regiões de elevada incerteza sobre o valor do atributo são dois esquemas possíveis para adensamento amostral. A idéia dessa tese é investigar qual a configuração de amostras (compreendendo dados iniciais e dados adicionais) mais eficiente na redução da incerteza sobre a função. A questão sobre quando adotar cada configuração é pesquisada. Para um mesmo número de dados iniciais, os desempenhos da configuração de amostras que compreende dados adicionados regularmente e daquela que compreende os dados adicionados nas regiões de elevada incerteza são avaliados, assim como, os fatores que influenciam nesses desempenhos. Simulação geoestatística é usada para acesso à incerteza sobre a função. Além disso, essa tese propõe um algoritmo para construção das configurações de amostras que visa agilizar a análise dos desempenhos dessas configurações, em cada caso.

Abstract

Mineral deposits evaluation is based on estimated values for relevant attributes along the deposit. Sampling of these attributes at each location is commonly carried out by drilling with core extraction, channel sampling or trenches. However, these types of sampling, specially drilling with core extraction, entail high costs, and due to these costs the number of samples is limited. Generally, the sampled information is used for estimates or simulations at unsampled locations to create the so called block models. After created, the block models can feed transfer functions as net present value, which require the values of the attribute along the deposit. It is known there is an uncertainty associated to the estimates or simulations, and consequently, to the function's value derived from this grade block models. In order to reduce the uncertainty about the function, new samples are required. As the number of drillings is limited, these new samples should be preferentially located in order to bring the maximum benefit in terms of uncertainty reduction. The location of additional samples following a regular grid and the location of samples at regions of high uncertainty related to the attribute's value are two possible schemes to be used. This thesis investigates which sample pattern (comprising initial and additional data) is the most efficient to reduce the uncertainty about a certain transfer function selected. The question about which sampling pattern should be used is investigated. Considering the same number of initial data, the efficiency of the two patterns, i.e. the one which add data regularly spaced and the one which add data at regions of high uncertainty are evaluated. Also, the factors which influence this efficiency are investigated. Geostatistical simulation is used to assess the uncertainty related to the block grades. Additionally, this thesis proposes an algorithm to build a sample pattern, thus making faster the evaluation of this pattern's efficiency.

Índice

Lista de Figuras	IX
Lista de Tabelas	XVII
Resumo	XX
Abstract	XXI
CAPÍTULO 1 - INTRODUÇÃO	1
1.1 - INTRODUÇÃO	1
1.2 - META	5
1.3 - OBJETIVOS	6
1.4 - METODOLOGIA	8
1.5 - ESTADO DA ARTE	13
1.5.1 - Contribuições da Tese	18
1.6 - ORGANIZAÇÃO DA TESE	19
CAPÍTULO 2 – SIMULAÇÃO E ÍNDICES DE INCERTEZA	21
2.1 - VARIÁVEIS REGIONALIZADAS E ESTACIONARIEDADE	23
2.2 - SIMULAÇÃO ESTOCÁSTICA	24
2.2.1 - Simulação Seqüencial	26
2.2.1.1 - Simulação Seqüencial Gaussiana	29
2.2.2 - Simulação por Bandas Rotativas	44
2.2.3 - Simulação Annealing	47
2.3 - ÍNDICES DE INCERTEZA	49
2.3.1 - Intervalo de Probabilidade	49
2.3.2 - Entropia Local	50

2.3.3 - Variância Condicional	51
2.3.4 - Coeficiente de Variação	52
2.3.5 - Intervalo entre Quartis	53
2.4 – FUNÇÕES DE TRANSFERÊNCIA	53
2.4.1 - Mudança de Suporte	54
2.4.2 – Função de Transferência - Tonelagem Lavrável	56
2.4.3 – Função de Transferência - Valor Presente Líquido (VPL)	57
CAPÍTULO 3 – IMPLEMENTAÇÃO DA METODOLOGIA – ESTUDO DE CASO EM 2D	61
3.1 – VARIÁVEL V (ppm)	62
3.1.1 – Dados Adicionais Obtidos pelo Cenário Real (exaustivo)	74
3.1.2 – Dados Adicionais Obtidos pela Realização 12	82
3.1.3 – Dados Adicionais Obtidos pela Realização 54	90
3.1.4 – Dados Adicionais Obtidos pela Realização 57	95
3.1.5 - Resultados Para um Menor Número de Amostras Adicionais	98
3.1.6 - Resultados Para um Menor Número de Dados Iniciais	99
3.1.7 - Cálculo do VPL com Seqüência Ótima de Lavra	107
3.1.8 - Cálculo do VPL com Diferentes Valores Econômicos	109
3.2 – ANÁLISE DOS RESULTADOS DE OTIMIZAÇÃO AMOSTRAL PARA A VARIÁVEL V	111
CAPÍTULO 4 – IMPLEMENTAÇÃO DA METODOLOGIA – ESTUDO DE CASO EM 3D	115
4.1 – VARIÁVEL TEOR DE COBRE (%)	115
4.1.1. – MODIFICAÇÃO DO CENÁRIO DE ONDE SÃO OBTIDOS OS VALORES DAS AMOSTRAS ADICIONAIS	136
4.2 – ANÁLISE DOS RESULTADOS PARA A VARIÁVEL Cu (%)	142
CAPÍTULO 5 – ALGORITMO PARA LOCAÇÃO DE AMOSTRAS ADICIONAIS	146
5.1 – ALGORITMO PARA CONSTRUÇÃO DAS CONFIGURAÇÕES	147

5.2 - AUTOMAÇÃO DA CONSTRUÇÃO DAS CONFIGURAÇÕES E SIMULAÇÃO – <i>JOURNAL FILE</i> (ISATIS [®])	158
5.3 – ROTINA COMPUTACIONAL PARA O CÁLCULO DAS FUNÇÕES DE TRANSFERÊNCIA	163
5.4 - ANÁLISE SOBRE O ALGORITMO PARA LOCAÇÃO DE AMOSTRAS	169
CAPÍTULO 6 – CONCLUSÕES E TRABALHOS FUTUROS	174
6.1 - CONCLUSÕES	174
6.2 – TRABALHOS FUTUROS	179
REFERÊNCIAS BIBLIOGRÁFICAS	182
ANEXO A	190
A.1 - ESCOLHA DOS NÓS DO <i>GRID1</i> QUE DEVEM RECEBER AMOSTRAS ADICIONAIS – CONFIGURAÇÃO 1	190
A.2 - LOCAÇÃO DE AMOSTRAS (3D) E ADIÇÃO DESSAS AMOSTRAS AOS DADOS INICIAIS - CONFIGURAÇÃO 1	193
A.3 - CRIAÇÃO DO <i>GRID</i> PARA DETERMINAR ESPAÇAMENTO MÍNIMO (EM X E Y) ENTRE AMOSTRAS – CONFIGURAÇÃO 2	194
A.4 - AUTOMAÇÃO DA CONSTRUÇÃO DAS CONFIGURAÇÕES E SIMULAÇÃO – <i>JOURNAL FILE</i> (ISATIS [®])	198
ANEXO B	213
ANEXO C	214

Lista de Figuras

Figura 1.1 – Esquema da metodologia utilizada para análise de diferentes configurações de amostras.	9
Figura 2.1 - Esquema da normalização dos dados pela metodologia descrita. Modificado de Goovaerts (1997).	31
Figura 2.2 - Esquema da projeção do vetor h em uma das linhas.	46
Figura 2.3 – Exemplo de distribuição de valores de VPL obtidos para diferentes realizações.	60
Figura 3.1 – (a) Mapa para o cenário real da variável V . (b) Histograma para o cenário real da variável V .	63
Figura 3.2 – (a) Mapa de localização para os dados iniciais da variável V . (b) Histograma para os dados iniciais da variável V .	63
Figura 3.3 – (a) Mapa para a realização que apresentou maior variância dos valores de V . (b) Histograma para a realização que apresentou maior variância dos valores de V .	65
Figura 3.4 – (a) Mapa para a realização que apresentou menor variância dos valores de V . (b) Histograma para a realização que apresentou menor variância dos valores de V .	66
Figura 3.5 – Variograma para a variável V . As linhas azuis representam os variogramas para as realizações que apresentaram maior e menor variância dos valores de V . Os pontos representam o variograma experimental para os dados iniciais (dados originais). A linha preta representa o modelo variográfico (Tabela 3.1) para a variável V .	66
Figura 3.6 – Distribuição acumulada das médias (a) e valores de desvio padrão (b) das realizações obtidas a partir dos dados iniciais. Cada circunferência representa o valor de média (ou desvio padrão) para uma realização. A linha representa a média (ou desvio padrão) para os dados iniciais.	67
Figura 3.7 – Número de realizações versus desvio padrão dos valores da	

- função. (a) Gráfico para tonelagem de minério acima do teor de corte. Foram impressos os resultados para quatro teores de corte. (b) Gráfico para VPL (rota 1). (c) Gráfico para VPL (rota 2). 69
- Figura 3.8 – (a) Gráfico de tonelagem acima do teor de corte. Pontos pretos representam o valor mínimo, o valor máximo, o quartil inferior e o superior da distribuição dos valores de tonelagens, obtidos pelos diversos cenários, para cada teor de corte. (b) Valores de VPL (rota de lavra 1) para as realizações geradas com os dados iniciais. (c) Valores de VPL (rota de lavra 2) para as realizações geradas com os dados iniciais. 71
- Figura 3.9 – Média de V para os 300 blocos. Considerando os blocos com os 300 maiores valores de V no cenário exaustivo. 71
- Figura 3.10 – (a) Mapa para a Configuração 1. (b) Histograma para os dados da Configuração 1. 75
- Figura 3.11 – (a) Mapa para a Configuração 2. (b) Histograma para os dados da Configuração 2. 75
- Figura 3.12 – (a) Mapa para a Configuração 3. (b) Histograma para os dados da Configuração 3. 76
- Figura 3.13 – (a) Mapa para a Configuração 4. (b) Histograma para os dados da Configuração 4. 76
- Figura 3.14 – Desvio padrão para a função tonelagem acima de 40 ppm. Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional. 80
- Figura 3.15 - Desvio padrão para a função tonelagem acima de 150 ppm. Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional. 80
- Figura 3.16 - Desvio padrão para a função tonelagem acima de 300 ppm. Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional. 80
- Figura 3.17 - Desvio padrão para a função tonelagem acima de 500 ppm. Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional. 81
- Figura 3.18 - Desvio padrão para a função VPL (Rota 1). Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de

informação adicional.	81
Figura 3.19 - Desvio padrão para a função VPL (Rota 2). Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional.	82
Figura 3.20 – (a) Mapa para a realização 12, obtida com os dados iniciais. (b) Histograma para a realização 12, obtida com os dados iniciais.	83
Figura 3.21 – (a) Mapa para a Configuração 1 (realização 12 como fonte de informação). (b) Histograma para a Configuração 1 (realização 12 como fonte de informação).	83
Figura 3.22 – (a) Mapa para a Configuração 2 (realização 12 como fonte de informação). (b) Histograma para a Configuração 2 (realização 12 como fonte de informação).	84
Figura 3.23 – (a) Mapa para a Configuração 3 (realização 12 como fonte de informação). (b) Histograma para a Configuração 3 (realização 12 como fonte de informação).	84
Figura 3.24 – (a) Mapa para a Configuração 4 (realização 12 como fonte de informação). (b) Histograma para a Configuração 4 (realização 12 como fonte de informação).	85
Figura 3.25 – Mapa e histograma para a Configuração 5 (realização 12 como fonte de informação).	88
Figura 3.26 – Mapa e histograma para a Configuração 6 (realização 12 como fonte de informação).	88
Figura 3.27 – (a) Mapa para a realização 54, obtida com os dados iniciais. (b) Histograma para a realização 54, obtida com os dados iniciais.	91
Figura 3.28 – (a) Mapa para a Configuração 1 (realização 54 como fonte de informação). (b) Histograma para a Configuração 1 (realização 54 como fonte de informação).	91
Figura 3.29 – (a) Mapa para a Configuração 2 (realização 54 como fonte de informação). (b) Histograma para a Configuração 2 (realização 54 como fonte de informação).	92
Figura 3.30 – (a) Mapa para a Configuração 4 (realização 54 como fonte de informação). (b) Histograma para a Configuração 4 (realização 54 como fonte de informação).	92
Figura 3.31 - Desvio padrão para a função tonelagem acima de 40 ppm.	

Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.	93
Figura 3.32 - Desvio padrão para a função tonelagem acima de 150 ppm. Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.	93
Figura 3.33 - Desvio padrão para a função tonelagem acima de 300 ppm. Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.	94
Figura 3.34 - Desvio padrão para a função tonelagem acima de 500 ppm. Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.	94
Figura 3.35 - Desvio padrão para a função VPL (Rota 1). Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.	95
Figura 3.36 - Desvio padrão para a função VPL (Rota 2). Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.	95
Figura 3.37 – (a) Mapa para a realização 57, obtida com os dados iniciais. (b) Histograma para a realização 57, obtida com os dados iniciais.	96
Figura 3.38 – (a) Mapa para a Configuração 1 (realização 57 como fonte de informação). (b) Histograma para a Configuração 1 (realização 57 como fonte de informação).	96
Figura 3.39 – (a) Mapa para a Configuração 2 (realização 57 como fonte de informação). (b) Histograma para a Configuração 2 (realização 57 como fonte de informação).	97
Figura 3.40 – (a) Mapa para a Configuração 4 (realização 57 como fonte de informação). (b) Histograma para a Configuração 4 (realização 57 como fonte de informação).	97
Figura 3.41 – Desvio padrão (VPL, rota 1) versus número de amostras.	100
Figura 3.42 – Mapa e histograma para a configuração com 30 dados.	100
Figura 3.43 – Mapa e histograma para a configuração com 60 dados.	101
Figura 3.44 – Mapa e histograma para a configuração com 99 dados.	101
Figura 3.45 – (a) Configuração com 225 dados. (b) Configuração com 275 dados.	102

- Figura 3.46 – Desvio padrão versus número de realizações para a configuração com 30 dados. 102
- Figura 3.47 – Mapa e histograma para os dados da Configuração 2, considerando 39 dados adicionados a 60 dados iniciais. 104
- Figura 3.48 - Realização 111 como fonte de informação adicional. (a) Histograma com os dados da Configuração 1. (b) Histograma com os dados da Configuração 2. 104
- Figura 3.49 - Desvio padrão para a função VPL (Rota 1, considerando 2830 blocos minerados). Considerando o cenário exaustivo (cinza), a realização 111 (verde) e a realização 74 (amarelo) como fonte de informação adicional. 105
- Figura 3.50 – Histograma dos valores de IEQ. (a) Simulação com base nos 60 dados da configuração inicial, valores de IEQ na região onde foram adicionadas amostras (até o limite dos dados da configuração inicial). (b) Simulação com base nos 195 dados da configuração inicial. 107
- Figura 3.51 – Seqüência de lavra. Cada período compreende uma seqüência com 368 blocos (5 m x 5 m, nas direções x e y). 110
- Figura 3.52 – Mapa com valores de IEQ nos locais simulados a partir dos 60 dados iniciais (na figura são mostrados os valores de IEQ na região onde foram adicionadas amostras, até o limite dos dados iniciais). Os valores mais altos de IEQ estão marcados em azul (aproximadamente os 27% mais altos). 114
- Figura 4.1 – (a) Vista em perspectiva dos teores de Cu regularizados (coordenadas em metro). (b) Histograma desagrupado para os teores de Cu regularizados. 117
- Figura 4.2 – (a) Vista em perspectiva (mostrando seções em $x=776,5$, $y=1516,5$ e $z=1184,5$) para a realização que apresentou maior variância dos teores de Cu. (b) Histograma para a realização que apresentou maior variância dos teores de Cu. 119
- Figura 4.3 – (a) Vista em perspectiva (mostrando seções em $x=776,5$, $y=1516,5$ e $z=1184,5$) para a realização que apresentou menor variância dos teores de Cu. (b) Histograma para a realização que apresentou menor variância dos teores de Cu. 119

- Figura 4.4 – Correlogramas para os teores de Cu. As linhas pontilhadas representam os correlogramas para as realizações que apresentaram maior e menor variância dos teores de Cu. Os pontos verdes representam o correlograma experimental para os dados originais e o número ao lado de cada ponto representa o número de pares encontrados para a respectiva distância. A curva preta contínua representa o modelo de continuidade espacial utilizado na simulação dos teores de Cu. 120
- Figura 4.5 – Esquema do plano horizontal com a localização dos furos de sonda da Configuração Completa (círculos vazios), dos furos de sonda selecionados para compor os dados iniciais (círculos cheios) e da região A (linha tracejada). 121
- Figura 4.6 – Vista em perspectiva e histograma desagrupado para os dados iniciais. 122
- Figura 4.7 – (a) Vista em perspectiva (mostrando seções em $x= 777$, $y= 1517$ e $z= 1185$) para a realização que apresentou maior variância dos teores de Cu. (b) Histograma para a realização que apresentou maior variância dos teores de Cu. 123
- Figura 4.8 – (a) Vista em perspectiva (mostrando seções em $x= 777$, $y= 1517$ e $z= 1185$) para a realização que apresentou menor variância dos teores de Cu. (b) Histograma para a realização que apresentou menor variância dos teores de Cu. 123
- Figura 4.9 – Correlogramas para os teores de Cu. As linhas pontilhadas representam os correlogramas para a realização que apresentou maior variância dos teores de Cu, para a realização que apresentou variância intermediária e para a realização que apresentou menor variância. A curva preta contínua representa o modelo de continuidade espacial utilizado na simulação dos teores de Cu. 124
- Figura 4.10 - Vistas em perspectiva dos blocos que pertencem à cava (azul), da linha que delimita a região A (linha branca) e da superfície topográfica (marrom). 128
- Figura 4.11 - Número de realizações versus desvio padrão de VPL. 129
- Figura 4.12 – (a) Esquema mostrando os furos de sonda que serão

- adicionados na Configuração 1 (triângulos) e na Configuração 2 (quadrados). Os dados iniciais são apresentados como círculos pretos. (b) Vista superior e em perspectiva dos blocos que pertencem à cava, os pontos e linhas em verde representam os furos de sonda que são adicionados pelas configurações 1 e 2. 130
- Figura 4.13 – (a) Vista em perspectiva para os dados da Configuração 1. (b) Histograma desagrupado para os dados da Configuração 1. Círculos pretos indicam os furos adicionados aos dados iniciais. 131
- Figura 4.14 – Vista em perspectiva e histograma desagrupado para os dados da Configuração 2. Círculos pretos indicam os furos adicionados aos dados iniciais. 131
- Figura 4.15 - Número de realizações versus desvio padrão de VPL. 132
- Figura 4.16 – (a) Valores Q2,5 e Q97,5, apresentados na Tabela 4.6. (b) Esquema da precisão necessária para um estudo conceitual, de pré-viabilidade e um estudo de viabilidade, segundo Lee (1984). 134
- Figura 4.17 - Número de realizações versus desvio padrão de número de blocos acima de 1%. 135
- Figura 4.18 - Vista em perspectiva e histograma dos teores dos blocos para o cenário real (mostrando seções em $x= 770$, $y= 1520$ e $z= 1186$). Apenas os blocos minerados são considerados. 137
- Figura 4.19 - Vista em perspectiva e histograma dos teores dos blocos para a realização 1 (mostrando seções em $x= 770$, $y= 1520$ e $z= 1186$). Apenas os blocos minerados são considerados. 137
- Figura 4.20 – (a) Vista em perspectiva para os dados da Configuração 1b. (b) Histograma desagrupado para os dados da Configuração 1b. 138
- Figura 4.21 – (a) Vista em perspectiva para os dados da Configuração 2b. (b) Histograma desagrupado para os dados da Configuração 2b. 139
- Figura 4.22 – Histograma dos valores de média de IEQ obtidos por simulação com base nos dados da configuração inicial. 144
- Figura 4.23 - Média de valores de IEQ obtidos com a simulação dos dados iniciais, os locais com valores mais altos estão em destaque azul (aproximadamente 24% mais altos valores). 145
- Figura 5.1 – Exemplo da construção do *grid1* (*grid* criado para posicionar as sondagens adicionais formando uma malha aproximadamente

regular com as sondagens iniciais).	150
Figura 5.2 – Esquema da estimativa dos nós do <i>grid1</i> com base na variável <i>x</i> do arquivo COLLAR.	151
Figura 5.3 – Locais do <i>gridsim</i> amostrados pelas novas sondagens.	152
Figura 5.4 – Esquema da busca de dados para interpolação, por Média Móvel, dos valores do índice de incerteza no <i>grid2</i> .	156
Figura 5.5 - <i>Grid2</i> criado com nós além do <i>gridsim</i> , na direção <i>x</i> .	156
Figura 5.6 – Exemplo da construção do <i>grid3</i> (criado para local as sondagens que devem ser adicionadas nas regiões de elevada incerteza) e escolha dos nós que devem receber sondagem.	158
Figura 5.7 – Tabela com campos do <i>Journal File</i> que podem ser preenchidos pelo usuário.	162
Figura 5.8 – Formulário que deve ser preenchido para cálculo do VPL pela rotina desenvolvida.	164
Figura 5.9 – Esquema do arquivo de texto que deve ser informado no formulário.	165
Figura 5.10 – Exemplo de cálculo do VPL pela rotina desenvolvida.	167
Figura A.1 – Esquema da busca de dados para estimativa no <i>gridsim</i> por Vizinho Mais Próximo. Os locais das amostras adicionais no <i>gridsim</i> recebem valor 1.	193
Figura A.2 – Exemplo da construção do <i>grid3</i> e escolha dos nós que devem receber sondagem.	197
Figura A.3 – Definição do ângulo de rotação.	202
Figura C.1 – Histogramas considerando pesos de desagrupamento obtidos pelo programa <i>Declus</i> . (a) Histograma para a Configuração 1. (b) Histograma para a Configuração 2. (c) Histograma para a Configuração 3. (d) Histograma para a Configuração 4.	215

Lista de Tabelas

Tabela 3.1 - Modelo variográfico utilizado na simulação para os dados no espaço original e normalizado	64
Tabela 3.2 – Teste de binormalidade para os dados iniciais	65
Tabela 3.3 – Estatística para realizações no suporte de blocos	67
Tabela 3.4 – Incerteza sobre a tonelagem acima do teor de corte, para diferentes configurações de amostras e quatro teores de corte	79
Tabela 3.5 – Mínima e máxima tonelagem acima do teor de corte, para diferentes configurações de amostras e quatro teores de corte	81
Tabela 3.6 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras	81
Tabela 3.7 – Mínimo e máximo VPL (Rota 1), para diferentes configurações de amostras	82
Tabela 3.8 – Incerteza sobre o VPL (Rota 2), para diferentes configurações de amostras	82
Tabela 3.9 – Mínimo e máximo VPL (Rota 2), para diferentes configurações de amostras	82
Tabela 3.10 – Incerteza sobre a tonelagem acima do teor de corte, para diferentes configurações de amostras e quatro teores de corte	86
Tabela 3.11 – Mínima e máxima tonelagem acima do teor de corte, para diferentes configurações de amostras e quatro teores de corte	86
Tabela 3.12 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras	87
Tabela 3.13 – Mínimo e máximo VPL (Rota 1), para diferentes configurações de amostras	87
Tabela 3.14 – Incerteza sobre o VPL (Rota 2), para diferentes configurações de amostras	87
Tabela 3.15 – Mínimo e máximo VPL (Rota 2), para diferentes configurações de amostras	87

Tabela 3.16 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras	89
Tabela 3.17 – Mínimo e máximo VPL (Rota 1), para diferentes configurações de amostras	89
Tabela 3.18 – Incerteza sobre o VPL (Rota 2), para diferentes configurações de amostras	89
Tabela 3.19 – Mínimo e máximo VPL (Rota 2), para diferentes configurações de amostras	89
Tabela 3.20 – Incerteza sobre a tonelagem acima do teor de corte, para diferentes configurações de amostras e 4 teores de corte	93
Tabela 3.21 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras	94
Tabela 3.22 – Incerteza sobre o VPL (Rota 2), para diferentes configurações de amostras	95
Tabela 3.23 – Incerteza sobre a tonelagem acima do teor de corte, para diferentes configurações de amostras e 4 teores de corte	98
Tabela 3.24 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras	98
Tabela 3.25 – Incerteza sobre o VPL (Rota 2), para diferentes configurações de amostras	98
Tabela 3.26 – Incerteza sobre o VPL para diferentes números de amostras adicionais	99
Tabela 3.27 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras	105
Tabela 3.28 – Incerteza sobre o VPL (seqüência ótima), para diferentes configurações	108
Tabela 3.29 – Incerteza sobre o VPL (seqüência ótima), para diferentes configurações	108
Tabela 3.30 – Incerteza sobre o VPL (seqüência ótima), para diferentes configurações	111
Tabela 3.31 – Incerteza sobre o VPL (seqüência ótima), para diferentes configurações	111
Tabela 4.1 – Estatística para o atributo comprimento (m) das amostras obtidas após a regularização	116

Tabela 4.2 – Estatística para os atributos Cu (%) e comprimento (m) das amostras obtidas após a regularização. As amostras foram divididas em dois grupos: amostras com comprimento maior que 1,5 m e com comprimento menor que 1,5 m	117
Tabela 4.3 - Modelos de continuidade espacial para os dados originais e normalizados	118
Tabela 4.4 – Teste de binormalidade	125
Tabela 4.5 – Estatística de teores de Cu (%) para realizações no suporte de blocos	126
Tabela 4.6 – Estatística para o VPL	132
Tabela 4.7 – Estatística para o número de blocos acima do teor 1%	136
Tabela 4.8 – Estatística para o VPL	139
Tabela 4.9 – Estatística para o número de blocos acima do teor 1%	140
Tabela 5.1 – Tempo de processamento para o cálculo da média do índice de incerteza na direção da sondagem	172
Tabela 5.2 – Tempo de processamento para a simulação geoestatística	173
Tabela A.1 – Esquema de arquivo para o <i>grid1</i> apresentado na Figura 5.2	192
Tabela A.2 – Esquema de arquivo para o <i>grid3</i> apresentado na Figura A.2	197
Tabela A.3 - Lista com o nome das variáveis definidas no arquivo <i>Samples3D</i>	199

Capítulo 1 - Introdução

1.1 - Introdução

Comumente, na mineração, a coleta de amostras é realizada por meio de sondagem. Uma amostra coletada por sondagem é analisada quimicamente e essa análise fornece, com uma certa precisão, o valor de um ou mais atributos de interesse no local amostrado. Contudo, a coleta de amostras por sondagem e a análise química dessas amostras envolvem elevados custos, por isso, duas questões importantes envolvem a coleta de novas amostras por sondagem: a relação custo/ benefício para um determinado número de sondagens adicionais e a locação das novas sondagens. Essas duas questões dependem do objetivo da amostragem. O objetivo da amostragem pode estar relacionado, por exemplo, com alguma função de transferência. Em geoestatística e planejamento de lavra várias operações são representadas genericamente por funções. Entre estas funções (também chamadas funções de transferência), algumas são usadas para atribuir um valor econômico ao empreendimento, como por exemplo, o valor presente líquido (VPL). Cada função considera o valor de um ou mais atributos (e.g. teor de metal), ao longo do depósito. Em geral, quanto maior o número de amostras, maior a certeza sobre a distribuição do atributo ao longo do depósito e, assim, maior a certeza sobre a função. Por isso, o objetivo da coleta de novas informações pode ser reduzir a incerteza sobre a função.

O VPL mostra o resultado financeiro de um empreendimento com base em receitas e despesas previstas durante um período de mineração. Isso auxilia na decisão sobre iniciar ou continuar a mineração do depósito. E, além disso, o VPL pode modificar o valor da empresa detentora dos direitos minerários no mercado financeiro, porque o lucro que pode ser obtido pela empresa influencia no valor de suas ações. Outras funções, como a tonelagem de minério, podem influenciar (além

do valor financeiro da empresa) o valor da *commodity* relativa ao depósito, no mercado mundial, porque o aumento das reservas mundiais de um bem mineral pode reduzir o preço de mercado desse bem. Ou seja, as funções que caracterizam o depósito e o empreendimento podem exercer influência no mercado financeiro em escala mundial, mas para que os valores dessas funções sejam aceitos publicamente é preciso que se tenha um nível de confiança aceitável sobre esses valores. Quando se fala em confiança, refere-se à elevada probabilidade do valor publicado da função estar correto, ou seja, próximo do real. Assim, o valor publicado pode ser modificado com a redução da incerteza sobre a função (aumento da confiança).

Por exemplo, as possíveis reservas de um depósito de carvão estão entre 5×10^6 t e 40×10^6 t, e só pode-se dizer, com a confiança necessária, que a reserva é superior a 5×10^6 toneladas. Isso pode representar uma sub-estimativa significativa no valor financeiro da empresa, para uma reserva de 23×10^6 toneladas. Porém, se a faixa de incerteza sobre a função é reduzida e os possíveis valores de reserva estão, agora, entre 20×10^6 t e 25×10^6 t, pode-se dizer que a reserva é superior a 20×10^6 toneladas. Nesse caso, o valor financeiro da empresa se torna mais próximo do que deve ser. Esse é um exemplo da importância da redução da incerteza sobre o valor da função de interesse.

Cada função usada para caracterizar o depósito ou empreendimento depende da distribuição dos valores de um ou mais atributos ao longo do depósito. Por exemplo, o VPL de um depósito de cobre está relacionado com o teor de cobre para cada bloco de minério que será lavrado, visto que quanto maior o teor de cobre em um bloco, mais metal será recuperado e maior será a receita associada àquele bloco. Outro exemplo é a dependência do VPL de um depósito de carvão em relação aos atributos espessura da camada e de sua densidade. O produto espessura da camada de carvão pela densidade da camada, para cada bloco, produz a tonelagem de carvão, e o valor econômico atribuído a cada bloco depende da tonelagem de carvão que será obtida com a lavra e beneficiamento do bloco. Considerando a dependência do valor da função a um ou mais atributos do depósito, pode-se dizer que a incerteza sobre o valor da função depende da incerteza sobre a distribuição dos valores do atributo ao longo do depósito que influenciam esta função.

A incerteza sobre a distribuição dos valores do atributo de interesse (atributo

considerado no valor da função) ao longo do depósito está relacionada com o nível de informação sobre essa distribuição, informação essa, que provém das amostras coletadas ao longo do depósito. Ou seja, para se reduzir a incerteza sobre a distribuição dos valores do atributo é necessário a coleta de novas amostras.

Nem sempre o aumento de informação sobre a distribuição dos valores do atributo garante um benefício econômico maior que o custo envolvido na coleta e análise química das novas amostras. Por exemplo, considere que a reserva de um depósito de carvão era superior a 20×10^6 t e após a nova campanha de amostragem passa a ser superior a $20,1 \times 10^6$ toneladas. A variação da reserva em $0,1 \times 10^6$ pode não aumentar o valor financeiro da empresa no mercado, embora a empresa tenha empregado recursos financeiros na nova campanha de amostragem. Por isso, a relação custo/ benefício deve ser avaliada para o número de sondagens que se deseja adicionar. Convém lembrar, que o benefício da amostragem adicional depende da função ou parâmetro que está sendo analisado. No exemplo mencionado, a amostragem adicional pode não trazer benefício na valorização da empresa no mercado financeiro, porém, pode reduzir a incerteza sobre o teor de cinza em alguns blocos. Isso pode melhorar a seleção dos blocos que devem ser beneficiados, e talvez, reduzir o custo com o beneficiamento. A redução do custo de beneficiamento pode compensar o custo envolvido na amostragem.

Outra questão relacionada ao benefício da amostragem, além do número de sondagens adicionais, é a locação espacial destas novas sondagens. Pode ser, que para um mesmo número de amostras adicionais, ou seja, para um mesmo custo, uma certa configuração de amostras adicionais forneça mais benefício do que outra. Comumente, após a realização de sondagens iniciais (por exemplo, sondagens coletadas em uma primeira campanha de amostragem, distribuídas regularmente ao longo do depósito), novas sondagens são realizadas formando uma malha aproximadamente regular com as sondagens iniciais, tentando contemplar igualmente todas regiões do depósito. Contudo, a incerteza sobre os valores do atributo de interesse pode ser maior em algumas regiões do depósito do que em outras. Por exemplo, a amplitude do intervalo de possíveis valores do atributo em uma região é maior do que em outras. Quando isso ocorre, a adição das sondagens nas regiões onde a incerteza sobre o valor do atributo de interesse é maior pode ser mais eficiente na redução de incerteza sobre uma função (relacionado com o atributo em questão) do que a adição das sondagens formando uma malha regular. Ou seja,

um esquema de locação de sondagens pode ser mais eficiente do que o outro na redução da incerteza sobre uma função de interesse, porém, se existe ou não diferença significativa entre os resultados obtidos para os dois esquemas e os fatores que influenciam essa diferença são questões que devem ser estudadas.

Convém lembrar que a escolha de um esquema de locação de amostras depende do objetivo da amostragem. Por exemplo, quando a classificação de recursos é baseada na distância mínima entre um bloco do modelo geológico e a sondagem, as novas sondagens devem ser locadas de maneira a minimizar a distância máxima ou distância média entre os blocos e as sondagens. O objetivo da amostragem também pode ser definir os limites do corpo de minério, as novas sondagens podem ser adicionadas nas bordas do modelo geológico. Nesses casos, o melhor esquema de locação de amostras a fim de aumentar os recursos medidos ou delimitar o corpo de minério pode ser diferente do melhor esquema de locação de amostras para reduzir a incerteza sobre uma função. No contexto dessa tese, o objetivo da amostragem é reduzir a incerteza sobre uma função de transferência.

Até aqui, se falou nos objetivos da redução de incerteza sobre uma função de interesse e nas configurações de amostras que podem ser eficientes na redução dessa incerteza. Contudo, para se avaliar a eficiência de configurações de amostras na redução de incerteza sobre uma função, essa incerteza deve ser acessada.

A geoestatística fornece algoritmos de simulação (Matheron, 1973; Journel, 1974; Isaaks, 1990; Alabert, 1987) que permitem o acesso à incerteza sobre uma função de interesse. Algoritmos de simulação geoestatística fornecem possíveis cenários para a distribuição dos valores do atributo de interesse, baseando-se nas informações do atributo obtidas a partir de amostragem. Cada cenário gerado possui um valor para a função. Assim, os diversos cenários geram uma distribuição dos possíveis valores da função, e uma medida do espalhamento dessa distribuição pode ser usada como aproximação da incerteza sobre o valor real da função. Além disso, os algoritmos de simulação fornecem uma distribuição dos possíveis valores do atributo em cada local não amostrado. O conhecimento dessa distribuição permite que se conheçam os locais de maior incerteza sobre o valor do atributo de interesse.

Resumindo os aspectos abordados, o objetivo da amostragem pode ser reduzir a incerteza sobre uma função, visto que a redução dessa incerteza pode

trazer benefícios econômicos para o minerador. Um número de amostras adicionais pode reduzir significativamente a incerteza sobre a função. Contudo, para um mesmo número de amostras adicionais, a locação das novas amostras deve ser estudada a fim de garantir o máximo benefício em relação aos objetivos. Nesse contexto, essa tese estuda a relação entre os esquemas de locação de amostras mencionados com a redução da incerteza sobre a função de transferência, usando simulação geoestatística para o acesso à incerteza.

1.2 - Meta

Conforme explicado anteriormente, quando n sondagens devem ser adicionadas às sondagens iniciais, a configuração de amostras em que as n sondagens são posicionadas nas regiões de elevada incerteza sobre o valor do atributo de interesse (chamada Configuração 2) pode ser mais eficiente, na redução da incerteza sobre uma função, do que a configuração de amostras em que as n sondagens são adicionadas em uma malha regular (Configuração 1). Uma das metas dessa tese é analisar a eficiência de cada tipo de configuração de amostras. A comparação dos resultados obtidos para as diferentes configurações é feita a fim de verificar se a diferença entre os resultados é significativa. Os fatores que influenciam a diferença entre os resultados obtidos para ambas configurações também são analisados.

A fim de melhorar a compreensão do texto, alguns termos mencionados anteriormente devem ser definidos no contexto da tese:

- “sondagens iniciais” - Representam sondagens coletadas em uma primeira campanha de amostragem (com grande distância entre sondagens, considerando a continuidade do atributo de interesse), visando uma primeira avaliação de recursos e, geralmente, distribuídas regularmente ao longo de todo depósito. Essas sondagens são comuns às Configurações 1 e 2. A chamada Configuração Inicial é formada pelas amostras ou dados obtidos pelas sondagens iniciais;
- “eficiência” - A menos que se diga o contrário, entende-se por eficiência (ou

desempenho) de uma Configuração A: a diferença entre a incerteza sobre a função obtida a partir da Configuração Inicial e a incerteza sobre a função obtida a partir da Configuração A. Sendo que a Configuração A é composta pelas amostras iniciais e uma quantidade pré-definida de amostras adicionais. Em geral, quanto maior a diferença entre a incerteza obtida a partir da Configuração Inicial e a obtida a partir da Configuração A, maior a eficiência da Configuração A.

A diferença entre os desempenhos das duas configurações de amostras pode variar de acordo com o comportamento do atributo de interesse ao longo do depósito em estudo. Assim, o ideal é que a eficiência das diferentes configurações seja analisada em cada caso, não apenas para auxiliar na escolha de uma configuração para próxima campanha de amostragem, como para analisar se o custo com a adição de um determinado número de amostras é recompensado. Muitas vezes, na indústria mineira, a decisão a respeito do número e posição das amostras adicionais deve ser realizada em um curto espaço de tempo. A automação da metodologia para análise da eficiência das duas configurações permite uma rápida aplicação dessa metodologia. Assim, outra meta dessa tese é a elaboração de um algoritmo que permita a construção das configurações de amostras, automaticamente, em softwares comuns do ramo da mineração e a elaboração de uma rotina computacional que agilize o cálculo da incerteza sobre a função de transferência.

1.3 - Objetivos

Os objetivos relacionados às metas propostas são listados a seguir:

i. Analisar a eficiência de cada configuração (1 e 2) para o caso bi-dimensional e tri-dimensional, sendo que bi-dimensional é o caso onde os valores do atributo podem ser identificados por, apenas, duas coordenadas. Por exemplo, cada valor do atributo espessura da camada de carvão pode ser identificado pelas coordenadas leste (x) e norte (y), visto que cada valor representa a espessura da camada para

uma sondagem iniciada nas coordenadas x e y da superfície do terreno. A análise da eficiência das configurações para os casos bi e tri-dimensional permite exemplificar a aplicação da metodologia e pesquisar possíveis falhas dessa metodologia para ambos casos;

ii. Desenvolver uma técnica para considerar a incerteza ao longo das três dimensões na locação das novas sondagens. Em cada local simulado para os dados iniciais se obtém uma distribuição dos possíveis valores do atributo. A partir dessa distribuição, pode-se aproximar a incerteza sobre o valor do atributo em cada local. Assim, se conhecem os locais de maior incerteza sobre o valor do atributo, o que é necessário para construção da Configuração 2. As medidas usadas para aproximar a incerteza sobre o valor do atributo, ou a incerteza sobre o valor do atributo estar acima de um teor de corte, serão chamadas índices de incerteza (Pilger et al., 2000). Para o caso bi-dimensional, as novas sondagens podem ser locadas nas regiões do plano x - y com elevada incerteza sobre o valor do atributo. Contudo, para o caso tri-dimensional as sondagens devem ser locadas de forma a contemplar um grande número de locais, ao longo das três dimensões, com elevada incerteza, por isso, é necessário desenvolver uma técnica para considerar a incerteza ao longo das três dimensões na locação das novas sondagens;

iii. Elaborar um algoritmo computacional para construção de configurações de amostras que permita o uso de uma distância mínima definida pelo usuário. As configurações de amostras adicionais devem obedecer à condição de distância mínima entre amostras, para que o agrupamento dessas não leve à baixa eficiência da campanha de amostragem (informações redundantes que não contribuem para redução da incerteza a respeito dos valores simulados), como já havia sido exemplificado por Van Meirvenne e Goovaerts (2001);

iv. Analisar o impacto, no valor da incerteza, causado pela escolha de diferentes cenários como fornecedores de cada informação "real". Cada configuração de amostra adicional compreende os locais das sondagens iniciais e das sondagens adicionais. Na etapa de simulação, para cada configuração testada, deve se conhecer os valores do atributo nos locais dos dados iniciais e nos locais das amostras adicionais. Os valores nos locais das amostras adicionais são obtidos a partir de um cenário simulado para os dados iniciais, escolhido como referência. Ou seja, valores previamente simulados são utilizados para prover a informação "real"

que seria obtida se amostrássemos nestes locais. Para que a incerteza sobre a função, obtida pela metodologia, reproduza o que ocorreria na realidade, a incerteza deve ser semelhante para diferentes cenários escolhidos como referência. Por isso, deve-se analisar o impacto, no valor da incerteza, causado pela escolha de diferentes cenários como fornecedores de cada informação “real”.

1.4 - Metodologia

Uma das metas desse trabalho é a análise do desempenho de diferentes configurações de amostras na redução da incerteza sobre a função de interesse. Para a realização dessa análise, a metodologia seguida pode ser dividida nas seguintes etapas (Figura 1.1):

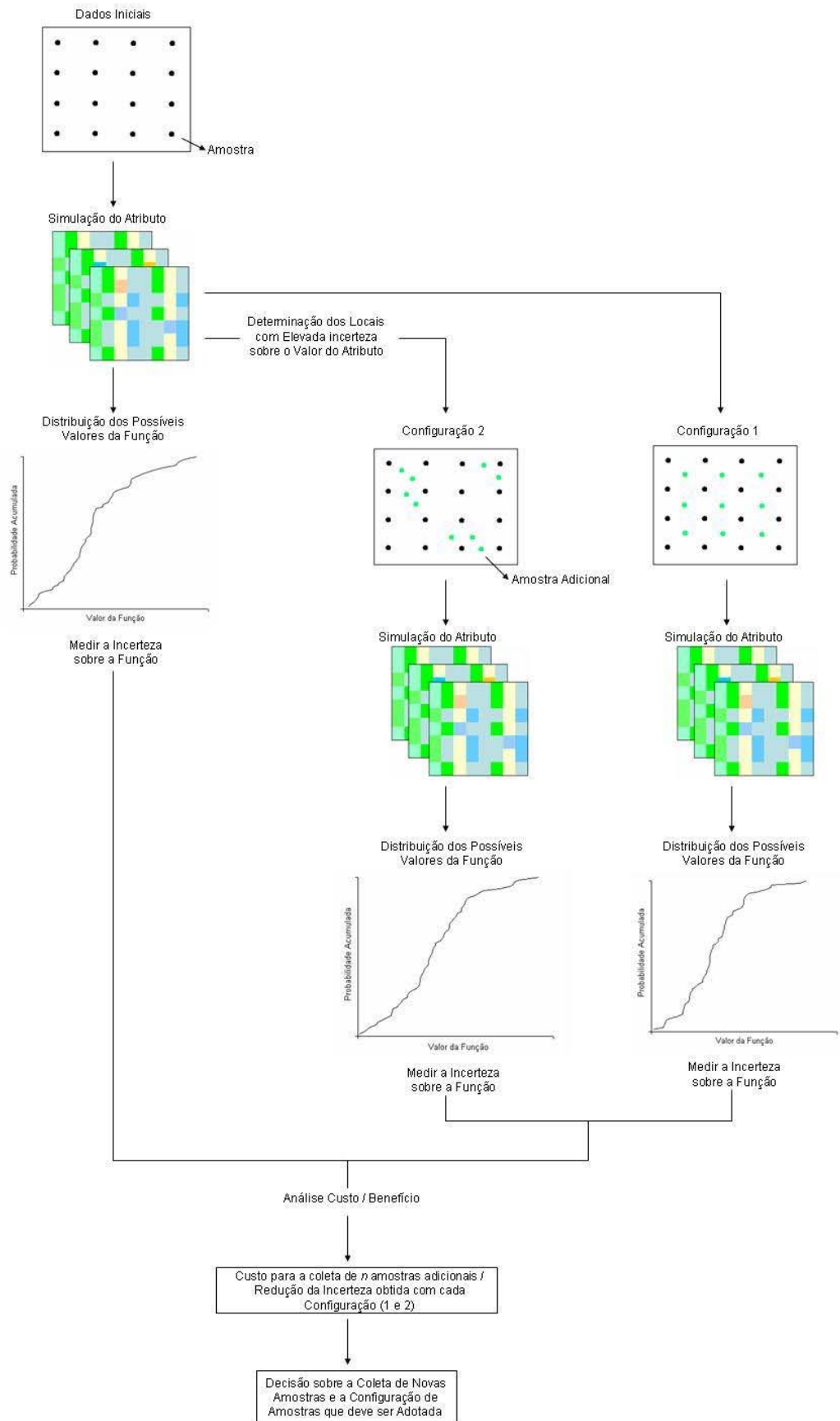


Figura 1.1 – Esquema da metodologia utilizada para análise de diferentes configurações de amostras.

- Definição da função de transferência cuja incerteza associada deseja-se reduzir;
- Realização da análise estatística. Considerando os dados iniciais, realizar a análise estatística dos valores do atributo de interesse e o mapeamento da continuidade desse atributo ao longo da região em estudo, visto que essas informações devem ser utilizadas no processo de simulação. O atributo de interesse é aquele que deve ser considerado no valor da função de transferência;
- Geração de diversos cenários do atributo de interesse, usando os dados iniciais, por simulação geoestatística;
- Cálculo do valor da função para cada cenário gerado, obtendo uma distribuição dos possíveis valores da função. Verificação da incerteza associada a essa distribuição. Essa incerteza será comparada com a incerteza obtida após a campanha de adensamento amostral;
- Utilização de um índice de incerteza. Com a simulação geoestatística, diversos possíveis valores do atributo de interesse são simulados em um dado local. Deve-se utilizar um índice de incerteza (Pilger et al., 2000) para medir a incerteza associada aos valores simulados em cada local;
- Definição do número (n) de amostras que serão adicionadas na próxima campanha de adensamento amostral;
- Construção de uma configuração de amostras (Configuração 1) contendo os dados iniciais e n amostras adicionadas em uma malha aproximadamente regular;
- Construção de uma configuração de amostras (Configuração 2) contendo os dados iniciais e n amostras adicionadas nos locais com elevada incerteza para o atributo de interesse. Esses locais são escolhidos de acordo com o valor do índice de incerteza em cada local simulado. Os locais das amostras adicionais respeitam uma distância mínima entre amostras;
- Escolha de um cenário simulado com os dados iniciais para servir como fonte de informações adicionais ou referência. Ou seja, os valores das amostras adicionadas nas configurações 1 e 2 são obtidos a partir do cenário que foi escolhido como referência;
- Geração de diversos cenários do atributo de interesse, usando os dados da configuração 1, por simulação geoestatística;

- Geração de diversos cenários do atributo de interesse, usando os dados da configuração 2, por simulação geoestatística;
- Análise da incerteza associada aos valores da função de transferência obtida para a configuração inicial, para a configuração 1 e para a configuração 2. Essa análise permite a escolha da configuração de amostras que deve ser construída na próxima campanha de amostragem e permite avaliar se o benefício da nova campanha de amostragem é significativo.

Essa metodologia poderá ser utilizada em diversas situações onde adensamento de sondagens seja necessário e onde seja viável a utilização do uso de algoritmos de simulação geoestatística. Contudo, algumas suposições e dificuldades da aplicação da metodologia são comentadas a seguir:

- Para aplicação da metodologia, o depósito deve ter sido previamente sub-amostrado e deve se ter conhecimento sobre a continuidade do atributo de interesse;
- A locação de sondagens para construção da Configuração 2 é guiada pelas regiões de elevada incerteza de um atributo. Em muitos depósitos minerais, o atributo de interesse pode estar presente em diferentes domínios geológicos, além disso, a distribuição dos valores e a continuidade espacial do atributo podem ser diferentes em cada domínio. Por isso, os domínios geológicos devem ser simulados separadamente, e as regiões de elevada incerteza são mapeadas para cada domínio geológico. Nesses casos, aumenta a dificuldade da construção da Configuração 2. Por exemplo, considerar regiões de elevada incerteza, apenas, para o domínio mais importante economicamente ou contemplar regiões de elevada incerteza para diferentes domínios? Para o caso tri-dimensional, quando os domínios são um acima do outro, deve se locar cada sondagem tentando contemplar, simultaneamente, regiões de elevada incerteza para diferentes domínios ou considerar apenas o domínio mais importante economicamente? As mesmas questões surgem no caso em que existe mais de um atributo de interesse. A locação das novas sondagens deve considerar regiões de elevada incerteza para os diferentes atributos? Respostas para essas questões podem ser encontradas

construindo a Configuração 2 de acordo com algumas alternativas e avaliando a eficiência de cada alternativa. Porém, é importante ressaltar que essas dificuldades encontradas na aplicação da metodologia não são abordadas na pesquisa realizada nessa tese. Para cada caso estudado ao longo da tese, são consideradas as regiões de elevada incerteza de um único atributo, pertencentes a um único domínio geológico.

Convém ressaltar que essa metodologia pode, além de auxiliar na escolha da melhor configuração de amostras (1 ou 2), indicar o número mínimo de amostras que deve ser adicionado para se alcançar o nível de incerteza desejado sobre a função. Ou seja, a metodologia pode ser aplicada diversas vezes, considerando diferentes números de amostras adicionais, a fim de se definir o número mínimo de amostras adicionais que produz o nível de incerteza necessário sobre a função (para uma dada configuração).

Adicionalmente, é importante comentar a possibilidade do uso da metodologia para melhorar a classificação de recursos de um depósito. Segundo alguns códigos para reportar recursos e reservas, como JORC (AusIMM, 2004), a classificação de recursos é baseada no nível de incerteza sobre a geologia do depósito. No sentido do menor para maior nível de certeza, os recursos podem ser classificados como inferido, indicado ou medido, respectivamente. Nessa tese, avalia-se a incerteza sobre uma função para diferentes configurações de amostras. Essa função é baseada em atributos geológicos (ex. a função que expressa tonelagem de carvão é baseada em estimativas de espessura da camada de carvão). Assim, o nível de incerteza sobre a função, que reflete a incerteza sobre um ou mais atributos geológicos, pode ser utilizado na classificação de recursos. Baseando-se nessa idéia, ao aplicar a metodologia abordada na tese, o empreendedor pode avaliar se a redução da incerteza sobre a função, obtida com a coleta de novas amostras, é capaz de melhorar a classificação de recursos e avaliar qual configuração de amostras fornece a melhor classificação. A locação de amostras adicionais e o cálculo da função podem considerar, apenas, uma região do depósito previamente classificada como inferida ou indicada. Nesse caso, o estudo mostra se o aumento do número de amostras pode melhorar a classificação dos recursos dessa região.

1.5 - Estado da Arte

Diversos esquemas de locação de amostras têm sido propostos ao longo do tempo com diferentes objetivos. Esquemas e metodologias originados em diferentes áreas do conhecimento podem ser aplicados em todas as ciências que compartilham o problema de locação de amostras, como, meio ambiente (e.g. coleta de amostras de contaminantes) e mineração (e.g. coleta de amostras para análise de teores de minério). Contudo, algumas metodologias são baseadas na teoria das variáveis regionalizadas (Matheron, 1963) e, por isso, devem ser aplicadas, apenas, para estes tipos de variáveis.

Quando a amostragem tem o objetivo de pesquisa inicial, o uso de amostragem regular é comumente aceito, visto que pode contemplar todas as regiões do depósito com probabilidade semelhante. Myers (1996) exemplifica alguns esquemas de amostragem regular, chamadas pelo autor de Amostragem Sistemática. O autor também exemplifica a Amostragem Aleatória Estratificada, onde a posição de cada amostra é definida de maneira aleatória dentro de limites previamente estabelecidos.

Muitos trabalhos abordam a questão da locação de amostras (amostras iniciais ou adicionadas a dados previamente coletados) com o objetivo de reduzir a incerteza sobre as estimativas do atributo, que serão obtidas a partir dos valores amostrados. Algumas medidas de incerteza sobre a estimativa não consideram os valores das amostras que serão coletadas, apenas, a posição dessas amostras. Assim, a malha de amostragem é construída visando alcançar o nível de incerteza desejado. Exemplos dessas medidas são: variância de extensão (Matheron, 1963) e variância de krigagem. Porém, essas medidas utilizam informações do modelo variográfico (modelo de continuidade) da variável de interesse, assim, um modelo deve ser inferido para a região em estudo. Em alguns casos, a chamada cruz geoestatística pode ser utilizada para auxiliar na obtenção do modelo variográfico. A cruz geoestatística pode ser entendida como uma primeira amostragem, envolvendo a coleta de poucas amostras ao longo de duas direções. Andriotti (2003) explica que o número de amostras não deve ser muito grande, mas suficiente para se definir o modelo de continuidade, e sugere que as direções amostradas devam coincidir com

as direções de maior e menor variabilidade do atributo (essas direções são inferidas a partir das informações sobre a origem do corpo de minério). Em geral, a cruz geoestatística é realizada para auxiliar na definição de uma malha de amostragem que envolve toda região em estudo e um grande número de amostras. No contexto dessa Tese, os dados coletados com a cruz geoestatística e essa primeira malha de amostragem são chamados dados iniciais. A construção da Configuração 1 e da Configuração 2 ocorre após a coleta dos dados iniciais.

Armstrong (1998) apresenta um exemplo do uso da variância de extensão para a otimização do tamanho da malha de amostragem.

Matheron (1963) apresentou o algoritmo de krigagem, após desenvolver estudos iniciados por Krige (1951). A krigagem é capaz de fornecer uma estimativa do valor do atributo de interesse em um dado local não amostrado a partir das amostras existentes no entorno. A variância do erro da estimativa de krigagem é a chamada variância de krigagem. Embora a variância de krigagem seja referida ao longo dessa tese como uma medida de incerteza da estimativa, deve se considerar que essa medida apresenta uma limitação no acesso à incerteza, essa limitação é comentada a seguir.

McBratney et al. (1981) observaram o uso da variância de krigagem para otimização do tamanho da malha de amostragem, e relatam que essa variância não depende dos valores das amostras utilizadas na estimativa. Esses autores comentam o uso do valor máximo de variância de krigagem como critério de otimização. McBratney et al. (1981) sugerem ainda a coleta de amostras ao longo de algumas direções, na região em estudo, a fim de obter um modelo variográfico para a variável. Seguindo o uso da variância de krigagem, McBratney and Webster (1981) apresentam um programa relacionando algumas configurações de amostras e variância de krigagem.

Olea (1984) propõe um algoritmo para escolha de uma configuração de amostras considerando a variância de krigagem universal. Isaaks e Srivastava (1989) comentam que quando a variabilidade dos valores do atributo de interesse depende da magnitude desses valores, a medida de incerteza local deve considerar esse fato. Por isso, Isaaks e Srivastava (1989) apresentam um exemplo com o uso da variância de krigagem e considerando essas condições locais. Atkinson e Lloyd (2007) também abordam essa questão para determinar o espaçamento entre

amostras.

Outro exemplo do uso de variância de krigagem para otimização de esquemas de amostragem é apresentado em Van Groenigen et al. (1999) e Van Groenigen (2000). Esses autores utilizam *Spatial Simulated Annealing* (SSA). Van Groenigen et al. (1999) comentam que SSA é baseado no algoritmo de otimização chamado Simulação Annealing (SA).

Muitas vezes, a estimativa de um atributo considera o modelo de continuidade dos valores do atributo, como no caso da krigagem. A locação de amostras pode visar a melhoria das futuras estimativas pela obtenção de um variograma melhor definido (melhor mapeamento da continuidade). Warrick e Myers (1987) apresentam um algoritmo de locação de amostras para melhorar a definição de variogramas experimentais.

Suponha que estimativas para valores do atributo já foram realizadas, baseadas nos valores de amostras previamente coletados. O objetivo da nova campanha de amostragem é melhorar as estimativas que apresentam elevada incerteza. Ou seja, as novas amostras devem ser locadas nas regiões de elevada incerteza sobre o valor do atributo. Assim, uma medida de incerteza é usada para guiar a locação de amostras. Conforme comentado, vários autores (e.g. Journel, 1986; Isaaks e Srivastava, 1989) observaram que a variância de krigagem não considera no seu cálculo os valores dos dados, apenas a configuração desses e o modelo variográfico. Nesse contexto, outras medidas de incerteza foram propostas, a chamada variância de interpolação (Froidevaux, 1993; Yamamoto e Conde, 1999) e a variância combinada (Arik, 1999).

Rocha e Yamamoto (2000) apresentam um estudo comparando o desempenho da variância de krigagem e da variância de interpolação para medir a incerteza associada às estimativas realizadas por krigagem ordinária. Para o estudo apresentado, os autores comentam que a variância de interpolação foi uma medida mais confiável em relação à variância de krigagem. Souza (2007) utiliza a incerteza associada à estimativa de um bloco para classificá-lo como medido, indicado ou inferido. O autor usa as variâncias de krigagem, de interpolação e combinada para medir a incerteza associada à estimativa e os resultados mostram que a classificação do bloco pode modificar com a variância utilizada.

Semelhante a idéia de adicionar amostras nos locais onde as estimativas

apresentam maior incerteza é a idéia proposta por McBratney et al. (1999). Esses autores propõem a adição de mais amostras nas regiões com maior variabilidade dos valores do atributo. Os autores apresentam um algoritmo que auxilia nesse procedimento e um modelo previamente gerado para os valores do atributo é usado no cálculo da variabilidade.

Existem casos onde é necessário saber, apenas, se o valor do atributo em um local é acima de um dado limite. Nesses casos, deve-se avaliar a incerteza do valor ser acima deste limite. Alguns trabalhos abordam a locação de amostras segundo esses valores de incerteza. Journel (1988) propõe a locação de amostras adicionais por meio da geoestatística não-paramétrica. O autor apresenta um exemplo do uso de krigagem da probabilidade para conhecer os locais onde a incerteza sobre o valor do atributo estar acima de um limite é elevada.

Ainda, ilustrando o uso de geoestatística não-paramétrica para obter a distribuição dos valores do atributo em cada local, Van Meirvenne e Goovaerts (2001) adicionam amostras simultaneamente em regiões de elevada incerteza e consideram uma distância mínima entre as novas amostras para aumentar a eficiência da amostragem.

Os locais de maior incerteza sobre o valor do atributo também podem ser mapeados por simulação geoestatística. Algoritmos de simulação geoestatística fornecem vários possíveis cenários para a distribuição dos valores do atributo e, assim, uma distribuição dos possíveis valores para o atributo em cada local simulado. Por meio dessa distribuição, pode se acessar a incerteza sobre o valor do atributo ou a probabilidade desse valor estar acima de um limite. Pilger (2000) utilizou simulação geoestatística para acessar a incerteza em cada local. A incerteza foi quantificada por diferentes índices de incerteza: variância condicional, coeficiente de variação e intervalo entre quartis. Cada sondagem adicional foi locada de acordo com os valores do índice de incerteza. Pilger et al. (2002) adicionam amostras, uma de cada vez, nas regiões de elevada incerteza e mostram o efeito na redução da incerteza, após a adição de cada informação, nos recursos medidos de um depósito de carvão.

Em Pilger (2000) e Pilger et al. (2002), as informações foram adicionadas uma de cada vez, ou seja, após a inserção de cada informação adicional um novo processo de simulação foi realizado e os índices de incerteza recalculados. Um dos

cenários simulados foi utilizado como fonte de informação sobre os valores do atributo. Contudo, dependendo do número de informações que se pretende adicionar na próxima campanha de amostragem, a geração de cenários simulados após a adição de cada amostra pode exigir muito tempo, o que pode dificultar a aplicação da metodologia na indústria da mineração, principalmente quando a simulação deve ser feita para o caso tri-dimensional. A adição de várias amostras de cada vez pode facilitar a aplicação da metodologia. Richmond et al. (2007) também utilizam simulação para mapear os locais de elevada incerteza e propõem um algoritmo que adiciona todas amostras, simultaneamente, nas regiões de elevada incerteza. Os autores utilizam uma distância mínima entre as novas amostras.

O objetivo da amostragem pode ser minimizar o valor de uma função econômica. Englund e Heravi (1992) apresentam um algoritmo para encontrar o número de amostras que minimiza uma função econômica (que expressa custo).

Diferente de minimizar o valor da função, o objetivo da amostragem pode ser reduzir a incerteza sobre uma função. Conforme explicado anteriormente, com o uso de simulação geoestatística pode se aproximar a incerteza associada ao real valor de uma função (e.g. VPL), que considera os valores do atributo simulado. Autores como Boucher et al. (2004) e Sullivan (2003) usam simulação para avaliar a influência da amostragem adicional na redução da incerteza sobre uma função. Sullivan (2003) avalia o desempenho de uma configuração de amostras na redução da incerteza sobre a tonelagem acima do teor de corte, além disso, apresenta um método para calcular a diferença no valor do VPL causada pela diferença de incerteza obtida com e sem amostras adicionais. Boucher et al. (2004) comparam o desempenho de diferentes malhas regulares (com diferentes espaçamento entre amostras e, assim, diferentes números de amostras) para amostragem adicional na redução de incerteza sobre a função.

Li et al. (2004, 2005) e Emery et al. (2008) também usam simulação para avaliar o desempenho de diferentes configurações de amostras. Li et al. (2004, 2005) utilizam um algoritmo para avaliar o desempenho de diferentes malhas regulares e ressaltam o fato de que, por causa das áreas com diferentes variabilidades para o atributo, diferentes malhas deveriam ser usadas em cada região. Emery et al. (2008) apresenta um algoritmo que otimiza a configuração de amostras usando simulação annealing. Ou seja, o algoritmo permite o teste de

configurações onde a posição das amostras é irregular. Os resultados são apresentados para o caso bi-dimensional. A incerteza sobre o valor de cada bloco é aproximada e a otimização da configuração considera o valor médio e máximo das incertezas para todos blocos.

1.5.1 - Contribuições da Tese

Considerando os trabalhos publicados sobre o tema semelhante ao da tese, algumas contribuições da tese são explicadas a seguir:

- Análise da diferença entre as configurações 1 e 2 - Uma das metas dessa tese é comparar o desempenho das configurações 1 e 2 na redução da incerteza sobre a função. A aproximação da incerteza sobre o atributo em cada local é realizada com o uso de simulação geoestatística, que fornece a distribuição para o atributo naquele local considerando os valores das amostras utilizadas na simulação, além da posição espacial e continuidade dos valores das amostras. Conforme comentado, alguns trabalhos têm usado simulação para avaliar o desempenho de um ou outro tipo de configuração na redução da incerteza sobre uma função como o VPL ou sobre algum parâmetro (e.g. média das incertezas sobre os valores dos blocos). Ou seja, alguns trabalhos comparam a eficiência entre malhas regulares, com diferentes densidades de amostras e outros mostram resultados, apenas, para a configuração que visa adicionar amostras nos locais de elevada incerteza. Contudo, quando se pretende investir em um número de sondagens adicionais já definido, deve-se avaliar se a diferença entre os desempenhos das configurações 1 e 2 é significativa e escolher entre as duas. Uma das metas dessa tese foi proposta a fim de descobrir se a diferença entre as configurações 1 e 2 pode ser significativa e auxiliar na escolha de uma dessas configurações;
- Implementação de um processo de otimização simplificado - Outras configurações de amostras, além das duas mencionadas (1 e 2), poderiam ser testadas antes de uma nova campanha de amostragem. O algoritmo apresentado por Emery (2008) permite o teste de diferentes configurações pelo processo de

otimização. Contudo, a comparação entre as configurações 1 e 2 é proposta para economia de tempo, principalmente para o caso tri-dimensional, onde um grande número de locais deve ser simulado. A idéia abordada nessa tese é que a configuração encontrada por um processo de otimização é uma das possíveis formas de construir as configurações 1 e 2. Por isso, esses dois tipos de configurações são estudados;

- Elaboração de um algoritmo para o caso 3D - Conforme comentado, os algoritmos propostos avaliam o desempenho de configurações para o caso bi-dimensional, uma das metas dessa tese é elaborar um algoritmo que permita a construção automática das configurações 1 e 2 para o caso tri-dimensional;
- Elaboração de um algoritmo que possa ser executado em larga escala - Mesmo com os trabalhos que têm sido realizados sobre análise da eficiência de configurações, essa análise ainda é pouco realizada na indústria mineira, uma das metas dessa tese visa elaborar um algoritmo para construção das configurações 1 e 2 que possa ser executado em softwares comuns do ramo da mineração. Dessa forma, os usuários desses softwares podem entender e executar o algoritmo quando precisarem.

Koppe et al. (2008a, 2008b, 2009) apresentam resultados oriundos da pesquisa realizada nessa Tese. Koppe et al. (2009) apresentam o algoritmo elaborado nessa tese para construção das configurações 1 e 2.

1.6 - Organização da Tese

Essa tese está dividida em 6 capítulos. Os conteúdos dos capítulos 2 a 6 são descritos a seguir:

- Capítulo 2 – Apresentação dos aspectos teóricos sobre algoritmos de simulação geoestatística que podem ser usados na metodologia adotada nessa tese. Um desses algoritmos foi escolhido para ser utilizado ao longo da tese. Em cada estudo

de caso, esse algoritmo foi utilizado antes e depois da inserção de amostras adicionais no banco de dados. Nesse capítulo, também, serão apresentados aspectos teóricos sobre os índices que podem ser utilizados para medir a incerteza a respeito do valor do atributo de interesse nos locais simulados. Os valores de um ou mais índices de incerteza são utilizados, ao longo da tese, como guia para a locação de amostras adicionais. Além disso, as funções de transferência, que são utilizadas ao longo da tese, são apresentadas nesse capítulo.

- Capítulo 3 – Análise da eficiência de diferentes configurações de amostras adicionais na redução de incerteza sobre uma função e os fatores que influenciam essa eficiência. A análise será ilustrada por um banco de dados bi-dimensional.
- Capítulo 4 – Análise da eficiência de diferentes configurações de amostras adicionais na redução de incerteza sobre uma função. A análise será ilustrada por um banco de dados tri-dimensional.
- Capítulo 5 – Apresentação do algoritmo que permite a construção automática das configurações de amostras, abordadas ao longo da tese, em um software comum do ramo da mineração. Nesse capítulo, também é apresentada uma rotina computacional criada para agilizar a aproximação do valor de incerteza sobre a função de transferência, obtido para cada configuração testada.
- Capítulo 6 – Apresentação das conclusões obtidas ao longo do trabalho e sugestões para trabalhos futuros baseados no tema da Tese.

Capítulo 2 – Simulação e Índices de Incerteza

Nesse trabalho, serão necessárias medidas de incerteza local a respeito dos valores do atributo de interesse além da geração de diversos cenários para distribuição dos valores desse atributo ao longo do depósito. As medidas de incerteza local serão utilizadas para definir os locais sujeitos à sondagem na próxima campanha de adensamento amostral, enquanto que os diversos cenários gerados alimentarão a função de transferência de interesse e possibilitarão a verificação da incerteza associada aos valores dessa função.

A geoestatística possui ferramentas, como os algoritmos de simulação estocástica, que são capazes de fornecer as medidas de incerteza local e os diversos cenários que serão necessários nesse trabalho. Nesse capítulo, encontram-se aspectos teóricos sobre alguns dos algoritmos de simulação geoestatística que podem ser usados na metodologia adotada nessa tese: simulação seqüencial Gaussiana (Isaaks, 1990) e simulação por bandas rotativas (Matheron, 1973).

Algoritmos de krigagem linear, como krigagem ordinária, também podem fornecer medidas da incerteza associada a cada estimativa por meio da variância de krigagem. Conforme explica Goovaerts (1997), a geração de uma distribuição gaussiana centrada no valor da estimativa e com variância igual à variância de krigagem é uma aproximação comumente utilizada para o modelo de incerteza associado à estimativa, contudo, o mesmo autor comenta as deficiências dessa aproximação.

O algoritmo de simulação seqüencial Gaussiana (SGS) é baseado em algoritmos de krigagem linear, contudo é mais robusto quanto às carências apresentadas por esses algoritmos no acesso à incerteza local e global. A simulação seqüencial Gaussiana é capaz de gerar diferentes cenários para a distribuição dos valores do atributo de interesse ao longo da área em estudo, cada um reproduzindo o modelo de continuidade espacial desse atributo. Esse algoritmo de simulação é classificado como paramétrico, visto que o modelo de incerteza associado aos

valores simulados depende de parâmetros como a média e a variância de krigagem.

Outro ramo da geoestatística também pode ser usado na obtenção de medidas da incerteza associada a cada estimativa: a geoestatística não-linear. Algoritmos de krigagem não-linear são capazes de fornecer a probabilidade acumulada para valores de um atributo em cada local interpolado. Baseando-se na distribuição de probabilidades obtida com a krigagem não-linear, medidas da incerteza associada a um dado valor do atributo (ou estimativa) podem ser geradas.

A simulação seqüencial dos indicadores (SIS) (Alabert, 1987) é baseada em um importante algoritmo de krigagem não-linear: krigagem dos indicadores (Journel, 1982). Ao contrário da simulação seqüencial Gaussiana, esse algoritmo de simulação é classificado como não-paramétrico. Goovaerts (1997), baseado em autores como Journel e Alabert (1988), explica que a simulação seqüencial Gaussiana não reproduz correlação significativa para valores extremos, nos casos em que essa correlação deve ser reproduzida, a simulação seqüencial dos indicadores é uma alternativa.

Assim, algoritmos de simulação geoestatística paramétrica e não-paramétrica podem ser escolhidos para realização da metodologia. Além disso, a simulação por bandas rotativas é um algoritmo de simulação não-seqüencial (porém paramétrica), ao contrário da SGS e SIS. Por isso, a metodologia adotada nesse estudo pode ser realizada com algoritmos teoricamente diferentes, e bastante utilizados na obtenção de cenários equiprováveis para a distribuição de variáveis regionalizadas.

Além da SGS e da simulação por bandas rotativas, também é abordado, ao longo do capítulo, o algoritmo de simulação *annealing* (e.g. Aarst e Korst, 1989), que pode ser utilizado em metodologias que visam a otimização de uma função ou parâmetro. Ainda, esse capítulo descreve aspectos teóricos sobre índices de incerteza que podem ser utilizados para medir a incerteza de uma distribuição de valores.

2.1 - Variáveis Regionalizadas e Estacionariedade

Os algoritmos de simulação estocástica, que serão utilizados nesse estudo, baseiam-se no fato de que a variável de interesse pode ser classificada como Variável Regionalizada. Algumas características podem ser relacionadas a uma Variável Regionalizada (Matheron (1963)): a distribuição dos valores ocorre ao longo de uma região, no caso da mineração, ao longo do depósito mineral; a variável pode apresentar uma continuidade dos valores no espaço, com alguns desvios entre valores vizinhos; e a variável pode apresentar uma direção de maior continuidade de seus valores. Journel e Huijbregts (1978) explicam que uma Variável Regionalizada apresenta, globalmente, um aspecto estruturado e, localmente, um aspecto aleatório.

Chilès e Delfiner (1999) comentam que a geoestatística se apóia na idéia da Variável Regionalizada em estudo ser uma realização de uma função aleatória. Conforme explicam Chilès e Delfiner (1999) e Armstrong (1998), o valor de uma Variável Regionalizada Z , em um local u , é considerado uma realização de uma variável aleatória $Z(u)$. Os autores ressaltam que, matematicamente, o conjunto de variáveis aleatórias é considerado uma função aleatória (FA).

De acordo com Isaaks e Srivastava (1989), as técnicas geoestatísticas mais utilizadas para realização de estimativas baseiam-se nos modelos estacionários de funções aleatórias. Armstrong (1998) explica que estacionariedade é verificada quando todos momentos estatísticos não variam com a translação. Contudo, a autora comenta que, na prática, apenas os dois primeiros momentos estatísticos (média e covariância) são considerados, por isso a hipótese assumida é chamada estacionariedade de segunda ordem. A autora também explica que, mesmo quando a média é constate ao longo da região em estudo, a covariância pode não existir. Nesse caso, a hipótese intrínseca pode ser utilizada. Conforme explica a autora, a hipótese intrínseca considera que a média e variância dos incrementos da função aleatória, para um dado h , existem e são independentes da posição espacial. Os incrementos para cada valor de h , podem ser definidos como a diferença entre o valor da variável aleatória em um local (u) e o valor da variável aleatória em um local ($u+h$).

As variáveis que serão utilizadas nesse estudo são Variáveis Regionalizadas e a estacionariedade de cada função aleatória será verificada a fim de garantir o uso das ferramentas geoestatísticas apresentadas nesse capítulo.

2.2 - Simulação Estocástica

Os algoritmos de simulação estocástica são mais indicados do que os algoritmos de krigagem quando o estudo visa a análise da incerteza de valores de um atributo ao longo do depósito ou a análise da incerteza a respeito dos valores de uma função de transferência.

Os algoritmos de krigagem têm por finalidade gerar a “melhor” estimativa para uma variável em um dado local (u). Conforme explica Armstrong (1998), a palavra “melhor” é utilizada no sentido de que o algoritmo busca minimizar a variância do erro (erro pode ser definido como a diferença entre o valor real e a estimativa em um local u) de interpolação. A estimativa obtida por krigagem também é chamada de não-tendenciosa, visto que a esperança matemática do erro da estimativa é nula. Contudo, apesar da krigagem fornecer as melhores estimativas locais, o mapa dessas estimativas apresenta um aspecto suavizado. O que não é comum para variáveis da área de mineração, como por exemplo, teor de um metal. A krigagem tem por objetivo a “melhor” estimativa sem respeitar algum parâmetro estatístico global (Deutsch e Journel, 1998).

Os algoritmos de simulação estocástica fornecem realizações ou cenários da distribuição dos valores do atributo de interesse ao longo do depósito. Cada realização corresponde ao conjunto dos n -valores simulados em diferentes locais (u). Os valores simulados, em cada realização, devem honrar os dados originais (amostrais) em seus respectivos locais (valores simulados condicionais aos dados originais), apresentar um histograma semelhante ao histograma desagrupado dos dados amostrais e apresentar um modelo de continuidade espacial próximo ao modelo de continuidade (modelo variográfico) ajustado aos dados originais.

Deutsch e Journel (1998) explicam que os valores simulados em cada local (u) consideram um componente de erro, ao contrário das estimativas resultantes da

krigagem. Cada valor simulado em um dado local (u) é a soma de um valor de erro simulado (r) e o valor da estimativa realizada por krigagem. Nesse caso, o erro é um resíduo adicionado ao valor da estimativa, que é responsável pelo aspecto localmente aleatório da Variável Regionalizada.

Os diversos cenários gerados pelo processo de simulação estocástica podem ser utilizados na análise da incerteza associada à distribuição espacial dos valores de um atributo, em uma determinada porção de um depósito mineral. A incerteza a respeito da distribuição dos valores do atributo pode ser útil, por exemplo, na etapa de beneficiamento de um minério metálico. A faixa de teores que serão lavrados, em uma porção do depósito, indica a necessidade de ajuste na concentração dos reagentes utilizados para o beneficiamento daquele minério e/ou o grau de homogeneização do minério antes que esse alimente a planta de beneficiamento. Cada cenário, gerado por simulação, apresenta uma distribuição de teores de minério que serão lavrados. A incerteza associada à faixa de possíveis teores que alimentarão a planta de beneficiamento, nesse período de tempo, pode ser acessada pela análise dos diversos cenários.

Os cenários gerados por simulação estocástica também podem ser utilizados na análise da incerteza associada aos valores de uma função de transferência global. Por exemplo, na classificação de recursos de um depósito mineral. Para o caso em que a função de transferência é a tonelagem de minério acima de um dado teor de corte, cada cenário simulado resultará em um valor para tonelagem de minério e a incerteza associada aos valores de tonelagem pode classificar as reservas como prováveis ou provadas.

Os algoritmos de simulação estocástica também permitem o acesso à incerteza local sobre os valores de um atributo, visto que, o conjunto de realizações fornece L (número de realizações) valores simulados em cada local. Cada valor simulado se relaciona com valores simulados em outros locais preservando uma continuidade espacial semelhante à continuidade dos dados originais.

Conforme comenta Goovaerts (1997), o histograma e variograma dos valores simulados, em cada realização, raramente são exatamente iguais aos modelos de histograma e variograma obtidos a partir dos dados originais. Esse autor explica que as flutuações da distribuição acumulada dos dados simulados e do variograma dos dados simulados em torno do modelo de distribuição acumulada dos

dados originais e do modelo variográfico dos dados originais, respectivamente, são chamadas de flutuações ergódicas. O mesmo autor também observa que quando os dados originais são poucos ou agrupados, ou seja, não representativos da população, as flutuações ergódicas permitem que a incerteza a respeito da estatística dos dados originais seja considerada.

Segundo Deutsch e Journel (1998), ergodicidade é alcançada quando o valor de um parâmetro estatístico da realização se aproxima do valor do mesmo parâmetro para os valores da Função Aleatória.

Alguns fatores influenciam na magnitude das flutuações ergódicas (Deutsch e Journel, 1992; Goovaerts 1997), por exemplo: o aumento de dados originais utilizados na simulação diminui as flutuações ergódicas; e a relação entre parâmetros do variograma, como o alcance e o efeito pepita, e o tamanho da área a ser simulada influenciam nessas flutuações. Os autores explicam que, geralmente, um variograma de grande alcance em relação ao tamanho do *grid* a ser simulado produz flutuações ergódicas significativas e os autores ressaltam a contribuição de um pequeno efeito pepita, em relação ao patamar, para essas flutuações.

Os algoritmos de simulação estocástica podem ser classificados em seqüenciais e não-seqüenciais de acordo com a forma com que garantem a semelhança entre os modelos de continuidade espacial dos valores simulados e o obtido a partir dos dados originais do atributo de interesse (Z).

2.2.1 - Simulação Seqüencial

A simulação seqüencial realiza a simulação em cada local (u) condicionada aos valores dos dados amostrais para o atributo de interesse e aos valores simulados anteriormente, em outros nós de um *grid*.

Em um primeiro local a ser simulado, é construída uma distribuição acumulada para valores do atributo em estudo, essa distribuição é construída condicionada aos valores dos dados originais. Sorteia-se, então, um valor dessa distribuição, que corresponde ao valor simulado para esse primeiro local. No segundo local a ser simulado, é construída, novamente, uma distribuição acumulada

para valores do atributo, essa distribuição é construída condicionada aos valores dos dados originais e ao valor simulado anteriormente. Sorteia-se, então, um valor dessa distribuição, que corresponde ao valor simulado nesse segundo local. A simulação dos próximos locais ocorre da mesma maneira, em cada local a ser simulado, uma distribuição acumulada é construída condicionada aos valores dos dados originais e aos valores previamente simulados e um valor é sorteado dessa distribuição, esse valor corresponde ao valor simulado para aquele local.

As probabilidades acumuladas dos valores de um atributo Z em um dado local, condicionadas aos dados originais e aos valores simulados, são valores de uma função: Função de Distribuição Acumulada Condicional (ou do inglês *Conditional Cumulative Distribution Function – ccdf*).

Goovaerts (1997) comenta que os algoritmos de simulação seqüencial garantem a reprodução do modelo de continuidade espacial, obtido a partir dos dados originais, devido ao condicionamento de cada distribuição acumulada pelos dados previamente simulados, além do condicionamento pelos dados originais.

A fim de que os algoritmos de simulação seqüencial possam ser aplicados em bancos de dados reais, oriundos de depósitos minerais, plumas de contaminação, etc., algumas implementações devem ser adotadas na prática desses algoritmos. Essas implementações facilitam a aplicação da teoria de simulação seqüencial em regiões que necessitem um grande número de locais simulados (nós do *grid*) e permitem que os cenários obtidos por simulação seqüencial reproduzam as características de continuidade espacial do atributo de interesse.

Deutsch e Journel (1998) comentam algumas implementações, como por exemplo, a utilização, apenas, dos dados mais próximos (dados originais e previamente simulados) para a simulação. É recomendado que seja utilizada a distância variográfica (considerando a continuidade espacial) dos dados em relação ao local a ser simulado. Na prática, apenas os dados mais próximos (de acordo com a continuidade espacial do atributo) é que contribuem de forma significativa na simulação. Goovaerts (1997) explica que devido ao fato de que as distribuições acumuladas são construídas utilizando-se algoritmos de krigagem, o aumento do número de dados condicionantes, causado pelo aumento do número de locais previamente simulados, dificulta a realização da krigagem e, conseqüentemente, a construção da distribuição acumulada, muitas vezes, tornando essa operação

impossível computacionalmente. Ao longo desse capítulo, serão comentados os algoritmos de krigagem utilizados pela simulação seqüencial Gaussiana.

Ainda no tópico de estratégias de busca para simulação seqüencial, Deutsch e Journel (1998) comentam que à medida que aumenta o número de locais previamente simulados próximos do local a ser simulado, ocorre uma tendência de um maior número de valores simulados condicionarem a simulação, em comparação com o número de dados originais. Nesses casos, pode ser utilizado um número mínimo de dados originais para cada nó simulado. Goovaerts (1997) comenta, baseado em Isaaks (1990), que a seqüência de locais que serão simulados, em cada realização, deve ser aleatória para evitar a geração de uma continuidade, que não condiz com a realidade, ao longo de uma direção. Além disso, também é recomendado que diferentes seqüências (ou caminhos) de locais a serem simulados sejam utilizados em cada realização a fim de evitar a geração de realizações muito semelhantes (*random path*).

Outra implementação importante para a prática dos algoritmos de simulação seqüencial é encontrada na literatura: o chamado *multiple-grid* (Gómez-Hernández, 1991). O *grid* a ser simulado é dividido em *grids* com diferentes espaçamentos entre os nós. O *grid* com maior espaçamento entre os nós é simulado primeiro e, assim, sucessivamente até a simulação daquele com menor espaçamento. Os valores simulados são considerados na simulação do próximo *grid*. Aquele com maior espaçamento permite a reprodução da continuidade de longo alcance apresentada pelo atributo. O número de etapas, em que deve ser simulado o *grid* de interesse, depende do número de estruturas, em que pode ser dividida a continuidade do atributo em estudo, e do espaçamento do *grid* de interesse.

Conforme foi explicado nessa seção, os algoritmos de simulação seqüencial constroem uma *ccdf* para valores do atributo de interesse em cada local a ser simulado. Esses algoritmos podem ser diferenciados de acordo com o modo que cada distribuição acumulada é construída. A seguir, um tipo de simulação seqüencial é abordado: a simulação seqüencial Gaussiana, que se baseia na construção da distribuição acumulada de forma Gaussiana por algoritmos da krigagem linear.

2.2.1.1 - Simulação Seqüencial Gaussiana

A simulação seqüencial Gaussiana (Isaaks, 1990) constrói uma distribuição acumulada para valores do atributo de interesse em cada local a ser simulado. A construção dessa distribuição acumulada é baseada no chamado *modelo multiGaussiano*.

Sabe-se que o modelo multiGaussiano assume, entre outras coisas, que para um atributo normal padrão (ou seja, para um atributo cuja distribuição dos valores, distribuídos ao longo da região em estudo, apresenta um histograma normal padrão) qualquer combinação de valores desse atributo apresenta distribuição normal. A distribuição acumulada de possíveis valores do atributo em um local u , considerando os dados da vizinhança, é normal e definida por parâmetros condicionados aos dados: esperança matemática e variância.

De acordo com Journel e Huijbregts (1978), no modelo multiGaussiano a esperança matemática e variância da distribuição normal acumulada são iguais a estimativa por krigagem simples e variância de krigagem simples, respectivamente, considerando que os dados vizinhos são utilizados na interpolação. O termo aproximação paramétrica é utilizado porque a estimativa de parâmetros (média e variância) é necessária para a definição de cada distribuição normal.

A utilização do modelo multiGaussiano para construção das distribuições normais assume que o atributo em estudo seja normal padrão. Nem sempre o atributo de interesse apresenta valores seguindo essa distribuição, nesse caso, a transformação da distribuição dos dados originais em uma distribuição normal é necessária para a simulação seqüencial Gaussiana. Para cada dado original, será atribuído um valor correspondente no espaço normal, esse processo é chamado normalização dos dados. O tópico “Normalização dos Dados Originais” descreve uma das metodologias que pode ser utilizada no processo de transformação da distribuição original em distribuição normal. Após a SGS, os valores normais devem ser transformados nos correspondentes valores do espaço original, aspectos teóricos sobre essa operação serão abordados no tópico “Retro-Transformação”.

Como será visto no tópico “Metodologia da SGS”, na prática, além da

krigagem simples, outros tipos de krigagem podem ser utilizados para inferência da esperança matemática de cada distribuição normal acumulada. Aspectos teóricos sobre krigagem simples e outros algoritmos de krigagem (Matheron, 1963) são apresentados nessa seção.

A. Normalização dos Dados Originais

Para utilização do modelo multiGaussiano na simulação seqüencial Gaussiana, a distribuição dos valores do atributo em estudo deve ser normal padrão. Rivoirard (1990) descreve uma metodologia para transformação dos dados originais em dados normais. De acordo com o autor, a partir da função de distribuição acumulada dos valores de Z ($F(z)$) e da função de distribuição acumulada de uma variável normal padrão Y ($G(y)$), se obtém, para cada valor de Z , o valor correspondente no espaço normal. O respectivo valor de Z no espaço normal é aquele que apresenta a mesma probabilidade acumulada do valor de Z no espaço original. A Figura 2.1 apresenta um esquema da normalização dos dados pela metodologia descrita. Na literatura geoestatística, a metodologia descrita é traduzida matematicamente da seguinte forma:

$$y = G^{-1}(F(z)) \quad (\text{Equação 2.1})$$

onde:

y = valor do espaço normal correspondente a um dado original z ;

$G(y)$ = função de distribuição acumulada para um valor normal padrão y ;

$F(z)$ = função de distribuição acumulada para um valor original z ;

$G^{-1}(F(z))$ = valor normal padrão cuja probabilidade acumulada é igual a $F(z)$.

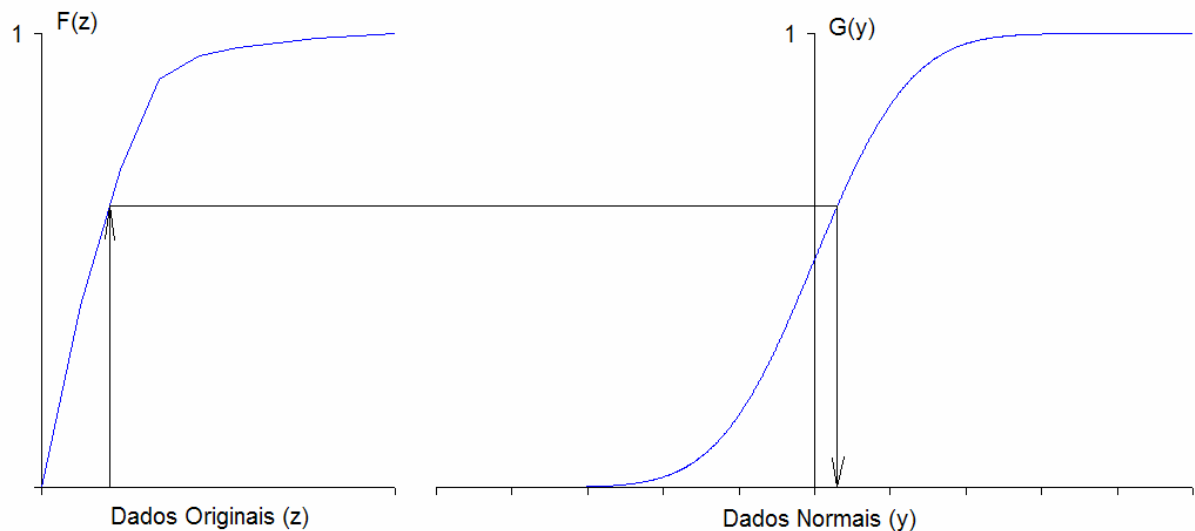


Figura 2.1 - Esquema da normalização dos dados pela metodologia descrita. Modificado de Goovaerts (1997).

De acordo com Deutsch e Journel (1998), se os dados originais apresentam agrupamento ao longo da área em estudo, pesos de desagrupamento devem ser utilizados para construção da função de distribuição acumulada ($F(z)$) que será considerada na normalização.

De acordo com Rivoirard (1990), a função que relaciona os dados originais com os dados normalizados é chamada função de anamorfose e pode ser escrita da seguinte forma:

$$z = \phi(y) \quad (\text{Equação 2.2})$$

onde:

z = dado original;

y = dado normal para o respectivo valor z .

Conforme comenta Rivoirard (1990), nem sempre a distribuição dos valores originais pode ser transformada em uma distribuição normal. O autor cita um exemplo onde a distribuição para uma dada concentração apresenta 50% dos valores das amostras iguais a 0 e explica a dificuldade de associar diferentes valores gaussianos à cada valor 0.

Sabe-se que modelo multiGaussiano assume multinormalidade da variável aleatória a ser simulada, ou seja, não apenas o histograma dos dados deve ser normal como o histograma para pares de dados (binormalidade), trios, etc. Segundo Goovaerts (1997), é difícil verificar a multinormalidade dos dados, por isso, se a binormalidade dos dados é verificada, simplificada assume-se que o modelo multiGaussiano pode ser utilizado.

Alguns testes para verificação da binormalidade são descritos na literatura. Deutsch e Journel (1992) descrevem um teste que consiste em comparar o variograma dos indicadores (ver equação 2.17) da variável normal (Y), para um teor de corte (y_p), com o variograma dos indicadores para uma variável binormal padrão, considerando o mesmo teor de corte. Esses variogramas devem ser semelhantes.

Outro teste para verificação da binormalidade consiste em dividir o valor do variograma para pares de valores normais, considerando um *lag* h , pelo valor da raiz quadrada do variograma para os mesmo pares de valores normais. O quociente encontrado deve ser semelhante ao valor 0,564 (Matheron, 1982).

B. Algoritmos de Krigagem

A simulação seqüencial Gaussiana utiliza krigagem simples para construção da distribuição normal acumulada em cada local a ser simulado. A krigagem simples assume que a média da variável normal Y é conhecida e constante ao longo da região em estudo. Contudo, a Krigagem Ordinária e Krigagem com um Modelo de Tendência podem ser utilizadas a fim de considerar a não-estacionariedade da média, quando necessário (Deutsch e Journel, 1992; Goovaerts 1997).

Considerando a estimativa de um atributo Z em um local u , a equação de krigagem simples pode ser definida por:

$$Z^*_{SK}(u) - m(u) = \sum_{i=1}^{n(u)} \lambda_i^{SK}(u) \cdot [Z(u_i) - m(u_i)] \quad (\text{Equação 2.3})$$

onde:

$Z^*_{SK}(u)$ = estimativa de krigagem simples em um local u ;

$m(u)$ = esperança matemática (média) dos valores de Z em um local u ;

$\lambda_i^{SK}(u)$ = peso atribuído a i -ésima amostra usada na estimativa para o local u ;

$Z(u_i)$ = valor da i -ésima amostra usada na estimativa;

$m(u_i)$ = esperança matemática dos valores de Z no local da i -ésima amostra;

$n(u)$ = número de amostras usadas na estimativa para o local u .

A expressão $[Z(u_i)-m(u_i)]$ na equação 2.3 define o valor da variável resíduo (R), no local da i -ésima amostra.

A minimização da variância do erro (diferença entre o valor real e a estimativa, em um local u) é definida em termos de um sistema de krigagem. Esse sistema considera o modelo de covariância da variável R e os pesos de krigagem $\lambda_i(u)$, com $i=1,\dots,n(u)$. O modelamento dos covariogramas (ou variogramas) para a variável R permite o cálculo dos pesos de krigagem. O sistema de krigagem simples é descrito a seguir:

$$\sum_{j=1}^{n(u)} \lambda_j^{SK}(u) \cdot C(u)_{R(ij)} = C(u)_{R(i0)} \quad i = 1, \dots, n \quad (\text{Equação 2.4})$$

onde:

$C(u)_{R(ij)}$ = covariância entre os i -ésimo e j -ésimo valores de resíduo usados para estimativa em um local u ;

$C(u)_{R(i0)}$ = covariância entre o i -ésimo valor de resíduo usado para estimativa em um local u e o ponto a ser interpolado;

$\lambda_j^{SK}(u)$ = peso atribuído a j -ésima amostra usada na estimativa para o local u ;

$n(u)$ = número de amostras usadas na estimativa para o local u .

Conforme explica Goovaerts (1997), quando existe uma amostra no local a ser interpolado, a resolução do sistema de krigagem atribui peso 1 para essa

amostra, ou seja, o valor da amostra será o valor da estimativa. Por isso, a krigagem simples (assim como a krigagem ordinária e krigagem com um modelo de tendência) é um interpolador exato.

Na prática, como a krigagem simples assume que as médias $m(u)$ e $m(u_i)$ são iguais e constantes em todos locais pertencentes à região em estudo, Goovaerts (1997) lembra que o modelo de covariância para a variável R é igual ao modelo de covariância para o atributo Z . Assim, os valores de covariância entre resíduos utilizados no sistema de krigagem podem ser substituídos pelos valores de covariância entre valores da variável Z , nos locais considerados.

A variância do erro da estimativa obtida por krigagem simples ou variância de krigagem simples é definida pela seguinte equação:

$$\sigma^2_{SK}(u) = \sigma^2 - \sum_{i=1}^{n(u)} \lambda_i^{SK}(u) C(u)_{(i0)} \quad (\text{Equação 2.5})$$

onde:

$\sigma^2_{SK}(u)$ = variância de krigagem simples para a estimativa em u ;

σ^2 = variância dos valores do atributo Z ;

$C(u)_{(i0)}$ = covariância entre o i -ésimo valor de Z usado para estimativa em um local u e o ponto a ser interpolado;

$\lambda_i^{SK}(u)$ = peso atribuído a i -ésima amostra usada na estimativa para o local u ;

$n(u)$ = número de amostras usadas na estimativa para o local u .

À medida que aumenta a covariância entre as amostras de Z e o ponto a ser interpolado, ou seja, à medida que aumenta a correlação entre as amostras e o ponto, diminui a incerteza associada à estimativa, medida pela variância de krigagem. Contudo, conforme comentado no Capítulo 1, muitos autores criticam o uso da variância de krigagem como medida de incerteza, visto que essa não considera a variabilidade dos dados utilizados na estimativa.

A equação de krigagem simples (2.3) pode ser re-escrita da seguinte forma:

$$Z^*_{sk}(u) = \sum_{i=1}^{n(u)} \lambda_i^{sk}(u) \cdot Z(u_i) + \lambda_m^{sk}(u) \cdot m \text{ sendo } \lambda_m^{sk}(u) = 1 - \sum_{i=1}^{n(u)} \lambda_i^{sk}(u) \quad (\text{Equação 2.6})$$

A equação 2.6 evidencia que a krigagem simples atribui um peso para a média estacionária além dos pesos para as amostras. Segundo explica Goovaerts (1997), para uma mesma correlação entre as amostras, quanto menor a correlação das amostras com o ponto a ser interpolado, menor o somatório dos pesos atribuídos para as amostras e maior o peso atribuído para a média.

A krigagem ordinária (Matheron, 1963) não considera a média de Z estacionária ao longo da região em estudo. Deutsch e Journel (1998) explicam que a média constante considerada na krigagem simples é substituída, na krigagem ordinária, por uma média interpolada dependente de cada local. Goovaerts (1997) comenta que quando as mesmas amostras são utilizadas na estimativa de dois locais diferentes, a média interpolada para esses locais é a mesma. Por isso, esse autor observa que a krigagem ordinária considera a média constante em uma vizinhança local.

A fim de considerar variações da média em diferentes vizinhanças, a krigagem ordinária impõe uma restrição aos pesos atribuídos para as amostras:

$$\sum_{i=1}^{n(u)} \lambda_i^{OK}(u) = 1 \quad (\text{Equação 2.7})$$

onde:

$\lambda_i^{OK}(u)$ = peso atribuído a i -ésima amostra usada na estimativa para o local u ;

$n(u)$ = número de amostras usadas na estimativa para o local u .

A equação da krigagem ordinária (Matheron, 1963) é escrita da seguinte forma:

$$Z^*_{OK}(u) = \sum_{i=1}^{n(u)} \lambda_i^{OK}(u) \cdot Z(u_i) \quad (\text{Equação 2.8})$$

onde:

$Z^*_{OK}(u)$ = estimativa de krigagem ordinária em um local u ;

$\lambda_i^{OK}(u)$ = peso atribuído a i -ésima amostra usada na estimativa para o local u ;

$Z(u_i)$ = valor da i -ésima amostra usada na estimativa;

$n(u)$ = número de amostras usadas na estimativa para o local u .

Os pesos de krigagem ordinária são conhecidos com a resolução do seguinte sistema:

$$\begin{cases} \sum_{j=1}^n \lambda_j^{OK}(u) C(u)_{R(ij)} + \mu_{OK}(u) = C(u)_{R(i0)} & i = 1, \dots, n \\ \sum_{j=1}^n \lambda_j^{OK}(u) = 1 \end{cases} \quad (\text{Equação 2.9})$$

onde:

$C(u)_{R(ij)}$ = covariância entre os i -ésimo e j -ésimo valores de resíduo usados para estimativa em um local u ;

$C(u)_{R(i0)}$ = covariância entre o i -ésimo valor de resíduo usado para estimativa em um local u e o ponto a ser interpolado;

$\lambda_j^{OK}(u)$ = peso atribuído a j -ésima amostra usada na estimativa para o local u ;

$n(u)$ = número de amostras usadas na estimativa para o local u ;

$\mu_{OK}(u)$ = parâmetro de Lagrange.

Na prática, a vizinhança de busca de amostras é a mesma para qual a média é considerada estacionária, por isso, o modelo de covariância para a variável R pode ser substituído pelo modelo de covariância atribuído para a variável Z .

O sistema de krigagem ordinária apresenta $(n(u)+1)$ equações, considerando a restrição aos pesos de krigagem. Isaaks e Srivastava (1989) explicam que o sistema também deve apresentar $(n(u)+1)$ incógnitas a fim de garantir uma única solução ao sistema. Os autores observam que esse número de incógnitas é alcançado com a introdução do parâmetro de Lagrange na equação da variância do erro da estimativa (a minimização dessa variância é definida na forma do sistema de krigagem) de forma a não modificar o valor dessa variância.

A variância do erro da estimativa obtida por krigagem ordinária ou variância de krigagem ordinária (Matheron, 1963) é definida pela seguinte equação:

$$\sigma^2_{OK}(u) = \sigma^2 - \left(\sum_{i=1}^{n(u)} \lambda_i^{OK}(u) \cdot C(u)_{(io)} + \mu_{OK}(u) \right) \quad (\text{Equação 2.10})$$

onde:

$\sigma^2_{OK}(u)$ = variância de krigagem ordinária para a estimativa em u ;

σ^2 = variância dos valores do atributo Z ;

$C(u)_{(io)}$ = covariância entre o i -ésimo valor de Z usado para estimativa em um local u e o ponto a ser interpolado;

$\lambda_i^{OK}(u)$ = peso atribuído a i -ésima amostra usada na estimativa para o local u ;

$n(u)$ = número de amostras usadas na estimativa para o local u ;

$\mu_{OK}(u)$ = parâmetro de Lagrange.

A krigagem com um modelo de tendência, ou krigagem universal (Matheron, 1969), diferente da krigagem ordinária, considera a média dos valores de Z variável dentro de cada vizinhança.

C. Retro-Transformação

A SGS assume que os dados originais apresentam histograma normal. A

transformação do histograma dos dados em um histograma normal pode ser feita pela metodologia descrita no tópico “Normalização dos dados originais”. Assim, a cada dado original é atribuído um valor no espaço normal.

Cada realização da SGS fornece valores no espaço normal, simulados em diferentes pontos da região de interesse. Esses valores devem ser transformados nos respectivos valores do espaço original. Essa operação pode ser chamada retro-transformação (ou do inglês *Back Transformation*). A retro-transformação pode ser realizada pelo inverso da normalização descrita anteriormente. Ou seja, para cada valor no espaço normal (y), atribuir um valor no espaço original (z) que apresenta a mesma probabilidade acumulada ($F(z)=G(y)$).

Geralmente, o número de dados originais não permite o conhecimento da função de distribuição acumulada $F(z)$ para qualquer valor de Z . O conhecimento dos valores de $F(z)$ para qualquer valor de z pode ser obtido por interpolação dos pontos previamente conhecidos e extrapolações a partir dos pontos correspondentes ao maior e ao menor teor até um valor máximo e mínimo, respectivamente. Os valores máximo e mínimo são aqueles que delimitam a faixa de possíveis valores para a variável Z naquele local.

O conhecido modelo de potência pode ser utilizado para interpolação e extrapolação entre e a partir, respectivamente, dos pontos previamente conhecidos. O valor $F(z)$, para um teor de corte z pertencente ao intervalo $(z_1, z_2]$, pode ser obtido pelo modelo de potência da seguinte forma (Deutsch e Journel, 1992; Goovaerts, 1997):

$$F(z) = F(z_1) + \left[\frac{z - z_1}{z_2 - z_1} \right]^w \cdot [F(z_2) - F(z_1)] \quad \forall z \in (z_1, z_2] \quad (\text{Equação 2.11})$$

onde:

$w > 0$

$F(z)$ = valor da função de distribuição acumulada para o teor de corte z ;

$F(z_1)$ = valor da função de distribuição acumulada para o teor de corte z_1 ;

$F(z_2)$ = valor da função de distribuição acumulada para o teor de corte z_2 ;

Goovaerts (1997) comenta que é comum a utilização do modelo linear para a interpolação. O modelo linear é um caso do modelo de potência onde $w=1$ (Deutsch e Journal, 1992; Goovaerts, 1997).

Outro modelo que pode ser usado para extrapolação é o chamado modelo hiperbólico. Deutsch e Journal (1992) ressaltam que esse modelo deve ser usado, apenas, quando o atributo em estudo apresenta uma distribuição com assimetria positiva e para extrapolação da cauda superior (direção do aumento dos teores de corte).

O valor $F(z)$, para um teor de corte z maior que z_2 , pode ser obtido pelo modelo hiperbólico da seguinte forma (Deutsch e Journal, 1992):

$$F(z) = 1 - \frac{[z_2^w \cdot (1 - F(z_2))]}{z^w}, \quad w \geq 1, \quad [z_2^w \cdot (1 - F(z_2))] > 0 \quad (\text{Equação 2.12})$$

onde:

$F(z)$ = valor da função de distribuição acumulada para o teor de corte z ;

$F(z_2)$ = valor da função de distribuição acumulada para o teor de corte z_2 ;

Conforme mencionado anteriormente, os valores da função de distribuição acumulada ($F(z)$), para qualquer valor de z , podem ser obtidos por interpolação e extrapolação entre e a partir dos valores conhecidos de $F(z)$. Os valores obtidos por interpolação e extrapolação dependem dos modelos utilizados (e.g. modelo de potência), assim como, dos parâmetros de cada modelo. Nesses casos, por meio da operação de retro-transformação, diferentes valores simulados podem ser encontrados para diferentes modelos (e parâmetros), por isso, a importância da escolha desses modelos (e parâmetros) em cada caso estudado.

A retro-transformação, ou seja, atribuição de um valor no espaço original (z) para cada valor no espaço normal (y), também pode ser realizada com o uso dos polinômios de Hermite, visto que a função de anamorfose $\phi(y)$, para qualquer valor gaussiano, pode ser definida em termos desses polinômios.

Conforme Rivoirard (1990), geralmente, qualquer função de uma variável normal padrão Y , em um local u , pode ser definida em termos dos polinômios de Hermite:

$$F[Y(u)] = \sum_{n=0}^{\infty} f_n \cdot H_n[Y(u)] \quad (\text{Equação 2.13})$$

onde:

$Y(u)$ = variável gaussiana padrão Y , em um local u ;

$F[Y(u)]$ = função de uma variável $Y(u)$;

f_n = coeficiente do polinômio de Hermite de grau n ;

$H_n[Y(u)]$ = polinômio de Hermite de grau n para $Y(u)$.

Assim, a função de anamorfose gaussiana $\phi(y)$, para um valor gaussiano y , também pode ser definida em termos dos polinômios de Hermite.

Rivoirard (1990) comenta que os polinômios de Hermite são definidos pela fórmula de Rodrigues:

$$H_n(y) = \frac{1}{\sqrt{n!} \cdot g(y)} \cdot \frac{d^n g(y)}{dy^n} \quad n \geq 0 \quad (\text{Equação 2.14})$$

onde:

y = valor gaussiano;

$H_n(y)$ = polinômio de Hermite de grau n , calculado para o valor gaussiano y ;

$g(y)$ = função de distribuição de probabilidade para o valor gaussiano y .

De acordo com Rivoirard (1990), os polinômios de Hermite para $n=0$ e $n=1$ são iguais a 1 e $-y$, respectivamente, enquanto os demais polinômios de Hermite

podem ser encontrados por:

$$H_{n+1}(y) = -\frac{1}{\sqrt{n+1}} \cdot y \cdot H_n(y) - \sqrt{\frac{n}{n+1}} \cdot H_{n-1}(y) \quad n > 0 \quad (\text{Equação 2.15})$$

Rivoirard (1990) demonstra que os coeficientes dos polinômios de Hermite são definidos pela seguinte fórmula:

$$f_n = \int f(y) \cdot H_n(y) \cdot g(y) \cdot dy \quad (\text{Equação 2.16})$$

Considerando que $f(y)$ na equação 2.16 é a função de anamorfose gaussiana $\phi(y)$, o coeficiente do polinômio de Hermite de grau n pode ser calculado a partir dos valores experimentais. Conhecendo os coeficientes dos polinômios de Hermite para cada grau n , o valor da função de anamorfose para qualquer valor gaussiano y pode ser calculado.

D. Metodologia da SGS

A simulação seqüencial Gaussiana (Isaaks, 1990) de uma variável contínua Z compreende as seguintes etapas:

- i. Transformar a distribuição dos dados originais desagrupados em uma distribuição normal padrão, quando necessário. A cada dado original (z) deve ser atribuído um valor no espaço normal (y), que será usado na simulação. Verificar binormalidade dos valores normalizados, se essa condição for verificada, prosseguir com a simulação;
- ii. Definir um caminho aleatório em que cada local a ser simulado (u_j , $j=1, \dots, N$), onde N é o número de locais a serem simulados ao longo da região em estudo, seja visitado uma vez;

- iii. Em um local u_j , assumir uma distribuição normal para os valores y . Realizar krigagem simples da variável normal Y , condicionada aos dados normais (y) e aos dados normais previamente simulados, pertencentes à vizinhança de busca de u_j . A krigagem simples utiliza valores de covariância para a variável normal (Y). A estimativa e variância de krigagem são a média e variância, respectivamente, da distribuição normal local. Assim, a *ccdf* para valores de Y no local u_j pode ser construída;
- iv. Sortear um valor da *ccdf* para valores de Y no local u_j . O valor sorteado corresponde ao valor simulado para o local u_j , para aquela realização;
- v. De acordo com o caminho aleatório, seguir para o próximo local e repetir as etapas iii a v;
- vi. Repetir as etapas iii a v até que os N locais tenham sido simulados;
- vii. Retro-transformar (considerando os pesos de desagrupamento, quando utilizados na normalização) os valores normais em valores originais;
- viii. Considerando L o número de realizações desejado para o estudo. Repetir as etapas ii a vii L vezes. Assim, serão gerados L cenários para distribuição dos valores de Z .

Na SGS, a krigagem simples pode ser substituída por krigagem ordinária e krigagem com um modelo de tendência quando a média do atributo a ser simulado não for estacionária ao longo da região em estudo. Porém, outros algoritmos além da krigagem simples devem ser utilizados, apenas, para a estimativa da média da distribuição normal, no local u . A variância da distribuição normal deve ser a variância de krigagem simples (Journel, 1980; Goovaerts, 1997).

E. SIS e SGS

Apesar da simulação seqüencial Gaussiana (SGS) ser mais simples computacionalmente que a simulação seqüencial dos indicadores (SIS), Goovaerts (1997) explica, citando autores como Journel e Alabert (1988), que devido a uma propriedade do modelo multiGaussiano, a SGS não reproduz a continuidade para

valores extremos. Para casos em que esses valores sejam bem correlacionados, a utilização desse algoritmo poderia subestimar ou superestimar a função de transferência baseada nos valores simulados. O autor apresenta um exemplo mostrando que a simulação seqüencial dos indicadores permitiu uma melhor reprodução da continuidade para baixos valores de uma variável, do que a simulação seqüencial Gaussiana.

Outros inconvenientes da SGS, também são citados por Goovaerts (1997): a construção da distribuição acumulada, para valores normais em um dado local u , é baseada na variância de krigagem, que não considera os valores dos dados utilizados na estimativa. Além disso, o autor comenta que a SGS exige que os dados normalizados sejam multinormais, hipótese que é assumida se os dados forem binormais, se a multinormalidade não for confirmada, o modelo multinormal não pode ser utilizado para construção de cada distribuição acumulada. Ou seja, se o modelo multinormal não pode ser utilizado, a SGS não se aplica para o caso em estudo.

A SIS (não-paramétrica) é uma alternativa à SGS quando algum dos inconvenientes desse último algoritmo não permitem sua aplicação para os dados em estudo. Além disso, a SIS pode ser utilizada para dados categóricos (e.g. tipo de rocha) e contínuos (e.g. teor de um metal), como descrevem Deutsch e Journel (1998).

A simulação seqüencial dos indicadores (Alabert, 1987) baseia-se na krigagem dos indicadores (Journel, 1982). Conforme Journel (1982), um indicador pode ser definido da seguinte maneira:

$$i(u; z_k) = \begin{cases} 1, & \text{para } z(u) \leq z_k \\ 0, & \text{outros casos} \end{cases} \quad k = 1, \dots, K \quad (\text{Equação 2.17})$$

onde:

$i(u; z_k)$ = indicador em um local u considerando um teor limite z_k ;

$z(u)$ = valor da variável Z em um local u ;

z_k = teor limite (ou teor de corte) para a classe k ;

K = número de classes de indicadores ou número de teores de corte.

Demais informações sobre a simulação seqüencial dos indicadores podem ser encontradas em Goovaerts (1997) e em Deutsch e Journel (1998).

2.2.2 - Simulação por Bandas Rotativas

O algoritmo de simulação utilizado para os estudos de caso dos Capítulos 3 e 4 foi Bandas Rotativas (Matheron, 1973; Journel 1974). Assim como, na simulação seqüencial Gaussiana, a distribuição dos valores de Z deve ser normal padrão ou essa distribuição deve ser transformada em normal padrão antes da simulação por Bandas Rotativas. Contudo, ao contrário da SGS, esse algoritmo realiza a simulação não-seqüencial de valores de um atributo Z .

Imagine Q linhas (concêntricas) dividindo o espaço. Para um número suficiente de linhas, o modelo de covariância, isotrópico, para uma variável normal (Y) tri-dimensional pode ser relacionado ao modelo de covariância para uma variável normal definida em uma dimensão (1D), ou seja, definida ao longo de uma linha que divide o espaço (Matheron, 1973; Journel 1974; Lantuéjoul, 2002):

$$C_3(h) = \frac{1}{Q} \cdot \sum_{q=1}^Q C_1(\langle h, q \rangle) \quad (\text{Equação 2.18})$$

onde:

$C_3(h)$ = covariância para um par de valores, em 3D, separados por um vetor h ;

$C_1(\langle h, q \rangle)$ = covariância para um par de valores, em 1D, separados pela projeção do vetor h na q -ésima linha que divide o espaço;

Q = número total de linhas que dividem o espaço.

Considere, agora, a simulação de uma variável normal Y ao longo de cada

linha, respeitando o modelo de covariância $C_1(h)$. Qualquer ponto de um plano perpendicular à linha possui valor igual ao da simulação de Y no ponto de interseção da linha com o plano, que é o que ocorreria se a linha fosse deslocada ao longo do plano. Assim, de acordo com o teorema do limite central, cada simulação tri-dimensional pode ser obtida a partir das simulações das linhas da seguinte forma (Journel e Huijbregts, 1978; Lantuéjoul, 2002):

$$Y^s(u_1) = \frac{1}{\sqrt{Q}} \cdot \sum_{q=1}^Q Y(u_1)_q \quad (\text{Equação 2.19})$$

onde:

$Y^s(u_1)$ = simulação de Y no local u_1 (3D);

$Y(u_1)_q$ = simulação de Y no local da projeção ortogonal de u_1 na q -ésima linha.

Conforme a equação 2.18, os valores simulados em 3D, obtidos pela equação 2.19, seguem o modelo de covariância $C_3(h)$.

A Figura 2.2 apresenta um esquema da projeção do vetor h em uma das linhas.

Assim, a simulação por Bandas Rotativas consiste em simular uma variável gaussiana Y , com um modelo de covariância $C_1(h)$, ao longo de Q linhas. O valor da variável gaussiana em cada local, definido em 3D, é obtido a partir de simulações realizadas em 1D, ao longo de diferentes linhas. Com o mesmo raciocínio da equação 2.18, Matheron (1973) mostra uma derivada onde o modelo de covariância $C_1(h)$ pode ser obtido a partir do modelo de covariância $C_3(h)$, assim, para a realização da simulação é suficiente conhecer o modelo de covariância para a variável tri-dimensional.

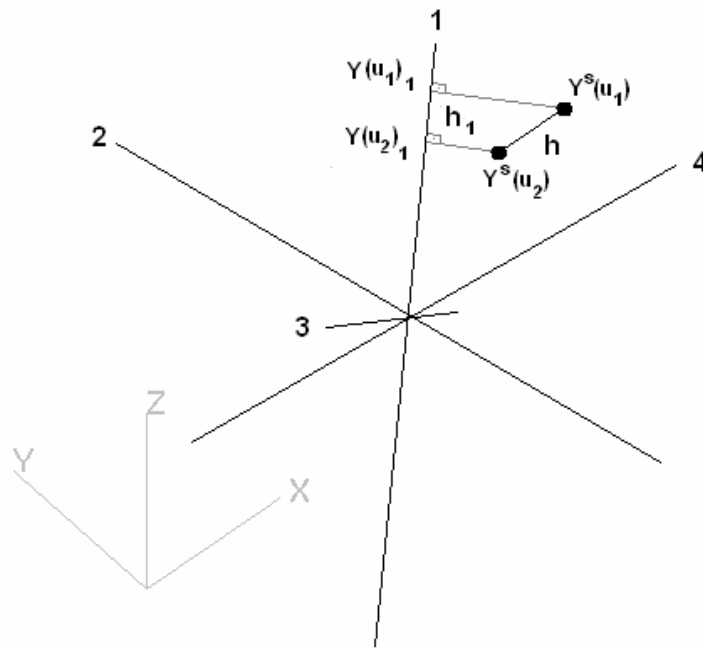


Figura 2.2 - Esquema da projeção do vetor h em uma das linhas.

Journal (1974) e Lantuéjoul (2002) explicam alguns métodos para se obter a simulação ao longo das linhas, respeitando o modelo de covariância $C_1(h)$. A simulação por Bandas Rotativas, utilizada nos estudos de caso dos Capítulos 3 e 4, foi realizada no software Isatis[®] (Bleines et al., 2001). Nesse software, o método utilizado para simulação ao longo das linhas, depende do tipo de modelo de continuidade $C_3(h)$ adotado para a variável em estudo (se esse modelo é esférico, exponencial...).

A equação 2.18 considera que o modelo de covariância para três dimensões $C_3(h)$ é isotrópico (representa a mesma continuidade para a variável em todas as direções). Segundo Emery (2006), uma forma de considerar um modelo de covariância anisotrópico é modificar a posição dos locais a serem simulados em cada direção, de acordo com a distância variográfica (de acordo com o modelo de continuidade), para a realização da simulação.

Até aqui, se explicou como são geradas as simulações da variável Gaussiana em 3D. Contudo essas simulações não são condicionadas aos dados originais (amostras). A fim de se realizar o condicionamento dos dados, as simulações não condicionais devem ser geradas nos nós do *grid* de simulação e nos locais dos dados originais. Assim, o condicionamento dos valores simulados pode

ser realizado da seguinte forma (Journal, 1974):

$$Y^s_c(u_1) = Y_k(u_1) + [Y^s_k(u_1) - Y^s(u_1)] \quad (\text{Equação 2.20})$$

onde:

$Y^s_c(u_1)$ = simulação de Y no local u_1 (3D), condicionada aos dados (normais);

$Y_k(u_1)$ = estimativa de krigagem de Y no local u_1 , condicionada aos dados;

$Y^s_k(u_1)$ = estimativa de krigagem de Y no local u_1 , condicionada aos valores simulados (não-condicionais) nos locais dos dados;

$Y^s(u_1)$ = simulação não-condicional de Y no local u_1 .

Conforme explica Journal (1974), após o condicionamento, a simulação por Bandas Rotativas honra os valores dos dados, ou seja, o valor simulado no local de um dado é igual ao valor do dado. Conforme comentado previamente, a Variável Regionalizada possui um aspecto aleatório em cada local. A krigagem produz uma estimativa ótima (um valor médio dos possíveis valores) em cada local, enquanto a simulação gera um possível valor de erro (resíduo) para essa estimativa ótima (em cada cenário), criando o aspecto aleatório (localmente) da Variável Regionalizada. Na equação 2.20, o valor entre parênteses é o resíduo obtido por simulação (Journal, 1974).

Após se obter os valores simulados, condicionados aos dados originais, esses valores gaussianos são transformados nos respectivos valores do espaço original. Métodos disponíveis para essa transformação são descritos no tópico “Retro-Transformação”.

2.2.3 - Simulação *Annealing*

A simulação estocástica de um atributo Z pode ser realizada pela chamada Simulação *Annealing*. Conforme Deutsch (1992), o termo simulação *annealing* é

utilizado para qualquer procedimento de otimização que se baseia na forma com que as moléculas se organizam durante o processo de resfriamento de metais, a fim de alcançarem o estado de menor energia para o sistema. O autor cita Geman e Geman (1984) como uns dos primeiros a aplicarem simulação *annealing* para um fenômeno espacial. A simulação *annealing* busca a otimização de uma função objetivo para geração de cada realização (Deutsch (1992) e Goovaerts (1997)). A função objetivo expressa a diferença entre um parâmetro estatístico calculado para os dados originais e para os valores de um determinado cenário simulado.

Na simulação *annealing*, um cenário inicial para a distribuição dos valores de um atributo Z é perturbado, criando um novo cenário. Segundo Deutsch (1992), uma regra que pode ser usada para decisão de aceitar ou rejeitar um novo cenário é: diante de uma perturbação favorável (a função objetivo para o novo cenário é menor ou igual à função objetivo para o antigo cenário) aceitar o novo cenário; diante de uma perturbação desfavorável, aceitar ou rejeitar o novo cenário de acordo com uma distribuição de probabilidade exponencial negativa (distribuição de Boltzmann (e.g. Aarst e Korst, 1989), equação 2.21). Diferentes perturbações são testadas até que um critério de convergência seja alcançado, quando, então, uma realização é obtida. Goovaerts (1997) cita alguns critérios de convergência, como o alcance de um valor para a função objetivo.

$$\text{Prob.} = \begin{cases} 1, & \text{se } O(c) \leq O(c-1) \\ \exp\left(\frac{[O(c-1) - O(c)]}{t(c)}\right), & \text{outros casos} \end{cases} \quad (\text{Equação 2.21})$$

onde:

Prob = probabilidade de aceitar ou rejeitar um novo cenário;

$O(c-1)$ = função objetivo para o cenário (c-1);

$O(c)$ = função objetivo para o c-ésimo cenário;

$t(c)$ = temperatura para o c-ésimo cenário.

O parâmetro temperatura controla a probabilidade de uma perturbação

desfavorável ser aceita. Conforme Deutsch (1992), a temperatura deve ser inicialmente alta e ser reduzida ao longo da simulação. Segundo o autor, se a temperatura for reduzida muito rápido, ocorre o risco de um cenário sub-otimizado ser mantido até o final da simulação, muitas vezes, não alcançando o critério de convergência com esse cenário, enquanto que a lenta redução da temperatura aumenta o tempo computacional para cada realização.

No Capítulo 1, alguns trabalhos citados, utilizam simulação *annealing* para otimizar a configuração de amostras. Assim, nos trabalhos citados, a simulação *annealing* não foi utilizada para simulação dos valores de Z, mas sim para otimizar uma função objetivo utilizada para definição da configuração espacial de amostras adicionais.

2.3 - Índices de Incerteza

A metodologia a ser desenvolvida nesse trabalho exige uma quantificação da incerteza associada aos valores simulados. Alguns índices que podem ser utilizados como medida da incerteza associada a valores de uma distribuição serão descritos nessa seção: Intervalo de Probabilidade, Entropia Local, Variância Condicional, Coeficiente de Variação e Intervalo entre Quartis. Pilger et al. (2000) chamou esses índices de índices de incerteza.

2.3.1 - Intervalo de Probabilidade

O intervalo de probabilidade (Journel, 1988) fornece a probabilidade do valor z , em um dado local u , estar em um certo intervalo. Por exemplo, conhecendo a Função de Distribuição Acumulada Condicional (*ccdf*) para valores de um atributo Z em u , a probabilidade de z estar entre o intervalo $(1,2]$ pode ser determinada da seguinte forma:

$$\text{Prob.} = F(u;2) - F(u;1) \quad (\text{Equação 2.22})$$

onde:

Prob. = probabilidade de z estar entre o intervalo (1,2];

$F(u;2)$ = probabilidade do valor z no local u ser menor ou igual a 2;

$F(u;1)$ = probabilidade do valor z no local u ser menor ou igual a 1.

Quanto menor o tamanho do intervalo definido e maior a probabilidade do valor z estar entre esse intervalo, menor a incerteza associada à distribuição dos valores de Z em u.

2.3.2 - Entropia Local

Considere a distribuição dos possíveis valores de Z, em cada local u, representada por um histograma com Q classes de mesma amplitude. A entropia local, a qual mede a incerteza associada aos valores da distribuição, é máxima quando todas classes apresentam mesma frequência (Shannon, 1948; Goovaerts, 1997). Por outro lado, a mínima entropia local ocorre quando uma classe apresenta 100% de frequência (Shannon, 1948; Goovaerts, 1997). A entropia local, para distribuição dos valores de Z, em u, é definida como [(Shannon (1948) definiu entropia como função de um logaritmo de base 10, enquanto Goovaerts (1997), assim como Christakos (1990), descreveram essa medida como função de um logaritmo natural)]:

$$H(u) \cong -\sum_{c=1}^C [\ln s_c(u)] \cdot s_c(u) \geq 0 \quad C \leq Q \quad (\text{Equação 2.23})$$

onde:

$H(u)$ = entropia local para distribuição dos valores de Z em u;

C = número de classes que não apresentam 0% de frequência;

$s_c(u)$ = frequência para a c -ésima classe;

Q = número total de classes do histograma.

De acordo com a equação 2.23, o mínimo valor de entropia local para qualquer distribuição é 0, enquanto o máximo valor é $\ln Q$ (Shannon, 1948; Goovaerts, 1997). Por isso, uma medida padronizada de entropia local, para um histograma com C classes, é dada por (Shannon, 1948; Goovaerts, 1997):

$$H_B(u) = \frac{H(u)}{\ln Q} \in [0,1] \quad (\text{Equação 2.24})$$

onde:

$H_B(u)$ = medida padronizada de entropia local para uma distribuição em u ;

Q = número total de classes do histograma.

2.3.3 - Variância Condicional

A variância condicional é uma medida do espalhamento dos valores de uma distribuição em torno da média dessa distribuição. Considere que a Função de Distribuição Acumulada Condicional (*ccdf*) para valores de um atributo Z , em um local u , é representada por um histograma com Q classes de mesma amplitude. A variância condicional, para essa distribuição, pode ser aproximada da seguinte forma (modificado de Goovaerts, 1997):

$$\sigma^2(u) = \sum_{q=1}^Q [z_q - m(u)]^2 \cdot [s_q] \quad (\text{Equação 2.25})$$

onde:

$\sigma^2(u)$ = variância condicional para distribuição dos valores de Z em um local u;

Q = número total de classes do histograma;

z_q = média aritmética dos valores de Z pertencentes a classe q;

s_q = frequência para a q-ésima classe;

$m(u) = \sum_{q=1}^Q z_q \cdot s_q$ = aproximação da média aritmética da distribuição para valores de

Z, em u.

O desvio padrão de uma distribuição é a raiz quadrada da variância da distribuição (e.g. Christmann, 1978).

2.3.4 - Coeficiente de Variação

De acordo com a estatística (e.g. Christmann, 1978), o coeficiente de variação é o resultado da divisão do desvio padrão de uma população pela média aritmética dessa população. Considere que a Função de Distribuição Acumulada Condicional (*ccdf*) para valores de um atributo Z, em um local u, é representada por um histograma com Q classes de mesma amplitude. O coeficiente de variação, para essa distribuição, pode ser aproximado da seguinte forma:

$$CV(u) = \frac{\sqrt{\sigma^2(u)}}{m(u)} \quad (\text{Equação 2.26})$$

onde:

$CV(u)$ = coeficiente de variação para distribuição dos valores de Z em um local u;

$\sigma^2(u)$ = variância condicional para distribuição dos valores de Z em um local u;

$m(u)$ = aproximação da média aritmética da distribuição para valores de Z, em u.

Conforme comentado anteriormente, exemplos da utilização do coeficiente de variação como índice de incerteza associada aos valores simulados foram apresentados por Pilger (2000) e Pilger et al. (2000).

2.3.5 - Intervalo entre Quartis

O Intervalo entre Quartil (IEQ) para a distribuição dos valores de um atributo Z , em um local u , é definido pela subtração do valor do quartil superior (0,75) pelo valor do quartil inferior (0,25) dessa distribuição. O IEQ é menos sensível a valores extremos do que a variância (ou desvio padrão) (Isaaks e Srivastava (1989)).

2.4 – Funções de Transferência

A metodologia desenvolvida na tese visa a redução da incerteza associada aos valores de uma dada função de transferência. Diferentes funções de transferência são consideradas nesse estudo, sendo exemplificadas ao longo do presente capítulo.

O cálculo do valor de cada função de transferência é baseado nos valores de um atributo Z para os blocos minerados ou SMUs (do inglês *selective mining units*). Por isso, os valores de cada realização, fornecida pela simulação estocástica, devem representar os blocos minerados. Contudo, quando os dados utilizados na simulação representam suportes quase-pontuais, como é o caso do suporte das amostras obtidas por sondagem, cada valor simulado também representa esse pequeno volume, que é menor do que o volume dos blocos minerados. Assim, uma operação de mudança de suporte deve ser utilizada para que simulações dos blocos sejam obtidas a partir das simulações quase-pontuais. Esse capítulo também descreve a operação de mudança de suporte que será utilizada nesse estudo.

2.4.1 - Mudança de Suporte

Nesse trabalho, a simulação dos valores de cada atributo de interesse, ao longo do depósito em estudo, visa o melhor conhecimento dos riscos associados a certas decisões na lavra deste depósito. Os blocos ou SMUs são as menores unidades a serem selecionadas nas operações de lavra. Por isso, os valores fornecidos pelas realizações devem contemplar ou permitir o acesso aos valores dos blocos do depósito para os diferentes atributos.

Goovaerts (2001) cita algumas metodologias que podem ser utilizadas para obtenção da função de distribuição acumulada condicional (*ccdf*), considerando um atributo Z , para um bloco B do depósito. Entre elas, aquela citada por Journel e Huijbregts (1978): em cada realização, o valor do bloco B corresponde à média dos valores simulados nos pontos (ou locais) que discretizam o respectivo bloco. Assim, os valores do bloco B , obtido por diferentes realizações fornece a *ccdf* para esse bloco B . Essa metodologia (considerando média aritmética dos valores simulados nos pontos que discretizam o bloco) será utilizada nesse trabalho e pode ser representada matematicamente pela seguinte equação:

$$F(B; z) = \frac{1}{L} \cdot \sum_{l=1}^L i_l(B; z) \quad (\text{Equação 2.27})$$

onde:

$F(B; z)$ = probabilidade do valor do atributo Z para o bloco B ser menor que o teor de corte z ;

L = número de realizações;

$i_l(B; z)$ = valor do l -ésimo indicador para o bloco B e o teor de corte z . Esse indicador pode ser definido da seguinte forma:

$$i_l(B; z) = \begin{cases} 1, & \text{para } z_l(B) \leq z \\ 0, & \text{outros casos} \end{cases} \quad (\text{Equação 2.28})$$

onde:

z = valor do teor de corte;

$z_l(B)$ = valor do atributo Z para o bloco B considerando a l -ésima realização. Esse valor pode ser calculado pela seguinte equação:

$$z_l(B) = \frac{1}{D} \cdot \sum_{d=1}^D z_l(u_d) \quad (\text{Equação 2.29})$$

onde:

D = número de pontos simulados que discretizam o bloco B ;

$z_l(u_d)$ = valor do atributo Z simulado em um local u_d para a l -ésima realização;

u_d = d -ésimo local simulado pertencente ao bloco B .

Após o cálculo do valor de cada bloco, obtém-se a distribuição de blocos para o depósito. A variância da distribuição dos valores dos blocos deve ser menor que a variância da distribuição dos pontos simulados (que representam volume quasi-pontual), pois os blocos representam um volume maior do que os pontos simulados. Segundo a Relação de Krige (Journel e Huijbregts, 1978), a variância da distribuição de blocos pode ser calculada da seguinte forma:

$$\sigma^2(p/A) = \sigma^2(B/A) + \sigma^2(p/B) \quad (\text{Equação 2.30})$$

onde:

$\sigma^2(p/A)$ = variância dos pontos (p) em todo depósito (A);

$\sigma^2(B/A)$ = variância dos blocos (B) em todo depósito (A);

$\sigma^2(p/B)$ = média dos valores de variância dos pontos dentro de cada bloco (B).

A variância dos pontos dentro do bloco ($\sigma^2(p/B)$) pode ser aproximada considerando o modelo variográfico (modelo de continuidade) para os valores do atributo e o número de pontos (q) distribuídos dentro de cada bloco (B) (Journel e Huijbregts, 1978; Isaaks e Srivastava, 1989):

$$\sigma^2(p/B) \cong \frac{1}{q^2} \sum_{i=1}^q \sum_{j=1}^q \gamma(h_{ij}) \quad (\text{Equação 2.31})$$

onde:

q = número de pontos distribuídos dentro de um bloco B ;

γ = variograma para a distância h , entre os pontos i e j pertencentes ao bloco B ;

h_{ij} = distância entre o ponto i e o ponto j .

Sabendo a variância dos pontos em todo depósito ($\sigma^2(p/A)$), aproximada pela variância dos dados amostrais, e a variância dos pontos dentro do bloco ($\sigma^2(p/B)$), pode-se encontrar a variância da distribuição dos blocos ao longo do depósito ($\sigma^2(B/A)$).

A variância da distribuição de blocos obtida pela Relação de Krige será chamada variância teórica da distribuição de blocos ao longo desse estudo. Para cada caso estudado, os cenários simulados devem gerar uma distribuição de blocos cuja variância reproduz (considerando as flutuações ergódicas) a variância teórica. Assim, a média dos pontos simulados dentro do bloco é considerada representativa do valor do bloco e o número de nós simulados dentro de cada bloco é considerado suficiente.

2.4.2 – Função de Transferência - Tonelagem Lavrável

Um dos principais objetivos da avaliação de recursos/reservas de um depósito mineral é o conhecimento da tonelagem lavrável ou tonelagem de minério.

Essa tonelagem é definida em função de um teor de corte.

Ao longo da tese, a eficiência de cada configuração de amostras na redução da incerteza sobre os valores de tonelagem de minério acima de um teor de corte será avaliada.

A configuração de amostras adicionais deve reduzir o desvio padrão (O) dos valores obtidos para a função de transferência (toneladas de minério acima de um teor de corte), conforme:

$$O = \sqrt{\frac{1}{L} \cdot \sum_{l=1}^L (TG(z_k)_l - E[TG(z_k)])^2} \quad (\text{Equação 3.32})$$

onde:

O = desvio padrão dos valores de tonelagem acima de um teor de corte;

$TG(z_k)_l$ = tonelagem considerando o teor de corte z_k para a l -ésima realização;

L = número de realizações;

$E[TG(z_k)]$ = valor esperado ou média aritmética dos valores de tonelagem, para o teor de corte z_k , obtidos para as L realizações.

2.4.3 – Função de Transferência - Valor Presente Líquido (VPL)

Muitas vezes, a mineração de um depósito requer elevados investimentos. Por isso, o retorno financeiro desses investimentos é um aspecto importante que pode ser analisado na etapa de avaliação econômica de reservas do depósito. Em qualquer empreendimento mineiro, informações como o tempo necessário para o retorno do investimento (*payback*) ou o lucro líquido, após um determinado tempo de mineração, auxiliam, não apenas, na decisão sobre a mineração do depósito, como na decisão sobre a escala de produção de minério, na compra de equipamentos, etc.

O parâmetro econômico que será analisado nesse trabalho é o chamado

Valor Presente Líquido (VPL). Esse parâmetro expressa o lucro líquido, trazido para a data zero ou do início do empreendimento, que será obtido após um certo período de mineração. Interpretando o que explicam Hustrulid e Kuchta (1995), se um empréstimo é realizado nos dias atuais (Valor Presente - VP) com uma determinada taxa de juros (i) ao ano, o valor que deve ser pago no fim de um ano (Valor Futuro - VF) é determinado pela seguinte equação:

$$VF = VP \cdot (1+i) \quad (\text{Equação 2.33})$$

Se o empréstimo for quitado, apenas, ao final de n anos, o valor que deverá ser pago no final do n -ésimo ano (VF) é calculado da seguinte forma:

$$VF = VP \cdot (1+i)^n \quad (\text{Equação 2.34})$$

O cálculo do valor futuro pode ser feito para diferentes unidades de tempo, por isso, o valor n pode expressar dias, meses, anos, etc.

A equação 2.34 pode ser utilizada para se obter o valor presente, nesse caso, a equação 2.34 é escrita da seguinte forma:

$$VP = \frac{VF}{(1+i)^n} \quad (\text{Equação 2.35})$$

Para o caso de um empreendimento mineiro, considerando a subtração da receita pelo custo operacional obtidos em um certo ano (n) de mineração, igual ao valor futuro (VF), o valor presente para aquele ano (VP) pode ser obtido pela equação 2.35. O somatório dos valores VP obtidos para diferentes períodos de mineração, corresponde ao VPL.

Considerando i a taxa de juros para cada período de tempo de mineração e I o investimento exigido para o início da mineração, o VPL obtido após o n -ésimo período de mineração pode ser calculado pela seguinte equação:

$$VPL = -I + \frac{VF_1}{(1+i)^1} + \frac{VF_2}{(1+i)^2} + \dots + \frac{VF_n}{(1+i)^n} \quad (\text{Equação 2.36})$$

Geralmente, o cálculo do valor futuro para cada período de mineração, considera, além de valores de receita e custo operacional, valores como a depreciação de equipamento e impostos que incidem sobre a atividade de mineração.

Nesse estudo, diversos cenários para distribuição do atributo de interesse, ao longo do depósito, serão fornecidos por um algoritmo de simulação estocástica. Dependendo do atributo em estudo, cada cenário pode apresentar diferentes teores de minério e/ou tonelagem de minério em um mesmo período de mineração. Assim, cada cenário apresentará diferentes valores de receita para um mesmo período de mineração e, conseqüentemente, diferentes valores de VPL. Uma das funções de transferência que será considerada nesse estudo é o VPL (Figura 2.3) e a configuração de amostras adicionais deve reduzir o desvio padrão (O) dos valores dessa função, obtidos para diferentes realizações (ou cenários), conforme:

$$O = \sqrt{\frac{1}{L} \cdot \sum_{l=1}^L (VPL_l - E[VPL])^2} \quad (\text{Equação 2.37})$$

onde:

O = desvio padrão dos valores de VPL para diferentes realizações;

L = número de realizações;

VPL_l = valor presente líquido para a l-ésima realização;

$E[VPL]$ = valor esperado ou média aritmética dos valores de VPL para as L realizações.

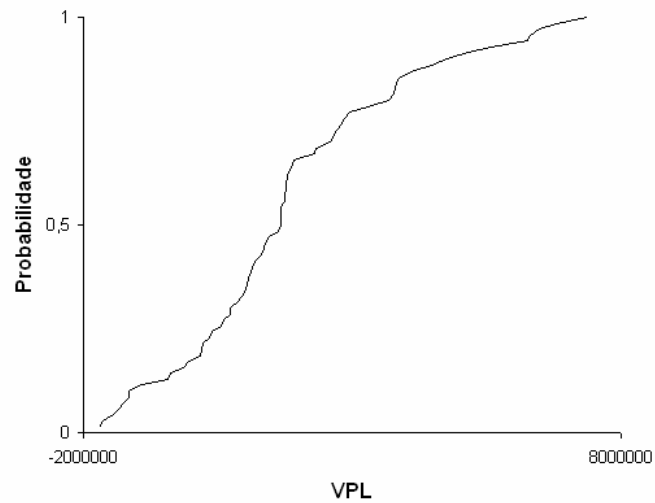


Figura 2.3 – Exemplo de distribuição de valores de VPL obtidos para diferentes realizações.

A mesma cava ótima e sequenciamento de lavra serão considerados para o cálculo do VPL de diferentes cenários de modelos de teores (ou outro atributo que afete a receita).

Capítulo 3 – Implementação da Metodologia – Estudo de Caso em 2D

Conforme salientado anteriormente, quando a nova campanha de amostragem visa reduzir a incerteza sobre uma função de transferência, a locação de amostras adicionais baseada nas regiões de maior incerteza sobre o atributo de interesse pode trazer um maior benefício (maior redução da incerteza) do que a adição de amostras em uma malha regular. Porém, algumas questões devem ser avaliadas. Por exemplo, para quantas amostras iniciais, ou qual atributo, passa a haver diferença significativa na redução de incerteza para as diferentes configurações de amostras? Ou seja, quando a otimização da configuração de amostras passa a ser realmente útil para reduzir o risco associado a uma determinada função de transferência que depende de uma dada variável?

As configurações de amostras adicionais propostas nesse e no próximo capítulo são comparadas aos resultados advindos do uso da amostragem regular na redução da incerteza sobre os resultados das funções de transferência. A análise dos desempenhos das configurações de amostras é baseada em um cenário hipotético para distribuição dos valores do atributo, ao longo da região em estudo. Assim, outra questão importante a ser analisada é a reprodução dos resultados encontrados na realidade pelos resultados baseados no cenário hipotético.

O estudo apresentado a seguir serve de exemplo da aplicação da metodologia para atributos bi-dimensionais, ou seja, atributos cujos valores podem ser posicionados por, apenas, duas coordenadas. Como exemplo desses atributos pode-se citar espessura de uma camada de carvão. Por exemplo, no caso de uma camada de carvão horizontal, cada sondagem vertical informa a espessura da camada nas coordenadas leste (x) e norte (y) em que a sondagem foi realizada. Assim, cada dado de espessura é locado, apenas, pelas coordenadas x e y da respectiva sondagem, após a simulação se obtêm valores de espessura da camada para outras coordenadas x e y. Também se pode trabalhar em duas dimensões com

um atributo em que cada valor representa toda espessura da camada, como teor de cinza da camada da camada de carvão, teor de um contaminante na camada de solo, entre outros. Contudo, quando o atributo é teor, os dados devem representar um mesmo suporte (amostras de mesma massa) para serem simulados. Se a camada apresenta espessura e densidade variável ao longo da área em estudo, a variável inicial deve ser transformada em uma variável “acumulada” (por exemplo, Gambin, 2003), antes da simulação, mas ainda assim, a nova variável pode ser simulada em duas dimensões.

Nesse capítulo, os desempenhos das diferentes configurações são comparados para o caso bi-dimensional. Para cada configuração, a incerteza sobre a função de transferência é aproximada. Além disso, diferentes funções de transferência são consideradas.

3.1 – Variável V (ppm)

A primeira variável usada para análise da metodologia é a variável V (ppm), do conhecido banco de dados *Walker Lake* (Isaaks e Srivastava, 1989). Esse banco de dados bi-dimensional possui uma malha de dados em 1 x 1 m nas direções Leste-Norte (x) e Norte-Sul (y). Por apresentar-se exaustivamente amostrado, esse conjunto de dados será usado como referência ou cenário real da variável V. O mapa e histograma para o cenário real da variável V são apresentados na Figura 3.1.

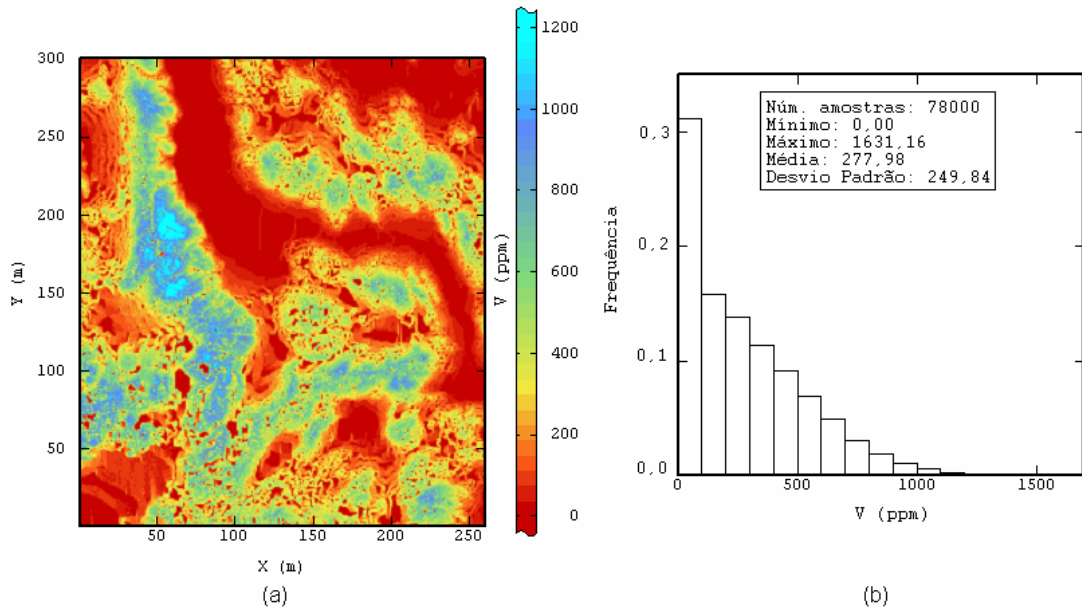


Figura 3.1 – (a) Mapa para o cenário real da variável V. (b) Histograma para o cenário real da variável V.

Um banco de dados com amostras extraídas do conjunto exaustivo foi criado, representando dados iniciais. O conjunto de dados de referência foi amostrado com uma malha de 20 x 20 m ao longo das direções Leste-Norte (x) e Norte-Sul (y), imitando o que seria um conjunto de dados típico de uma campanha inicial de amostragem. O mapa de localização e histograma para os dados iniciais são apresentados na Figura 3.2.

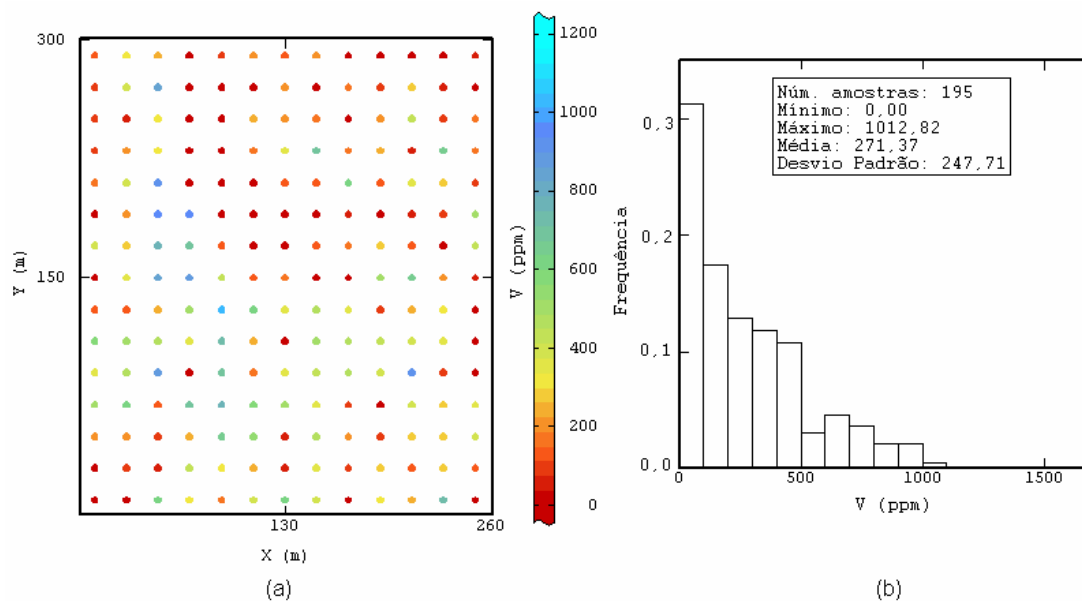


Figura 3.2 – (a) Mapa de localização para os dados iniciais da variável V. (b) Histograma para os dados iniciais da variável V.

Calculou-se a incerteza sobre os resultados obtidos com cada função de interesse, aplicadas sobre o modelo de teores simulados usando apenas os dados iniciais. A simulação dos teores foi obtida usando o método de Bandas Rotativas.

A simulação foi realizada em pontos com espaçamento de 1 x 1 m nas direções x e y. Conforme explicado no Capítulo 2, a simulação necessita informações sobre a continuidade espacial (modelo variográfico) dos valores da variável de interesse. Assim, um modelo variográfico foi ajustado aos variogramas experimentais para os valores normalizados da variável V e esse modelo (Tabela 3.1) foi utilizado para condicionar a simulação. Os parâmetros de busca de amostras utilizados na simulação de cada nó de grid são iguais ao alcance do modelo de continuidade em cada direção. Além disso, um número mínimo de 3 dados originais foi selecionado para simulação de cada nó com no máximo 2 amostras por octante de busca.

Conforme comentado no Capítulo 2, recomenda-se a verificação da binormalidade dos dados normalizados para verificar a aderência dos dados às hipóteses requeridas para o emprego do método de simulação por Bandas Rotativas. A Tabela 3.2 apresenta um teste de binormalidade para os dados iniciais, e os resultados encontrados não permitem rejeitar a hipótese de binormalidade.

Tabela 3.1 - Modelo variográfico utilizado na simulação para os dados no espaço original e normalizado

	Primeira Estrutura					Segunda Estrutura			
	c0	Modelo	c1	a (D1)	a (D2)	Modelo	c1	a (D1)	a (D2)
Original	4400	Esférico	37300	38 m	35 m	Esférico	19662,4	150 m	36 m
Normal	0,07	Esférico	0,61	38 m	35 m	Esférico	0,32	150 m	36 m

D1 - Direção de Maior Continuidade - N157,5°.

D2 - Direção de Menor Continuidade – N67,5°.

c1 – Patamar.

a – Alcance.

c0 - Efeito pepita.

Os diversos cenários (realizações) simulados devem reproduzir (considerando as flutuações ergódicas) o mapa, histograma e variograma obtidos para os dados iniciais. Os mapas, histogramas e variogramas para as realizações que apresentaram maior e menor variância para os valores de V são apresentados nas Figuras 3.3, 3.4 e 3.5. Os mapas, histogramas e variogramas para essas

realizações são semelhantes aos construídos para os dados iniciais.

Tabela 3.2 – Teste de binormalidade para os dados iniciais

Dados Iniciais - Omnidirecional Horizontal		
Número de Pares	Distância Média	Madograma / Raiz Quadrada do Variograma
698	24	0,546
954	43	0,556
1160	61	0,561
2034	81	0,565
1578	103	0,563
1960	122	0,564
1710	141	0,572
1748	160	0,572
2022	181	0,577
1330	202	0,572
1324	221	0,575
898	240	0,568
750	260	0,572
425	280	0,578

Os parâmetros estatísticos média e desvio padrão das diferentes realizações flutuaram ao redor dos respectivos parâmetros para os dados iniciais, conforme esperado. A Figura 3.6 apresenta a distribuição acumulada de cada um desses parâmetros, obtidos para os diversos cenários.

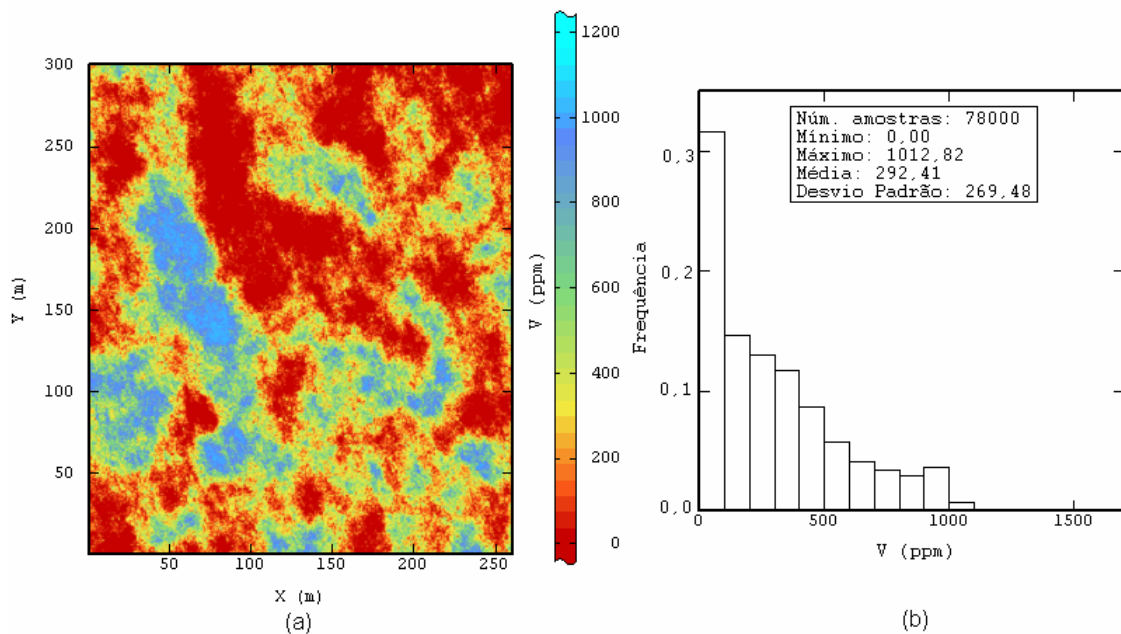


Figura 3.3 – (a) Mapa para a realização que apresentou maior variância dos valores de V. (b) Histograma para a realização que apresentou maior variância dos valores de V.

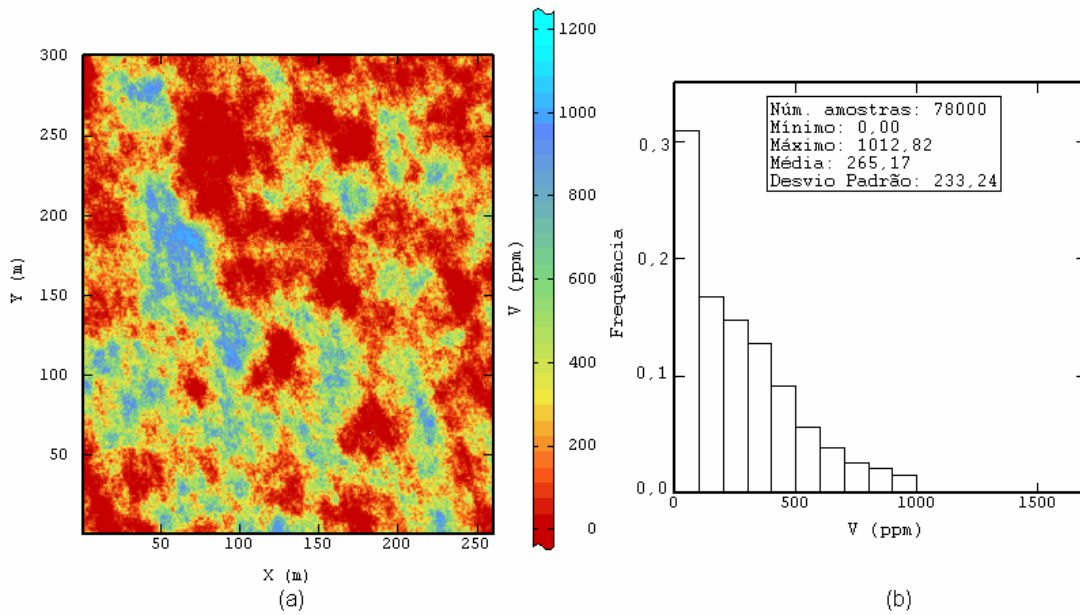


Figura 3.4 – (a) Mapa para a realização que apresentou menor variância dos valores de V. (b) Histograma para a realização que apresentou menor variância dos valores de V.

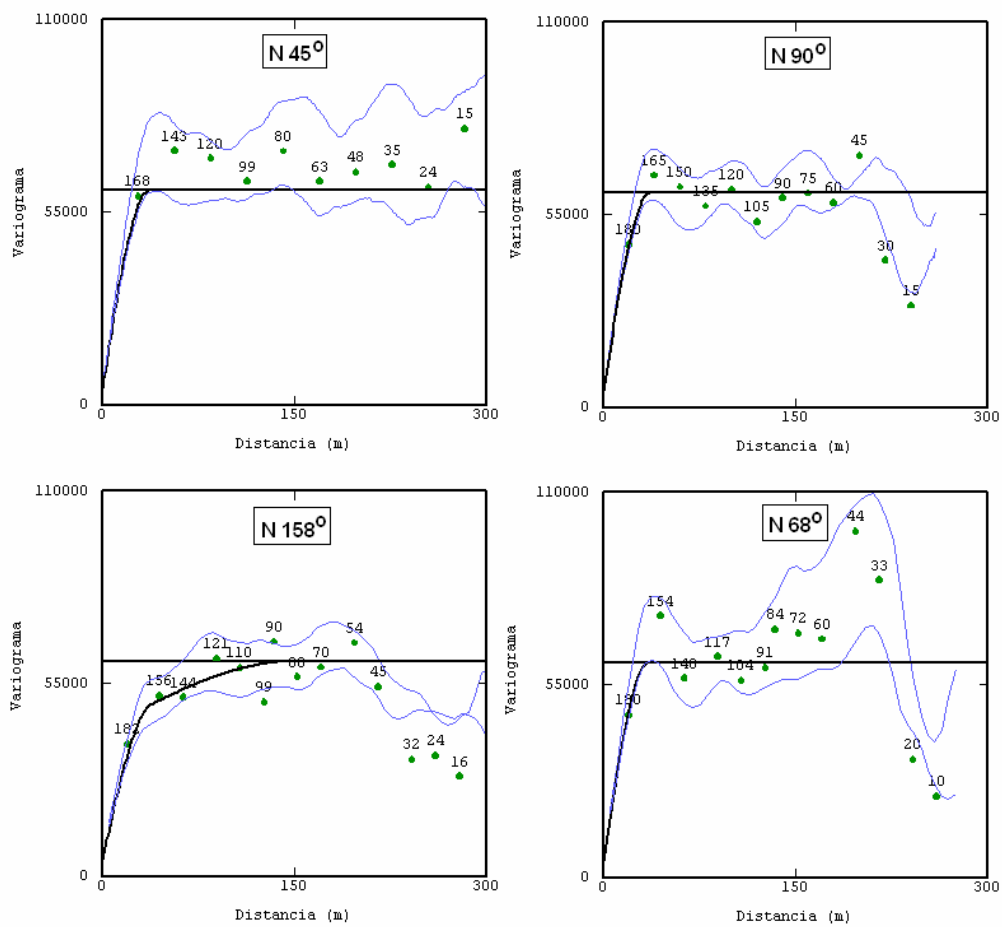


Figura 3.5 – Variograma para a variável V. As linhas azuis representam os variogramas para as realizações que apresentaram maior e menor variância dos valores de V. Os pontos representam o variograma experimental para os dados iniciais (dados originais). A linha preta representa o modelo variográfico (Tabela 3.1) para a variável V.

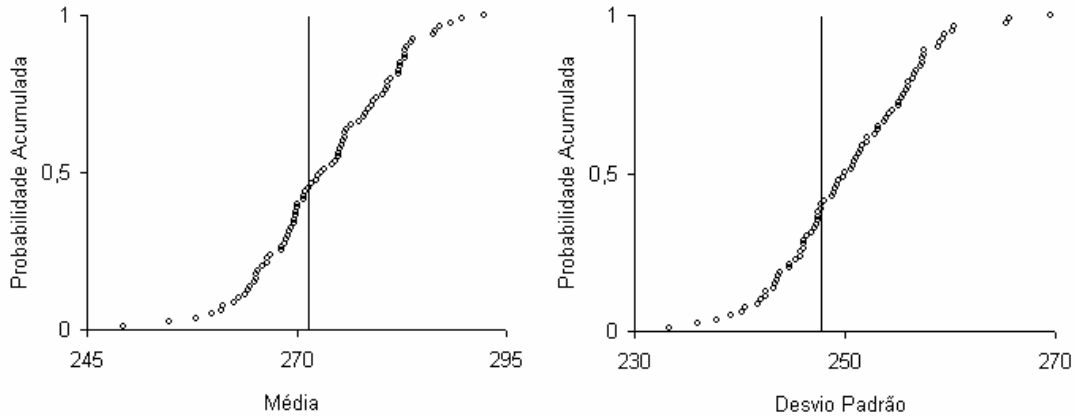


Figura 3.6 – Distribuição acumulada das médias (a) e valores de desvio padrão (b) das realizações obtidas a partir dos dados iniciais. Cada circunferência representa o valor de média (ou desvio padrão) para uma realização. A linha representa a média (ou desvio padrão) para os dados iniciais.

A partir dos valores simulados, em cada realização, foram calculados os valores de V para 3120 blocos de 5×5 m nas direções x e y , visto que o cálculo das funções de interesse é baseado nos valores dos blocos que serão lavrados. A média aritmética dos valores de V para pontos que pertencem a um mesmo bloco (no total de 25 pontos por bloco) corresponde ao valor de V para o respectivo bloco. A distribuição dos teores dos blocos deve apresentar variância próxima (considerando as flutuações ergódicas) da variância teórica da distribuição dos blocos (segundo a relação de Krige, conforme comentado no Capítulo 2), desse modo o número de pontos simulados dentro do bloco é considerado suficiente para que o valor do bloco seja aproximado pela média dos valores dos pontos.

Os valores de variância dos teores de blocos para as diferentes realizações obtidas com os dados iniciais flutuam ao redor da variância teórica dos blocos (aproximadamente 51669 ppm^2), conforme esperado. A Tabela 3.3 apresenta a estatística para as realizações no suporte de blocos que apresentaram variância próxima à mediana, a menor variância e a maior variância.

Tabela 3.3 – Estatística para realizações no suporte de blocos

	Mínimo	Máximo	Média	Variância
Realização	0	1001	265	43850
	0	1008	258	52174
	0	1006	292	61907

Cada cenário simulado produz um resultado distinto para a função de interesse. A distribuição dos valores da função para os diversos cenários simulados permite o acesso à incerteza sobre o valor da função.

O número de cenários simulados deve ser suficiente para garantir o acesso ao espaço de incerteza sobre a função de interesse. Assim, para cada função analisada, foi construído um gráfico: número de realizações versus desvio padrão dos valores da função obtidos para as realizações. O número de cenários para o qual o desvio padrão dos valores da função torna-se aproximadamente estável foi considerado o número mínimo de cenários para mapear o espaço de incerteza desta dada função.

Nesse estudo, duas funções de transferência foram consideradas: tonelagem de minério acima do teor de corte e o valor presente líquido (VPL) para uma cava projetada seguindo duas rotas de lavra. A Figura 3.7 mostra o número de realizações versus desvio padrão dos valores da função. Foram geradas 80 realizações, obtidas com os dados iniciais. Esse número de realizações foi considerado suficiente para o acesso ao espaço de incerteza, sobre cada função, visto que os gráficos da Figura 3.7 mostram estabilização do desvio padrão para um número de realizações inferior a 80.

Para simplificação do cálculo da função tonelagem de minério acima do teor de corte, a tonelagem de cada bloco foi considerada uma unidade de massa. A tonelagem dos blocos que apresentam teor maior ou igual ao teor de corte representa o valor da função para o respectivo teor de corte. Os teores de corte 40 ppm, 150 ppm, 300 ppm e 500 ppm foram considerados.

Uma distribuição dos valores de tonelagem acima do teor de corte pode ser construída considerando os diversos cenários obtidos a partir dos dados iniciais. A Figura 3.8 (a) apresenta o valor mínimo, o valor máximo, o quartil inferior e o superior (pontos pretos) da distribuição dos valores de tonelagem, para cada teor de corte considerado. Essa figura, também, apresenta os valores de tonelagem, obtidos usando o modelo exaustivo de V (referência), para cada teor de corte. Analisando-se a Figura 3.8 (a), percebe-se que, para os teores de corte 40 ppm, 300 ppm e 500 ppm, o valor de tonelagem obtido para a referência pertence ao intervalo limitado pelos quartis superior e inferior da distribuição de valores da função. Isso não ocorre para o teor de corte 150 ppm (porém está incluído dentro do intervalo min-max

derivados das múltiplas simulações), onde existe uma tendência dos valores de tonelagem obtidos por simulação serem menores que a tonelagem real, mostrando uma dificuldade da Configuração Inicial de reproduzir em média a tonelagem real para esse teor de corte.

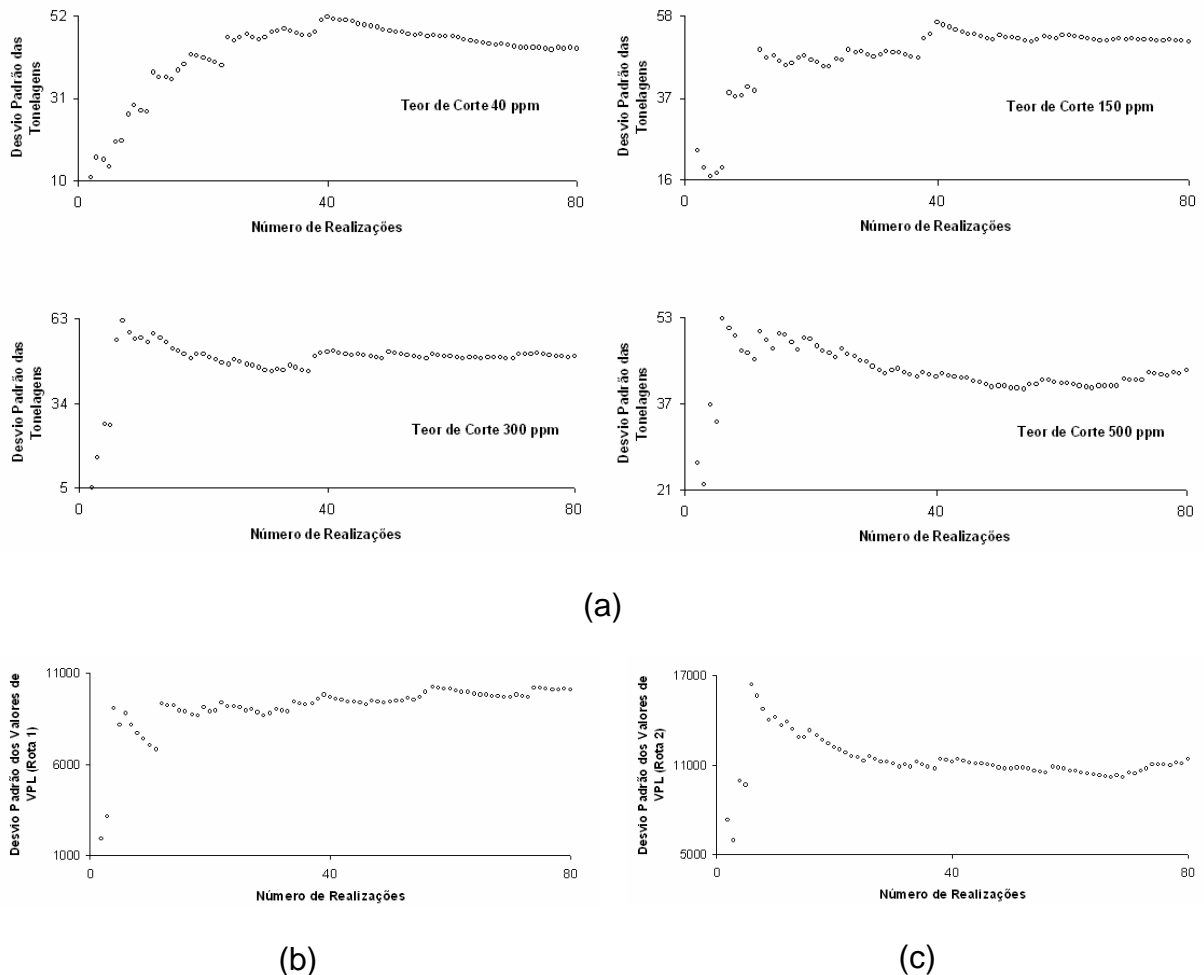
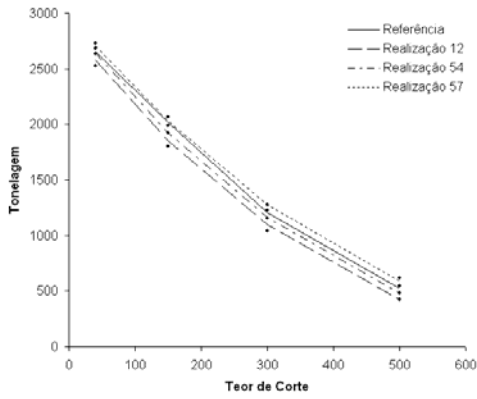


Figura 3.7 – Número de realizações versus desvio padrão dos valores da função. (a) Gráfico para tonelagem de minério acima do teor de corte. Foram impressos os resultados para quatro teores de corte. (b) Gráfico para VPL (rota 1). (c) Gráfico para VPL (rota 2).

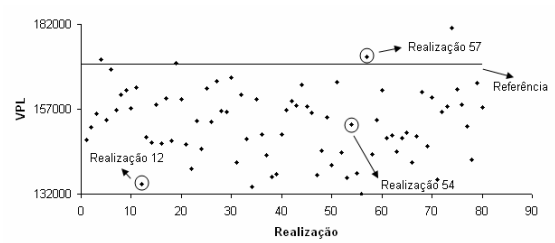
A função VPL foi calculada para duas rotas de lavra. A rota 1 consiste da lavra ao longo da direção y, avançando da esquerda para direita no depósito, sendo que duas linhas de blocos são lavradas em cada ano. A rota 2 consiste da lavra na direção x, iniciando no extremo sul do depósito, sendo que duas linhas de blocos são lavradas em cada ano. Para ambos os casos, o custo operacional foi considerado o mesmo para todos os anos e a receita bruta (sem custo operacional) com a venda de cada bloco foi considerado igual ao teor (ppm) do bloco. Além disso,

a taxa de desconto foi considerada 12% ao ano. As Figuras 3.8 (b) e (c) apresentam os valores de VPL, respectivamente para as rotas 1 e 2, para as realizações obtidas com os dados iniciais. Essas figuras também apresentam os valores de VPL obtidos para a referência.

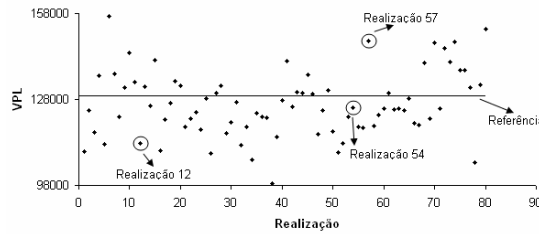
Analisando-se a Figura 3.8 (b) observa-se que existe uma tendência dos valores de VPL para a rota 1, obtidos pela realizações, serem menores que o valor de VPL (rota 1) para a referência. Isso evidencia a suavização de valores extremos, nos cenários simulados, na porção noroeste do depósito, que possui grande impacto no valor do VPL calculado para a rota 1. Essa suavização não causa tanto impacto nos valores de VPL calculados para a rota 2 (Figura 3.8 (c)), já que esses valores são mais influenciados pelos teores dos blocos pertencentes à porção sul do depósito. Esse fato, também evidencia que a amostragem perdeu alguns valores extremos, o que contribuiu para essa suavização, os histogramas para os dados da referência e para as amostras iniciais mostram uma menor frequência de valores extremos para as amostras iniciais. Essa suavização pode ser observada na Figura 3.9. Selecionou-se no cenário exaustivo os 300 blocos com maiores valores de V e calculou-se a média desses blocos. Considerando os mesmos 300 blocos, calculou-se a média de V para cada realização (obtida a partir dos dados iniciais). A Figura 3.9 mostra que para todas as realizações a média dos 300 blocos é inferior àquela encontrada para o cenário exaustivo.



(a)



(b)



(c)

Figura 3.8 – (a) Gráfico de tonelage acima do teor de corte. Pontos pretos representam o valor mínimo, o valor máximo, o quartil inferior e o superior da distribuição dos valores de tonelagens, obtidos pelos diversos cenários, para cada teor de corte. (b) Valores de VPL (rota de lavra 1) para as realizações geradas com os dados iniciais. (c) Valores de VPL (rota de lavra 2) para as realizações geradas com os dados iniciais.

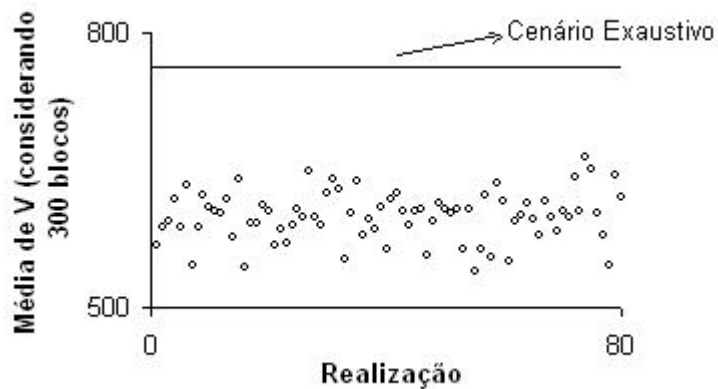


Figura 3.9 – Média de V para os 300 blocos. Considerando os blocos com os 300 maiores valores de V no cenário exaustivo.

Para o desenvolvimento da metodologia de otimização da configuração de amostras, visando à redução da incerteza sobre a função de interesse, uma questão importante deve ser analisada: a incerteza sobre a função de interesse modifica significativamente com a mudança do cenário simulado escolhido como fonte para a busca de informações (amostras) adicionais?

A idéia inicial é que a configuração otimizada de amostras resultante do processo desenvolvido não seja modificada significativamente se for utilizado um cenário simulado diferente como referência para a busca de informações (amostras) adicionais. Assim, para qualquer cenário escolhido para representar a realidade dos teores *in situ*, a adição de amostras nos locais determinados pelo processo de otimização resultaria em uma eficiente redução de incerteza. Por isso, para análise dessa questão alguns cenários simulados foram escolhidos para serem utilizados como fonte de informações adicionais e os resultados obtidos para esses cenários foram comparados aos resultados obtidos para quando buscou-se amostras adicionais utilizando como referência os teores exaustivamente amostrados (o cenário real).

A fim de avaliar o comportamento da metodologia para cenários significativamente diferentes quanto ao valor da função de interesse, três cenários foram escolhidos como fonte de informação adicional: realização 12, realização 54 e realização 57. Os valores das funções consideradas, calculados sobre os modelos de teores obtidos para esses cenários, são apresentados nas Figuras 3.8 (a) a (c). As realizações 12, 57 e 54 apresentam, respectivamente, baixos, altos e valores intermediários para função tonelagem (teores de corte 40, 150, 300 e 500 ppm) e VPL (rota 1 e 2), em relação à maioria das realizações obtidas com os dados iniciais.

Quatro configurações de amostras são testadas nesse estudo. Em cada configuração, 160 valores de V (ppm) serão adicionados aos 195 valores de V , obtidos na amostragem inicial.

A Configuração 1 foi elaborada de modo que os 160 dados adicionais formassem com os 195 dados existentes uma malha aproximadamente regular. Assim, essa configuração representaria um tipo de amostragem comumente aceito.

A Configuração 2 foi construída para que as amostras fossem distribuídas próximas aos locais (nós do grid) com elevado IEQ (Intervalo entre Quartis) para os possíveis teores. Em cada local simulado, os valores de IEQ foram calculados

considerando os 80 possíveis valores do atributo, gerados para as realizações obtidas com os dados iniciais. A idéia dessa configuração de amostras foi avaliar a eficiência da adição simultânea de várias amostras, nos locais de elevado IEQ, na redução da incerteza para a função de transferência. Contudo, a configuração 2 pode não ser tão eficiente na redução da incerteza sobre a função devido à adição simultânea de todas amostras próximas aos locais de elevado IEQ. A adição de uma porção das amostras em uma região de elevado IEQ pode reduzir a incerteza nessa região de forma que outras regiões passem a ser mais carentes de amostras. Nesse caso, as amostras restantes devem ser adicionadas nas novas regiões de elevada incerteza. Esse procedimento pode garantir uma maior redução na incerteza sobre uma função global, do que a adição simultânea de todas amostras. Por isso, com intuito de melhorar os resultados obtidos para Configuração 2, sugeriu-se a Configuração 3.

A Configuração 3, assim como a Configuração 2, foi construída para que as amostras fossem distribuídas próximas aos locais com elevado IEQ. Porém, nessa configuração, a adição das amostras foi feita em etapas. Na primeira etapa, 30 amostras foram distribuídas próximas aos locais de elevado IEQ, obtidos com a simulação para os dados iniciais. Realizou-se a simulação para os 225 dados de V. Na segunda etapa, 50 amostras foram distribuídas nos locais de elevado IEQ, obtidos com a simulação para os 225 dados de V. Realizou-se a simulação para os 275 dados de V. Na terceira etapa, 40 amostras foram distribuídas nos locais de elevado IEQ, obtidos com a simulação para os 275 dados de V. Realizou-se a simulação para os 315 dados de V. Na quarta etapa, 40 amostras foram distribuídas nos locais de elevado IEQ, obtidos com a simulação para os 315 dados de V. Realizou-se a simulação para os 355 dados de V, que formam a chamada configuração 3.

A Configuração 4 foi construída para que as amostras fossem distribuídas nos locais de baixo IEQ. Teoricamente, essa configuração de amostras deve ser menos eficiente na redução da incerteza sobre a função de transferência do que as configurações 1, 2 e 3, já que as amostras serão adicionadas nas regiões de baixa incerteza sobre os teores locais. Essa configuração foi proposta para permitir a avaliação do benefício das outras configurações de amostras, ou seja, se a configuração de amostras realmente precisa ser otimizada ou se a redução de incerteza para qualquer configuração, com o mesmo número de amostras adicionais,

não apresenta diferença significativa.

As quatro configurações de amostras foram testadas para a referência e a realização 12, usadas como fonte de informação adicional. A configuração de amostras que resultaria em uma maior redução da incerteza (para a referência exaustiva) espera-se ser a mesma que seria obtida considerando qualquer realização como fonte de informação. Assim, garantir-se-ia que a configuração de amostras mais eficiente, obtida pelo processo de otimização, seja a mesma que seria obtida se tivéssemos acesso aos valores reais nos pontos escolhidos para receber amostras adicionais.

3.1.1 – Dados Adicionais Obtidos pelo Cenário Real (exaustivo)

As quatro configurações de amostras testadas para a referência são apresentadas nessa seção. Essas configurações apresentam 160 dados, obtidos a partir da referência, adicionados aos 195 dados iniciais. As Figuras 3.10 a 3.13 apresentam os mapas e histogramas para as configurações 1 a 4, respectivamente. Os histogramas para as configurações 1 a 4 consideram pesos de desagrupamento gerados pelo software Isatis[®] (o peso de cada amostra é inversamente proporcional ao número de amostras que pertencem ao quadrado de 20 x 20 m, nas direções x e y, centrado na amostra). Para cada configuração, os dados foram normalizados considerando os pesos de desagrupamento, para serem usados na etapa de simulação geoestatística.

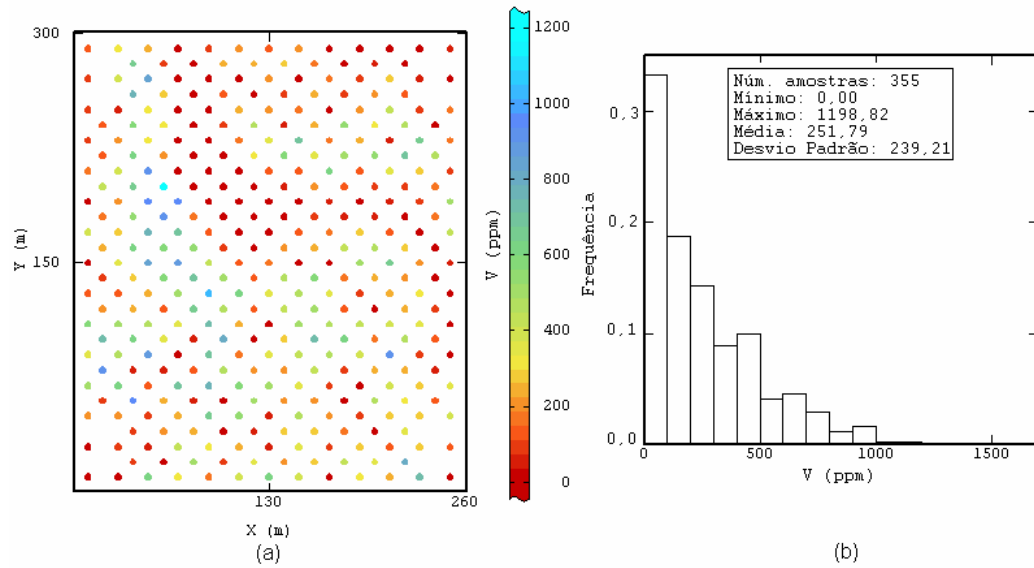


Figura 3.10 – (a) Mapa para a Configuração 1. (b) Histograma para os dados da Configuração 1.

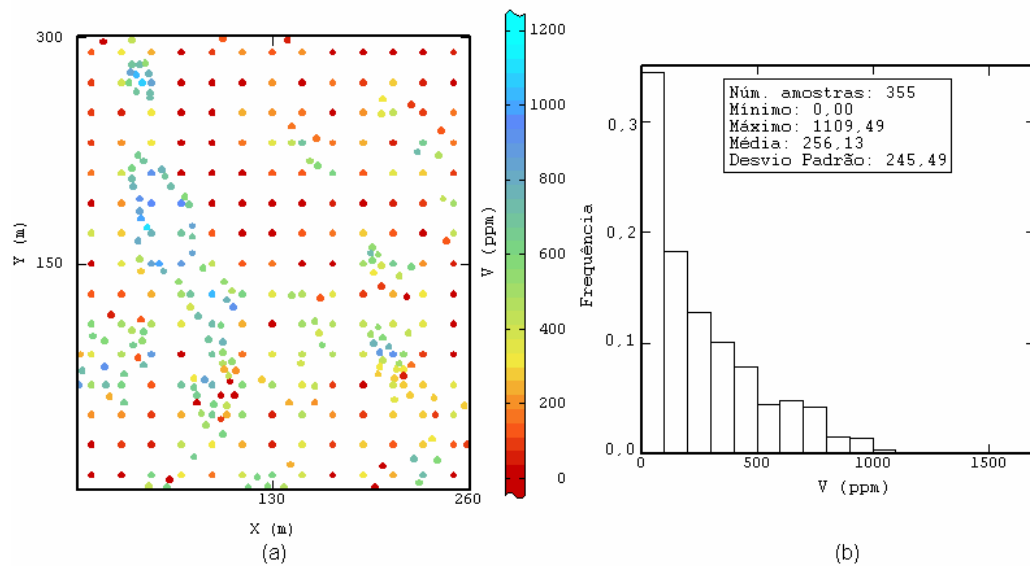


Figura 3.11 – (a) Mapa para a Configuração 2. (b) Histograma para os dados da Configuração 2.

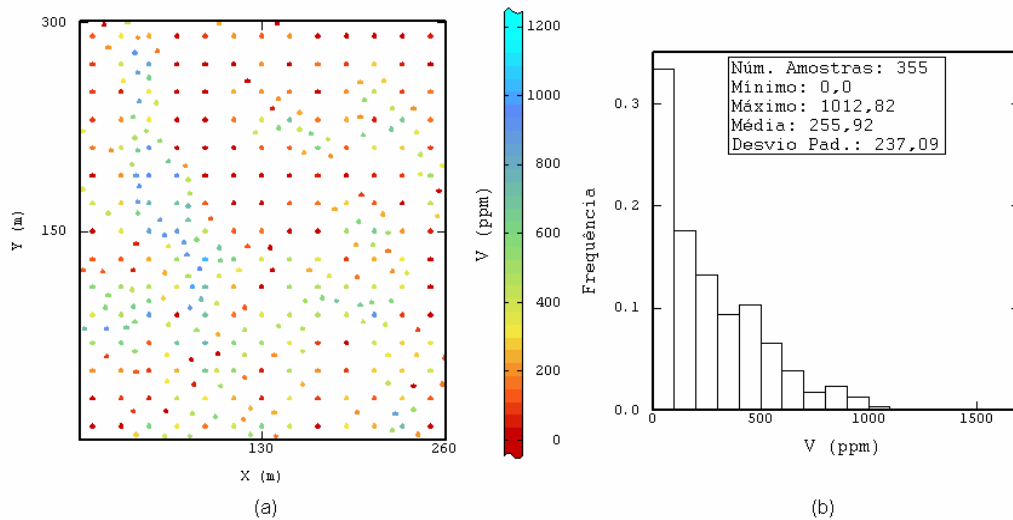


Figura 3.12 – (a) Mapa para a Configuração 3. (b) Histograma para os dados da Configuração 3.

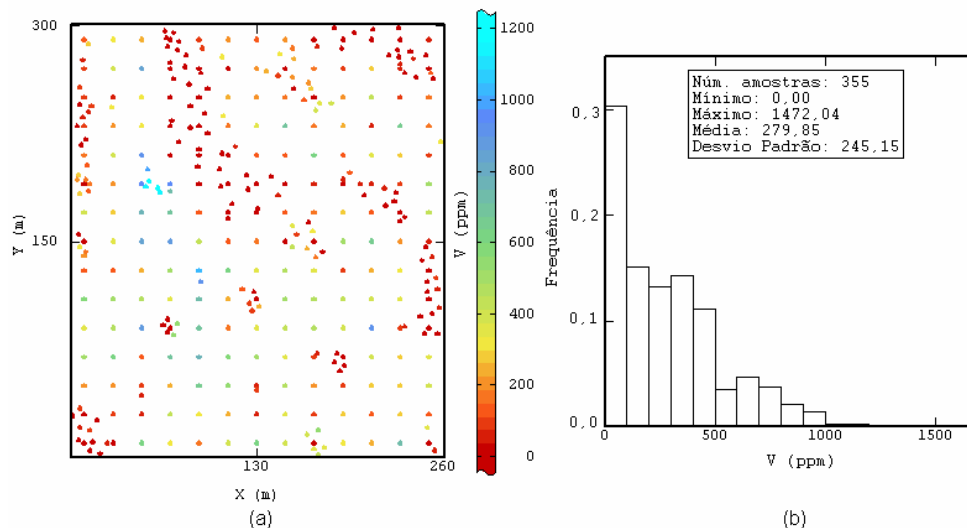


Figura 3.13 – (a) Mapa para a Configuração 4. (b) Histograma para os dados da Configuração 4.

Oitenta novas realizações foram geradas para cada configuração de amostra adicional. Para todas configurações testadas, as estatísticas para cada função de transferência são apresentadas nas Tabelas 3.4 a 3.9 e Figuras 3.14 a 3.19. Essas estatísticas compreendem medidas de acurácia (reprodução em média do valor obtido com o modelo de teores real - exaustivo) e precisão (incerteza). Para cada configuração, a média dos valores das funções de transferência, obtidos para as 80 realizações, foi escrita como porcentagem do valor real da função, ou valor da função utilizando o modelo de teores exaustivo ou de referência. Quanto mais próxima de 100%, mais acurada é a média. A média dos valores da função também será chamada valor esperado para a função ou $E[F]$. A incerteza sobre os valores

das funções, obtidos para as 80 realizações, foi medida pelo desvio padrão (σ) e IEQ.

Para cada configuração testada, um intervalo de possíveis valores da função de transferência é gerado. Assim, se supõe que o verdadeiro valor da função de transferência pertença à esse intervalo. O desvio padrão e o IEQ medem o espalhamento desse intervalo. Quanto menor o intervalo, maior a certeza sobre o verdadeiro valor da função. Portanto, a média dos valores da função para a configuração e a precisão desses valores devem ser analisadas simultaneamente. A combinação desses parâmetros (média \pm erro) deve gerar um intervalo de possíveis valores da função que englobe o valor real.

As Tabelas 3.5, 3.7 e 3.9 apresentam o mínimo e máximo da distribuição dos valores da função para cada configuração (dados adicionais obtidos a partir da referência). Os valores mínimo e máximo limitam o intervalo onde se encontra 100% dos possíveis valores da função, ou teoricamente, o intervalo que possui 100% de chance de conter o valor real. Na maioria dos casos apresentados, o valor da função obtido com o modelo de teores exaustivo ou de referência pertence a esse intervalo. Contudo, existem casos onde o valor da função obtido com o modelo de teores exaustivo ou de referência, apesar de próximo de um dos limites, encontra-se fora desse intervalo. Isto, expressa dificuldade (não garantia) do intervalo obtido a partir dos cenários simulados englobarem o verdadeiro valor da função.

Primeiro, a função tonelagem será analisada. Apesar da diferença entre os índices de incerteza (desvio padrão e IEQ) não ser muito significativa na maioria dos casos, alguns aspectos podem ser observados. A configuração 1 reduziu a incerteza com eficiência para todos teores de corte, em comparação com as configurações 2, 3 e 4. As configurações 2 e 3 ao contemplar áreas de elevado IEQ, adicionaram a maioria dos dados na região de altos valores, reduzindo a incerteza para o teor de corte 500 ppm de forma semelhante a configuração 1. Ainda sobre as configurações 2 e 3, a incerteza para 150 ppm continuou semelhante aos valores encontrados para os dados iniciais, o que mostra que se o objetivo da amostragem fosse a redução da incerteza sobre a tonelagem acima de 150 ppm, essas configurações significariam desperdício de amostras. A configuração 4 adicionou dados nas regiões de baixa incerteza (IEQ) para V, que coincidem na maioria dos locais a baixos valores de V, ajudando a reduzir a incerteza (desvio padrão) para a tonelagem acima do teor de

corde 40 ppm. Por outro lado, essa configuração modificou pouco a incerteza para a tonelagem acima de 500 ppm, em relação aos resultados obtidos usando somente os dados iniciais.

Analisando-se a função VPL, rotas 1 e 2, as configurações 1 a 3 tendem a apresentar menor incerteza sobre o valor do VPL do que a configuração 4, como era esperado. Isso pode ser explicado pelo fato de que as configurações 1 a 3 adicionaram um maior número de amostras na região de alta incerteza (IEQ) (região oeste), onde o valor de cada bloco pode mudar significativamente em cada cenário, e assim, o VPL calculado. Além disso, apesar da pequena diferença, a configuração 3 apresenta valores das medidas de incerteza menor do que a configuração 2 para ambas as rotas de lavra. Isso mostra o benefício da adição de amostras em etapas nos locais de elevado IEQ.

Comparando os resultados de incerteza sobre o VPL, rotas 1 e 2, obtidos para as configurações 1 e 3 observa-se que existe uma tendência dos valores de incerteza para a configuração 3 serem menor do que os valores para a configuração 1. Contudo, a diferença entre as medidas de incerteza para essas configurações é pouco significativa. Por exemplo, considere a distribuição dos valores de VPL, para cada configuração testada, uma distribuição próxima à normal. Para uma distribuição normal, a expressão $(E[F]+2\sigma)$ define o limite superior do intervalo de confiança (confiança de 95%) e $(E[F]-2\sigma)$ define o limite inferior (Christmann, 1978). Esse intervalo de confiança é aquele que possui 95% de probabilidade de incluir o valor do cenário real. A fim de comparar o desempenho das duas configurações na redução da incerteza sobre a função, o valor de erro, expresso por 2σ , foi dividido pelo valor da função para a referência: $[\pm(2\sigma/fr)]$, onde fr é o valor da função para a referência. Desse modo, o erro em relação ao valor da função, para o VPL (rota 1), é aproximadamente 8,3% para a configuração 1 e aproximadamente 7,6% para a configuração 3, a diferença entre o erro obtido para as duas configurações em relação ao valor da função é 0,7%, que é considerado pouco significativo. No caso da rota 2, essa diferença também é pequena: aproximadamente 1%.

A diferença entre as eficiências das configurações na redução de incerteza sobre o VPL pode ser avaliada de uma outra forma. O desvio padrão para a distribuição de valores de VPL, obtido para cada configuração com amostras adicionais, pode ser dividido pelo desvio padrão obtido para a Configuração Inicial.

Ou seja, a diferença entre as configurações é avaliada considerando a incerteza que pode ser reduzida com a adição de novas amostras. Considerando o VPL, rota 1, o desvio padrão para a Configuração 1 representa 70% do desvio padrão para a Configuração Inicial, enquanto o desvio padrão para a Configuração 3 representa 64%. Dessa forma a diferença entre as configurações é um pouco maior.

Analisando os resultados da Tabela 3.6, por exemplo, observa-se que o desvio padrão para configuração 3 é menor que para configuração 1, porém o IEQ é um pouco maior ou praticamente igual ao da configuração 1. O IEQ mede o espalhamento dos valores centrais da distribuição e o desvio padrão é sensível a valores extremos da distribuição. Nesse exemplo, os valores de desvio e IEQ mostram que a diferença das distribuições de VPL para as configurações 1 e 3 está no espalhamento dos valores mais extremos, supõem-se que a maioria desses valores estejam mais afastados do centro da distribuição para a configuração 1.

Analisando-se os resultados para tonelagem acima do teor de corte e VPL, observa-se que algumas configurações geraram um intervalo de possíveis valores da função que não englobou o valor real (ou de referência) dessa função. Contudo, a diferença do valor de referência para o limite do intervalo (limite mais próximo) é menor que 5% em todos casos e os intervalos obtidos ainda são considerados bons guias sobre o valor real.

Tabela 3.4 – Incerteza sobre a tonelagem acima do teor de corte, para diferentes configurações de amostras e quatro teores de corte

Teor de Corte (ppm)	Configuração Inicial			Configuração 1			Configuração 2			Configuração 3			Configuração 4		
	E[F]/fr	σ	IEQ	E[F]/fr	σ	IEQ	E[F]/fr	σ	IEQ	E[F]/fr	σ	IEQ	E[F]/fr	σ	IEQ
40	100%	44	54	100%	31	36	100%	41	54	100%	40	54	101%	29	46
150	97%	51	64	95%	40	42	95%	52	66	96%	49	61	100%	45	62
300	98%	50	72	93%	38	56	96%	45	65	97%	39	51	104%	47	52
500	98%	43	66	89%	29	38	99%	30	42	91%	29	35	99%	43	60

Nota: Tonelagem acima do teor de corte para a referência

2653 (**40**); 2016 (**150**); 1207 (**300**); 521 (**500**).

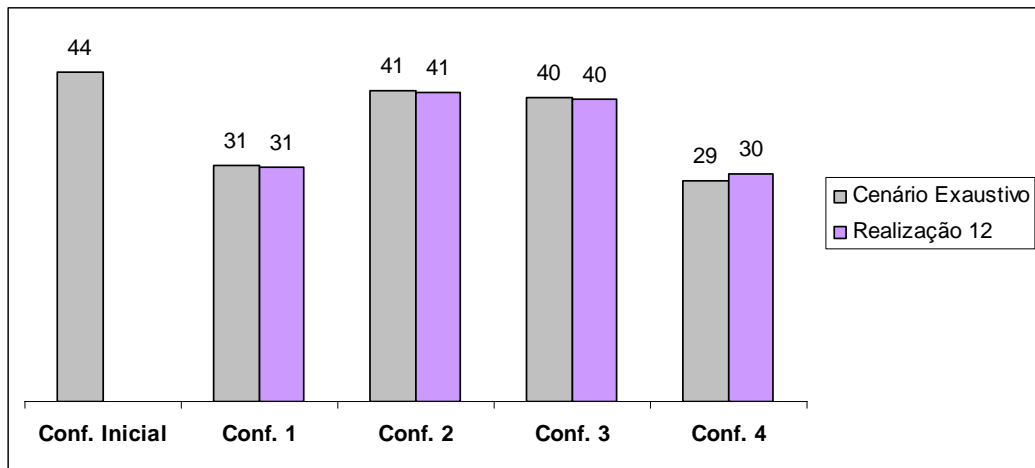


Figura 3.14 – Desvio padrão para a função tonelagem acima de 40 ppm. Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional.

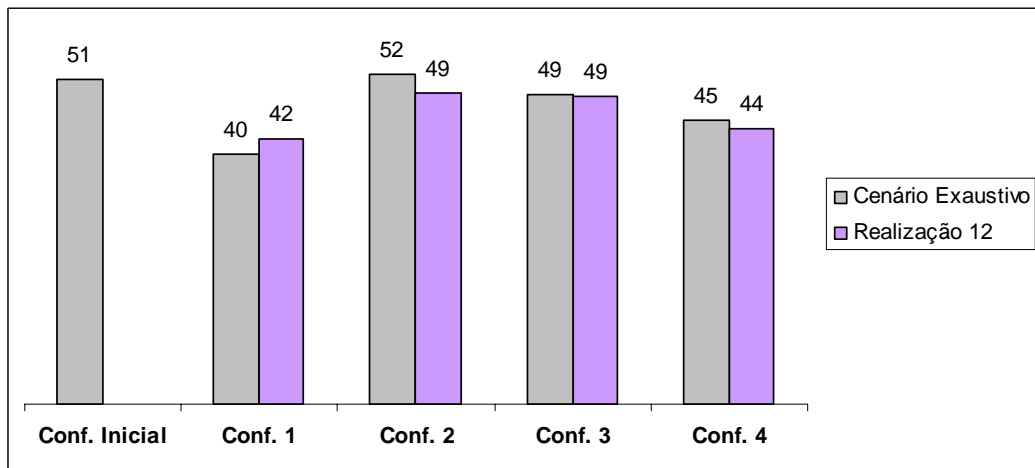


Figura 3.15 - Desvio padrão para a função tonelagem acima de 150 ppm. Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional.

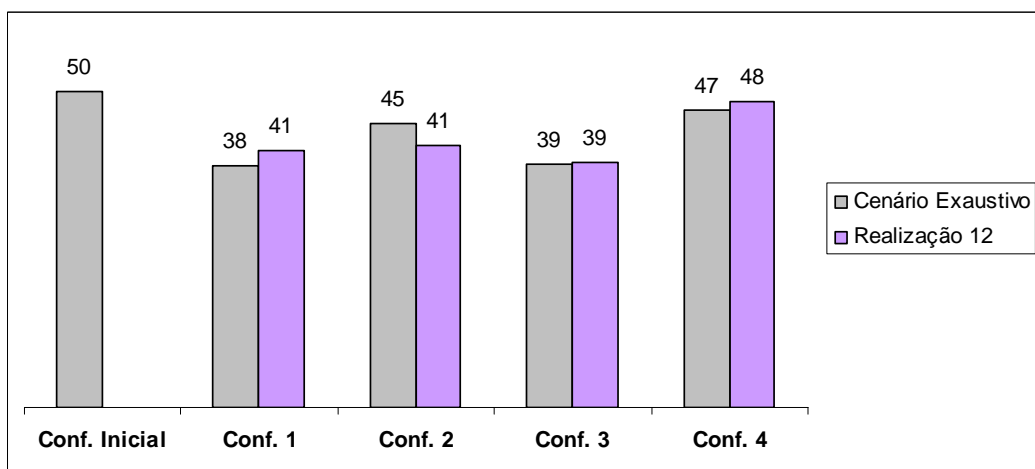


Figura 3.16 - Desvio padrão para a função tonelagem acima de 300 ppm. Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional.

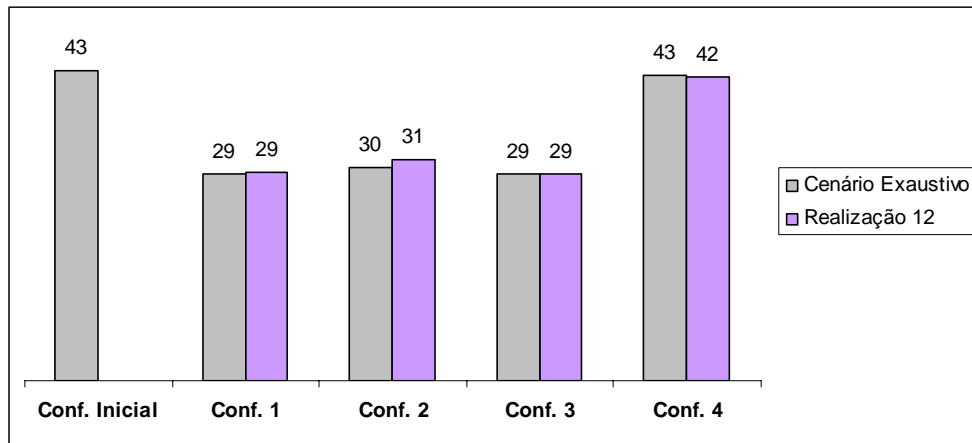


Figura 3.17 - Desvio padrão para a função tonelagem acima de 500 ppm. Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional.

Tabela 3.5 – Mínima e máxima tonelagem acima do teor de corte, para diferentes configurações de amostras e quatro teores de corte

Teor de Corte (ppm)	Configuração Inicial		Configuração 1		Configuração 2		Configuração 3		Configuração 4	
	mín.	máx.	mín.	máx.	mín.	máx.	mín.	máx.	mín.	máx.
40	2526	2729	2589	2756	2557	2741	2545	2735	2605	2737
150	1802	2070	1820	2052	1801	2079	1813	2059	1887	2160
300	1039	1277	1042	1221	1044	1287	1069	1244	1113	1341
500	423	619	397	541	451	592	409	559	421	625

Tabela 3.6 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras

Estatística para VPL	Configuração Inicial	Configuração 1	Configuração 2	Configuração 3	Configuração 4
E[F]/ fr	90%	85%	91%	89%	95%
σ	10086	7065	7161	6486	9078
IEQ	14196	8709	10085	8863	13564

Nota: O VPL para a referência é 170209.

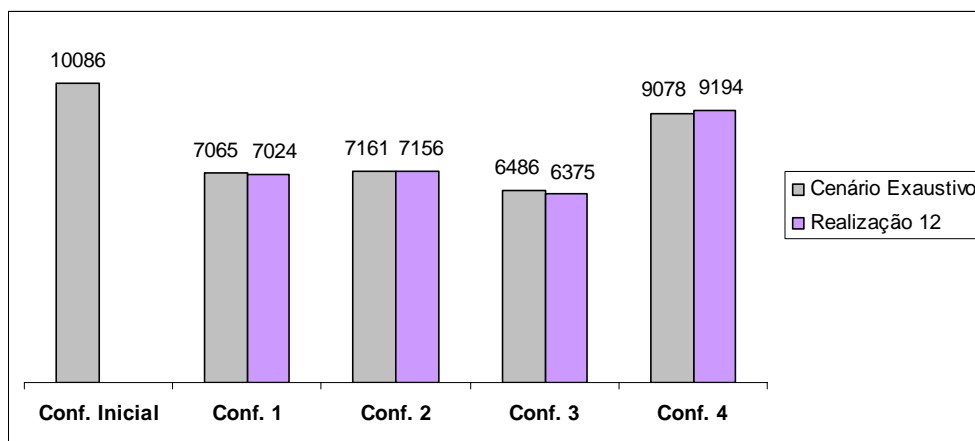


Figura 3.18 - Desvio padrão para a função VPL (Rota 1). Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional.

Tabela 3.7 – Mínimo e máximo VPL (Rota 1), para diferentes configurações de amostras

VPL	Configuração Inicial	Configuração 1	Configuração 2	Configuração 3	Configuração 4
mín.	132065	131307	136825	137026	140900
máx.	181037	162766	174649	166588	185008

Tabela 3.8 – Incerteza sobre o VPL (Rota 2), para diferentes configurações de amostras

Estatística para VPL	Configuração Inicial	Configuração 1	Configuração 2	Configuração 3	Configuração 4
E[F]/ fr	98%	82%	91%	88%	106%
σ	11435	7590	7874	6928	10715
IR	14271	9511	12259	8667	14616

Nota: O VPL para a referência é 129149.

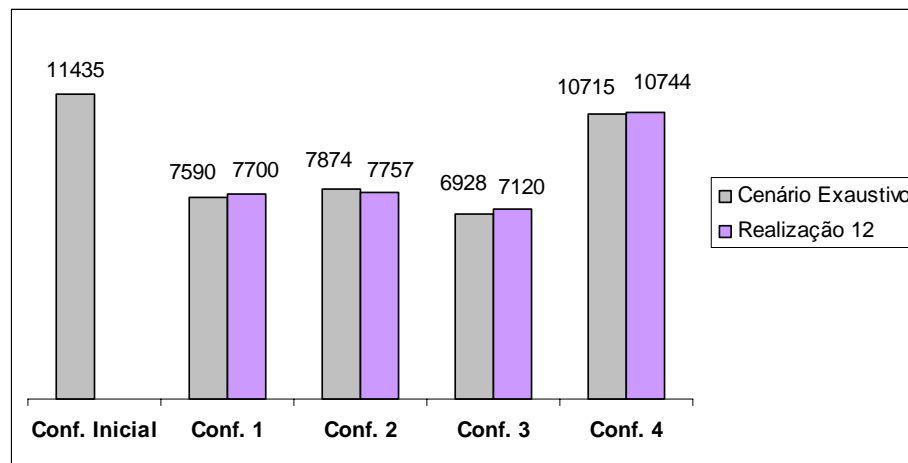


Figura 3.19 - Desvio padrão para a função VPL (Rota 2). Considerando o cenário exaustivo (cinza) e a realização 12 (violeta) como fonte de informação adicional.

Tabela 3.9 – Mínimo e máximo VPL (Rota 2), para diferentes configurações de amostras

VPL	Configuração Inicial	Configuração 1	Configuração 2	Configuração 3	Configuração 4
mín.	98558	86157	98244	93859	112767
máx.	156973	127968	135504	126524	165972

3.1.2 – Dados Adicionais Obtidos pela Realização 12

Conforme mencionado anteriormente, as estatísticas da resposta de cada função foi analisada para diferentes cenários (realizações) como fonte para obter o valor das amostras adicionadas. Os resultados para as funções de transferência, usando a realização 12 como fonte de informação, são apresentados nas Tabelas

3.10 a 3.15 e Figuras 3.14 a 3.19. As quatro configurações de amostras testadas para a realização 12 são apresentadas nessa seção. Essas configurações apresentam 160 dados, obtidos a partir da realização 12, adicionados aos 195 dados iniciais. A Figura 3.20 apresenta o mapa e o histograma para a realização 12, obtida a partir dos dados iniciais, que será fonte de informações para as configurações testadas. As Figuras 3.21 a 3.24 apresentam os mapas e histogramas para as configurações 1 a 4, respectivamente.

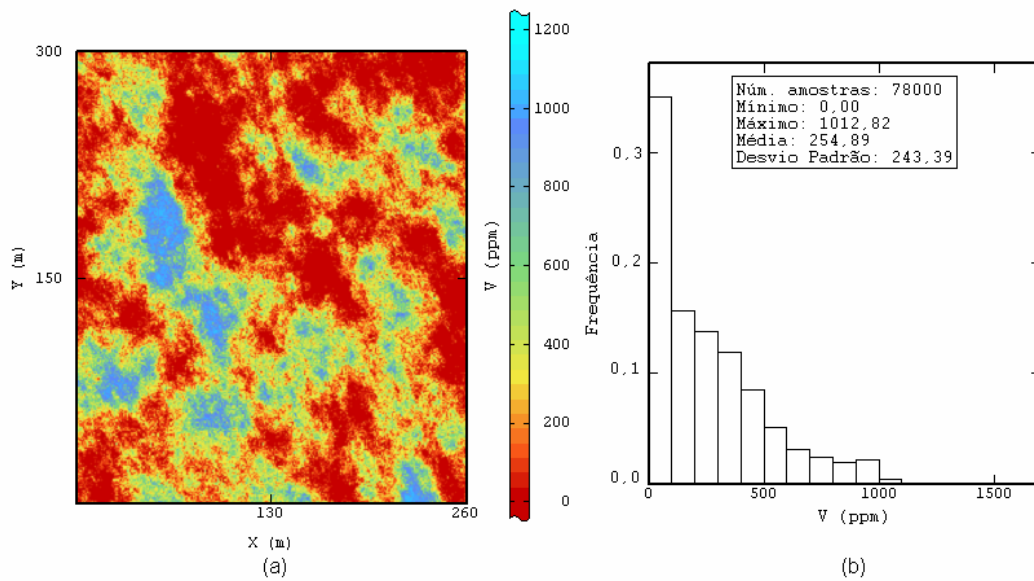


Figura 3.20 – (a) Mapa para a realização 12, obtida com os dados iniciais. (b) Histograma para a realização 12, obtida com os dados iniciais.

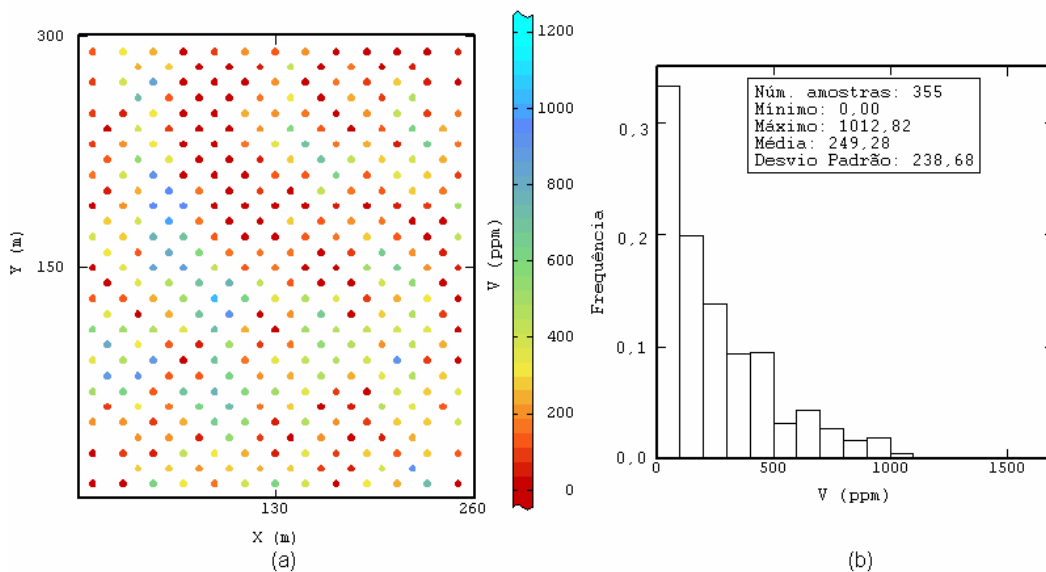


Figura 3.21 – (a) Mapa para a Configuração 1 (realização 12 como fonte de informação). (b) Histograma para a Configuração 1 (realização 12 como fonte de informação).

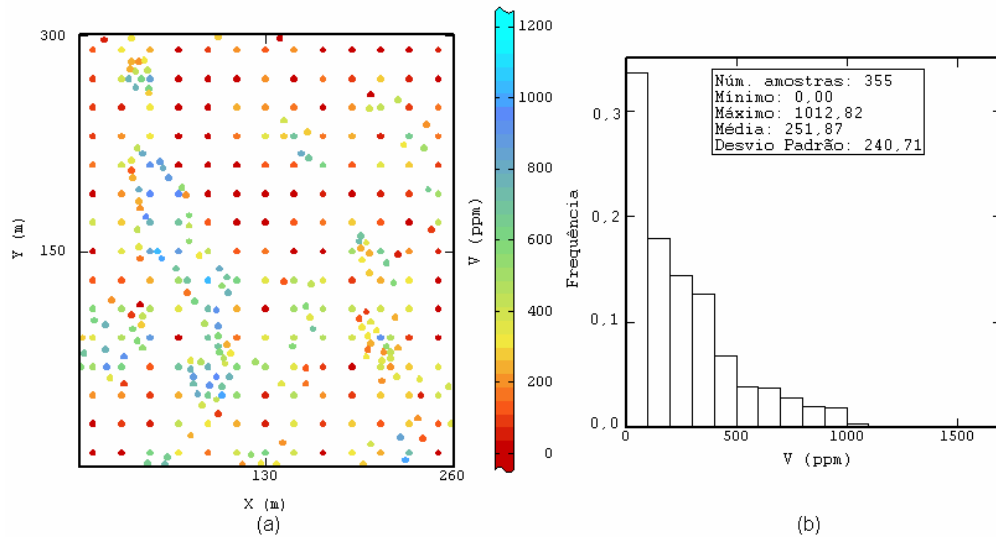


Figura 3.22 – (a) Mapa para a Configuração 2 (realização 12 como fonte de informação). (b) Histograma para a Configuração 2 (realização 12 como fonte de informação).

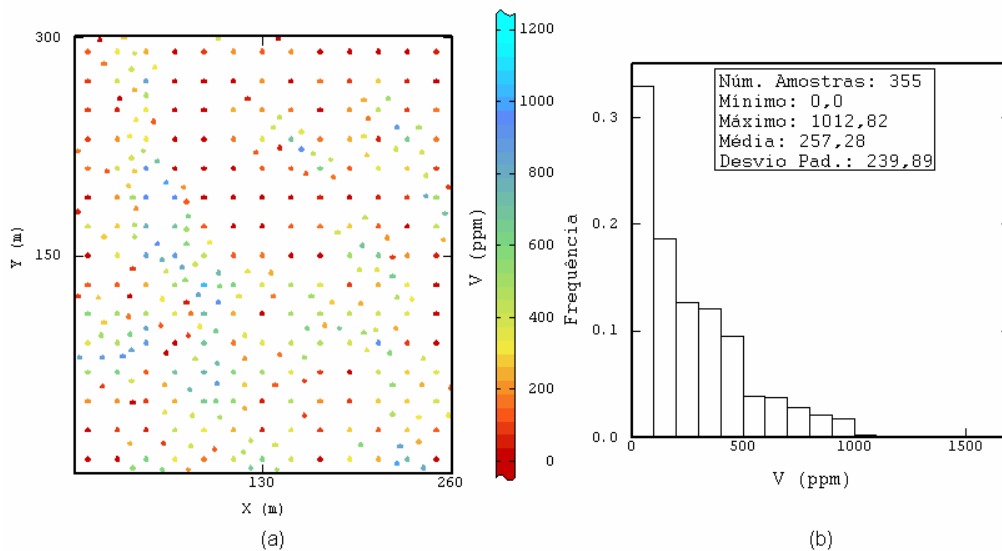


Figura 3.23 – (a) Mapa para a Configuração 3 (realização 12 como fonte de informação). (b) Histograma para a Configuração 3 (realização 12 como fonte de informação).

Assim como o que foi feito para o cenário real, 80 realizações foram geradas para cada configuração de amostra adicional testada. Para cada configuração, a média dos valores das funções de transferência, obtidos para as 80 realizações, foi escrita como porcentagem do valor real da função, ou valor da função para a realização 12.

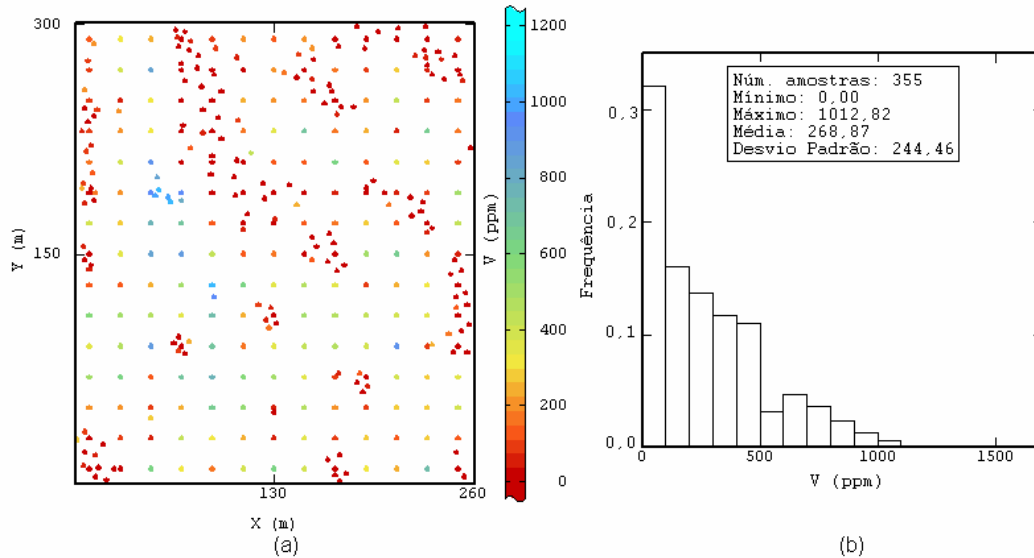


Figura 3.24 – (a) Mapa para a Configuração 4 (realização 12 como fonte de informação). (b) Histograma para a Configuração 4 (realização 12 como fonte de informação).

Para o VPL (rotas 1 e 2), o comportamento dos valores de incerteza para as diferentes configurações é semelhante ao encontrada quando utilizou-se os valores das amostras lidos na referência exaustiva como fonte de informação.

Para a função de transferência tonelagem acima de um dado teor de corte, as mesmas observações sobre as medidas de incerteza feitas para a referência como fonte de informação adicional podem ser feitas usando a realização 12 como fonte de informação. A diferença entre os índices de incerteza obtidos para as diferentes configurações não é significativa. Em geral, quando uma configuração se destaca positiva ou negativamente na redução da incerteza para um teor de corte isso também ocorre utilizando diferentes cenários fontes de informação. Assim, a análise dos resultados para tonelagem e VPL indica que um processo de otimização, para encontrar a configuração que reduz mais incerteza para uma função, pode ser baseado em qualquer cenário simulado a partir dos dados iniciais como fonte de obtenção do valor da amostra adicionada. Em suma, a redução da incerteza não é influenciada pelo valor da amostra adicionada em uma dada localização, mas sim, pelo fato de termos conhecimento do valor naquele local. É irrelevante de onde extrai-se o valor de uma futura amostra, desde que este venha de um modelo simulado condicionalmente e não necessariamente seja igual ao valor real no dado local.

Analisando os valores de média da função de transferência (tonelagem e

VPL) obtidos para as diversas configurações utilizando a realização 12 como fonte de informação adicional, observa-se que as configurações que melhor reproduziram o valor da função ($(E[F]/fr)$ mais próximo de 100%) são diferentes dos resultados utilizando a referência exaustiva como base de dados. Ou seja, a utilização de um cenário hipotético simulado para o depósito não permite encontrar a configuração que produz resultados mais acurados para a função. No entanto, conforme comentado, a configuração de amostras mais eficiente na redução de incerteza sobre a função não precisa ser a mais acurada desde que permita a inclusão do real valor da função no intervalo de possíveis valores gerados pela simulação.

As Tabelas 3.11, 3.13 e 3.15 apresentam o mínimo e máximo da distribuição dos valores de tonelagem, VPL (rota 1) e VPL (rota 2), respectivamente, para cada configuração testada, usando a realização 12 como fonte de informação. A realização 12 é um cenário simulado, apresenta uma suavização dos valores extremos em relação ao cenário exaustivo. Por isso, para a função VPL, que é bastante influenciada pelos valores extremos, é mais fácil incluir o VPL para a realização 12 no intervalo obtido por simulação (Tabelas 3.13 e 3.15) do que incluir o VPL para o cenário exaustivo no respectivo intervalo (Tabelas 3.7 e 3.9).

Tabela 3.10 – Incerteza sobre a tonelagem acima do teor de corte, para diferentes configurações de amostras e quatro teores de corte

Teor de Corte (ppm)	Configuração Inicial			Configuração 1			Configuração 2			Configuração 3			Configuração 4		
	E[F]/fr	σ	IEQ	E[F]/fr	σ	IEQ	E[F]/fr	σ	IEQ	E[F]/fr	σ	IEQ	E[F]/fr	σ	IEQ
40	103%	44	54	102%	31	35	103%	41	58	103%	40	60	101%	30	39
150	106%	51	64	103%	42	47	105%	49	67	106%	49	67	103%	44	54
300	108%	50	72	101%	41	51	102%	41	60	107%	39	53	107%	48	66
500	121%	43	66	113%	29	38	113%	31	38	114%	29	41	121%	42	63

Nota: Tonelagem acima do teor de corte para a realização 12

2573 (**40**); 1849 (**150**); 1099 (**300**); 423 (**500**).

Tabela 3.11 – Mínima e máxima tonelagem acima do teor de corte, para diferentes configurações de amostras e quatro teores de corte

Teor de Corte (ppm)	Configuração 1		Configuração 2		Configuração 3		Configuração 4	
	mín.	máx.	mín.	máx.	mín.	máx.	mín.	máx.
40	2543	2708	2558	2747	2544	2740	2536	2663
150	1801	2050	1826	2088	1859	2068	1791	2041
300	1016	1196	1025	1231	1069	1258	1051	1276
500	417	571	399	564	388	548	426	623

Tabela 3.12 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras

Estatística para VPL	Configuração Inicial	Configuração 1	Configuração 2	Configuração 3	Configuração 4
E[F]/ fr	113%	104%	109%	103%	113%
σ	10086	7024	7156	6375	9194
IEQ	14196	8387	8893	8455	13643

Nota: O VPL para a realização 12 é 134926.

Tabela 3.13 – Mínimo e máximo VPL (Rota 1), para diferentes configurações de amostras

VPL	Configuração 1	Configuração 2	Configuração 3	Configuração 4
mín.	125719	129019	125020	133381
máx.	158562	166237	153664	177129

Tabela 3.14 – Incerteza sobre o VPL (Rota 2), para diferentes configurações de amostras

Estatística para VPL	Configuração Inicial	Configuração 1	Configuração 2	Configuração 3	Configuração 4
E[F]/ fr	112%	101%	110%	114%	109%
σ	11435	7700	7757	7120	10744
IEQ	14271	9158	12180	9782	14693

Nota: O VPL para a realização 12 é 112665.

Tabela 3.15 – Mínimo e máximo VPL (Rota 2), para diferentes configurações de amostras

VPL	Configuração 1	Configuração 2	Configuração 3	Configuração 4
mín.	94178	104391	107650	99276
máx.	134705	142105	141690	153750

O mapa de localização dos dados da Configuração 2 mostra que alguns dados se encontram muito próximos um do outro. A distância mínima entre amostras é de aproximadamente 3 m para essa configuração, isso pode ter contribuído para reduzir sua eficiência. Assim, uma nova configuração (Configuração 5) foi construída adicionando 160 dados nas regiões de alto IEQ e obedecendo uma distância mínima entre amostras de 7 m (a Configuração 5 foi construída seguindo o algoritmo apresentado no Capítulo 5). A Figura 3.25 apresenta o mapa e o histograma para a Configuração 5.

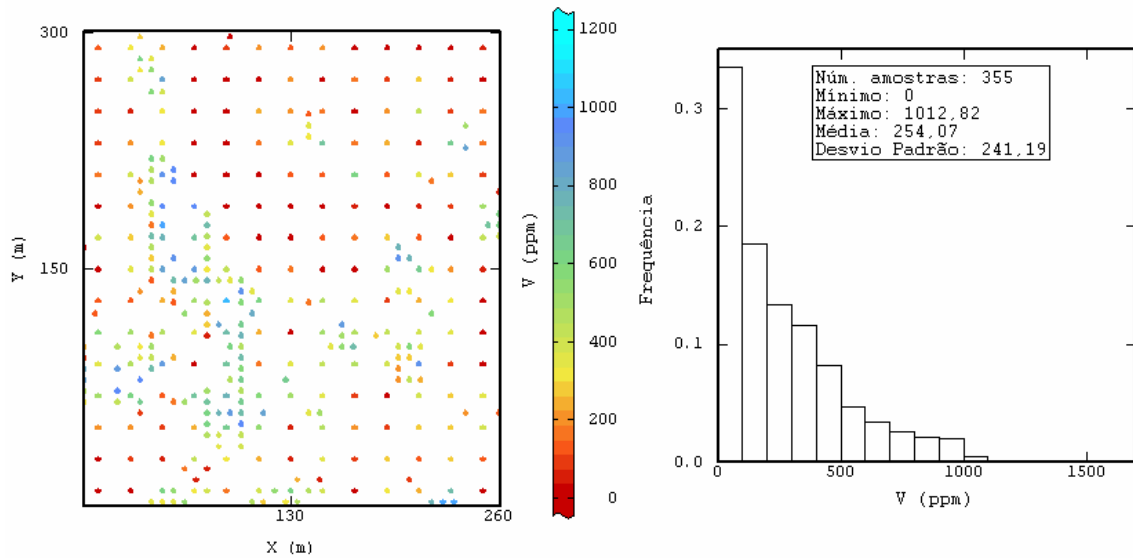


Figura 3.25 – Mapa e histograma para a Configuração 5 (realização 12 como fonte de informação).

Também, se construiu uma configuração adicionando 160 dados nas regiões de alto IEQ e obedecendo a uma distância mínima entre amostras de 9,9 m (Configuração 6). Embora a normalização dos dados da Configuração 6 e a simulação geoestatística tenham sido realizadas no software Isatis[®] (Bleines et al., 2001), os pesos de desagrupamento utilizados para os dados da Configuração 6 foram obtidos pelo programa *Declus* do *GSLIB* (Deutsch e Journel, 1998) (tamanho de célula 20 x 20 m nas direções x e y). A Figura 3.26 apresenta o mapa e o histograma para a Configuração 6.

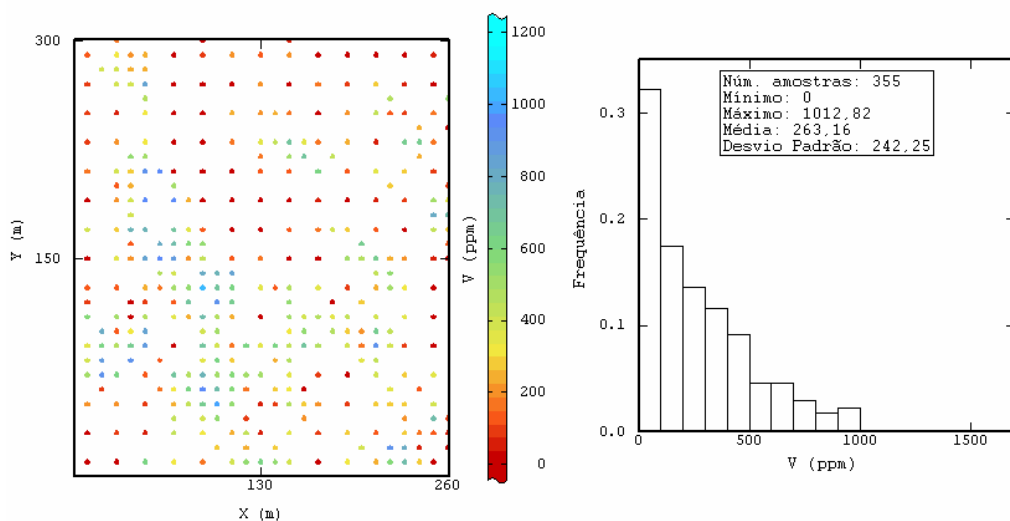


Figura 3.26 – Mapa e histograma para a Configuração 6 (realização 12 como fonte de informação).

A estatística para os valores de VPL (rotas 1 e 2), obtidos com as Configurações 5 e 6, usando a realização 12 como cenário fonte de informação é apresentada nas Tabelas 3.16 a 3.19.

Tabela 3.16 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras

Estatística para VPL	Configuração 5	Configuração 6
E[F]/ fr	109%	105%
σ	6905	7123
IEQ	9094	9510

Nota: O VPL para a realização 12 é 134926.

Tabela 3.17 – Mínimo e máximo VPL (Rota 1), para diferentes configurações de amostras

VPL	Configuração 5	Configuração 6
min.	125212	127264
max.	162662	163928

Tabela 3.18 – Incerteza sobre o VPL (Rota 2), para diferentes configurações de amostras

Estatística para VPL	Configuração 5	Configuração 6
E[F]/ fr	110%	108%
σ	8129	8192
IEQ	11336	11107

Nota: O VPL para a realização 12 é 112665.

Tabela 3.19 – Mínimo e máximo VPL (Rota 2), para diferentes configurações de amostras

VPL	Configuração 5	Configuração 6
min.	102941	98761
max.	141370	141548

Anteriormente, se comparou a eficiência das Configurações 2 e 3 na redução da incerteza sobre o VPL, a fim de avaliar o benefício da adição de amostras em 4 etapas nas regiões de elevado IEQ em relação a adição de amostras em uma única etapa. Contudo, a distância mínima entre amostras era muito diferente para as Configurações 2 e 3 e o maior agrupamento de algumas amostras na Configuração 2 poderia estar reduzindo o desempenho dessa configuração. O benefício da adição de amostras em etapas pode ser melhor avaliado comparando os desempenhos das Configurações 3 e 5, visto que a menor distância entre amostras da Configuração 3 é próxima de 7,1 m (semelhante à menor distância para a Configuração 5). Observando os resultados das Tabelas 3.16 a 3.19 pode-se dizer

que a Configuração 3 é um pouco mais eficiente que a Configuração 5 na redução da incerteza sobre o VPL, mostrando o benefício da adição de amostras em etapas nas regiões de elevado IEQ, quando um grande número de amostras deve ser adicionado.

3.1.3 – Dados Adicionais Obtidos pela Realização 54

A estatística para cada função foi analisada para outros dois cenários utilizados como fonte de informação adicional: realização 54 e realização 57 (obtidas a partir de simulações geradas com os dados iniciais). Os resultados para as funções de transferência, usando a realização 54 como fonte de informação são apresentados nas Tabelas 3.20 a 3.22 (e Figuras 3.31 a 3.36) e os resultados usando a realização 57 são apresentados nas Tabelas 3.23 a 3.25 (e Figuras 3.31 a 3.36). As mesmas observações sobre as medidas de incerteza feitas para a realização 12 podem ser feitas para as realizações 54 e 57 como fonte de informação adicional. Confirma-se, que um processo de otimização, para encontrar a configuração que reduz mais incerteza para uma dada função de transferência, pode ser baseado em qualquer cenário simulado a partir dos dados de iniciais para obter os valores das amostras a serem adicionadas.

As três configurações de amostras testadas para a realização 54 são apresentadas nessa seção. Essas configurações apresentam 160 dados, obtidos a partir da realização 54, adicionados aos 195 dados iniciais. A Figura 3.27 apresenta o mapa e o histograma para a realização 54, obtida a partir dos dados iniciais, que será fonte de informações para as configurações testadas. As Figuras 3.28 a 3.30 apresentam os mapas e histogramas para as configurações 1, 2 e 4, respectivamente.

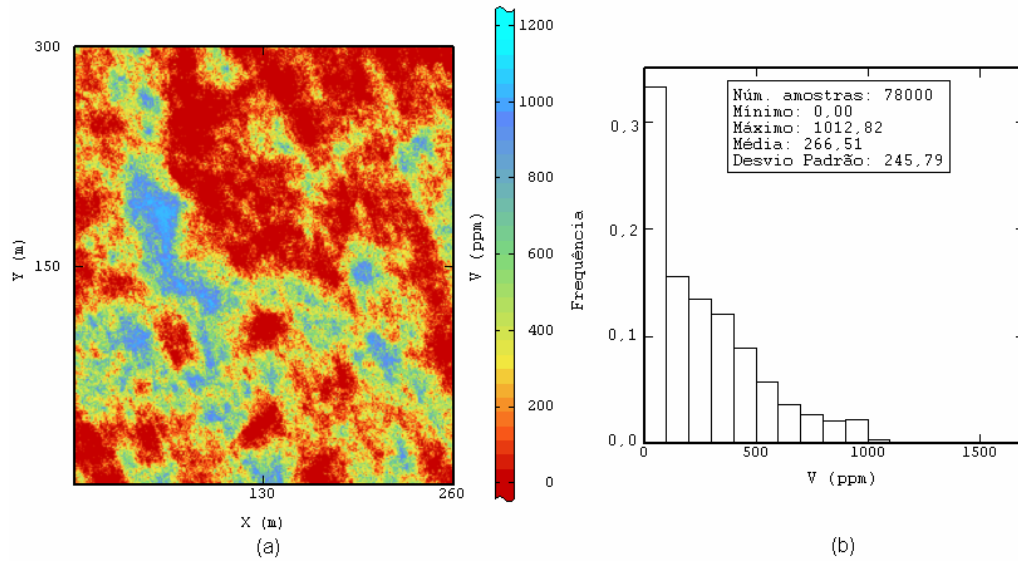


Figura 3.27 – (a) Mapa para a realizao 54, obtida com os dados iniciais. (b) Histograma para a realizao 54, obtida com os dados iniciais.

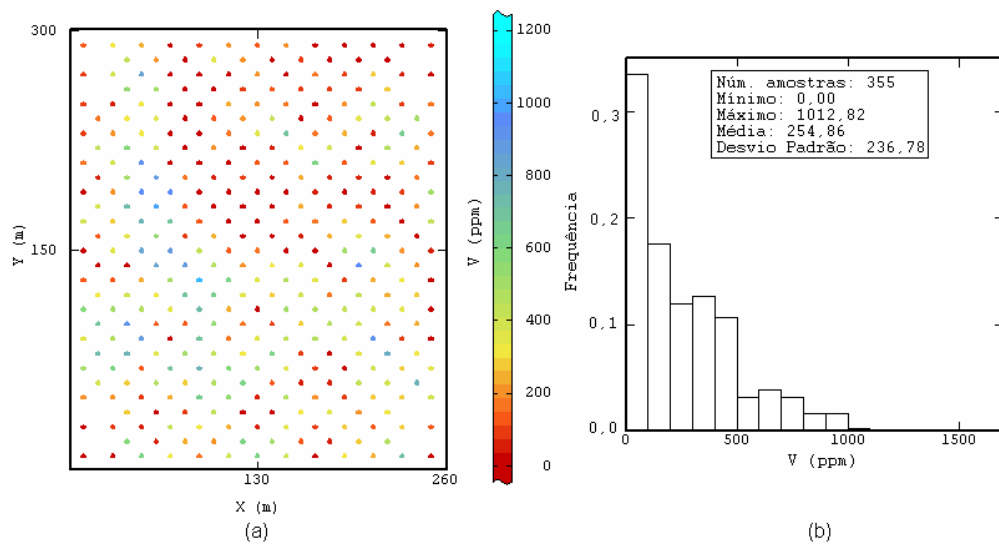


Figura 3.28 – (a) Mapa para a Configurao 1 (realizao 54 como fonte de informao). (b) Histograma para a Configurao 1 (realizao 54 como fonte de informao).

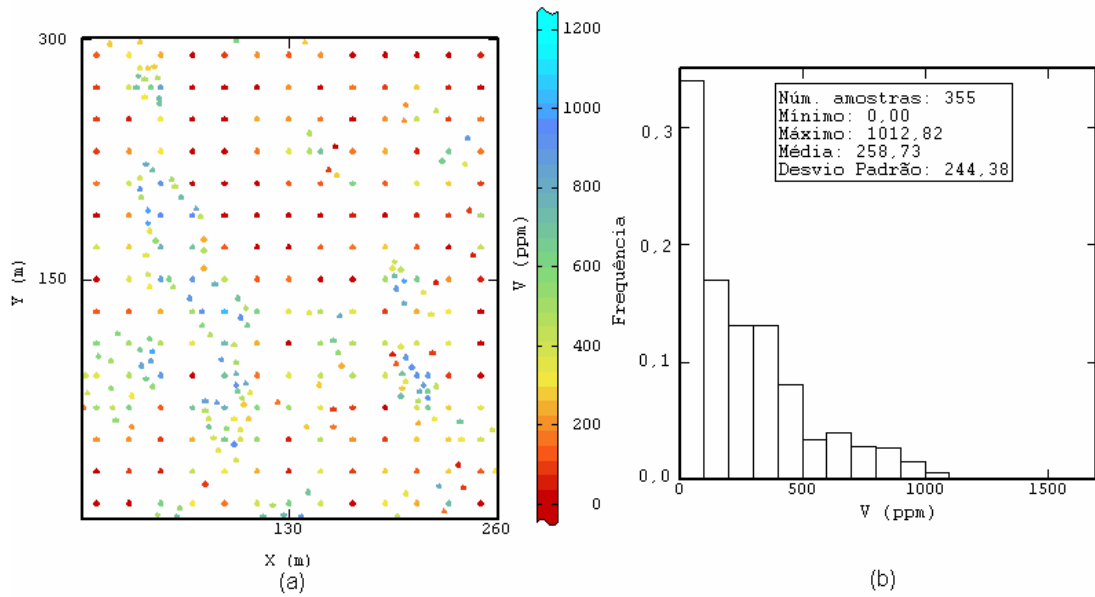


Figura 3.29 – (a) Mapa para a Configuração 2 (realização 54 como fonte de informação). (b) Histograma para a Configuração 2 (realização 54 como fonte de informação).

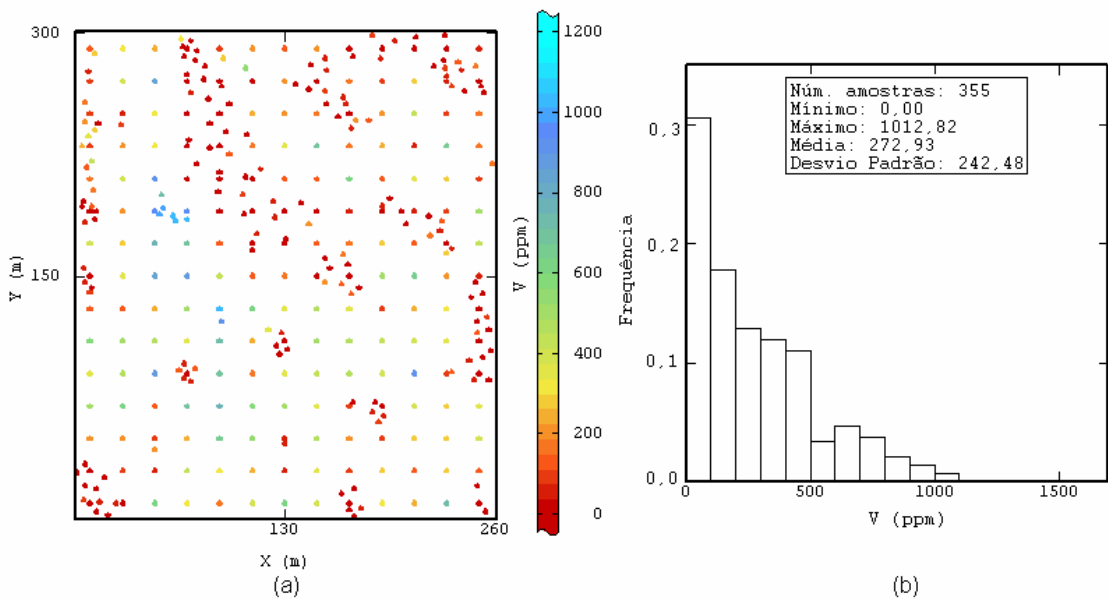


Figura 3.30 – (a) Mapa para a Configuração 4 (realização 54 como fonte de informação). (b) Histograma para a Configuração 4 (realização 54 como fonte de informação).

Para todas configurações testadas, as estatísticas sobre cada função de transferência são apresentadas nas Tabelas 3.20 a 3.22.

Tabela 3.20 – Incerteza sobre a tonelagem acima do teor de corte, para diferentes configurações de amostras e 4 teores de corte

Teor de Corte (ppm)	Configuração Inicial			Configuração 1			Configuração 2			Configuração 4		
	E[F]/ fr	σ	IEQ	E[F]/ fr	σ	IEQ	E[F]/ fr	σ	IEQ	E[F]/ fr	σ	IEQ
40	101%	44	54	100%	33	43	101%	42	56	101%	30	45
150	101%	51	64	101%	42	52	102%	51	65	101%	42	58
300	102%	50	72	101%	38	44	102%	38	57	102%	48	53
500	106%	43	66	99%	28	35	105%	30	36	106%	42	58

Nota: Tonelagem acima do teor de corte para realização 54 2640 (40); 1926 (150); 1165 (300); 484 (500).

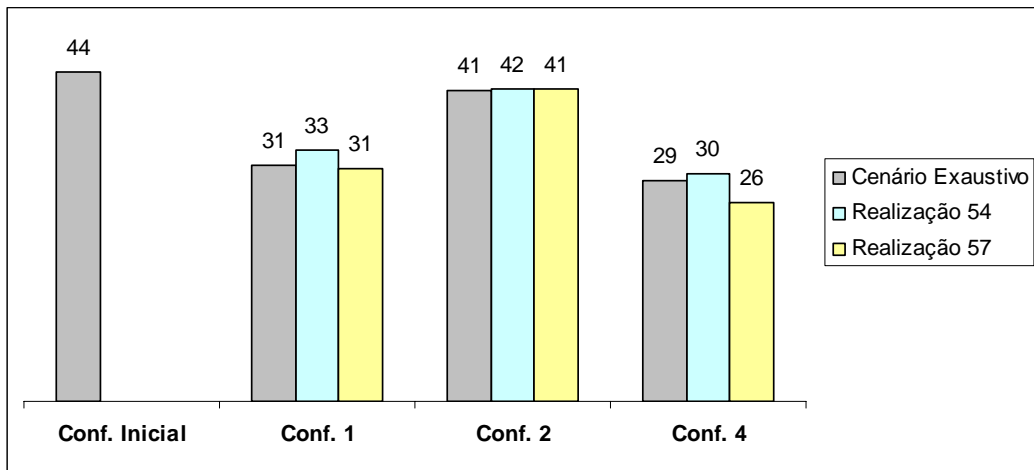


Figura 3.31 - Desvio padrão para a função tonelagem acima de 40 ppm. Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.

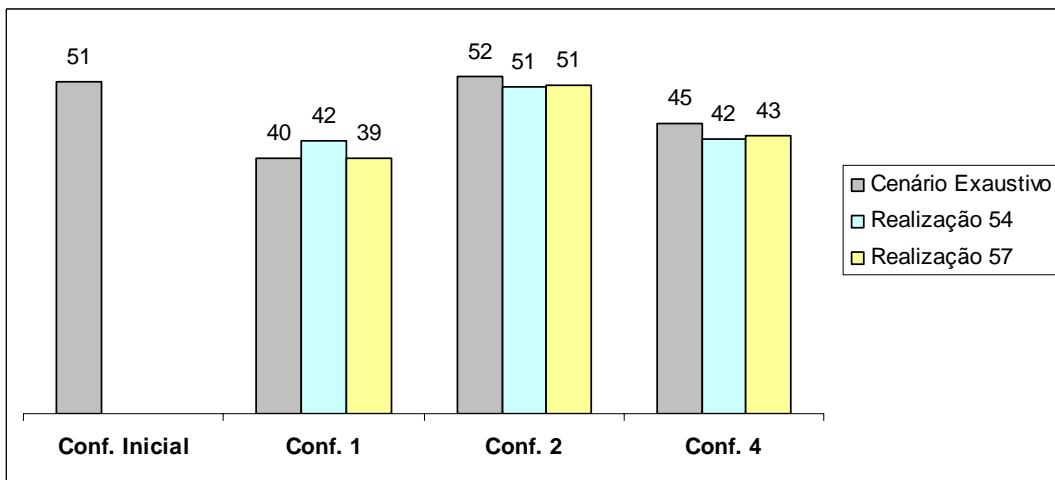


Figura 3.32 - Desvio padrão para a função tonelagem acima de 150 ppm. Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.

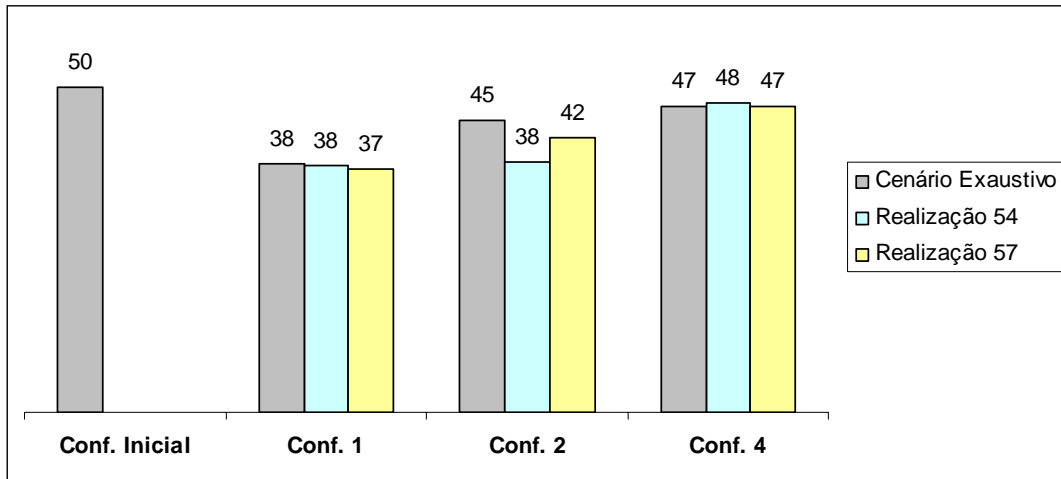


Figura 3.33 - Desvio padrão para a função tonelagem acima de 300 ppm. Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.

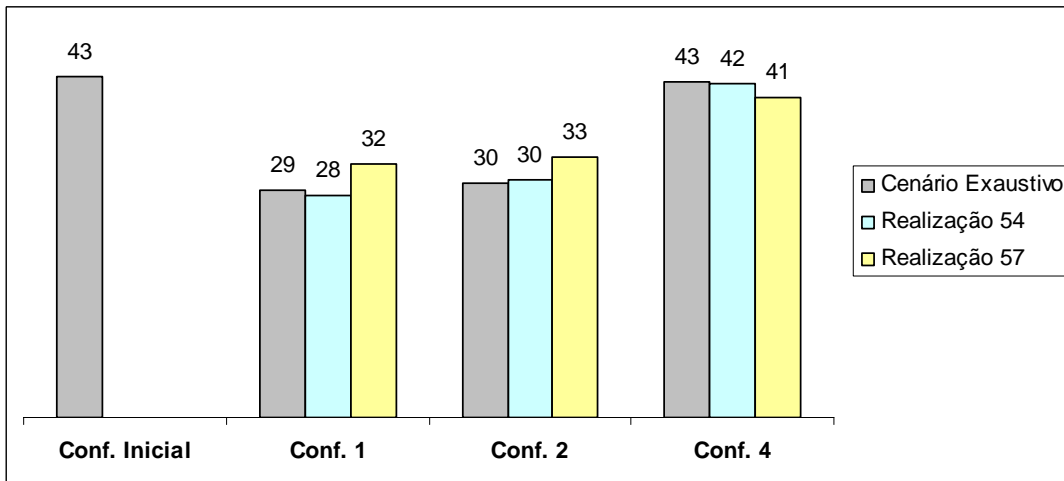


Figura 3.34 - Desvio padrão para a função tonelagem acima de 500 ppm. Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.

Tabela 3.21 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras

Estatística para VPL	Configuração Inicial	Configuração 1	Configuração 2	Configuração 4
E[F]/ fr	100%	100%	99%	101%
σ	10086	6997	7205	9082
IEQ	14196	8930	10068	13086

Nota: O VPL para a realização 54 é 152591.

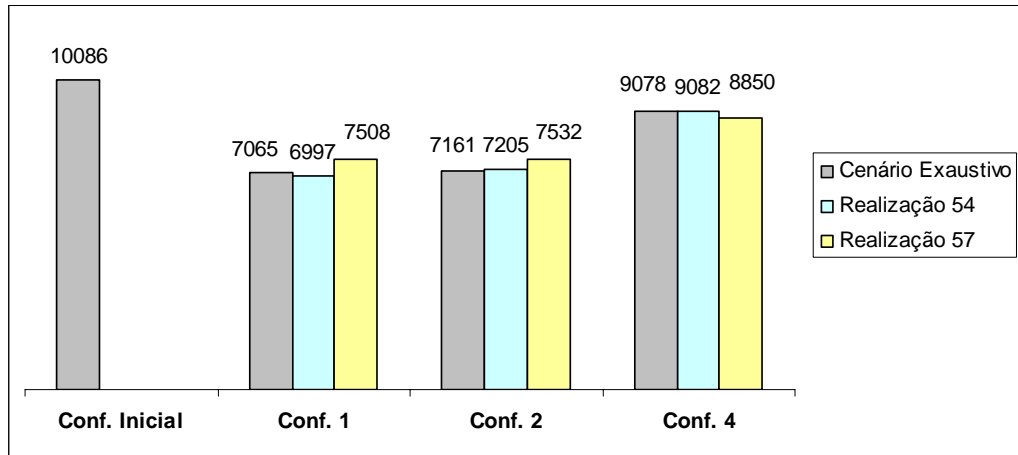


Figura 3.35 - Desvio padrão para a função VPL (Rota 1). Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.

Tabela 3.22 – Incerteza sobre o VPL (Rota 2), para diferentes configurações de amostras

Estatística para VPL	Configuração Inicial	Configuração 1	Configuração 2	Configuração 4
E[F]/ fr	101%	98%	101%	100%
σ	11435	7646	7837	10669
IEQ	14271	8803	12008	14025

Nota: O VPL para a realização 54 é 124993.

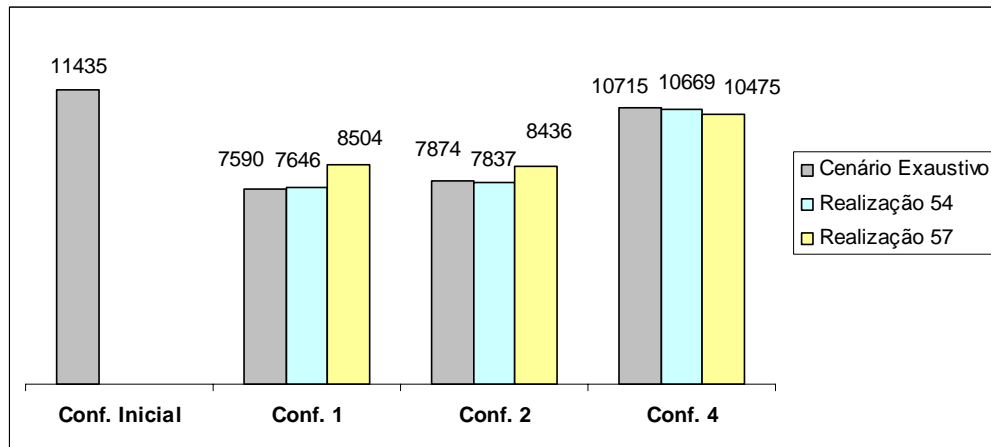


Figura 3.36 - Desvio padrão para a função VPL (Rota 2). Considerando o cenário exaustivo (cinza), a realização 54 (verde) e a realização 57 (amarelo) como fonte de informação adicional.

3.1.4 – Dados Adicionais Obtidos pela Realização 57

As três configurações de amostras testadas para a realização 57 são apresentadas nessa seção. Essas configurações apresentam 160 dados, obtidos a

partir da realização 57, adicionados aos 195 dados iniciais. A Figura 3.37 apresenta o mapa e o histograma para a realização 57, obtida a partir dos dados iniciais, que será fonte de informações para as configurações testadas. As Figuras 3.38 a 3.40 apresentam os mapas e histogramas para as configurações 1, 2 e 4, respectivamente.

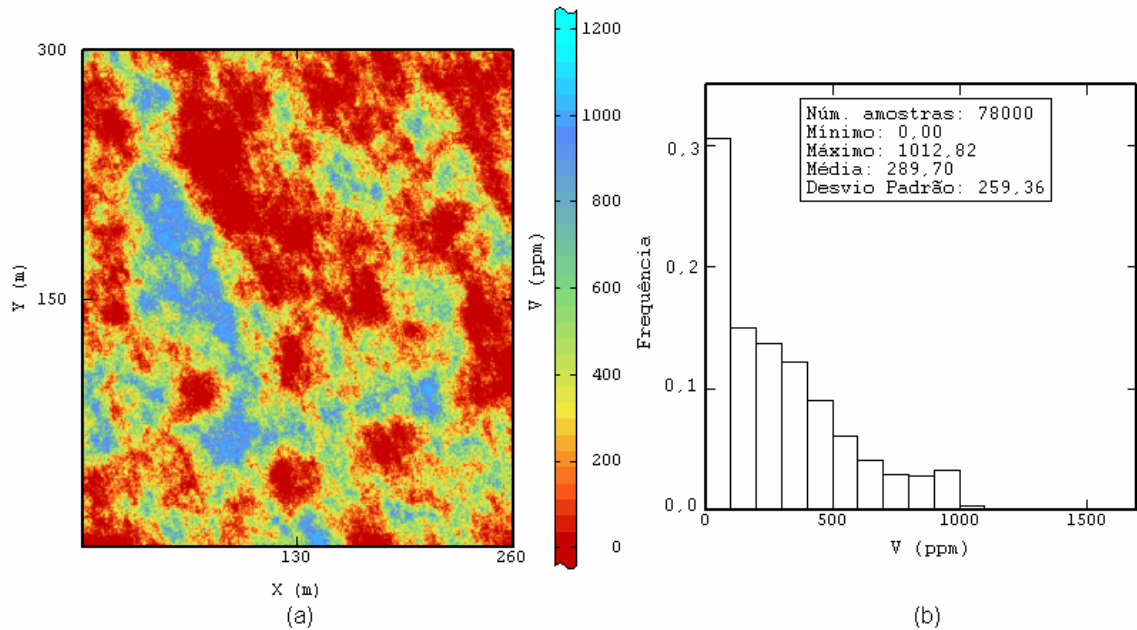


Figura 3.37 – (a) Mapa para a realização 57, obtida com os dados iniciais. (b) Histograma para a realização 57, obtida com os dados iniciais.

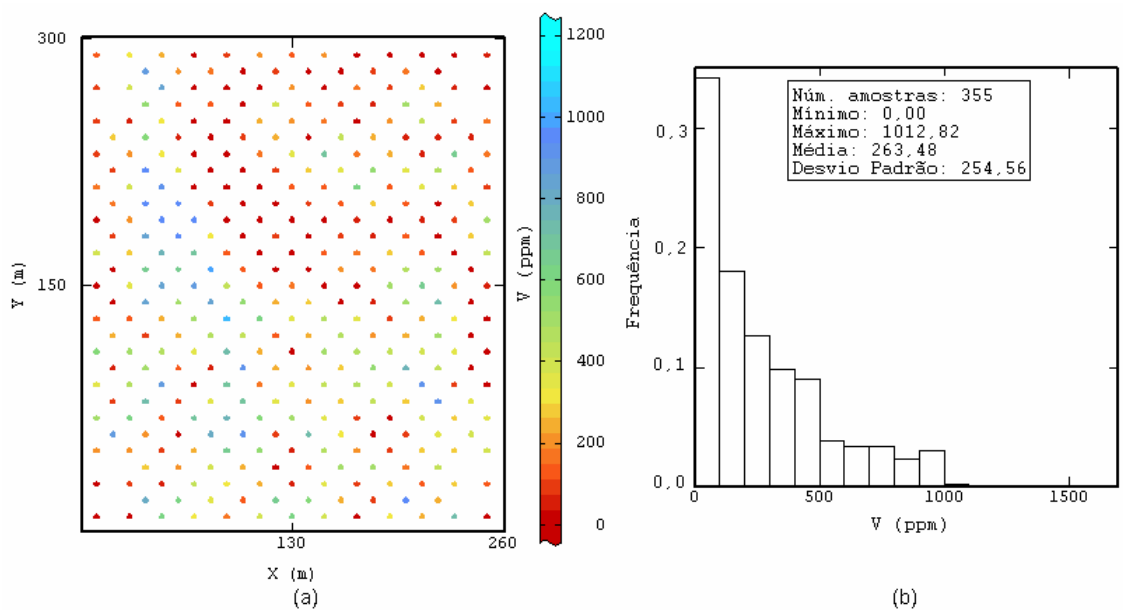


Figura 3.38 – (a) Mapa para a Configuração 1 (realização 57 como fonte de informação). (b) Histograma para a Configuração 1 (realização 57 como fonte de informação).

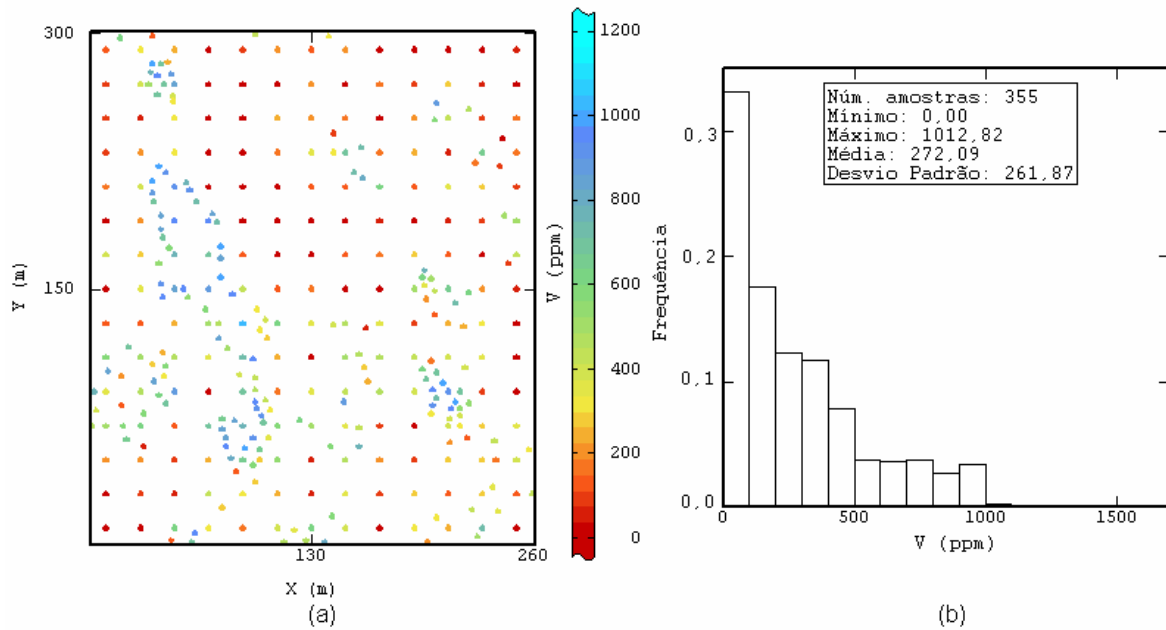


Figura 3.39 – (a) Mapa para a Configuração 2 (realização 57 como fonte de informação). (b) Histograma para a Configuração 2 (realização 57 como fonte de informação).

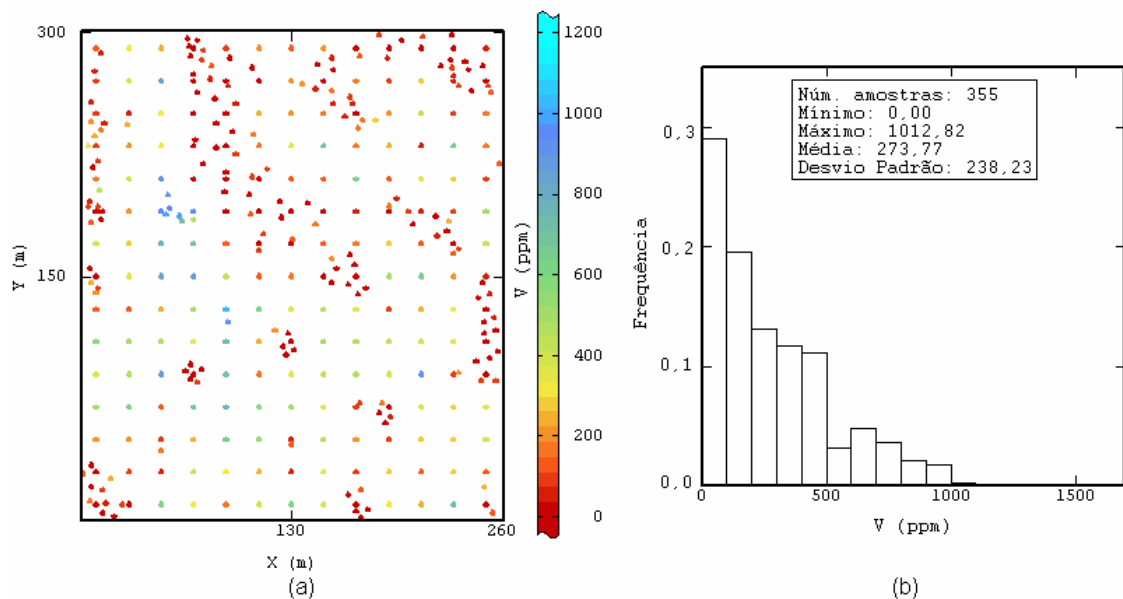


Figura 3.40 – (a) Mapa para a Configuração 4 (realização 57 como fonte de informação). (b) Histograma para a Configuração 4 (realização 57 como fonte de informação).

Para todas configurações testadas, as estatísticas sobre cada função de transferência são apresentadas nas Tabelas 3.23 a 3.25.

Tabela 3.23 – Incerteza sobre a tonelagem acima do teor de corte, para diferentes configurações de amostras e 4 teores de corte

Teor de Corte (ppm)	Configuração Inicial			Configuração 1			Configuração 2			Configuração 4		
	E[F]/ fr	σ	IEQ	E[F]/ fr	σ	IEQ	E[F]/ fr	σ	IEQ	E[F]/ fr	σ	IEQ
40	98%	44	54	98%	31	38	99%	41	62	101%	26	37
150	96%	51	64	94%	39	42	96%	51	64	97%	43	54
300	93%	50	72	93%	37	50	95%	42	62	92%	47	59
500	87%	43	66	95%	32	42	99%	33	39	85%	41	59

Nota: Tonelagem acima do teor de corte para realização 57 2703 (**40**); 2034 (**150**); 1277 (**300**); 587 (**500**).

Tabela 3.24 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras

Estatística para VPL	Configuração Inicial	Configuração 1	Configuração 2	Configuração 4
E[F]/ fr	89%	92%	98%	89%
σ	10086	7508	7532	8850
IEQ	14196	9958	10345	13105

Nota: O VPL para a realização 57 é 172335.

Tabela 3.25 – Incerteza sobre o VPL (Rota 2), para diferentes configurações de amostras

Estatística para VPL	Configuração Inicial	Configuração 1	Configuração 2	Configuração 4
E[F]/ fr	85%	91%	90%	86%
σ	11435	8504	8436	10475
IEQ	14271	10217	12889	13549

Nota: O VPL para a realização 57 é 148444.

3.1.5 - Resultados Para um Menor Número de Amostras Adicionais

A fim de verificar se a diferença nos resultados para as configurações 1 e 3 pode aumentar com um menor número de dados adicionais, a diferença nos valores de incerteza sobre o VPL (rotas 1 e 2) foi verificada para 30 amostras adicionais e para 80 amostras adicionais. Para a configuração 3, as 30 amostras foram adicionadas em uma etapa e as 80 amostras foram adicionadas em duas etapas (após a adição das 30 amostras a incerteza sobre o valor de V foi re-calculada em cada local e 50 amostras foram adicionadas nas regiões de elevada incerteza).

A Tabela 3.26 apresenta os valores de incerteza sobre o VPL para os diferentes números de amostras adicionais.

Tabela 3.26 – Incerteza sobre o VPL para diferentes números de amostras adicionais

Função	Estatística	Configuração Inicial	30 amostras adicionais		80 amostras adicionais		160 amostras adicionais	
			Conf. 1	Conf. 3	Conf. 1	Conf. 3	Conf. 1	Conf. 3
VPL Rota 1	E[F]	152902	150198	146071	144728	144891	144797	151859
	σ	10086	9310	8870	7775	7518	7065	6486
	IEQ	14196	11833	12239	11323	10593	8709	8863
	Mínimo	132065	129265	127050	128559	128452	131307	137026
	Máximo	181037	176282	171990	160874	161193	162766	166588
VPL Rota 2	E[F]	126208	120006	115577	111524	111322	105376	114293
	σ	11435	10302	10329	8444	7873	7590	6928
	IEQ	14271	13069	14178	12869	10629	9511	8667
	Mínimo	98558	98443	91805	92477	92939	86157	93859
	Máximo	156973	144990	145879	130756	127471	127968	126524

Observando os resultados da Tabela 3.26 pode-se dizer que a diferença entre as configurações continua pouco significativa para 30 e 80 amostras adicionais.

3.1.6 - Resultados Para um Menor Número de Dados Iniciais

Analisando-se os resultados anteriores observa-se que mesmo com a adição de 160 amostras aos 195 dados iniciais obtém-se uma pequena redução no valor do desvio padrão para o VPL (rotas 1 e 2). Para um menor número de dados iniciais, a redução da incerteza sobre a função poderia ser maior. A Figura 3.41 apresenta o gráfico de desvio padrão sobre a função (VPL, rota 1) versus o número de amostras usados na simulação (Boucher et al. (2004) também apresentam uma análise de incerteza versus malha de amostragem). O gráfico da Figura 3.41 considera malhas aproximadamente regulares para cada número de amostras. As configurações com 195 e 355 dados foram apresentadas nas Figuras 3.2 e 3.10, as demais configurações são apresentadas nas Figuras 3.42 a 3.45. Quando existe um pequeno número de dados, um maior número de realizações é necessário para o mapeamento do espaço de incerteza sobre a função de transferência. Os valores de desvio padrão para 30 e 60 dados consideram 120 realizações, os demais valores

de desvio padrão consideram 80 realizações. A Figura 3.46 apresenta o gráfico desvio padrão versus número de realizações para a configuração com 30 dados.

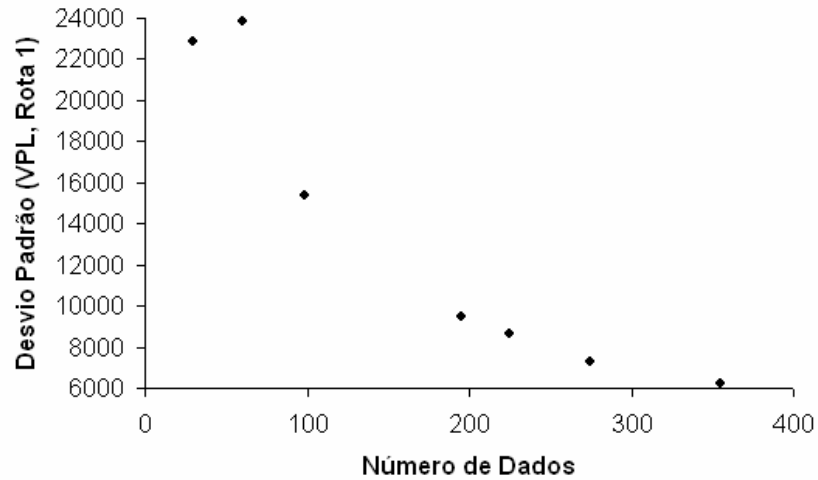


Figura 3.41 – Desvio padrão (VPL, rota 1) versus número de amostras.

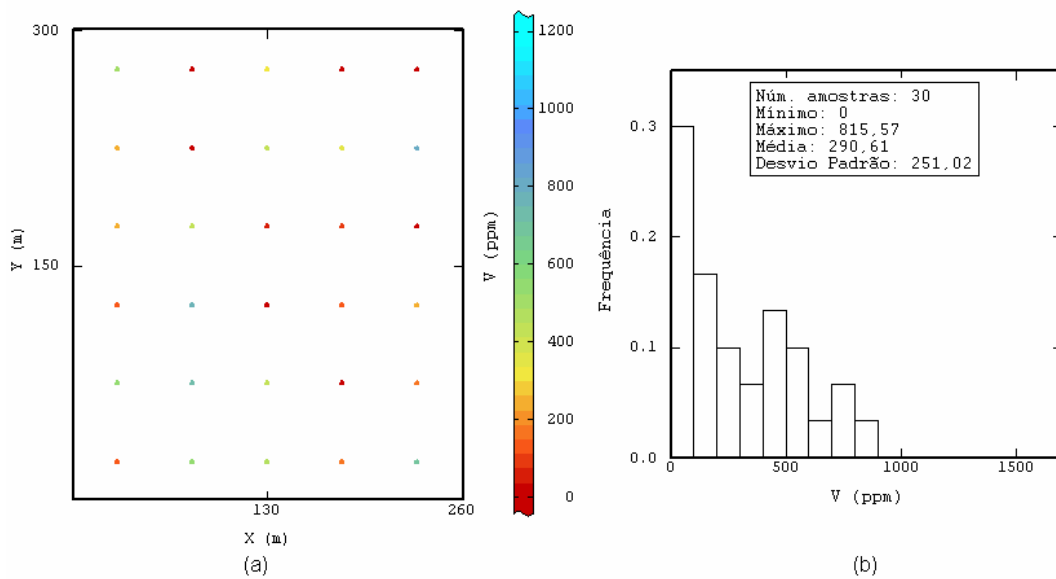


Figura 3.42 – Mapa e histograma para a configuração com 30 dados.

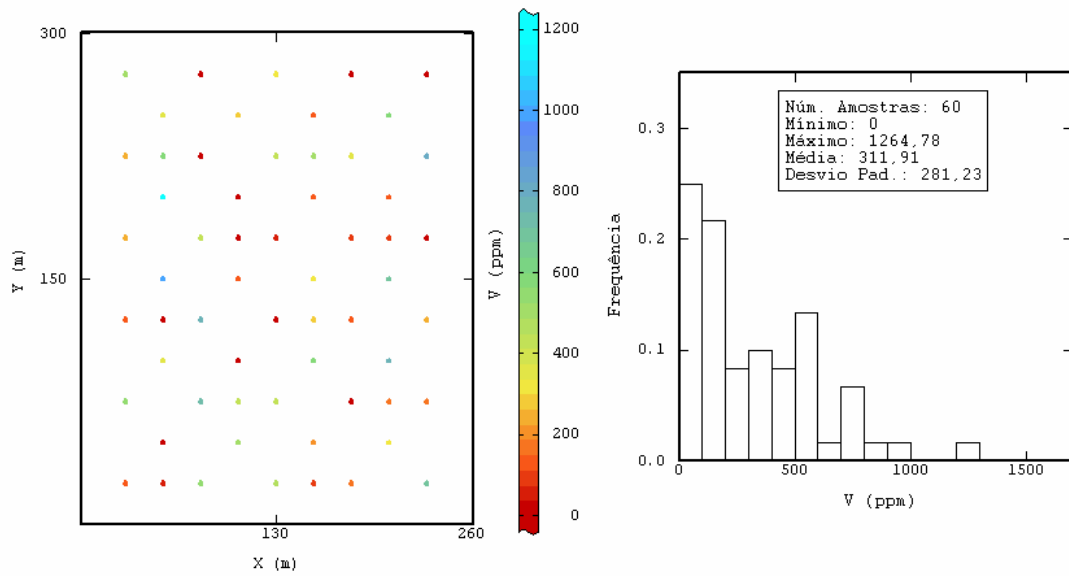


Figura 3.43 – Mapa e histograma para a configuração com 60 dados.

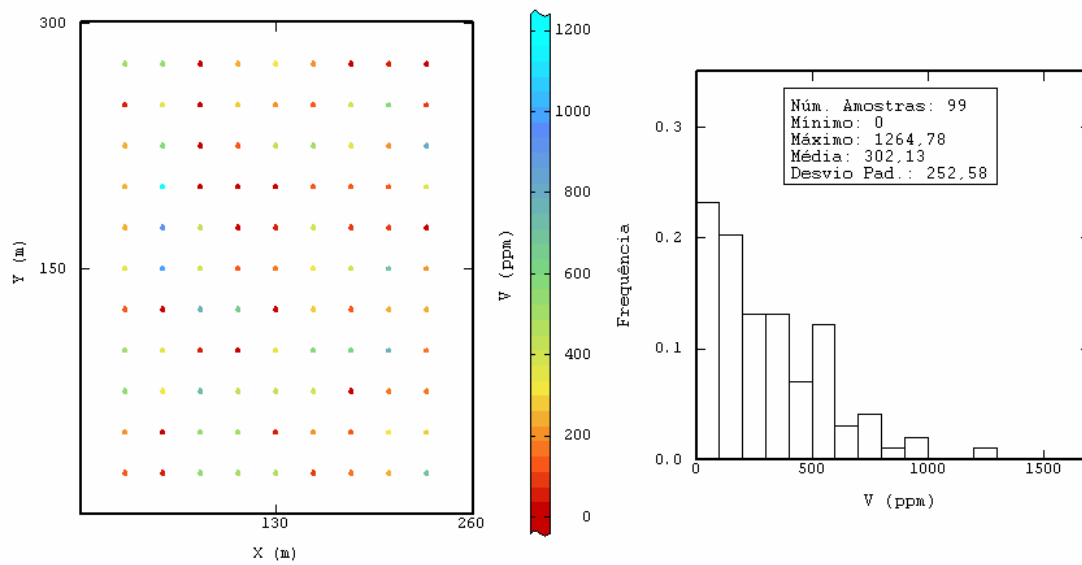


Figura 3.44 – Mapa e histograma para a configuração com 99 dados.

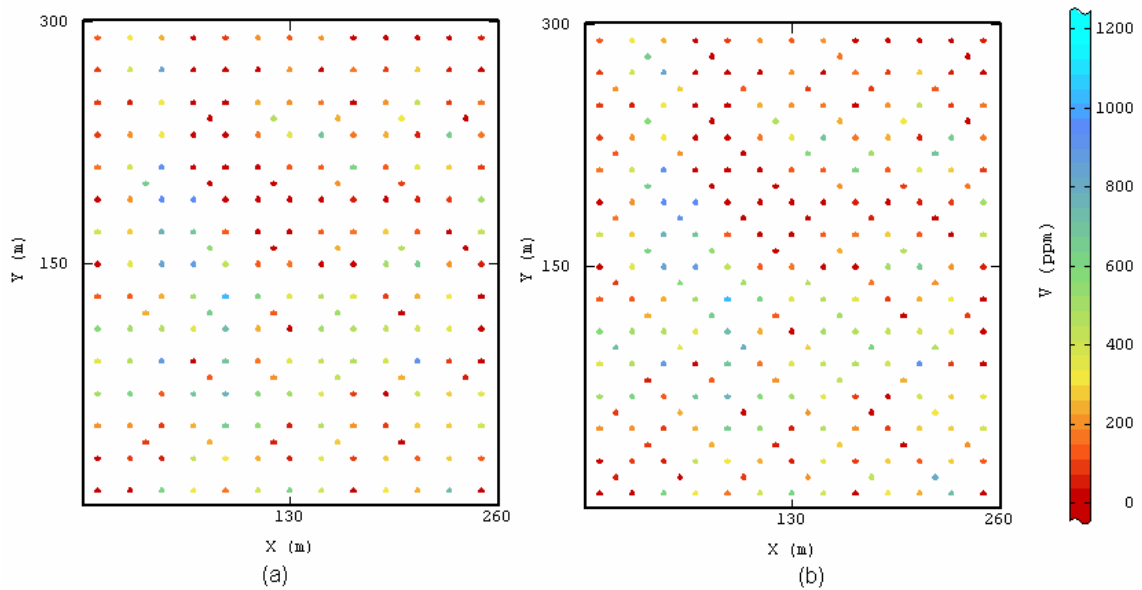


Figura 3.45 – (a) Configuração com 225 dados. (b) Configuração com 275 dados.

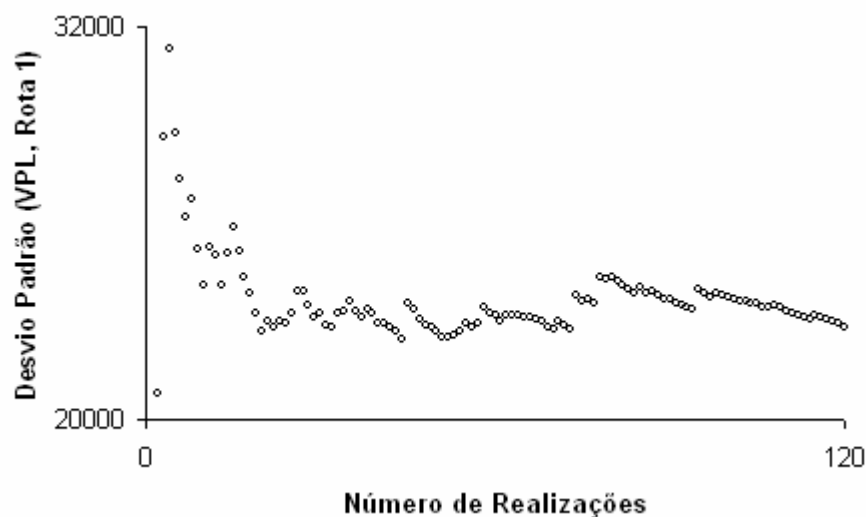


Figura 3.46 – Desvio padrão versus número de realizações para a configuração com 30 dados.

De acordo com a Figura 3.41, o benefício da nova campanha de amostragem depende do número de amostras já existentes. Por exemplo, a adição de 96 amostras a 99 dados iniciais pode reduzir o desvio padrão em aproximadamente US\$ 6000, enquanto a adição de 130 amostras a 225 dados iniciais pode reduzir o desvio em apenas US\$ 2468. A metodologia adotada nessa tese deve ser usada, não só para avaliar a melhor configuração de amostras em cada estudo, mas para avaliar se novas amostras podem trazer benefício na redução do desvio padrão. Porém, quando o número de dados iniciais é muito pequeno, o valor da incerteza sobre a função pode ser muito sensível aos valores

das novas amostras, ou seja, o desvio padrão sobre a função pode depender da realização escolhida como fonte de informação. Nesse caso, a real incerteza sobre a função não pode ser aproximada com o uso de um cenário hipotético como fonte de informação adicional. Um exemplo disso será comentado a seguir.

Para todos os pontos mostrados na Figura 3.41, a etapa de simulação considerou os mesmos parâmetros de continuidade e vizinhança de busca de amostras que os adotados para o caso dos 195 dados iniciais. No caso em que 30 dados foram usados na simulação, alguns locais do depósito não foram simulados por falta de dados dentro da vizinhança de busca e 2830 blocos foram considerados no cálculo do desvio padrão. Por isso, os valores de desvio padrão, apresentados na Figura 3.41, para os demais números de amostras, consideraram os mesmos 2830 blocos. Assim, o cálculo do VPL considerado no gráfico da Figura 3.41 é um pouco diferente do cálculo adotado anteriormente, utilizou-se um número diferente de blocos minerados por ano e custo operacional por bloco.

A comparação entre os desempenhos das configurações considerou que amostras foram adicionadas a 195 dados iniciais. Nessa comparação, a diferença entre os desempenhos das configurações na redução da incerteza não foi considerada significativa. Contudo, se as configurações são comparadas para um menor número de dados iniciais, a diferença entre as configurações pode ser maior, visto que, nesse caso, espera-se uma maior redução da incerteza, mesmo com a coleta de um menor número de dados adicionais. Assim, as configurações foram comparadas para o caso onde existem 60 dados iniciais. A Figura 3.43 apresenta o mapa e o histograma dos 60 dados iniciais considerados nessa comparação. Uma das configurações considera a adição de 39 amostras a esses 60 dados iniciais formando uma malha regular (Configuração 1). A outra configuração (Configuração 2) considera a adição de 39 dados nos locais de elevada incerteza (acessada pelo IEQ), obtidos a partir da simulação com os 60 dados iniciais.

A Configuração 1 não possui dados adicionais além dos limites dos dados iniciais, assim, para uma melhor comparação entre as configurações, os locais de elevada incerteza, além dos limites dos dados iniciais, não foram considerados na construção da Configuração 2. A Figura 3.44 mostra o mapa e o histograma para os dados da Configuração 1, enquanto a Figura 3.47 mostra o mapa e o histograma para os dados da Configuração 2.

As configurações 1 e 2 foram comparadas para três situações: uma considerando o cenário exaustivo como fonte de informação adicional, as outras considerando um cenário qualquer (realização 111) e o cenário com maior valor de VPL (realização 74), obtido com a simulação dos 60 dados iniciais. As configurações foram comparadas para as três situações para se ter certeza que os resultados obtidos dependem da posição das amostras adicionais e são pouco influenciados pelos valores das novas amostras. A Figura 3.48 apresenta os histogramas para as configurações 1 e 2 considerando a realização 111 como fonte de informação adicional.

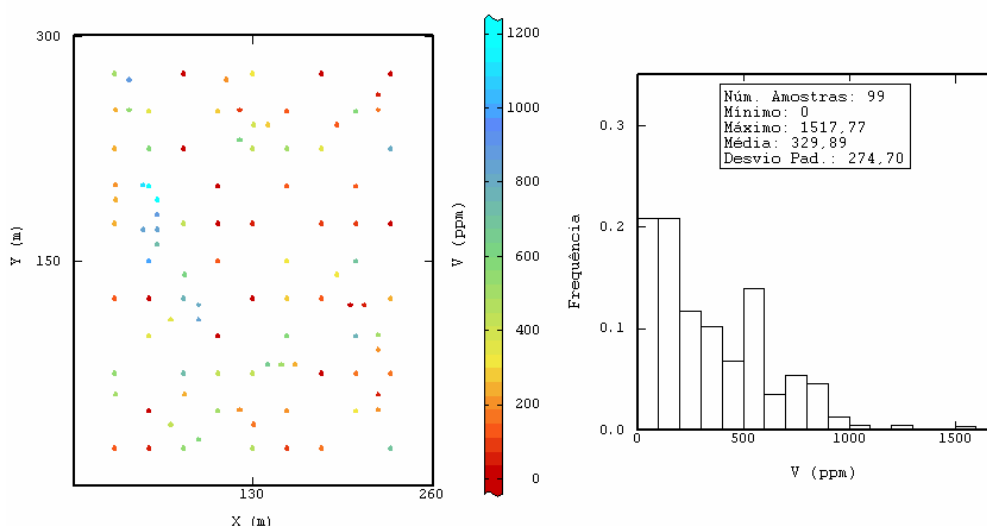


Figura 3.47 – Mapa e histograma para os dados da Configuração 2, considerando 39 dados adicionados a 60 dados iniciais.

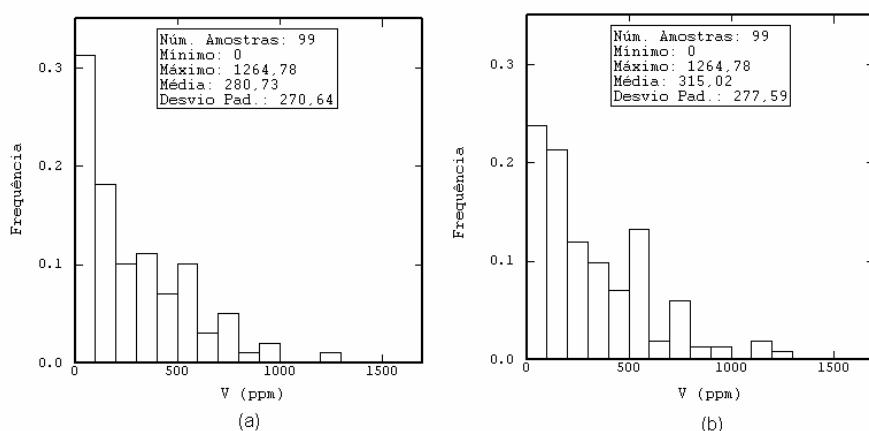


Figura 3.48 - Realização 111 como fonte de informação adicional. (a) Histograma com os dados da Configuração 1. (b) Histograma com os dados da Configuração 2.

A Tabela 3.27 (e Figura 3.49) apresenta a estatística para os valores de VPL (Rota 1, considerando 2830 blocos minerados) obtida com a simulação dos 60 dados iniciais e das Configurações 1 e 2.

Tabela 3.27 – Incerteza sobre o VPL (Rota 1), para diferentes configurações de amostras

Estatística para VPL (Rota 1)	Configuração Inicial (60 dados)	Cenário Exhaustivo		Realização 111 (obtida a partir dos 60 dados iniciais)		Realização 74 (obtida a partir dos 60 dados iniciais)	
		Conf. 1	Conf. 2	Conf. 1	Conf. 2	Conf. 1	Conf. 2
E[F]	157947	153812	169917	140801	144908	210611	219501
σ	23853	15375	19315	16563	19327	18937	22956
IEQ	29911	21785	24090	22185	22200	26479	27880
Mínimo	99479	104085	120777	90146	97332	156246	165172
Máximo	223149	200071	227190	191504	205336	266752	287877

Nota: O VPL para o cenário exaustivo é 157410.

O VPL para a realização 111 é 169881.

O VPL para a realização 74 é 223149.

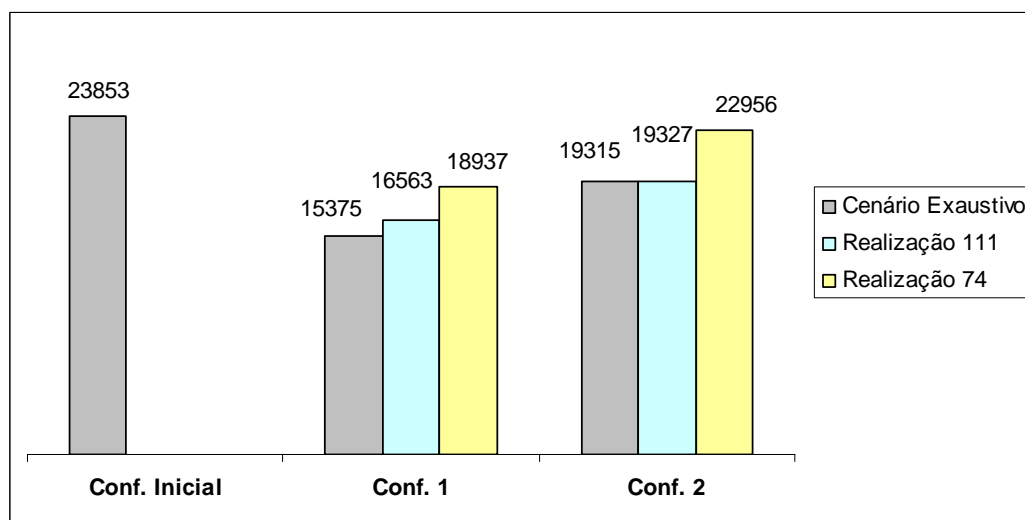


Figura 3.49 - Desvio padrão para a função VPL (Rota 1, considerando 2830 blocos minerados). Considerando o cenário exaustivo (cinza), a realização 111 (verde) e a realização 74 (amarelo) como fonte de informação adicional.

O histograma para a Configuração 2 (Figura 3.47) considera pesos de desagrupamento gerados pelo software Isatis[®] (o peso de cada amostra é inversamente proporcional ao número de amostras que pertencem ao quadrado de 20 x 20 m, nas direções x e y, centrado na amostra). Uma melhor opção para cálculo dos pesos de desagrupamento poderia ser, nesse estudo, a utilização de um quadrado com dimensões 35 m x 35 m, nas direções x e y. Contudo, tanto para o

cenário exaustivo, como para a realização 74 como fonte de informação adicional, a estatística para o VPL (Rota 1) não modifica significativamente em virtude do desagrupamento com um quadrado de 20 x 20 m ou 35 x 35 m.

Analisando a Tabela 3.27, observa-se que existe uma tendência da diferença entre os valores de desvio padrão para as Configurações 1 e 2 ser maior em comparação ao caso onde existem 195 dados iniciais. Isto é, a diferença entre os desempenhos das configurações pode depender do número de dados iniciais. Contudo, essa diferença ainda é sensível aos valores adicionados, varia com o cenário fonte de informação, ou seja, não se pode dizer ao certo quanto maior é a diferença entre as configurações no caso onde existem 60 dados iniciais. Além disso, o valor de desvio padrão sobre VPL, encontrado para uma configuração, usando uma realização como fonte de informação, não pode ser transportado para a realidade, pois esse valor ainda é dependente da realização.

Ainda observando os resultados da Tabela 3.27, pode-se dizer que a Configuração 2 tende a ser menos eficiente na redução da incerteza do que a Configuração 1, ao contrário do que se observou para o caso onde 160 dados foram adicionados aos 195 dados iniciais (comparando a Configuração 1 com a Configuração 3). Para o caso onde existem, apenas, 60 dados iniciais, o IEQ nos locais de elevada incerteza não é muito mais elevado que o IEQ nas demais regiões do depósito. Por isso, a adição de amostras, apenas, nas regiões de elevada incerteza (Configuração 2) pode não compensar a redução de incerteza em todas as regiões do depósito (Configuração 1), quando uma função global como o VPL é o objetivo da amostragem.

Tanto para o caso onde existem 60 dados iniciais, como para o caso onde existem 195 dados, realizou-se a simulação com os dados iniciais. Em cada local simulado, uma distribuição dos possíveis valores de V foi obtida e calculou-se o IEQ dessa distribuição. A Figura 3.50 apresenta o histograma dos valores de IEQ obtido com a simulação dos 60 dados iniciais e o histograma dos valores de IEQ obtido com a simulação dos 195 dados iniciais.

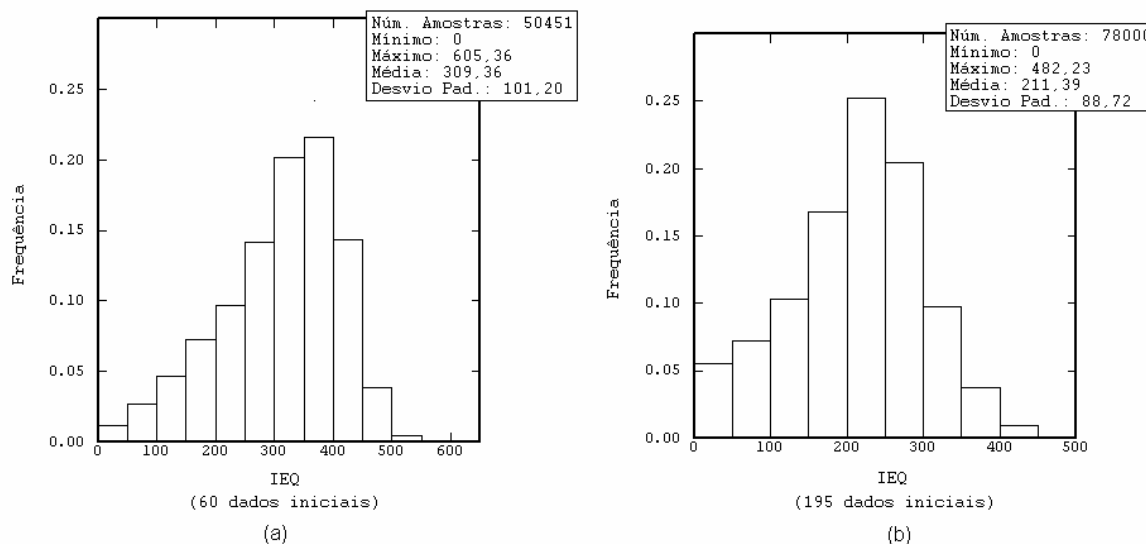


Figura 3.50 – Histograma dos valores de IEQ. (a) Simulação com base nos 60 dados da configuração inicial, valores de IEQ na região onde foram adicionadas amostras (até o limite dos dados da configuração inicial). (b) Simulação com base nos 195 dados da configuração inicial.

Os histogramas apresentados na Figura 3.50 mostram que, para o caso onde existem 60 dados iniciais, a frequência de locais com elevado IEQ é alta, isso ajuda a amostragem regular contemplar muitos locais com elevado IEQ e ser mais eficiente que a adição de amostras, apenas, nas regiões de IEQ mais extremos. Para o caso onde existem 195 dados iniciais, a frequência dos locais com elevado IEQ é menor e os valores mais extremos são mais afastados da média. Em vista desses resultados, pode-se dizer que quando a distribuição do índice de incerteza possui assimetria positiva, é mais provável que a Configuração 2 apresente um melhor desempenho que a configuração 1 na redução de incerteza sobre a função global.

3.1.7 - Cálculo do VPL com Seqüência Ótima de Lavra

Quando o sequenciamento é determinado de acordo com os valores dos blocos, visando otimizar o VPL, as regiões com valores mais elevados são, em geral, mineradas antes das demais (essa seqüência de mineração dos blocos será chamada seqüência ótima). No caso do depósito Walker Lake, a configuração 2 adicionou amostras nas regiões de elevado IEQ, que correspondem às regiões de

alto V (ppm). A eficiência da configuração 2 pode aumentar com a realização da seqüência ótima, pois a configuração 2 tende a reduzir os valores de incerteza sobre os blocos que são minerados primeiro (altos valores de V), que exercem grande influência na incerteza sobre o VPL.

A fim de avaliar a diferença entre as configurações para a seqüência ótima, o VPL foi re-calculado. O sequenciamento utilizado é explicado no próximo tópico e é esquematizado na Figura 3.51. Os novos valores de VPL foram re-calculados para dois casos anteriores: o caso onde existem 195 dados iniciais (Tabela 3.28) e o caso onde existem 60 dados iniciais (Tabela 3.29). Anteriormente, para o caso onde existem 60 dados iniciais, alguns blocos não foram considerados no cálculo do VPL. Porém, 133 desses blocos pertencem à seqüência ótima. Assim, atribui-se valor econômico próximo de US\$ 0 a esses 133 blocos.

Tabela 3.28 – Incerteza sobre o VPL (seqüência ótima), para diferentes configurações

Estatística para VPL (Seqüência Ótima)	Configuração Inicial (195 dados)	Cenário Exaustivo	
		Configuração 1	Configuração 3
E[F]	244189	237223	249077
σ	12461	8361	6839
IEQ	18372	11391	9431
Mínimo	216974	218455	234035
Máximo	275185	259342	264565

Nota: O VPL para o cenário exaustivo é 299621.

Tabela 3.29 – Incerteza sobre o VPL (seqüência ótima), para diferentes configurações

Estatística para VPL (Seqüência Ótima)	Configuração Inicial (60 dados)	Cenário Exaustivo	
		Configuração 1	Configuração 2
E[F]	197473	200709	213710
σ	21411	11614	15976
IEQ	26927	13505	22375
Mínimo	143596	156820	166461
Máximo	277728	231677	243834

Nota: O VPL para o cenário exaustivo é 299621.

Analisando os resultados da Tabela 3.28, observa-se, para todas configurações, que o intervalo de possíveis valores de VPL não engloba o VPL para o cenário exaustivo. A dificuldade dos valores de VPL obtidos por simulação englobarem o VPL exaustivo já havia sido observada na Tabela 3.7, onde a rota 1 foi

o sequenciamento adotado. Os cenários simulados possuem dificuldade em reproduzir valores extremos do cenário exaustivo, por isso os valores de VPL simulados tendem a ser mais baixos que o real. Essa dificuldade é realçada para o caso onde a seqüência ótima é utilizada, pois os blocos com valores extremos do cenário exaustivo são os primeiros a serem minerados.

Anteriormente, um erro relativo foi definido a fim de comparar os desempenhos das configurações: $[\pm(2\sigma/fr)]$. Calculando-se esse erro relativo para o caso onde existem 195 dados iniciais, observa-se que a diferença entre as configurações encontrada com o uso da seqüência ótima continua pouco significativa e semelhante à encontrada com o uso da rota 1 como seqüência de lavra.

Analisando-se os resultados da Tabela 3.29, também pode se observar a dificuldade dos valores de VPL obtidos por simulação englobarem o VPL exaustivo. Contudo, o VPL para o cenário exaustivo é um pouco menor (US\$ 283.371) quando os 133 blocos mencionados anteriormente recebem valor econômico próximo a US\$ 0 para o cenário exaustivo, assim como foi feito para as configurações. Calculando-se o erro relativo para o caso onde existem 60 dados iniciais, observa-se que a diferença entre as configurações encontrada com o uso da seqüência ótima é semelhante à encontrada com o uso da rota 1 como seqüência de lavra.

3.1.8 - Cálculo do VPL com Diferentes Valores Econômicos

No Capítulo 4, a diferença entre os desempenhos das configurações foi avaliada para um depósito de cobre. A fim de comparar os resultados obtidos nesse capítulo com os resultados obtidos no próximo capítulo, o valor econômico de cada bloco (5 m x 5 m, em x e y) foi calculado novamente, de acordo com o valor econômico dos blocos do depósito de cobre (equação 4.1). O valor do atributo V (ppm), para cada bloco, foi multiplicado por um fator (0,003), a fim de tornar a ordem de grandeza de cada valor próxima à encontrada para blocos de mesmo tamanho (em x e y) para o depósito de cobre. Assim, o novo valor de V corresponde ao teor de cobre (%). Para o depósito estudado no Capítulo 4, cada bloco de 20 x 20 x 4 m,

em x, y e z (vertical), respectivamente, possui uma massa de 4000 t. Considerando esse valor, adotou-se 250 t de massa para cada bloco de 5 m x 5 m, em x e y, do depósito Walker Lake (assim, é como se 1 bloco do depósito de cobre compreendesse 16 blocos de 5 x 5 x 4 m, em x, y e z, respectivamente, do depósito Walker Lake).

Para o cálculo do VPL, os valores econômicos dos blocos do cenário exaustivo foram utilizados na determinação da cava ótima e seqüência de mineração (obtidas utilizando o software *NPV scheduler*®). A mesma seqüência de mineração foi adotada para todos cenários gerados. Seguindo o estudo com o depósito de cobre, a produção mensal adotada foi de 92.000 t (368 blocos do depósito Walker Lake). A Figura 3.51 mostra os blocos minerados em cada período (mês) de acordo com a seqüência de mineração adotada.

O cálculo do VPL considerou uma taxa de juros de 1% ao mês e um investimento de US\$ 2.000.000, assim como foi adotado para o depósito de cobre.

De acordo com os novos parâmetros econômicos, o VPL foi re-calculado para dois casos anteriores: o caso onde existem 195 dados iniciais (Tabela 3.30) e o caso onde existem 60 dados iniciais (Tabela 3.31). Anteriormente, para o caso onde existem 60 dados iniciais, alguns blocos não foram considerados no cálculo do VPL. Porém, 133 desses blocos pertencem ao novo sequenciamento (Figura 3.51). Assim, atribui-se valor econômico próximo de US\$ 0 a esses 133 blocos.

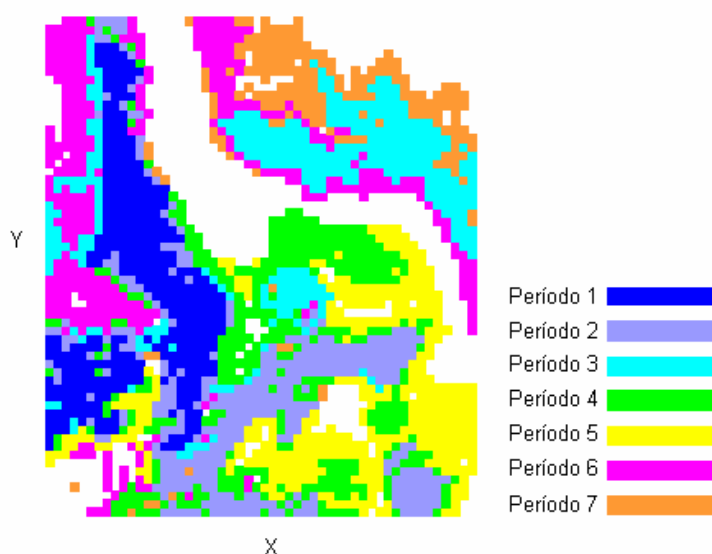


Figura 3.51 – Seqüência de lavra. Cada período compreende uma seqüência com 368 blocos (5 m x 5 m, nas direções x e y).

Tabela 3.30 – Incerteza sobre o VPL (seqüência ótima), para diferentes configurações

Estatística para VPL (Seqüência Ótima)	Configuração Inicial (195 dados)	Cenário Exaustivo	
		Configuração 1	Configuração 3
E[F]	8137446	7776062	7902542
σ	428540	280595	264144
IEQ	611196	345846	343654
Mínimo	6943578	7081231	7132180
Máximo	9235756	8540859	8435905

Nota: O VPL para o cenário exaustivo é 9204080.

Tabela 3.31 – Incerteza sobre o VPL (seqüência ótima), para diferentes configurações

Estatística para VPL (Seqüência Ótima)	Configuração Inicial (60 dados)	Cenário Exaustivo	
		Configuração 1	Configuração 2
E[F]	8321256	8163493	8370882
σ	849556	472959	671705
IEQ	891821	600566	1024539
Mínimo	5758365	6297939	6105651
Máximo	11027041	9184244	9544335

Nota: O VPL para o cenário exaustivo é 9204080 (8711265)*.

*VPL considerando valor econômico US\$ 0 para os 133 blocos.

Calculando-se o erro relativo para o caso onde existem 195 dados iniciais, observa-se que a diferença entre as configurações é pouco significativa, assim como à encontrada no caso onde o VPL foi calculado com os valores originais de V. Para o caso onde existem 60 dados iniciais, a diferença entre as configurações é semelhante à encontrada no caso onde o VPL foi calculado com os valores originais de V.

3.2 – Análise dos Resultados de Otimização Amostral para a Variável V

Ao longo do estudo com a variável V foram gerados diversos cenários simulados para cada configuração amostral testada. Cada cenário gerou um valor para a função de transferência. A distribuição dos valores da função para cada configuração foi analisada estatisticamente pelos valores E[F], desvio padrão e IEQ. Em alguns casos, foram apresentados o mínimo e o máximo de cada distribuição

dos valores da função de transferência. Diferentes cenários foram utilizados como fonte de informação para as configurações que continham, além dos dados iniciais, amostras adicionais.

Cada função de transferência estudada tem seus resultados dependentes dos valores de uma ou mais variáveis associadas aos blocos que serão minerados. Teoricamente, a configuração mais eficiente na redução de incerteza sobre a função de transferência é aquela que reduz a incerteza sobre as variáveis de *input* consideradas na função. Por exemplo, a função VPL é baseada no valor de V em cada bloco minerado. No caso onde existem 195 dados iniciais, a Configuração 3 adicionou amostras nos locais de alta incerteza (medida por IEQ) para V, por isso, essa configuração foi uma das mais eficiente na redução da incerteza sobre o VPL. Além disso, essa configuração adicionou algumas amostras próximas às regiões de alta incerteza de V estar acima de 500 ppm, colaborando para a redução da incerteza sobre a tonelagem acima desse teor de corte. Assim, antes de definir o índice de incerteza (e.g. IEQ, probabilidade do valor do ponto ou bloco estar acima de um teor de corte), e a variável à qual será aplicado esse índice, para ser usado como guia na locação de amostras adicionais, se define a função cuja incerteza deve ser reduzida. Onde deve ser amostrado depende do que se espera com a amostragem adicional. Richmond et al. (2007) mostram a variação da configuração de amostras adicionais considerando a incerteza sobre diferentes atributos.

Após analisar o desempenho das diferentes configurações de amostras adicionais na redução da incerteza sobre a função, a estatística para cada função foi analisada para diferentes cenários simulados como fonte de busca sobre o valor da amostra a ser adicionada. Os resultados mostraram que, quando uma configuração se destaca positiva ou negativamente na redução da incerteza para a função, esse fato se repete para os diferentes cenários usados como fonte de informação. Isso mostra que um cenário hipotético, simulado a partir dos dados iniciais, pode ser usado na otimização da configuração de amostras para redução da incerteza sobre uma função. Contudo, quando o número de dados iniciais é pequeno e poucas amostras são adicionadas aos dados iniciais, a incerteza sobre a função pode ser muito dependente dos valores adicionados, assim, a metodologia adotada nessa tese não pode prever o desempenho que a configuração apresentará se adotada na realidade. Mas, é importante ressaltar que para o caso onde existem 60 dados iniciais, apesar do valor de desvio padrão sobre o VPL ainda ser sensível ao cenário

usado como fonte de informação, a configuração mais eficiente na redução da incerteza sobre o VPL foi a mesma para os diferentes cenários testados, o que indica que para um banco de dados iniciais com mais de 60 dados e um atributo com variabilidade semelhante ou menor que V , provavelmente, a metodologia já possa auxiliar na escolha da configuração de amostras.

A configuração de amostras adicionais, que será adotada no depósito real, é otimizada para garantir os melhores resultados possíveis quanto à redução da incerteza (obtida por simulação geoestatística) sobre a função de transferência. As medidas de incerteza adotadas nesse estudo revelam o espalhamento da distribuição que engloba os possíveis valores da função. Assim, o ideal é que o valor da função para o depósito real esteja contido nesse intervalo. Nos exemplos apresentados, o valor real da função pertence ao intervalo ou é próximo a um dos limites do intervalo.

Convém lembrar, que o bom desempenho das configurações amostrais, está associado à amostragem inicial que contemplou as principais zonas do depósito. Por exemplo, se a amostragem inicial não tivesse contemplado as zonas de altos teores de V , o valor médio para o VPL (rota 1) seria mais baixo do que o apresentado e a incerteza sobre o VPL seria menor, visto que os blocos de alta variabilidade seriam reproduzidos como teores mais baixos. O valor real da função estaria fora dos valores de VPL obtidos com a simulação. Assim, é importante que a amostragem inicial envolva toda a extensão do depósito considerada na função e o espaçamento entre amostras permita que as principais zonas tenham sido amostradas. De acordo com essa questão, a locação de amostras em malha regular mostra a vantagem de poder encontrar algum setor de valores extremos não pego pela amostragem inicial.

De acordo com os resultados apresentados, quando um grande número de amostras deve ser adicionado, a locação de amostras em etapas, nas regiões de elevada incerteza sobre o atributo, pode ser mais eficiente do que a locação de todas amostras ao mesmo tempo. Esse procedimento aliado com utilização de uma distância mínima entre amostras (Van Meirvenne e Goovaerts, 2001) garante uma maior eficiência desse esquema de locação de amostras e economiza tempo computacional em comparação com a adição de amostras uma a uma (adotada em Pilger, 2000).

Um fator que pode contribuir para um melhor desempenho da configuração

que adiciona amostras seguindo uma malha regular é a forma com que os locais de elevada incerteza, obtidos a partir da simulação dos dados iniciais, distribuem-se ao longo do depósito. No caso onde existem 60 dados iniciais, a Figura 3.52 mostra que os locais de elevado IEQ se distribuem em diferentes regiões do depósito, isso aumenta a probabilidade da malha regular também contemplar regiões de elevada incerteza sobre o valor do atributo.

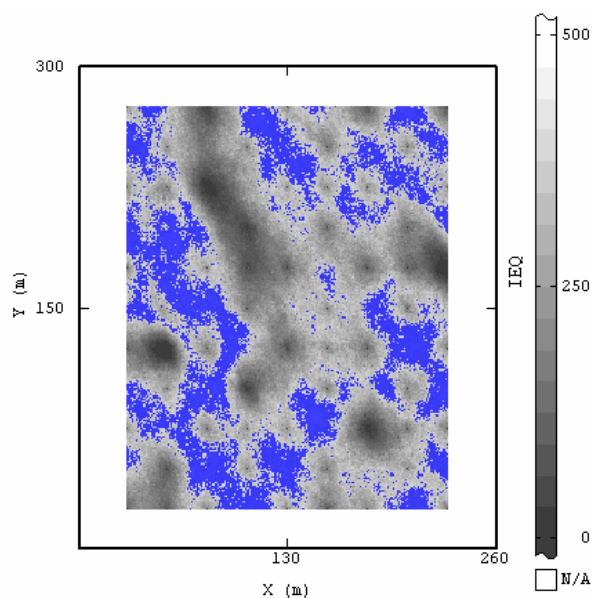


Figura 3.52 – Mapa com valores de IEQ nos locais simulados a partir dos 60 dados iniciais (na figura são mostrados os valores de IEQ na região onde foram adicionadas amostras, até o limite dos dados iniciais). Os valores mais altos de IEQ estão marcados em azul (aproximadamente os 27% mais altos).

Capítulo 4 – Implementação da Metodologia – Estudo de Caso em 3D

Da mesma forma que no Capítulo 3, nesse capítulo comparam-se os desempenhos das diferentes configurações de amostras, na redução da incerteza sobre a função de interesse. Porém, nesse capítulo, essa comparação é feita para o caso tri-dimensional.

4.1 – Variável Teor de Cobre (%)

O atributo utilizado nesse estudo é o teor de Cu (%) de um depósito localizado no estado do Arizona (EUA). O banco de dados sobre esse depósito é apresentado em Hustrulid e Kuchta (2006a). Esse atributo foi escolhido por apresentar uma distribuição assimétrica e alto coeficiente de variação. Provavelmente, existem regiões do depósito onde a incerteza sobre o valor do atributo seja maior que outras, regiões estas onde se encontram os valores extremos. Ou seja, teoricamente, existem regiões mais carentes de amostras do que outras e, assim, diferentes configurações de amostras podem gerar valores significativamente diferentes de redução de incerteza sobre a função de transferência. Dessa forma, a importância da otimização da configuração de amostras pode ser analisada para esses casos extremos.

Esse banco de dados apresenta teores de Cu (%) em uma situação tri-dimensional, coletados ao longo de 40 furos de sonda (todos verticais). Amostras de diferentes comprimentos foram analisadas para obtenção dos teores de Cu, o comprimento médio das amostras é aproximadamente 1,4 metros. Porém, na etapa de simulação é necessário que as amostras representem um mesmo suporte (a média aritmética dos teores de amostras deve representar o teor médio quando as

amostras são homogeneizadas, uma explicação sobre essa questão (aditividade) é encontrada em Armstrong (1998)). Por isso, as amostras iniciais foram transformadas em amostras com suporte semelhante.

O suporte mencionado é a massa de cada amostra. O teor é a massa do metal pela massa da amostra, ou seja, o teor deve ser referente a uma mesma massa para todas amostras. Como a densidade é considerada igual para todas amostras, basta que as amostras possuam mesmo volume para apresentarem mesma massa. A área da seção é igual para todas amostras, pois todas foram coletadas com o mesmo equipamento, assim, basta que as amostras possuam mesmo comprimento para que o volume seja o mesmo.

Uma operação de regularização foi realizada pelo software Isatis[®] para gerar amostras com comprimentos semelhantes. A partir do início de cada furo de sonda, calculou-se a média dos teores das amostras, ponderada pela espessura de cada amostra, para cada intervalo de 2 m de comprimento. Quando a soma dos comprimentos das amostras analisadas no intervalo de 2 m é menor que 1,5 m (e.g. devido à falta de parte do testemunho coletado ou devido o fim do furo) nenhum teor é atribuído àquele intervalo. As Tabelas 4.1 e 4.2 apresentam a estatística para todas amostras obtidas com a regularização e para as amostras que foram descartadas, ou seja, que possuem comprimento menor que 1,5 m. Essas tabelas mostram que, apenas, uma pequena porcentagem das amostras foi descartada. A utilização de um comprimento mínimo de 1,5 m permite que as amostras utilizadas na simulação apresentem suporte muito parecido além de garantir que poucas amostras sejam descartadas.

Tabela 4.1 – Estatística para o atributo comprimento (m) das amostras obtidas após a regularização

Estatística / Atributo	Total de Amostras Após Regularização Comprimento (m)
Número de Amostras	1451
Soma	2820,33
Média	1,94
Desvio Padrão	0,26
Máximo	2,00
Quantil 6%	2,00
Quantil 5%	1,73
Quantil 3%	1,09
Quantil 1%	0,45
Mínimo	0,01

Tabela 4.2 – Estatística para os atributos Cu (%) e comprimento (m) das amostras obtidas após a regularização. As amostras foram divididas em dois grupos: amostras com comprimento maior que 1,5 m e com comprimento menor que 1,5 m

Estatística / Atributo	Amostras menores que 1,5 m		Amostras maiores que 1,5 m	
	Comprimento (m)	Cu (%)	Comprimento (m)	Cu (%)
Número de Amostras	59	59	1392	1392
Soma	43,93	-	2776,40	-
Média	0,74	0,100	1,99	0,172
Desvio Padrão	0,41	0,235	0,04	0,384
Máximo	1,44	1,690	2,00	3,663
Quartil 0,75	1,09	0,082	2,00	0,139
Mediana	0,78	0,040	2,00	0,050
Quartil 0,25	0,47	0,020	2,00	0,020
Mínimo	0,01	0,000	1,51	0,000

As amostras regularizadas, com comprimento maior que 1,5 m, obtidas a partir dos 40 furos de sonda, compõem a configuração que é chamada “Configuração Completa” ao longo desse capítulo. A Figura 4.1 apresenta a vista em perspectiva e o histograma desagrupado dos teores de cobre para a Configuração Completa. O desagrupamento (realizado pelo software Isatis[®] (Bleines et al., 2001)) gerou um peso para cada amostra inversamente proporcional ao número de amostras que pertencem ao cubo de 70 x 70 x 2 m (nas direções leste, norte e vertical, respectivamente) centrado nessa amostra.

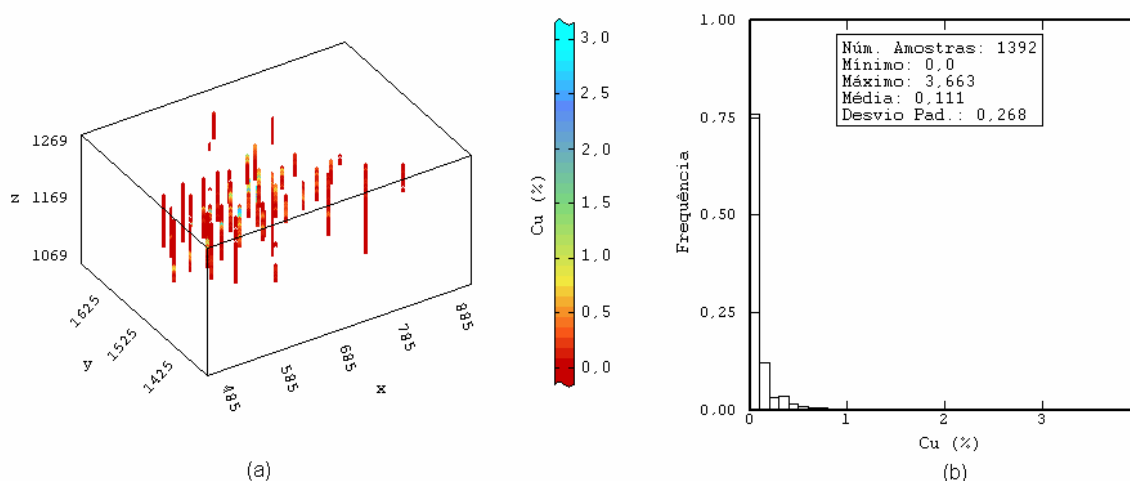


Figura 4.1 – (a) Vista em perspectiva dos teores de Cu regularizados (coordenadas em metro). (b) Histograma desagrupado para os teores de Cu regularizados.

A análise da continuidade espacial do atributo teor de cobre foi feita pela construção de correlogramas (considerando pesos de desagrupamento) dos teores

de cobre (melhores definidos que os variogramas) e um modelo de continuidade para a variável foi ajustado aos correlogramas. A fim de simular por Bandas Rotativas um cenário para representar o depósito real (referência), a distribuição dos dados originais foi normalizada e um modelo de continuidade foi ajustado aos correlogramas dos dados normalizados (Tabela 4.3).

Tabela 4.3 - Modelos de continuidade espacial para os dados originais e normalizados

	Primeira Estrutura						Segunda Estrutura				
	c0	Modelo	c1	a (D1)	a (D2)	a (D3)	Modelo	c1	a (D1)	a (D2)	a (D3)
Original	0,3	Esférico	0,5	230	50	10	Esférico	0,2	250	80	60
Normal	0,2	Esférico	0,6	190	100	60	Esférico	0,2	200	110	65

D1 – Direção de Maior Continuidade – N247,5° Dip22,5°.

D2 – Direção de Continuidade Intermediária - N157° Dip0°.

D3 – Direção de Menor Continuidade - N247,5° Dip-67,5°.

c1 – Patamar.

a – Alcance.

c0 - Efeito pepita.

Os parâmetros de busca de amostras utilizados na simulação de cada nó de *grid* consideraram o raio de busca de amostras igual ao alcance do modelo de continuidade espacial em cada direção. Além disso, um número mínimo de três dados condicionantes foi usado para simulação de cada nó e no máximo dois dados por octante de busca permitido. A avaliação da qualidade do modelo de continuidade e parâmetros de busca, que serão utilizados na simulação da referência, foi feita com a geração de cinco cenários em um grid de 2 x 2 x 2 m (direções norte (y), leste (x) e vertical (z)). Os mapas, histogramas e correlogramas desses cenários devem reproduzir (considerando as flutuações ergódicas) o mapa, histograma e correlogramas dos dados da Configuração Completa, respectivamente.

Os mapas, histogramas e correlogramas para as realizações que apresentaram maior e menor variância para os teores de Cu são apresentados nas Figuras 4.2 a 4.4. Os mapas, histogramas e correlogramas para essas realizações são semelhantes aos construídos com os dados da Configuração Completa.

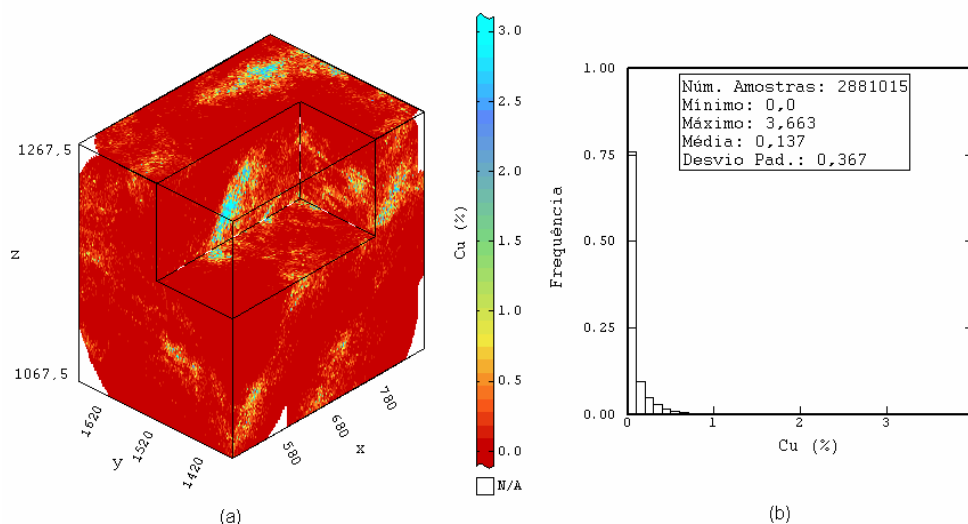


Figura 4.2 – (a) Vista em perspectiva (mostrando seções em $x=776,5$, $y=1516,5$ e $z=1184,5$) para a realização que apresentou maior variância dos teores de Cu. (b) Histograma para a realização que apresentou maior variância dos teores de Cu.

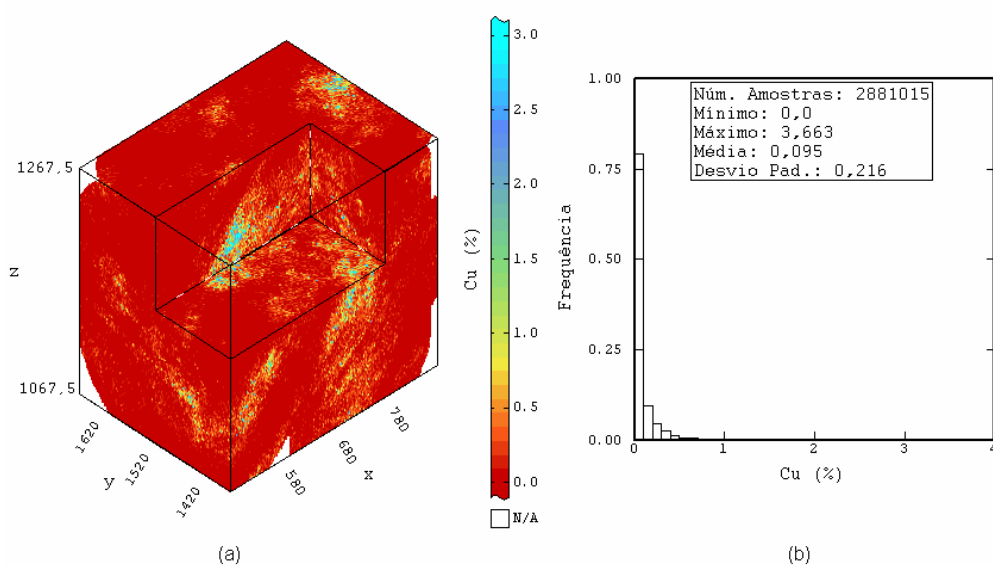


Figura 4.3 – (a) Vista em perspectiva (mostrando seções em $x=776,5$, $y=1516,5$ e $z=1184,5$) para a realização que apresentou menor variância dos teores de Cu. (b) Histograma para a realização que apresentou menor variância dos teores de Cu.

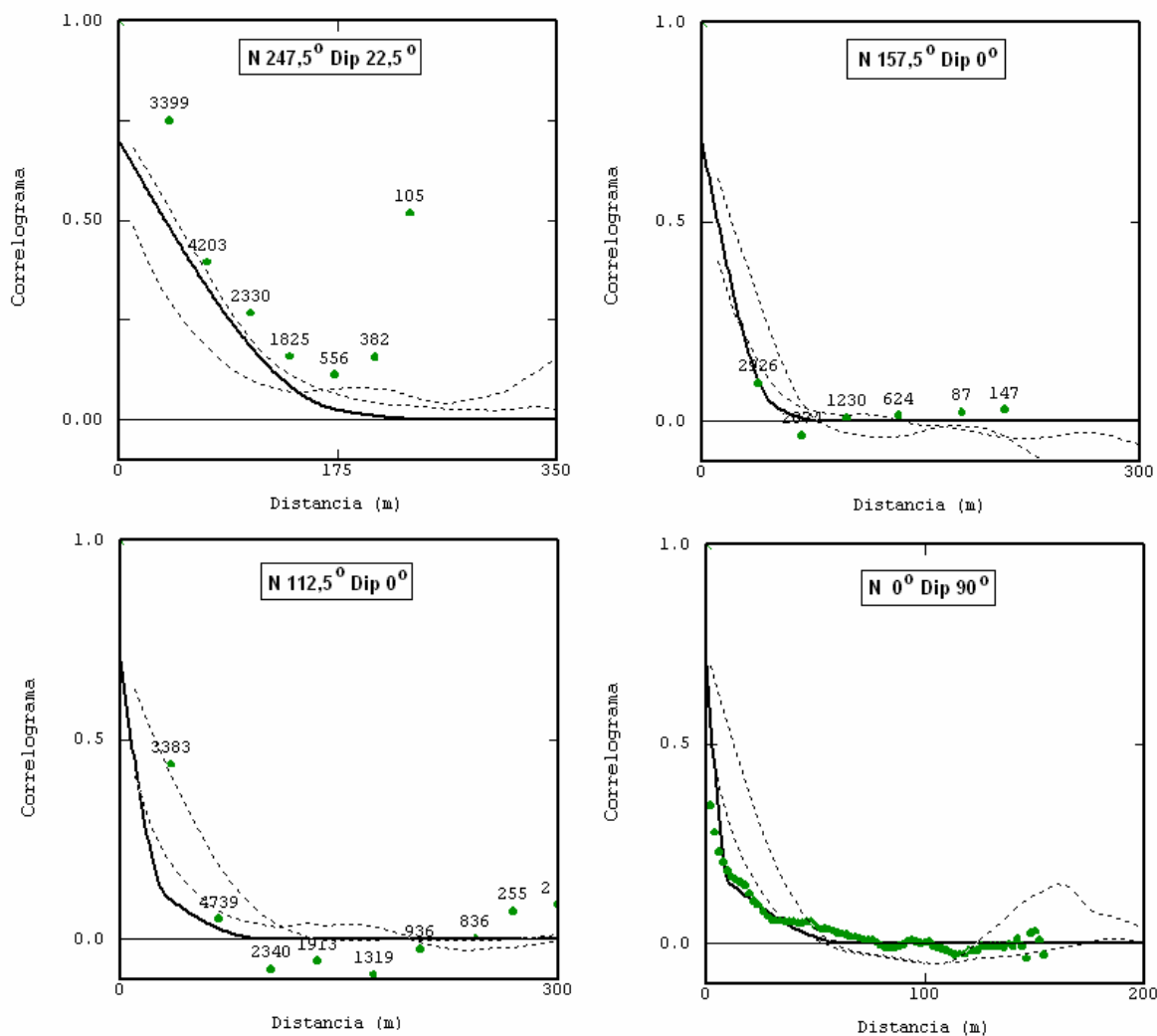


Figura 4.4 – Correlogramas para os teores de Cu. As linhas pontilhadas representam os correlogramas para as realizações que apresentaram maior e menor variância dos teores de Cu. Os pontos verdes representam o correlograma experimental para os dados originais e o número ao lado de cada ponto representa o número de pares encontrados para a respectiva distância. A curva preta contínua representa o modelo de continuidade espacial utilizado na simulação dos teores de Cu.

O cenário gerado para representar o depósito real serve como fonte de informação para as configurações de amostras adicionais que serão testadas e como valor de referência da função de transferência. Conforme comentado no Capítulo 3, o valor da função de transferência é baseado nos valores dos blocos minerados. A simulação do depósito real foi realizada no suporte das amostras (*quasi-pontual*) e a média aritmética dos pontos simulados dentro de cada bloco de 20 x 20 x 4 m em x, y e z, respectivamente, corresponde ao valor do bloco. Para que a média aritmética fosse calculada com um número significativo de pontos, o cenário referência (real) foi gerado em um *grid* com espaçamento 1 m nas três dimensões (somente os blocos calculados com 1600 pontos foram considerados no estudo). O

número de pontos simulados dentro de cada bloco é considerado suficiente quando a distribuição de blocos resultante apresenta variância próxima da variância teórica dos blocos (segundo a relação de Krige, conforme comentado no Capítulo 2). Mais adiante, nesse capítulo, será mostrado que o número de pontos simulados dentro de cada bloco do cenário real foi considerado suficiente.

Alguns furos de sonda (15) foram escolhidos, a partir dos dados da Configuração Completa, para representarem os dados iniciais ou amostragem de longo prazo. Os furos foram escolhidos de acordo com dois critérios: espaçamento entre furos e região do depósito amostrada pelo furo. Os furos foram escolhidos respeitando um espaçamento próximo de 50 m, a fim de representar uma malha comum para fase de amostragem inicial de depósitos de cobre. Os furos escolhidos pertencem à região do depósito que apresentou os mais altos teores de cobre, onde deverá começar a mineração e onde deverão ser coletadas as amostras adicionais. A Figura 4.5 apresenta um esquema do plano horizontal com a localização dos furos de sonda da Configuração Completa, dos furos de sonda selecionados para compor os dados iniciais e da região (região A) onde serão coletados os furos de sonda adicionais.

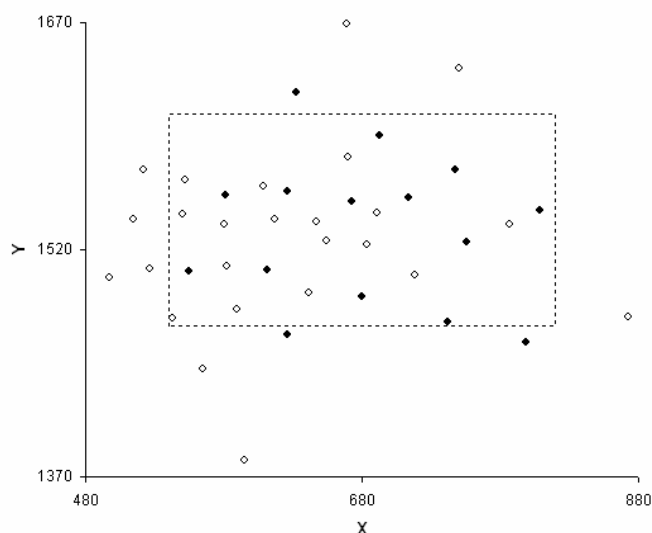


Figura 4.5 – Esquema do plano horizontal com a localização dos furos de sonda da Configuração Completa (círculos vazios), dos furos de sonda selecionados para compor os dados iniciais (círculos cheios) e da região A (linha tracejada).

A vista em perspectiva e o histograma desagrupado são apresentados na Figura 4.6. As mesmas dimensões do paralelogramo usado para o desagrupamento

dos dados da Configuração Completa foram usadas para desagrupamento dos dados iniciais e dos dados de cada configuração de amostra adicional testada.

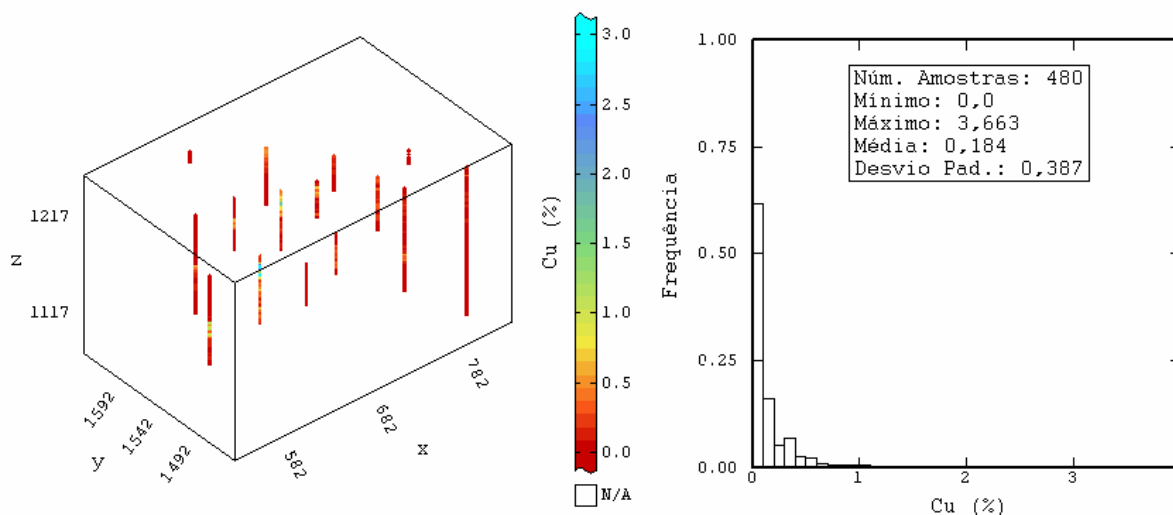


Figura 4.6 – Vista em perspectiva e histograma desagrupado para os dados iniciais.

A partir dos dados iniciais, foram gerados por bandas rotativas 80 cenários do teor de Cu (%) para a região A. O mesmo modelo de continuidade espacial e parâmetros de busca de amostras, usado para simulação do depósito real, foi usado na simulação com os dados iniciais e na simulação para as configurações de amostras adicionais que serão testadas. Como os cenários gerados, a partir dos dados iniciais, foram simulados para uma região diferente daquela considerada na simulação do depósito real, a qualidade desses cenários também foi verificada. Os mapas, histogramas e correlogramas para as realizações que apresentaram maior e menor variância para os teores de Cu são apresentados nas Figuras 4.7 a 4.9. Os mapas, histogramas e correlogramas para essas realizações são semelhantes ao mapa, histograma e modelo de continuidade espacial, respectivamente, para os dados iniciais.

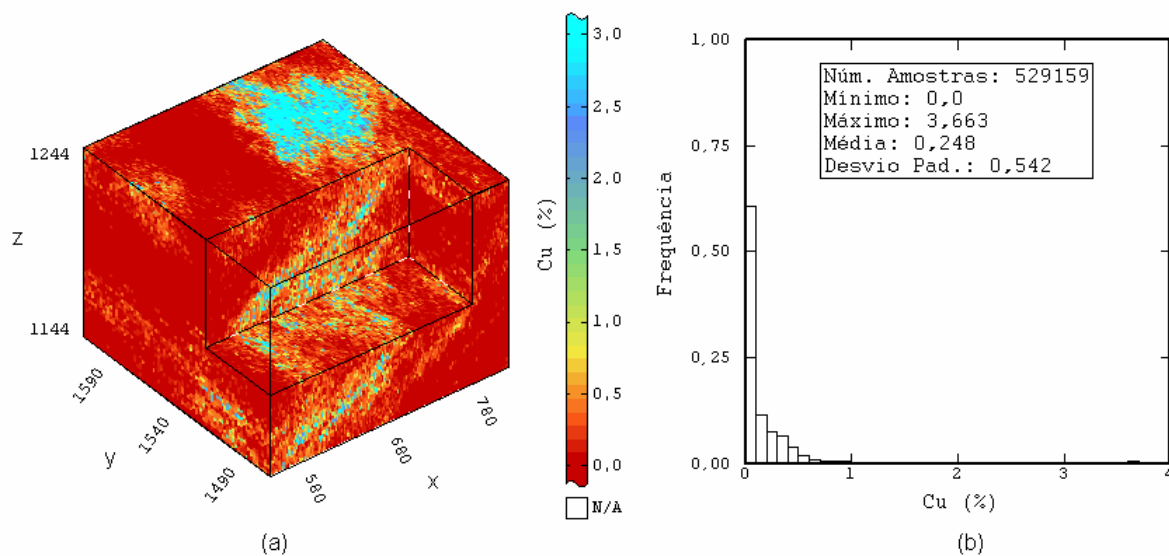


Figura 4.7 – (a) Vista em perspectiva (mostrando seções em $x=777$, $y=1517$ e $z=1185$) para a realização que apresentou maior variância dos teores de Cu. (b) Histograma para a realização que apresentou maior variância dos teores de Cu.

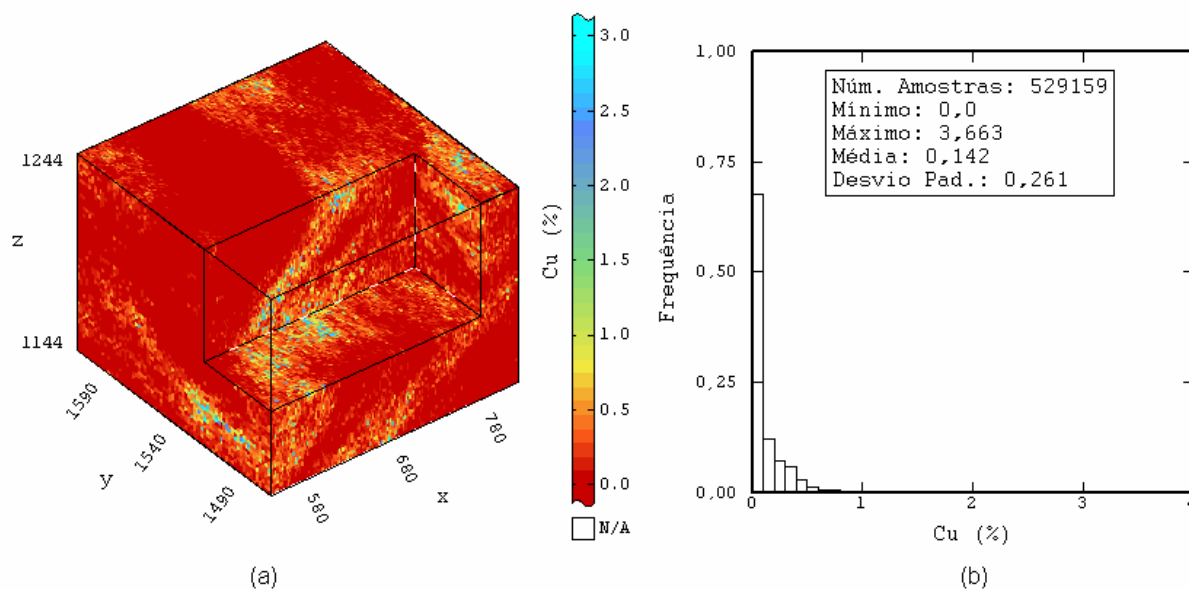


Figura 4.8 – (a) Vista em perspectiva (mostrando seções em $x=777$, $y=1517$ e $z=1185$) para a realização que apresentou menor variância dos teores de Cu. (b) Histograma para a realização que apresentou menor variância dos teores de Cu.

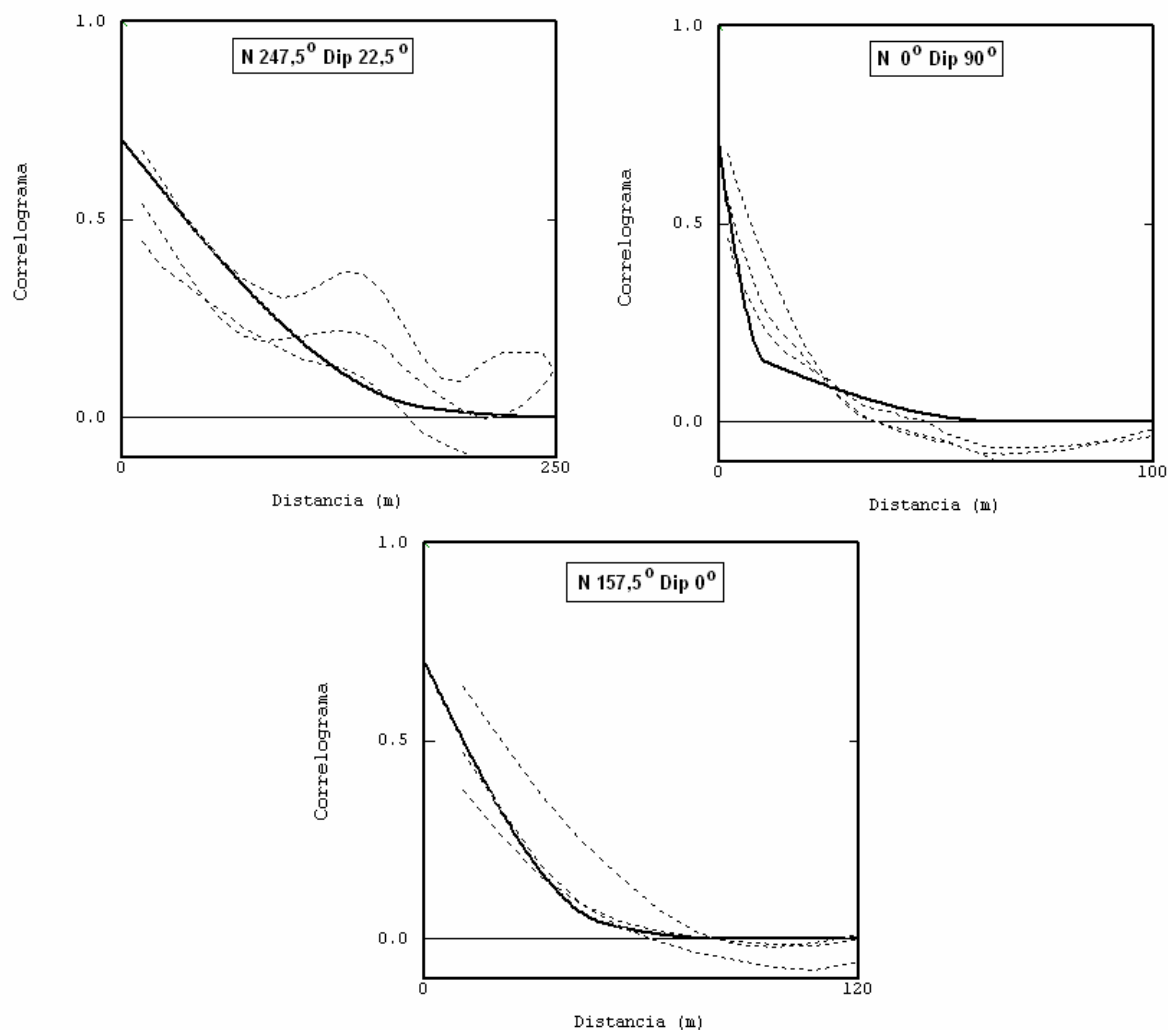


Figura 4.9 – Correlogramas para os teores de Cu. As linhas pontilhadas representam os correlogramas para a realização que apresentou maior variância dos teores de Cu, para a realização que apresentou variância intermediária e para a realização que apresentou menor variância. A curva preta contínua representa o modelo de continuidade espacial utilizado na simulação dos teores de Cu.

Conforme comentado no Capítulo 2, é necessário testar a binormalidade dos dados normalizados para que eles possam ser utilizados na simulação por bandas rotativas. A Tabela 4.4 apresenta um teste de binormalidade para os dados usados na construção do cenário real (dados da Configuração Completa) e para os dados iniciais, os resultados encontrados não permitem rejeitar a hipótese de binormalidade.

Tabela 4.4 – Teste de binormalidade

Dados da Configuração Completa - Omnidirecional Horizontal			
Número de Pares	Distância Média	Madograma / Raiz Quadrada do Variograma	
2795	4		0,49
13155	37		0,52
19581	70		0,55
17352	103		0,57
13429	138		0,57
10269	173		0,57
6403	208		0,58
4550	243		0,60
2001	278		0,56
1182	309		0,59

Dados Iniciais - Omnidirecional Horizontal			
Número de Pares	Distância Média	Madograma / Raiz Quadrada do Variograma	
913	3		0,52
926	46		0,56
2599	67		0,56
2263	104		0,56
1856	137		0,58
1306	176		0,55
615	199		0,59
578	242		0,56

Assim como para o cenário real, a simulação para os dados iniciais foi feita no suporte de pontos (amostras *quasi-pontuais*) e a média aritmética dos pontos dentro dos blocos corresponde ao valor do bloco. Os cenários obtidos pela simulações a partir dos dados iniciais foram gerados em um *grid* com espaçamento 2 m nas três dimensões. Essa malha de simulação garante que a distribuição de blocos (somente os blocos calculados com 200 pontos simulados foram considerados no estudo) apresente variância próxima da variância teórica dos blocos (considerando as flutuações ergódicas). A variância teórica da distribuição de blocos, considerando os dados iniciais, é aproximadamente 0,06 %². Assim, o número de pontos dentro de cada bloco do cenário real, mencionado anteriormente, também é considerado suficiente, visto que o mesmo modelo de continuidade espacial foi utilizado, a variância dos dados da Configuração Completa é inferior a dos dados iniciais e um maior número de pontos foi considerado dentro de cada bloco. A Tabela 4.5 apresenta a estatística do teor de Cu (%) para algumas realizações, obtidas a partir dos dados iniciais, no suporte de blocos. As realizações apresentadas possuem variância próxima ao quartil 25% (Q25) da distribuição dos

valores de variância, próxima ao quartil 50% (Q50), próxima ao quartil 75% (Q75), a menor variância e a maior variância.

Tabela 4.5 – Estatística de teores de Cu (%) para realizações no suporte de blocos

Realização	Mínimo	Máximo	Média	Variância
Menor Variância	0	1,44	0,14	0,02
Q25	0	2,25	0,16	0,04
Q50	0	1,64	0,19	0,05
Q75	0	1,76	0,17	0,06
Maior Variância	0	3,34	0,25	0,17

A primeira função de transferência considerada é o Valor Presente Líquido (VPL). A fim de calcular o VPL para cada cenário, um valor econômico (lucro líquido para o bloco) é atribuído para cada bloco do cenário, de acordo com o teor do bloco. A equação para determinação do valor do bloco é baseada naquela proposta por Hustrulid e Kuchta (2006a) (páginas 836 a 839). A equação 4.1 explica como foi calculado o valor econômico de cada bloco.

$$vb = \begin{cases} \left\{ \left[\left(\frac{Cu\%}{100} \right) \cdot (recup) \cdot (pv - cr) \right] - [cl + cb] \right\} \cdot tb & \text{se } vb > -(cl \cdot tb) \\ -(cl \cdot tb) & \text{outros casos} \end{cases} \quad (\text{Equação 4.1})$$

onde:

vb = valor econômico do bloco (US\$);

Cu% = teor (%) de cobre para o bloco;

recup = recuperação do metal na lavra e beneficiamento (até o final da etapa de concentração) = 90%;

pv = preço de venda por tonelada de metal a longo prazo = US\$ 2866/ t;

cr = custo do refino por tonelada de metal = US\$ 200/ t;

cl = custo de lavra (terceirizada) por tonelada de minério = US\$ 5/ t;

cb = custo de beneficiamento por tonelada de minério = US\$ 1,2/ t

t_b = massa do bloco = 4000 t, considerando densidade de todos blocos igual a 2,5 t/m³.

Os valores usados na equação 4.1 são baseados nos valores usados em Hustrulid e Kuchta (2006a) (páginas 836 a 839). Algumas modificações significativas foram realizadas na equação e nos valores utilizados na referência. Hustrulid e Kuchta (2006a) utilizaram um custo de lavra (incluindo lavra, transporte, custos gerais e administrativos) igual a US\$ 2/ t e um acréscimo no custo de acordo com aumento da profundidade, devido aumento na distância de transporte e dureza da rocha. Nesse estudo, por simplificação, o custo de lavra inicial foi elevado para servir como um valor médio para todos níveis (US\$ 2,5/ t), além disso, esse valor foi duplicado para representar o valor da lavra realizada por empresa terceirizada. Sabe-se inicialmente, que a mineração será de curta duração e a compra de equipamentos para lavra não será considerada. Outra modificação foi feita no preço de venda do metal, esse valor foi aproximado dos valores encontrados para o mês de dezembro de 2008.

O VPL é calculado baseado nos valores econômicos dos blocos que serão minerados, nesse caso, que pertencem a cava final. Quando a cava final é construída de maneira a otimizar o VPL, cada cenário resulta em uma cava final (e seqüência de mineração), visto que, o teor de metal de cada bloco (que determina o valor econômico do bloco) varia em cada cenário. Porém, nesse estudo, uma mesma cava final e seqüência de mineração será considerada para todos cenários gerados. A incerteza no valor do VPL irá considerar a incerteza dos teores dos blocos pertencentes à região que possivelmente será minerada.

A fim de obter uma cava final próxima àquelas que seriam obtidas para a maioria dos cenários, calculou-se o valor *E-type* (Deutsch e Journal, 1998) de cada bloco obtido a partir dos dados iniciais e o cenário dos valores *E-type* foi utilizado para construção da cava final. O valor *E-type* é a média aritmética dos teores obtidos para um mesmo bloco nos diferentes cenários gerados. A cava final foi projetada pelo software *NPV scheduler*[®], o qual realizou essa operação otimizando o Valor Presente Líquido (VPL). Esse software também gerou uma seqüência ótima de extração. Os blocos do modelo *E-type* foram limitados à topografia, ou seja apenas os blocos abaixo da superfície topográfica foram considerados no sequenciamento

de lava. A superfície topográfica foi obtida pelo método de estimativa Vizinho Mais Próximo (essa operação foi realizada no software *Surfer*® (Anon, 2002)), usando os valores de elevação do início de cada furo de sonda original. A Figura 4.10 apresenta a vista em perspectiva dos blocos que pertencem à cava final.

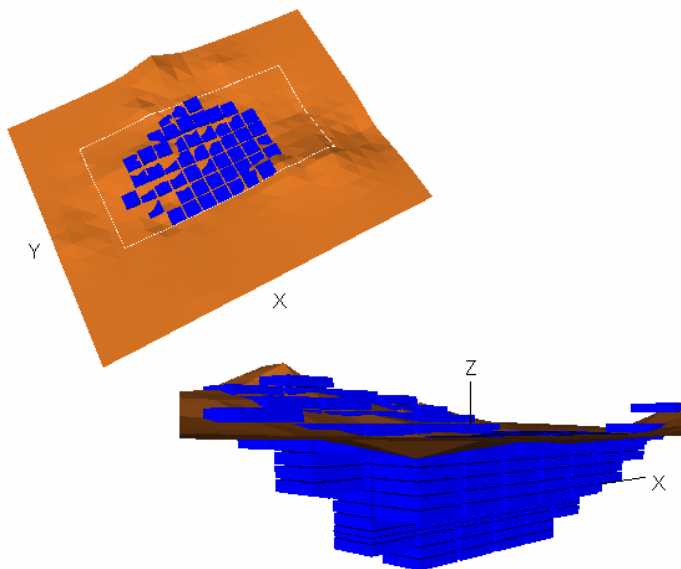


Figura 4.10 - Vistas em perspectiva dos blocos que pertencem à cava (azul), da linha que delimita a região A (linha branca) e da superfície topográfica (marrom).

Conforme a cava projetada, 277 blocos (20 x 20 x 4 m, em x, y e z, respectivamente) devem ser minerados. Como exemplo, esse estudo considerou a mineração de aproximadamente 23 blocos por mês. Além disso, a taxa de juros foi considerada 12% ao ano. Um investimento hipotético de US\$ 2.000.000 foi considerado.

O número de cenários simulados a partir dos dados iniciais deve ser suficiente para garantir o acesso ao espaço de incerteza sobre o VPL. Nesse estudo, 70 cenários foram considerados suficientes, visto que o desvio padrão dos valores de VPL torna-se aproximadamente estável antes desse número de realizações (Figura 4.11).

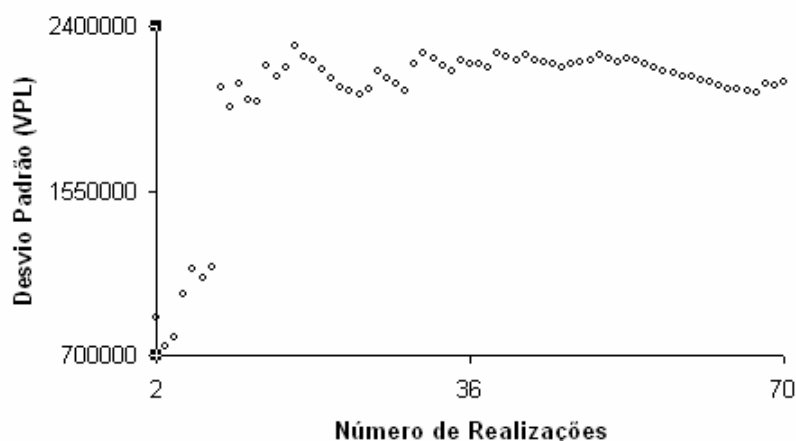


Figura 4.11 - Número de realizações versus desvio padrão de VPL.

Duas configurações de amostras adicionais foram testadas nesse estudo. Em cada configuração, seis furos de sonda foram utilizados para coletar teores de cobre do cenário real (que foi obtido a partir de uma simulação considerando todos dados existentes). Os furos de sonda coletaram amostras com espaçamento de 2 m na vertical, e esses dados novos foram adicionados aos teores de cobre existente no banco de dados iniciais. Todas sondagens adicionais possuem mergulho de 90°, assim como as sondagens inicialmente realizadas. A elevação mínima dos dados adicionados foi limitada a uma elevação (1126,5 m) maior que a mínima dos dados iniciais. A elevação máxima dos dados adicionados é um pouco menor que a elevação da topografia nas respectivas coordenadas x e y.

A Configuração 1 foi elaborada de modo que os seis furos de sonda formassem com os furos existentes uma malha aproximadamente regular (Figura 4.12 (a)). Assim, essa configuração representaria um tipo de amostragem comumente aceito. A Figura 4.13 apresenta a vista em perspectiva e o histograma desagrupado para os dados da Configuração 1.

A Configuração 2 foi construída para que as amostras fossem distribuídas nos locais com elevado IEQ (Intervalo entre Quartis). Em cada ponto simulado, os valores de IEQ foram calculados considerando os 80 possíveis valores do atributo, gerados para as realizações obtidas com os dados iniciais. A média aritmética dos valores IEQ com as mesmas coordenadas x e y foi calculada, cada média é atribuída para o local do plano horizontal com as respectivas coordenadas x e y. Assim, a região do plano horizontal com maiores valores de média de IEQ foi selecionada para receber furos de sonda adicionais. Os seis furos de sonda adicionais foram

distribuídos ao longo dessa região (Figura 4.12 (a)). Teoricamente, na região selecionada, o teor de cada bloco é mais variável para os diferentes cenários, e assim, o valor econômico de cada bloco deve ter consigo um maior grau de incerteza. A Configuração 2 visa reduzir a incerteza no VPL, reduzindo a incerteza sobre os teores dos blocos na região de elevado IEQ. A Figura 4.14 apresenta a vista em perspectiva e o histograma desagrupado para os dados da Configuração 2.

Apesar de dois furos da Configuração 1 não interceptarem blocos que pertencem à cava, todos furos de sonda adicionados pelas configurações 1 e 2 localizam-se próximos ou pertencem à cava (Figura 4.12 (b)). Ou seja, todas as sondagens foram adicionadas de forma a influenciarem nas simulações de blocos que são considerados no cálculo da função de transferência.

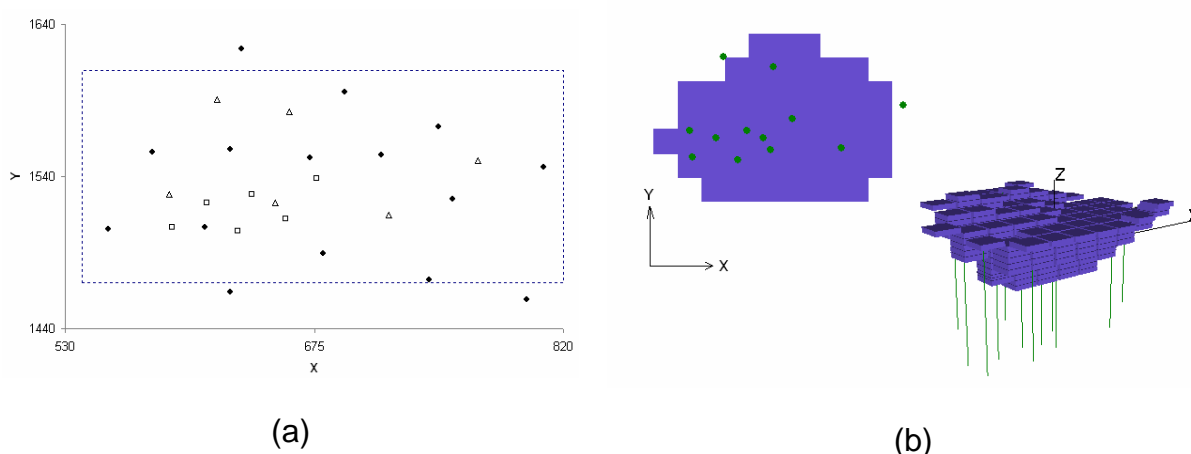


Figura 4.12 – (a) Esquema mostrando os furos de sonda que serão adicionados na Configuração 1 (triângulos) e na Configuração 2 (quadrados). Os dados iniciais são apresentados como círculos pretos. (b) Vista superior e em perspectiva dos blocos que pertencem à cava, os pontos e linhas em verde representam os furos de sonda que são adicionados pelas configurações 1 e 2.

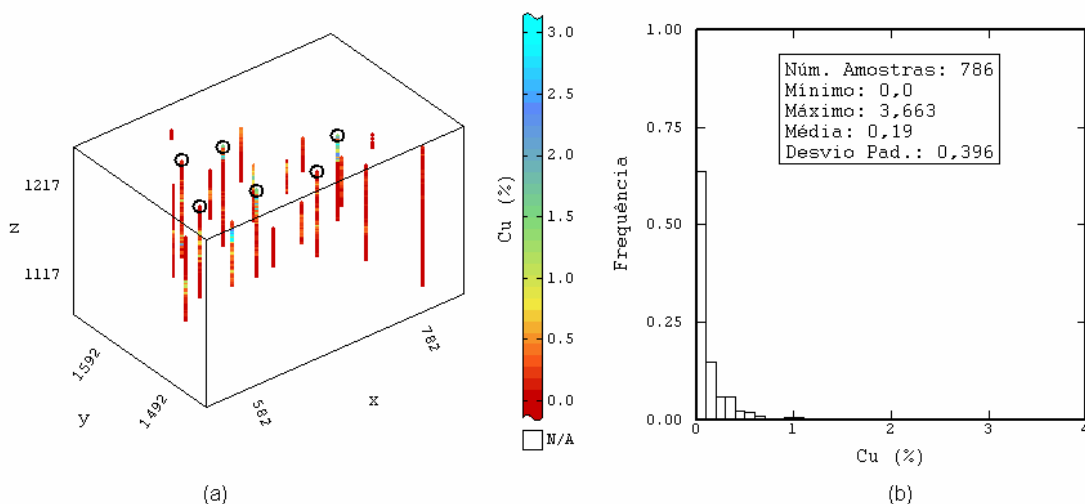


Figura 4.13 – (a) Vista em perspectiva para os dados da Configuração 1. (b) Histograma desagrupado para os dados da Configuração 1. Círculos pretos indicam os furos adicionados aos dados iniciais.

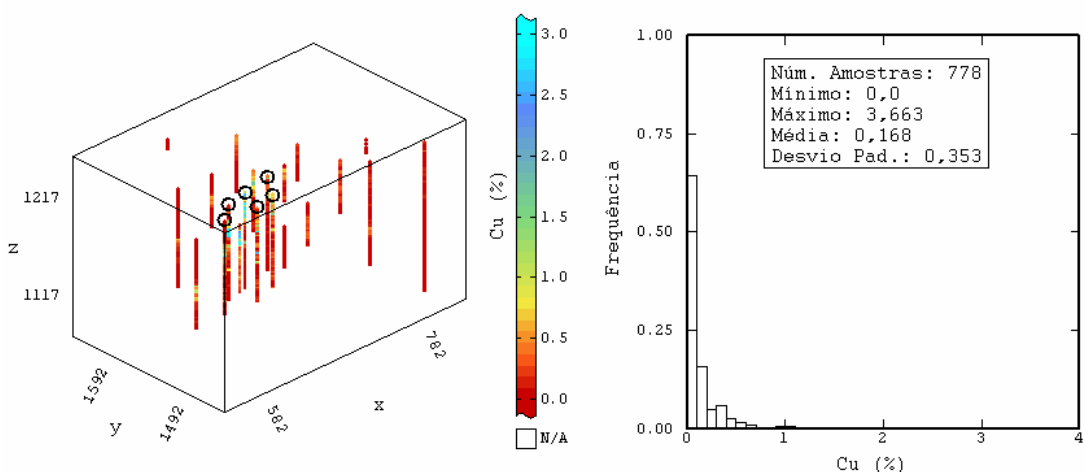


Figura 4.14 – Vista em perspectiva e histograma desagrupado para os dados da Configuração 2. Círculos pretos indicam os furos adicionados aos dados iniciais.

Conforme explicado anteriormente, 70 realizações geradas para os dados iniciais foram consideradas suficientes para mapear a incerteza sobre o VPL. Por isso, para as configurações testadas, também foram geradas 70 realizações para mapeamento da incerteza sobre a função. Para as configurações testadas, a estatística sobre o VPL é apresentada na Tabela 4.6. Essas estatísticas compreendem medidas de acurácia (reprodução em média do valor real) e precisão (incerteza sobre o valor médio). Para cada configuração, a média dos valores de VPL, obtidos para as 70 realizações, foi escrita como porcentagem do VPL para o cenário real (considerando a cava final obtida para o cenário *E-type*). Quanto mais

próxima de 100%, mais acurada é a média. A média dos valores da função também será chamada valor esperado da função ou $E[F]$. Conforme definido no Capítulo 3, o valor da função para a referência é representado por fr . A incerteza sobre os valores das funções, obtidos para as 70 realizações, foi medida pelo desvio padrão (σ) e IEQ. Para cada configuração, o quantil próximo de 97,5% (Q97,5) foi subtraído do quantil próximo de 2,5% (Q2,5), o valor resultante corresponde ao intervalo que engloba 95% dos valores de VPL.

Tabela 4.6 – Estatística para o VPL

Estatística para o VPL	Configuração Inicial	Configuração 1	Configuração 2
E[F] em US\$	1855799	3351904	3384007
E[F]/ fr	54%	98%	99%
Desvio Padrão	2107968	1573531	1306550
Q97,5-Q2,5 (95%)	8124375	6155783	4649631
Q97,5	6486899	6663632	5875639
Quartil Superior	2843617	4247570	3965144
Mediana	1660576	3212026	3222271
Quartil Inferior	577381	2295733	2388092
Q2,5	-1637476	507848	1226009
IEQ	2266236	1951837	1577052

Valor do VPL para o cenário real: 3419735.

Apesar da curva “desvio padrão de VPL x número de realizações” estabilizar próxima de 70 realizações, convém verificar se a diferença entre os valores de desvio padrão, obtidos para as configurações 1 e 2, é pouco variável para os últimos pontos do final da curva. A Figura 4.15 mostra que essa diferença é semelhante para os últimos pontos das curvas obtidas para as configurações 1 e 2.

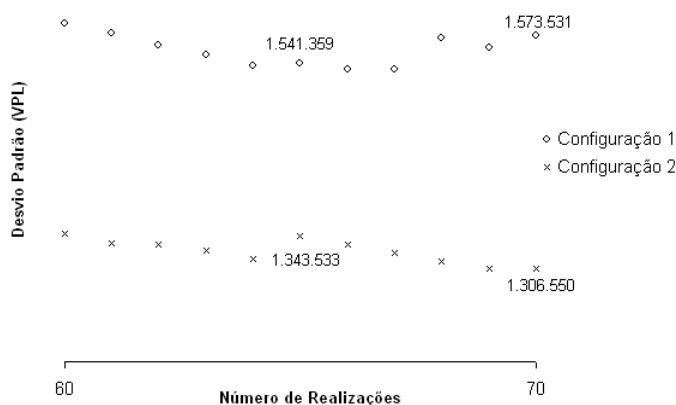


Figura 4.15 - Número de realizações versus desvio padrão de VPL.

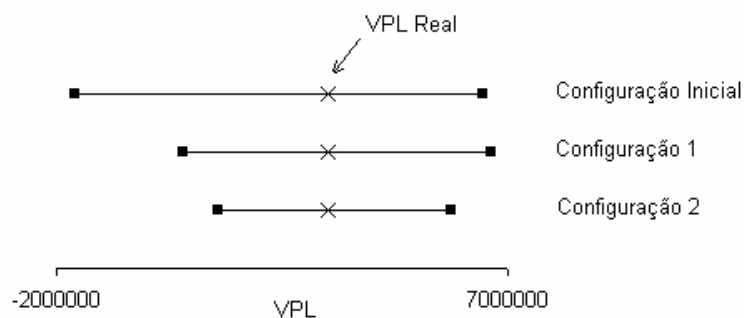
Analisando-se os resultados da Tabela 4.6, observa-se que todas configurações englobam o VPL do cenário real no intervalo de possíveis valores da função, obtido por simulação, como é esperado. As duas configurações de amostras adicionais reduziram a incerteza (desvio padrão e IEQ) em relação ao VPL, em comparação à incerteza obtida para os dados iniciais. As duas configurações de amostras adicionais mostram com que o VPL é positivo, isso não pode ser confirmado com 97,5% de certeza (Q2,5) para os dados iniciais. Ou seja, supondo que o custo com a amostragem adicional é aproximadamente US\$ 200.000 (US\$ 200 o metro de sondagem e análise química x 170 m de comprimento x 6 sondagens adicionais) é recompensado com uma maior certeza de que o empreendimento é lucrativo.

A configuração 2 reduziu mais incerteza do que a configuração 1. É difícil explicar o maior benefício da Configuração 2, se a maior redução da incerteza obtida para a Configuração 2 poderia ser decisiva para o início da mineração ou poderia modificar o valor da empresa mineradora no mercado financeiro. Contudo, pode-se afirmar que se a Configuração 1 tivesse sido implementada, seria preciso no mínimo uma sondagem a mais para se obter o grau de incerteza obtido com a Configuração 2. Ou seja, a Configuração 2 permitiu a economia de no mínimo US\$ 30.000.

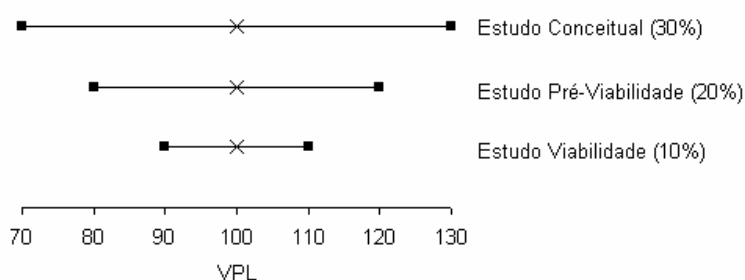
Assim como foi feito no capítulo 3, calculou-se o valor do erro em relação ao valor real de VPL. Esse erro relativo ($\pm(2\sigma/fr)$) é aproximadamente 92% para a configuração 1 e 76% para a configuração 2. Ou seja, existe uma diferença de mais de 10% no erro relativo entre os resultados das duas configurações. Essa diferença é mais significativa do que a encontrada no estudo do Capítulo 3. Por isso, nesse caso, é mais fácil afirmar que a configuração 2 traria um maior benefício se adotada na realidade.

A diferença entre o erro relativo obtido para as configurações é maior que 10%. Assim, a classificação de um estudo como conceitual, de pré-viabilidade ou de viabilidade pode depender da configuração de amostras adotada. Hustrulid e Kuchta (2006b) (baseados em Lee, 1984) definem níveis de precisão, sobre o valor da função, necessários para os diferentes estudos: estudo conceitual, precisão de 30%; estudo de pré-viabilidade, precisão de 20%; estudo de viabilidade, precisão de 10%. Por exemplo, se a Configuração 1 resulta em um erro relativo de 30% e a Configuração 2 resulta em um erro relativo de 20% para o VPL, o estudo poderia ser

classificado como conceitual com a adoção da Configuração 1 e de pré-viabilidade com a adoção da Configuração 2. Para melhor visualizar a incerteza sobre o VPL obtida para cada configuração, a Figura 4.16 mostra um gráfico com o intervalo Q2,5 a Q97,5 apresentado na Tabela 4.6. Essa figura também apresenta um esquema da precisão necessária para cada tipo de estudo, segundo Lee (1984).



(a)



(b)

Figura 4.16 – (a) Valores Q2,5 e Q97,5, apresentados na Tabela 4.6. (b) Esquema da precisão necessária para um estudo conceitual, de pré-viabilidade e um estudo de viabilidade, segundo Lee (1984).

Os blocos simulados para as duas configurações testadas também foram utilizados para o cálculo da segunda função de transferência: número de blocos (ou tonelagem, visto que foi considerada a mesma tonelagem para todos blocos) com teor acima de 1%. O número de cenários simulados a partir dos dados iniciais foi suficiente para garantir o acesso ao espaço de incerteza sobre a função (Figura 4.17). A estatística para essa função é apresentada na Tabela 4.7.

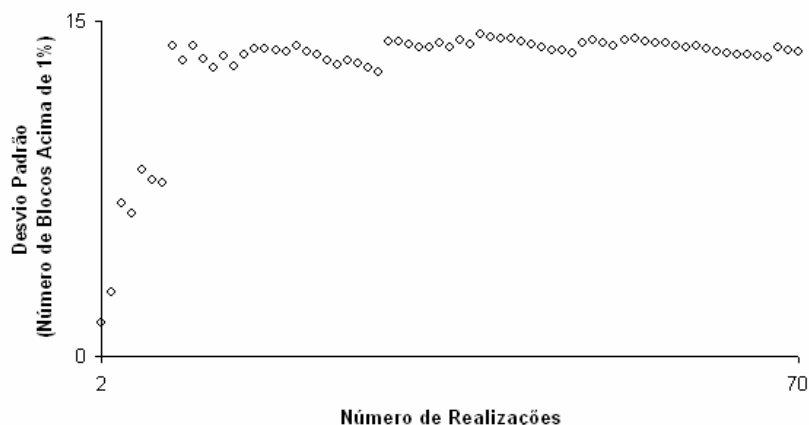


Figura 4.17 - Número de realizações versus desvio padrão de número de blocos acima de 1%.

As amostras adicionais pertencentes à Configuração 2 foram coletadas ao longo da região que apresentou maiores valores médios de IEQ (2D) e essa região é semelhante à região de elevada incerteza dos valores serem acima de 1%.

A região de elevada incerteza dos valores serem acima de 1% foi definida da seguinte forma: em cada ponto simulado, a probabilidade do teor de Cu (%) ser acima de 1% foi calculada, a probabilidade é o quociente entre o número de realizações maiores ou iguais a 1% e o número total de realizações (80); os pontos com probabilidade entre 0,4 e 0,6 (incluindo esses valores) foram convertidos em 1, os demais foram convertidos em 0 (probabilidades transformadas em indicadores, conforme explicado no Capítulo 2); a média dos valores 1 e/ou 0 com as mesmas coordenadas x e y foi calculada, cada média é atribuída para o local do plano horizontal com as respectivas coordenadas x e y e a região do plano horizontal com maiores valores de média é a região mencionada.

As duas configurações de amostras adicionais reduziram a incerteza (desvio padrão e IEQ) em relação ao número de blocos, em comparação à incerteza obtida para os dados iniciais. Novamente, a configuração 2 reduziu mais incerteza do que a configuração 1.

Tabela 4.7 – Estatística para o número de blocos acima do teor 1%

Estatística para o Número de Blocos Acima de 1%	Configuração Inicial	Configuração 1	Configuração 2
E[F]	21,2	31,8	31,4
E[F]/ fr	64%	96%	95%
Desvio Padrão	13,7	11,1	8,8
Q97,5-Q2,5 (95%)	51	44	33
Q97,5	53	56	50
Quartil Superior	28	40	36
Mediana	17	29	30
Quartil Inferior	12	24	25
Q2,5	2	12	17
IEQ	16	16	11

Número de blocos acima de 1% para o cenário real: 33.

4.1.1. – Modificação do Cenário de onde são Obtidos os Valores das Amostras Adicionais

O destaque da configuração 2, na redução da incerteza sobre a função, foi verificado considerando um outro cenário como fonte dos valores adicionais. Esse cenário (chamado realização 1) foi obtido a partir dos dados iniciais. Existe uma diferença significativa nos valores das funções (VPL e número de blocos acima de 1%) encontrados para a realização 1 e para o cenário real (Tabelas 4.6 a 4.9). Por isso, a realização 1 foi considerada nessa verificação. Assim, avalia-se se o melhor desempenho da configuração 2 não modifica com o cenário de onde são obtidos os valores adicionais, mesmo que os cenários comparados gerem valores significativamente diferentes para a função.

As Figuras 4.18 e 4.19 apresentam as vistas em perspectiva e os histogramas para os teores de cobre dos blocos obtidos a partir do cenário real e da realização 1 (gerada com os dados iniciais).

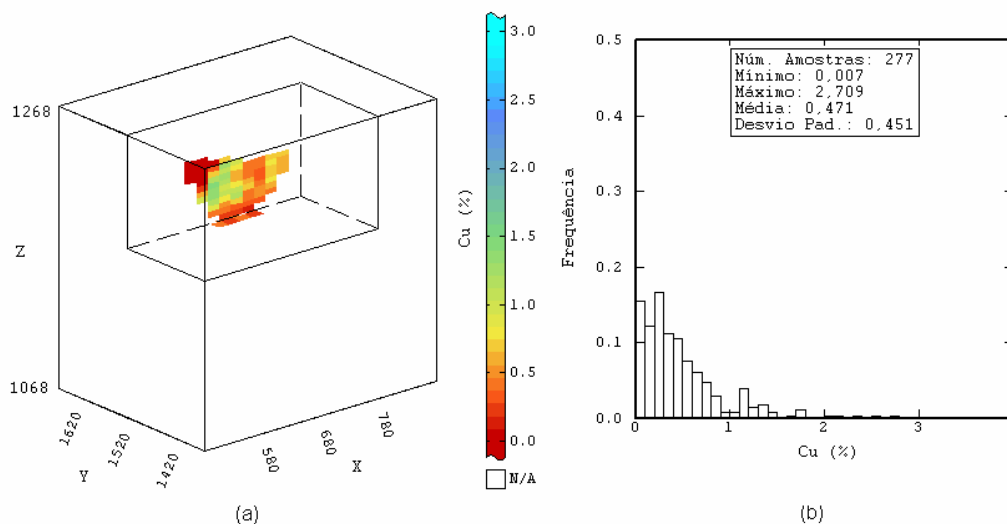


Figura 4.18 - Vista em perspectiva e histograma dos teores dos blocos para o cenário real (mostrando seções em $x=770$, $y=1520$ e $z=1186$). Apenas os blocos minerados são considerados.

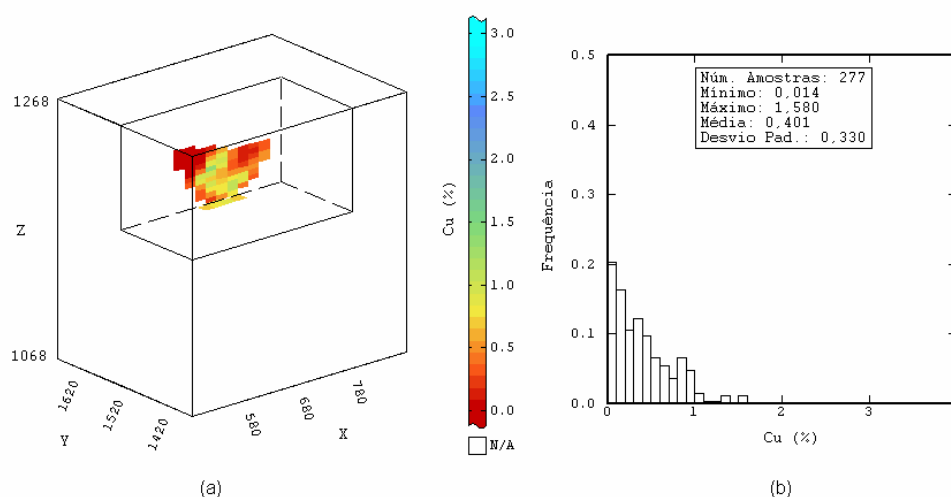


Figura 4.19 - Vista em perspectiva e histograma dos teores dos blocos para a realização 1 (mostrando seções em $x=770$, $y=1520$ e $z=1186$). Apenas os blocos minerados são considerados.

A idéia, nessa etapa, foi construir uma configuração 1b e uma configuração 2b com amostras nas mesmas posições das configurações 1 e 2, respectivamente, e atribuir aos locais das amostras adicionais os teores de cobre simulados para a realização 1, nos respectivos locais.

Antes de apresentar os resultados algumas observações sobre a posição das amostras nas configurações 1b e 2b devem ser mencionadas. Os *grids* onde foram simulados o cenário real e a realização 1 não possuem simulações exatamente nos mesmos locais, assim, as amostras adicionadas, agora, apresentam

um pequeno deslocamento (em z e um deslocamento máximo de 0,5 m em x e y) em relação as respectivas amostras, obtidas a partir do cenário real. Além disso, a elevação mínima dos dados adicionados foi limitada a elevação mínima do *grid* onde foi simulada a realização 1 (1141 m).

Assim como nas configurações 1 e 2, ambas configurações 1b e 2b compreendem seis furos de sonda adicionais, com espaçamento amostral de 2 m na vertical. A elevação máxima dos dados adicionados é um pouco menor que a elevação da topografia nas respectivas coordenadas x e y.

A Figura 4.20 apresenta a vista em perspectiva e o histograma desagrupado para os dados da Configuração 1b. A Figura 4.21 apresenta a vista em perspectiva e o histograma desagrupado para os dados da Configuração 2b.

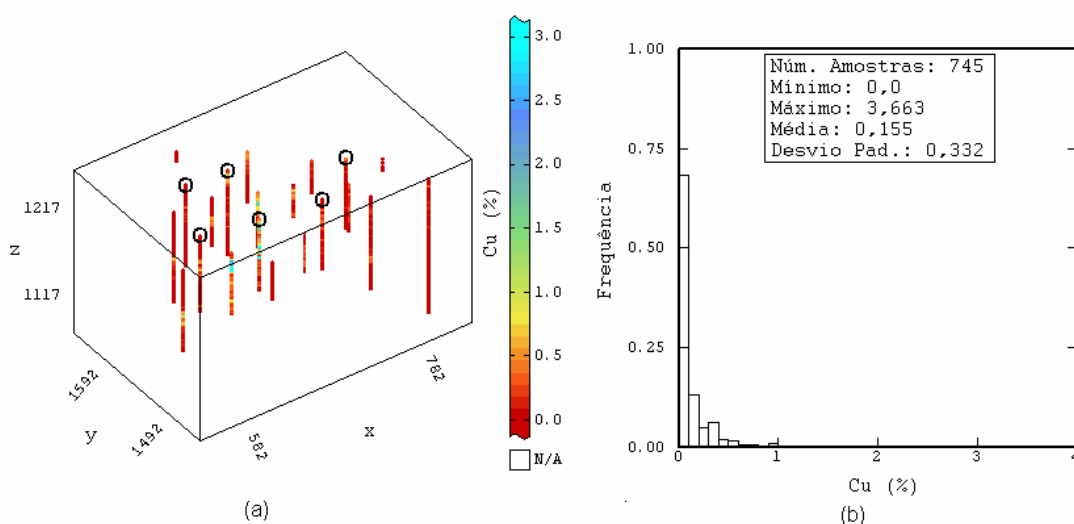


Figura 4.20 – (a) Vista em perspectiva para os dados da Configuração 1b. (b) Histograma desagrupado para os dados da Configuração 1b.

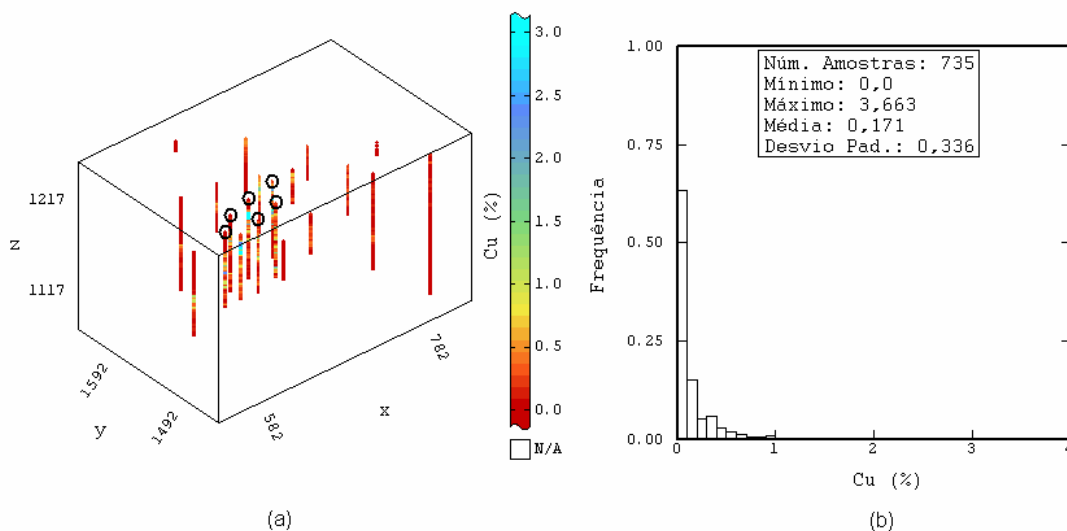


Figura 4.21 – (a) Vista em perspectiva para os dados da Configuração 2b. (b) Histograma desagrupado para os dados da Configuração 2b.

Para as configurações testadas, a estatística sobre o VPL (considerando 70 realizações) é apresentada na Tabela 4.8. A estatística sobre o número de blocos com teor acima de 1% (considerando 70 realizações) é apresentada na Tabela 4.9.

Tabela 4.8 – Estatística para o VPL

Estatística para o VPL	Configuração 1	Configuração 2
E[F] em US\$	2391553	1860324
E[F]/ fr	144%	112%
Desvio Padrão	1450053	1251753
Q97,5-Q2,5 (95%)	5397629	5255330
Q97,5	5273573	4766202
Quartil Superior	3339339	2516469
Mediana	2198912	1688181
Quartil Inferior	1519561	980964
Q2,5	-124056	-489129
IEQ	1819778	1535505

Valor do VPL para a realização 1: 1660870.

Tabela 4.9 – Estatística para o número de blocos acima do teor 1%

Estatística para o Número de Blocos Acima de 1%	Configuração 1	Configuração 2
E[F]	24,8	17,3
E[F]/ fr	207%	144%
Desvio Padrão	11,1	7,9
Q97,5-Q2,5 (95%)	45,0	28,0
Q97,5	50,0	34,0
Quartil Superior	30,0	21,8
Mediana	24,0	16,0
Quartil Inferior	16,3	12,0
Q2,5	5,0	6,0
IEQ	13,8	9,8

Número de blocos acima de 1% para a realização 1: 12.

Analisando-se os resultados das Tabelas 4.8 e 4.9, observa-se que todas configurações englobam o valor da função para a realização 1 no intervalo de possíveis valores da função, obtido por simulação, como é esperado. No caso onde o cenário real foi utilizado como fonte de informação, as configurações 1 e 2 produziram resultados, em média, mais acurados do que os dados iniciais, sendo acurado o valor próximo ao valor da função para o cenário real. Contudo, para o caso em que a realização 1 é a fonte de informação adicional, nem sempre as configurações de amostras adicionais tornaram os resultados mais acurados do que os obtidos com os dados iniciais, sendo acurado o valor próximo ao valor da função para a realização 1. Além disso, comparando as Configurações 1 e 2, nem sempre a configuração mais acurada no caso onde o cenário real é usado como fonte de informação adicional é a mais acurada para o caso onde a realização 1 é usada como fonte de informação adicional. A acurácia dos resultados obtidos para cada configuração depende dos valores adicionados no banco de dados e do cenário escolhido como fonte de informação. Assim, o uso de um cenário simulado para avaliar o desempenho de configurações de amostras adicionais não pode ser aplicado para questões de acuracidade.

Analisando-se as medidas de incerteza (desvio padrão e IEQ) em relação ao VPL e o número de blocos acima do teor 1%, observa-se que a configuração 2 reduziu mais incerteza do que a configuração 1 para ambas funções. O mesmo havia ocorrido para o caso em que o cenário real serviu como fonte dos valores

adicionais. Assim, o destaque da configuração 2 na redução da incerteza sobre as funções consideradas foi reproduzido para um outro cenário fonte de informação adicional. Além disso, as grandezas dos valores de desvio padrão e IEQ são muito semelhantes às encontradas para o caso onde a referência é fonte de informação adicional. Essas medidas de incerteza são mais afetadas pelo local das amostras adicionais do que pelo valor dessas amostras. Assim, usando um cenário simulado a partir dos dados iniciais como fonte de informação adicional, pode-se avaliar a eficiência que diferentes configurações de amostras apresentarão na redução da incerteza sobre a função, se forem aplicadas na realidade. Essa avaliação permite dizer se a diferença na redução da incerteza produzida por diferentes configurações de amostras é significativa e em caso positivo qual configuração deverá ser adotada na realidade, além disso, essa avaliação mostra, aproximadamente, qual a incerteza sobre a função que será obtida após a adição de novas amostras.

Considere que a distribuição dos valores de VPL é normal. A amplitude do intervalo que engloba 95% dos valores da função é 4σ . Suponha que a configuração 2 foi testada, usando a realização 1 como fonte de informação, para se obter uma aproximação da amplitude do intervalo que será obtido na realidade. De acordo com o resultado para a realização 1, após a nova amostragem, 95% dos valores de VPL vão pertencer a um intervalo de aproximadamente US\$ 5.000.000 de amplitude, contudo, não se pode dizer os limites inferior e superior desse intervalo.

É importante lembrar que, quando existem poucos dados iniciais, o desvio padrão para a distribuição dos valores da função pode ser sensível ao cenário fonte de informação (conforme observado no Capítulo 3). O ideal é que o resultado seja testado para cenários extremos como fonte de informação, a fim de verificar se o desvio padrão é semelhante para esses cenários. Além disso, a previsão do desvio padrão que será obtido após a nova amostragem considera o uso de um mesmo algoritmo para realizar novas simulações, pois alguns autores (e.g. Gotway e Rutherford, 1994) mostram que o espaço de incerteza sobre a função pode variar com o algoritmo.

Quando as configurações de amostras são testadas usando um cenário hipotético como fonte de informação, deve-se ter cuidado em aplicar a medida de erro relativo usada anteriormente ($\pm(2\sigma/fr)$). Essa medida depende do valor da função para a referência e esse valor pode ser bem diferente do valor da função

para o depósito real. Por exemplo, o empreendedor pode acreditar que existe uma diferença de aproximadamente 24% no erro relativo (para o VPL) obtido para as duas configurações, quando o valor da referência é US\$ 1660870, contudo a diferença entre as configurações seria 16% se fossem aplicadas no depósito real. Por isso, na análise da importância da diferença entre as configurações também deve se considerar a possível variação do valor da função para o depósito real.

4.2 – Análise dos Resultados para a Variável Cu (%)

Ao longo do estudo com a variável Cu (%) foram gerados diversos cenários simulados para cada configuração amostral testada. Cada cenário gerou um valor para a função de transferência. A distribuição dos valores da função para cada configuração foi analisada estatisticamente pelos valores de $E[F]$, desvio padrão e IEQ.

Assim como no estudo com o banco de dados Walker Lake, a função VPL calculada para o depósito de cobre é baseada no valor de Cu (%) em cada bloco minerado. A Configuração 2 adicionou amostras nos locais de alta incerteza (medida por IEQ) para Cu (%) e foi mais eficiente que a amostragem regular (configuração 1) para redução da incerteza sobre o VPL. Além disso, a configuração 2 adicionou amostras próximas aos locais de elevada incerteza dos valores de Cu (%) serem acima de 1% e foi mais eficiente que a amostragem regular (configuração 1) para redução da incerteza sobre o número de blocos com teor acima de 1%. Assim, no estudo para o depósito de cobre duas hipóteses levantadas no estudo da variável V foram confirmadas: a utilização dos índices de incerteza para locação de amostras adicionais pode ser mais eficiente que uma amostragem regular para redução da incerteza sobre a função. Quando a densidade de amostras é igual a comumente usada na prática de campo e o atributo em estudo apresenta alta variabilidade (coeficiente de variação para os dados originais próximo de 2,1); a diferença na redução da incerteza obtida para diferentes configurações de amostras adicionais pode ser significativa economicamente.

Conforme comentado no estudo para a variável V, o ideal é que o valor da função de transferência para o depósito real esteja incluído no intervalo de possíveis

valores da função obtido para cada configuração. Nos exemplos apresentados, esse fato foi observado para as duas configurações de amostras adicionais testadas.

Outra questão que foi abordada no estudo para o depósito de cobre é a adição de furos de sonda em um depósito, tentando contemplar a região do espaço tri-dimensional que apresenta maior variabilidade para o atributo de interesse. Dois procedimentos foram propostos para definir a região do plano (bi-dimensional) onde a sondagem deve ser realizada. O local do plano horizontal onde a sondagem deve ser realizada deve permitir que um grande número de pontos simulados, com elevados valores para o índice de incerteza, seja contemplado. Nesse caso, um procedimento foi adotado para local neste plano cada índice de incerteza: IEQ e probabilidade do teor no local estar acima de um teor limite.

As sondagens adicionadas no estudo do cobre possuem mergulho de 90° , respeitando o padrão das sondagens originais. Porém, em alguns casos, dependendo do mergulho do corpo de minério, as sondagens devem apresentar um mergulho diferente de 90° . Por exemplo, é importante a coleta de informações na direção de menor continuidade do atributo de interesse cuja variabilidade dos teores é maior do que outras direções, para uma mesma distância. Nos casos em que a direção de maior continuidade do corpo de minério possui mergulho acentuado em relação a horizontal, a direção de menor continuidade possui mergulho diferente de 90° . Os procedimentos propostos para locação de sondagem no plano horizontal podem ser aplicados nos casos que a sondagem deve ser inclinada em relação a horizontal.

No caso apresentado, a locação de amostras é baseada nos valores de média de índices de incerteza ou de indicadores de índices de incerteza. Cada valor de média é calculado para pontos simulados em coordenadas que apresentem mesmas coordenadas x e y e diferentes elevações, ou seja, pontos simulados em uma "linha vertical". Os valores de médias são atribuídos às respectivas coordenadas x e y no plano horizontal e os maiores valores de média geram os locais candidatos a receberem sondagem adicional. Quando a sondagem for inclinada em relação a horizontal, os valores de média devem ser calculados para pontos simulados em uma linha com inclinação próxima à inclinação que a sondagem deva apresentar. O valor da média é atribuído para o ponto de interseção da linha inclinada com a topografia do terreno. Os maiores valores de média locados

na superfície topográfica são candidatos à receberem sondagem adicional.

De acordo com comentários anteriores, dois fatores podem influenciar no melhor desempenho da Configuração 2: a distribuição dos valores dos índices de incerteza e o espalhamento das regiões de elevada incerteza.

Após a simulação dos dados iniciais, uma média dos valores de IEQ foi atribuída a pontos do plano horizontal, a fim de localizar as novas sondagens da Configuração 2. O histograma desses valores de média de IEQ (Figura 4.22) mostra uma assimetria positiva e valores de média de IEQ muito elevados em relação aos demais. Assim, a adição de sondagens nesses locais de elevada média de IEQ pode contribuir na redução de grande parte da incerteza sobre o VPL.

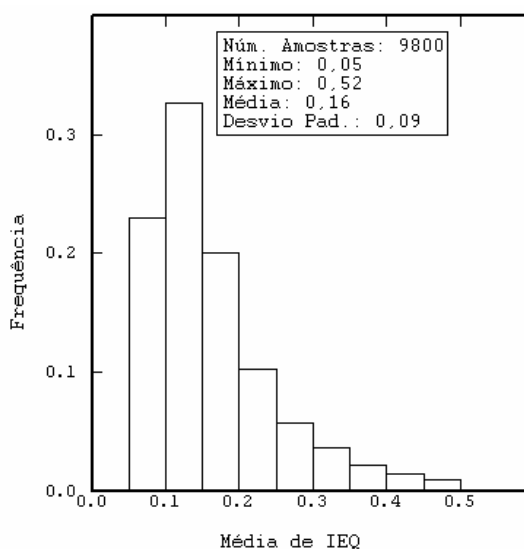


Figura 4.22 – Histograma dos valores de média de IEQ obtidos por simulação com base nos dados da configuração inicial.

O espalhamento dos locais de elevada incerteza sobre o valor do atributo, obtido a partir da simulação com base nos dados iniciais, influencia na diferença entre os resultados das configurações porque representa a chance que a amostragem em malha regular possui de contemplar regiões de elevada incerteza. Quando os locais de elevada incerteza são mais dispersos ao longo do depósito, a diferença entre a eficiência das configurações na redução da incerteza pode ser menor do que a diferença encontrada quando os locais de elevada incerteza pertencem a uma única região do depósito. Para o caso do teor de Cu, os locais de elevada incerteza pertencem a uma única região do depósito (Figura 4.23), apenas

duas sondagens adicionadas pela malha regular contemplam essa região.

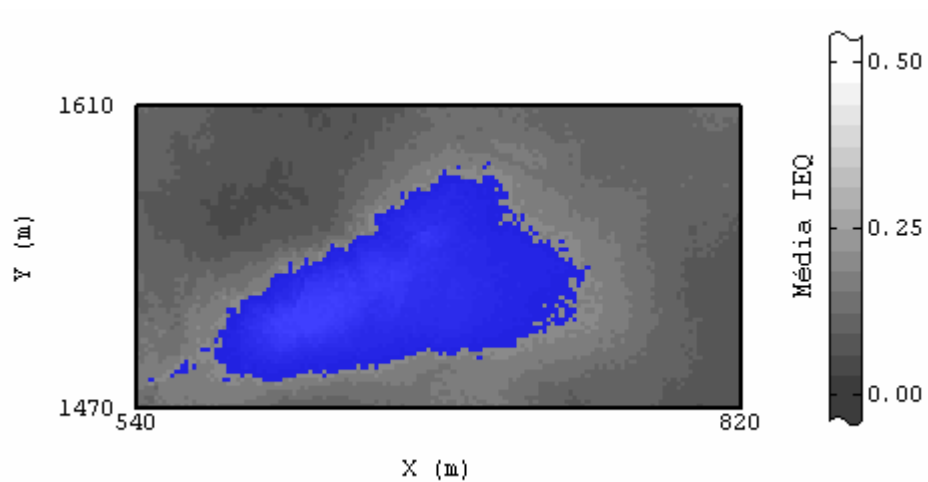


Figura 4.23 - Média de valores de IEQ obtidos com a simulação dos dados iniciais, os locais com valores mais altos estão em destaque azul (aproximadamente 24% mais altos valores).

Capítulo 5 – Algoritmo para Locação de Amostras Adicionais

Nos capítulos 3 e 4, comparou-se o desempenho de dois procedimentos para locação de amostras visando a redução da incerteza sobre uma função de transferência. Os procedimentos comparados analisam a locação de amostras adicionais em uma malha regular e a locação de amostras adicionais guiada pela incerteza sobre o valor do atributo, em cada região do depósito considerado. No capítulo 4, conforme os comentários realizados, houve diferença significativa entre os valores de incerteza sobre a função, obtidos para as diferentes configurações de amostras adicionais. Porém, no capítulo 3, essa diferença foi irrelevante e a adoção de uma malha regular, comumente aceita, poderia ser indicada. Nesse caso, a amostragem regular poderia trazer tanto benefício quanto outra configuração de amostra, em relação ao objetivo da amostragem. Esta amostragem regular poderia revelar algum setor de valores extremos do atributo, não contemplado pela amostragem inicial.

Conforme foi explicado anteriormente, existem alguns fatores que influenciam na diferença entre o benefício da amostragem regular e o da locação de amostras adicionais nas regiões de elevada incerteza sobre o valor do atributo. Por isso, é importante que os dois procedimentos para locação de amostras sejam testados em cada estudo.

Contudo, a indústria da mineração requer decisões tomadas agilmente e, muitas vezes, o profissional responsável pela locação das amostras adicionais não dispõe do tempo que é necessário para construir as diferentes configurações de amostras adicionais e avaliar o desempenho de cada uma na redução da incerteza sobre a função de interesse. Assim, o desenvolvimento de um algoritmo que permita a construção das configurações de amostras adicionais, automaticamente, ajudaria a acelerar a avaliação do benefício da incorporação de novos dados e a análise dos diferentes procedimentos de locação de amostras, para as etapas de planejamento

de longo e médio prazo em um empreendimento mineiro.

Esse capítulo apresenta um algoritmo para construção de dois tipos de configurações de amostras, mencionados ao longo da tese, que utiliza funções comuns em softwares de mineração. Assim, a construção das configurações pode ser gerada automaticamente em softwares que apresentem todas as funções necessárias, como por exemplo, Isatis[®] (Bleines et al., 2001) e Datamine Studio[®] (Anon, 2003). Softwares como o Isatis[®] e o Datamine Studio[®] apresentam uma ferramenta para construção de um arquivo que armazena a sequência de rotinas utilizadas (e.g. *Journal File* no Isatis[®]) e que quando executado realiza de forma sequencial os processos empregados. Por edição dos arquivos automáticos, permite-se ao usuário, que antes da execução do arquivo, este possa definir os parâmetros que devem ser usados em cada processo. Assim, a sequência de operações necessárias para construção das configurações pode ser realizada automaticamente nesses softwares.

Ambos softwares, Isatis[®] e Datamine Studio[®], apresentam funções para geração de simulação geoestatística. Assim, todas as funções necessárias para implementação do método de construção de configurações amostrais e simulação usando cada configuração podem ser encontradas em um mesmo software. A seguir, é apresentado um exemplo de *Journal File* (Isatis[®]) que compreende as funções mencionadas. As duas configurações de amostras são criadas para o caso tri-dimensional, ou seja, as configurações determinam a posição dos furos de sonda (verticais ou inclinados).

As configurações construídas correspondem à locação de n furos de sonda em uma malha regular e de n furos nas regiões de elevada incerteza para o atributo de interesse.

5.1 – Algoritmo para Construção das Configurações

Os passos do algoritmo para construção das configurações de amostras, usando funções comuns em softwares do ramo da mineração, são descritos a seguir. Os passos i a iv são realizados para construção da configuração que propõe

a adição de amostras formando uma malha aproximadamente regular com os dados iniciais (Configuração 1). Os passos v a viii são realizados para construção da configuração que visa adicionar amostras nas regiões de alta incerteza para valores do atributo (Configuração 2).

i. Criação de um *grid* para adicionar amostras regularmente – Configuração 1

Um *grid* (*grid1*) é criado para locar as amostras adicionais. Cada nó do *grid1* é um local candidato a receber sondagem adicional. Os nós do *grid1* devem ser distribuídos formando uma malha regular com os dados iniciais (que, comumente, são coletados em uma malha aproximadamente regular). O usuário deve definir a origem do *grid1*, espaçamento entre nós e número de nós em x (leste) e y (norte). Apenas um nível deve ser criado em z (vertical). Contudo, o usuário deve seguir algumas restrições:

- Os nós do *grid1* devem ser localizados no meio dos furos iniciais. Para isso, o espaçamento entre nós do *grid1* em x e y deve ser semelhante ao espaçamento entre os dados iniciais nessas direções. Dependendo do número de sondagens que se pretende adicionar, ou seja, do número de sondagens adicionais que devem ser distribuídas regularmente ao longo da região, o espaçamento entre nós do *grid1* pode ser duas vezes (ou um múltiplo maior) o espaçamento entre os dados iniciais. Essas restrições garantem que os locais candidatos a receberem sondagem adicional formem uma malha regular com os dados iniciais (Figura 5.1, o *grid1* é definido por linhas pretas);
- Quando se pretende adicionar n sondagens, o número de nós do *grid1* deve ser igual ou um pouco maior que n . Alguns furos iniciais podem não obedecer à malha regular e serem localizados muito próximos de alguns dos nós do *grid1*. Quando isso ocorre, cada nó do *grid1*, próximo ao furo inicial, deixa de ser candidato a receber sondagem. Nesses casos, um número de nós para o *grid1* um pouco maior que n é recomendado. A distância mínima permitida entre o local candidato a receber sondagem e o furo inicial pode ser definida pelo usuário (ver passo ii). Na Figura 5.1, o *grid1* possui 16 nós e se deseja adicionar 12 sondagens. Um nó do *grid1* é muito

próximo de um furo inicial, assim, 15 nós são candidatos a receberem sondagem. Apesar disso, a locação de 12 sondagens é garantida;

- A origem do *grid1* deve coincidir com um local simulado para o cenário de referência e o espaçamento do *grid1* deve ser múltiplo do espaçamento do *grid* simulado para a referência, naquela direção. De acordo com a metodologia descrita ao longo da tese, a referência para a busca de dados adicionais é um dos cenários simulados usando os dados iniciais. Os locais selecionados para receberem amostras adicionais recebem os valores simulados para a referência nos respectivos locais. Assim, a restrição mencionada nesse item garante que existam valores da referência nos locais escolhidos para receber sondagem. A Figura 5.1 mostra que os nós do *grid1* coincidem com os nós do *grid* (*gridsim*) usado na simulação dos dados iniciais;
- A elevação (z) do *grid1* deve ser igual ao maior nível do *gridsim*. Essa restrição será explicada nos próximos passos;

No exemplo da Figura 5.1, três nós do *grid1* foram descartados, para que apenas 12 sondagens fossem adicionadas (círculos cheios azuis). Quando o número de nós candidatos à sondagem é um pouco maior que o número de sondagens que se pretende adicionar, a escolha dos nós que devem receber sondagem depende de como o software enumera os nós de um *grid*. Por exemplo, o Isatis[®] enumera os nós de um *grid* primeiro em ordem crescente de x , depois em ordem crescente de y e depois em ordem crescente de z . No exemplo da Figura 5.1, os 12 primeiros nós são escolhidos para receberem sondagem, para o software Isatis[®] os 12 primeiros nós (candidatos a receberem sondagem) correspondem aos círculos azuis cheios.

Pelo algoritmo descrito nessa seção, o número de nós do *grid1* não pode ser muito maior do que o número de sondagens adicionais, pois desse modo, as sondagens adicionais são locadas, apenas, em uma região do depósito e não distribuídas regularmente como é o objetivo da Configuração1.

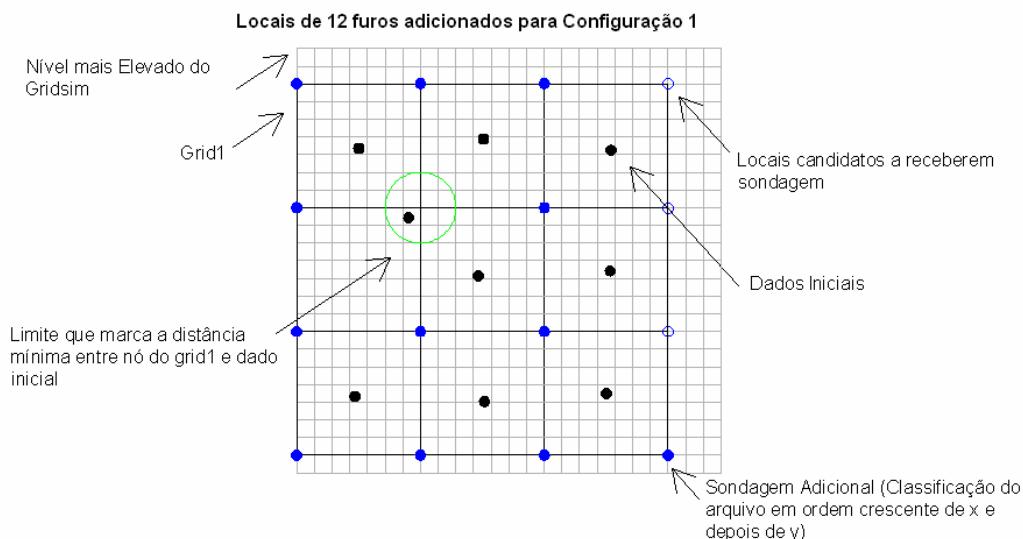


Figura 5.1 – Exemplo da construção do *grid1* (*grid* criado para posicionar as sondagens adicionais formando uma malha aproximadamente regular com as sondagens iniciais).

ii. Vizinho próximo para determinar os nós do *grid1* que não devem receber amostras adicionais – Configuração 1

Antes da execução do algoritmo para construção das configurações, um arquivo (COLLAR) com os valores de x e y do começo de cada furo inicial deve ser importado para o software. Esse arquivo será usado em algumas estimativas. Para os próximos passos, é necessário que um mesmo valor de elevação seja associado a todos os pares x,y . O valor de elevação deve ser igual ao maior nível do *gridsim*.

Os valores da coordenada x do arquivo COLLAR são atribuídos ao *grid1* pelo método Vizinho Mais Próximo (atribui ao nó de *grid* o valor do atributo que estiver na posição mais próxima). A distância de busca de amostras em x e y deve ser igual à distância mínima permitida entre amostras. Os nós do *grid1* que forem interpolados (círculo laranja na figura 5.2) não são candidatos a receberem furos adicionais, pois estão a uma distância menor do que a permitida dos furos iniciais. A Figura 5.2 apresenta os valores da coordenada x do arquivo COLLAR interpolados no *grid1* (mesmo exemplo apresentado na Figura 5.1).

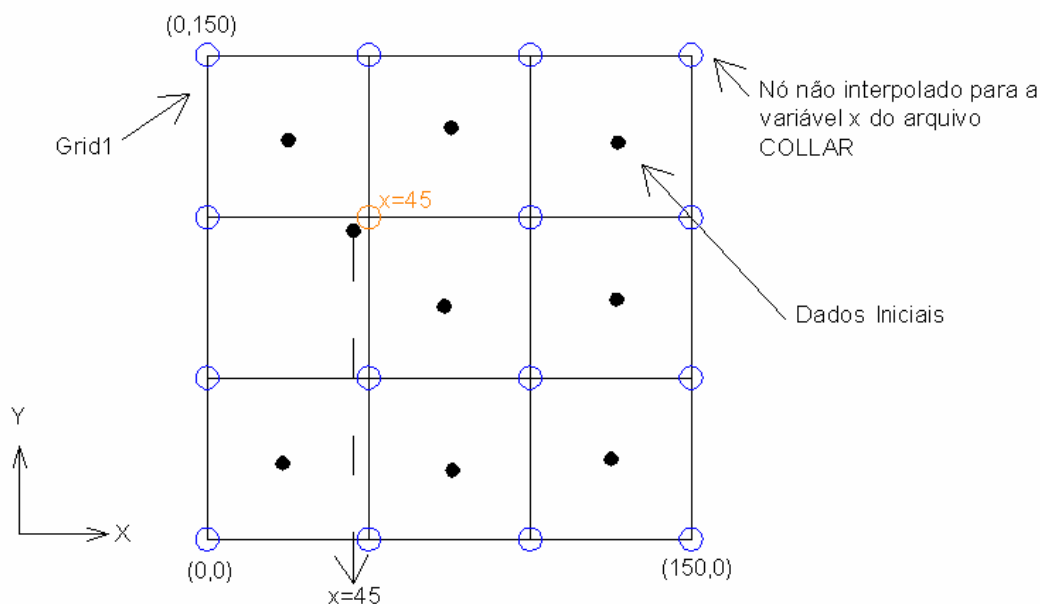


Figura 5.2 – Esquema da estimativa dos nós do *grid1* com base na variável *x* do arquivo COLLAR.

iii. Escolha dos nós do *grid1* que devem receber amostras adicionais – Configuração 1

No passo ii, os nós do *grid1* que não foram interpolados para a variável *x* do arquivo COLLAR correspondem aos nós candidatos a receberem sondagem. No passo iii, *n* nós são escolhidos entre os candidatos a receberem sondagem, sendo *n* o número de sondagens que se pretende adicionar. Os nós escolhidos demarcam os locais das novas sondagens no plano *x-y*.

As funções necessárias para realizar computacionalmente a escolha dos nós que devem receber sondagem são detalhadas no Anexo A – A.1.

iv. Locação de amostras (3D) e adição dessas amostras aos dados iniciais - Configuração 1

Nos passos anteriores, foram definidos os nós do *grid1* que demarcam os locais do plano *x-y* onde as novas sondagens devem ser realizadas. Agora, deve-se definir os locais do espaço que serão amostrados pelas novas sondagens, coletar os

valores do cenário de referência nos locais das novas amostras e adicionar as novas amostras aos dados iniciais.

Os locais do espaço tridimensional que serão amostrados pelas sondagens são nós do *gridsim*. A Figura 5.3 apresenta um esquema da seleção dos nós do *gridsim* onde serão adicionadas novas amostras. A Figura 5.3 é baseada no exemplo da Figura 5.1 e considera que a inclinação das novas sondagens é de aproximadamente 72° em relação a horizontal.

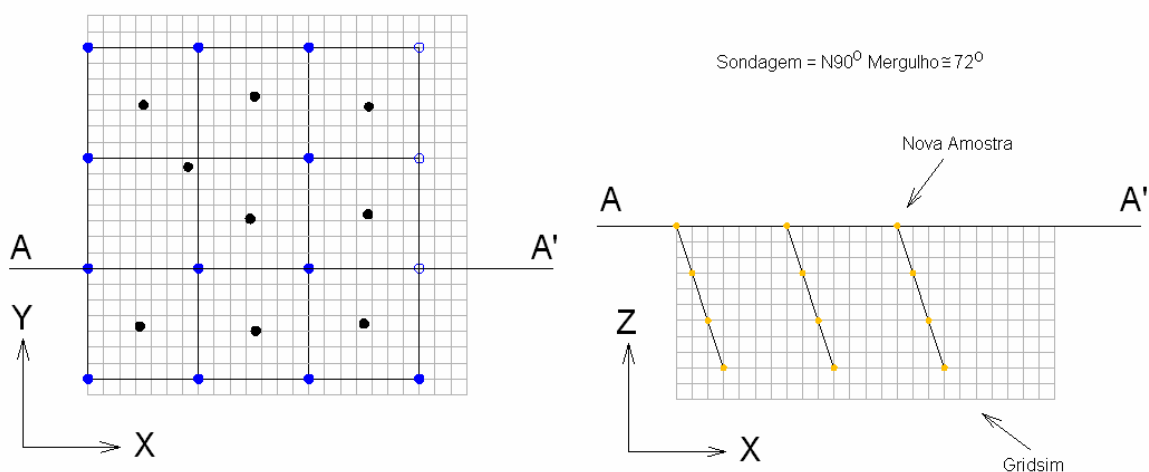


Figura 5.3 – Locais do *gridsim* amostrados pelas novas sondagens.

As novas sondagens podem ser inclinadas em relação a horizontal, pois, conforme mencionado no capítulo 4, a inclinação das sondagens depende das direções de maior e menor continuidade do atributo. No algoritmo para construção das configurações, supõe-se que cada sondagem adicional possui mesma inclinação ao longo de sua extensão e que todas sondagens adicionais possuem mesma inclinação em relação a horizontal.

Dependendo da malha do *gridsim* nas três dimensões, algumas direções do espaço podem não interceptar pontos simulados. Ou seja, a sondagem pode ser projetada em uma direção que não intercepta nós do *gridsim*. Nesse caso, não existiriam valores para as amostras adicionais no cenário de referência. Por isso, nesse caso, deve-se escolher uma direção para a sondagem (fictícia) mais próxima da que a sondagem deve apresentar quando for executada no campo.

As funções necessárias para selecionar, computacionalmente, os nós do *gridsim* onde serão adicionadas novas amostras e para atribuir o valor do cenário de referência nos locais das novas amostras são detalhadas no Anexo A – A.2.

A Configuração 1 possui os valores para o atributo de interesse coletados na campanha inicial somados aos valores coletados nos nós simulados do *gridsim* na campanha de adensamento de amostras, seguindo uma malha regular para adição das sondagens.

Para que o banco de dados da Configuração 1 possa ser criado automaticamente dentro do software, esse software deve apresentar funções que permitam a união de dados de arquivos diferentes. Porque, os valores do atributo de interesse para as amostras da campanha inicial, que estão no arquivo com os dados iniciais, devem ser combinados aos valores do atributo de interesse nos locais selecionados para buscar as amostras adicionais, e que estão no arquivo com os dados do *gridsim*. Tanto o Datamine Studio[®], quanto o Isatis[®] possuem funções que permitem a criação automática de uma variável com os valores do atributo de interesse para os dados iniciais e para as amostras adicionais. Ou seja, uma variável que compreende os dados da Configuração 1 e que será usada na etapa subsequente de nova simulação.

v. Cálculo dos índices de incerteza em cada local simulado para os dados iniciais – Configuração 2

Na Configuração 2, as amostras são adicionadas nas regiões de elevada incerteza para o valor do atributo. Para adicionar as amostras nessas regiões, é preciso acessar a incerteza sobre o valor do atributo em cada região do depósito. Essa incerteza é acessada por meio de simulação geoestatística. Desse modo, primeiro, propõe-se medir a incerteza sobre o valor do atributo em cada local simulado utilizando somente os dados iniciais. Com a simulação geoestatística para os dados iniciais, n possíveis valores do atributo simulado são gerados para cada nó do *gridsim*, e a incerteza do valor do atributo é baseada na distribuição destes n valores. Conforme explicado no Capítulo 3, se o objetivo da amostragem é a redução da incerteza sobre o VPL, a incerteza sobre os n valores é medida pelo IEQ. Se a função de transferência for a tonelagem de minério acima de um teor de

corde, a incerteza deve ser medida pela probabilidade do valor do atributo estar acima do teor naquele local.

O depósito em estudo pode apresentar regiões com elevado IEQ onde a faixa de valores do atributo não interessa economicamente. Não é necessário reduzir a incerteza sobre o teor nessas regiões, pois os blocos ali removidos não são minério, mesmo com o aumento de informação o valor do VPL não modificará significativamente. Assim, sugere-se o cálculo, em cada local simulado, da probabilidade do valor do atributo estar acima de um teor limite (TL), junto com o cálculo do IEQ. Nesse caso, os locais com elevada probabilidade do valor do atributo estar acima do teor limite e com elevado IEQ serão considerados candidatos a receberem amostra adicional. Os valores de probabilidade são transformados em indicadores para serem usados nos próximos passos:

$$I(u;X) = \begin{cases} 1 & \text{para } \text{prob}(Z(u) \geq TL) > X \\ 0 & \text{outros casos} \end{cases} \quad (\text{Equação 5.1})$$

onde:

$I(u;X)$ = indicador no local u para uma probabilidade limite X ;

$Z(u)$ = valor do atributo Z no local u ;

$\text{prob}(Z(u) \geq TL)$ = probabilidade do valor do atributo Z no local u ser maior ou igual a TL ;

X = probabilidade limite para considerar que o local é candidato a receber sondagem adicional. No *Journal File* criado, foi adotado 0,7. O valor 0,7 pode ser modificado, dependendo do valor do TL , dependendo se os nós precisam ter elevada probabilidade de estarem acima de TL para serem considerados candidatos.

vi. Transformação da incerteza medida tri-dimensionalmente para uma solução bi-dimensional – Configuração 2

No último passo, determinou-se o índice de incerteza para cada ponto simulado no espaço. A idéia da Configuração 2 é que as novas sondagens acessem

pontos com elevada incerteza. Contudo, as sondagens contemplam um número limitado de locais do espaço, por isso, essas devem ser locadas no plano x-y de forma a contemplar o maior número possível de pontos com elevada incerteza. Para avaliar os locais do plano x-y que permitem que a sondagem (vertical ou inclinada) locada ali alcance um grande número de pontos com elevada incerteza, a seguinte rotina é proposta:

- I. Constrói-se um *grid* (*grid2*) semelhante ao *grid* simulado para os dados iniciais (*gridsim*). Nas direções x e y, o *grid2* possui os mesmos parâmetros (origem, número de nós e espaçamento entre nós) do *gridsim*. Na direção z, o *grid2* possui um único nível com elevação igual a do maior nível do *gridsim*;
- II. Interpola-se as medidas de incerteza (IEQ e probabilidade) dos pontos do *gridsim* no *grid2*. A operação de interpolação pode ser Média Móvel (*Moving Average* como é implementada no software Isatis[®]), que calcula a média dos dados dentro do alcance de busca. Conforme esquema da Figura 5.4 (seguindo exemplo da Figura 5.3), o elipsóide de busca de dados para interpolação deve apresentar o maior eixo na direção das novas sondagens e esse eixo deve alcançar todos níveis do *gridsim*, para garantir que todos dados naquela direção sejam usados na interpolação. Além disso, os demais eixos do elipsóide de busca devem ser pequenos, para que só permitam que nós do *gridsim* na direção da sondagem, ou na direção muito próxima a sondagem, sejam considerados no cálculo da média.

A interpolação por Média Móvel é realizada para a variável IEQ e para os indicadores criados pela equação 5.1. O resultado da interpolação para a variável corresponde à média dos índices de incerteza encontrados na direção da sondagem. Teoricamente, os locais do *grid2* com média elevada para o IEQ, permitem que a sondagem ali locada contemple um grande número de pontos com elevada incerteza. Assim, os maiores valores de *média de IEQ* para o *grid2* correspondem aos locais do plano x-y candidatos à sondagem. Já, a *média dos indicadores* é calculada para mostrar os locais que não devem receber sondagem, os locais que possuem baixa média para os indicadores não são candidatos à sondagem. Teoricamente, os nós do *grid2* com baixa *média de indicadores* apontam que a sondagem ali locada alcança muitos nós que não são interessantes

economicamente, ou seja, nós que apresentam baixa probabilidade de estarem acima do teor limite.

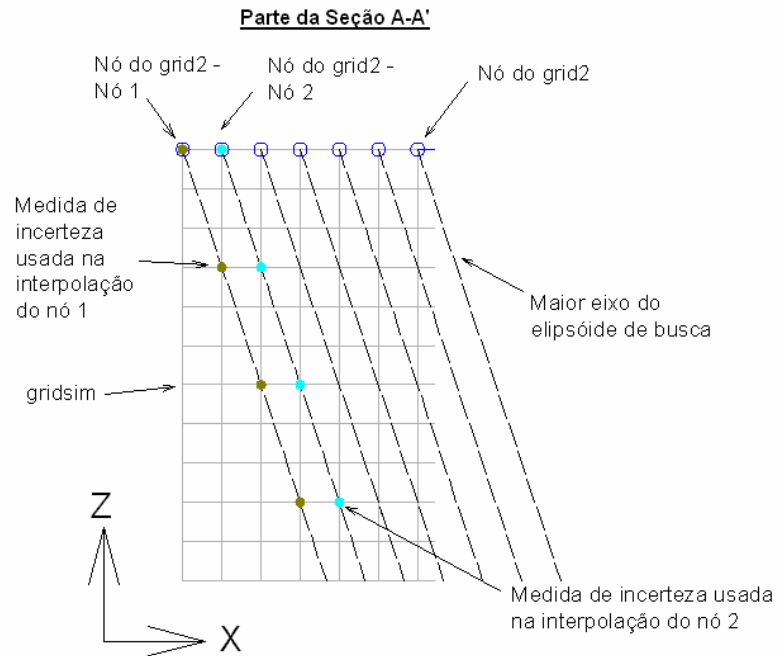


Figura 5.4 – Esquema da busca de dados para interpolação, por Média Móvel, dos valores do índice de incerteza no *grid2*.

Nos casos em que o ângulo entre a sondagem e o plano x-y é pequeno, o *grid2* pode ser criado com um maior número de nós nas direções x e/ou y do que o *gridsim*, visando contemplar pontos simulados mais profundos, conforme mostra Figura 5.5.

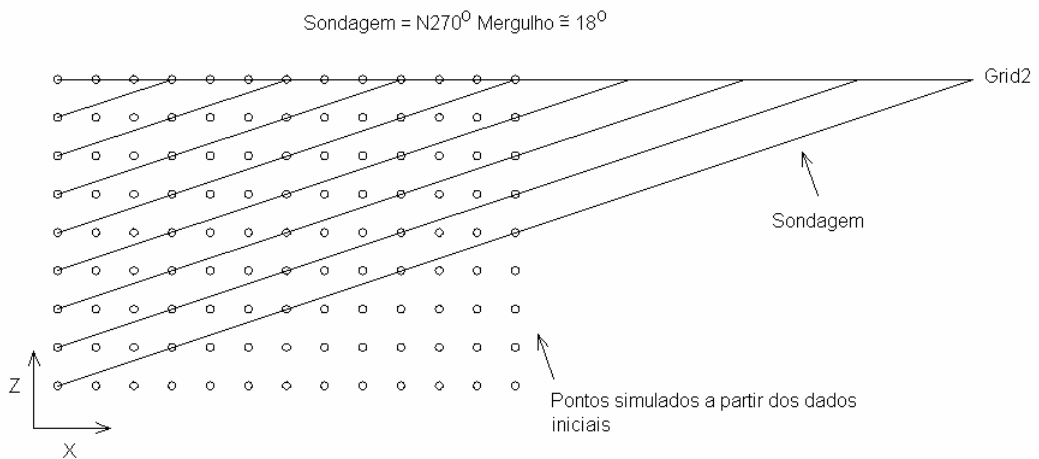


Figura 5.5 - *Grid2* criado com nós além do *gridsim*, na direção x.

vii. Criação do *grid* para determinar espaçamento mínimo (em x e y) entre amostras – Configuração 2

No passo vi, construiu-se o *grid2* e atribuiu-se valores de *média de IEQ* e *indicadores* para cada nó desse *grid*. A ideia da Configuração 2 é adicionar as sondagens nos locais com elevada *média de IEQ* e, ao mesmo tempo, suficiente valor de *média de indicadores*, ou seja, locais que permitam que a sondagem alcance um número suficiente de nós do *gridsim* candidatos a receberem amostras. Sendo assim, bastaria localizar as amostras no *grid2* seguindo esses critérios.

Contudo, o espaçamento entre nós do *grid2*, que corresponde ao mesmo espaçamento usado para simulação dos dados iniciais, pode ser menor que o espaçamento mínimo permitido entre amostras. Ou seja, se duas sondagens são localizadas em nós consecutivos (em uma mesma direção), porque ambos nós apresentam média elevada de IEQ, essas sondagens estariam a uma distância menor que a permitida e poderiam representar informação redundante para simulação, o que significa desperdício de amostras. A fim de localizar as sondagens adicionais no plano x-y, com uma distância apropriada uma da outra, cria-se o *grid3*. Sendo assim, os nós do *grid3* se apresentam como possíveis candidatos a receber a sondagem adicional.

Nas direções x e y, o *grid3* possui nós que coincidem com os nós do *grid2*, porém o espaçamento entre nós é diferente do *grid2*. No *grid3*, o objetivo do espaçamento entre nós é estabelecer uma distância mínima, no plano x-y, entre sondagens adicionais. Na direção z, o *grid3* possui um único nível com elevação igual a do *grid2* (ver Figura 5.6, seguindo exemplo da Figura 5.1).

Como os nós do *grid3* coincidem com os nós do *grid2*, atribuem-se os valores de *média de IEQ* e *média dos indicadores* para os nós do *grid3*. Além disso, assim como foi feito para o *grid1*, determina-se os nós do *grid3* que estão a uma distância menor que a permitida dos dados iniciais. Isto é, conhecendo os valores de *média de indicadores* em cada nó e os nós que estão muito próximos dos dados iniciais, definem-se os nós do *grid3* que são candidatos a receberem sondagem. Por fim, adicionam-se as novas sondagens nos locais candidatos e com maiores médias de IEQ.

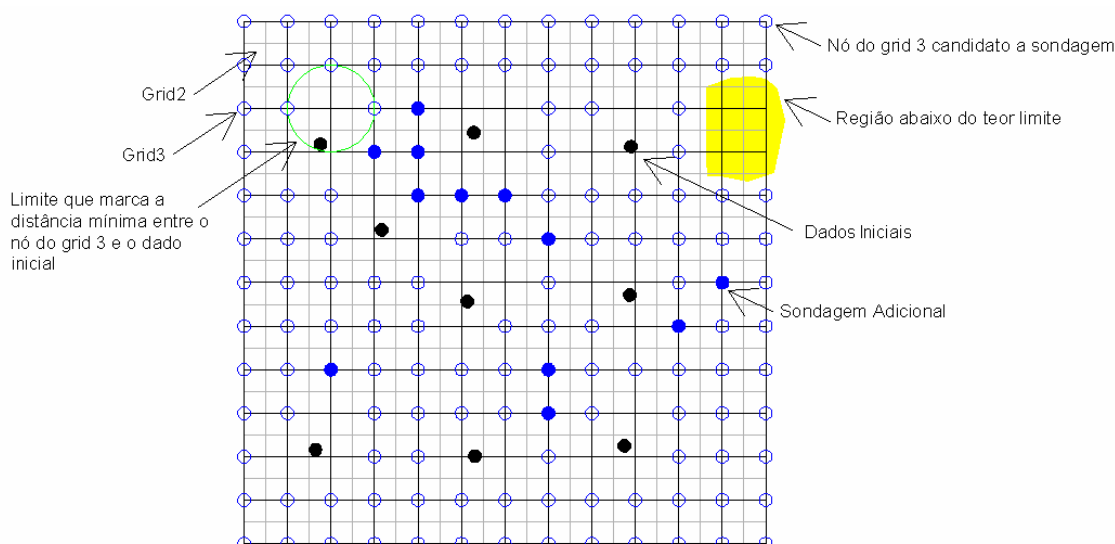


Figura 5.6 – Exemplo da construção do *grid3* (criado para locar as sondagens que devem ser adicionadas nas regiões de elevada incerteza) e escolha dos nós que devem receber sondagem.

As funções necessárias para realizar, computacionalmente, a escolha dos nós do *grid3* que devem receber sondagem são detalhadas no Anexo A – A.3.

viii. Locação de amostras (3D) e adição dessas amostras aos dados iniciais - Configuração 2

Esse passo é realizado da mesma forma que o passo iv (feito para construção da Configuração 1).

5.2 - Automação da Construção das Configurações e Simulação – *Journal File* (Isatis[©])

O *Journal File* é um arquivo que, quando executado no software Isatis[©], realiza uma sequência de operações. O *Journal File* criado nessa tese será chamado *Samples3D*. As operações compreendidas em *Samples3D* envolvem simulação com os dados iniciais, execução do algoritmo de construção de configurações (seção 5.1), simulação com cada configuração de amostras e construção de modelos de blocos para os cenários gerados. Ou seja, a execução de

Samples3D, permite a construção automática de configurações de amostras e geração de diversos cenários, para cada configuração. Além disso, o *Samples3D* compreende operações de exportação dos modelos de blocos para arquivos de texto. Esses arquivos de texto são criados para serem utilizados por uma rotina computacional que não pertence ao Isatis[®]. Essa rotina computacional foi criada na linguagem *Visual Basic*[®] (Microsoft Press, 1999) e calcula o valor da função de transferência para cada modelo de bloco gerado.

Antes da execução de *Samples3D*, três arquivos devem ser importados para o Isatis[®]. O arquivo com os dados para a variável de interesse, o arquivo COLLAR (descrito anteriormente) e o arquivo com a seqüência dos blocos de minério que serão lavrados. A seqüência dos blocos de minério que serão lavrados pode ter sido obtida a partir de um modelo interpolado usando os dados iniciais (e.g. interpolação por krigagem) ou, até mesmo, baseado em cenários previamente simulados usando os dados iniciais.

A seqüência de mineração dos blocos será exportada junto com os modelos de blocos para ser usada no cálculo da função de transferência (e.g. VPL), por isso, o arquivo com a seqüência de blocos que serão lavrados deve conter as coordenadas x,y,z de cada bloco e o número do bloco na seqüência de mineração. Ou seja, se o bloco for o quarto bloco a ser lavrado o número do bloco será 4. A contagem dos blocos na seqüência de mineração deve ser de 1 a *B*, sendo *B* o número total de blocos lavrados. Além disso, um sinal negativo (-) pode ser adicionado antes do número do bloco para aqueles blocos que são lavrados, mas considerados estéreis, independente do valor do atributo de interesse. Por exemplo, em um depósito de cobre, o valor econômico do bloco para o cálculo do VPL depende do teor de cobre, contudo os blocos que pertencem a uma região que possui alto teor de um contaminante, apesar de pertencerem a cava final, devem ser considerados estéreis. Outro exemplo, em um depósito de carvão, a tonelagem de cada bloco é o atributo de interesse para o cálculo do VPL, contudo, os blocos que estão em uma região de alto teor de enxofre, apesar de pertencerem a cava, não podem ser comercializados e em um cenário extremo serão considerados estéreis.

Os três arquivos devem ser importados para um estudo novo (um projeto dentro do software sem dados antigos) onde será executado o *Samples3D*. Recomenda-se a criação de um estudo novo porque outros arquivos são gerados

durante a execução do *Samples3D* e a utilização de um estudo novo evita que arquivos antigos sejam sobrescritos.

Além disso, antes da execução do *Samples3D*, a distribuição original deve ser transformada em uma distribuição gaussiana, essa operação deve ser realizada dentro do Isatis[®], para que a variável gaussiana e o arquivo de anamorfose (ver Capítulo 2), que são necessários para a simulação e retro-transformação da variável gaussiana, possam ser acessados durante a execução do *Samples3D*. O modelo de continuidade para os dados normalizados também deve ser inserido no software.

Os arquivos, variáveis e modelo, que devem ser criados no software pelo usuário, são reconhecidos pelo *Samples3D* por nomes previamente definidos, assim como, o diretório que contém esses arquivos. Por isso, o usuário deve identificá-los com os seguintes nomes:

- Arquivo com os dados para a variável de interesse:

Diretório: *data*

Arquivo: *dataset*

Variável no domínio original: *orig*

Variável gaussiana: *gaus*

- Arquivo COLLAR:

Diretório: *data*

Arquivo: *collar*

Coordenada X: *x*

-Seqüência dos blocos de minério que serão lavrados:

Diretório: *data*

Arquivo: *sequence*

Variável com a ordem de mineração dos blocos: *seq*

Obs: As coordenadas de cada dado correspondem ao centro do respectivo bloco.

- Arquivos de Parâmetros:

Anamorfose: *anamdat* (*variável que foi normalizada: orig*)

Modelo de continuidade: *modnorm*

Obs: O arquivo de anamorfose e o modelo de continuidade para os dados gaussianos são salvos como arquivo de parâmetros (*Parameter File*) dentro do software.

Quando o *Samples3D* é executado, uma tabela (Figura 5.7) aparece na tela do computador com campos que podem ser preenchidos pelo usuário. Esses campos correspondem a parâmetros necessários em algumas das operações realizadas. Os rótulos desses campos estão escritos na língua inglesa.

Questions

Number Questions in Selected Bulletins: 87

1	X Origin (gridsim)	N	541 m	Default
2	Y Origin (gridsim)	N	1471 m	Default
3	Z Origin (gridsim)	N	1247 m	Default
4	X Mesh (gridsim)	N	2 m	Default
5	Y Mesh (gridsim)	N	2 m	Default
6	Z Mesh (gridsim)	N	1 m	
7	X Nodes (gridsim)	N	140	
8	Y Nodes (gridsim)	N	70	
9	Z Nodes (gridsim)	N	1	
10	X Origin (gridbl)	N	541 m	
11	Y Origin (gridbl)	N	1471 m	
12	Z Origin (gridbl)	N	1247 m	
13	X Mesh (gridbl)	N	2 m	
14	Y Mesh (gridbl)	N	2 m	
15	Z Mesh (gridbl)	N	1 m	
16	X Nodes (gridbl)	N	140	
17	Y Nodes (gridbl)	N	70	
18	Z Nodes (gridbl)	N	1	
19	Number of Bands	N	1000	
20	Number of Real.	N	1	
21	Rot. around Z (Search)	A	-45 degrees	
22	Rot. around Y (Search)	A	0 degrees	
23	Rot. around X (Search)	A	0 degrees	
24	Max. Dist. X (Search)	N	200 m	
25	Max. Dist. Y (Search)	N	100 m	
26	Max. Dist. Z (Search)	N	50 m	
27	Ang. Sectors (Search)	N	8	
28	Min. Data (Search)	N	3	
29	Data Sector (Search)	N	2	
30	Points inside block (a)	N	199.5	
31	Points inside block (b)	N	200.5	
32	Cutoff (a) (Index)	N	1	
33	Cutoff (b) (Index)	N	1	
34	X Origin (grid1)	N	541 m	
35	Y Origin (grid1)	N	1471 m	
36	Z Origin (grid1)	N	1247 m	
37	X Mesh (grid1)	N	2 m	
38	Y Mesh (grid1)	N	2 m	
39	Z Mesh (grid1)	N	1 m	
40	X Nodes (grid1)	N	140	
41	Y Nodes (grid1)	N	70	
42	Z Nodes (grid1)	N	1	
43	Min. Dist. X (data)	N	10 m	
44	Min. Dist. Y (data)	N	10 m	
45	Min. X (collar)	N	-10	
46	Max. X (collar)	N	5000	
47	Replace 11 for adic. data	M	v2=ifelse(v1>11,0,1)	
48	Rot. around Z (hole)	A	0 degrees	
49	Rot. around Y (hole)	A	-45 degrees	
50	Rot. around X (hole)	A	0 degrees	
51	Max. Dist. X (hole)	N	2000 m	
52	Max. Dist. Y (hole)	N	0.001 m	
53	Max. Dist. Z (hole)	N	0.001 m	
54	All Data (hole)	N	10	
55	Replace 01 for Refer.	A	v2 =simdat[00001]	
56	X Origin (grid2)	N	541 m	
57	Y Origin (grid2)	N	1471 m	
58	Z Origin (grid2)	N	1247 m	
59	X Mesh (grid2)	N	2 m	
60	Y Mesh (grid2)	N	2 m	
61	Z Mesh (grid2)	N	1 m	
62	X Nodes (grid2)	N	140	
63	Y Nodes (grid2)	N	70	
64	Z Nodes (grid2)	N	1	
65	X Origin (grid3)	N	545 m	
66	Y Origin (grid3)	N	1475 m	
67	Z Origin (grid3)	N	1247 m	
68	X Mesh (grid3)	N	10 m	
69	Y Mesh (grid3)	N	10 m	
70	Z Mesh (grid3)	N	1 m	
71	X Nodes (grid3)	N	28	
72	Y Nodes (grid3)	N	14	
73	Z Nodes (grid3)	N	1	
74	Min. X (collar)	N	-10	
75	Max. X (collar)	N	5000	
76	Prop. Limit	N	0.05	
77	Replace 9 for adic. data	M	v2=ifelse(v1>9,0,1)	
78	Replace 01 for Refer.	A	v2 =simdat[00001]	
79	Number of Bands	N	1000	
80	Number of Real.	N	1	
81	Number of Bands	N	1000	
82	Number of Real.	N	1	
83	Max. Sequence	N	200	
84	Output (expl)	A	E:\backup\save.txt	
85	Output (conf1)	A	E:\backup\save.txt	
82	Number of Real.	N	1	
83	Max. Sequence	N	200	
84	Output (expl)	A	E:\backup\save.txt	
85	Output (conf1)	A	E:\backup\save.txt	
86	Output (conf2)	A	E:\backup\save.txt	

Figura 5.7 – Tabela com campos do *Journal File* que podem ser preenchidos pelo usuário.

A sequência de operações realizadas pelo *Samples3D* e os parâmetros necessários em cada operação são descritos no Anexo A – A.4.

5.3 – Rotina Computacional para o Cálculo das Funções de Transferência

O objetivo do desenvolvimento de um algoritmo de construção de configurações amostrais é permitir que cenários de disposição espacial de amostras adicionais possam ser construídos e avaliados de forma automática e rápida. Pode-se usar esta ferramenta para o processo de tomada de decisão sobre a locação de amostras adicionais que proporcionem o máximo benefício. Assim, agiliza-se o processo de análise do desempenho dessas configurações em relação aos objetivos da nova campanha de amostragem. O objetivo da amostragem, abordado ao longo da tese, é a redução da incerteza sobre o valor de uma função de transferência e a incerteza é acessada pela geração de cenários do atributo de interesse, para cada configuração construída.

A execução do *Journal File* “*Samples3D*” realiza automaticamente a execução do algoritmo de construção de configurações e a geração de cenários do atributo de interesse, para cada configuração de amostras construída. Contudo, o cálculo da função de transferência é feito fora do Isatis[®]. Uma rotina computacional foi desenvolvida para agilizar o cálculo da função VPL para cada um dos cenários, ou modelos de blocos, gerados. Com a distribuição dos valores de VPL gerados para uma mesma configuração de amostras, a rotina desenvolvida acessa a incerteza sobre essa distribuição.

A Figura 5.8 apresenta um formulário com campos que devem ser preenchidos pelo usuário antes da execução da rotina. Esses campos referem-se aos parâmetros que são usados no cálculo do VPL para cada modelo de blocos. Os rótulos desses campos estão escritos na língua inglesa.

The screenshot shows the NPV software interface with the following sections and values:

- Input Data:**
 - File with blocks: (File in the folder of program)
 - Format of File (no header): Columns - X,Y,Z,Real1,...,RealN,Sequence (N is the number of realizations)
 - Number of Columns: (The number of columns is (N+4))
 - Check the box if the variable is grade (%)
- If Grade:**
 - Recovery: Price for Sell: Refining Cost:
 - Mining Cost: Processing Cost: Block Mass:
- If Tonnage per Unit Area:**
 - Mining Cost (ton): Profit (ton): Block Area:
- Mining Sequence:**
 - Investment = + (×)
 - Rate for the Period (decimal): Mined per Period:
 - Number of blocks if variable is grade:
 - Tonnage otherwise (tonnage slightly higher will be considered):
- Statistics:**
 - E[VPL]: Standard Deviation: Column:
 - Minimum: Maximum: VPL for Column:

Buttons: Calculate VPL, Exit

Figura 5.8 – Formulário que deve ser preenchido para cálculo do VPL pela rotina desenvolvida.

O usuário informa, no formulário, o arquivo de texto que contém os modelos de blocos para uma das configurações testadas. O arquivo de texto não deve conter cabeçalho. As três primeiras colunas do arquivo de texto podem ser usadas para as coordenadas do centro de cada bloco, embora os valores das três primeiras colunas não influenciem no cálculo do VPL. Depois, o arquivo deve conter L colunas, cada uma com o valor do bloco para o respectivo cenário. A última coluna deve conter o valor da seqüência de mineração do bloco. A execução do *Samples3D* gera arquivos de texto com esse formato. A Figura 5.9 apresenta um esquema do arquivo de texto que deve ser informado no formulário.

X	Y	Z	Cenário 1	Cenário 2	Cenário 3	Sequência de Lavra	(sinal negativo indica que o bloco é estéril)
1	1	1	10	11	9	1	
2	1	1	12	10	11	4	
1	2	1	8	9	10	-3	
2	2	1	9	11	10	2	

← Arquivo

Figura 5.9 – Esquema do arquivo de texto que deve ser informado no formulário.

Os campos do formulário referentes ao arquivo de texto que deve ser usado são listados a seguir:

- File with blocks – Nome e extensão do arquivo de texto que contém os modelos de blocos;
- Number of Columns – Número de colunas no arquivo de texto.

O cálculo do VPL pode ser feito de duas formas, a forma que será seguida depende do atributo que é considerado no cálculo do VPL. Se o atributo for teor de um metal (%), o valor econômico de cada bloco que será lavrado é calculado pela equação 4.1, descrita no Capítulo 4. Além disso, se o valor da seqüência de mineração do bloco possui um sinal negativo (-) na frente, o bloco é considerado estéril e seu valor econômico é o produto do custo de lavra da unidade de massa pela massa do bloco e pelo valor (-1). O valor econômico dos blocos que não foram simulados (blocos não simulados possuem valor -999 no arquivo de texto exportado pelo *Samples3D*) é igual a 0. Os campos do formulário referentes ao cálculo do valor econômico quando o atributo é teor (%) são:

- Check the box if the variable is grade (%) – Marcar a caixa ao lado para informar que o atributo é teor (%);
- Recovery – Recuperação do metal (recup, conforme equação 4.1);
- Price for Sell – Preço de venda (pv);

- Refining Cost – Custo do refino (cr);
- Mining Cost – Custo de lavra (cl);
- Processing Cost – Custo do beneficiamento (cb);
- Block Mass – Massa do bloco (tb).

Quando o atributo é o teor de um metal (%), o cálculo do VPL considera que um mesmo número de blocos é minerado em cada período de tempo. O usuário informa no formulário o número de blocos minerados em cada período, que pode ser um mês, um ano, etc. Além disso, o usuário deve informar a taxa de juros para aquele período. A rotina computacional procura os blocos que devem ser minerados no mesmo período, de acordo com a seqüência de mineração dos blocos, e calcula o valor presente para aquele período. Depois, os valores presentes para todos períodos são somados e o valor do investimento (informado pelo usuário) é deduzido dessa soma. Assim, o resultado é o VPL, cuja fórmula para o cálculo foi descrita no Capítulo 2. Os campos do formulário referentes ao cálculo do VPL quando o atributo é teor (%) são:

- Number of blocks if variable is grade – Número de blocos minerados em um período.

A Figura 5.10 apresenta um exemplo de como é calculado o VPL para o arquivo apresentado na Figura 5.9.

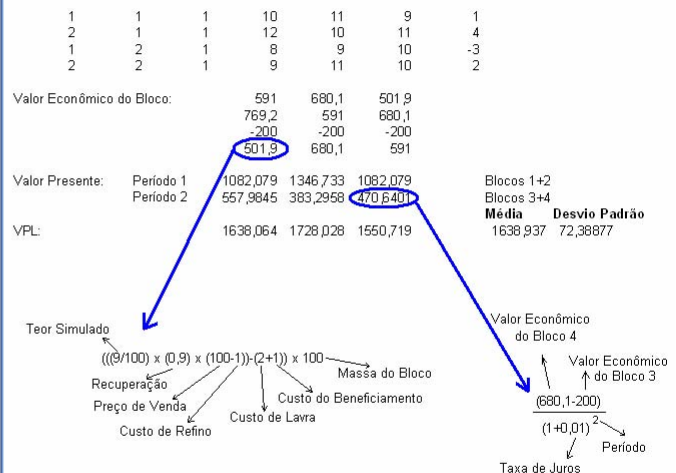
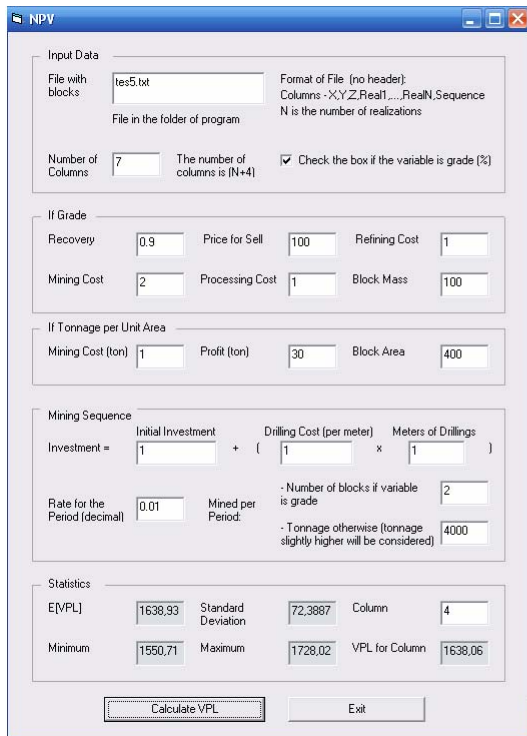


Figura 5.10 – Exemplo de cálculo do VPL pela rotina desenvolvida.

Considere, agora, que o objetivo da simulação foi obter a tonelagem de cada bloco, em cada cenário. O atributo simulado pode ter sido tonelagem por unidade de área. Os dados usados na simulação podem ter sido obtidos com o produto da densidade pela espessura de cada amostra da sondagem. Ou seja, esses dados representam tonelagem pela unidade de área, e todos valores representam um mesmo suporte, visto que as amostras foram obtidas em furos de sonda com mesmo diâmetro da seção transversal. Se o atributo simulado é tonelagem por unidade de área (e.g. t/m², quando a área é metro (m²)), o valor econômico de cada bloco que será lavrado é calculado de acordo com a Equação 5.2. O bloco que não foi simulado para esse atributo recebe valor econômico 0 e tonelagem igual a 0.

$$Econ(u) = \begin{cases} \left[-\left(\frac{t}{a^2}(u) \right) \cdot (ablock) \cdot (cl(ton)) \right] & \text{se } seqbl(u) < 0 \\ \left[\left(\frac{t}{a^2}(u) \right) \cdot (ablock) \cdot (prof(ton)) \right] & \text{outros casos} \end{cases} \quad \text{(Equação 5.2)}$$

onde:

$Econ(u)$ = valor econômico do bloco u ;

$t/a^2(u)$ = valor simulado para o bloco u ;

$ablock$ = área do bloco, todos blocos possuem mesma área;

$cl(ton)$ = custo de lavra por tonelada;

$prof(ton)$ = lucro (receita - custo) com a venda de uma tonelada;

$seqbl$ = número da seqüência de mineração do bloco u (quando possui um sinal negativo na frente o bloco é estéril).

Os campos do formulário referentes ao cálculo do valor econômico quando o atributo é tonelagem por unidade de área são:

- Mining Cost (ton) – Custo de lavra por tonelada ($cl(ton)$, na equação 5.2);
- Profit (ton) – Lucro com a venda de uma tonelada ($prof(ton)$);
- Block Area – Área do bloco ($ablock$).

Quando o atributo é tonelagem por unidade de área, o cálculo do VPL considera que a tonelagem minerada em cada período de tempo é semelhante de um período para o outro. O usuário informa no formulário a tonelagem aproximada que deve ser minerada em um período. A rotina computacional procura os blocos que devem ser minerados em um mesmo período, de acordo com a seqüência de mineração dos blocos, e com a tonelagem que deve ser minerada. A seqüência de mineração dos blocos vai sendo seguida e as massas dos blocos vão sendo somadas, assim que a soma passa a ser maior que a tonelagem indicada pelo usuário, a rotina calcula o valor presente para aqueles blocos. Pode acontecer que os blocos minerados no último período somem uma tonelagem menor do que a indicada pelo usuário, mesmo assim, o valor presente é calculado para essa tonelagem menor. Os campos do formulário referentes ao cálculo do VPL quando o atributo é tonelagem por unidade de área são:

- Tonnage otherwise (tonnage slightly higher will be considered) – tonelagem minerada em um período. Uma tonelagem um pouco maior será considerada em cada período, contudo, uma tonelagem menor pode ser considerada no último.

Os seguintes campos são usados para ambos tipos de atributos:

- Investment – Investimento. O investimento é a soma do investimento inicial, definido no campo “Initial Investment”, com o investimento em sondagem. O investimento em sondagem é o custo de um metro de sondagem (“Drilling Cost (per meter)”) multiplicado pelo número de metros de sondagens realizados (“Meters of Drillings”);
- Rate for the Period (decimal) – Taxa de juros para o período (em decimal).

Os resultados exibidos pela rotina são:

- E[VPL] – Valor esperado da função ou média do VPL para os diversos cenários;
- Standard Deviation – Desvio padrão para os diversos cenários;
- Minimum – Menor valor de VPL para os cenários;
- Maximum – Maior valor de VPL para os cenários;
- VPL for Column – VPL para o cenário que está na coluna definida em “Column”.

5.4 - Análise Sobre o Algoritmo para Locação de Amostras

Esse capítulo apresentou um algoritmo para facilitar a construção de configurações de amostras e uma rotina para calcular a incerteza sobre a função VPL. Algumas vantagens do algoritmo podem ser destacadas:

- A forma como o algoritmo constrói a configuração 2 permite que o usuário escolha o nível de agrupamento das sondagens adicionais nas regiões de elevada incerteza sobre o atributo de interesse. Por exemplo, o usuário pode querer construir uma configuração de amostras em que as sondagens sejam próximas ou pertençam a regiões de elevada incerteza, porém o agrupamento entre as sondagens seja pequeno de forma que essas contemplem uma maior área do depósito. Para isso, o usuário aumenta o espaçamento entre nós do *grid3*. Ou seja, o usuário pode testar uma configuração 2 mais ou menos parecida com a configuração 1, aumentando ou diminuindo o espaçamento entre nós do *grid3*, respectivamente;
- A forma com que o algoritmo adiciona sondagens nos locais de elevada incerteza sobre o valor do atributo permite que as novas sondagens sejam adicionais em etapas. Conforme comentado no Capítulo 3, quando muitas sondagens devem ser adicionadas aos dados iniciais essas sondagens devem ser adicionadas em etapas. O algoritmo pode ser modificado de forma a construir a configuração 2 duas ou três vezes. Por exemplo, o número de sondagens adicionais (n) é dividido em 3 partes. O algoritmo constrói a configuração 2 considerando $n/3$ sondagens adicionais. Depois, o algoritmo considera os $n/3$ dados adicionados como dados iniciais e constrói a configuração 2, novamente, com mais $n/3$ dados adicionais, e assim por diante. O *grid2* e o *grid3* são iguais para todas vezes que o algoritmo constrói a configuração 2, a diferença é que durante esse processo aumenta o número de nós do *grid3* que devem receber sondagem até que o número total de novas sondagens esteja locado nesse *grid*;
- O algoritmo pode ser aplicado nos casos em que o usuário deseja usar um *grid* girado em relação ao plano horizontal, para a simulação a partir dos dados iniciais. Ou seja, os planos que contém os nós do *gridsim*, em cada dimensão, são perpendiculares entre si, mas não concordam com o plano horizontal. Nesse caso, o *grid1*, o *grid2* e o *grid3* devem ser definidos na mesma direção de um desses planos, seguindo o raciocínio necessário para o funcionamento do algoritmo. Convém lembrar que, após a execução do algoritmo e escolha de uma das configurações de amostras para ser adotada na realidade, o usuário deve calcular a posição do início da sondagem na superfície do terreno. Para isso basta estender a sondagem, mantendo a inclinação, até a superfície do terreno.

Quando o terreno tem um desnível acentuado, pode ser conveniente excluir os valores simulados para os nós do *gridsim* que pertencem à região já erodida. A exclusão desses valores não permite que influenciem no cálculo do valor médio do índice de incerteza. Isso evita que zonas de alta incerteza que não existem mais sejam contempladas com sondagem adicional. Além disso, a exclusão desses valores não permite que influenciem na função de transferência. O algoritmo pode ser adaptado para esses casos. Após a execução da simulação, tanto a partir dos dados iniciais como para os dados das configurações 1 e 2, deve-se executar uma função para atribuir um símbolo de valor inexistente aos nós do *gridsim* que pertencem à região erodida.

Existem casos em que as sondagens, tanto para a Configuração 1 como para Configuração 2, devem ser adicionadas em, apenas, um painel de mineração (amostragem de curto prazo). O algoritmo apresentado no Capítulo 5 permite que as Configurações 1 e 2 adicionem amostras, apenas, nesse painel. Para isso, o usuário cria os *grids* 1 e 3 de forma que seus nós não pertençam a locais além dos limites do painel.

Na metodologia adotada na tese as informações das sondagens adicionais são obtidas a partir de um cenário simulado (simulado considerando os dados iniciais), escolhido como referência. Quando o espaçamento entre os nós do *grid* onde foi simulado o cenário de referência é grande na direção da sondagem, o espaço entre os dados coletados pelas novas sondagens também é grande. Assim, em alguns casos, o número de dados coletados pela sondagem adicional pode ser muito menor do que o coletado pelas sondagens iniciais, principalmente quando a sondagem adicional é inclinada. Nesses casos, o número de dados coletado pela sondagem fictícia não representa o número de dados que seriam coletados na realidade. Por isso, o espaçamento entre os nós do *grid* onde o cenário de referência é simulado deve ser definido tentando amenizar essa limitação da metodologia.

Na análise do desempenho das configurações, as etapas de simulação geoestatística e cálculo da média do índice de incerteza na direção da sondagem (passo vi do algoritmo de construção de configurações), provavelmente, demandam o maior tempo de processamento. Alguns fatores influenciam no tempo de processamento para essas etapas, como: número de locais simulados,

características do computador, algoritmo de simulação, entre outros. As Tabelas 5.1 e 5.2 apresentam exemplos de tempos de processamento necessários para essas etapas, quando realizadas sobre determinadas condições.

Tabela 5.1 – Tempo de processamento para o cálculo da média do índice de incerteza na direção da sondagem

Características do Computador	Espaço entre Nós/ Número de Nós						Inclinação da Sondagem	Tempo
	X Gridsim	Y Gridsim	Z Gridsim	X Grid2	Y Grid2	Z Grid2		
Windows XP Professional (2002) (32bits) Intel® Core™ 2 Duo Processador T7300	2 m/ 140	2 m/ 70	2 m/ 54	2 m/ 140	2 m/ 70	2 m/ 1	Azimuth 0°/ Mergulho 45°	~ 1 min 9 seg
Processador - 2 GHz, 780 MHz RAM - 1,99 Gb								
Windows XP Professional (2002) (32bits) Intel® Core™ 2 Duo Processador T7300	2 m/ 140	2 m/ 70	2 m/ 54	2 m/ 140	2 m/ 70	2 m/ 1	Azimuth 0°/ Mergulho 90°	~ 2 seg
Processador - 2 GHz, 780 MHz RAM - 1,99 Gb								

Obs: Nós do Grid2 coincidem com nós do Gridsim.

Interpolações por Média Móvel realizadas no software Isatis® v8.04.

Tabela 5.2 – Tempo de processamento para a simulação geoestatística

Banco de Dados	Características do Computador	Número de Nós Simulados (por 1 realização)/ Número de Nós do Grid	Número de Cenários	Dados Usados na Simulação	Raio de Busca de Dados – Maior Continuidade/ Continuidade Intermediária/ Menor Continuidade	Tempo
Walker Lake	Windows XP Professional (2002) (32bits) Intel® Core™ 2 Duo Processor T7300 Processador - 2 GHz, 780 MHz RAM - 1,99 Gb	76.816/ 78.000	90	99	150 / 36	~ 2 min 8 seg
Cobre	Windows XP Professional (2002) (32bits) Intel® Core™ 2 Duo Processor T7300 Processador - 2 GHz, 780 MHz RAM - 1,99 Gb	529.159/ 529.200	70	735	200 / 110 / 65	~ 15 min 14 seg
Cobre	Windows XP Professional (2002) (64bits) Intel® Core™ 2 Duo CPU E6850 Processador - 3 GHz, 2,99 GHz RAM - 3 Gb	529.162/ 529.200	70	778	200 / 110 / 65	~ 12 min 9 seg
Cobre	Windows XP Professional (2002) (64bits) Intel® Core™ 2 Duo CPU E6850 Processador - 3 GHz, 2,99 GHz RAM - 3 Gb	23.042.799/ 24.000.000	1	1392	200 / 110 / 65	~ 28 min 55 seg

Obs: Processos realizados no software Isatis® v8.04.
Todas simulações geradas por Bandas Rotativas.

Capítulo 6 – Conclusões e Trabalhos Futuros

Nesse capítulo, são lembrados as metas e objetivos traçados para esta tese e as conclusões relacionadas a cada item. A seguir, também são apresentadas idéias de trabalhos futuros que podem ser realizados no tema da tese.

6.1 - Conclusões

Uma das metas dessa tese consistia em analisar a eficiência de duas configurações de amostras quanto à redução da incerteza sobre uma função de transferência. Ambas configurações possuíam, em comum, os dados iniciais e o número de sondagens (amostras) adicionais. A diferença entre elas era a posição das sondagens adicionais. Na chamada Configuração 1, as sondagens foram adicionadas de forma a comporem uma malha aproximadamente regular com os dados iniciais. Na chamada Configuração 2, as sondagens foram adicionadas nas regiões de elevada incerteza sobre o atributo de interesse. Essa análise foi ilustrada por dois estudos: um considerando a variável V (Capítulo 3) e o outro considerando o teor de cobre (Capítulo 4).

Ao longo do estudo, uma medida de erro relativo ($[\pm(2\sigma/fr)]$, onde σ é o desvio padrão sobre os valores da função e fr é o valor da função para o cenário de referência) foi adotada para avaliar o que significa economicamente a incerteza obtida com cada configuração e avaliar a importância da diferença entre as configurações em relação à ordem de grandeza do valor da função. A diferença entre as configurações mostrou ser mais significativa economicamente para o caso em que a variável foi o teor de cobre. Contudo, se os próprios dados iniciais já fornecem uma incerteza pequena em relação ao valor da função para a referência, provavelmente, a diferença entre as configurações também será pequena. A

incerteza fornecida pelos dados iniciais em relação ao valor da referência, depende, além do número de dados iniciais, da variabilidade do atributo. Os dados iniciais para o depósito de cobre possuem alta variabilidade (elevado coeficiente de variação).

De acordo com os resultados obtidos para os dois estudos, a configuração mais eficiente na redução da incerteza pode variar em cada caso. Um dos fatores que influenciam na melhor eficiência de uma configuração é a distribuição dos valores do índice de incerteza. Pode-se dizer, que quanto mais assimétrica é a distribuição desses valores, maior é a chance da Configuração 2 apresentar um melhor desempenho. A distribuição dos valores dos índices de incerteza, por sua vez, depende do número de dados iniciais. Essa distribuição pode se tornar mais assimétrica com o aumento dos dados.

Considere que a Configuração 2 possui melhor desempenho, observando os fatores que influenciam os desempenhos das configurações, pode se dizer que dois fatores podem influenciar no aumento do desempenho da Configuração 2 em relação a Configuração 1: a variabilidade do atributo e o espalhamento das regiões com elevada incerteza sobre o valor do atributo:

- Variabilidade do atributo – Quanto maior a variabilidade do atributo, provavelmente, a distribuição dos valores desse atributo será mais assimétrica, assim como, a distribuição dos valores do índice incerteza relacionado a esse atributo. Nesse caso, a redução de um valor extremo do índice de incerteza, com a adição de uma amostra próxima do local desse valor, pode contribuir muito mais na redução da incerteza sobre a função do que a redução de valores medianos do índice de incerteza, contemplados por uma malha de amostragem regular. Assim, a eficiência da Configuração 2 em relação a 1 pode aumentar com o aumento da variabilidade do atributo.
- Espalhamento das regiões com elevada incerteza – Quando as regiões de elevada incerteza são mais dispersas ao longo do depósito, a amostragem regular tem mais chance de contemplar essas regiões, aumentando sua eficiência na redução da incerteza. Assim, quando os valores extremos de índice de incerteza são mais agrupados, a amostragem regular tem menos chance de contemplar esses locais, que contribuem na maior porção da incerteza sobre a função, e a eficiência

da Configuração 2 em relação a 1 pode aumentar.

Convém lembrar, que o atributo de interesse depende de qual função de transferência está sendo avaliada. Se o objetivo do estudo para o depósito de cobre (Capítulo 4) fosse reduzir a incerteza sobre uma outra função, que se baseia na distribuição de valores de um atributo com menor variabilidade do que o teor de cobre, a diferença entre as configurações poderia ser pouco significativa economicamente. Aliás, considerando um outro atributo, as regiões de maior incerteza poderiam ser diferentes das encontradas para o teor de cobre e, assim, a locação das amostras adicionais na Configuração 2. Além disso, o índice de incerteza utilizado para localizar as amostras adicionais, na Configuração 2, depende da função de transferência. Por exemplo, se a função de transferência fosse a tonelagem acima de um dado teor de corte, o índice de incerteza deve identificar os locais onde a incerteza sobre o valor do atributo estar acima de um teor de corte é elevada.

A análise das configurações para dois estudos permitiu o conhecimento dos fatores que podem influenciar a diferença entre os desempenhos das configurações. Além disso, os resultados para o estudo em que a variável era o teor de cobre evidenciaram que, às vezes, a Configuração 2 pode trazer um benefício econômico significativamente maior. Ou seja, as duas configurações devem ser testadas para cada caso. Assim, o empreendedor pode avaliar os resultados e escolher entre uma ou outra configuração. Se as duas configurações apresentam desempenho semelhante, o empreendedor pode adotar a Configuração 1, que além de comumente aceita, não representa desperdício de informação, para o caso mencionado, e pode contemplar zonas de valores extremos não descobertas pela campanha inicial de amostragem.

Os objetivos relacionados à meta de analisar os desempenhos das duas configurações de amostras e as conclusões referentes a esses objetivos são comentados a seguir:

- Avaliar o valor de incerteza sobre a função, obtido para cada configuração, utilizando diferentes cenários como fonte de informação: a metodologia adotada nessa tese considera um cenário simulado a partir dos dados iniciais como

referência, ou seja, o cenário simulado fornece os valores para as amostras adicionais. Para uma configuração de amostras, a incerteza sobre a distribuição dos possíveis valores da função de interesse (medida pelo desvio padrão), deve ser semelhante para diferentes cenários escolhidos como referência. No estudo do teor de cobre, que apresentou uma diferença significativa entre os desempenhos das configurações 1 e 2 na redução da incerteza sobre a função, a diferença entre as configurações permaneceu semelhante com a mudança do cenário fonte de informação adicional. Além disso, o valor da incerteza sobre a função, obtido para cada configuração, não sofreu alteração importante com a mudança da referência. Esses resultados comprovaram que a metodologia pode indicar a configuração que apresentará um melhor desempenho, quando adotada na realidade, além disso, pode fornecer uma aproximação da incerteza sobre a função que será obtida após a nova amostragem (utilizando um mesmo algoritmo de simulação, ou ferramenta de acesso à incerteza, após a nova amostragem);

- Realizar a análise dos desempenhos das configurações para os casos bi e tri-dimensional: o estudo considerando a variável V representou a análise para o caso 2D, enquanto o estudo considerando o teor de cobre representou a análise para o caso 3D. O estudo de caso bi-dimensional ilustrou uma ressalva sobre os resultados da metodologia na presença de poucos dados iniciais. Quando existem poucos dados iniciais, o desvio padrão sobre a função, para cada configuração, pode ser sensível aos valores das amostras adicionais. Nesse caso, a metodologia não pode fornecer a redução de incerteza que será encontrada na realidade, pois o valor do desvio padrão varia com o cenário fonte de informação. Porém, apesar da variação do desvio padrão obtido para cada configuração, se a mesma configuração apresenta melhor desempenho para os diferentes cenários, pode-se considerar que essa é a melhor configuração a ser adotada. Contudo, os resultados da metodologia podem ser tão sensíveis ao cenário fonte de informação que não se pode dizer nem qual é a melhor configuração, pois essa varia de acordo com o cenário. Assim, existem casos em que os resultados da metodologia não podem ser considerados. Para se ter certeza que os resultados da metodologia são poucos sensíveis aos valores do cenário fonte de informação, deve-se testar as configurações para alguns cenários bem diferentes, por exemplo, que possuem valores extremos para a função. Dessa forma, pode se ter certeza que os resultados obtidos com a metodologia representam o que aconteceria se as configurações fossem adotadas

na realidade;

- Desenvolver uma técnica para considerar a incerteza ao longo das três dimensões na locação das sondagens: a diferença da aplicação da metodologia para o caso tri-dimensional em relação ao caso bi-dimensional é que mais de um dado, sobre o valor do atributo, é coletado ao longo de cada sondagem. Assim, para a Configuração 2, cada sondagem deve ser locada de forma a coletar um grande número de dados em locais de elevada incerteza. O Capítulo 4 descreve a técnica utilizada para locar as novas sondagens dessa forma e apresenta um exemplo onde o índice de incerteza é o IEQ e um exemplo onde o índice de incerteza é a probabilidade do valor no local estar acima de um teor de corte.

Outra meta dessa tese consistia em elaborar um algoritmo que pudesse ser executado em softwares comuns na indústria de mineração e construísse automaticamente as configurações 1 e 2. Além disso, uma rotina computacional deveria ser elaborada para calcular a incerteza sobre a função de transferência, obtida para cada cenário. O Capítulo 5 descreve o algoritmo elaborado, assim como, a rotina computacional. Tanto para a Configuração 1 como para a 2, o algoritmo apresentado constrói um *grid* com apenas um nível em Z, para demarcar os locais das novas sondagens. Os nós desses *grids* são os locais onde as novas sondagens podem ser coletadas. O algoritmo permite que o usuário defina uma distância mínima entre as sondagens iniciais e as novas sondagens, os nós do *grid* separados das sondagens iniciais a uma distância menor que a permitida não recebem sondagem adicional.

O algoritmo também é preparado para o caso em que as novas sondagens devam ser inclinadas. Nesse caso, se uma das configurações for adotada na realidade, o usuário deve fazer a correspondência entre o início da sondagem (local x, y e z), demarcado pelo algoritmo, e o início da sondagem na superfície do terreno.

Um dos objetivos relacionados à meta de elaborar um algoritmo era considerar uma distância mínima entre novas sondagens na construção das configurações 1 e 2. O algoritmo descrito no Capítulo 5, adiciona as novas sondagens considerando uma distância mínima entre elas. Essa distância é definida pelo usuário e corresponde ao espaçamento entre os nós do *grid* criado para locar as novas sondagens.

Conforme comentado no Capítulo 1, a idéia dessa tese é que a configuração mais otimizada quanto à redução da incerteza sobre a função é uma das possíveis formas de construir as configurações 1 e 2. O algoritmo apresentado permite que o usuário construa as configurações 1 e 2 de diferentes formas, variando o espaçamento, em cada direção, entre os nós do *grid* criado para locar as amostras.

Ao longo da tese se observou que a melhor configuração de amostras (1 ou 2) para reduzir a incerteza sobre uma função depende de alguns fatores, isso reforça a importância de um algoritmo que agilize a análise do desempenho de cada configuração, em cada caso que o empreendedor necessite planejar uma nova campanha de amostragem.

Sobre a metodologia adotada na tese, é importante lembrar que essa pode indicar, para um determinado número de amostras adicionais, a configuração mais eficiente na redução da incerteza sobre a função e pode indicar o número de amostras que deve ser coletado (para uma determinada configuração) a fim de se obter o nível de incerteza necessário.

6.2 – Trabalhos Futuros

A seguir são apresentadas sugestões para trabalhos futuros relacionados com a pesquisa realizada na tese:

- Estudo em um depósito com reserva maior - O estudo para o teor de cobre mostrou uma diferença significativa entre as configurações 1 e 2 na redução da incerteza sobre cada função analisada. Contudo, o depósito estudado possui uma reserva pequena. Assim, um novo estudo em um depósito com uma reserva maior, e considerando um atributo com alta variabilidade, pode ajudar a mostrar o impacto econômico da diferença entre as configurações;
- Estudo considerando diferentes tipos de minério - O estudo para o teor de cobre considera um tipo de minério. Assim, a construção da configuração 2 poderia ser ilustrada por um depósito com diferentes tipos de minérios ou domínios geológicos, onde cada nova sondagem deve contemplar a incerteza para diferentes atributos;

- Criação de um processo de otimização para construir as Configurações 1 e 2 de diferentes formas - Conforme comentado no Capítulo 5, alguns softwares podem executar, automaticamente, a simulação com os dados iniciais do atributo, o algoritmo de construção das configurações, a simulação com os dados da Configuração 1 e a simulação com os dados da Configuração 2. Arquivos de texto com os modelos de blocos simulados para todas configurações são exportados do software e uma rotina computacional, que não pertence ao software, calcula o desvio padrão da distribuição de VPL, obtida para uma dada configuração. O usuário pode, então, avaliar os desempenhos das Configurações 1 e 2. Contudo, essas configurações podem ser construídas considerando diversos valores para o espaçamento mínimo entre sondagens e a criação de um processo de otimização pode direcionar a construção das configurações de diferentes formas. Além disso, diversos números de sondagens podem ser adicionados às sondagens iniciais. Estudos futuros podem agrupar o algoritmo de construção de configurações, a etapa de simulação com as configurações e a rotina para cálculo do VPL. O número de nós do *grid1* (construído para posicionar as novas sondagens formando uma malha regular com os dados iniciais) e o espaçamento mínimo entre as sondagens podem ser modificados buscando o menor desvio padrão da distribuição de VPL para cada configuração;
- Estudo considerando a incerteza na delimitação dos corpos de minério - Em muitos casos, a melhor delimitação dos corpos de minério pode modificar significativamente a avaliação de recursos. Assim, a locação das sondagens para a Configuração 2, deve considerar, além das regiões de elevada incerteza para o atributo, as regiões de elevada incerteza sobre a existência de um determinado tipo de minério. Por isso, é importante o desenvolvimento de estudos ou um algoritmo de construção de configurações considerando a incerteza sobre o atributo de interesse e sobre a geologia do depósito;
- Estudos mostrando o número de sondagens a partir do qual a redução da incerteza sobre a função não compensa os custos da nova amostragem - Existe um número a partir do qual a adição de amostras produz pequena redução de incerteza, ou seja, a adição de amostras, nesses casos, pode não trazer benefício em relação ao custo da sondagem. É importante que o usuário defina, em cada caso, esse número, isto é, o número máximo de sondagens que devem ser adicionadas. Como exemplo, estudos complementares à tese podem mostrar para o depósito Walker

Lake (Capítulo 3) e para o depósito de cobre (Capítulo 4) o número de sondagens a partir do qual a redução da incerteza sobre a função não compensa os custos da nova campanha de amostragem;

- Estudos considerando os erros de amostragem e análise química nos valores das amostras adicionais - Os erros de amostragem e análise química somados aos valores das amostras adicionais é uma questão que pode ser estudada em trabalhos complementares à tese. Em cada caso estudado na tese, valores do atributo nos locais das amostras adicionais foram obtidos em um cenário de referência para serem considerados na etapa de simulação. O cenário de referência foi considerado o cenário real. Após a simulação, se obteve a distribuição dos possíveis valores para a função. Porém, o valor atribuído a cada amostra coletada na realidade é o resultado do verdadeiro valor da amostra somado aos erros de amostragem e análise química. A fim de avaliar o impacto desses erros nos resultados, ou seja, na distribuição dos possíveis valores para a função, um erro (soma dos erros de amostragem e análise química) pode ser inferido e somado ao valor de cada amostra adicional, antes da simulação. No final, deve se comparar os resultados obtidos considerando esse erro no valor de cada amostra com os resultados que não consideram esse erro;
- Estudo sobre métodos de preencher as sondagens adicionais, nos casos em que poucos dados adicionais são coletados no cenário de referência - Conforme comentado no Capítulo 5, quando as sondagens adicionais não são verticais, essas sondagens podem coletar poucos dados no cenário de referência, ou seja, o número de dados adicionais para as Configurações 1 e 2 não representam o número de dados que seriam coletados na realidade. Assim, trabalhos complementares à tese podem estudar formas de preencher cada sondagem adicional.

Referências Bibliográficas

Aarts, E. e Korst, J. 1989. Simulated Annealing and Boltzmann Machines. John Wiley & Sons, New York, 272p.

Alabert, F. 1987. Stochastic Imaging of Spatial Distributions Using Hard and Soft Information. MSc. Thesis, Stanford University, California.

Andriotti, J. L. S. 2003. Fundamentos de Estatística e Geoestatística. Editora Unisinos, São Leopoldo, Brasil, 165p.

Anon 2000. NPV Scheduler 3.1 Help. Earthworks Corporation Pty. Ltd. Software Help Online.

ANON 2002. GOLDEN SOFTWARE INC. Surfer[®] Version 8 – Help Tutorial. Golden Software Inc., Colorado.

Anon 2003. DATAMINE STUDIO[®] *Version 2.0* – Help Tutorial. Mineral Industries Computing Limited.

Arik, A. 1999. An Alternative Approach to Resource Classification. In: Proceedings of the 28th International Symposium on Computer Applications in the Mineral Industries (APCOM'99). Colorado School of Mines, Golden, Colorado, EUA, pp. 45-53

Armstrong, M. 1998. Basic Linear Geostatistics. Springer, Berlin, 153p.

Atkinson, P.M., Lloyd C. D. 2007. Non-Stationary Variogram Models for Geostatistical Sampling Optimisation: An Empirical Investigation Using Elevation Data. Computers & Geosciences 33, pp. 1285-1300.

AusIMM, 2004. Australasian Code for Reporting of Exploration Results, Mineral

Resources and Ore Reserves (The JORC Code 2004 Edition). Joint Ore Reserves Committee of The Australasian Institute of Mining and Metallurgy, Australasian Institute of Geoscientists and Minerals Council of Australia. Este artigo foi encontrado na internet via http://www.jorc.org/pdf/jorc2004print_v2.pdf. Capturado em 13/11/09.

Bleines, C., Deraisme, J., Geffroy, F., Perseval, S., Rambert, F., Renard, D., Touffait, Y. 2001. ISATIS software manual. Geovariances and École des Mines de Paris, 531p.

Boucher, A., Dimitrakopoulos, R. e Vargas-Guzman, J.A. 2004. Joint Simulations, Optimal Drillhole Spacing and the Role of the Stockpile. Geostatistics Banff 2004, pp. 35-44.

Chilès, J.-P. e Delfiner, P. 1999. Geostatistics: Modeling Spatial Uncertainty. Wiley-Interscience, USA, 695p.

Christakos, G. 1990. A Bayesian/Maximum-Entropy View to the Spatial Estimation Problem. Mathematical Geology, Vol. 22, No. 7, pp. 763-777.

Christmann, R. U. 1978. Estatística Aplicada. Edgard Blücher Ltda, São Paulo, 135p.

David, M. 1977. Geostatistical Ore Reserve Estimation. Elsevier Scientific Publishing Company, Developments in Geomathematics 2, Netherlands, 345p.

Deutsch, C. V. 1992. Annealing Techniques Applied to Reservoir Modeling and The Integration of Geological and Engineering (Well Test) Data. PhD Thesis, Stanford University, USA.

Deutsch, C. V. e Journel, A. G. 1992. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press, New York, 340p.

Deutsch, C. V. e Journel, A. G. 1998. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press, New York, 2^a edição, 369p.

Englund, E.J. e Heravi, N. 1992. Conditional Simulation: Practical Application For Sampling Design Optimization. Geostatistics Troia'92, pp. 613-624.

Emery, X., Lantuéjoul, C. 2006. TBSIM: A Computer Program for Conditional Simulation of Three-Dimensional Gaussian Random Fields via The Turning Bands Method. *Computers & Geosciences* 32, pp. 1615-1628.

Emery, X., Hernández, J., Corvalán, P. e Montaner, D. 2008. Developing a Cost-Effective Sampling Design for Forest Inventory. In: *Geostats 2008*, Santiago. Proceedings of the Eighth International Geostatistics Congress. Gecamin Ltd, Vol. 2, pp. 1001-1010.

Froidevaux, R. 1993. Constrained Kriging as an Estimator of Local Distribution Functions. Capasso V., Girone V., Posa D. *Statistics of Spatial Processes: Theory and Applications*, pp.106-118.

Gambin, F. 2003. Aperfeiçoamento da Estratégia de Homogeneização de Minérios Utilizando Simulação Geoestatística. Dissertação de Mestrado. Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais (PPGEM), Universidade Federal do Rio Grande do Sul, 96p.

Geman, S., Geman D. 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI, Vol. 6, No. 6, pp. 721-741.

Gómez-Hernández, J. 1991. A Stochastic Approach to the Simulation of Block Conductivity Fields Conditioned upon Data Measured at a Smaller Scale. PhD Thesis, Stanford University, USA.

Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, 483p.

Goovaerts, P. 2001. Geostatistical Modelling of Uncertainty in Soil Science. *Geoderma*, Vol. 103, pp. 3-26.

Gotway, C. A. e Rutherford, B. M. 1994. Stochastic Simulation for Imaging Spatial Uncertainty: Comparison and Evaluation of Available Algorithms. In: *Proceedings of Geostatistical Simulations*. Armstrong, M. e Dowd, P. A. (Eds.), Kluwer Academic

Publisher, pp. 1-21.

Hustrulid, W., Kuchta, M. 1995. Open Pit Mine Planning & Design Fundamentals. A. A. Balkema, Rotterdam, Netherlands, Vol. 1, 636p.

Hustrulid, W. e Kuchta, M. 2006a. Open Pit Mine Planning & Design. CSMine Software Package and Orebody Case Example. Taylor & Francis, Londres, Vol.2, 235p.

Hustrulid, W. e Kuchta, M. 2006b. Open Pit Mine Planning & Design. 1. Fundamentals. Taylor & Francis, Londres, 2^a edição, 735p.

Isaaks, E. H. e Srivastava, R. M. 1989. Applied Geostatistics. Oxford University Press, New York, 560p.

Isaaks, E. H. 1990. The Application of Monte Carlo Methods to the Analysis of Spatially Correlated Data. PhD. Thesis, Stanford University, USA, 213p.

Journel, A. G. 1974. Geostatistics for Conditional Simulation of Ore Bodies. Economic Geology, Vol. 69, No. 5, pp. 673-687.

Journel, A. G. 1980. The Lognormal Approach to Predicting Local Distributions of Selective Mining Unit Grades. Mathematical Geology, Vol. 12, No. 4, pp. 285-303.

Journel, A.G. 1982. The Indicator Approach to Estimation of Spatial Distributions. In Proceedings of the 17th APCOM (International Symposium on the Application of Computers and Mathematics in the Mineral Industry), SME-AIME, Golden, Colorado, EUA, pp. 793-806.

Journel, A. G. 1986. Geostatistics: Models and Tools for the Earth Sciences. Mathematical Geology, Vol. 18, No. 1, pp. 119-140.

Journel, A. G. 1988. Geostatistics: Nonparametric Geostatistics for Risk and Additional Sampling Assessment. American Chemical Society, New York, pp. 45-72.

Journel, A. G. e Alabert F. G. 1988. Focusing on Spatial Connectivity of Extreme

Valued Attributes: Stochastic Indicator Models of Reservoir Heterogeneities. SPE, No. 18324.

Journel, A. G. e Huijbregts, Ch. J. 1978. Mining Geostatistics. Academic Press, London, 600p.

Koppe, V. C., Costa, J. F. C. L., Koppe, J. C. 2008a. O Impacto da Configuração de Amostras Sobre a Função de Transferência. In: V Congresso Brasileiro de Mina a Céu Aberto & V Congresso Brasileiro de Mina Subterrânea e Workshop Barragens de Rejeito. CD-Rom, IBRAM.

Koppe, V. C., Costa, J. F. C. L., Koppe, J. C. 2008b. The Benefit of Additional Data Accessed Via Transfer Functions. In: Geostats 2008, Santiago. Proceedings of the Eighth International Geostatistics Congress. Gecamin Ltd, Vol. 2, pp. 1179-1184.

Koppe, V. C., Costa, J. F. C. L., Koppe, J. C. 2009. Algorithm to Help Additional Drilling Location. In: Application of Computers and Operations Research in the Mineral Industry, 2009, Vancouver. CD-Rom, Canadian Institute of Mining, Metallurgy and Petroleum.

Krige, D. G. 1951. A Statistical Approach to Some Mine Valuations and Allied Problems at the Witwatersrand. MSc. Thesis, University of Witwatersrand, Johannesburg, South Africa.

Lantuéjoul, C. 2002. Geostatistical Simulation: Models and Algorithms. Springer, Berlin, 256p.

Lee, T. D. 1984. Planning and Mine Feasibility Study – An Owners Perspective. In: Proceedings of the 1984 NWMA Short Course ‘Mine Feasibility – Concept to Completion’. G. E. McKelvey, compiler, Spokane, WA.

Li, S., Dimitrakopoulos, R., Scott, J. e Dunn, D. 2004. Quantification of Geological Uncertainty and Risk Using Stochastic Simulation and the Applications in Coal Mining Industry. In: Orebody Modelling and Strategic Mine Planning: Uncertainty and Risk Management, 2004, Perth. Proceedings of Orebody Modelling and Strategic Mine

Planning. The Australasian Institute of Mining and Metallurgy, Spectrum Series, Vol. 14, pp. 233-240.

Li, S., Scott, J., Dimitrakopoulos, R. 2005. Stochastic Simulation-Based Risk Assessment For Drilling Optimisation. In: Application of Computers and Operations Research in the Mineral Industry, 2005, Tucson, EUA. Proceedings of the 32nd International Symposium on the Application of Computers and Operations Research in the Mineral Industry. Londres: A. A. Balkema Publishers, 2005, pp. 475-478.

Matheron, G. 1963. Principles of Geostatistics. Economic Geology, No. 58, pp. 1246-1266.

Matheron, G. 1969. Le Krigeage Universel. Le Cahiers du Centre de Morphologie Mathématique de Fontainebleau, Fasc. 1, Ecole des Mines de Paris. Este artigo foi encontrado na internet via http://cg.ensmp.fr/bibliotheque/1969/Matheron/Ouvrage/Doc_00131/Matheron_Ouvrage_00131.pdf. Capturado em 26/04/07.

Matheron, G. 1973. The Intrinsic Random Functions and Their Applications. In Adv. Appl. Prob., Vol. 5, pp. 439-468. Este artigo foi encontrado na internet via <http://cg.ensmp.fr/bibliotheque/index.html>. Capturado em 8/12/06.

Matheron, G. 1982. La Destructuration des Hautes Teneurs et le Krigeage des Indicatrices. Report N-761, CGMM, Fontainebleau. Este artigo foi encontrado na internet via http://cg.ensmp.fr/bibliotheque/1982/Matheron/Rapport/Doc_00231/Matheron_Rapport_00231.pdf. Capturado em 18/04/07.

McBratney, A. B., Webster, R. e Burgess, T. M. 1981. The Design of Optimal Sampling Schemes for Local Estimation and Mapping of Regionalized Variables – I. Theory and Method. Computers & Geosciences, Vol. 7, No. 4, pp. 331-334.

McBratney, A. B., Webster, R. 1981. The Design of Optimal Sampling Schemes for Local Estimation and Mapping of Regionalized Variables – II. Program and Examples. Computers & Geosciences, Vol. 7, No. 4, pp. 335-365.

McBratney, A. B., Whelan, B. M., Walvoort, D. J. J., Minasny, B. 1999. A Purposive Sampling Scheme For Precision Agriculture. In: Precision Agriculture '99 (J. V. Stafford Ed.), Sheffield Academic Press, pp. 101-110.

Microsoft Press, 1999. Microsoft Visual Basic 6.0. Guia do Programador. Makron Books do Brasil Editora Ltda, São Paulo, 921p.

Myers, J. C. 1996. Geostatistical Error Management: Quantifying Uncertainty for Enviromental Sampling and Mapping. Van Nostrand Reinhold, New York, 545p.

Pilger, G. G. 2000. Critérios para Locação Amostral Baseados em Simulação Estocástica. Dissertação de Mestrado. Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais (PPGEM), Universidade Federal do Rio Grande do Sul, 102p.

Pilger, G. G., Costa, J. F. C. L., Koppe, J. C. 2000. Projeto de Amostragem: Uma Proposição Baseada em Simulação Estocástica. REM – Revista da Escola de Minas de Ouro Preto, Ouro Preto, Vol. 53, No. 4, pp. 265-274.

Pilger, G. G., Costa, J. F. C. L., Koppe, J. C. 2002. Optimizing The Value of a Sample. In: Application of Computers and Operations Research in the Mineral Industry, 2002, Phoenix. Proceedings of the 30th International Symposium. Littleton: Society for Mining, Metallurgy and Exploration, Inc. (SME), 2002, Vol. 1, pp. 85-94.

Remy, N., Boucher, A. e Wu, J. 2006. SgeMS User's Guide. 128p.

Richmond, A., Waltho, A., Bateman, J. e Fergusson, D. 2007. Locating Drill Holes to Reduce Multivariate Risk. In: Proceedings of the APCOM 2007: 33rd International Symposium on the Application of Computers and Operations Research in the Mineral Industry, pp. 65-69.

Rivoirard, J. 1990. Introduction to Disjunctive Kriging and Nonlinear Geostatistics. Centre de Géostatistique, Ecole des Mines de Paris, 89p. Este polígrafo foi encontrado na internet via <http://cg.ensmp.fr/biblioteque/index.html#1990>.

Rocha, M. N. da, Yamamoto, J. K. 2000. Comparison Between Variance and

Interpolation Variance as Uncertainty Measurements in the Capanema Iron Mine, State of Minas Gerais – Brazil. *Natural Resources Research*, Vol. 9, No. 3, pp. 223-235.

Shannon, C. E. 1948. A Mathematical Theory of Communication. Uma versão do artigo apresentado no “The Bell System Technical Journal” foi encontrada na Internet via <http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf>. Capturado em 18/04/07.

Souza, L. E. 2007. Proposição Geoestatística Para Quantificação do Erro em Estimativas de Tonelagens e Teores. Tese de Doutorado. Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais (PPGEM), Universidade Federal do Rio Grande do Sul, 193p.

Sullivan, J. 2003. Determining The Value of Additional Drilling. In: *Application of Computers and Operations Research In the Mineral Industry*, 2003, Cape Town. Proceedings of the 31st International Symposium. South African Institute of Mining and Metallurgy, Symposium Series S31, 2003, pp. 445-452.

Van Groenigen, J. W. 2000. The Influence of Variogram Parameters on Optimal Sampling Schemes for Mapping by Kriging. *Geoderma*, Vol. 97, pp. 223-236.

Van Groenigen, J. W., Siderius, W., Stein, A. 1999. Constrained Optimisation of Soil Sampling for Minimisation of Kriging Variance. *Geoderma*, Vol. 87, pp. 239-259.

Van Meirvenne, M., Goovaerts, P. 2001. Evaluating the Probability of Exceeding a Site-specific Soil Cadmium Contamination Threshold. *Geoderma*, Vol. 102, pp. 75-100.

Warrick, A. W., Myers, D. E. 1987. Optimization of Sampling Locations for Variogram Calculations. *Water Resources Research*, Vol. 23, No. 3, pp. 496-500.

Yamamoto, J. K., Conde, R. P. 1999. Classificação de Recursos Minerais Usando a Variância de Interpolação. *Revista Brasileira de Geociências*, Vol 29, No. 3, pp. 349-356.

Anexo A

Esse anexo contém a descrição detalhada de algumas etapas do algoritmo para construção de configurações de amostras apresentado no Capítulo 5.

A.1 - Escolha dos nós do *grid1* que devem receber amostras adicionais – Configuração 1

Para realizar computacionalmente a escolha dos nós que devem receber sondagem, propõe-se a seguinte rotina: atribuir valor 1 para os nós candidatos a receberem sondagem, enumerar os nós que receberam valor 1 (o Isatis[®] apresenta uma função para ordenar os valores de um atributo em ordem crescente ou decrescente, contudo se os valores do atributo são iguais, esse ordenamento dependerá da forma que o software ordena os nós do *grid* em um arquivo) e escolher os nós que receberam numeração de 1 a n .

Essa rotina pode ser realizada no software Isatis[®] por uma função de seleção de dados de um atributo, uma função de cálculo com valores de atributos e uma função de enumeração de valores de um atributo:

- I. Cria-se uma seleção para os nós do *grid1* que não podem receber sondagem adicional e uma seleção para os nós que são candidatos à sondagem;
- II. Cria-se um atributo onde os nós selecionados como candidatos à sondagem recebem valor 1;
- III. Enumera-se os valores de 1;
- IV. Seleciona-se os valores de 1 que receberam numeração de 1 a n .

A mesma rotina pode ser realizada usando, apenas, a função de cálculo com valores de atributos e a função de enumeração de valores de um atributo. A seguir é descrito um exemplo de como a rotina pode ser realizada no Datamine Studio[®]:

I. Atribui-se valor 1 aos nós do *grid1* que são candidatos à sondagem e valor 0 para os nós que não são candidatos. A explicação de como isso é feito computacionalmente é apresentada a seguir:

Ao final da interpolação dos valores de x no *grid1*, um valor de x é atribuído para cada nó que não é candidato à sondagem. O menor valor de x que pode ser atribuído a cada nó é igual à menor coordenada x ($xmin$) do arquivo COLLAR. Assim, uma fórmula matemática que envolve condição pode ser usada para criar um atributo (*Excl*), onde os nós candidatos à sondagem possuem valor 1 e os demais possuem valor 0:

$$\text{Excl}(u) = \begin{cases} 0 & \text{se } X(u) \geq xmin \\ 1 & \text{outro caso} \end{cases} \quad (\text{Equação A.1})$$

onde:

$X(u)$ = valor interpolado de x para o nó do *grid1* no local u ;

$xmin$ = menor valor da coordenada x para os furos iniciais;

$\text{Excl}(u)$ = valor que indica se o nó do *grid1*, no local u , é candidato a receber sondagem adicional.

Nota 1: Quando o software não permite interpolação usando a variável x , dois campos com os dados x podem ser criados no arquivo COLLAR. Um para servir como coordenada e o outro como variável.

Nota 2: Antes da interpolação dos valores de x , cria-se um atributo (“Aux”) com valores numéricos (quaisquer) para cada nó do *grid1*. Os resultados da interpolação do atributo x no *grid1* são armazenados no mesmo arquivo com os

valores de “Aux”. A criação de “Aux” permite que valores 1 sejam atribuídos aos locais não interpolados para a variável x (equação A.1).

II. Ordenam-se os valores do atributo (*Excl*) em ordem decrescente. Os nós do *grid1* são ordenados no arquivo de modo que o atributo (*Excl*) seja apresentado em ordem decrescente (ver Tabela A.1);

Tabela A.1 – Esquema de arquivo para o *grid1* apresentado na Figura 5.2

X (grid1)	Y (grid1)	Z (grid1)	x (Interpolado)	Excl	Num	Sond
0	0	0	-	1	1	1
0	50	0	-	1	2	1
0	100	0	-	1	3	1
0	150	0	-	1	4	1
50	0	0	-	1	5	1
50	50	0	-	1	6	1
50	150	0	-	1	7	1
100	0	0	-	1	8	1
100	50	0	-	1	9	1
100	100	0	-	1	10	1
100	150	0	-	1	11	1
150	0	0	-	1	12	1
150	50	0	-	1	13	0
150	100	0	-	1	14	0
150	150	0	-	1	15	0
50	100	0	45	0	16	0

III. Cria-se um novo atributo (*Num*) onde os nós de *grid1* são enumerados de acordo com a ordem em que são armazenados no arquivo (Tabela A.1);

IV. Cria-se um novo atributo (*Sond*) onde os valores de 1 a n do atributo (*Num*) recebem valor 1 e os demais recebem valor 0. O valor n é o número de sondagens adicionais, assim, os nós do *grid1* que devem receber sondagem possuem valor 1 para o atributo (*Sond*).

A.2 - Locação de amostras (3D) e adição dessas amostras aos dados iniciais - Configuração 1

Para selecionar, computacionalmente, os nós do *gridsim* onde serão adicionadas novas amostras utiliza-se o método de estimativa Vizinho Mais Próximo:

- I. Considere que o atributo (*Sond*), criado no passo iii, possui valor 1 para os nós do *grid1* que devem receber sondagem adicional e 0 para os demais;
- II. Os valores do atributo (*Sond*) são atribuídos no *gridsim* por Vizinho Mais Próximo. Conforme esquema da Figura A.1 (segundo exemplo da Figura 5.3), o elipsóide de busca de dados para interpolação deve apresentar o maior eixo na direção das novas sondagens e esse eixo deve alcançar a elevação do *grid1*, isto deve ocorrer para estimativa de todos nós do *gridsim*. Além disso, os demais eixos do elipsóide de busca devem ser pequenos, para que só permitam que nós do *gridsim* na direção da sondagem, ou na direção muito próxima a sondagem, recebam o respectivo valor do atributo (*Sond*), ou seja, o valor do atributo (*Sond*) que se encontra mais próximo do maior eixo do elipsóide de busca.

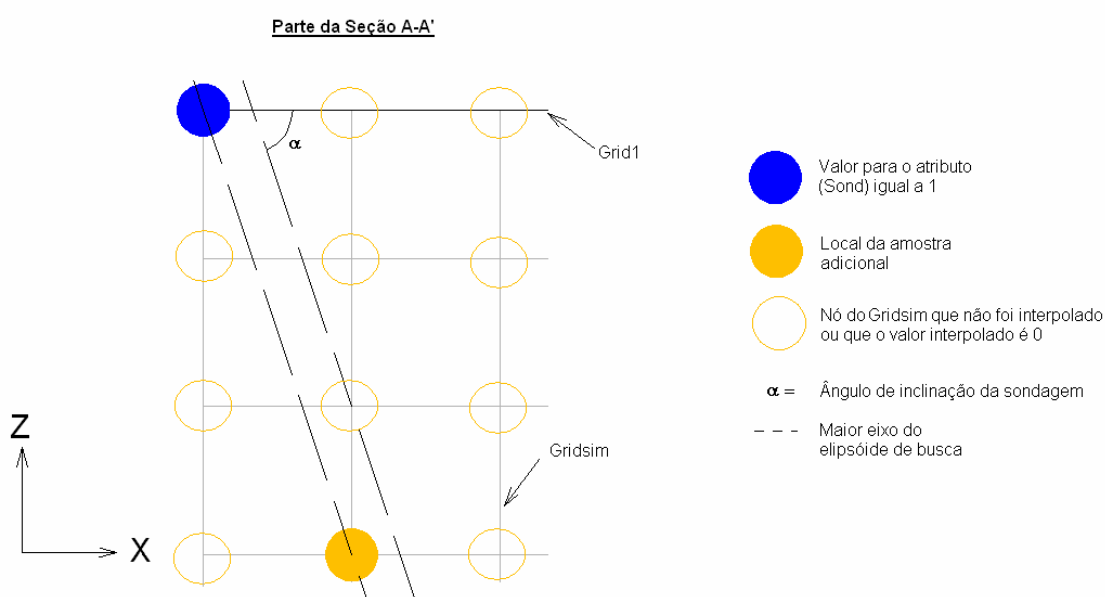


Figura A.1 – Esquema da busca de dados para estimativa no *gridsim* por Vizinho Mais Próximo. Os locais das amostras adicionais no *gridsim* recebem valor 1.

III. Os nós do *gridsim* podem receber valor 1 ou 0 como resultado da atribuição de valores, ou ainda não serem atribuídos. Considere que os resultados dessa estimativa são valores do atributo (*Amostras*) para o *gridsim*. Os nós do *gridsim* que recebem valor 1 para o atributo (*Amostras*) são os locais das amostras adicionais.

Conhecendo os locais onde devem ser adicionadas novas amostras, deve-se “coletar” os valores do cenário de referência nesses locais. Isso pode ser feito computacionalmente usando a fórmula matemática que envolve condição, para os nós do *gridsim*:

$$Ad(u) = \begin{cases} Ref(u) & \text{se } Amostras(u) = 1 \\ -999 & \text{outros casos} \end{cases} \quad (\text{Equação A.2})$$

onde:

$Ad(u)$ = valor do atributo (*Ad*) no local u . O atributo (*Ad*) contém os valores das amostras adicionais para a Configuração 1;

$Ref(u)$ = valor do cenário de referência no local (u);

$Amostras(u)$ = valor do atributo (*Amostras*) no local u .

O valor -999 para o atributo (*AD*) significa valor não existente e não é considerado nas etapas seguintes. Os valores do atributo (*AD*) são adicionados aos dados iniciais, formando a Configuração 1.

A.3 - Criação do *grid* para determinar espaçamento mínimo (em x e y) entre amostras – Configuração 2

Para realizar computacionalmente a escolha dos nós do *grid3* que devem receber sondagem, propõe-se uma rotina que envolve funções de atribuição de valores aos nós por Vizinho Mais Próximo, cálculo com valores de atributos e

enumeração de valores de um atributo. A seguir, é exibido um exemplo de como a rotina pode ser implementada no software Datamine Studio[®]:

I. Estimativa por Vizinho Mais Próximo dos valores de x (arquivo COLLAR) no *grid3* - Os nós do *grid3* que estão a uma distância dos dados iniciais menor que a permitida não são candidatos à sondagem. Para identificar os nós do *grid3* que estão muito próximos dos dados iniciais, a variável x do arquivo COLLAR é interpolada no *grid3* pelo método Vizinho Mais Próximo. A distância máxima de busca de amostras deve ser igual à menor distância permitida entre sondagens no plano x - y . Desta forma, os nós do *grid3* que forem interpolados nesse raio de influência estão a uma distância dos furos iniciais menor do que a permitida, portanto, esses nós não devem receber sondagem adicional;

II. Estimativa por Vizinho Mais Próximo da *média dos indicadores (grid2)* no *grid3* - Além dos nós do *grid3* que estão muito próximos dos dados iniciais, os nós correspondentes a locais com baixa probabilidade de serem minério também não são candidatos à sondagem. Esses nós são identificados pelo baixo valor de *média dos indicadores*. O valor de *média dos indicadores* para cada nó do *grid3* já foi calculado no passo anterior, visto que os nós do *grid3* coincidem com nós do *grid2*. A variável *média dos indicadores* criada para o *grid2* é atribuída ao *grid3* por Vizinho Mais Próximo. Essa função é utilizada para copiar os valores de *média dos indicadores* do *grid2* para o respectivo nó no *grid3*. Os valores de média muito pequenos (o usuário define um limite para esse valor) identificam locais que não devem receber sondagem adicional. A Figura 5.6 apresenta um esquema da escolha de nós candidatos à sondagem pelo algoritmo proposto. Observa-se que os nós que não são candidatos à sondagem, porque estão muito próximos dos dados iniciais ou porque correspondem a locais que não interessam economicamente, não estão marcados com uma circunferência azul;

III. Estimativa por Vizinho Mais Próximo da *média de IEQ (grid2)* no *grid3* - Para a configuração 2, os nós que são candidatos à sondagem e apresentam os n maiores valores de *média de IEQ* são escolhidos para receber sondagem. Sendo n o número de sondagens que se deseja adicionar. A fim de se conhecer os valores de *média de IEQ* para os nós do *grid3*, a variável *média de IEQ* criada para o *grid2* é transferida

ao *grid3* por Vizinho Mais Próximo. Essa função é utilizada para transferir os valores de *média de IEQ* do *grid2* para o respectivo nó no *grid3*;

IV. Cálculo com valores de atributos (exclusão dos nós do *grid3* que não são candidatos à sondagem) - Os valores de *média de IEQ* para os nós do *grid3* devem ser organizados em ordem decrescente. Desse modo, os nós com os *n* maiores valores de IEQ podem ser selecionados para receberem sondagem adicional. Contudo, antes de se ordenar os valores de *média de IEQ*, deve-se excluir os nós que não são candidatos a receber sondagem. O que se tem até essa etapa, é um arquivo com os valores de *média de indicadores*, com os valores da variável *x* (obtido a partir dos dados do arquivo COLLAR) e *média de IEQ* para cada nó do *grid3*. Os valores de *média de IEQ* para os nós do *grid3* que não são candidatos à sondagem devem ser substituídos por um valor que signifique *média de IEQ* inexistente naquele local (e.g. -999). Isso pode ser feito computacionalmente usando a fórmula matemática que envolve condição, para os nós do *grid3*:

$$\text{CandIEQ}(u) = \begin{cases} -999 & \text{se } X(u) \geq x_{\min} \text{ ou } \text{Medind}(u) < L \\ \text{MedIEQ}(u) & \text{outros casos} \end{cases} \quad (\text{Equação A.3})$$

onde:

$\text{CandIEQ}(u)$ = média de IEQ naquele local, se *u* é um local candidato à sondagem.
Valor -999 se *u* não é candidato à sondagem;

$X(u)$ = valor de *x* assumido a partir dos dados do arquivo COLLAR;

x_{\min} = menor valor da coordenada *x* para os furos iniciais;

$\text{Medind}(u)$ = média dos indicadores no nó *u*;

L = porcentagem mínima (definida pelo usuário) dos nós do *gridsim*, ao longo de uma sondagem, que devem ser candidatos a receberem sondagem adicional;

$\text{MedIEQ}(u)$ = média dos valores de IEQ para o nó *u*.

V. Cálculo com valores de atributos (organização dos valores de *média de IEQ* em ordem decrescente e enumeração desses valores) – Nessa etapa, a variável *CandIEQ*, que possui os valores de *média de IEQ* para os nós do *grid3* candidatos à

sondagem e o valor -999 para os demais nós, é organizada em ordem decrescente. Cria-se um novo atributo (*Num2*), onde, os nós do *grid3* são enumerados de acordo com a ordem que são armazenados no arquivo (a Figura A.2 apresenta um exemplo de *grid3*, diferente do que foi apresentado nas figuras anteriores, e a Tabela A.2 apresenta os valores para as variáveis mencionadas, considerando o *grid3* da Figura A.2);

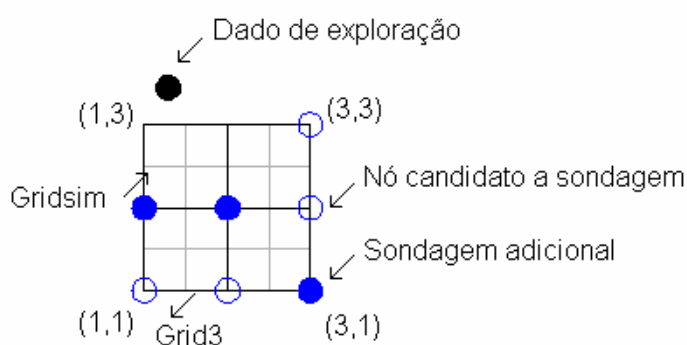


Figura A.2 – Exemplo da construção do *grid3* e escolha dos nós que devem receber sondagem.

Tabela A.2 – Esquema de arquivo para o *grid3* apresentado na Figura A.2

X (grid3)	Y(grid 3)	Z (grid 3)	X (Interpolado)	Medind	MedIEQ	CandIEQ	Num2	Sond2
1	2	1	-	1	10	10	1	1
2	2	1	-	0,7	9	9	2	1
3	1	1	-	0,6	8	8	3	1
1	1	1	-	0,8	6,5	6,5	4	0
3	3	1	-	0,5	6	6	5	0
3	2	1	-	0,9	5,8	5,8	6	0
2	1	1	-	0,9	3,2	3,2	7	0
1	3	1	1,3	0,9	5,4	-999	8	0
2	3	1	1,3	0,9	3,9	-999	9	0

Obs: Um nó do *grid3* com valor de média dos indicadores acima de 0,1 é candidato à sondagem.

Nota: A coluna "x (Interpolado)" é a estimativa obtida por Vizinho mais Próximo da variável x (arquivo COLLAR).

A coluna "Medind" é a estimativa obtida por Vizinho mais Próximo da média dos indicadores (*grid2*) no *grid3*.

A coluna "MedIEQ" é a estimativa obtida por Vizinho mais Próximo da média de IEQ (*grid2*) no *grid3*.

A coluna "CandIEQ" é o valor obtido pela equação A.3.

A coluna "Num2" é a numeração dos nós do *grid3* de acordo com a ordem que estão armazenados no arquivo (após ordenamento de "CandIEQ" em ordem decrescente).

A coluna "Sond2" é igual a 1 para os valores de "Num2" que pertencem ao intervalo [1,n], onde n é o número de sondagens adicionais, e 0 para os demais.

VI. Cálculo com valores de atributos (atribuir valor 1 para os nós do *grid3* que devem receber sondagem) - Cria-se um novo atributo (*Sond2*) onde os valores de 1 a *n* do atributo (*Num2*) recebem valor 1 e os demais recebem valor 0. O valor *n* é o número de sondagens adicionais, assim, os nós do *grid3* que devem receber sondagem possuem valor 1 para o atributo (*Sond2*).

A rotina para escolha dos nós do *grid3* que devem receber sondagem pode ser implementada de forma um pouco diferente, usando uma função de seleção de dados de um atributo. A seguir é exibido um exemplo de como a rotina pode ser implementada no software Isatis[®]:

- I. Estimativa por Vizinho Mais Próximo dos valores de *x* (arquivo COLLAR) no *grid3*;
- II. Estimativa por Vizinho Mais Próximo da *média dos indicadores (grid2)* no *grid3*;
- III. Criação de uma seleção para os nós do *grid3* que não podem receber sondagem adicional e uma seleção para os nós que são candidatos à sondagem;
- IV. Estimativa por Vizinho Mais Próximo da *média de IEQ (grid2)* no *grid3* (a estimativa é realizada apenas nos nós selecionados como candidatos à sondagem);
- V. Cálculo com valores de atributos (enumeração dos valores de *média de IEQ* do *grid3* em ordem decrescente);
- VI. Cálculo com valores de atributos (atribuir valor 1 para os nós do *grid3* que receberam numeração de 1 a *n* na etapa anterior, sendo *n* o número de sondagens adicionais, e 0 para os demais).

A.4 - Automação da Construção das Configurações e Simulação – Journal File (Isatis[®])

A sequência de operações realizadas pelo *Samples3D* e os parâmetros necessários em cada operação são descritos a seguir. Para facilitar a compreensão das operações descritas, uma lista com o nome das variáveis mencionadas, e que são definidas no arquivo *Samples3D*, é apresentada na Tabela A.3.

Tabela A.3 - Lista com o nome das variáveis definidas no arquivo *Samples3D*

Variável definida no Samples3D	Explicação
gridsim	Grid para simulação dos dados iniciais
gridbl	Modelo de blocos para o cálculo da função de transferência
data	Diretório para os dados iniciais
dataset	Arquivo com os dados iniciais
gaus	Dados iniciais normalizados
simdat	Simulação para os dados iniciais no domínio Gaussiano
modnorm	Modelo de continuidade para os dados iniciais normalizados
nsim	Parâmetros da vizinhança de busca para simulação
bksimdat	Simulação para os dados iniciais no domínio original
anamdat	Arquivo com os parâmetros de anamorfose
orig	Dados iniciais no domínio original
numpt	Número de pontos simulados do gridsim dentro do bloco do gridbl
selnumpt	Seleção dos blocos do gridbl com o numpt dentro do intervalo definido pelo usuário
bkdatbl	Média dos pontos do gridsim dentro do bloco do gridbl – Considerando simulação para dados iniciais
quant	Quartis da simulação dos dados iniciais para cada nó do gridsim
above	Probabilidade de cada nó do gridsim estar acima do Teor Limite – Considerando simulação dos dados iniciais
IR	IEQ para cada nó do gridsim - Considerando simulação dos dados iniciais
abcut	Indicador para probabilidade do nó do gridsim estar acima do Teor Limite - Considerando simulação dos dados iniciais
grid1	Grid1
collar	Arquivo COLLAR com as coordenadas do começo de cada furo inicial
X	Coordenada x do arquivo COLLAR
collarx (grid1)	Interpolação da coordenada x do arquivo COLLAR no grid1
mindist	Vizinhança de busca para interpolação da coordenada x do arquivo COLLAR no grid1
selmin	Seleção dos nós do grid1 interpolados para coordenada x do arquivo COLLAR
seldist	Seleção dos nós do grid1 não interpolados para coordenada x do arquivo COLLAR
cand	Variável com valor 1 para os nós do grid1 que pertencem a seleção seldist e valor inexistente para os demais nós
sortcand	Numeração dos nós do grid1 que são candidatos à sondagem
Ad	Atributo com valor 1 para nós do grid1 que devem receber sondagem e 0 para os demais
selad	Seleção dos nós do grid1 que devem receber sondagem
ponto1	Variável com valor 1 para os nós do gridsim que devem receber amostra adicional e valor inexistente para os demais - Configuração 1
neighad	Vizinhança de busca para interpolação dos valores de ad pertencentes à selad no gridsim. O resultado da interpolação é ponto1
adreg	Variável com valor do cenário de referência para os nós do gridsim que devem receber amostra adicional - Configuração 1
seladreg	Seleção para os nós do gridsim que devem receber amostra adicional - Configuração 1
simulation	Diretório com os arquivos gerados durante a execução de SAMPLES3D
conf1	Arquivo com as amostras para Configuração 1
var1	Atributo de interesse para os dados da Configuração 1
grid2	Grid2
IRmed	Média dos valores de IR na direção da sondagem adicional
abctmed (grid2)	Média dos valores de abcut na direção da sondagem adicional
grid3	Grid3

Tabela A.3 - (Continuação)

Variável definida no Samples3D	Explicação
collarx (grid3)	Interpolação da coordenada x do arquivo COLLAR no grid3
abctmed (grid3)	Média dos valores de abcut na direção da sondagem adicional
selA	Seleção dos nós do grid3 interpolados para coordenada x do arquivo COLLAR
selB	Seleção dos nós do grid3 com baixo valor de abctmed, nós não candidatos à sondagem
selC	Seleção dos nós do grid3 que pertencem à seleção selA e selB
selD	Seleção dos nós do grid3 candidatos à sondagem
medIR (grid3)	Valor de IRmed (grid2) atribuído para os nós do grid3 candidatos à sondagem
sortmedIR	Numeração de medIR (grid3) em ordem decrescente
adIR (grid3)	Atributo com valor 1 para nós do grid3 que devem receber sondagem e 0 para os demais
selE	Seleção dos nós do grid3 que devem receber sondagem
adIR (gridsim)	Variável com valor 1 para os nós do gridsim que devem receber amostra adicional e valor inexistente para os demais - Configuração 2
adIRref	Variável com valor do cenário de referência para os nós do gridsim que devem receber amostra adicional - Configuração 2
seladref	Seleção para os nós do gridsim que devem receber amostra adicional - Configuração 2
conf2	Arquivo com as amostras para Configuração 2
var2	Atributo de interesse para os dados da Configuração 2
simconf1	Simulação para Configuração 1 no domínio Gaussiano
bkconf1	Simulação para Configuração 1 no domínio original
bkconf1bl	Média dos pontos do gridsim dentro do bloco do gridbl – Considerando simulação para Configuração 1
simconf2	Simulação para Configuração 2 no domínio Gaussiano
bkconf2	Simulação para Configuração 2 no domínio original
bkconf2bl	Média dos pontos do gridsim dentro do bloco do gridbl – Considerando simulação para Configuração 2
sequence	Arquivo com sequência de mineração dos blocos
seq	Variável com sequência de mineração dos blocos
seqbl (gridbl)	Número da sequência de mineração do bloco para cada bloco do gridbl
selseqbl	Seleção dos blocos que são usados no cálculo da função de transferência

I. Criação do Grid para simulação dos dados iniciais

Criação de um *grid* (*gridsim*) onde será realizada a simulação com base nos dados iniciais. Os parâmetros do *grid* devem seguir a definição adotada no Isatis[®]. O usuário deve definir os campos:

1. X Origin (*gridsim*) – Coordenada x da origem.
2. Y Origin (*gridsim*) - Coordenada y da origem.
3. Z Origin (*gridsim*) - Coordenada z da origem.
4. X Mesh (*gridsim*) – Espaçamento entre nós em x.

5. Y Mesh (gridsim) - Espaçamento entre nós em y.
6. Z Mesh (gridsim) - Espaçamento entre nós em y.
7. X Nodes (gridsim) – Número de nós em x.
8. Y Nodes (gridsim) - Número de nós em y.
9. Z Nodes (gridsim) - Número de nós em z.

II. Criação do Modelo de Blocos que será Usado no Cálculo das Funções de Transferência

No software Isatis[®], assim como no Datamine Studio[®], cada *grid* criado é entendido como um modelo de blocos. O espaçamento entre nós de *grid* corresponde ao tamanho do bloco naquela direção e o nó de *grid* corresponde ao centro do bloco. O objetivo da operação 2 é criar um *grid* (*gridbl*) para representar o modelo de blocos que será usado no cálculo da função de transferência. O valor de cada bloco minerado será a média dos pontos simulados (*gridsim*) que são englobados pelo bloco. Para realizar operações com *grids* no Isatis[®], o *gridbl* deve ser construído de forma que as arestas dos blocos coincidam com as arestas dos blocos delimitados pelo *gridsim*. O usuário deve definir os campos:

10. X Origin (gridbl)
11. Y Origin (gridbl)
12. Z Origin (gridbl)
13. X Mesh (gridbl)
14. Y Mesh (gridbl)
15. Z Mesh (gridbl)
16. X Nodes (gridbl)
17. Y Nodes (gridbl)
18. Z Nodes (gridbl)

III. Simulação por Bandas Rotativas dos Dados Iniciais

Simulação por Bandas Rotativas com base nos dados iniciais. Parâmetros:

19. Number of Bands – Número de bandas que devem ser utilizadas na simulação.
20. Number of Real. – Número de realizações (cenários).

21. Rot. Around Z (Search) – Definição da elipse de busca de dados. Antes da rotação dos eixos, o eixo x representa a direção leste, o y a direção norte e o z a direção vertical. Esse parâmetro define a rotação (ângulo) dos eixos x e y ao redor de z. De acordo com as convenções do Isatis[®] que foram escolhidas para o *Samples3D*, a Figura A.3 define quando o ângulo de rotação é negativo. Os eixos rotacionados são chamados x' e y' , respectivamente.

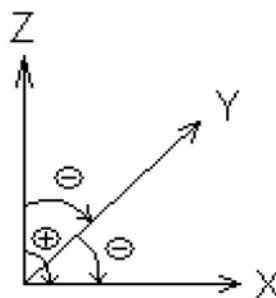


Figura A.3 – Definição do ângulo de rotação.

22. Rot. Around Y (Search) - Esse parâmetro define a rotação dos eixos x' e z ao redor de y' . Os eixos rotacionados são chamados x'' e z' , respectivamente.

23. Rot. Around X (Search) – Esse parâmetro define a rotação dos eixos y' e z' ao redor de x'' . Os eixos rotacionados são chamados y'' e z'' , respectivamente.

24. Max. Dist. X (Search) – Distância máxima de busca de dados na direção x'' .

25. Max. Dist. Y (Search) – Distância máxima de busca de dados na direção y'' .

26. Max. Dist. Z (Search) – Distância máxima de busca de dados na direção z'' .

27. Ang. Sectors (Search) – Número de setores angulares para busca de dados. De acordo com definição do Isatis[®].

28. Min. Data (Search) – Número mínimo de dados usados na simulação.

29. Data Sector (Search) – Número máximo de dados em cada setor angular.

Obs: O esquema de rotação dos eixos descrito aqui é usado para a definição dos ângulos de rotação que serão mencionados nas próximas operações.

IV. Retro-Transformação da Variável Gaussiana Obtida com a Simulação para os Dados Iniciais

Os valores gaussianos simulados com base nos dados iniciais são retro-transformados em valores do espaço original. O arquivo de anamorfose (*anamdat*) é utilizado.

V. Cálculo do Número de Pontos Simulados dentro dos Blocos

Para cada cenário simulado com os dados iniciais, calcula-se o número de nós do *gridsim* que são englobados por cada bloco do *gridbl*. Dependendo da posição dos dados originais, alguns pontos do *gridsim* não são simulados (não possuem dados dentro da vizinhança de busca), assim, o valor de cada bloco (*gridbl*) pode ser obtido a partir de um número diferente de simulações. Na próxima operação, o usuário pode definir o número mínimo de pontos simulados dentro do bloco para que esse seja usado no cálculo da função de transferência.

VI. Seleção dos Blocos que Serão Usados no Cálculo da Função de Transferência

Cria-se uma seleção dos blocos que serão usados no cálculo das funções de transferência. O usuário define o número mínimo e máximo de pontos simulados dentro do bloco para que esse seja considerado no cálculo das funções de transferência.

30. Points inside block (a) – Mínimo de pontos simulados dentro do bloco.

31. Points inside block (b) – Máximo de pontos simulados dentro do bloco.

VII. Cálculo do Valor de Cada Bloco – Média dos Pontos Simulados dentro do Bloco

Para cada cenário simulado com os dados iniciais, calcula-se a média dos pontos simulados do *gridsim* dentro de cada bloco do *gridbl*. Apenas os blocos selecionados pela operação VI são considerados.

VIII. Pós-Processamento da Simulação com base nos Dados Iniciais

Calcula-se, em cada nó do *gridsim*, os quartis da distribuição condicional, dos valores simulados em todas realizações, condicionados pelos dados iniciais. Essa operação é utilizada para construção do índice de incerteza em cada nó simulado. Além disso, calcula-se a probabilidade do valor do nó ser igual ou maior que um teor de limite (conforme mencionado no passo v do algoritmo de construção de configurações). A probabilidade é o número de realizações igual ou maior ao teor

dividido pelo total de realizações. O usuário define o valor do teor limite. Por causa do formato dessa operação no software, o teor é inserido duas vezes:

32. Cutoff (a) (Index) – Teor limite.

33. Cutoff (b) (Index) – Teor limite.

IX. Cálculo do IEQ em Cada Local Simulado com base nos Dados Iniciais

Calcula-se, em cada nó do *gridsim*, o IEQ (intervalo entre quartis), da distribuição de valores simulados em todas realizações, condicionados pelos dados iniciais.

X. Criação de Indicadores (Equação 5.1)

De acordo com o passo v do algoritmo de construção de configurações, atribui-se valor 1 para os nós do *gridsim* cuja probabilidade do valor ser igual ou maior ao teor limite é superior a 0,7. O valor 0 é atribuído para os demais nós. A variável criada é *abcut*.

XI. Criação do Grid1

De acordo com o passo i do algoritmo de construção das configurações, cria-se o *grid1*. Os parâmetros do *grid1* são definidos pelo usuário:

34. X Origin (grid1)

35. Y Origin (grid1)

36. Z Origin (grid1)

37. X Mesh (grid1)

38. Y Mesh (grid1)

39. Z Mesh (grid1)

40. X Nodes (grid1)

41. Y Nodes (grid1)

42. Z Nodes (grid1)

XII. Interpolação por Vizinho Próximo da Variável x (arquivo COLLAR) no Grid1

A coordenada *x* do arquivo COLLAR é interpolada por Vizinho Próximo no *grid1*. A máxima distância de busca de amostras deve ser igual à distância mínima entre sondagens (ver passo ii do algoritmo de construção de configurações).

43. Min. Dist. X (data) – Mínima distância entre sondagens na direção leste.

44. Min. Dist. Y (data) – Mínima distância entre sondagens na direção norte.

XIII. Seleção de Nós do *Grid1* Interpolados para a Variável *x*

Cria-se uma seleção dos nós do *grid1* que foram interpolados pela operação XII. Ou seja, uma seleção dos nós do *grid1* que não são candidatos a receberem sondagem adicional.

45. Min. X (collar) – Mínima coordenada x do arquivo COLLAR.

46. Max. X (collar) - Máxima coordenada x do arquivo COLLAR.

XIV. Seleção de Nós do *Grid1* Candidatos a Receber Sondagem

Cria-se uma seleção excluindo os nós do *grid1* que não são candidatos a receberem sondagem adicional.

XV. Atribuir Valor 1 para os Nós do *Grid1* Candidatos à Sondagem

Atribui-se valor 1 para os nós do *grid1* que são candidatos a receberem sondagem adicional.

XVI. Enumeração dos Valores de 1

Os nós do *grid1* que são candidatos a receberem amostras adicionais (possuem valor 1) são ordenados. No Isatis[®], os nós dos *grids* são classificados em ordem crescente de *x*, depois de *y* e depois de *z*. A contagem dos valores de 1 será feita em ordem crescente seguindo essa classificação.

XVII. Escolha dos Nós do *Grid1* que Devem Receber Sondagem Adicional

Os nós do *grid1* que receberam valor 1 a *n* pela operação XVI recebem, agora, valor 1 (variável *ad*) e são os locais onde serão adicionadas sondagens. O valor *n* é o número de sondagens adicionais definido pelo usuário. Os nós do *grid1* que não vão receber sondagens recebem, agora, o valor 0 (variável *ad*).

47. Replace 11 for adic. data - campo - v2=ifelse(v1>11,0,1) – Substituir o valor 11 pelo número de sondagens adicionais para a Configuração 1.

XVIII. Seleção dos Nós do *Grid1* que Devem Receber Sondagem Adicional

Cria-se uma seleção dos nós do *grid1* com os valores da variável *ad* iguais a 1.

XIX. Escolha dos Nós do *Gridsim* que Devem Receber Amostra Adicional – Configuração 1

Os valores da variável *ad*, do *grid1*, iguais a 1 são atribuídos por Vizinho Próximo no *gridsim*. Essa operação é usada para atribuir valor 1 aos nós do *gridsim* que devem ceder os valores para a amostra adicional. Conforme descrito no passo vi do algoritmo de construção de configurações, a elipse de busca dos dados é definida de acordo com a inclinação e comprimento das sondagens adicionais:

48. Rot. around Z (hole)

49. Rot. around Y (hole)

50. Rot. around X (hole)

51. Max. Dist. X (hole)

52. Max. Dist. Y (hole)

53. Max. Dist. Z (hole)

54. All Data (hole) - Máximo de dados por setor usados na interpolação. Esse parâmetro não influencia essa operação, pois o valor do dado mais próximo será atribuído ao nó do *gridsim*. Contudo, esse parâmetro de busca também é usado na operação XXIV, onde todos os dados dentro da vizinhança de busca devem ser considerados na interpolação. Por isso, o valor desse parâmetro deve ser suficiente para garantir essa condição.

XX. Atribuir Valor Obtido no Modelo de Referência no Local da Amostra Adicional

Cria-se a variável *adreg* no *gridsim* com o valor gaussiano obtido no modelo simulado escolhido como referência nos locais que vão receber amostras adicionais e o valor -99 nos demais locais.

55. Replace 01 for Refer. – campo – v2=simdat[00001] – O valor 00001 deve ser substituído pelo índice da realização que será escolhida como referência (fonte de informações

adicionais). O usuário pode executar várias vezes o Samples3D e escolher diferentes cenários como referência.

XXI. Seleção das Amostras Adicionais para a Configuração 1

Cria-se uma seleção com os valores *adreg* do *gridsim* pertencentes ao intervalo -90 a 100. Ou seja, só os dados adicionais são selecionados.

XXII. Mesclar as Amostras Adicionais com os Dados Iniciais – Configuração 1

Cria-se um arquivo com os valores gaussianos que devem ser adicionados para a configuração 1 e com os dados gaussianos obtidos pela transformação dos dados iniciais.

XXIII. Criação do *Grid2*

De acordo com o passo vi do algoritmo de construção das configurações, cria-se o *grid2*. Os parâmetros do *grid2* são definidos pelo usuário:

56. X Origin (*grid2*)

57. Y Origin (*grid2*)

58. Z Origin (*grid2*)

59. X Mesh (*grid2*)

60. Y Mesh (*grid2*)

61. Z Mesh (*grid2*)

62. X Nodes (*grid2*)

63. Y Nodes (*grid2*)

64. Z Nodes (*grid2*)

XXIV. Cálculo da Média de IEQ ao Longo da Sondagem que Passa pelo Nó do *Grid2*

Os valores de IEQ do *gridsim* são interpolados no *grid2* por média móvel. O objetivo é atribuir, a cada nó do *grid2*, a média dos valores de IEQ ao longo da sondagem que passa pelo respectivo nó (ver passo vi do algoritmo de construção das configurações). A variável criada no *grid2* é chamada *IRmed*.

XXV. Cálculo da Média de Indicadores ao Longo da Sondagem que Passa pelo Nó do Grid2

Os valores de *abcut* (variável criada pela operação X) do *gridsim* são interpolados no *grid2* por média móvel. O objetivo é atribuir a cada nó do *grid2* a média dos valores de *abcut* ao longo da sondagem que passa pelo respectivo nó (ver passo vi do algoritmo de construção das configurações). A variável criada no *grid2* é chamada *abctmed*.

XXVI. Criação do Grid3

De acordo com o passo vii do algoritmo de construção das configurações, cria-se o *grid3*. Os parâmetros do *grid3* são definidos pelo usuário:

65. X Origin (grid3)

66. Y Origin (grid3)

67. Z Origin (grid3)

68. X Mesh (grid3)

69. Y Mesh (grid3)

70. Z Mesh (grid3)

71. X Nodes (grid3)

72. Y Nodes (grid3)

73. Z Nodes (grid3)

XXVII. Interpolação por Vizinho Próximo da Variável x (arquivo COLLAR) no Grid3

A coordenada *x* do arquivo COLLAR é interpolada por Vizinho Próximo no *grid3*. A máxima distância de busca de amostras deve ser igual à distância mínima entre sondagens (ver passo vii do algoritmo de construção de configurações).

XXVIII. Interpolação por Vizinho Próximo da Média dos Indicadores (Grid2) no Grid3

Os valores de *abctmed* (variável criada pela operação XXV) do *grid2* são interpolados no *grid3* por Vizinho Próximo (ver passo vii do algoritmo de construção de configurações).

XXIX. Seleção de Nós do Grid3 Interpolados para a Variável x

Cria-se uma seleção para os valores da coordenada x atribuídos ao *grid3*. O usuário deve definir:

74. Min. X (collar) – Mínima coordenada x do arquivo COLLAR.

75. Max. X (collar) - Máxima coordenada x do arquivo COLLAR.

XXX. Seleção de Nós do *Grid3* com Baixo Valor de Média dos Indicadores

Cria-se uma seleção para os valores de *abctmed* interpolados no *grid3*. A variável *abctmed* corresponde à proporção de pontos simulados, na direção da sondagem, que são candidatos a serem amostrados. O objetivo é selecionar os nós onde *abctmed* é baixo, ou seja, selecionar os locais que não são candidatos a receberem sondagem. O usuário deve definir:

76. Prop. Limit – O menor valor de *abctmed* (excluindo esse valor) necessário para considerar o nó do *grid3* candidato a receber sondagem.

XXXI. Seleção dos Nós do *Grid3* que Não São Candidatos à Sondagem

Cria-se uma seleção para os nós do *grid3* que foram selecionados nas operações XXIX e XXX. Ou seja, uma seleção para os nós do *grid3* que não são candidatos a receberem sondagem.

XXXII. Seleção de Nós do *Grid3* Candidatos a Receber Sondagem

Cria-se uma seleção excluindo os nós do *grid3* que não são candidatos a receberem sondagem.

XXXIII. Interpolação por Vizinho Próximo da Média de IEQ (*Grid2*) no *Grid3*

Os valores de *IRmed* (variável criada pela operação XXIV) do *grid2* são interpolados no *grid3* por Vizinho Próximo, apenas nos locais candidatos a receberem sondagem adicional (ver passo vii do algoritmo de construção de configurações). A variável criada no *grid3* é *medIR*.

XXXIV. Enumeração da Média de IEQ no *Grid3* em Ordem Decrescente

Os valores de *medIR* do *grid3* são enumerados em ordem decrescente.

XXXV. Escolha dos Nós do *Grid3* que Devem Receber Sondagem Adicional

Os locais dos n maiores valores de *medIR* do *grid3* recebem valor 1 e os demais locais recebem valor 0. Onde n é o número de sondagens que devem ser adicionadas. A variável criada é chamada *adIR*.

77. Replace 9 for adic. data – campo – $v2=ifelse(v1>9,0,1)$ – O valor 9 deve ser substituído pelo número de sondagens que deve ser adicionado para Configuração 2.

XXXVI. Seleção dos Nós do *Grid3* que Devem Receber Sondagem Adicional

Os valores 1 da variável *adIR* no *grid3* são selecionados. Ou seja, os nós do *grid3* que devem receber sondagem adicional são selecionados.

XXXVII. Escolha dos Nós do *Gridsim* que Devem Receber Amostra Adicional – Configuração 2

Os valores da variável *adIR*, do *grid3*, iguais a 1 são atribuídos por Vizinho Próximo ao *gridsim*. Essa operação é usada para atribuir valor 1 aos nós do *gridsim* que devem ser selecionados para obter-se os valores das amostras adicionais.

XXXVIII. Atribuir Valor da Simulação de Referência no Local Selecionado para Obter-se a Amostra Adicional – Configuração 2

Cria-se a variável *adIRref*, no *gridsim*, com o valor gaussiano obtido na simulação escolhida de referência nos locais que vão receber amostras adicionais e o valor -99 nos demais locais.

78. Replace 01 for Refer. – campo – $v2=simdat[00001]$ – O valor 00001 deve ser substituído pelo índice da realização que foi escolhida como referência (fonte de informações adicionais).

XXXIX. Seleção das Amostras Adicionais para a Configuração 2

Cria-se uma seleção com os valores *adIRref* do *gridsim* pertencentes ao intervalo -90 a 100. Ou seja, só os dados adicionais são selecionados.

XL. Mesclar as Amostras Adicionais com os Dados Iniciais – Configuração 2

Cria-se um arquivo com os valores gaussianos que devem ser adicionados para a configuração 2 e com os dados gaussianos transformados a partir dos dados iniciais.

XLI. Simulação por Bandas Rotativas para Configuração 1

Simulação por Bandas Rotativas para a Configuração 1. Parâmetros:

79. Number of Bands

80. Number of Real

XLII. Retro-Transformação da Variável Gaussiana Obtida com a Simulação para Configuração 1

Os valores gaussianos para cada cenário gerado com os dados da Configuração 1 são retro-transformados em valores do espaço original.

XLIII. Cálculo do Valor de Cada Bloco – Média dos Pontos Simulados dentro do Bloco

Para cada cenário simulado com os dados da Configuração 1, calcula-se a média das simulações do *gridsim* dentro de cada bloco do *gridbl*. Apenas os blocos selecionados pela operação VI são considerados.

XLIV. Simulação por Bandas Rotativas para Configuração 2

Simulação por Bandas Rotativas para a Configuração 2. Parâmetros:

81. Number of Bands

82. Number of Real

XLV. Retro-Transformação da Variável Gaussiana Obtida com a Simulação para Configuração 2

Os valores gaussianos para cada cenário gerado com os dados da Configuração 2 são retro-transformados em valores do espaço original.

XLVI. Cálculo do Valor de Cada Bloco – Média dos Pontos Simulados dentro do Bloco

Para cada cenário simulado com os dados da Configuração 2, calcula-se a média dos valores dos nós simulados do *gridsim* dentro de cada bloco do *gridbl*. Apenas os blocos selecionados pela operação VI são considerados.

XLVII. Atribuir o Número da Seqüência de Mineração para Cada Bloco

Os valores da variável *seq* (do arquivo importado pelo usuário com a seqüência de mineração dos blocos) são atribuídos por Vizinho Próximo ao *gridbl*. Essa operação é usada para criar no arquivo *gridbl* uma variável (*seqbl*) com a ordem de mineração de cada bloco.

XLVIII. Seleção dos Blocos que serão usados no Cálculo da Função de Transferência

Cria-se uma seleção para os blocos do *gridbl* cujo valor de *seqbl* pertence ao intervalo definido pelo usuário. Os blocos selecionados são aqueles que serão usados no cálculo da função de transferência. Essa seleção é criada para que só os blocos que serão usados no cálculo da função de transferência sejam exportados para o arquivo de texto.

83. Max. Sequence - Valor máximo (incluindo esse valor) de *seqbl* para que o bloco seja considerado no cálculo da função de transferência.

XLIX. Exportação dos Modelos de Blocos para os Dados Iniciais

Exportação dos modelos de blocos obtidos a partir dos dados iniciais. Exporta-se um arquivo com as coordenadas *x,y,z* dos blocos, o valor dos blocos para cada cenário e a seqüência de mineração dos blocos.

84. Output (expl) – Localização e nome do arquivo de saída com os modelos de blocos.

L. Exportação dos Modelos de Blocos para a Configuração 1

Exportação dos modelos de blocos obtidos a partir da Configuração 1.

85. Output (conf1) – Localização e nome do arquivo de saída com os modelos de blocos.

LI. Exportação dos Modelos de Blocos para a Configuração 2

Exportação dos modelos de blocos obtidos a partir da Configuração 2.

86. Output (conf2) – Localização e nome do arquivo de saída com os modelos de blocos.

Anexo B

Esse anexo apresenta as referências dos artigos publicados durante o período do Doutorado. Alguns artigos são relacionados com o tema abordado nessa Tese.

Koppe, V. C., Costa, J. F. C. L., Koppe, J. C., Gambin, F. 2006. Geoestatística Não-Linear Versus Linear: Uma Comparação entre Krigagem Disjuntiva, Krigagem dos Indicadores e Krigagem Ordinária. In: IV Congresso Brasileiro de Mina a Céu Aberto e IV Congresso Brasileiro de Mina Subterrânea. CD-Rom, IBRAM.

Koppe, V. C., Costa, J. F. C. L., Koppe, J. C., Gambin, F., Fallon, G., Davies, N. 2007. Geostatistical Simulation of Acoustic Log Data for Seismic Depth Conversion. Natural Resources Research, Vol. 16, No. 4, pp. 293-303.

Koppe, V. C., Costa, J. F. C. L., Koppe, J. C. 2008. O Impacto da Configuração de Amostras Sobre a Função de Transferência. In: V Congresso Brasileiro de Mina a Céu Aberto & V Congresso Brasileiro de Mina Subterrânea e Workshop Barragens de Rejeito. CD-Rom, IBRAM.

Koppe, V. C., Costa, J. F. C. L., Koppe, J. C. 2008. The Benefit of Additional Data Accessed Via Transfer Functions. In: Geostats 2008, Santiago. Proceedings of the Eighth International Geostatistics Congress. Gecamin Ltd, Vol. 2, pp. 1179-1184.

Koppe, V. C., Costa, J. F. C. L., Koppe, J. C. 2009. Algorithm to Help Additional Drilling Location. In: Application of Computers and Operations Research in the Mineral Industry, 2009, Vancouver. CD-Rom, Canadian Institute of Mining, Metallurgy and Petroleum.

Anexo C

No Capítulo 3, as Figuras 3.10 a 3.13 mostram os histogramas para as quatro configurações de amostras. Cada uma dessas configurações compreende os 195 dados iniciais e 160 dados adicionais da variável V , obtidos a partir do cenário exaustivo. Todas as configurações possuem dados que representam um mesmo cenário (cenário exaustivo) e a estatística de cada configuração foi obtida considerando pesos de desagrupamento para os dados. Assim, os valores de média e desvio padrão (considerando pesos de desagrupamento) dos dados de V devem ser semelhantes para as quatro configurações.

Contudo, a Configuração 4 apresenta valores de média e desvio padrão um pouco diferentes das demais configurações (a média é um pouco mais elevada e o desvio padrão é um dos mais altos) e mais próximos dos respectivos parâmetros para o cenário exaustivo. Esse fato é um efeito dos pesos de desagrupamento calculados pelo software *Isatis*[®]. Quando esses pesos são calculados pelo programa *Declus* do *GSLIB* (Deutsch e Journel, 1998), os valores de média e desvio padrão de todas as configurações são mais semelhantes entre si (Figuras C.1 (a) a C.1 (d)).

Porém, a tendência dos parâmetros média e desvio padrão, obtidos para a Configuração 4 serem maiores que e os obtidos para as demais configurações, que pode ter sido causada pelo método de cálculo dos pesos de desagrupamento, exerceu pouca influência nos resultados do estudo (nas medidas de incerteza sobre as funções). Por exemplo, quando a realização 57 foi escolhida como cenário fonte de informação (Figuras 3.37 a 3.40), a Configuração 4 apresentou média mais semelhante às médias das demais configurações e o menor desvio padrão (ao contrário do que havia sido observado quando as informações adicionais foram obtidos a partir do cenário exaustivo), e mesmo assim, se uma configuração se destaca na redução da incerteza sobre a função, a mesma configuração se destaca no caso em que o cenário exaustivo foi usado como fonte de informação.

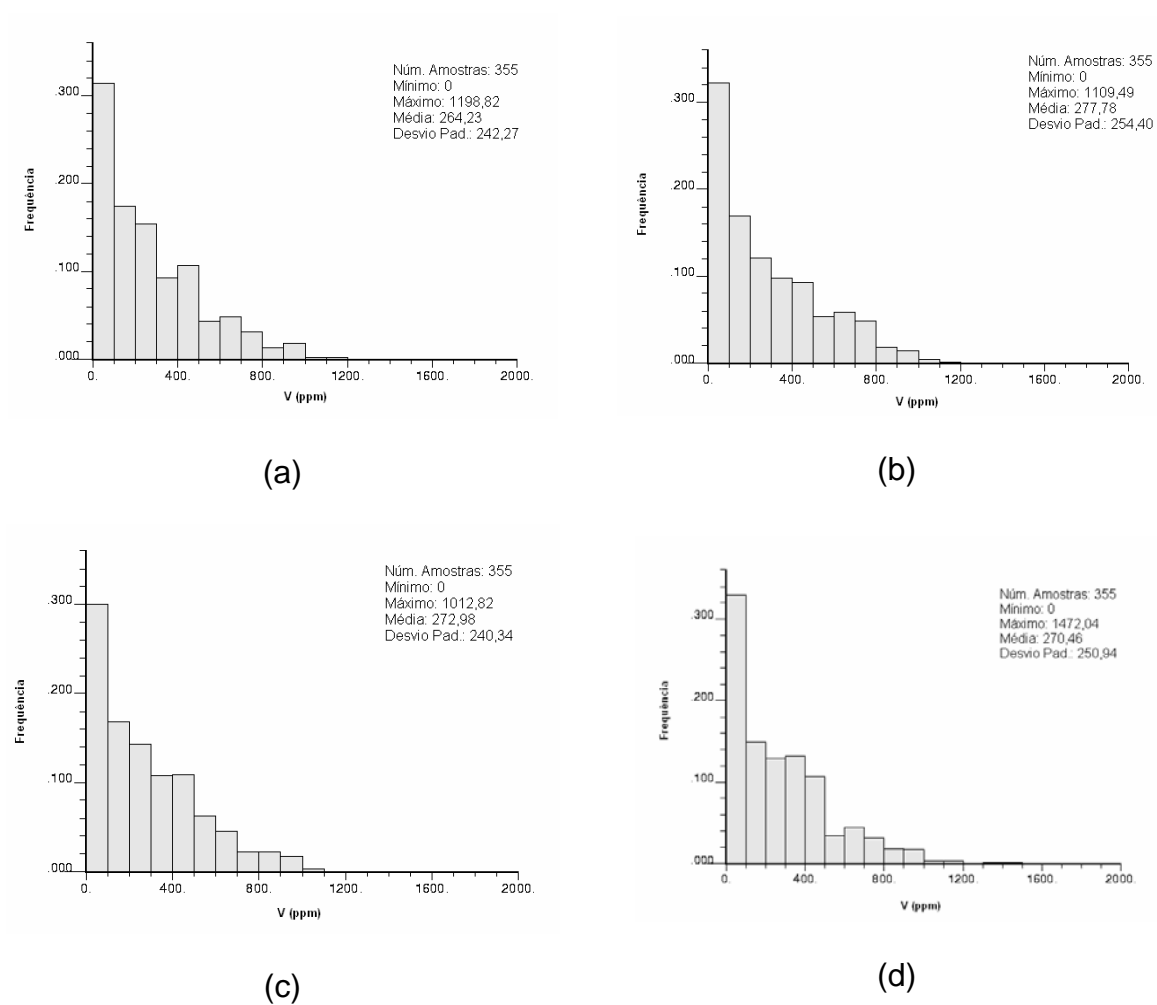


Figura C.1 – Histogramas considerando pesos de desagrupamento obtidos pelo programa *Declus*. (a) Histograma para a Configuração 1. (b) Histograma para a Configuração 2. (c) Histograma para a Configuração 3. (d) Histograma para a Configuração 4.