

Marcelo Batelochi

**Simulação geoestatística e regressão  
multivariada aplicadas na modelagem  
geometalúrgica de um depósito de zinco**

Porto Alegre

2018



Marcelo Batelochi

**Simulação geoestatística e regressão multivariada  
aplicadas na modelagem geometalúrgica de um depósito  
de zinco**

Esta dissertação foi analisada e julgada adequada para a obtenção do título de Mestre em Engenharia, área de concentração de Tecnologia Mineral, e aprovada em sua forma final pelo Orientador e pela Banca Examinadora designada pelo Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais da Universidade Federal do Rio Grande do Sul.

Universidade Federal do Rio Grande do Sul - UFRGS

Escola de Engenharia

Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais

Orientador: Prof. Dr. João Felipe Coimbra Leite Costa

Porto Alegre

2018

---

Marcelo Batelochi

Simulação geoestatística e regressão multivariada aplicadas na modelagem geometalúrgica de um depósito de zinco/ Marcelo Batelochi. – Porto Alegre, 2018-81 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. João Felipe Coimbra Leite Costa

Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul - UFRGS  
Escola de Engenharia

Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais,  
2018.

1. geometalurgia. 2. modelos geometalúrgicos. 3. regressão linear multivariada  
I. Prof. Dr. João Felipe Coimbra Leite Costa. II. Universidade Federal do Rio Grande do Sul. III. Faculdade de Engenharia. IV. Previsão da recuperação do minério in situ por métodos de simulação e modelagem geológica.

CDU 02:141:005.7

---

Marcelo Batelochi

**Simulação geoestatística e regressão multivariada  
aplicadas na modelagem geometalúrgica de um depósito  
de zinco**

Esta dissertação foi analisada e julgada adequada para a obtenção do título de Mestre em Engenharia, área de concentração de Tecnologia Mineral, e aprovada em sua forma final pelo Orientador e pela Banca Examinadora designada pelo Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais da Universidade Federal do Rio Grande do Sul.

Trabalho aprovado. Porto Alegre, XX de XXX de 2017:

---

**Prof. Dr. João Felipe Coimbra Leite  
Costa**  
Orientador

---

**Professor**  
Convidado 1

---

**Professor**  
Convidado 2

Porto Alegre  
2018



**À minha família.**



# Agradecimentos

Ao Professor João Felipe pelos ensinamento e pela recepção de braços abertos no laboratório. Aos colegas do LPM que contribuíram de alguma forma para o desenvolvimento desse trabalho, Mara Vieira, Roberto Rolo, Ricardo Radkte, Péricles Machado. Aos demais colegas de laboratório, em especial, Ryu Okada, Lucas Roncarati, Marcel Bassani, Cristina Araújo, David Drumond e Augusto Torres.

Aos meus pais pelo incentivo aos meus estudos e suporte financeiro. Ao meu irmão pelo companheismo. À Ane Catarine pelo companherismo e pela contribuição com o desenvolvimento desta dissertação.

Aos meus irmãos da República MataBurro pelos bons momentos.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) pela bolsa concedida.



*“Se fracassar, ao menos que  
fracasse ousando grandes feitos,  
de modo que a sua postura não seja  
nunca a dessas almas frias e tímidas  
que não conhecem nem a vitória  
nem a derrota.”*  
*(Theodore Roosevelt)*



# Resumo

A recuperação de metal na planta metalúrgica é o fator mais importante para o sucesso econômico de uma mina. A recuperação de massa da planta de beneficiamento é geralmente associada à teores de alimentação, porém os circuitos do processamento mineral são complexos e dependem da interação de uma diversidade de características da rocha. Com a redução dos teores nas operações de mina, a geometurgia vem ganhando destaque em questões de redução dos riscos de produção. Testes geometúrgicos foram realizados usando minério de 104 minas de um grande depósito de zinco. Com o auxílio de métodos de simulação geoestatística e regressão linear multivariada foram elaborados modelos da incerteza geometúrgica, fornecendo previsões in situ sobre os riscos associados à recuperação de massa nos blocos de lavra. O processo é laborioso e dispendioso em termos computacionais, mas os resultados da geometurgia provaram ser úteis para o planejamento de mineração e otimização metalúrgica. A metodologia é ilustrada usando uma operação principal de zinco subterrâneo.

**Palavras-chaves:** geometurgia, regressão multivariada, simulação geoestatística.



# Abstract

The metal recovery at the metallurgical plant is the most important factor for the economic success of a mine. Plant mass recovery is usually associated with head grades, however mineral processing circuits are complex and depend on the interaction of a variety of rock characteristics. As grades of most mines are declining pushes geometallurgy importance to help in reducing production risks. Geometallurgical tests were realized using ore from 104 mined stopes from a major Zinc deposit. Additionally, geostatistical simulations and multivariate linear regression were applied to build uncertainty models of concentrate mass recovery , providing in situ predictions about the risks associated with mass recovery at every mining block. The process is laborious and computational expensive, but the geometallurgy results proved to be useful for mining planning and metallurgical optimization. The methodology is illustrated using a major underground zinc operation.

**Key-words:** geometallurgy, multivariate linear regression, geostatistical simulation.



# Lista de ilustrações

Figura 1 – Disciplinas envolvidas no estudo metalúrgico. . . . .	24
Figura 2 – Modelo de regressão. A variável dependente $\mathbf{Y}$ pode ser obtida por meio de uma função da variável independente $\mathbf{X}$ . . . . .	27
Figura 3 – Distribuição normal e seu intervalo de confiança $1 - \alpha$ . . . . .	30
Figura 4 – Em i e ii estão ilustrados exemplos de situações em que a hipótese $H_0$ não é rejeitada e portanto não há relação linear entre as variáveis. Em iii e iv estão exemplificadas situações em que $H_0$ é rejeitada pode indicar a existência de uma correlação linear. . . . .	33
Figura 5 – Exemplo de diagrama de tornado. As variáveis são ranqueadas de acordo com o módulo do valor de seu coeficiente de sensibilidade. . . . .	37
Figura 6 – Diferenças entre o modelo krigado e os modelos simulados. . . . .	38
Figura 7 – Simulações por bandas rotativas. Em [i] utilizou-se apenas 15 linhas e em [ii] foram utilizadas 1000 linhas. . . . .	42
Figura 8 – . Simulação do ponto x pelo método de bandas rotativas em 2D. . . . .	43
Figura 9 – Coluna estratigráfica do Grupo Vazante. Em destaque unidades presentes na região. . . . .	46
Figura 10 – Corte transversal da zona de minério Vazante mostrando as relações espaciais entre os metabasaltos, a sequencia hospedeira, o minério willemítico e a zona de cisalhamento Vazante. . . . .	48
Figura 11 – Ilustração representativa da definição de amostras de testemunhos para realização dos testes metalúrgicos. . . . .	49
Figura 12 – Histograma e informações estatísticas da recuperação metalúrgica do Zinco, em porcentagem nos testes. . . . .	49
Figura 13 – Histograma dos teores médios de alimentação do Zinco nos testes metalúrgicos. . . . .	50
Figura 14 – Distribuição das litologias presentes nos 104 testes metalúrgicos. . . . .	51
Figura 15 – Distribuição e localização das amostras contendo análise química para o Zinco. A extensão da mina na direção $N50E$ é de aproximadamente 5,500 metros. . . . .	52
Figura 16 – Histograma da variável Zinco. . . . .	53
Figura 17 – Gráfico de probabilidade ilustrando a probabilidade dos teores de zinco contemplando todos os domínios geometalúrgicos. . . . .	54
Figura 18 – Histograma da distribuição das litologias reclassificadas de acordo com a classificação das rochas utilizadas no teste de bancada. . . . .	55
Figura 19 – Histograma da distribuição das litologias reclassificadas de acordo com a classificação das rochas utilizadas no teste de bancada. . . . .	56

Figura 20 – Correlograma não ergódico do banco de dados original do Zinco. . . . .	57
Figura 21 – Correlograma não ergódico do banco de dados original da litologia Brecha Dolomítica. . . . .	58
Figura 22 – Matriz de correlação das variáveis do teste metalúrgico. . . . .	59
Figura 23 – Matriz de correlação das variáveis do teste metalúrgico. . . . .	60
Figura 24 – Histograma dos dados do Zinco desagrupados e normalizados. . . . .	61
Figura 25 – Correlograma não ergódico dos dados desagrupados e normalizados pelo método de <i>normal score</i> . . . . .	62
Figura 26 – Grid rotacionado, reduzindo assim o número de blocos e diminuído o tempo de processamento. . . . .	64
Figura 27 – Número de realizações versus o desvio padrão das médias das simulações de cada uma das variáveis. O diagrama (a) ilustra a curva de dispersão do desvio padrão para a variável Zinco e em (b), o diagrama da dispersão para o litotipo BXD. . . . .	64
Figura 28 – Em [i] está ilustrado o sólido da envoltória dos blocos simulados. Em [ii] a seção do modelo de blocos do primeiro cenário gerado pela simulação do zinco. . . . .	65
Figura 29 – Validação da reprodução da distribuição dos teores do Zinco. Em preto estão representadas as distribuições das realizações e em vermelho o a distribuição desagrupada do Zinco. . . . .	66
Figura 30 – Validação da continuidade espacial dos cenários simulados. Em (A) o variograma da direção de maior continuidade [1, 1, 0], em (B) a direção de média continuidade [-1, 1, 0] e em (C) a direção de menor continuidade [-1, 0, -1]. . . . .	67
Figura 31 – Em [i] está ilustrado a envoltória dos blocos simulados e em [ii] está ilustrado a seção A-A' realizada no modelo de blocos do cenário número 1. . . . .	68
Figura 32 – Validação dos cenários simulados da variável BXD. As simulações mantiveram as proporções litológicas desagrupadas. Em preto, os 50 cenários. Em vermelho, a proporção desagrupada pelo método vizinho mais próximo. . . . .	69
Figura 33 – Validação da continuidade espacial das realizações medida pelo correlogramas não ergódico. . . . .	70
Figura 34 – Em [i] está ilustrado a envoltória dos blocos simulados e em [ii] o modelo do primeiro cenário da recuperação metalúrgica do Zinco obtido pela regressão multivariada. . . . .	71
Figura 35 – Gráfico de probabilidade ilustrando o teste de normalidade do erro. O valor-p superior à 0.05 confirma a suposição de normalidade pelo teste de Anderson Darling. . . . .	72

Figura 36 – Ilustração da normalidade do erro da regressão linear no caso de um modelo univariado. O $\mu$ representa o valor médio da recuperação de zinco obtida pelo modelo de regressão linear. . . . .	73
Figura 37 – Nova distribuição normal gerado pelos valores da média da recuperação metalúrgica e erro de regressão calculado. O círculo verde ilustra o valor sorteado aleatoriamente dentro do intervalo P90 da distribuição. . . . .	74
Figura 38 – Validação dos modelos regredidos pela comparação com os dados laboratoriais. Os gráficos de caixas representam as faixas incertezas das variáveis de entrada adicionadas à incerteza do modelo de regressão para um dos estopes. Os pontos em vermelho representam os testes de bancadas aos quais o modelo geometalúrgico foram aprovado, e cada um dos pontos em azul representam os reprovados. . . . .	75



# Lista de tabelas

Tabela 1 – Tabela de resultados da ANOVA . . . . .	35
Tabela 2 – Análise de variância (ANOVA) para testar a significância da regressão.	60
Tabela 3 – Resultados obtidos pela divisão do madograma pela raiz do variograma da variável normalizada do Zinco. . . . .	63



# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>23</b>
<b>1.1</b>	<b>Meta desta dissertação</b>	<b>25</b>
<b>1.2</b>	<b>Objetivos específicos</b>	<b>25</b>
<b>1.3</b>	<b>Metodologia</b>	<b>25</b>
<b>1.4</b>	<b>Organização da dissertação</b>	<b>26</b>
<b>2</b>	<b>REVISÃO BIBLIOGRÁFICA</b>	<b>27</b>
<b>2.1</b>	<b>Análise de regressão</b>	<b>27</b>
2.1.1	Princípios básicos da regressão linear	27
2.1.2	Correlação	28
2.1.3	Regressão linear simples	29
2.1.4	Regressão linear multivariada	29
<b>2.2</b>	<b>Intervalo de confiança</b>	<b>30</b>
2.2.1	Intervalo de confiança da regressão simples	30
2.2.2	Intervalo de confiança da regressão multivariada	31
<b>2.3</b>	<b>Validação da regressão</b>	<b>32</b>
2.3.1	Teste de hipótese	32
2.3.2	Análise de variância - ANOVA	34
2.3.3	Coeficiente de determinação ( $R^2$ )	35
2.3.4	Análise sensibilidade	35
<b>2.4</b>	<b>Simulação estocástica</b>	<b>37</b>
2.4.1	Simulação sequencial gaussiana (ISAAKS, 1990)	38
2.4.2	Simulação sequencial dos indicadores (ALABERT, 1987)	40
2.4.3	Simulação por bandas rotativas ( <i>Turning Bands</i> )	41
2.4.4	Mudança de suporte	43
<b>2.5</b>	<b>Observações finais</b>	<b>44</b>
<b>3</b>	<b>ESTUDO DE CASO</b>	<b>45</b>
<b>3.1</b>	<b>Geologia Local</b>	<b>45</b>
<b>3.2</b>	<b>Análise exploratória dos dados</b>	<b>48</b>
3.2.1	Banco de dados geometalúrgico	48
3.2.2	Análise exploratória dos dados de sondagem	51
3.2.2.1	Modelo de continuidade espacial do Zinco	56
3.2.2.2	Modelo de continuidade espacial da litologia BXD	57
<b>3.3</b>	<b>Análise estatística multivariada</b>	<b>58</b>
<b>3.4</b>	<b>Aplicação dos métodos de simulação</b>	<b>60</b>

3.4.1	Transformação gaussiana variável Zinco . . . . .	61
3.4.2	Multigaussianidade . . . . .	62
3.4.3	<i>Grid</i> . . . . .	63
<b>3.5</b>	<b>Simulação dos teores Zinco por bandas rotativas . . . . .</b>	<b>64</b>
3.5.1	Validação das realizações . . . . .	66
<b>3.6</b>	<b>Simulação do litotipo BXD por indicadores . . . . .</b>	<b>67</b>
3.6.1	Validação das realizações . . . . .	68
3.6.2	Tratamento dos modelos simulados . . . . .	70
<b>3.7</b>	<b>Modelagem da recuperação metalúrgica por regressão . . . . .</b>	<b>71</b>
<b>3.8</b>	<b>Validação do modelo geometalúrgico . . . . .</b>	<b>72</b>
<b>3.9</b>	<b>Obtenção do modelo de incertezas da recuperação metalúrgica para novos estopes . . . . .</b>	<b>76</b>
<b>4</b>	<b>CONSIDERAÇÕES FINAIS . . . . .</b>	<b>77</b>
4.1	Recomendações para trabalhos futuros . . . . .	78
	<b>REFERÊNCIAS . . . . .</b>	<b>79</b>

# 1 Introdução

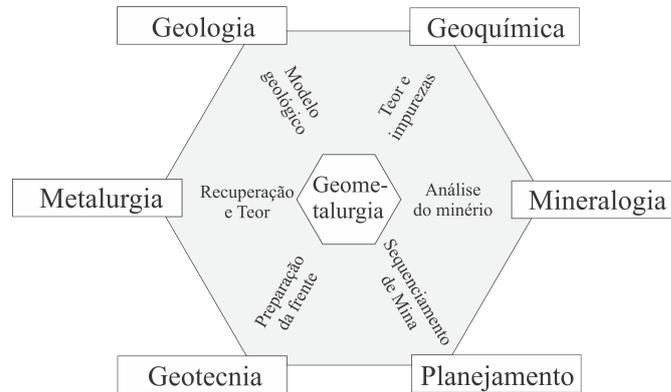
Os modelos geometalúrgicos são considerados hoje uma importante ferramenta para a indústria mineral. Combina informações de diversas áreas de conhecimentos (Figura 1), é possível criar previsões geometalúrgicas para o processamento mineral. A redução contínua nos teores das minas em operação é um dos fatores que colocou em evidência a importância do controle geometalúrgico (LAMBERG, 2011). Teor de rejeito baixo significam menores perdas durante o processo de beneficiamento, podendo ser um importante diferencial na vida útil de muitas operações.

Os corpos de minério são um fenômeno natural que ocorrem com muita similaridade entre diferentes tipos de mineralização, porém com pouca consistência. Ocorrências de um mesmo domínio metálico em diferentes localidades podem apresentar sutis diferenças nas características litoquímicas e mineralógica. Estas pequenas alterações são, muitas vezes, capazes de alterar o comportamento do minério na planta metalúrgica, podendo causar grandes prejuízos ou até mesmo inviabilizar um projeto (ADAMS, 2016; LAMBERG, 2011).

Além de ser um empreendimento com riscos, a mineração ainda requer um investimento financeiro substancial com retorno de longo prazo. Portanto, é essencialmente importante conhecer os riscos inerentes ao investimento e tentar minimizá-los. Um dos fatores de risco citados pela literatura, é o *Risco Técnico*, que envolve etapas desde a avaliação e desenvolvimento do projeto até o planejamento da mina e a performance do beneficiamento metalúrgico, sendo a performance da planta o maior responsável pela variabilidade na massa final do metal extraído. Para reduzir estes riscos, as mineradoras vêm utilizando a geometalurgia como forma de entender e medir o comportamento do minério (*in situ*) e sua resposta na planta metalúrgica. O conhecimentos sobre a variabilidade espacial do fenômeno mineralizante e suas características petrofísicas/petroquímicas são de suma importância para a modelagem geometalúrgica (ADAMS, 2016).

Os estudos geometalúrgicos inciam com a amostragem e testes de bancada retiradas de testemunhos de sondagem. Primeiramente, é realizado um estudo geológico, o qual permitirá a identificação de melhores áreas para amostragem, como uma litologia que historicamente é conhecida por apresentar problemas no beneficiamento. As amostras são então submetidas aos testes metalúrgicos laboratoriais que permitirão a identificação dos domínios geometalúrgicos. Os dados são tratados e submetidos a análises estatísticas multivariadas, estabelecendo relações matemáticas importantes para predição do modelo. Definidas as variáveis de interesse, confecciona-se então o modelo geometalúrgico (LAMBERG, 2011; BOISVERT et al., 2013).

Figura 1 – Disciplinas envolvidas no estudo metalúrgico.



Modificado de [Williams e Richardson \(2004\)](#).

Os testes de bancada normalmente incluem informações metalúrgicas tais como: o rendimento da planta, a recuperação do metal contido, energia necessária (índices de cominuição) e o consumo de reagentes ([DEUTSCH, 2015](#)). O resultado destes testes causam problemas para a modelagem geoestatística convencional, pois além do comportamento não aditivo de algumas variáveis, estas podem apresentar relações multivariadas complexas. Algumas variáveis podem apresentar também redundância de informação, gerando ruídos e reduzindo a qualidade do modelo.

[Deutsch \(2013\)](#) e [Lamberg \(2011\)](#) citam a dificuldade de relacionar os testes metalúrgicos com as variáveis comumente analisadas pela indústria mineira. O fato dos testes de bancada serem muito caros ocorre normalmente uma subamostragem, gerando dados dispersos no espaço. Outra dificuldade está associada ao suporte amostral. As análises metalúrgicas exigem uma quantidade de massa muito grande em relação as massas necessárias pra análises de teores, por exemplo.

Uma alternativa plausível para prever o comportamento do minério (*in situ*) é a confecção de um modelo preditivo geometalúrgico. Devido a não aditividade de algumas das variáveis, recorre-se ao modelo de maneira indireta, aplicando-se técnicas de regressão às propriedades físico-químicas de comportamento aditivos. Neste trabalho, pretende-se acessar a incerteza da recuperação metalúrgica aplicando técnicas de simulação geoestatística e regressão linear multivariada.

Recorreu-se aos métodos de simulação visando reproduzir a variabilidade e a continuidade espacial do conjunto de dados de entrada. Os cenários simulados são modelos igualmente prováveis, permitindo o acesso às incertezas locais atribuídas pela incerteza das variáveis de entrada ([ROSSI; DEUTSCH, 2013](#)).

As informações utilizadas neste trabalho provém de uma importante mina de zinco

brasileira. O estudo foi motivado pelos problemas enfrentados na recuperação do metal em planta. Das diversas variáveis obtidas pelos testes de bancada, este trabalho limitou-se somente ao estudo da recuperação mássica do Zinco.

A evolução do processamento computacional contribuiu para a aplicação dos métodos multivariados, proporcionando a criação de modelos de regressão em campos  $n$  dimensionais.

## 1.1 Meta desta dissertação

A partir de técnicas estatísticas multivariada e simulação estocástica, deseja-se mensurar o impacto das incertezas das variáveis no modelo de recuperação do Zinco e avaliar a qualidade das informações previamente geradas.

## 1.2 Objetivos específicos

Por meio de um estudo de caso em um importante depósito de Zinco, esta dissertação propõe atingir a sua meta cumprindo os seguintes objetivos específicos:

- Identificar por meio de regressão multivariada as características do minério que melhor explicam a incerteza da recuperação do Zinco em planta;
- Aplicar métodos de simulação *in situ* das variáveis que afetam o comportamento do minério na planta de beneficiamento;
- Avaliar os resultados obtidos pela regressão múltipla das simulações, comparando-as aos resultados reais medidos em blocos lavrados.

## 1.3 Metodologia

A metodologia deste trabalho inicia com a obtenção de um modelo de regressão para a recuperação metalúrgica, seguido pela aplicação de simulação sequencial para as variáveis explicativas da recuperação do Zinco e aplicação da regressão multivariada aos modelos simulados. Ao final, será realizado uma reconciliação com os dados de recuperação medidos no estopes já lavrados.

Os seguintes itens enumerados resumem a metodologia deste trabalho:

1. Definição das variáveis geológicas que melhor expliquem a recuperação metalúrgica a partir da análise de sensibilidade e correlação entre elas. Posterior modelagem da equação de regressão multivariada com as variáveis explicativas selecionadas;

2. Aplicação da simulação geoestatística às variáveis explicativas identificadas pelo passo anterior;
3. Aplicação da regressão multivariada nos nós do *grid* simulado usando os valores das realizações das variáveis explicativas previamente simuladas;
4. Validar os modelos geometalúrgicos através de uma reconciliação com os dados de estopes já lavrados.

## 1.4 Organização da dissertação

Os capítulos desta dissertação estão organizados conforme a sequência:

O Capítulo 1 realiza uma breve introdução a geometalúrgia, abordando suas principais aplicações, a importância para a indústria mineral e desafios para geração desses modelos. Em seguida, são apresentadas as metas, os objetivos específicos e a sequência metodológica adotada neste trabalho.

No Capítulo 2, são apresentados os fundamentos teóricos necessários para compreensão dos métodos estatísticos e geoestatísticos associados à modelagem geometalúrgica.

O Capítulo 3 contempla o estudo de caso, onde serão aplicadas as técnicas para a obtenção do modelo geometalúrgico.

No Capítulo 4, estão contidas as conclusões do trabalho e recomendações para futuras pesquisas.

## 2 Revisão bibliográfica

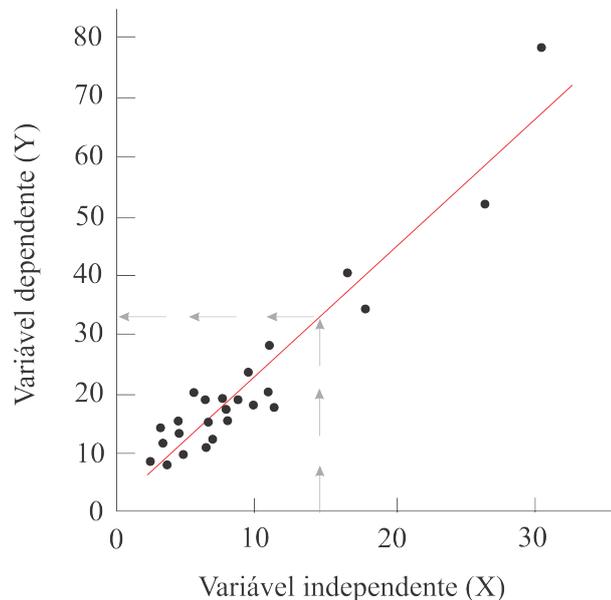
Este capítulo trata de uma breve revisão teórica de métodos estatísticos, estimativa de teores em bancos de dados multivariados e métodos de simulação geoestatística.

### 2.1 Análise de regressão

#### 2.1.1 Princípios básicos da regressão linear

A regressão linear é uma metodologia empregada na modelagem de relações entre variáveis, podendo ser utilizada na predição de uma variável dependente  $Y$  por meio de uma ou mais variáveis independentes  $X$ . Uma variável independente é entendida como aquela que possui o valor fixo, independe de qualquer evento externo. A variável dependente é a variável que pode ser explicada por uma ou mais variáveis independentes através de uma função ( $f(x)$ ) (Figura 2) (KUTNER; NACHTSHEIM; NETER, 2004; HAIR et al., 2009; MONTGOMERY; PECK; VINING, 2012).

Figura 2 – Modelo de regressão. A variável dependente  $Y$  pode ser obtida por meio de uma função da variável independente  $X$ .



Duas vertentes dividem os métodos de análise de regressão, a simples e a múltipla. A regressão linear simples é caracterizada por apresentar somente uma variável explicativa independente, e o modelo de regressão multivariado é caracterizada pela relação casual da

variável dependente  $Y$  com duas ou mais variáveis independentes  $X_1, X_2, \dots, X_n$  (MONTGOMERY; PECK; VINING, 2012). Neste trabalho, a análise de regressão será utilizada na obtenção indireta do modelo de recuperação do zinco, utilizando simulações de variáveis explicativas da variável de entrada.

### 2.1.2 Correlação

Triola (2008) cita que a correlação entre duas variáveis existe quando os valores de uma variável são de alguma forma associados aos valores da outra variável. Denotado pelo símbolo  $\rho$ , o coeficiente de correlação é a primeira abordagem utilizada para verificar a relação entre as variáveis dependente e independente. O coeficiente de correlação sempre retorna um valor entre -1 e 1, sendo que, quando o coeficiente é 0 (zero) diz-se que as variáveis em análise são "não correlacionadas". A Equação 2.1 ilustra o cálculo da correlação ( $\rho$ ).

$$\rho(X, Y) = \frac{C_{i,j}(x, y)}{\sqrt{\sigma^2(x) * \sigma^2(y)}} \quad (2.1)$$

Em que a covariância entre as variáveis  $\mathbf{X}$  e  $\mathbf{Y}$  ( $C_{i,j}(X, Y)$ ) pode ser obtida pela Equação 2.2:

$$C_{i,j}(x, y) = E[(x - E[x]) * (y - E[y])] \quad (2.2)$$

As variâncias de  $\mathbf{X}$  ( $\sigma^2(x)$ ) de  $\mathbf{Y}$  ( $\sigma^2(y)$ ) são obtidas pelas Equações 2.3 e 2.4:

$$\sigma^2(x) = E[(x - E[x])^2] \quad (2.3)$$

$$\sigma^2(y) = E[(y - E[y])^2] \quad (2.4)$$

A esperança matemática (E) corresponde a média dos dados de uma variável aleatória. Se duas variáveis aleatórias  $X$  e  $Y$  são independentes, então  $X$  e  $Y$  não estão correlacionados e portanto, o coeficiente de correlação  $\rho(X, Y)$  será 0. Outro cenário em que  $\rho(X, Y) = 0$ , ocorre caso as variância de  $X$  ( $\sigma^2(X)$ ) ou  $Y$  ( $\sigma^2(Y)$ ) tenham valor zero (DEKKING, 2005).

Estudos de correlação foram realizados entre as variáveis do banco de dados metalúrgicos a fim de identificar as propriedades petrofísicas e petroquímicas que apresentassem correlação com o valores de recuperação do Zinco.

### 2.1.3 Regressão linear simples

O modelo de regressão linear simples apresenta somente uma variável regressora e uma variável resposta. A Equação 2.5 representa o modelo de regressão linear simples.

$$Y = b_0 + b_1X + \varepsilon \quad (2.5)$$

Em que:

$b_0$  é o termo independente da equação, o intercepto;

$b_1$  é o coeficiente angular;

$\varepsilon$  é um componente do erro aleatório.

Os coeficientes  $b_1$  e  $b_0$  são obtidos pelo método dos mínimos quadrados, representados pelas Equações 2.6 e 2.7. Esse método estima os coeficientes de modo a encontrar o melhor ajuste para um conjunto de dados, tentando minimizar a diferença entre a soma dos quadrados dos valor estimado e dos dados observados (MONTGOMERY; PECK; VINING, 2012).

$$b_0 = \bar{Y} - b_1\bar{X} \quad (2.6)$$

$$b_1 = \frac{\sum(x_i - \bar{X})(y_i - \bar{Y})}{\sum(x_i - \bar{x})^2} \quad (2.7)$$

Um modelo de regressão é denominado multivariado quando esse apresenta mais de uma variável independente  $\mathbf{X}$ .

### 2.1.4 Regressão linear multivariada

O modelo de regressão que envolve mais de uma variável de regressão é definido como Modelo de Regressão Multivariada (MONTGOMERY; PECK; VINING, 2012). A Equação 2.8 ilustra um modelo multivariado apresentando  $k$  variáveis regressoras:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k + \varepsilon \quad (2.8)$$

$Y$  é a variável dependente,

$b_j, j = 0, 1, 2, \dots, k$  são os coeficientes de regressão,

$X_j, j = 1, 2, 3, \dots, k$  são as variáveis independentes,

$\varepsilon$  é um componente do erro aleatório.

O modelo multivariado é mais complexo e exige um maior poder computacional no processamento da regressão em  $n$  dimensões, sendo  $n$  o número de variáveis envolvidas.

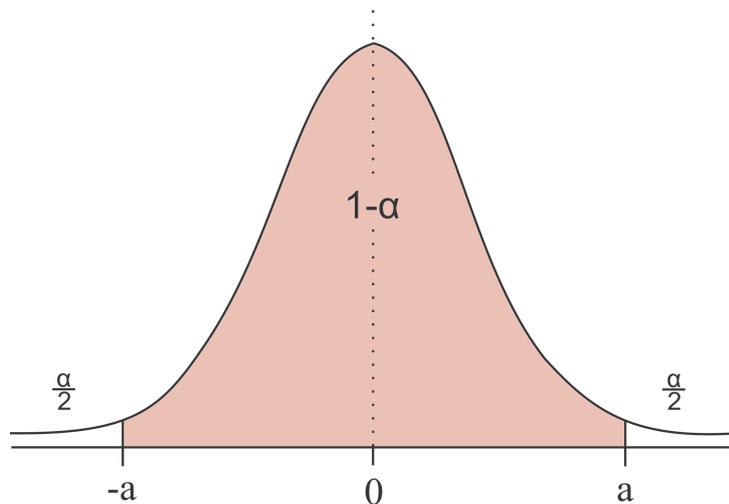
A grande vantagem deste método é a maximização da qualidade da predição da variável dependente (HAIR et al., 2009).

## 2.2 Intervalo de confiança

O intervalo de confiança é definido como o intervalo que possui  $100(1 - \alpha)\%$  de probabilidade de ocorrer, ou seja, é a probabilidade de  $100(1 - \alpha)\%$  de se encontrar uma amostra da população no intervalo (Figura 3). O intervalo está relacionado também com a qualidade global do modelo de regressão, sendo o parâmetro  $\alpha$  correspondente a soma das probabilidades inferior e superior do intervalo de confiança.

Para um número amostral  $n$  suficientemente grande, uma média amostral  $\mu$  segue uma distribuição que se aproxima da Gaussiana. Deve-se considerar que o erro também é independente e normalmente distribuído, apresentando média 0 e variância  $\sigma^2$  (CUNHA; CARVAJAL, 2009; MONTGOMERY; PECK; VINING, 2012; COX; HINKLEY, 1979)

Figura 3 – Distribuição normal e seu intervalo de confiança  $1 - \alpha$ .



O intervalo de confiança será utilizado na validação dos modelos de incerteza da recuperação metalúrgica do Zinco obtidos pela regressão. A análise do intervalo de confiança compreende duas vertentes, o intervalo unidimensional e o multidimensional.

### 2.2.1 Intervalo de confiança da regressão simples

O principal uso do modelo de regressão é na obtenção de valores médios para a variável resposta ( $Y$ ) em função da variável regressora ( $X$ ). O intervalo de confiança  $100(1 -$

$\alpha$ )%, para uma variável estimada, pode ser obtido pela Equação 2.9 (MONTGOMERY; PECK; VINING, 2012; CUNHA; CARVAJAL, 2009).

$$\hat{Y}_0 \pm t_{\alpha/2, n-2} = \sqrt{\hat{\sigma}^2 \left[ 1 + \frac{1}{n} + \frac{(x_0 - \bar{X})^2}{\sum_{i=1}^n (x_i - \bar{X})^2} \right]} \quad (2.9)$$

$\bar{X}$  é a média amostral,

$t_{\frac{\alpha}{2}, n-2}$  é o valor da tabela t-student com n-2 graus de liberdade e

$\hat{\sigma}^2$  é o estimador da variância.

### 2.2.2 Intervalo de confiança da regressão multivariada

Assim como na regressão simples, a análise do intervalo de confiança da regressão multivariada exerce um importante papel. Para definir o intervalo de confiança da regressão linear múltipla, mantém-se a premissa de que a distribuição do erro é normal e independentemente distribuída com média zero e variância 1.

O intervalo de confiança para um modelo de regressão multivariado pode ser obtido pela Equação 2.10. Maiores detalhes sobre a obtenção do intervalo de confiança multivariada pode ser encontrada na página 506 de [Montgomery e Runger \(2010\)](#).

$$\hat{Y}_0 \pm t_{\alpha/2, n-2} = \sqrt{\hat{\sigma}^2 (1 + x_0^T (X^T X)^{-1} x_0)} \quad (2.10)$$

Em que  $x_0$  é o vetor

$$x_0 = \begin{bmatrix} 1 \\ x_{01} \\ x_{02} \\ \cdot \\ \cdot \\ \cdot \\ x_{0k} \end{bmatrix}$$

e  $X^T X$  é a seguinte matriz.

$$X^T X = \begin{bmatrix} n & \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i2} & \cdots & \sum_{i=1}^n x_{ik} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & \sum_{i=1}^n x_{i1}x_{i2} & \cdots & \sum_{i=1}^n x_{i1}x_{ik} \\ \sum_{i=1}^n x_{i2} & \sum_{i=1}^n x_{i2}x_{i1} & \sum_{i=1}^n x_{i2}^2 & \cdots & \sum_{i=1}^n x_{i2}x_{ik} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \sum_{i=1}^n x_{ik} & \sum_{i=1}^n x_{ik}x_{i1} & \sum_{i=1}^n x_{ik}x_{i2} & \cdots & \sum_{i=1}^n x_{ik}^2 \end{bmatrix}$$

Sendo  $n$  o número total de observações e  $x_k$ ,  $k = 1, 2, \dots, K$  as  $K$  variáveis independentes  $X$ .

O intervalo de confiança será utilizada na obtenção do erro do modelo de regressão múltipla. O erro será considerado no valor final da simulação do modelo da recuperação do Zinco.

## 2.3 Validação da regressão

O modelo de regressão deve passar por avaliações de adequação, que consistem em identificar potenciais erros e falhas na entrada das variáveis, ou na inclusão indevida de variáveis redundantes. [Montgomery, Peck e Vining \(2012\)](#) citam que a validação é uma etapa essencial para assegurar a confiabilidade do modelo de regressão. Nesta seção, serão apresentados alguns métodos utilizados na validação do modelo de regressão da variável geometalúrgica.

### 2.3.1 Teste de hipótese

Os testes de hipótese permitem provar uma afirmação feita sobre alguma característica de uma variável aleatória qualquer. O teste de hipótese é um procedimento de decisão que permite escolher entre duas afirmações mutuamente excludentes: Hipótese Nula ( $H_0$ ) e a Hipótese Alternativa ( $H_1$ ) ([PINHEIRO et al., 2013](#); [TRIOLA, 2008](#)). A hipótese  $H_0$  sempre será tratada como uma hipótese simples, já a hipótese alternativa  $H_1$  é, em geral, uma hipótese composta apresentando mais de uma distribuição de probabilidade. As decisões tomadas no teste são baseadas em informações parciais aleatoriamente amostradas, portanto, há dois tipos possíveis de erro de decisão:

- Erro Tipo I: Rejeitar  $H_0$ , quando  $H_0$  é verdadeira;
- Erro Tipo II - Aceitar  $H_0$ , quando  $H_0$  é falsa.

O teste de hipótese também pode ser utilizado na validação de um modelo de regressão linear, sugerindo as seguintes hipóteses para o coeficiente angular ( $\beta_1$ ):

$$H_0 : \beta_1 = 0; H_1 : \beta_1 \neq 0.$$

Se a hipótese nula ( $H_0$ ) for rejeitada, conclui-se que a variável independente  $X$  é capaz de explicar a variância na variável dependente  $Y$ . Se a hipótese  $H_0$  for aceita, conclui-se que não há relação linear entre as variáveis. Um dos procedimentos utilizados para realizar o teste de hipótese é a estatística  $t$ , utilizando o valor  $t_{crítico}$  da tabela de *Student* para comparar os resultados de  $T$  calculados. O teste  $T$  ( $T_0$ ) para  $\beta_0$  pode ser

obtido pela Equação 2.11 e  $T_0$  para  $\beta_1$  é calculado pela Equação 2.12 (MONTGOMERY; PECK; VINING, 2012).

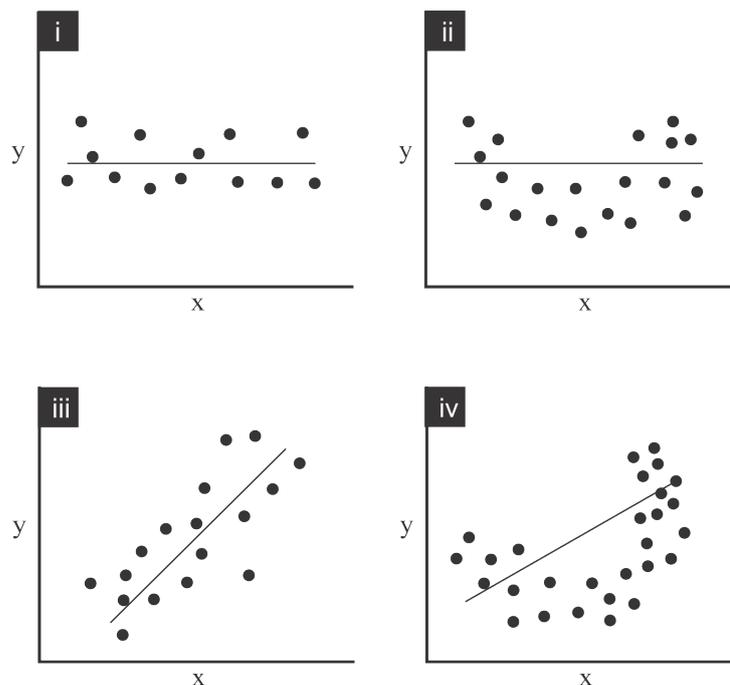
$$T_0 = \frac{\hat{\beta}_0}{\sqrt{QME \left( \frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^n (x_i - \bar{X})^2} \right)}} \quad (2.11)$$

$$T_0 = \frac{\hat{\beta}_1}{\sqrt{\frac{QME}{\sum_{i=1}^n (x_i - \bar{X})^2}}} \quad (2.12)$$

Em que  $\beta_1$  é o declive da reta de regressão,  $\beta_0$  é o intercepto,  $\bar{X}$  é a média amostral,  $x_i$  é cada termo da variável X e QME é o quadrado médio dos erros.

Em ambos os casos,  $H_0$  será rejeitado com um nível de confiança de  $(1 - \alpha)100\%$  caso o valor medido de  $T_0$  seja maior do que o valor de  $t_{critico}$  na tabela de *Student*. Na Figura 4, os diagramas **i** e **ii** ilustram uma situações em que não existe relação linear entre as variáveis X e Y, e portanto é mais provável que a hipótese nula não seja rejeitada. Já os diagrama **iii** e **iv** ilustram situações em que a a hipótese nula seria rejeitada, confirmando assim uma relação explicativa da variável X na Y.

Figura 4 – Em i e ii estão ilustrados exemplos de situações em que a hipótese  $H_0$  não é rejeitado e portanto não há relação linear entre as variáveis. Em iii e iv estão exemplificadas situações em que  $H_0$  é rejeitada pode indicar a existência de uma correlação linear.



Modificado de Montgomery, Peck e Vining (2012)

### 2.3.2 Análise de variância - ANOVA

A análise de variância (ANOVA) testa a hipótese de que duas ou mais populações são iguais, é um meio de avaliar a importância dos fatores nas médias da variável resposta. A hipótese nula da ANOVA afirma que todas as médias de população são iguais, enquanto a hipótese alternativa afirma que pelo menos uma é diferente. O cálculo da ANOVA está ilustrado pela Equação 2.13 (MONTGOMERY; PECK; VINING, 2012; TRIOLA, 2008).

$$\sum_{i=1}^n (y_i - \bar{Y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{Y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.13)$$

Em que:

$y_i$  é cada termo presente nas amostras;

$\hat{y}_i$  é a média dentro de cada um dos grupos e;

$\bar{Y}$  é a média total, podendo ser obtida pela média das médias dos grupos.

A Equação 2.13 pode ser expressa simbolicamente pela equação 2.14. O parâmetro SQT representa a soma dos quadrados das variações das populações, SQR é a soma dos quadrados das amostras de cada grupo, e o SQG é a soma dos quadrados entre os grupos.

$$SQT = SQG + SQR \quad (2.14)$$

A ANOVA também é utilizada nos testes de significância da regressão. No modelo de regressão, o parâmetro SQR pode ser entendido como o resíduo, e o SQG reporta o quanto da variável dependente pode ser explicada pelo modelo de regressão. O teste  $F$  é o teste de hipótese utilizado na ANOVA do modelo de regressão linear, rejeitando ou aceitando a existência de uma relação entre as variáveis. Consideramos as hipótese nula ( $H_0$ ) ou a hipótese alternativa ( $H_1$ ) a seguir:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \beta_j \neq 0 \quad \forall \quad j = 1, \dots, p$$

Se a hipótese  $H_0$  for rejeitada, conclui-se que no modelo de regressão pelo menos uma variável independente apresenta influência significativa na variável dependente (Y). Como estratégia de simplificação da notação, o cálculo de  $F_{medido}$  pode ser expresso pelos termos simbólicos da ANOVA (Equação 2.15).

$$F_{medido} = \frac{\frac{SQG}{p}}{\frac{SQR}{(n-p)}} \quad (2.15)$$

Em que  $p$  são os graus de liberdade da soma dos quadrados entre os grupos e  $n$  são os graus de liberdade totais.

Os valores de  $F_{medido}$  e  $F_{crítico}$  devem ser comparados, respeitando o nível de significância e os graus de liberdade da análise de variância. As seguintes considerações são feitas para validar o teste de hipótese:

- Caso  $F_{medido} \geq F_{crítico}$ , a hipótese  $H_0$  é rejeitada no nível de significância considerada para o fenômeno e então o modelo de regressão proposto está adequado;
- Caso  $F_{medido} < F_{crítico}$ ,  $H_0$  não é rejeitado no nível de significância e então o modelo não explica o fenômeno.

A Tabela 1 ilustra uma análise de variância (ANOVA), preenchida com os valores obtidos pela Equação 2.14.

Tabela 1 – Tabela de resultados da ANOVA

Fonte de Variação	Soma dos quadrados	Graus de liberdade	Média dos quadrados	Teste F
Entre Grupos	SQG	k-1	MQG	MQG/MQR
Dentro Grupos	SQR	n-k	MQR	
Total	SQT	n-1		

A tabela da ANOVA será fundamental para a análise dos dados realizada pelo programa SABOR (*Sensitivity Analysis Based On Regression*), presente no estudo de caso deste trabalho.

### 2.3.3 Coeficiente de determinação ( $R^2$ )

O coeficiente de determinação é a medida de ajustamento de um modelo estatístico linear generalizado, como uma regressão linear, por exemplo. Com ele é possível avaliar o quanto da proporção da variação do modelo de regressão pode ser explicada pela variável regressora (X). O coeficiente pode ser obtido pela razão expressa na Equação 2.16.

$$R^2 = \frac{SQG}{SQT} \quad (2.16)$$

O resultado desta razão é um valor entre 0 e 1. Valores de  $R^2$  próximos 1 implicam que a maior parte da variabilidade em Y é explicada pelo modelo de regressão (MONTGOMERY; PECK; VINING, 2012).

### 2.3.4 Análise sensibilidade

A análise de sensibilidade aborda a incerteza de um modelo (numérico ou não), permitindo a identificação e a divisão desta incerteza entre as diferentes variáveis de

entrada. Além da quantificação, ela ainda exhibe a proporção da incerteza da variável que é capaz de explicar a incerteza total do modelo, ou seja, quantificar a influência dos valores de entrada no modelo final regredido (SALTELLI et al., 2004; ZAGAYEVSKIY; DEUTSCH, 2011).

A análise de sensibilidade fornece informações ao modelador quanto ao comportamento da equação de regressão, ajudando a escolher corretamente as variáveis de forma a reduzir a incerteza do modelo. Petropoulos e Srivastava (2016) citam que a análise é útil para tornar os dados visíveis perante ao impacto nos resultados do modelo da regressão. O coeficiente de sensibilidade é obtido pela razão entre as derivadas parciais do modelo e a variável de entrada (Equação 2.17) .

$$\frac{\partial y}{\partial x_0} \Big|_{x_0} \quad (2.17)$$

O fato da análise de sensibilidade local examinar valores restritos de uma variável, a coloca em favoritismo quando comparada com a sensibilidade global, pois na local quase sempre existem mais informações disponíveis a respeito do valor esperado para a variável de entrada (ZAGAYEVSKIY; DEUTSCH, 2011; SALTELLI et al., 2004). As Equações 2.18 e 2.19 são provenientes da Equação 2.17 e representam os coeficientes de primeira e segunda ordem, nas quais  $x_0$  especifica o raio no qual sensibilidade local será estimada e a derivada parcial ( $\partial$ ) representa as pequenas perturbações feitas no modelo.

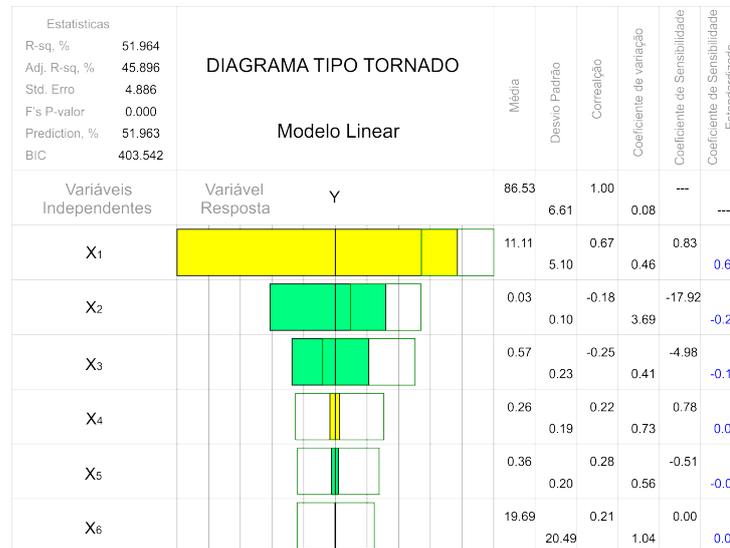
$$S_i = \frac{\partial y}{\partial x_i} \Big|_{x_k = \bar{x}_k, k=1, \dots, p, k \neq j} \quad (2.18)$$

$$S_i = \frac{\partial^2 y}{\partial x_i \partial x_j} \Big|_{x_k = \bar{x}_k, k=1, \dots, p, k \neq i, k \neq j} \quad (2.19)$$

Os resultados da análise de sensibilidade são ranqueados de acordo com o grau de influência das variáveis no modelo de regressão, auxiliando na identificação e escolha das variáveis significativas. A exclusão de algumas variáveis menos importantes não significa que o modelo ficará pior, em alguns casos ocorre uma melhora significativa. Outro ganho com a remoção de variáveis é que, com um número menor de variáveis ocorre uma simplificação do modelo final de regressão (ZAGAYEVSKIY; DEUTSCH, 2011).

Um gráfico de tornado (Figura 5) pode ser usado para ilustrar os resultados da análise de sensibilidade. O nível de relevância de uma variável é medido pelo coeficiente de sensibilidade. Quanto maior for o valor deste coeficiente, em módulo, maior a importância desta variável em relação a variável resposta.

Figura 5 – Exemplo de diagrama de tornado. As variáveis são ranqueadas de acordo com o módulo do valor de seu coeficiente de sensibilidade.



O diagrama foi obtido pelo programa SABOR de [Zagayevskiy e Deutsch \(2011\)](#)

Neste trabalho, foi utilizado o software SABOR (*Sensibility analysis based on regression*) da biblioteca do GSLIB ([DEUTSCH; JOURNAL, 1998](#)) para cálculos do coeficiente de regressão. O software ainda reporta um diagrama de tornado, com base no qual foram feitas a devidas análises e exclusões de variáveis menos relevantes.

## 2.4 Simulação estocástica

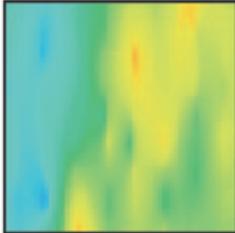
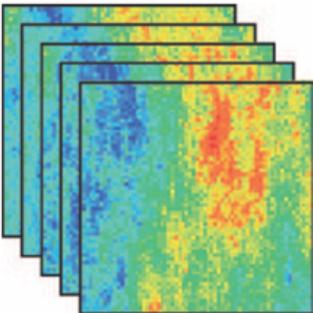
Na indústria mineral, o método de interpolação mais utilizado ainda é a krigagem ordinária (KO). Conhecida por fornecer os melhores modelos lineares não enviesados, a krigagem ordinária estima uma combinação linear dos pesos dos dados disponíveis; o modelo não é enviesado pois o estimador tenta aproximar a média residual de zero ( $m_r = 0$ ) e é considerado o melhor modelo pois a variância dos erros ( $\sigma_R^2$ ) é a menor possível. Por outro lado, a minimização da variância dos erros acaba por suavizar os resultados de maneira não uniforme, sendo esta inversamente proporcional a densidade amostral. Regiões mais densamente amostradas tendem a apresentar uma menor suavização ([GOOVAERTS, 1997; OLEA, 2012](#)).

Outro método conhecido mas pouco utilizado na indústria mineral é a simulação estocástica. Este método visa a reprodução fiel da variabilidade e distribuição de probabilidades do atributo de interesse do depósito evidenciado pelas amostras. Os métodos de simulação, assim como a KO, apresentam limitações quanto ao sua utilização ([Figura 6](#)). Os modelos gerados pela simulação, quando individualmente analisados, apresentam uma

variabilidade maior do que o modelo krigado, portanto a escolha entre a krigagem ordinária e os métodos de simulação deve ser feita com base no objetivo do trabalho de modelagem:

- caso o objetivo seja atingir o mínimo de erros de estimativa locais, a KO seria o estimador mais indicado;
- caso o objetivo seja o de modelar corretamente a continuidade espacial quantificando incertezas do modelo, o método mais indicado seria a simulação estocástica.

Figura 6 – Diferenças entre o modelo krigado e os modelos simulados.

Krigagem	Simulação
	
Versão suavizada da realidade	Imagens equiprováveis de mesma variabilidade espacial dos dados originais
Ruim para modelar valores extremos	Bom para modelar valores extremos

Modificado de Rossi e Deutsch (2013)

A simulação é uma importante ferramenta para fornecer modelos de incertezas e análises de riscos em diferentes fases de um projeto de mineração (OLEA, 2012; ROSSI; DEUTSCH, 2013). Existem diversas técnicas de simulações divididas em dois tipos: as sequenciais e as espectrais. No presente trabalho, utilizou-se técnicas de simulação sequencial dos indicadores (SIS) e simulação por bandas rotativas (*Turing Bands*).

#### 2.4.1 Simulação sequencial gaussiana (ISAKS, 1990)

O objetivo da simulação sequencial é simplesmente imitar as variações espaciais da variáveis de entrada. Os algoritmos das simulações condicionais garantem a reprodução da variabilidade e distribuição honrando as informações dos dados originais e também informações pontuais previamente simulados. O procedimento inicia-se com a normalização dos dados originais, caso estes não apresentem uma distribuição gaussiana. O processo é feito através de uma transformação não linear, associando valores dos quantis dos dados

originais desagrupados aos quantis da distribuição gaussiana, deixando a distribuição com média 0 e desvio padrão ( $\sigma = 1$ ) (GOOVAERTS, 1997; CHILES; DELFINER, 2009; ROSSI; DEUTSCH, 2013).

O algoritmo da simulação sequencial tem seu início marcado pela construção de uma distribuição acumulada (ccdf) para o primeiro nó do grid ( $u_1$ ) a partir dos dados originais do atributo. Em seguida, sorteia-se aleatoriamente um número na ccdf, sendo este o valor final atribuído ao nó  $u_1$ . No próximo nó do grid ( $u_2$ ), outra ccdf é modelada condicionada aos dados originais e também ao valor previamente simulado para o local  $u_1$ . Sorteia-se um novo número aleatório na distribuição acumulada, atribuindo este valor simulado para o local  $u_2$ . O processo ocorre sucessivamente para os  $n$  locais  $u_n$ , seguindo a mesma sequência: [1] em cada nó gera-se uma nova ccdf condicionada tanto aos dados originais quanto aos previamente simulados e [2] da distribuição acumulada (ccdf) sorteia-se um valor randômico para cada nó ( $u_n$ ) do grid.

O condicionamento da distribuição é explicada pelo Axioma de Bayes, o qual propõe que uma ccdf de dois pontos pode ser expressa pelo produto de uma ccdf de cada um dos pontos. O Axioma, expresso pela Equação 2.20, gera o par de pontos  $z^{(l)}(u'_1), z^{(l)}(u'_2)$  em duas etapas: [1] o valor de  $z^{(l)}(u'_1)$  é primeiramente extraído da ccdf, [2] então a ccdf na posição  $u'_2$  é condicionada à realização  $z^{(l)}(u'_1)$  além dos dados originais e sua amostragem produz o valor correlacionado  $z^{(l)}(u'_2)$  (GOOVAERTS, 1997)

$$F(u'_1, u'_2; z_1, z_2 | (n)) = F(u'_2; z_2 | (n+1)) \cdot F(u'_1; z_1 | (n)) \quad (2.20)$$

O atributo  $(n+1)$  da Equação 2.20 denota a condicionante tanto para os  $n$  valores de  $(z(u_a))$  quanto para a realização  $(Z(u) = z^{(l)}(u_1))$ .

No que diz respeito ao elipsoide de busca das amostras, Deutsch e Journel (1998) recomendam que a procura deve ser realizada por pelo menos até a distância máxima a qual o variograma deva ser reproduzido. A limitação de vizinhança implica que as propriedades estatísticas do conjunto  $(N + n)$  das variáveis randômicas serão reproduzidas apenas até a distância máxima da amostra encontrada na vizinhança. Goovaerts (1997) cita que os nós de *grid*  $u$  a serem simulados devem ser escolhidos aleatoriamente a fim de evitar a geração de uma possível continuidade artificial.

A técnica de *multigrid* (TRAN, 1994) é uma outra estratégia para dificultar a manifestação de artefatos na simulação. O conceito permite reproduzir estruturas de correlação de longo alcance, utilizando vizinhança de busca grande com poucos dados de condicionamento. O procedimento para a estratégia de busca por *multigrid* pode ser explicado em duas etapas: [1] construir um *grid* com blocos de maiores dimensões do que os blocos do *grid* original e utilizar uma vizinhança de busca igualmente grande, garantindo a reprodução das estruturas do variograma numa maior escala, [2] os nós restantes são

simulados utilizando tamanhos de blocos e estratégia de busca cada vez menores, até que atinja os parâmetros iniciais de *grid* e raio de busca (GOOVAERTS, 1997; DEUTSCH; JOURNAL, 1998).

Após finalizar as realizações da simulação, basta aplicar uma retro-transformação nos valores gaussianos para que eles voltem ao seu espaço original. Este procedimento é feito pelo processo inverso da transformação inicial, associando os valores do campo gaussiano à distribuição original dos dados.

#### 2.4.2 Simulação sequencial dos indicadores (ALABERT, 1987)

Os indicadores são utilizados com intuito de reduzir os  $K$  valores possíveis de uma variável categórica, para apenas duas possibilidades de um determinado evento ocorrer ou não. As informações qualitativas podem ser tratadas por algoritmos de indicadores desde que codificados em valores binários, digamos 1 se a categoria estiver presente e 0 caso esteja ausente (GOOVAERTS, 1997; JOURNAL, 1983).

Seja  $S$  um atributo categórico com  $K$  estados possíveis  $s_k$ ,  $k=1, 2, \dots, K$ . Os estados de  $K$  são exaustivos e mutuamente exclusivos no sentido de que um e somente um estado  $s_k$  ocorre em cada local  $u$ . O padrão de variação de uma categoria  $s_k$  pode ser caracterizado por semivariogramas, definido num código indicador da presença ou ausência dessa categoria (GOOVAERTS, 1997).

Diferentemente das técnicas baseadas no espaço gaussiano, os algoritmos não paramétricos não assumem qualquer forma particular ou expressão analítica para as distribuições condicionais. Em vez disso, o valor da função  $F(u, z|(n))$  (Equação 2.21) é determinada para uma série de valores categóricos de  $K$  discretizados nas  $z_k s$  (GOOVAERTS, 1997):

$$F(u, Z_k|(n)) = Prob[Z(u) \leq Z_k|(n)] \quad k = 1, 2, \dots, K \quad (2.21)$$

A função de probabilidade é igual a esperança condicional das variáveis categóricas aleatórias  $I(u; z_k)$ , em relação aos dados condicionantes  $i(u_\alpha; z_k)$ .

A vantagem do método não paramétrico é que este não requer transformação não linear dos dados, simulando diretamente o valor no espaço original a partir das distribuições condicionais derivadas da krigagem local dos indicadores. Outra vantagem das técnicas de simulação de indicadores é a sua flexibilidade na incorporação de informações com incertezas associada sob o formato de probabilidades locais anteriores (ROSSI; DEUTSCH, 2013; GOOVAERTS, 1997).

Goovaerts (1997) descreve os passos para a realização da simulação sequencial de indicadores:

- i Discretize o intervalo de variação de  $z$  em classes ( $K + 1$ ) usando valores  $z_k$  das  $K$  categorias. Em seguida, transformar cada dado  $z(u_\alpha)$  em um vetor de dados indicadores originais (Equação 21);

$$f(x) = \begin{cases} 1 & \text{se } z(u_\alpha) = z_k; \\ 0 & \text{caso contrario.} \end{cases}$$

- ii Definir um caminho aleatório visitando cada nó do *grid*;
- iii Determinar em cada nó  $u'$  os valores das  $K$  cdfs  $[F(u'_i; z_k | (n))]$ \* usando qualquer um dos algoritmos de krigagem de indicadores. Em seguida deve-se corrigir desvio de relação de ordem. Construir um modelo cdf completo  $F(u'_i; z_k | (n)), \forall z$ , usando os algoritmos de interpolação ou extrapolação das probabilidades estimadas. Um valor simulado de  $z^{(l)}(u')$  deve ser sorteado aleatoriamente da cdf previamente construída. Os valores simulados são então adicionados ao conjunto de dados de condicionamento;
- iv Repetir os procedimentos anteriores para os próximos nós do *grid*, seguindo um caminho aleatório de visita aos blocos para evitar artefatos de continuidade espacial.

Os algoritmos baseados em indicadores garantem a reprodução aproximada da proporção da classe  $K$  e do semivariograma correspondente ao indicador.

### 2.4.3 Simulação por bandas rotativas (*Turning Bands*)

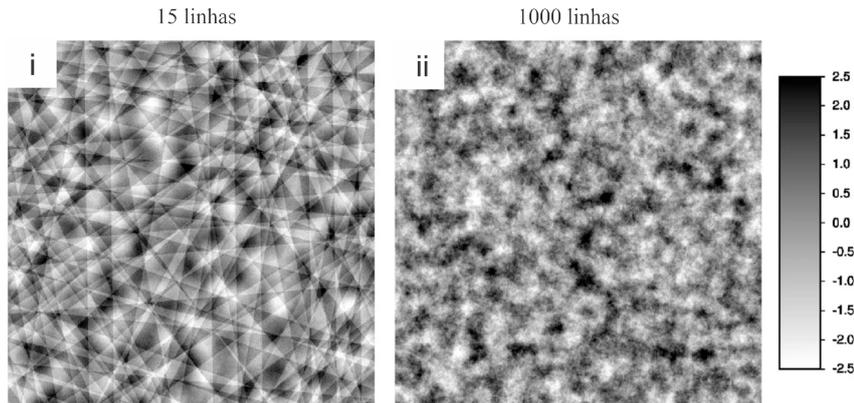
A simulação por bandas rotativas é um método que visa a simplificação de uma simulação no espaço multidimensional ao custo computacional de simulações de linhas no espaço unidimensional. O método consiste em  $N$  linhas (1D) concêntricamente rotacionadas no espaço 2D ou 3D (LANTUÉJOUL, 2002; EMERY; LANTUÉJOUL, 2006; CHILES; DELFINER, 2009).

Este método de simulação também exige que os dados de entrada apresentem uma distribuição gaussiana, afim de garantir a reprodução da continuidade espacial. Caso a distribuição não apresente tal comportamento, deve-se aplicar a transformação não linear *normal score* descrita no início desta seção. O algoritmo da simulação por bandas rotativas executa basicamente dois processos: (1) gerar um modelo não condicional respeitando as condições de variabilidade espacial e distribuição de probabilidade original do banco de dados; (2) condicionar a simulação das bandas rotativa através de um pós-processamento usando krigagem (ROSSI; DEUTSCH, 2013).

A simulação não condicional pode ser realizada por diversos métodos, sendo os dois mais tradicionais o por médias móveis e os métodos espectrais contínuos. Os demais métodos de simulação são expostos por Emery e Lantuéjoul (2006). Os autores ainda afirmam que, para o procedimento, 1000 é um valor suficiente de linhas conseguindo

atender a maioria dos casos sem gerar artefatos. A [Figura 7](#) ilustra a diferença entre a simulação realizadas com um pequeno número de linhas, apenas 15 e outra com o número considerável de 1000 linhas. Percebe-se que a simulação obtida com o maior número de linhas, apresenta-se mais suavizada e sem as feições espaciais, como no exemplo com 15 linhas, em que aparecem artefatos que se assemelham a um espelho quebrado.

Figura 7 – Simulações por bandas rotativas. Em [i] utilizou-se apenas 15 linhas e em [ii] foram utilizadas 1000 linhas.



Modificado de [Emery e Lantuéjoul \(2006\)](#)

O número de linhas não é um fator limitante no quesito computacional, visto que as simulação em uma dimensão são extremamente rápidas. O grande gasto computacional fica por conta do condicionamento dos dados ([EMERY; LANTUÉJOUL, 2006](#)).

[Chiles e Delfiner \(2009\)](#) expõem um exemplo de simulação em 2D utilizando apenas 8 linhas concentricamente rotacionadas ([Figura 8](#)). Para simular o ponto  $x$ , primeiramente deve-se identificar as projeções deste ponto nas oito linhas. Ao ser identificado os valores da projeção, a simulação no ponto  $x$  pode ser facilmente obtida pela Equação 2.22.

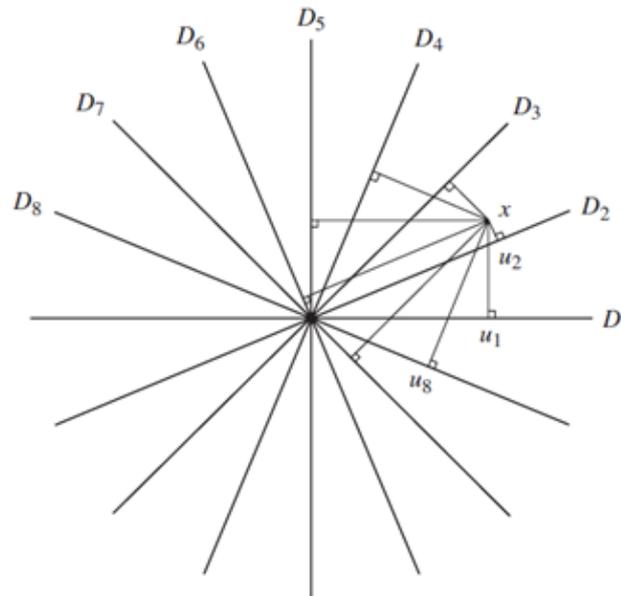
$$S(x) = \frac{1}{\sqrt{n_D}} \sum_{t=1}^{n_D} S_t(S_t) \quad (2.22)$$

Onde  $n_D$  representa o número de bandas rotativas e  $S_t(S_t)$  são as simulações não condicionais.

O método honra exatamente os dados de condicionamento e também preserva a variabilidade das realizações da simulação incondicional. O condicionamento dos dados é realizado através da krigagem simples dos resíduos, que são obtidos entre os valores de entrada e os valores simulados coincidentemente geoposicionados. No locais onde ocorre a sobreposição de dados, é realizada a substituição do valor interpolado pelo valor do dado de entrada ([JOURNAL, 1974; ROSSI; DEUTSCH, 2013](#)).

Os modelos simulados devem ser retro-transformados para as unidades originais, para que sejam realizados os procedimentos de validação da continuidade espacial e da distribuição de frequência do atributo.

Figura 8 – . Simulação do ponto x pelo método de bandas rotativas em 2D.



Fonte: Chiles e Delfiner (2009).

#### 2.4.4 Mudança de suporte

SMU é definido como o menor volume de bloco que pode ser seletivamente extraído como minério ou estéril. Baseado em critérios de espaçamento entre amostras, Journel e Huijbregts (1978) propõem que as dimensões dos blocos devem corresponder de 30 à 50% do espaçamento amostral. Essa prática visa minimizar erros causados por suporte de bloco inadequado. Blocos grandes demais podem mascarar a resolução da malha de sondagem, e blocos muito pequenos produzirão uma suavização artificial no modelo (ROSSI; DEUTSCH, 2013). Pela variância da dispersão, é possível quantificar a redução da variância para um aumento específico de volume do bloco. A variância da distribuição pode ser calculada pela Equação 2.23.

$$\sigma^2\left(\frac{p}{A}\right) = \sigma^2\left(\frac{B}{A}\right) + \sigma^2\left(\frac{p}{B}\right) \quad (2.23)$$

Em que:

$\sigma^2\frac{p}{A}$  é a variância do suporte ponto em A;

$\sigma^2 \frac{B}{A}$  é a variância do suporte bloco em A;

$\sigma^2 \frac{p}{B}$  é a variância dos pontos dentro dos blocos.

A equação descreve que a variância das amostras do depósito pode ser obtida pela adição entre as variâncias dos pontos discretizados dos blocos e as variância entre blocos.

Através da estatística básica das amostras é possível obter a variância dos pontos. Já, a variância dentro dos blocos é obtida pela discretização do bloco em  $n$  pontos calculando o variograma médio destes pontos dentro do bloco. A discretização dos blocos normalmente compreende 125 pontos (5x5x5) igualmente distribuídos dentro do bloco. [Rossi e Deutsch \(2013\)](#) citam que este valor de discretização é suficiente para se obter uma estimativa robusta da variância dos pontos dentro dos blocos.

O suporte dos blocos influencia principalmente na variância da distribuição. No método de simulação, deve-se garantir a reprodução da variância original do depósito amostrado no suporte das amostras (pontos). Dessa maneira, o suporte dos blocos para as realizações da simulação deve ser suficientemente pequeno, aproximando-se ao suporte da variável. Utilizando blocos pequenos reduz-se a suavização da variância causada pela mudança de volume do suporte.

## 2.5 Observações finais

Neste capítulo, foram abordados os métodos estatísticos univariados e bivariados que contribuem com o modelo de regressão proposto para a recuperação do Zinco. As variáveis de entrada serão perturbadas pelos métodos de simulação estocásticas, possibilitando o acesso às incertezas locais associadas às variáveis de entrada. No próximo capítulo, será apresentado o estudo de caso, colocando em prática o conhecimento teórico exposto no presente capítulo.

## 3 Estudo de caso

O estudo de caso contempla os trabalhos de geometalurgia desenvolvidos em uma importante mina de Zinco brasileira, no intuito de compreender o comportamento dos diferentes tipos de minério na planta de beneficiamento. Para atingir os objetivos propostos pela dissertação, as variáveis relevantes frente à recuperação do Zinco serão perturbadas nos modelos de teores utilizando o método de simulação estocástica.

### 3.1 Geologia Local

A geologia da área é composta por uma sequência de rochas pelito-carbonáticas caracterizada por rochas pertencentes às formações Serra de Garrote e Serra do Poço Verde e Morro do Calcário ([Figura 9](#)), ambas inseridas no Grupo Vazante. Estas exibem comportamentos distintos, ligados à diferença de competência de seus litotipos. A Fm. Serra do Garrote é representada quase que exclusivamente por filitos, localmente associados a níveis de silexito e quartzíticos e, blocos de hematita, já a Fm. Serra do Poço Verde é composta de dolomitos com intercalações de níveis metapelíticos e margas ([DARDENNE et al., 1998](#); [DARDENNE et al., 2000](#)).

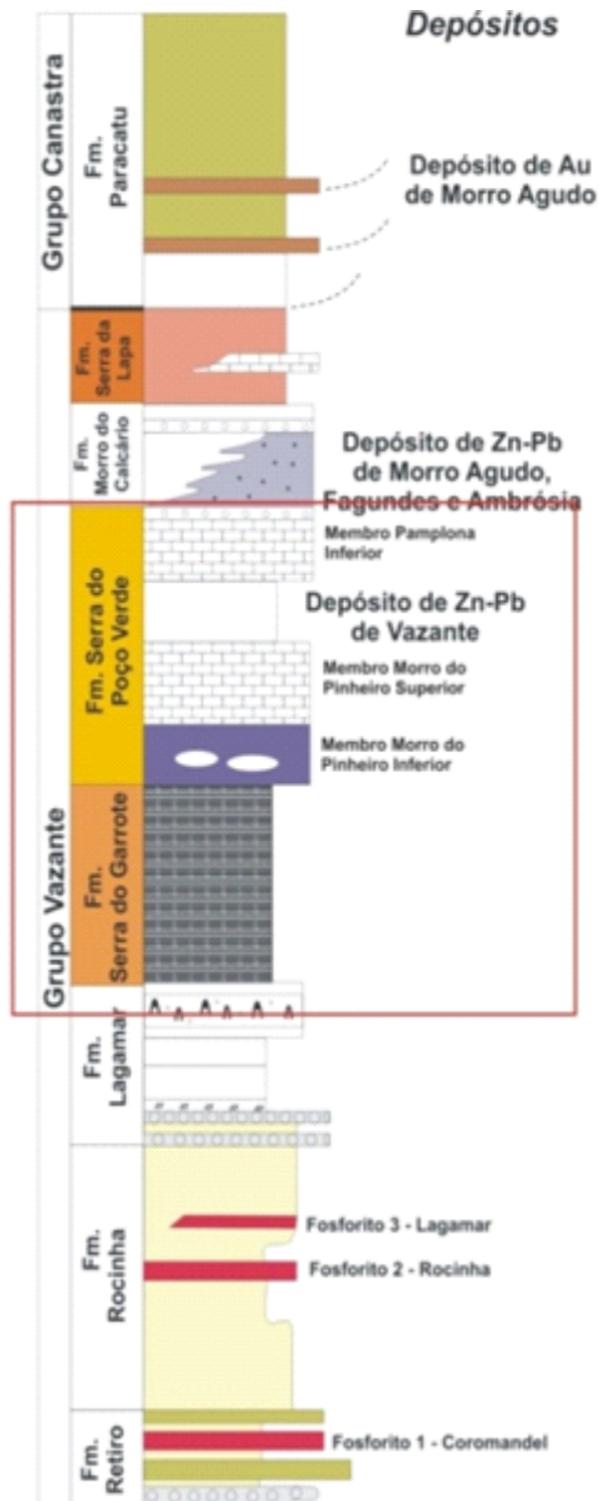
- Formação Serra do Garrote

É constituída quase que exclusivamente por filitos carbonosos que estão associados a brechas ferruginosas e blocos de hematita. Essas duas litologias são de origem tardia com relação aos filitos, estando a primeira associada a processos superficiais, e a segunda com passagem de fluídos (hidrotermalismo), desta forma serão abordadas mais adiante. São filitos carbonosos, de cor preta a cinza-chumbo quando fresco, e/ou coloridos quando intemperizados, bastante homogêneos, apenas com diferenças pontuais, sem distribuição espacial que possa influenciar em uma distinção litológica. Os afloramentos são abundantes, ocorrendo tanto na forma de lajedos, onde há passagem de água, quanto grandes paredões, esculpido nas encostas dos morros e em barrancos de corte de estrada ([BABINSKI et al., 2005](#)).

- Formação Serra do Poço Verde

Na área do Extremo Norte, esta formação é representada da base para o topo pelos membros Morro do Pinheiro Inferior, Morro do Pinheiro Superior e Pamplona Inferior. Os dois primeiros ocorrem a leste da Falha de Vazante, no contato de lapa com a brecha dolomítica hidrotermal e o último ocorre a oeste da falha, no contato de capa com a brecha.

Figura 9 – Coluna estratigráfica do Grupo Vazante. Em destaque unidades presentes na região.



Modificado de Babinski et al. (2005)

- Membro Morro do Pinheiro Inferior

Proposto por [Dardenne \(1978\)](#), este membro ocorre na base da sequência dolomítica da Formação Serra do Poço Verde na área do Extremo Norte. Formada por dolomitos cinza, por vezes rosada, com esteiras algais e raros bird's eyes. Ocorrem nesta unidade, dolorruditos, caracterizados por exibir clastos carbonáticos angulosos a sub-arredondados, de tamanhos variados, dispostos em uma matriz micrítica. É composto quase que exclusivamente por dolomita e por alguns raros cristais de opacos e quartzo. Mostra-se localmente silicificado e apresenta venulações de Fe-dolomita, quartzo e dolomita.

- Membro Morro do Pinheiro Superior

Definido por [Dardenne \(1978\)](#), este membro é caracterizado por um metadolomito fino, com granulação de fração argila a silte, exibindo estrutura maciça a levemente laminada. Mostra-se praticamente isenta de influência hidrotermal, mantendo-se sempre com sua cor cinza claro e/ou escura inalterada e suas estruturas primárias preservadas.

- Membro Pamplona Inferior

Definida por [Dardenne \(1978\)](#), esta litofácia situa-se a oeste da Falha de Vazante, e está contida no bloco que foi rebaixado pela Falha Vazante. Este metadolomito exibe tonalidade rosada, granulação fina, e é marcado pela presença de esteiras de algas, que definem seu acamamento sedimentar. Possui níveis margosos, de cor roxa a esverdeada, onde é possível identificar duas foliações metamórficas.

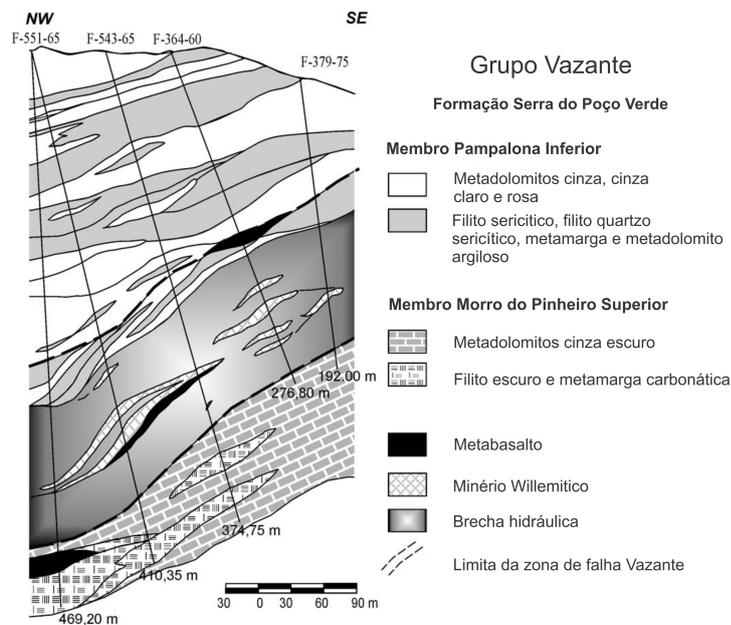
- Metadolomitos brechados (Brecha hidrotermal)

Os metadolomitos desta unidade estão contidos em uma faixa de direção NS associada à Falha de Vazante e cortam a área do Extremo Norte exibindo um espessamento próximo ao contato com a Formação Serra de Garrote. São caracterizados por uma intensa deformação (D3), que propiciou o desenvolvimento de sua estrutura brechada. Essa rocha exibe cor rosa a avermelhada e não preserva estruturas primárias que possam evidenciar o acamamento sedimentar, mostra-se intensamente venulada, com espessos veios de siderita, ankerita e dolomita. Seus clastos são arredondados a sub-arredondados devido à assimilação parcial ocorrida durante a passagem dos fluidos hidrotermais superaquecidos; além disso, mostram-se dispostos caoticamente em matriz dolomítica, totalmente recristalizada, exibindo estrutura em pente ([BABINSKI et al., 2005](#)).

O magmatismo associado ao Grupo Vazante é representado por pequenos corpos de rochas metabásicas, identificados principalmente na mina de zinco subterrânea Vazante ([Figura 10](#)). Neste depósito, os metabasaltos ocorrem na zona de cisalhamento Vazante,

tectonicamente imbricados com brechas hidráulicas, rochas alteradas por hidrotermalismo e o minério de zinco não sulfetado, composto principalmente de willemita.

Figura 10 – Corte transversal da zona de minério Vazante mostrando as relações espaciais entre os metabasaltos, a sequencia hospedeira, o minério willemítico e a zona de cisalhamento Vazante.



Modificado de Babinski et al. (2005 apud MONTEIRO et al., 1999)

## 3.2 Análise exploratória dos dados

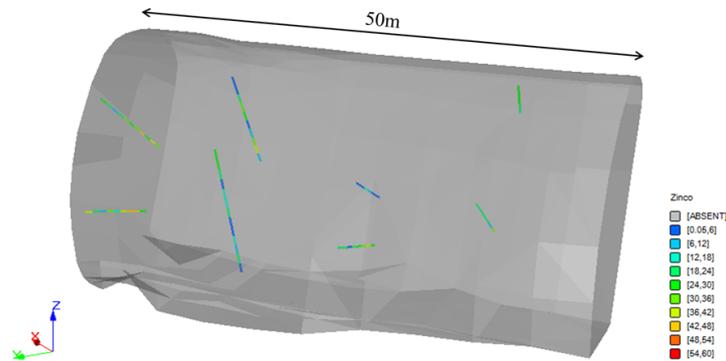
A análise exploratória dos dados está dividida em dois bancos de dados, sendo o primeiro deles o banco de dados geometalúrgico contendo 104 amostras e o segundo banco de dados contendo informações litoquímica (descrição geológica e informações de teores para o elemento zinco).

### 3.2.1 Banco de dados geometalúrgico

Os testes de bancada exigem uma quantidade considerável de massa de testemunhos de sondagem, entre 30 e 50kg para cada teste. Devida a esta dificuldade em gerar massa a partir da amostragem de testemunhos, a coleta de dados para os testes geometalúrgicos foram feitas em diversos furos de maneira a compreender o volume de um bloco de lavra (*estope*), como ilustrado pela Figura 11.

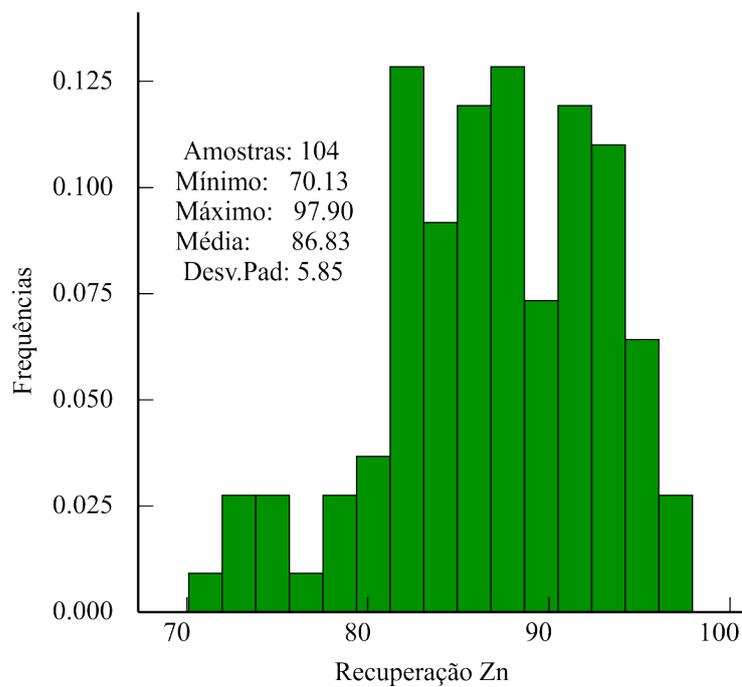
Foram coletadas um total de 1325 amostras de 257 testemunhos de sondagem compreendendo toda a extensão da mina. As amostras foram suficientes para gerar 104

Figura 11 – Ilustração representativa da definição de amostras de testemunhos para realização dos testes metalúrgicos.



testes metalúrgicos e a [Figura 12](#) ilustra o histograma dos resultados de recuperação mássica do concentrado (%).

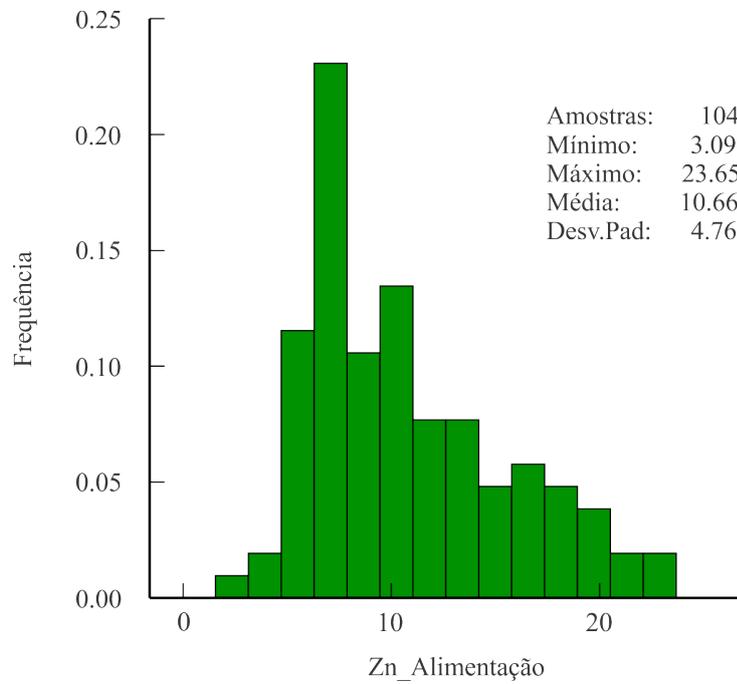
Figura 12 – Histograma e informações estatísticas da recuperação metalúrgica do Zinco, em percentagem nos testes.



O histograma da [Figura 13](#) contempla os dados dos teores de alimentação média de Zinco nos testes de bancada. A média das alimentações ficou em torno dos 10.66% de

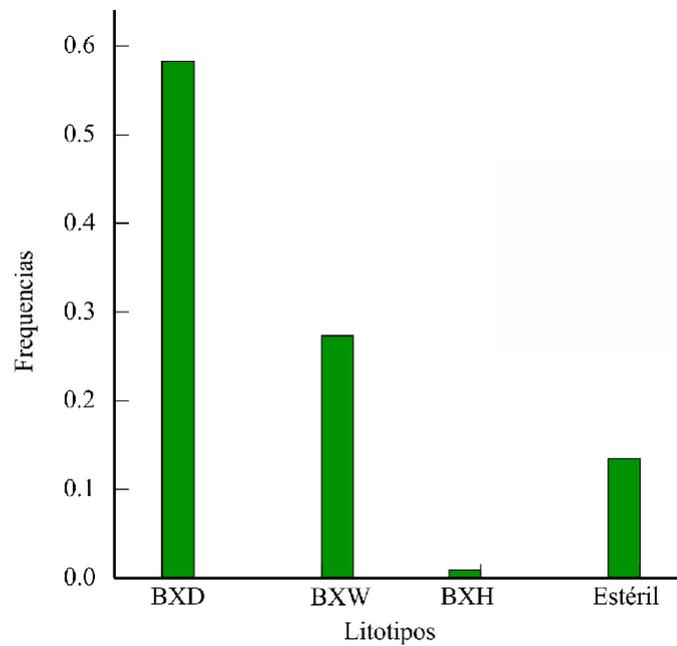
Zinco com um desvio padrão de 4.76%. O CV deste banco de dados é de aproximadamente 0.5, indicando uma erraticidade relativamente baixa.

Figura 13 – Histograma dos teores médios de alimentação do Zinco nos testes metalúrgicos.



A [Figura 14](#) ilustra a distribuição das litologias presentes nos testes metalúrgicos. É interessante observar, que além dos litotipos mineralizados, havendo uma predominância da Brecha Dolomítica nos testes geometalúrgicos. A presença de uma quantidade significativa de estéril, aproximadamente 15%, se faz presente devido à diluição planejada do bloco de lavra.

Figura 14 – Distribuição das litologias presentes nos 104 testes metalúrgicos.

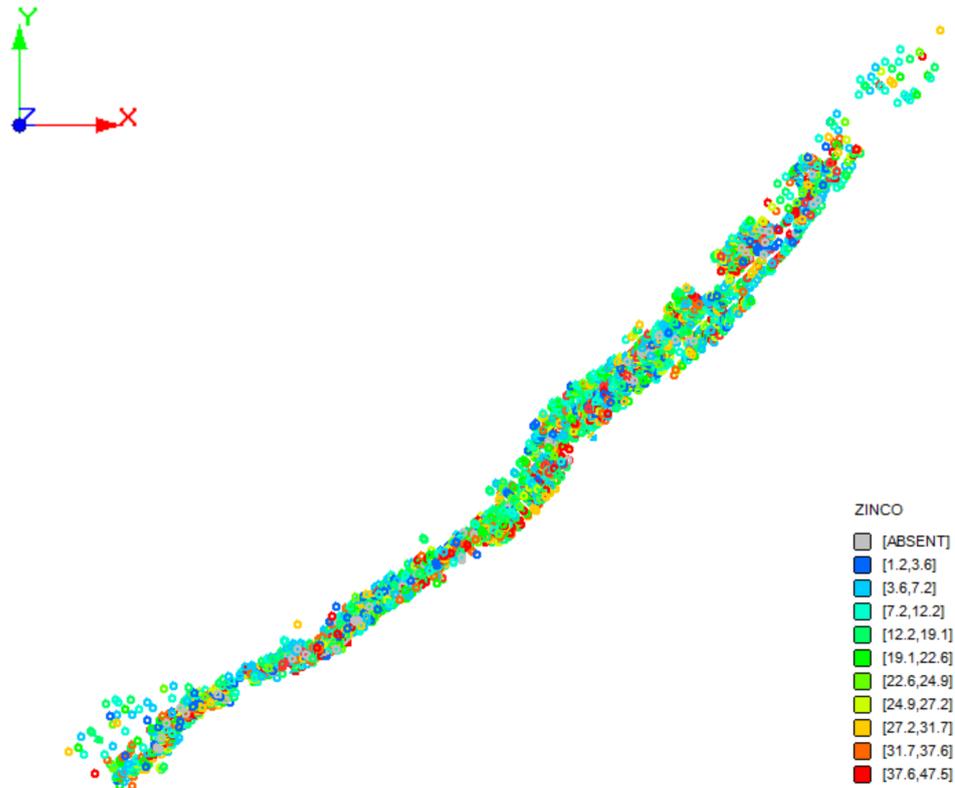


Os histograma abaixo ilustram os teores de zinco por domínio litológico a fim de avaliar algum viés na litologia.

### 3.2.2 Análise exploratória dos dados de sondagem

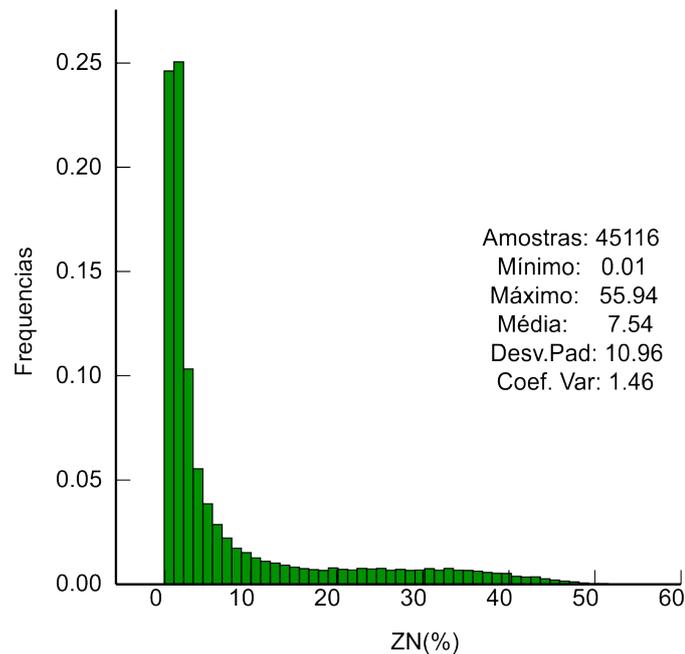
A mineralização é controlada por uma zona de falhas de atitude N50E/70NW, justificando a direção preferencial das campanhas de sondagem para este azimute ([Figura 15](#)).

Figura 15 – Distribuição e localização das amostras contendo análise química para o Zinco. A extensão da mina na direção  $N50E$  é de aproximadamente 5,500 metros.



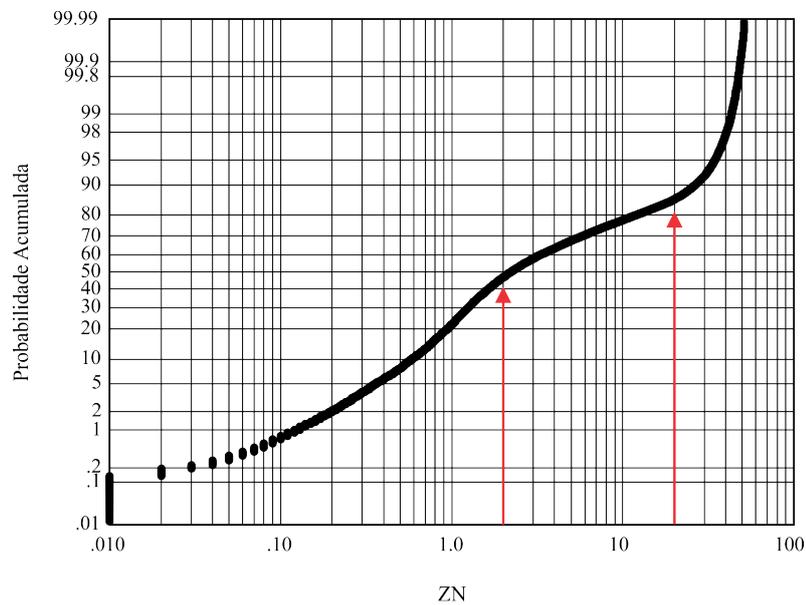
O banco de dados abrange 45,116 amostras de teores de Zinco, apresentando média de 7.54% e desvio padrão de  $11\%^2$ . O valor do coeficiente de variação (CV) é de 1.45 (Figura 16). Isaaks e Srivastava (1989) citam que um valor de CV acima de 1.00 pode indicar presença de alguns valores erráticos altos, sendo algumas vezes necessária a separação dos domínios mineralizados. Este tema será abordado nas próximas subseções. As amostras não possuem as mesmas dimensões e a fim de igualar o suporte amostral foi realizada uma regularização dos dados, colocando todas as amostras com comprimento de 2 metros.

Figura 16 – Histograma da variável Zinco.



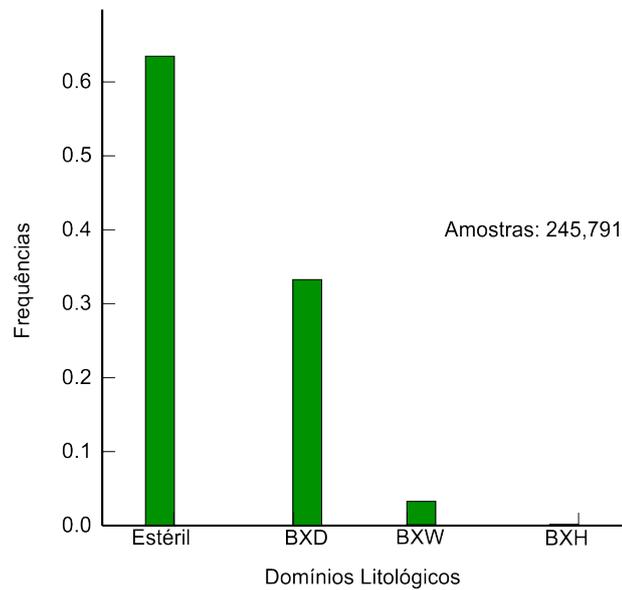
Um diagrama de probabilidade (Figura 17) foi elaborado contemplando todos os domínios para avaliar o comportamento dos teores de Zinco, de uma maneira geral. O diagrama ilustra algumas inflexões na curva log-gaussiana, indicadas pelas setas vermelhas. Estas inflexões podem sugerir uma mistura de populações com comportamentos estatísticos diferentes. Como objetivo deste trabalho é perturbar ao máximo os dados de entrada, não foi considerada nenhum tipo de manipulação dos dados originais, como por exemplo uma truncagem dos valores extremos.

Figura 17 – Gráfico de probabilidade ilustrando a probabilidade dos teores de zinco contemplando todos os domínios geometalúrgicos.



Com relação às litologias, o banco de dados compreende 245,791 amostras descritas em 26 diferentes litologias. Para os testes metalúrgicos foram consideradas três litologias e o restante foi reclassificado como estéril, pois são litotipos que apresentam teores muito baixos de Zinco, quando apresentam. Nos testes de bancada, os litotipos foram reclassificados em apenas quatro litotipos, sendo que as rochas que não são BXD, nem BXH e nem BXW foram renomeadas para estéril (Figura 18).

Figura 18 – Histograma da distribuição das litologias reclassificadas de acordo com a classificação das rochas utilizadas no teste de bancada.

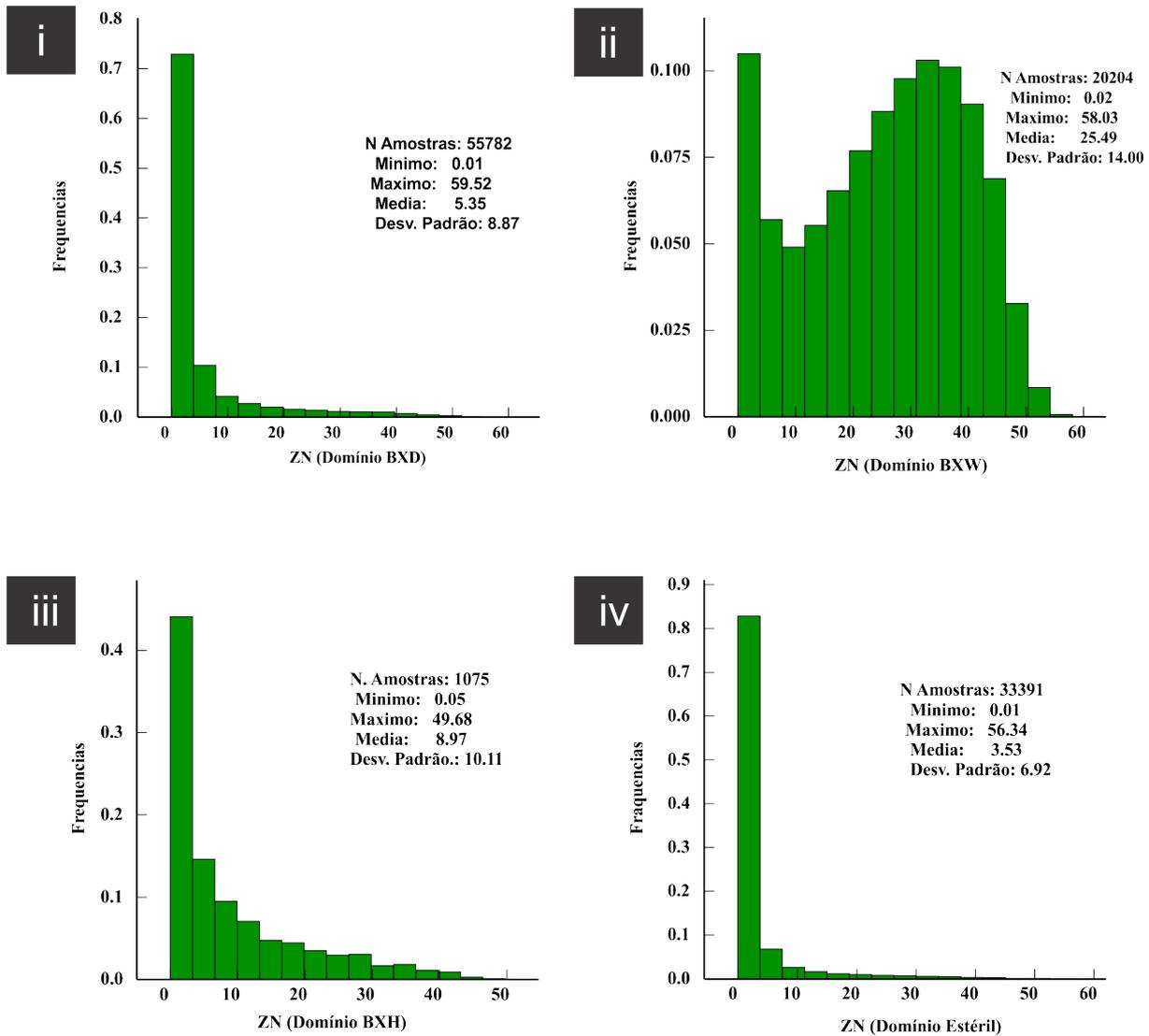


Os histogramas da [Figura 19](#) ilustram os teores de zinco em cada um dos domínios litológicos definidos para os testes geometalúrgicos. Dentro os domínios analisados, a Brecha Willemítica (BXW) apresenta o maior teor de 25% de zinco em média e é o domínio menos errático, com coeficiente de variação (CV) de 0.55.

O domínio da brecha dolomítica (BXD) apresenta média de teores de zinco de 5.35%, sendo que 73% das amostras apresentam teores de até 5%. A população deste domínio é a mais errática, sendo o seu CV superior a 1.50. Neste caso é recomendado realizar um estudo de restrição de valores extremos (*capping*). A brecha hematítica está pouco presente nos testes metalúrgicos, mas parece-se que este domínio apresenta um teor médio mais alto do que a brecha dolomítica e menos errático.

No caso do domínio do estéril temos um teor médio relativamente alto, provavelmente ocasionado por amostras em zonas de contato ou até mesmo teores perdidos no meio do estéril.

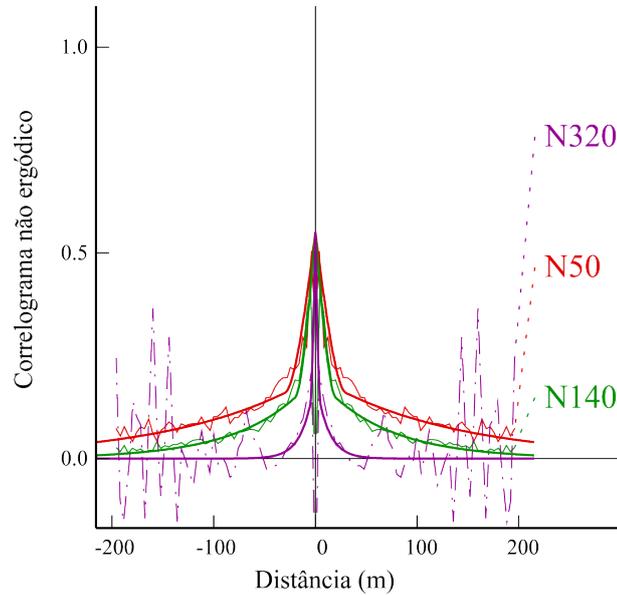
Figura 19 – Histograma da distribuição das litologias reclassificadas de acordo com a classificação das rochas utilizadas no teste de bancada.



### 3.2.2.1 Modelo de continuidade espacial do Zinco

A definição das direções de continuidade é importante para mensurar o nível de dissimilaridade entre os espaços não amostrados e os dados mais próximos (JOURNEL, 1989), para uma mesma distância  $h$ . O modelo de continuidade espacial foi elaborado e a Figura 20 ilustra o correlograma não ergódico da variável Zinco. A Equação 3.1 representa matematicamente o modelo correlográfico da variável Zinco.

Figura 20 – Correlograma não ergódico do banco de dados original do Zinco.



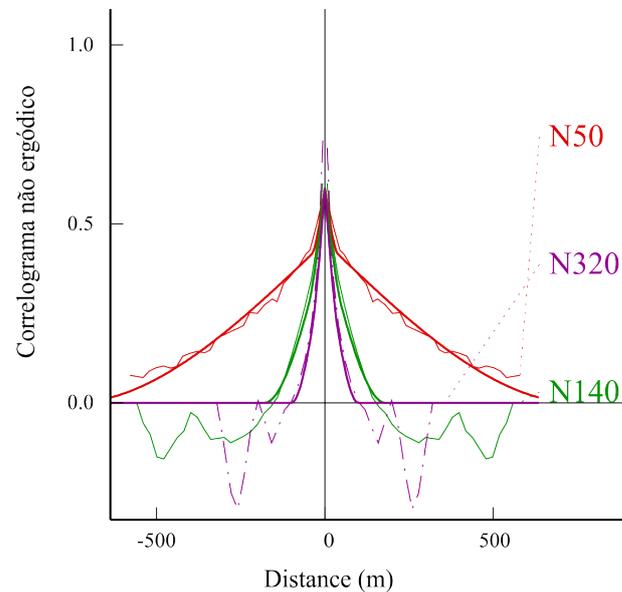
$$\begin{aligned} \rho(h) = & 0.45 + 0.35Sph \left( \frac{N50/0}{30m}, \frac{N320/60}{20m}, \frac{N140/30}{5m} \right) \\ & + 0.20Sph \left( \frac{N50/0}{400m}, \frac{N320/60}{200m}, \frac{N140/30}{36m} \right) \end{aligned} \quad (3.1)$$

### 3.2.2.2 Modelo de continuidade espacial da litologia BXD

O correlograma da [Figura 21](#) ilustra o modelo de continuidade da variável Brecha Dolomítica e a Equação 3.2 ilustra matematicamente o modelo correlográfico. O modelo de continuidade dos indicadores reproduziu uma direção de maior continuidade idêntica ao modelo dos teores de Zinco. Possivelmente, esta condição de repetição pode confirmar que a formação da rocha também está condicionada às falhas N50E/70NW, assim como a deposição do Zinco.

$$\begin{aligned} \rho(h) = & 0.40 + 0.15Sph \left( \frac{N50/0}{50m}, \frac{N320/60}{40m}, \frac{N140/30}{30m} \right) \\ & + 0.45Sph \left( \frac{N50/0}{750m}, \frac{N320/60}{180m}, \frac{N140/30}{100m} \right) \end{aligned} \quad (3.2)$$

Figura 21 – Correlograma não ergódico do banco de dados original da litologia Brecha Dolomítica.



### 3.3 Análise estatística multivariada

A recuperação metalúrgica apresenta um comportamento não aditivo, apresentando problema para sua estimativa com a geoestatística convencional. As técnicas de regressão linear foram utilizadas neste trabalho como alternativa para os problemas de não aditividade da variável geometalúrgica.

Alguns cuidados devem ser tomados quando se trabalha com regressões lineares, pois a presença de variáveis com baixa correlação ou informações redundantes podem gerar ruídos no modelo final. Utilizando o software *Corrmat* da Biblioteca de Geoestatística (GSLib) (DEUTSCH; JOURNAL, 1998), foi elaborada uma matriz de correlação exibindo as relações entre todas as variáveis metalúrgicas (Figura 22). A matriz de correlação é um dos métodos utilizados para avaliar a correlação entre as amostras mas sabe-se que esta não é a melhor metodologia para esta análise.

Figura 22 – Matriz de correlação das variáveis do teste metalúrgico.

**Matriz de Correlação**

Zn_Alim	1.00	0.41	-0.28	0.18	0.13	0.67
Pb_Alim	0.41	1.00	-0.10	-0.11	-0.06	0.28
%BXD	-0.28	-0.10	1.00	-0.02	-0.27	-0.42
%BXH	0.18	-0.11	-0.02	1.00	0.12	-0.07
%BXW	0.13	-0.06	-0.27	0.12	1.00	0.22
Zn_Rec	0.67	0.28	-0.42	-0.07	0.22	1.00
	Zn_Alim	Pb_Alim	%BXD	%BXH	%BXW	Zn_Rec

A partir da matriz e correlação apresentada, realizou-se uma análise de sensibilidade gerado pelo Software SABOR a Biblioteca de Geoestatística (GSLib) (DEUTSCH; JOURNAL, 1998), gerando o diagrama de tornado. Foi constatado que as variáveis mais significativas para a recuperação do Zinco são: [1] o teor de Zinco e [2] litotipo BXD. O valor de  $R_a^2$  ( $R^2$  ajustado) é de 48.07%, o que não é considerado um valor alto. Demais informações sobre desvio padrão, valor médio, erro, coeficiente de variação, de sensibilidade e coeficiente de sensibilidade estandardizado podem ser visualizados na Figura 23.

Figura 23 – Matriz de correlação das variáveis do teste metalúrgico.



Conforme mostra a análise de variância usada para testar a significância da regressão na Tabela 2, os resultados do teste de hipótese  $F$  sugerem que o modelo seja aceito.

Tabela 2 – Análise de variância (ANOVA) para testar a significância da regressão.

	Graus de Liberdade	Soma dos Quadrados	Média quadrática	$F$	$P$	Aceito em $\alpha = 0.05$
Regressão	5	1774.12	354.82	24.91	0	Sim
Resíduo(erro)	99	1762.4	17.8			
Total	104	3536.51				

O modelo de regressão linear multivariado para a recuperação do Zinco está representado pela Equação 3.3 e foi obtido também pelo Software SABOR.

$$Zn_{rec} = 0.70 * (Zn_{alim}\% - 11.12) - 6.90 * (Lito_{BXD}\% - 0.50) + 86.54 \quad (3.3)$$

O intervalo de confiança de 90% para a regressão da recuperação metalúrgica foi obtida pela Equação 2.10 apresentada no Capítulo 2, sendo os parâmetros  $t_{\alpha/2,42} = 1.30$  e  $\hat{\sigma}^2 = 21.10\%^2$ .

### 3.4 Aplicação dos métodos de simulação

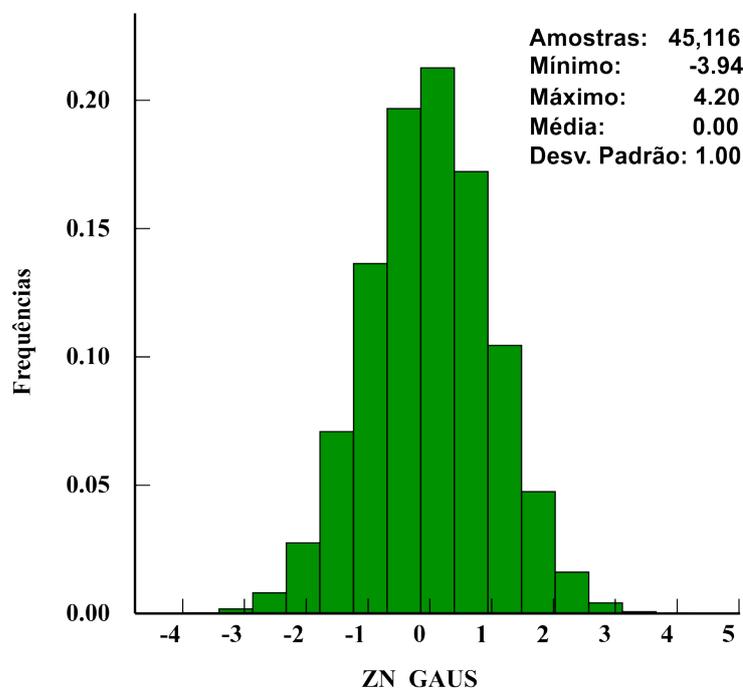
Como foi citado anteriormente, pode-se concluir que a recuperação de Zinco pode ser explicada pela regressão multivariada do teor de Zinco e proporção de BXD. Para acessar a incerteza da variável de entrada dos teores do Zinco, será utilizado o método

de simulação por bandas rotativas (*Turning Bands*), já para a variável litológica será abordado o método de simulação de indicadores categóricos.

### 3.4.1 Transformação gaussiana variável Zinco

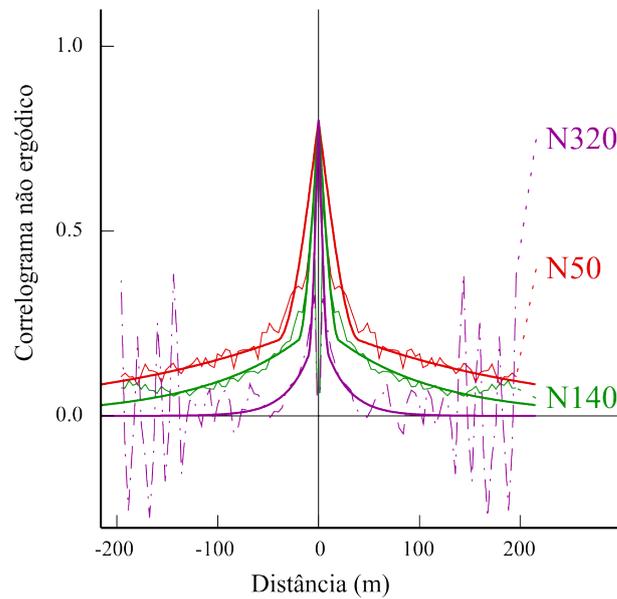
A transformação dos dados de Zinco para o espaço gaussiano se faz necessária no método de simulação por bandas rotativas. A normalização dos dados foi realizada com auxílio do algoritmo *Normal Score* do Software ISATIS©(Figura 24). Os dados de Zinco foram primeiramente desagrupados por células móveis e os pesos foram utilizado na elaboração da função de anamorfose.

Figura 24 – Histograma dos dados do Zinco desagrupados e normalizados.



O modelo de continuidade espacial da variável Zinco normalizada pode ser contemplado pela Figura 25. As direções de continuidade máxima, média e mínima são respectivamente: N50E/0, N320/60 e N140/30.

Figura 25 – Correlograma não ergódico dos dados desagrupados e normalizados pelo método de *normal score*.



### 3.4.2 Multigaussianidade

Goovaerts (1997) cita que a normalidade das funções da distribuição cumulativa (cdfs) de um e dois pontos é necessária, mas não suficiente para garantir que uma função randômica gaussiana seja multiGaussiana para modelar a distribuição espacial dos dados normalizados. Portanto, deve-se checar a normalidade das cdfs em todos os  $n$  pontos. Uma das maneiras de validar a multinormalidade do modelo gaussiano é checando se o valor 0.564 é reproduzido pela relação entre o madograma e raiz quadrada do variograma (Equação 3.5).

$$\frac{\gamma_1(h)}{\sqrt{\gamma_2(h)}} = 0.546 \quad (3.4)$$

Sendo

$\gamma_1(h)$  o madograma normalizado da variável Zinco e

$\gamma_2(h)$  o variograma desta variável.

A Tabela 3 ilustra os resultados obtidos pela Equação 3.5 nas distâncias  $h$ . Os valores das razões calculadas para as diversas distancias  $h$  se aproximam bem do valor padrão 0.564. Este fato não rejeita a multinormalidade dos dados de Zinco.

Tabela 3 – Resultados obtidos pela divisão do madograma pela raiz do variograma da variável normalizada do Zinco.

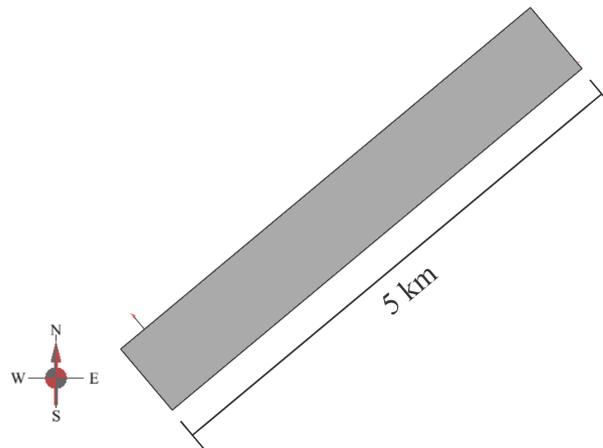
Número de Pares	Distância média (m)	Madograma/Raiz do variograma
163482	5.84	0.538
1456912	22.79	0.55
3434706	40.71	0.554
5557118	60.11	0.557
7218904	80.3	0.558
8460395	100.49	0.559
9436567	120.39	0.56
9862613	139.94	0.561
10660606	159.69	0.562
11099048	180.03	0.562
11246816	200.25	0.562
11207827	220.24	0.562
11042198	239.93	0.563
11218458	259.68	0.564
11327593	279.89	0.564
11137714	300.2	0.564
10838734	320.2	0.564
10500426	339.84	0.564
10607899	359.63	0.564
10561995	379.95	0.565

### 3.4.3 Grid

O suporte dos blocos do *grid* utilizado na simulação deve corresponder a uma escala puntual e assim garantir a reprodução da distribuição dos dados e sua variabilidade.

A definição das dimensões dos blocos do *grid* ocorreram de acordo com a malha de sondagem, aliado ao tamanho das amostras compostas. A malha de sondagem apresenta espaçamento irregular de 20x20 metros. Foram definidos tamanhos de blocos com dimensões de 4x4x3 metros, correspondentes à  $\pm\frac{1}{4}$  da malha de sondagem. Uma rotação de 50° foi aplicada ao modelo do *grid* de forma a reduzir a quantidade de blocos.

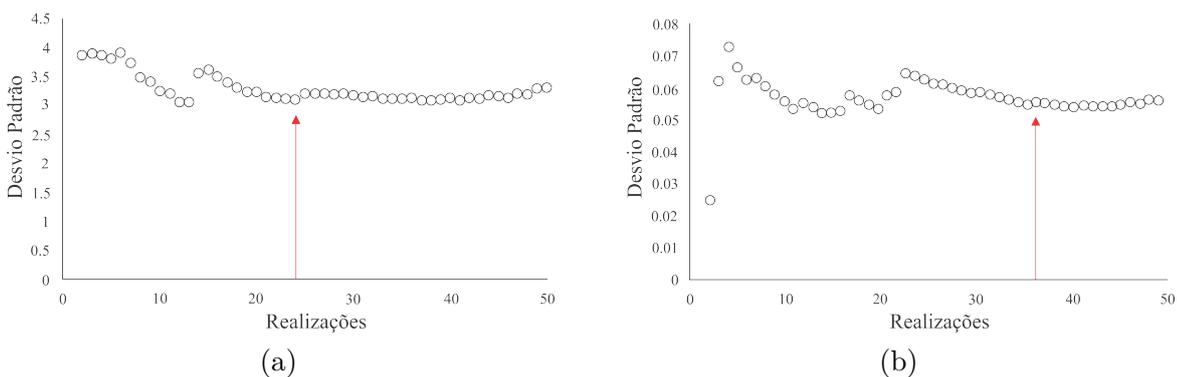
Figura 26 – Grid rotacionado, reduzindo assim o número de blocos e diminuindo o tempo de processamento.



### 3.5 Simulação dos teores Zinco por bandas rotativas

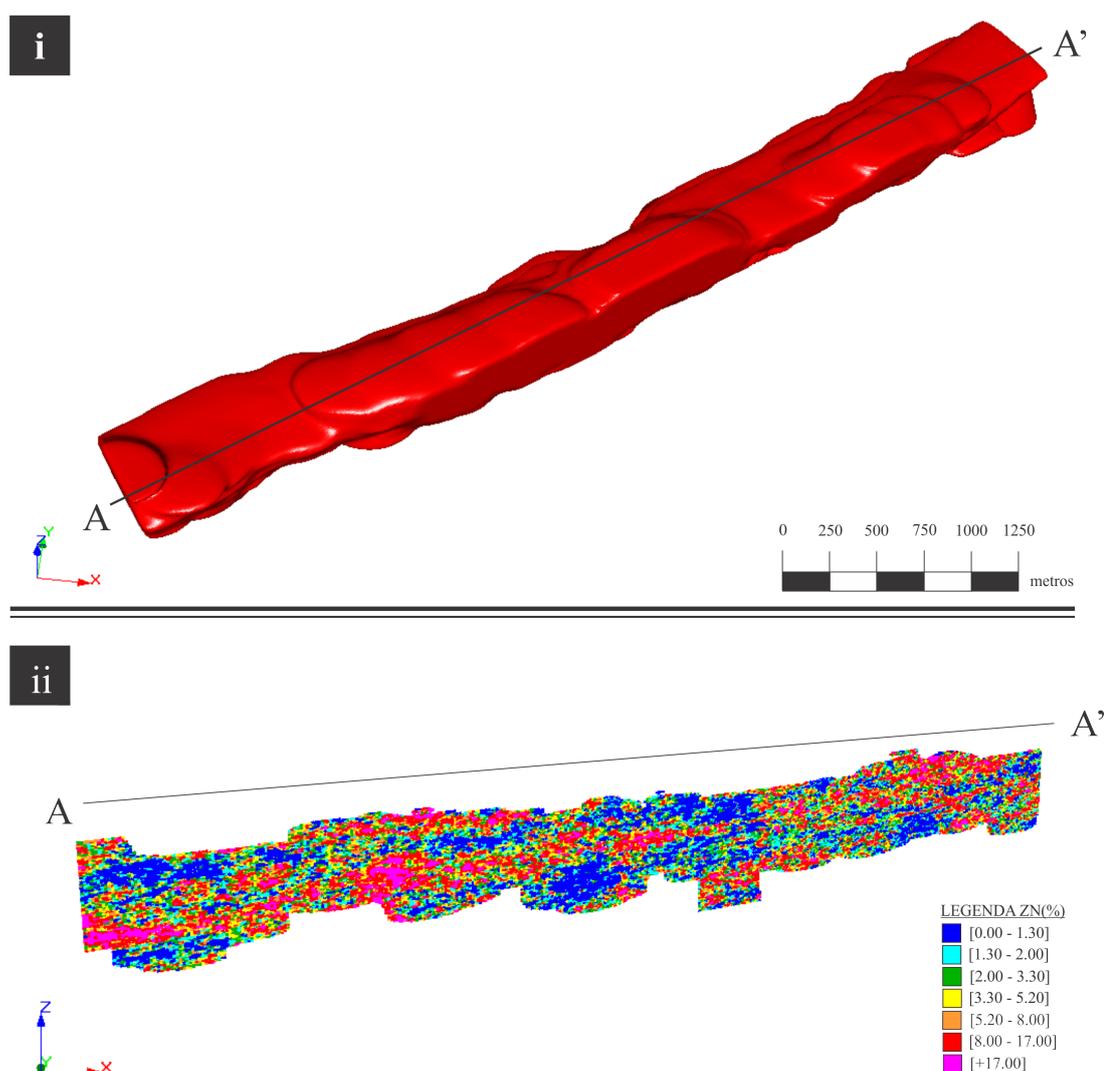
Nesta etapa, serão geradas as simulações dos  $n$  cenários, definidos pelo usuário. A quantidade de realização deve ser suficientemente capaz de mapear todo o espaço de incerteza da variável de entrada. Costa (1997) e Bonato (2000) propõem uma metodologia de análise do número de cenários versus variância das médias de cada simulação. As análises foram realizadas tanto para a variável Zinco (Figura 27a) quanto para a litologia BXD (Figura 27b). No caso da variável Zinco, a estabilização da variância ocorre a partir da 25ª realização, e para a variável litológica, a estabilização ocorre a partir da 35ª realização. Com base nestas informações, concluiu-se que 50 realizações seriam suficientes para mapear todo o espaço de incerteza das variáveis de entrada.

Figura 27 – Número de realizações versus o desvio padrão das médias das simulações de cada uma das variáveis. O diagrama (a) ilustra a curva de dispersão do desvio padrão para a variável Zinco e em (b), o diagrama da dispersão para o litotipo BXD.



A simulação dos teores de Zinco se deu pelo método de Bandas Rotativas (MATHERON, 1973; JOURNEL, 1974), utilizando o *Grid* apresentado na subseção 3.4.3. As variáveis de entrada utilizadas foram os dados de Zinco declusterizados normalizados apresentados na subseção 3.4.1. O elipsoide de busca foi concebido respeitando as mesmas direções de azimute e *dip* dos modelos de continuidade espacial, apresentando dimensões de busca de 150x30x15 metros. Foram utilizadas 1000 bandas rotativas na elaboração de cada cenário simulado, um número aceitável segundo a literatura. A semente inicial da função randômica utilizado nas realizações foi o número 423141, valor *default* do Software ISATIS©. A Figura 28 ilustra a seção A-A' do modelo de blocos da primeira realização para os teores de Zinco.

Figura 28 – Em [i] está ilustrado o sólido da envoltória dos blocos simulados. Em [ii] a seção do modelo de blocos do primeiro cenário gerado pela simulação do zinco.

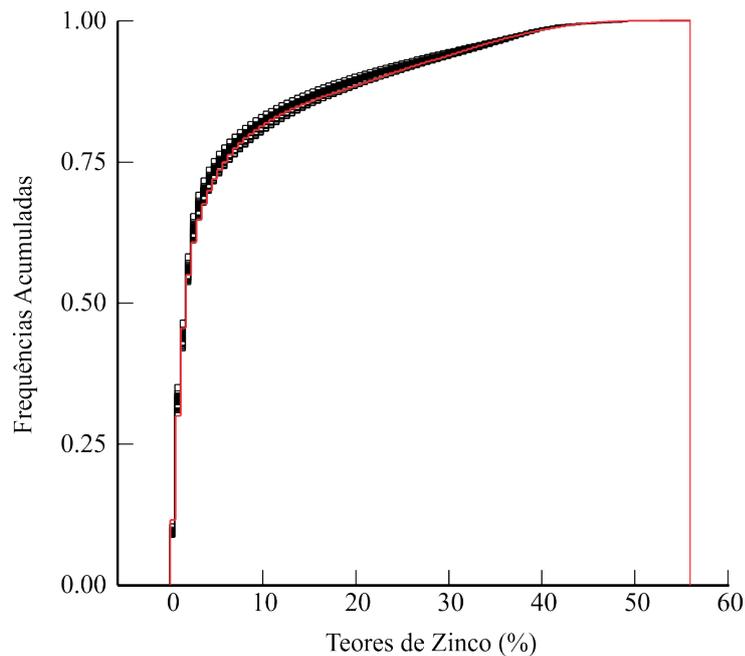


Realizada a simulação no espaço normal, os dados devem ser agora retransformados para o espaço original das amostras, aplicando a função inversa de anamorfose obtida pela transformação *normal score* apresentada na [subseção 3.4.1](#). Para checar as simulações aplica-se técnicas de validação das realizações, apresentadas na seguinte subseção.

### 3.5.1 Validação das realizações

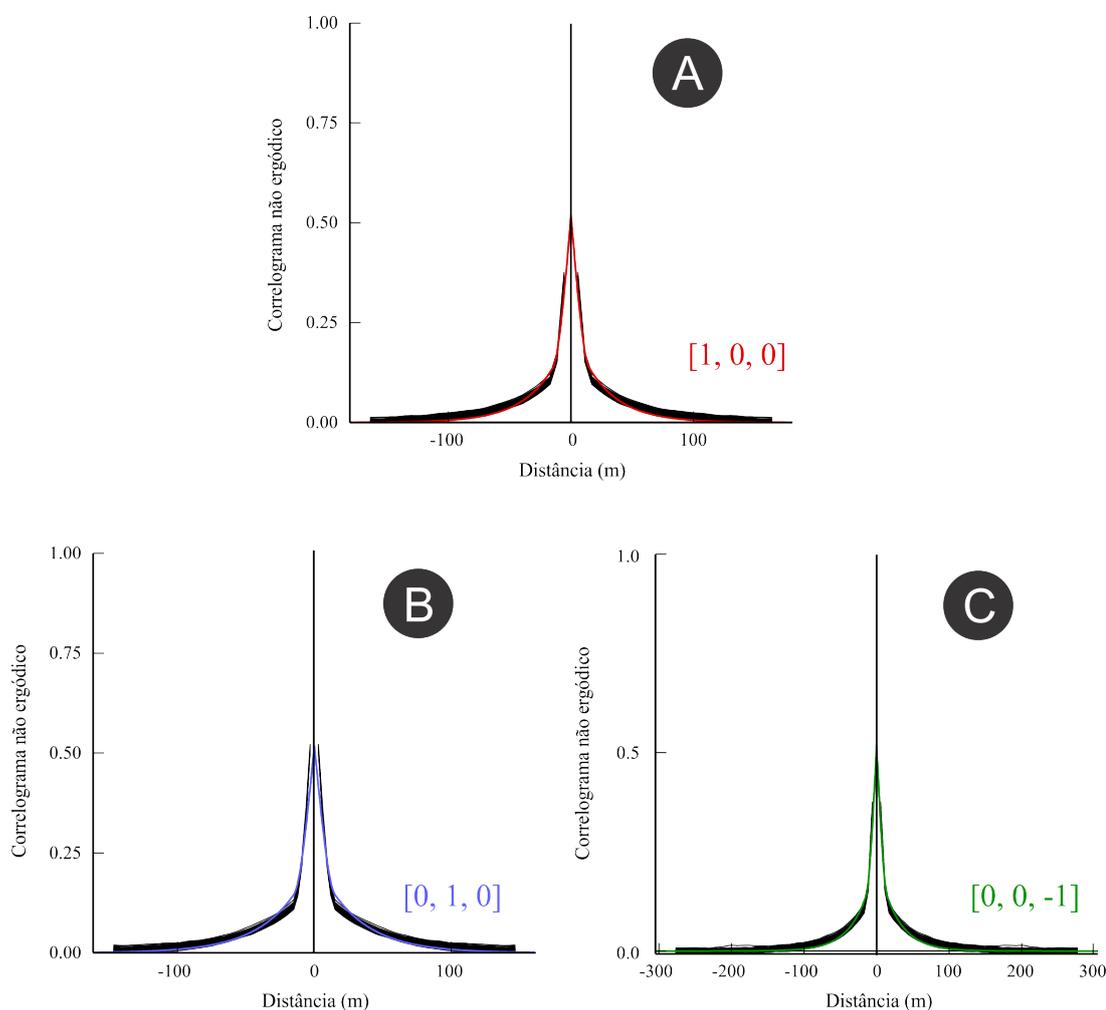
Realizadas as simulações, deve-se verificar a aderência dos cenários garantindo a reprodutibilidade das continuidades espaciais e proporções do fenômeno mineralizante, gerados pelas  $n$  realizações simuladas. A [Figura 29](#) ilustra a validação por comparação entre a distribuição desagrupada dos dados de Zinco (vermelho) e os 50 cenários simulados (preto). Verificou-se uma boa reprodutibilidade dos modelos.

Figura 29 – Validação da reprodução da distribuição dos teores do Zinco. Em preto estão representadas as distribuições das realizações e em vermelho o a distribuição desagrupada do Zinco.



A [Figura 30](#) ilustra a validação das continuidades espaciais medidas pelo correlograma não ergódico, nos quais: (A) maior, (B) média e (C) menor continuidade. A validação ocorreu no espaço normal, e pela análise destes, pode-se dizer que os modelos simulados honraram a continuidade espacial da variável de entrada.

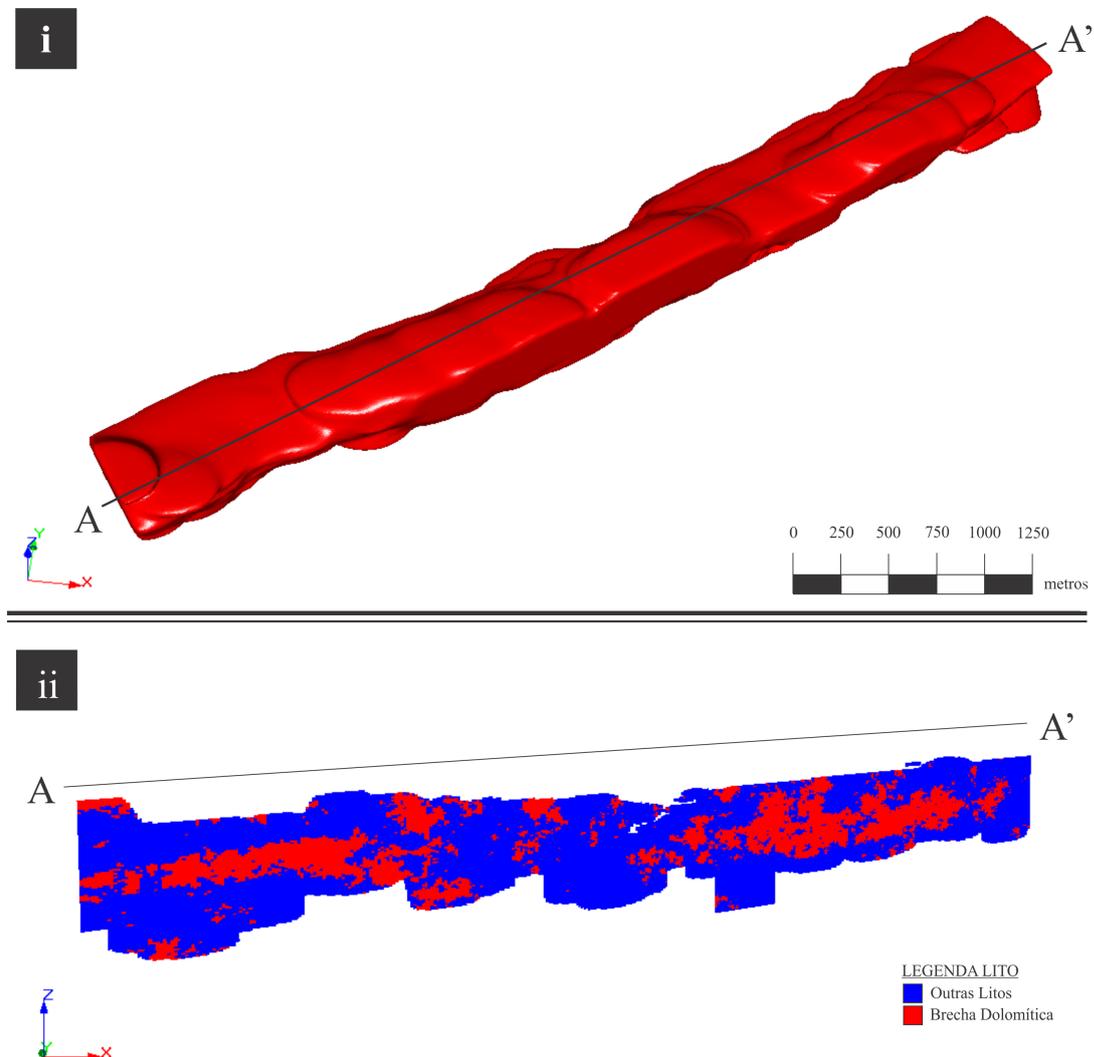
Figura 30 – Validação da continuidade espacial dos cenários simulados. Em (A) o variograma da direção de maior continuidade  $[1, 1, 0]$ , em (B) a direção de média continuidade  $[-1, 1, 0]$  e em (C) a direção de menor continuidade  $[-1, 0, -1]$ .



### 3.6 Simulação do litotipo BXD por indicadores

Diferentemente das simulações sequenciais baseadas em variáveis contínuas, a simulação dos indicadores não exige a normalização dos dados. Para este procedimento de simulação, foi utilizado o mesmo *grid* com dimensões de blocos  $4 \times 4 \times 3$  apresentado na [subseção 3.4.3](#) e os parâmetros introduzidos nas seções anteriores. Como resultado deste procedimento, obteve-se 50 modelos de blocos preenchidos com os valores 0 e 1, em que o valor 1 (vermelho) representa a litologia BXD e 0 (azul) representa uma outra litologia qualquer ([Figura 31](#)).

Figura 31 – Em [i] está ilustrado a envoltória dos blocos simulados e em [ii] está ilustrado a seção A-A' realizada no modelo de blocos do cenário número 1.

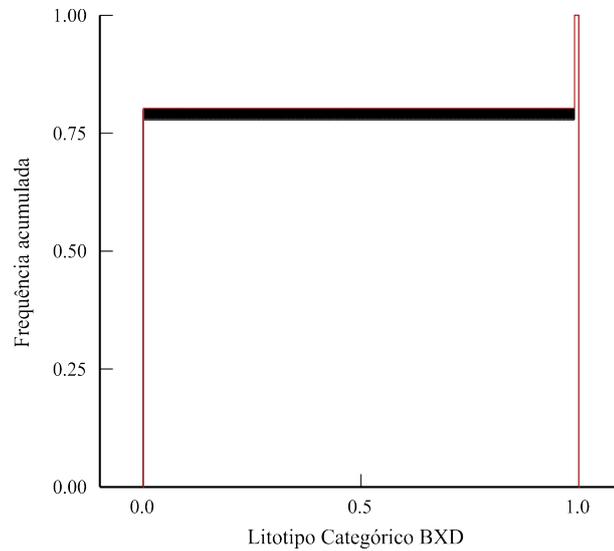


### 3.6.1 Validação das realizações

A validação das simulações ocorre pela comparação das distribuições de proporções e verificação da reprodutibilidade das continuidades espaciais, quando comparadas aos dados de amostra.

Tanto nas variáveis contínuas quanto nas categóricas, o procedimento de validação das realizações obtidas pela simulação dos indicadores, também compreendem a análise de reprodução do histograma e a análise da continuidade espacial nos cenários simulados. A [Figura 32](#) ilustra o histograma desagrupado dos dados de entrada da litologia BXD, em vermelho, e em preto, estão ilustrados os histogramas de cada um dos 50 cenários. Os

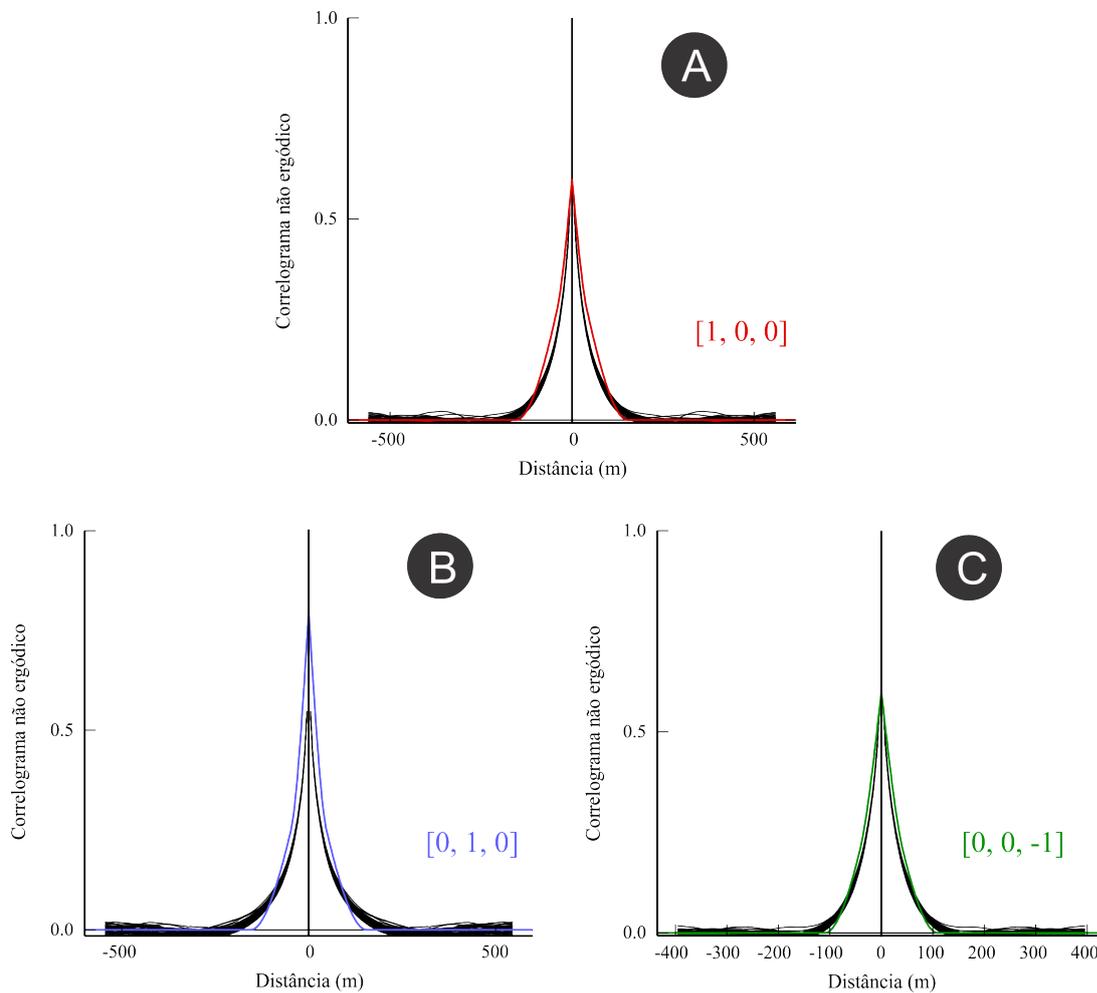
Figura 32 – Validação dos cenários simulados da variável BXD. As simulações mantiveram as proporções litológicas desagrupadas. Em preto, os 50 cenários. Em vermelho, a proporção desagrupada pelo método vizinho mais próximo.



cenários mantiveram, mais ou menos, a proporção das categorias dos dados de entrada.

A [Figura 33](#) ilustra a validação das continuidades espaciais medidas pelo correlograma não ergódico, nos quais: (A) maior, (B) média e (C) menor continuidade. Os modelos de continuidade indicam uma boa reprodução das continuidades espaciais, honrando os dados de entrada.

Figura 33 – Validação da continuidade espacial das realizações medida pelo correlogramas não ergódico.



### 3.6.2 Tratamento dos modelos simulados

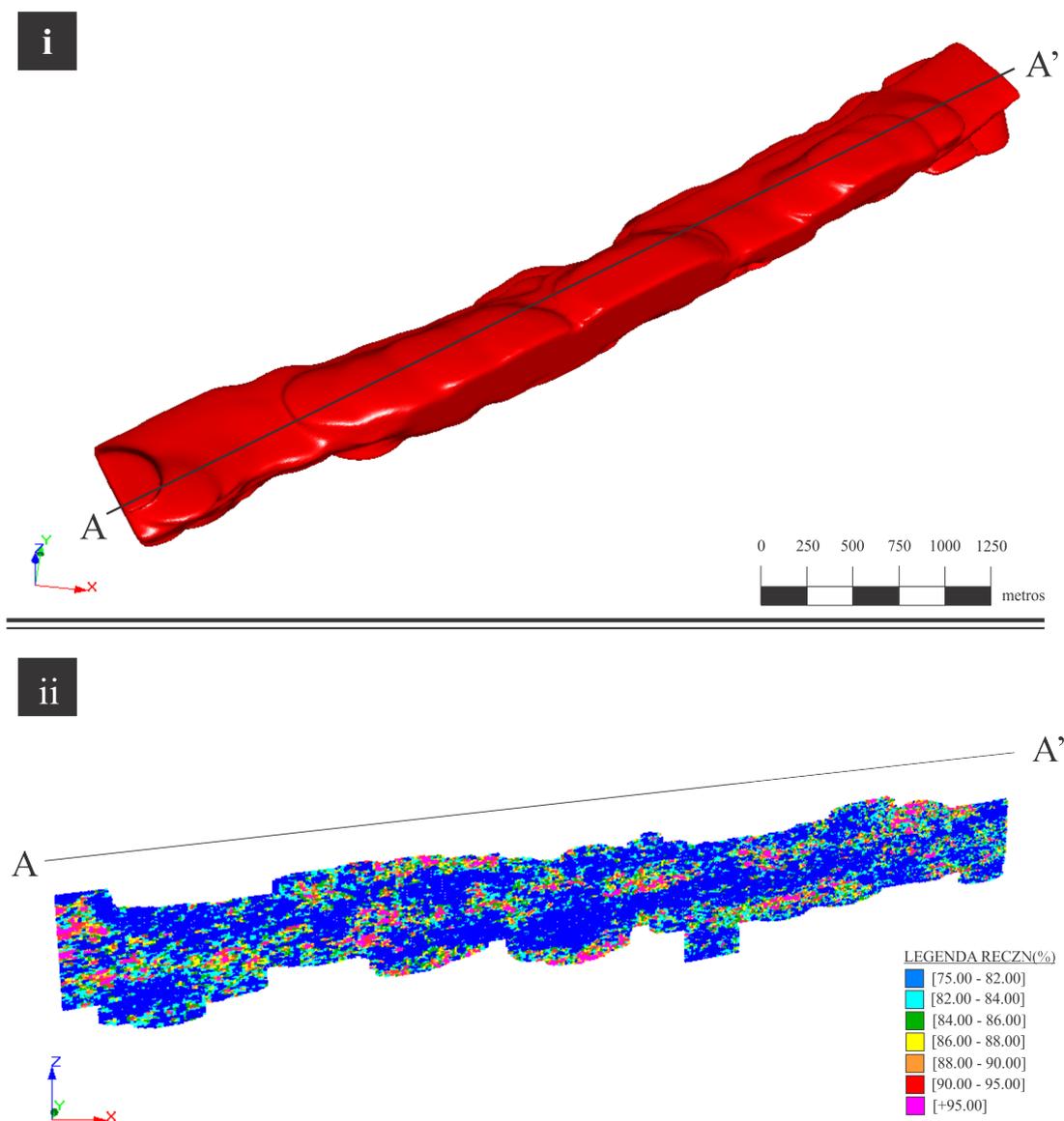
Uma mudança de suporte foi efetuada transferindo os dados simulados na escala ponto ( $4 \times 4 \times 3$ ) para blocos de dimensões  $12 \times 12 \times 3$ . Para cada um dos blocos maiores foram calculadas as médias dos blocos menores que o pertencem. Este procedimento suaviza as variâncias dos teores de Zinco simulados e nos cenários de BXD tem-se a proporção do litotipo em cada bloco.

A partir dos modelos reescalados foi possível aplicar a regressão linear multivariada, obtendo assim 50 cenários com informações geometalúrgicas.

### 3.7 Modelagem da recuperação metalúrgica por regressão

Após o procedimento de reescalonamento dos volumes dos blocos, foi aplicado aos modelos simulados de Zinco e BXD a regressão linear multivariada, apresentada na seção 3.3, gerando 50 cenários de recuperação do Zinco. A Figura 34 ilustra a primeira realização do modelo de recuperação metalúrgica do Zinco, obtido pela regressão linear multivariada.

Figura 34 – Em [i] está ilustrado a envoltória dos blocos simulados e em [ii] o modelo do primeiro cenário da recuperação metalúrgica do Zinco obtido pela regressão multivariada.

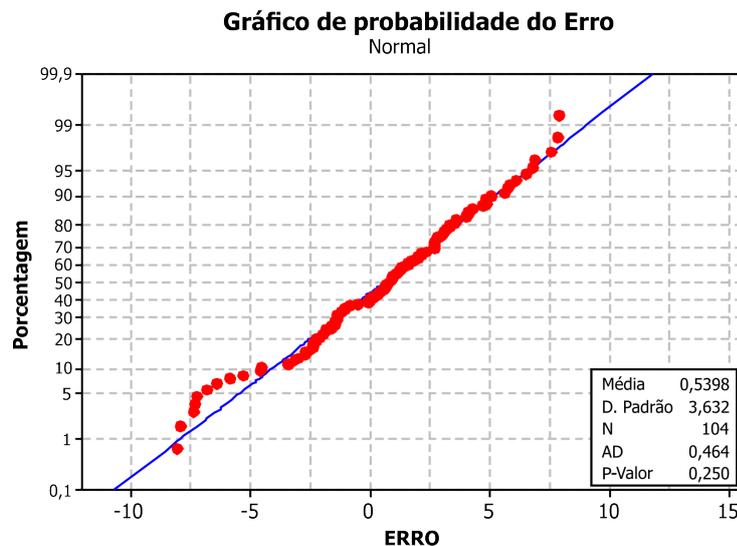


### 3.8 Validação do modelo geometalúrgico

Como forma de checar os resultados obtidos, realizou-se a validação dos modelos geometalúrgicos. O modelo obtido por regressão apresenta problemas com os métodos convencionais de validação (histograma e continuidade espacial) por questões de suporte amostral, uma vez que os testes laboratoriais foram efetuados numa escala muito grande e em blocos (estopes) de diferentes dimensões. Portanto, a reconciliação foi o método encontrado para avaliar os resultados do modelo regredido, confrontando os valores laboratoriais de cada teste de bancada com seus respectivos valores obtidos pelo modelo regredido.

Foi realizado o teste de normalidade do erro onde foi checado o modelo de regressão aplicado aos dados laboratoriais. O resultado esperado é que o valor-p seja maior do que 0.05, e assim podemos dizer que a suposição de normalidade é aceita pelo teste de Anderson Darling. Pela [Figura 35](#) pode-se afirmar que a normalidade do erro foi aceita.

Figura 35 – Gráfico de probabilidade ilustrando o teste de normalidade do erro. O valor-p superior à 0.05 confirma a suposição de normalidade pelo teste de Anderson Darling.

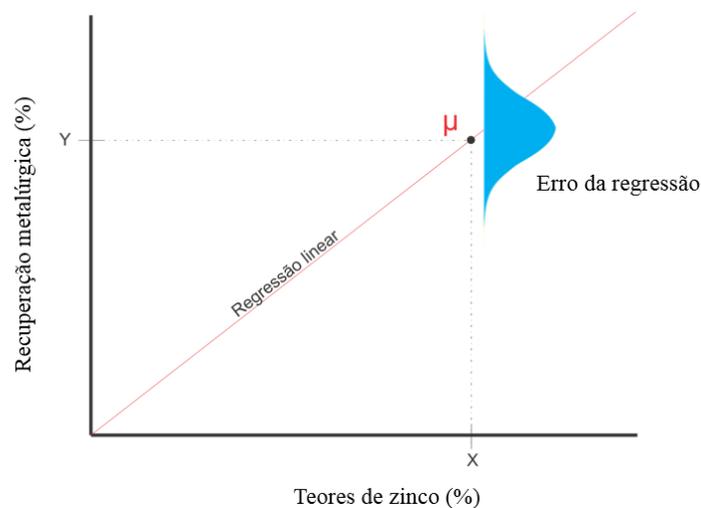


O modelo de incertezas foi obtido pelos métodos de simulação geoestatística. Às incertezas das variáveis de entrada se faz necessário adicionar também a incerteza da regressão linear, o erro da regressão. A ideia nesta etapa é realizar uma simulação do erro da regressão, e os itens enumerados a seguir descrevem o passo a passo para inserção do erro de regressão ao modelo de simulação regredido.

- i O erro da regressão foi obtido pela Equação 2.20 com 90% de confiança e apresenta desvio padrão de 4.59%;

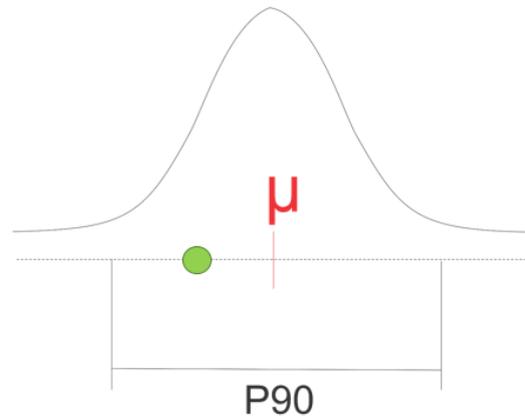
- ii [Montgomery, Peck e Vining \(2012\)](#) citam que a distribuição do erro é normal. Com os dados de desvio padrão e a média (valor obtido pela regressão das variáveis), elaborou-se uma distribuição gaussiana do erro de variância ( $\sigma^2$ ) calculada e média ( $\mu$ ) correspondente ao valor do modelo de regressão. A [Figura 36](#) ilustra um exemplo da normalidade do erro em um caso univariado..

Figura 36 – Ilustração da normalidade do erro da regressão linear no caso de um modelo univariado. O  $\mu$  representa o valor médio da recuperação de zinco obtida pelo modelo de regressão linear.



- iii Desta distribuição gaussiana do erro, foram sorteados aleatoriamente valores para a recuperação metalúrgica no intervalo P90 ([Figura 37](#));

Figura 37 – Nova distribuição normal gerado pelos valores da média da recuperação metalúrgica e erro de regressão calculado. O círculo verde ilustra o valor sorteado aleatoriamente dentro do intervalo P90 da distribuição.



- iv As etapas foram repetidas em todos os estopes onde haviam informações laboratoriais até que o processo compreendesse todos os cenários simulados.

Apesar dos 104 testes metalúrgicos, somente 96 blocos de estope puderam ser validados. A fim de preservar a confidencialidade dos dados, os estopes originais foram renomeados para e1, e2, ..., e96. Foi plotado um gráfico de caixas (Figura 38) com os valores sorteados da distribuição dos erros. Os pontos em vermelho representam os testes de laboratório nos quais o modelo de regressão foram aprovados e em azul estão ilustrados os modelos reprovados. Todos os valores extremos do modelo foram desconsiderados.

Os resultados da validação indicam uma boa confiabilidade do modelo de regressão para acessar a incerteza da recuperação metalúrgica, uma vez que dos 96 testes avaliados apenas 10 resultados foram incompatíveis. Esse resultado ratifica o modelo geometalúrgico obtido pela regressão.

Dentre as vantagens da utilização da simulação, vale destacar a obtenção de intervalos de incertezas atribuídas às variáveis de entrada do modelo de regressão. Desta forma, têm-se um modelo regredido contemplando tanto os erros da regressão quanto a influência locais das variáveis. As demais considerações e ideias para trabalhos futuros serão discutidas no próximo capítulo.



### 3.9 Obtenção do modelo de incertezas da recuperação metalúrgica para novos estopes

A obtenção do modelo de incerteza para qualquer bloco de lavra pode ser obtida seguindo os seguintes passos, considerando um modelo de 50 cenários:

- i Selecionar os blocos de cada cenário simulados dos teores e litologia (indicadores) dentro do estope em questão;
- ii Aplicar o modelo de regressão linear multivariada em cada um dos cenários, obtendo assim um valor médio de recuperação em cada cenário;
- iii Confeccionar para cada cenário uma distribuição normal do erro de regressão linear, utilizando a média do bloco e o desvio padrão do erro calculado;
- iv Sortear um valor da distribuição normal do erro com intervalo de 90% de confiança.

Ao final serão obtidos 50 valores de recuperação, o quais podem ser organizados e apresentado conforme ilustrado anteriormente pela [Figura 38](#). À estes valores de recuperação metalúrgica estão atribuídos tanto a incerteza das variáveis de entrada como a incerteza do regressão linear.

## 4 Considerações finais

Essa dissertação propôs como meta investigar a incerteza associado à recuperação metalúrgica do Zinco como método auxiliar planejamento de mina. Para tanto, foram apresentados dois objetivos específicos:

- i Identificar por meio de regressão multivariada um modelo que melhor explique a resposta geometalúrgica;
- ii Conduzir o estudo de caso aplicando o modelo de regressão aos modelos de simulação geoestatística quantificando a incerteza da resposta e medindo sua aderência aos valores de produção reais.

O modelo de regressão foi obtido e validado pelo métodos estatísticos multivariados, apresentados no Capítulo 2. No estudo de caso, o modelo de regressão da recuperação de Zinco foi ratificado, visto que na reconciliação do resultados, 90% dos blocos de estope estudados apresentaram-se condizentes com os resultados dos testes laboratoriais.

Pela análise do *boxplot* da [Figura 38](#) percebe-se que os *ranges* de incerteza do modelo regredido são elevados. Isso se deve à pequena quantidade de amostras e à deficiência de dados dos testes de bancada. Para melhorar e diminuir a incerteza do modelo seria necessário uma quantidade maior de informação e uma refinação do modelo de regressão.

A prática de estudos geometalúrgicos vêm se provando uma etapa importante na análise de riscos do empreendimento mineiro. O bom entendimento do comportamento mineral na planta de beneficiamento traz diversos benefícios em várias fases do empreendimento, desde auxiliar na decisão do investidor até nas etapas de sequenciamento de lavra. Na literatura, não é raro encontrar referências que citem a geometalurgia como uns dos principais responsáveis pelo sucesso ou fracasso de uma mina.

Na fase inicial do projeto, o entendimento dos riscos geometalúrgicos podem auxiliar na decisão do investidor e impactar no futuro do empreendimento. Nas fases de lavra, a informação da recuperação *in situ* se faz importante no planejamento e sequenciamento da mina, podendo serem feitos planos que busquem blocos com menores riscos no beneficiamento. Outro ganho pode ocorrer na própria planta, melhorando a qualidade do *blend* de minérios aumentando a massa de metal recuperado e reduzindo desperdícios de reagentes.

## 4.1 Recomendações para trabalhos futuros

Como recomendação para trabalhos futuros, sugere-se os seguintes itens:

- Separação dos domínios litológicos, gerando modelos simulados de teores condicionados aos domínios litológicos;
- Realizar mais testes geometalúrgicos;
- Analisar com mais detalhe as características do minério, como a descrição mineralógica e até mesmo incluir informações geotécnicas ao modelo;
- Utilizar informações de modelo do minério interpretado pelo time de geologia de mina e categorizá-lo pelo domínio do litotipo. Aplicar o modelo de categorias na equação de regressão.

# Referências

- ADAMS, M. D. *Gold Ore Processing: Project Development and Operations*. [S.l.]: Elsevier, 2016. v. 15. 1040 p. Citado na página 23.
- ALABERT, F. G. *Stochastic Imaging of Spatial Distributions Using Hard and Soft Information*. 416 p. Dissertação (Mestrado), 1987. Citado 2 vezes nas páginas 21 e 40.
- BABINSKI, M. et al. Isotope geochemistry of the mafic dikes from the vazante nonsulfide zinc deposit, brazil. *Journal of South American Earth Sciences*, Elsevier, v. 18, n. 3-4, p. 293–304, 2005. Citado 4 vezes nas páginas 45, 46, 47 e 48.
- BOISVERT, J. B. et al. Geometallurgical modeling at olympic dam mine, south australia. *Mathematical Geosciences*, Springer, v. 45, n. 8, p. 901–925, 2013. Citado na página 23.
- BONATO, F. Planejamento de lavra com auxílio de simulação geoestatística. *Universidade Federal do Rio Grande do Sul, Dissertação de Mestrado*, p. 87, 2000. Citado na página 64.
- CHILES, J.-P.; DELFINER, P. *Geostatistics: modeling spatial uncertainty*. [S.l.: s.n.], 2009. 734 p. Citado 4 vezes nas páginas 39, 41, 42 e 43.
- COSTA, J. F. C. L. *Developments in recoverable reserves estimation and ore body modelling*. 333 p. Tese (Doutorado), 1997. Citado na página 64.
- COX, D. R.; HINKLEY, D. V. *Theoretical statistics*. [S.l.: s.n.], 1979. 528 p. Citado na página 30.
- CUNHA, S. B. D.; CARVAJAL, S. *Estatística Basica - A Arte de Trabalhar com Dados*. [S.l.: s.n.], 2009. 288 p. Citado 2 vezes nas páginas 30 e 31.
- DARDENNE, M. et al. Evolução tectono-sedimentar do grupo vazante no contexto da faixa de dobramentos Brasília. In: RESUMOS, SBG BELO HORIZONTE. *Congresso Brasileiro Geologia*. [S.l.], 1998. v. 40, p. 26. Citado na página 45.
- DARDENNE, M. et al. The Brasília fold belt. *Tectonic Evolution of South America*, SBG Rio de Janeiro, v. 1, p. 231–263, 2000. Citado na página 45.
- DARDENNE, M. A. Síntese sobre a estratigrafia do grupo bambuí no Brasil central. In: *Congresso Brasileiro de Geologia*. [S.l.: s.n.], 1978. v. 30, n. 1978, p. 597–610. Citado na página 47.
- DEKKING, F. M. *A Modern Introduction to Probability and Statistics: Understanding why and how*. [S.l.: s.n.], 2005. 488 p. Citado na página 28.
- DEUTSCH, C. Geostatistical modelling of geometallurgical variables—problems and solutions. In: *The second AusIMM international geometallurgy conference*. [S.l.: s.n.], 2013. v. 30, p. 7–15. Citado na página 24.
- DEUTSCH, C. V.; JOURNAL, A. *Geostatistical software library and user's guide*. [S.l.]: Oxford University Press, New York, 1998. Citado 5 vezes nas páginas 37, 39, 40, 58 e 59.

- DEUTSCH, J. L. *Multivariate Spatial Modeling of Metallurgical Rock Properties*. Tese (Doutorado) — University of Alberta, 2015. Citado na página 24.
- EMERY, X.; LANTUÉJOUL, C. Tbsim: A computer program for conditional simulation of three-dimensional gaussian random fields via the turning bands method. *Computers & Geosciences*, Elsevier, v. 32, n. 10, p. 1615–1628, 2006. Citado 2 vezes nas páginas 41 e 42.
- GOOVAERTS, P. *Geostatistics for natural resources evaluation*. [S.l.: s.n.], 1997. 483 p. Citado 4 vezes nas páginas 37, 39, 40 e 62.
- HAIR, J. F. et al. *Análise multivariada de dados*. [S.l.]: Bookman Editora, 2009. Citado 2 vezes nas páginas 27 e 30.
- ISAAKS, E. H. *The application of Monte Carlo methods to the analysis of spatially correlated data*. 213 p. Tese (Doutorado), 1990. Citado 2 vezes nas páginas 21 e 38.
- ISAAKS, E. H.; SRIVASTAVA, R. M. *Applied geostatistics*. [S.l.]: Oxford University Press, 1989. 561 p. Citado na página 52.
- JOURNEL, A. G. Geostatistics for conditional simulation of ore bodies. *Economic Geology*, SecG, v. 69, n. 5, p. 673–687, 1974. Citado 2 vezes nas páginas 42 e 65.
- JOURNEL, A. G. Nonparametric estimation of spatial distributions. *Journal of the International Association for Mathematical Geology*, Springer, v. 15, n. 3, p. 445–468, 1983. Citado na página 40.
- JOURNEL, A. G. *Fundamentals of geostatistics in five lessons*. [S.l.: s.n.], 1989. v. 8. 40 p. Citado na página 56.
- JOURNEL, A. G.; HUIJBREGTS, C. J. *Mining geostatistics*. [S.l.]: Academic press, 1978. 600 p. Citado na página 43.
- KUTNER, M. H.; NACHTSHEIM, C.; NETER, J. *Applied linear regression models*. [S.l.: s.n.], 2004. 701 p. Citado na página 27.
- LAMBERG, P. Particles-the bridge between geology and metallurgy. In: *Konferens i mineralteknik 2011: 08/02/2011-09/02/2011*. [S.l.: s.n.], 2011. p. 1–17. Citado 2 vezes nas páginas 23 e 24.
- LANTUÉJOUL, C. *Geostatistical simulation: models and algorithms*. [S.l.: s.n.], 2002. 256 p. Citado na página 41.
- MATHERON, G. The intrinsic random functions and their applications. *Advances in applied probability*, Cambridge University Press, v. 5, n. 3, p. 439–468, 1973. Citado na página 65.
- MONTEIRO, L. V. S. et al. The vazante zinc mine, minas gerais, brazil; constraints in willemitic mineralization and fluid evolution. *Exploration and Mining Geology*, Canadian Institute of Mining, Metallurgy & Petroleum, v. 8, n. 1-2, p. 21–42, 1999. Citado na página 48.
- MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. *Introduction to linear regression analysis*. [S.l.: s.n.], 2012. v. 821. 645 p. Citado 10 vezes nas páginas 27, 28, 29, 30, 31, 32, 33, 34, 35 e 73.

- MONTGOMERY, D. C.; RUNGER, G. C. *Applied statistics and probability for engineers*. [S.l.: s.n.], 2010. 784 p. Citado na página 31.
- OLEA, R. A. *Geostatistics for engineers and earth scientists*. [S.l.: s.n.], 2012. 324 p. Citado 2 vezes nas páginas 37 e 38.
- PETROPOULOS, G.; SRIVASTAVA, P. K. *Sensitivity Analysis in Earth Observation Modelling*. [S.l.: s.n.], 2016. 448 p. Citado na página 36.
- PINHEIRO, J. et al. *Probabilidade e estatística: quantificando a incerteza*. [S.l.: s.n.], 2013. 568 p. Citado na página 32.
- ROSSI, M. E.; DEUTSCH, C. V. *Mineral resource estimation*. [S.l.: s.n.], 2013. 332 p. Citado 8 vezes nas páginas 24, 38, 39, 40, 41, 42, 43 e 44.
- SALTELLI, A. et al. *Sensitivity analysis in practice: a guide to assessing scientific models*. [S.l.: s.n.], 2004. 219 p. Citado na página 36.
- TRAN, T. T. Improving variogram reproduction on dense simulation grids. *Computers & Geosciences*, Elsevier, v. 20, n. 7-8, p. 1161–1168, 1994. Citado na página 39.
- TRIOLA, M. F. *Essentials of statistics*. [S.l.: s.n.], 2008. 696 p. Citado 3 vezes nas páginas 28, 32 e 34.
- WILLIAMS, S. R.; RICHARDSON, J. Geometallurgical mapping: A new approach that reduces technical risk. In: *Proceedings 36th Annual Meeting of the Canadian Mineral Processors*. [S.l.: s.n.], 2004. p. 241–268. Citado na página 24.
- ZAGAYEVSKIY, Y. V.; DEUTSCH, C. V. Updated code for sensitivity analysis based on regression. *CCG Annual Report*, v. 13, n. 40, p. 1–20, 2011. Citado 2 vezes nas páginas 36 e 37.