

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE MINAS, METALÚRGICA E DE MATERIAIS

MELISSA ABRÃO ZENI

ANÁLISE DO DESEMPENHO DA KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA  
NA PRESENÇA DE ERROS AMOSTRAIS E VALORES EXTREMOS

Porto Alegre

2019

MELISSA ABRÃO ZENI

ANÁLISE DO DESEMPENHO DA KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA  
NA PRESENÇA DE ERROS AMOSTRAIS E VALORES EXTREMOS

*Dissertação submetida ao Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais da Universidade Federal do Rio Grande do Sul, como requisito parcial à obtenção do título de Mestre em Engenharia, modalidade Acadêmica. Área de Concentração: Tecnologia Mineral.*

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

*Coorientador: Prof. Dr. Áttila Leães Rodrigues*

Porto Alegre

2019

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Rui Vicente Oppermann

Vice-Reitora: Jane Fraga Tutikian

ESCOLA DE ENGENHARIA

Diretor: Luiz Carlos Pinto da Silva filho

Vice-Diretora: Carla Schwengber ten Caten

PROGRAMA DE PÓS-GRADUAÇÃO

Coordenador: Afonso Reguly

Abrao Zeni, Melissa

Análise do Desempenho da Krigagem com Variância do Erro de Medida na Presença de Erros Amostrais e Valores Extremos / Melissa Abrao Zeni. -- 2019. 82 f.

Orientador: João Felipe Coimbra Leite Costa.

Coorientador: Áttila Leães Rodrigues.

Dissertação (Mestrado) -- Universidade Federal do Rio Grande do Sul, Escola de Engenharia, Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais, Porto Alegre, BR-RS, 2019.

1. Erro amostral. 2. Valores Extremos. 3. Krigagem. I. Coimbra Leite Costa, João Felipe, orient. II. Leães Rodrigues, Áttila, coorient. III. Título.

MELISSA ABRÃO ZENI

ANÁLISE DO DESEMPENHO DA KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA  
NA PRESENÇA DE ERROS AMOSTRAIS E VALORES EXTREMOS

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.

---

Prof. Dr. João Felipe Coimbra Leite Costa  
Orientador

---

Prof. Dr. Carlos Pérez Bergmann  
Coordenador do PPGE3M

Aprovado em 1º de Julho de 2019

BANCA EXAMINADORA:

Dr<sup>a</sup>. Camilla Zacché da Silva – LPM/UFRGS

Prof. Dr. Marcel Antonio Arcari Bassani – DEMIN/UFRGS

Prof<sup>a</sup>. Dr<sup>a</sup>. Vanessa Cerqueira Koppe – DEMIN/UFRGS

## **AGRADECIMENTOS**

Em primeiro lugar agradeço a Universidade Federal do Rio Grande do Sul, pela oportunidade proporcionada de forma gratuita e de excelente qualidade, desde a graduação até a conclusão deste mestrado. Agradeço também ao CNPq pelo apoio financeiro.

Agradeço com carinho meu orientador Prof. João Felipe, coorientador Prof. Áttila, e ao colega Victor Miguel Silva, pelo interesse em ensinar, por sempre dividirem seus conhecimentos, e por toda atenção que me foi dedicada.

Grata aos meus colegas do LPM, pelo companheirismo e bons momentos.

E em especial, agradeço aos meus amigos e a minha família, principalmente meus pais Maristéla e Edison, e minha irmã Marília, pelo suporte, incentivo e compreensão.

## RESUMO

Os recursos minerais são tipicamente mensurados por múltiplas fontes de dados. Esses dados possuem confiabilidade variada, e para utilizá-los na estimativa, é necessário metodologias que os integrem de maneira adequada. Nesta dissertação, é proposta a investigação da krigagem com variância do erro de medida, a qual penaliza com menor peso as amostras com maior incerteza associada, sendo uma alternativa à abordagem onde dados imprecisos são agrupados como dados secundários. A metodologia é analisada através da sua aplicação utilizando dados sintéticos com diferentes configurações do erro, elaborados a partir de um banco de dados original. A variância do erro de medida foi obtida através da modelagem das diferentes estruturas dos variogramas. Os resultados das estimativas foram comparados à krigagem ordinária, utilizando como modelo de referência a variável original reblocada no mesmo suporte das estimativas. Essa comparação é realizada através da análise dos valores máximos e mínimos estimados, diagramas de dispersão, coeficiente de correlação, análise de deriva e erro quadrático médio. De modo geral, a krigagem com variância do erro de medida forneceu uma melhor acurácia das estimativas comparada à krigagem ordinária. Dependendo da disposição espacial do erro, e de sua correlação com os teores, entretanto, poderá ocorrer enviesamento das estimativas. A investigação da krigagem com variância do erro de medida como alternativa ao se trabalhar com valores extremos apresentou bons resultados, uma vez que mantêm o teor de todas as amostras, e distribui peso a estas conforme sua confiabilidade.

Palavras-chave: Erro amostral. Valores Extremos. Krigagem.

## ABSTRACT

Mineral resources are measured by multiple data sources. These data have varied reliability, and to use them in the estimation requires methodologies that integrate them adequately. In this dissertation, it is proposed the investigation of kriging with variance of the measurement error, which penalizes with less weight the samples associated with greater uncertainty, being an alternative to the approach where imprecise data are grouped as secondary data. This methodology is analyzed through its application using synthetic data with different error configurations, generated from an original database. The variance of the measurement error was obtained by modeling the different structures of the variograms. The results of the estimations were compared to ordinary kriging, using as reference model the original variable reblocked in the same estimates support. This comparison is performed through the analysis of estimated maximum and minimum values, scatterplots, correlation coefficient, swath plots and mean square error. In general, kriging with variance of the measurement error provided a better accuracy of the estimates compared to ordinary kriging. However, depending on the spatial arrangement of the error, and its correlation with the grades, bias of the estimates may occur. The investigation of kriging with variance of the measurement error as an alternative to dealing with outliers presented good results, since it maintains the grades of all samples, and distributes weight to them according to their reliability.

Keywords: Sampling Error. Outliers. Kriging.

## LISTA DE FIGURAS

FIGURA 2.1 – RELAÇÃO ENTRE OS ERROS DE AMOSTRAGEM. FONTE: MODIFICADO DE MINNITT (2007). .....	7
FIGURA 2.2 – CONCEITOS DE PRECISÃO E ACURÁCIA: A) IMPRECISO E INACURADO; B) IMPRECISO E ACURADO; C) PRECISO E INACURADO; D) PRECISO E ACURADO. ....	8
FIGURA 2.3 – CARACTERÍSTICAS DE UM VARIOGRAMA. FONTE: MODIFICADO DE ROSSI E DEUTSCH (2014). ....	17
FIGURA 2.4 – A) DADOS SEM ERRO ( $i = 1, \dots, 5$ ), DADO COM ERRO ASSOCIADO ( $i = 6$ ), E VALOR A SER ESTIMADO $z(x_0)$ . B) INFLUÊNCIA DA INCERTEZA SOBRE OS PESOS. FONTE: MODIFICADO DE DELHOMME (1976). ....	21
FIGURA 2.5 – EXEMPLOS UNIDIMENSIONAIS DE INTERPOLAÇÃO E FITTING. A) DADOS SINTÉTICOS GERADOS A PARTIR DE UMA FUNÇÃO (LINHA TRACEJADA). B) RESULTADO DA INTERPOLAÇÃO POR KRIGAGEM. A EQUAÇÃO DE INTERPOLAÇÃO É PLOTADA COMO A LINHA SÓLIDA, ENQUANTO A VARIÂNCIA DO ERRO DE KRIGAGEM É MOSTRADA COMO A FAIXA SOMBREADA. C) O MESMO QUE EM (B) PORÉM COM EFEITO PEPITA. D) RESULTADO DA ESTIMATIVA AO SE AJUSTAR O ESTIMADOR AOS DADOS COM VARIÂNCIA DO ERRO DE MEDIDA. FONTE: MODIFICADO DE PRESS ET AL. (2008). ....	22
FIGURA 3.1 – A) HISTOGRAMA DO BANCO DE DADOS EXAUSTIVO DA VARIÁVEL V. B) HISTOGRAMA DA VARIÁVEL V APÓS A MODIFICAÇÃO PARA BLOCOS DE 10 X 10 METROS. ....	26
FIGURA 3.2 – A) DISTRIBUIÇÃO ESPACIAL DA VARIÁVEL EXAUSTIVA V. B) DISTRIBUIÇÃO ESPACIAL DE V APÓS TRANSFORMAÇÃO PARA SUPORTE BLOCO. ESCALA DE CORES EM PPM. ....	27
FIGURA 3.3 – A) HISTOGRAMA DO BANCO DE DADOS AMOSTRAL DA VARIÁVEL V. B) DISTRIBUIÇÃO ESPACIAL DAS AMOSTRAS DE V. ..	28
FIGURA 3.4 – A) LOCALIZAÇÃO DAS AMOSTRAS “HARD DATA”. B) LOCALIZAÇÃO DAS AMOSTRAS “SOFT DATA”. ....	29
FIGURA 3.5 – A) DIAGRAMA DE DISPERSÃO DOS 275 DADOS ORIGINAIS DE V VERSUS 275 DADOS DO CENÁRIO 5 ( <i>STD.DEV_1</i> ). B) DISTRIBUIÇÃO ESPACIAL DO CENÁRIO_5 ( <i>STD.DEV_1</i> ). ....	30
FIGURA 3.6 – A) DIAGRAMA DE DISPERSÃO DOS 275 DADOS ORIGINAIS DE V VERSUS 275 DADOS DO CENÁRIO 5 ( <i>STD.DEV_2</i> ). B) DISTRIBUIÇÃO ESPACIAL DO CENÁRIO_5 ( <i>STD.DEV_2</i> ). ....	30
FIGURA 3.7 – A) DIAGRAMA DE DISPERSÃO DOS 275 DADOS ORIGINAIS DE V VERSUS 275 DADOS DO CENÁRIO 5 ( <i>STD.DEV_3</i> ). B) DISTRIBUIÇÃO ESPACIAL DO CENÁRIO_5 ( <i>STD.DEV_3</i> ). ....	31
FIGURA 3.8 – VARIOGRAMA OMNIDIRECIONAL <i>HARD DATA</i> (195 AMOSTRAS), MOSTRANDO O EFEITO PEPITA RELATIVO A VARIÂNCIA EM MICROESCALA ( <i>cMS</i> ). ....	32
FIGURA 3.9 – VARIOGRAMAS OMNIDIRECIONAIS DE CENÁRIOS REFERENTES AO ERRO <i>STD.DEV_1</i> , INDICANDO O EFEITO PEPITA RELATIVO A VARIÂNCIA EM MICROESCALA ( <i>cMS</i> ) E RELATIVO A VARIÂNCIA DO ERRO DE MEDIDA ( <i>cME</i> ). ....	33
FIGURA 3.10 – VARIOGRAMAS OMNIDIRECIONAIS DE CENÁRIOS REFERENTES AO ERRO <i>STD.DEV_2</i> , INDICANDO O EFEITO PEPITA RELATIVO A VARIÂNCIA EM MICROESCALA ( <i>cMS</i> ) E RELATIVO A VARIÂNCIA DO ERRO DE MEDIDA ( <i>cME</i> ). ....	33
FIGURA 3.11 – VARIOGRAMAS OMNIDIRECIONAIS DE CENÁRIOS REFERENTES AO ERRO <i>STD.DEV_3</i> , INDICANDO O EFEITO PEPITA RELATIVO A VARIÂNCIA EM MICROESCALA ( <i>cMS</i> ) E RELATIVO A VARIÂNCIA DO ERRO DE MEDIDA ( <i>cME</i> ). ....	33
FIGURA 3.12 – A) ELIPSE DE BUSCA SOBRE AS AMOSTRAS DE UM CENÁRIO ESTIMADO POR KRIGAGEM ORDINÁRIA. AS AMOSTRAS, REPRESENTADAS POR QUADRADOS EM VERMELHO, INDICAM SEUS PESOS. B) INFORMAÇÕES SOBRE A ESTIMATIVA POR KRIGAGEM ORDINÁRIA DO BLOCO COM ÍNDICE $x = 22, y = 6$ E $z = 1$ . ....	36
FIGURA 3.13 – A) ELIPSE DE BUSCA SOBRE AS AMOSTRAS DE UM CENÁRIO ESTIMADO POR KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA. AS AMOSTRAS, REPRESENTADAS POR QUADRADOS EM VERMELHO, INDICAM SEUS PESOS. B) INFORMAÇÕES SOBRE A ESTIMATIVA POR KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA DO BLOCO COM ÍNDICE $x = 22, y = 6$ E $z = 1$ . ....	37
FIGURA 3.14 – A) MODELO DE REFERÊNCIA. B) KRIGAGEM ORDINÁRIA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_1</i> . C) KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_1</i> . ....	38
FIGURA 3.15 – A) MODELO DE REFERÊNCIA. B) KRIGAGEM ORDINÁRIA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_2</i> . C) KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_2</i> . ....	38
FIGURA 3.16 – A) MODELO DE REFERÊNCIA. B) KRIGAGEM ORDINÁRIA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_3</i> . C) KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_3</i> . ....	38
FIGURA 3.17 – DIAGRAMA DE DISPERSÃO: OK_CENÁRIOS VERSUS REFERÊNCIA (TRIÂNGULOS EM VERMELHO), E KVME_CENÁRIOS VERSUS REFERÊNCIA (CÍRCULOS EM AZUL). NÍVEL DE INCERTEZA <i>STD.DEV_1</i> . ....	40
FIGURA 3.18 – DIAGRAMA DE DISPERSÃO: OK_CENÁRIOS VERSUS REFERÊNCIA (TRIÂNGULOS EM VERMELHO), E KVME_CENÁRIOS VERSUS REFERÊNCIA (CÍRCULOS EM AZUL). NÍVEL DE INCERTEZA <i>STD.DEV_2</i> . ....	40
FIGURA 3.19 – DIAGRAMA DE DISPERSÃO: OK_CENÁRIOS VERSUS REFERÊNCIA (TRIÂNGULOS EM VERMELHO), E KVME_CENÁRIOS VERSUS REFERÊNCIA (CÍRCULOS EM AZUL). NÍVEL DE INCERTEZA <i>STD.DEV_3</i> . ....	41

FIGURA 3.20 – ERRO QUADRÁTICO MÉDIO DOS CENÁRIOS ( <i>STD.DV_3</i> ) POR FAIXA DE TEOR. LINHA EM AZUL REPRESENTA OS DESVIOS DA KVME, E LINHA EM VERMELHO DESVIOS DA OK.....	42
FIGURA 3.21 – ANÁLISE DE DERIVA DE UM CENÁRIO PERTENCENTE AO ERRO <i>STD.DEV_1</i> . A) ANÁLISE NA DIREÇÃO Y. B) ANÁLISE NA DIREÇÃO X.....	43
FIGURA 3.22 – ANÁLISE DE DERIVA DE UM CENÁRIO PERTENCENTE AO ERRO <i>STD.DEV_2</i> . A) ANÁLISE NA DIREÇÃO Y. B) ANÁLISE NA DIREÇÃO X.....	44
FIGURA 3.23 – ANÁLISE DE DERIVA DE UM CENÁRIO PERTENCENTE AO ERRO <i>STD.DEV_3</i> . A) ANÁLISE NA DIREÇÃO Y. B) ANÁLISE NA DIREÇÃO X.....	44
FIGURA 3.24 – CURVA DE PARAMETRIZAÇÃO TEOR X TONELAGEM. CURVA OBTIDA ATRAVÉS DA MÉDIA DOS CENÁRIOS PERTENCENTE AO ERRO <i>STD.DEV_3</i> . ....	46
FIGURA 3.25 – GRÁFICOS DE PROBABILIDADE DE DOIS CENÁRIOS REFERENTES AO ERRO <i>STD.DEV_3</i> . LINHA TRACEJADA EM LARANJA REPRESENTA CAPPING NO P95, E EM LILÁS P99. ....	47
FIGURA 3.26 – DIAGRAMA DE DISPERSÃO: OK_CENÁRIOS <i>VERSUS</i> REFERÊNCIA (TRIÂNGULOS EM AMARELO), E KVME_CENÁRIOS <i>VERSUS</i> REFERÊNCIA (CÍRCULOS EM AZUL). CAPPING P99.....	49
FIGURA 3.27 – DIAGRAMA DE DISPERSÃO: OK_CENÁRIOS <i>VERSUS</i> REFERÊNCIA (TRIÂNGULOS EM AMARELO), E KVME_CENÁRIOS <i>VERSUS</i> REFERÊNCIA (CÍRCULOS EM AZUL). CAPPING P95.....	50
FIGURA 3.28 – ERRO QUADRÁTICO MÉDIO DO CENÁRIOS ( <i>STD.DV_3</i> ) POR FAIXA DE TEOR. LINHA EM AZUL REPRESENTA OS DESVIOS DA KVME, E LINHA EM VERMELHO DESVIOS DA OK. ....	51
FIGURA 4.1 – A) V – <i>HARD DATA</i> . B) V – <i>SOFT DATA</i> . ....	53
FIGURA 4.2 – A) DIAGRAMA DE DISPERSÃO DOS 195 DADOS ORIGINAIS DE V <i>VERSUS</i> 195 DADOS DO CENÁRIO 5 ( <i>STD.DEV_1</i> ). B) DISTRIBUIÇÃO ESPACIAL DO CENÁRIO_5 ( <i>STD.DEV_1</i> ).....	54
FIGURA 4.3 – A) DIAGRAMA DE DISPERSÃO DOS 195 DADOS ORIGINAIS DE V <i>VERSUS</i> 195 DADOS DO CENÁRIO 5 ( <i>STD.DEV_2</i> ). B) DISTRIBUIÇÃO ESPACIAL DO CENÁRIO_5 ( <i>STD.DEV_2</i> ).....	55
FIGURA 4.4 – A) DIAGRAMA DE DISPERSÃO DOS 195 DADOS ORIGINAIS DE V <i>VERSUS</i> 195 DADOS DO CENÁRIO 5 ( <i>STD.DEV_3</i> ). B) DISTRIBUIÇÃO ESPACIAL DO CENÁRIO_5 ( <i>STD.DEV_3</i> ).....	55
FIGURA 4.5 – VARIOGRAMA OMNIDIRECIONAL <i>HARD DATA</i> (275 AMOSTRAS), MOSTRANDO O EFEITO PEPITA RELATIVO A VARIÂNCIA EM MICROESCALA ( <i>cMS</i> ). ....	57
FIGURA 4.6 – VARIOGRAMAS OMNIDIRECIONAIS DE CENÁRIOS REFERENTES AO ERRO <i>STD.DEV_1</i> (A), ERRO <i>STD.DEV_2</i> (B) E AO ERRO <i>STD.DEV_3</i> (C), INDICANDO O EFEITO PEPITA RELATIVO A VARIÂNCIA EM MICROESCALA ( <i>cMS</i> ) E RELATIVO A VARIÂNCIA DO ERRO DE MEDIDA ( <i>cME</i> ). ....	57
FIGURA 4.7 – A) MODELO DE REFERÊNCIA. B) KRIGAGEM ORDINÁRIA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_1</i> . C) KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_1</i> . ....	60
FIGURA 4.8 – A) MODELO DE REFERÊNCIA. B) KRIGAGEM ORDINÁRIA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_2</i> . C) KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_2</i> . ....	60
FIGURA 4.9 – A) MODELO DE REFERÊNCIA. B) KRIGAGEM ORDINÁRIA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_3</i> . C) KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA DO CENÁRIO_5 REFERENTE AO ERRO <i>STD.DEV_3</i> . ....	61
FIGURA 4.10 – DIAGRAMA DE DISPERSÃO: OK_CENÁRIOS <i>VERSUS</i> REFERÊNCIA (TRIÂNGULOS EM VERMELHO), E KVME_CENÁRIOS <i>VERSUS</i> REFERÊNCIA (CÍRCULOS EM AZUL). NÍVEL DE INCERTEZA <i>STD.DEV_1</i> . ....	62
FIGURA 4.11 – DIAGRAMA DE DISPERSÃO: OK_CENÁRIOS <i>VERSUS</i> REFERÊNCIA (TRIÂNGULOS EM VERMELHO), E KVME_CENÁRIOS <i>VERSUS</i> REFERÊNCIA (CÍRCULOS EM AZUL). NÍVEL DE INCERTEZA <i>STD.DEV_2</i> . ....	63
FIGURA 4.12 – DIAGRAMA DE DISPERSÃO: OK_CENÁRIOS <i>VERSUS</i> REFERÊNCIA (TRIÂNGULOS EM VERMELHO), E KVME_CENÁRIOS <i>VERSUS</i> REFERÊNCIA (CÍRCULOS EM AZUL). NÍVEL DE INCERTEZA <i>STD.DEV_3</i> . ....	63
FIGURA 4.13 – ERRO QUADRÁTICO MÉDIO DO CENÁRIOS ( <i>STD.DV_3</i> ) POR FAIXA DE TEOR. LINHA EM AZUL REPRESENTA OS DESVIOS DA KVME, E LINHA EM VERMELHO DESVIOS DA OK. ....	64
FIGURA 4.14 – ANÁLISE DE DERIVA DE UM CENÁRIO PERTENCENTE AO ERRO <i>STD.DEV_3</i> . A) ANÁLISE NA DIREÇÃO Y. B) ANÁLISE NA DIREÇÃO X.....	66
FIGURA 4.15 – CURVA DE PARAMETRIZAÇÃO TEOR X TONELAGEM. CURVA OBTIDA ATRAVÉS DA MÉDIA DOS CENÁRIOS PERTENCENTE AO ERRO <i>STD.DEV_3</i> . ....	67

## LISTA DE TABELAS

TABELA 3.1 – ESTATÍSTICA DESCRITIVA DOS CENÁRIOS SIMULADOS EM FUNÇÃO DO DESVIO PADRÃO ADICIONADO NA GERAÇÃO DOS ERROS. ....	31
TABELA 3.2 – VARIOGRAMAS REFERENTES AOS CENÁRIOS COM ERRO <i>STD.DEV_1</i> . (*VALOR ADICIONADO COMO VARIÂNCIA DO ERRO DE MEDIDA).....	34
TABELA 3.3 – VARIOGRAMAS REFERENTES AOS CENÁRIOS COM ERRO <i>STD.DEV_2</i> . (*VALOR ADICIONADO COMO VARIÂNCIA DO ERRO DE MEDIDA).....	35
TABELA 3.4 – VARIOGRAMAS REFERENTES AOS CENÁRIOS COM ERRO <i>STD.DEV_3</i> . (*VALOR ADICIONADO COMO VARIÂNCIA DO ERRO DE MEDIDA).....	35
TABELA 3.5 – VALORES DE REFERÊNCIA E MÉDIA DOS CENÁRIOS ESTIMADOS POR OK E KVME. MÍNIMO DE TODOS OS CENÁRIOS ESTIMADOS DE 0 PPM. ....	39
TABELA 3.6 – ERRO QUADRÁTICO MÉDIO DOS CENÁRIOS POR FAIXA DE TEOR. CÉLULAS NA COR VERDE REPRESENTAM ERRO QUADRÁTICO MÉDIO DA KVME < OK; EM AMARELO DESVIO KVME = OK; EM VERMELHO DESVIO KVME > OK. ....	42
TABELA 3.7 – TEORES ORIGINAIS MÁXIMOS NOS CENÁRIOS SIMULADOS (H_S CENÁRIOS), NA REFERÊNCIA (V ORIGINAL), E APÓS A REALIZAÇÃO DO <i>CAPPING</i> EM P99 E P95. ....	48
TABELA 3.8 – VALORES DE REFERÊNCIA E MÉDIA DOS CENÁRIOS ESTIMADOS POR OK (P95 E P99) E KVME. MÍNIMO DE TODOS OS CENÁRIOS ESTIMADOS DE 0 PPM. ....	48
TABELA 3.9 – ERRO QUADRÁTICO MÉDIO DOS CENÁRIOS POR FAIXA DE TEOR. CÉLULAS NA COR VERDE REPRESENTAM DESVIO QUADRÁTICO MÉDIO DA KVME < OK; EM AMARELO DESVIO KVME = OK; EM VERMELHO DESVIO KVME > OK. ....	51
TABELA 4.1 – ESTATÍSTICA DESCRITIVA DOS CENÁRIOS SIMULADOS EM FUNÇÃO DO DESVIO PADRÃO ADICIONADO NA GERAÇÃO DOS ERROS. ....	56
TABELA 4.2 – VARIOGRAMAS REFERENTES AOS CENÁRIOS COM ERRO <i>STD.DEV_1</i> . (*VALOR ADICIONADO COMO VARIÂNCIA DO ERRO DE MEDIDA).....	58
TABELA 4.3 – VARIOGRAMAS REFERENTES AOS CENÁRIOS COM ERRO <i>STD.DEV_2</i> . (*VALOR ADICIONADO COMO VARIÂNCIA DO ERRO DE MEDIDA).....	58
TABELA 4.4 – VARIOGRAMAS REFERENTES AOS CENÁRIOS COM ERRO <i>STD.DEV_3</i> . (*VALOR ADICIONADO COMO VARIÂNCIA DO ERRO DE MEDIDA).....	59
TABELA 4.5 – VALORES DE REFERÊNCIA E MÉDIA DOS CENÁRIOS ESTIMADOS POR OK E KVME. MÍNIMO DE TODOS OS CENÁRIOS ESTIMADOS DE 0 PPM. ....	61
TABELA 4.6 – ERRO QUADRÁTICO MÉDIO DOS CENÁRIOS POR FAIXA DE TEOR. CÉLULAS NA COR VERDE REPRESENTAM ERRO QUADRÁTICO MÉDIO DA KVME < OK; EM AMARELO DESVIO KVME = OK; EM VERMELHO DESVIO KVME > OK. ....	64

## SUMÁRIO

<b>1. INTRODUÇÃO</b>	<b>1</b>
1.1 ADICIONANDO INCERTEZA AMOSTRAL DURANTE AS ESTIMATIVAS	1
1.2 META DO TRABALHO E OBJETIVOS ESPECÍFICOS	3
1.3 METODOLOGIA	3
1.4 ESTRUTURA DA DISSERTAÇÃO	4
<b>2. REVISÃO BIBLIOGRÁFICA</b>	<b>5</b>
2.1 AMOSTRAGEM E CONTROLE DE QUALIDADE	5
2.1.1 ERROS DE AMOSTRAGEM	5
2.1.2 GARANTIA E CONTROLE DE QUALIDADE	7
2.2 ESTIMATIVA DE DEPÓSITOS	9
2.2.1 KRIGAGEM ORDINÁRIA	10
2.2.2 SEMI VARIOGRAMA	16
2.2.3 EFEITO PEPITA E A KRIGAGEM	17
2.2.4 KRIGAGEM COM VARIÂNCIA DO ERRO DE MEDIDA	18
2.3 <i>OUTLIERS</i> E <i>CAPPING</i>	24
<b>3. ESTUDO DE CASO – 1</b>	<b>26</b>
3.1 CENÁRIOS HARD – SOFT	28
3.2 ADICIONANDO IMPRECISÃO AOS DADOS	29
3.3 CONTINUIDADE ESPACIAL E VARIÂNCIA DO ERRO DE MEDIDA	32
3.4 ESTIMATIVAS	35
3.5 DISCUSSÕES	44
3.6 <i>CAPPING</i> E <i>KVME</i>	46
<b>4. ESTUDO DE CASO – 2</b>	<b>53</b>
4.1 CENÁRIOS HARD – SOFT	53
4.2 ADICIONANDO IMPRECISÃO AOS DADOS	54
4.3 CONTINUIDADE ESPACIAL E VARIÂNCIA DO ERRO DE MEDIDA	56
4.4 ESTIMATIVAS	59
4.5 DISCUSSÕES	66
<b>5. CONCLUSÕES</b>	<b>68</b>
5.1 TRABALHOS FUTUROS	69
<b>REFERÊNCIAS</b>	<b>70</b>

## 1. INTRODUÇÃO

A estimativa de recursos minerais é frequentemente realizada utilizando informações de múltiplas fontes. Essas informações podem ter sido adquiridas em diferentes campanhas de sondagens, e através de distintas técnicas de perfuração, como sondagem diamantada (*diamond drill hole* - DDH), e circulação reversa (RC). Além disso, dados de produção, como amostras de canaleta e de furo de desmonte (*blast hole*), também podem fazer parte do banco de dados.

Cada fonte de dado está associada a um nível de precisão e acurácia, dependendo de como estes foram adquiridos e dos procedimentos de preparação e análise das amostras. Se os dados forem aceitos sem uma verificação rigorosa de sua qualidade, serão impossibilitados de delinear e detalhar corretamente um depósito mineral (PITARD, 1993).

Programas de controle de qualidade (QA/QC - *Quality Assurance and Quality Control*), monitoram a incerteza dos dados durante sua coleta e análise. Quando aprovados, esses dados são utilizados na estimativa e se assume que são isentos de incerteza, apesar de ser de amplo conhecimento que o erro é inerente ao processo de amostragem (GY, 1982). Sabe-se ainda, que o erro pode variar em função das características do material (GY, 1982) e em relação ao teor da amostra (THOMPSON; HOWARTH, 1973).

Como o erro amostral é indissociável ao processo amostral, uma das decisões a ser tomada durante o processo de estimativa, é se todos os dados disponíveis devem ser incorporados para estimar o depósito, ou se alguns dados devem ser excluídos devido a sua incerteza. Segundo Journel & Huijbregts (1978), a inclusão de dados adicionais reduz o erro da estimativa devido ao efeito da informação, porém, ao se excluir uma ou mais fontes de dados com alta imprecisão, supõe-se que o erro da estimativa será reduzido.

Para se tomar a decisão ótima sobre a utilização ou não de dados incertos, deve-se considerar a análise de técnicas geoestatísticas que incorporem tais situações de incerteza. Diferentes métodos geoestatísticos estão disponíveis para serem empregados neste contexto, apresentando distintas abordagens teóricas.

### 1.1 Adicionando incerteza amostral durante as estimativas

O método de estimativa mais conhecido que trabalha com fontes distintas de dados é a cokrigagem (GOOVAERTS, 1997). A técnica estima uma variável com base nos seus valores medidos, em conjuntos com os dados de uma ou mais variáveis secundárias. Journel &

Huijbregts (1978) e Goovaerts (1997) sugerem que a cokrigagem traz vantagens apenas quando a correlação entre os atributos é alta, e a variável de interesse é subamostrada em relação às variáveis secundárias. O objetivo de utilizar esses dados secundários é a minimização da variância do erro da estimativa da variável de interesse (WACKERNAGEL, 2003). Isso resulta em melhores resultados se comparados a estimativa apenas com os dados primários, pois ela filtra os erros e minimiza os efeitos do viés condicional (MINNITT; DEUTSCH, 2014).

Para incorporar as informações dos dados secundários na cokrigagem, é necessário a modelagem dos covariograma diretos e cruzados entre as variáveis primária e secundárias. Em virtude disso, o método pode ser extremamente laborioso quando se utiliza mais de dois atributos. Isso se deve a necessidade de encontrar um modelo que garanta que o sistema seja matematicamente válido, sendo o modelo linear de correogionalização (MLC) o mais conhecido e empregado.

Uma outra alternativa para utilizar dados com incerteza associada, é a krigagem dos indicadores com *soft data*. Na técnica, uma variável pode ter sua probabilidade *a priori* também baseada em dados secundários, chamados de *soft data* (ALABERT, 1987; SULLIVAN, 1984). Nessa metodologia, dados categóricos ou dados contínuos podem substituir as informações indefinidas nos intervalos restritos, permitindo que diferentes tipos de informações sejam processados em conjunto. O método traz consigo as limitações da krigagem dos indicadores, como a falta de formalismo teórico, e problemas de relação de ordem na presença de múltiplos indicadores.

Na existência de erros associados aos dados, não necessariamente a melhor estimativa naquele ponto é o valor obtido na amostragem. Dessa forma, sugere-se aplicar métodos de estimativa que não interpoem os teores, e sim se ajustem aos dados e aos erros introduzidos. Metodologias como Krigagem Filtrada (CHRISTENSEN, 2011; CRESSIE, 1991), Krigagem *New-Value* (KRIVORUCHKO; GRIBOV; HOEF, 2006) e Krigagem com Variância do Erro de Medida (DELHOMME, 1976) são alternativas aos métodos de interpolação.

Devido ao grande avanço da utilização de rotinas de QA/QC na indústria mineral, informações sobre a incerteza associadas aos processos de coleta, preparação e análise amostral se tornaram amplamente disponíveis. Isso viabilizou o emprego efetivo de métodos capazes de considerar o erro individual associado a cada medição. Nesse sentido, diversos *softwares* comerciais vêm oferecendo em suas versões mais recentes ferramentas estatísticas

que podem estar acompanhadas por um atributo de incerteza, como por exemplo, a implementação da Krigagem com Variância do Erro de Medida (KVME - *Kriging With Measurement Error Variance*) no *software* ISATIS (GEOVARIANCES, 2018).

Em vista ao exposto acima, a Krigagem com Variância do Erro de Medida é uma alternativa para se trabalhar com bancos de dados compostos por medidas de diferentes qualidades. A metodologia leva em consideração as incertezas das amostras, penalizando com menor peso aquelas com maior incerteza associada. Entretanto, compreender como essa nova variável influencia no sistema de krigagem, e encontrar um valor de erro fundamentado na amostragem e que seja compatível com a metodologia, não é uma atividade trivial, além de ser pouco explorado, o que justifica um estudo mais aprofundado da técnica.

### 1.2 Meta do trabalho e objetivos específicos

A meta da dissertação é investigar a metodologia "Krigagem com Variância do Erro de Medida" a partir da utilização de bancos de dados com amostras apresentando diferentes erros amostrais.

Para atingir a meta, foram elaborados os seguintes objetivos específicos:

i) Com base em um *dataset* original sem erro, elaborar uma gama de cenários que adicionem imprecisão aos teores verdadeiros. Esses cenários devem abranger diferentes configurações espaciais de erro.

ii) Desenvolver formas de se obter a variância do erro de medida para os diferentes cenários elaborados em (i).

iii) Investigar possíveis ganhos de precisão nas estimativas devido a incorporação do erro amostral e estimativa pela krigagem com variância do erro de medida.

iv) Avaliar a aplicabilidade do método como alternativa a krigagem ordinária.

### 1.3 Metodologia

A metodologia é composta das seguintes etapas:

i) Baseado em um banco de dados original, são gerados banco de dados secundários através da perturbação dos teores. Nesse trabalho, existem dois estudos de caso, e cada um possui uma forma de selecionar quais dados serão perturbados, e como será o comportamento do ruído adicionado.

ii) Dentro de cada estudo de caso, com a sua respectiva configuração amostral e distribuição do erro, serão simulados cenários com diferentes magnitudes de erro. Isso se dá com o aumento gradual do desvio-padrão da distribuição normal utilizada para a simulação dos novos teores dos cenários.

iii) Para utilizar a krigagem com variância do erro de medida, é necessário que cada amostra possua associada um valor de variância do erro de medida. Para cada estudo de caso, foi utilizada uma sistemática para a obtenção desses valores, explicada detalhadamente em cada capítulo correspondente.

iv) Os cenários são estimados por krigagem ordinária, técnica que não leva em consideração o erro nas amostras, e krigagem com variância do erro de medida, que utilizará a variância do erro de medida obtido no item (iii).

iv) Os resultados das estimativas são analisados e comparados entre si e entre o modelo de referência (variável original reblocada no mesmo suporte das estimativas). Isso é realizado através da análise dos valores máximos e mínimos estimados, desvio padrão, diagramas de dispersão, *swath plots*, coeficiente de correlação e erro quadrático médio.

#### 1.4 Estrutura da dissertação

O Capítulo 1 contextualiza o problema e apresenta os objetivos da dissertação. Compreende também uma breve descrição da metodologia adotada no trabalho.

O Capítulo 2 começa com uma revisão dos erros de amostragem, mostrando suas origens, subdivisões e como estes se inter-relacionam. É apresentado como a qualidade das amostras pode ser quantificada e monitorada a partir de programas de QA/QC. O Capítulo 2 ainda revisa a teoria da krigagem ordinária, da variografia e do efeito pepita. Expõe como é possível incorporar amostras com diferentes níveis de precisão na estimativa a partir de modificações no sistema de krigagem.

No início do Capítulo 3, é apresentado um banco de dados sintético, o qual é a base para a geração dos dois estudos de caso, apresentados ao longo do Capítulo 3 e Capítulo 4. Os estudos de caso estão associados a distintos comportamentos de erro. Os resultados das estimativas, por krigagem ordinária e por krigagem com variância do erro de medida, são discutidos e comparados entre si e contra a referência.

O Capítulo 5 mostra as conclusões do trabalho, os ganhos e limitações da metodologia. Possíveis caminhos de pesquisa são sugeridos para trabalhos futuros.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 Amostragem e Controle de Qualidade

A amostragem está entre as atividades mais fundamentais na mineração, pois as informações contidas nas análises dessas amostras serão utilizadas em tomadas de importantes decisões. Essas decisões podem envolver grandes quantidades de capital, como por exemplo abrir ou encerrar as atividades de uma mina, ou a adição de custos marginais, selecionando se determinado bloco minerado deve ser enviado para a usina de processamento ou para a pilha de estéril. Segundo Minnitt (2007), dos três processos que compõem o desenvolvimento de um projeto, que são os estudos de exploração mineral, a análise técnica e a análise financeira, o primeiro deve receber proporcionalmente mais tempo, dedicação e investimento que as etapas seguintes.

Para aumentar a confiança sobre os resultados das análises das amostras é fundamental a implementação de um programa de garantia e controle de qualidade. A maioria dos relatórios de recursos minerais, nos padrões internacionais, exigem tal programa, no qual as estimativas de recursos devem ser acompanhadas de uma descrição e declaração da qualidade dos dados. Rossi & Deutsch (2014) citam que esse programa é um item básico em qualquer revisão por terceiros do modelo de recursos obtido, podendo ter um impacto significativo na percepção geral da qualidade deste.

#### 2.1.1 Erros de Amostragem

A amostragem é o processo de extração de uma porção determinada de um material, sendo este representativo das características do material a ser amostrado. Neste processo podem existir diferentes tipos de erros: aqueles que estão presentes devido às propriedades intrínsecas do material, e aqueles que surgem durante o procedimento de amostragem e de sua preparação inadequada. O trabalho de Pierre Gy (GY, 1982) na compreensão dos diferentes erros de amostragem é o primeiro passo para estabelecer um protocolo de amostragem ideal. Abzalov (2016) agrupa os erros em três tipos, dependendo dos fatores que os causam, conforme equação (2.1):

$$\textit{Erro Total} = \textit{Erros 1}^{\text{o}} \textit{Grupo} + \textit{Erros 2}^{\text{o}} \textit{Grupo} + \textit{Erros 3}^{\text{o}} \textit{Grupo} \quad (2.1)$$

No primeiro grupo estão combinados os erros relacionados aos procedimentos de extração e preparação das amostras, conhecido como protocolo de amostragem. O principal

erro nesse grupo é conhecido como Erro Fundamental (FE), o qual é uma função da heterogeneidade da constituição do material a ser amostrado (GY, 1982). A eliminação do FE não é possível porque os minérios não têm estrutura ou composição uniformes, entretanto, este erro pode ser minimizado através da otimização do protocolo de amostragem. No primeiro grupo também será adicionado o erro de agrupamento e segregação, que é gerado pela distribuição heterogênea do material amostrado, e reflete as diferenças no conteúdo de um metal de interesse entre grupos de fragmentos coletados (PITARD, 1993).

Os erros que formam o segundo grupo estão relacionados com a forma que o protocolo de amostragem foi desenvolvido e implementado. Esse grupo inclui erros de delimitação, extração, preparação e pesagem. Erros humanos, como identificação das amostras, também podem ser incluídos neste grupo. Esses erros podem ser minimizados através da atualização das práticas de extração, preparação e controle de amostras, que geralmente precisam de uma melhoria dos procedimentos de controle de qualidade, e geralmente exigem atualização dos equipamentos (ABZALOV, 2016).

O último grupo abrange os erros analíticos e instrumentais que ocorrem durante as atividades analíticas. O grupo inclui ensaios, análise de umidade, pesagem das alíquotas, análise de densidade e desempenho sub-ótimo dos instrumentos analíticos. Um caso especial de erro instrumental é dado pela escolha incorreta de técnicas analíticas, a qual não é eficiente para dado tipo de mineralização ou à faixa de teores de interesse.

Talvez o aspecto mais importante que requer atenção na detecção dos erros seja a variabilidade, geralmente medida em termos de precisão, visto que tende a afetar os procedimentos de amostragem. Minnitt (2007) cita que há boas evidências de que as variâncias são aditivas em cada etapa do processo de amostragem e sua posterior análise. No diagrama da Figura 2.1, é apresentado como os erros se relacionam entre si, os quais são variações que se acumulam em diferentes níveis. A heterogeneidade, o qual é uma combinação do efeito pepita, do erro fundamental, e do erro de agrupamento e segregação, é considerada como a variância principal. Os erros de delimitação, extração, preparação e pesagem, assim como o erro analítico são também contribuintes.

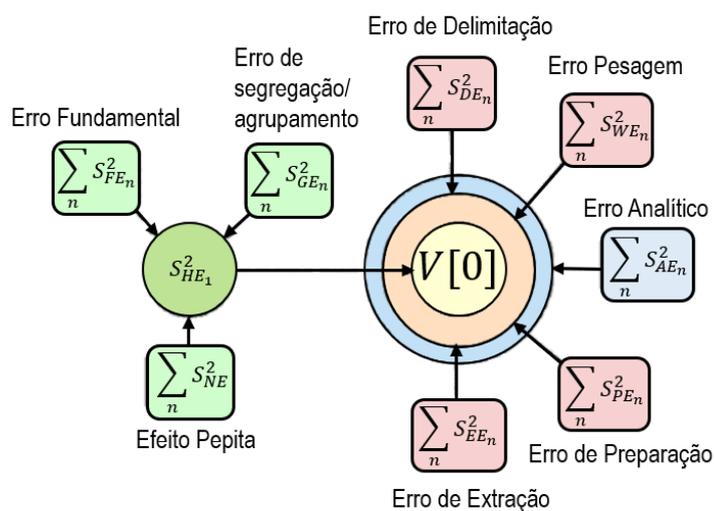


Figura 2.1 – Relação entre os erros de amostragem. Fonte: modificado de Minnitt (2007).

### 2.1.2 Garantia e Controle de Qualidade

Os procedimentos que envolvem a quantificação e monitoramento da qualidade das amostras são referidos como “QA/QC” (*Quality Assurance and Quality Control*) (CSA, 2005). Segundo Rossi & Deutsch (2014), o principal objetivo do programa de QA/QC é minimizar os erros introduzidos devido aos procedimentos de amostragem, preparação e análise das amostras. Esses procedimentos asseguram a integridade da informação, permitindo sua posterior utilização na estimativa de recursos, no controle de teores e para a reconciliação.

Enquanto a Garantia de Qualidade compreende ações sistemáticas e preestabelecidas orientadas a elevar o nível de confiança do programa de exploração, o Controle de Qualidade abarca os procedimentos rotineiros para monitorar a qualidade dos dados. Dessa forma, ao passo que a Garantia de Qualidade trabalha com a prevenção, o Controle de Qualidade se relaciona com a detecção dos problemas. É um processo contínuo, que compreende o diagnóstico dos erros e a identificação de suas fontes, fornecendo as informações necessárias para corrigir possíveis falhas.

A qualidade dos dados geoquímicos utilizados na exploração e desenvolvimento do depósito mineral é determinada e reportada em termos de precisão e acurácia (JORC, 2012). A acurácia é uma medida do grau de concordância do valor da amostra analisada com o seu valor verdadeiro. Já a precisão, ou reprodutibilidade, indica o quão próximo os valores das amostras replicadas estão uns dos outros. A Figura 2.2 ilustra os conceitos de precisão e acurácia utilizando uma analogia com um alvo e seus lançamentos.

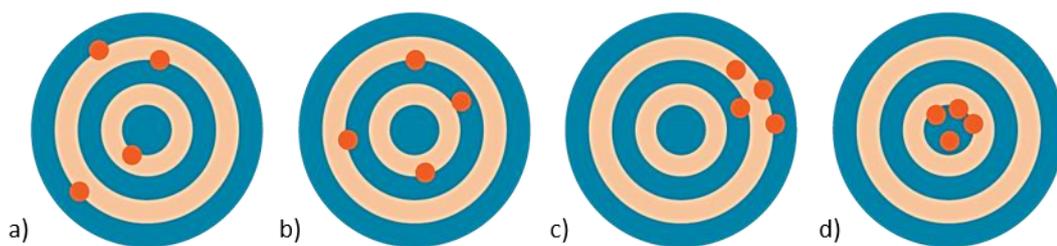


Figura 2.2 – Conceitos de precisão e acurácia: a) impreciso e inacurado; b) impreciso e acurado; c) preciso e inacurado; d) preciso e acurado.

### i) Precisão

A precisão analítica é monitorada usando análises pareadas. A primeira amostra é referida como original e a segunda é denominada duplicada. A duplicada é coletada no mesmo local e seguindo os mesmos processos utilizados para coletar a amostra original. Ambas são preparadas e analisadas no mesmo laboratório e utilizando os mesmos procedimentos. Variações nos resultados entre tais pares de amostras permitem estimar a precisão associada com a preparação e análises das amostras (ABZALOV, 2016).

Uma duplicata pode ser, por exemplo, uma segunda amostra de um furo de desmonte (*blast hole*) recolhida a partir do mesmo cone da primeira amostra, ou pode ser outra fração de um testemunho (*drill core*). Posteriormente a obtenção das amostras, haverá vários estágios de cominuição e subamostragem, e nestas etapas também deverão ser tomadas duplicatas. A vantagem nas duplicatas de campo é que a variância observada inclui todas as fontes de erro, desde aquela derivada da amostragem *in situ* e de sua primeira etapa de redução de tamanho.

O propósito das duplicatas de material grosseiro é quantificar as variâncias advindas de erros introduzidos em diferentes estágios de preparação das amostras, até o primeiro estágio de cominuição (ROSSI; DEUTSCH, 2014). Essas duplicatas fornecem uma medida da precisão dos procedimentos analíticos utilizados, a qual é uma soma das variâncias causadas pela análise e preparação da amostra. Duplicatas de polpa são obtidas no estágio final da preparação da amostra, permitindo uma medida de precisão dos procedimentos analíticos. Duplicatas de polpa enviadas para o mesmo laboratório primário fornecem uma estimativa da variância analítica daquele laboratório. Quando enviados para um segundo laboratório, as duplicatas de polpa quantificam a variância analítica entre os dois laboratórios.

## ii) Acurácia

Uma indicação da acurácia pode ser obtida através da inclusão de amostras de teores conhecidos junto ao lote a ser analisado. Estas amostras são chamadas de *standards* ou padrões, e são estatisticamente certificadas para alguns tipos de minério, sendo obtidas de laboratórios de instituições certificadas ao redor do mundo. Alternativamente, a equipe de QA/QC pode desenvolver seus próprios *standards*, comumente chamados de *matrix-matched standards*. Ao se optar por essa alternativa, o material utilizado para sua obtenção será tipicamente do mesmo depósito que as amostras a serem monitoradas, o que garante que as diferenças nas matrizes de amostra sejam mínimas. Ambos tipos de padrões envolvem em seu processo de certificação a análise do material por vários laboratórios, chamadas de análises “*Round Robin*”, as quais atestam o seu teor e correspondente limite de tolerância. Outro componente dos programas de controle e garantia de qualidade são os *blanks* ou brancos. Brancos são amostras cujo teor da variável de interesse é nulo ou muito baixo, geralmente abaixo do limite de detecção. Estes são inseridos também nas análises do lote com objetivo de verificar possíveis contaminações no laboratório e verificar o manuseio correto das amostras.

Brancos e padrões são chamados de materiais de referência e são utilizados para detectar um possível viés nos dados e para quantificar a acurácia do processo. A acurácia das análises laboratoriais é avaliada por testes estatísticos que são baseados em uma comparação da média aritmética das análises replicadas de um *standard* em relação à sua média certificada. Os testes estatísticos da repetitividade dos *standards* também podem ser baseados em análises “*Round Robin*”, o qual utiliza o mesmo método em diferentes equipamentos, ou uma variedade de métodos e equipamentos diferentes. Segundo Abzalov (2011), é considerado boa prática apoiar o teste de acurácia através da estimativa do desvio padrão das análises replicadas das amostras *standards*, uma vez que a baixa precisão destas análises podem impedir a detecção de viés analítico.

## 2.2 Estimativa de Depósitos

Em geral, menos de um bilionésimo do volume de um depósito é amostrado (ROSSI; DEUTSCH, 2014). Dessa forma, certos atributos geológicos devem ser estimados com o principal objetivo de prever o teor e a tonelagem do material que poderá ser extraído. A

estimativa correta dos recursos será de extrema importância para avaliar projetos e custos de lavra.

Nas ciências da terra em geral e, em particular, nos depósitos minerais, a continuidade espacial de propriedades geológicas representam um fenômeno natural. A geoestatística, tendo início na década de 50 (KRIGE, 1951; MATHERON, 1963), oferece formas de descrever e modelar a continuidade espacial de variáveis regionalizadas, permitindo integrar estes parâmetros com interpolações espaciais.

As aplicações das técnicas geoestatísticas na indústria de mineração são muito amplas, compreendendo a estimativa dos recursos minerais, a avaliação da incerteza do modelo, a quantificação dos riscos associados aos teores do minério ou a determinação de um modelo ideal de sondagem para os *grids* de amostragem. Os métodos mais comumente utilizados são os estimadores geoestatísticos lineares univariados, a Krigagem Ordinária e Krigagem Simples.

### 2.2.1 Krigagem Ordinária

A Krigagem Ordinária é um método geralmente associado ao acrônimo BLUE (*Best Linear Unbiased Estimator* – Melhor Estimador Linear Não Enviesado). É “linear” pois é uma combinação linear ponderada dos dados disponíveis, “não possui viés” pois a média dos erros é igual a zero, e é “melhor” porque procura minimizar a variância dos erros.

No sentido prático, como nunca será conhecida a média dos erros, não será possível garantir que ela seja exatamente zero. E também, não é possível conhecer a variância dos erros para poder minimizá-la. Logo, o que pode ser feito é criar um modelo com os dados em estudo e trabalhar com o erro médio e a variância dos erros do modelo. Na krigagem ordinária é utilizado um modelo probabilístico, e escolhidos pesos para as amostras mais próximas que assegurem que a média dos erros do modelo seja zero e que a variância do erro dos valores estimados no modelo seja minimizada (ISAAKS; SRIVASTAVA, 1989).

#### i) Condição de não viés

Em todos os pontos onde não há amostras, serão estimados os valores desconhecidos usando uma combinação linear das amostras disponíveis no entorno usando:

$$\hat{v} = \sum_{i=1}^n w_i \cdot v_i \quad (2.2)$$

Na equação (2.2),  $\hat{v}$  é o valor estimado,  $v_1, \dots, v_n$  são os  $n$  dados disponíveis na vizinhança local, e  $w_i$  será o peso atribuído para cada valor  $v_i$ . O conjunto de pesos pode ser alterado conforme se estima valores desconhecidos em locais diferentes.

O erro,  $r$ , de qualquer valor estimado é a diferença entre o valor estimado e o valor verdadeiro no mesmo local dado por:

$$r_i = \hat{v}_i - v_i \quad (2.3)$$

então o erro médio de um conjunto de  $k$  estimativas será:

$$m_r = \frac{1}{k} \sum_{i=1}^k r_i = \frac{1}{k} \sum_{i=1}^k \hat{v}_i - v_i \quad (2.4)$$

Como essa equação envolve valores verdadeiros  $v_i$  que são desconhecidos, não se pode fazer uso dela. Assim, a solução probabilística para este problema consiste em conceituar o fenômeno como realizações de um processo aleatório e resolver o problema para esse modelo conceitual. Para qualquer ponto em que se pretende estimar um valor desconhecido, o modelo será uma função aleatória estacionária, que consiste em uma série de variáveis aleatórias, uma para cada um dos locais amostrados,  $V(x_1) \dots, V(x_n)$ , e uma para o valor desconhecido no ponto que se deseja estimar,  $V(x_0)$ . Cada uma dessas variáveis aleatórias possui a mesma lei de probabilidade, e em todos os locais o valor esperado da variável aleatória é  $E\{V\}$ . Qualquer par de variáveis aleatórias tem uma distribuição conjunta que depende apenas da separação entre os dois pontos e não de suas localizações. A covariância entre pares de variáveis aleatórias separadas por uma determinada distância,  $h$ , é denotado por  $\tilde{C}_v(h)$  (ISAAKS; SRIVASTAVA, 1989).

Cada valor nesse modelo é visto como o resultado de uma variável aleatória, tanto as amostras como o valor desconhecido. A estimativa também é uma variável aleatória, pois é uma combinação linear ponderada das variáveis nos locais com amostras disponíveis. Da mesma forma, o erro de estimativa, definido como a diferença entre a estimativa e a variável aleatória que modela o valor real, é também uma variável aleatória. Pode-se escrever então o erro da estimativa,  $R(x_0)$ , conforme a equação (2.5):

$$R(x_0) = \sum_{i=1}^n w_i \cdot V(x_i) - V(x_0) \quad (2.5)$$

Para garantir que o erro em qualquer lugar específico tenha um valor esperado de zero, aplica-se a fórmula para o valor esperado de uma combinação linear à equação (2.5):

$$\begin{aligned} E \{R(x_0)\} &= E \left\{ \sum_{i=1}^n w_i \cdot V(x_i) - V(x_0) \right\} \\ &= \sum_{i=1}^n w_i E \{V(x_i)\} - E\{V(x_0)\} \end{aligned} \quad (2.6)$$

Como se trata de uma função aleatória estacionaria, o valor esperado pode ser expresso como:

$$E \{R(x_0)\} = \sum_{i=1}^n w_i E \{V\} - E\{V\} \quad (2.7)$$

O valor esperado do erro em qualquer localização particular  $E \{R(x_0)\}$  é referido como viés. Atribuindo zero para este valor esperado será garantido a condição de não viés, resultando na seguinte conclusão:

$$\begin{aligned} E \{R(x_0)\} = 0 &= \sum_{i=1}^n w_i E \{V\} - E\{V\} \\ E \{V\} \sum_{i=1}^n w_i &= E\{V\} \\ \sum_{i=1}^n w_i &= 1 \end{aligned} \quad (2.8)$$

Dessa forma, para estimativas sem viés, é preciso restringir a soma dos pesos das combinações lineares ponderadas para um. Esse resultado é geralmente chamado de condição de não tendenciosidade.

## ii) Variância mínima do erro

A variância do erro,  $\sigma_R^2$ , de um conjunto de  $k$  estimativas pode ser escrita como:

$$\sigma_R^2 = \frac{1}{k} \sum_{i=1}^k (r_i - m_R)^2 \quad (2.9)$$

$$= \frac{1}{k} \sum_{i=1}^k \left[ \hat{v}_i - v_i - \frac{1}{k} \sum_{i=1}^k (\hat{v}_i - v_i) \right]^2 \quad (2.10)$$

$v_1, \dots, v_n$  na fórmula são os valores reais (amostrais), e  $\hat{v}_1, \dots, \hat{v}_n$  são as estimativas correspondentes. Assumindo que se tem a média do erro igual a zero, a equação anterior pode ser simplificada:

$$\sigma_R^2 = \frac{1}{k} \sum_{i=1}^k (r_i - 0)^2 \quad (2.11)$$

$$= \frac{1}{k} \sum_{i=1}^k (\hat{v}_i - v_i)^2 \quad (2.12)$$

Como essa equação também exige o conhecimento dos valores verdadeiros, da mesma forma que anteriormente, será utilizada uma solução probabilística para minimizar a variância dos erros. Serão utilizadas  $n + 1$  variáveis aleatórias, no qual  $n$  variáveis aleatórias modelam o comportamento do fenômeno nas amostras mais próximas e a outra variável aleatória modelará o comportamento no local onde será realizada a estimativa. Como não é possível minimizar a variância dos erros, pois são desconhecidos os valores verdadeiros, então será minimizada a variância do modelo do erro  $R(x_0)$ . A minimização será realizada encontrando a expressão para modelar a variância do erro,  $\tilde{\sigma}_R^2$ , e atribuindo zero para as várias derivadas parciais dessa expressão (ISAACS; SRIVASTAVA, 1989).

A fórmula da variância dos erros como uma combinação linear ponderada é expressa na equação (2.13). Uma vez que se tenha escolhido os parâmetros do modelo, especificamente a variância,  $\tilde{\sigma}^2$ , e todas as covariâncias,  $\tilde{C}_{ij}$ , a equação abaixo fornece uma expressão para a variância do erro como uma função de  $n$  variáveis, ou seja, os pesos  $w_1, \dots, w_n$ .

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \tilde{C}_{ij} - 2 \sum_{i=1}^n w_i \tilde{C}_{i0} \quad (2.13)$$

A minimização de uma função de  $n$  variáveis é geralmente realizada atribuindo zero para as  $n$  primeiras derivadas parciais. Porém, esse processo não está completamente correto para a minimização da variância do erro, já que se tem uma limitação na solução desse problema. A condição de não viés requer que a possível solução seja restrita para que a soma

dos pesos seja igual a um. Ao ser realizada a minimização da equação anterior, ocorrerão algumas dificuldades, visto que atribuindo zero para as  $n$  primeiras derivadas parciais da variância, serão geradas  $n$  equações com  $n$  incógnitas. A condição de não viés adicionará uma nova equação sem adicionar uma incógnita. Isso leva a um sistema com  $n + 1$  equações e somente  $n$  incógnitas. Para resolver esse problema, será incorporada uma nova incógnita na equação da variância, e essa nova variável é chamada de  $\mu$ , parâmetro de Lagrange (ISAAKS; SRIVASTAVA, 1989). A integração do parâmetro de Lagrange na equação (2.13) gera a expressão abaixo:

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 + \sum_{i=1}^n \sum_{j=1}^n w_i w_j \tilde{C}_{ij} - 2 \sum_{i=1}^n w_i \tilde{C}_{i0} + 2\mu \left( \sum_{i=1}^n w_i - 1 \right) \quad (2.14)$$

A adição de uma nova variável a uma equação precisa de alguns cuidados para evitar que a igualdade seja prejudicada. A maneira que foi utilizada para adicionar a variável na equação anterior é segura já que o termo adicionado ao final é igual a zero conforme a condição de não viés, apresentada na equação (2.15). A solução desse problema, além de adicionar a condição de imparcialidade, fornece o parâmetro de Lagrange que será utilizado no cálculo da variância do erro.

$$\begin{aligned} \sum_{i=1}^n w_i &= 1 \\ \sum_{i=1}^n w_i - 1 &= 0 \\ 2\mu \left( \sum_{i=1}^n w_i - 1 \right) &= 0 \end{aligned} \quad (2.15)$$

A variação do erro para o modelo,  $\tilde{\sigma}_R^2$ , é agora função de  $n + 1$  variáveis, dos  $n$  pesos e do parâmetro Lagrange. A variância do erro será minimizada calculando as  $n + 1$  primeiras derivadas parciais da equação (2.13), e atribuindo zero para cada uma delas, equação (2.16).

$$\frac{\partial(\tilde{\sigma}_R^2)}{\partial w_i} = 2 \sum_{j=1}^n w_j \tilde{C}_{ij} - 2\tilde{C}_{i0} + 2\mu = 0 \quad (2.16)$$

O conjunto de pesos que minimizam o erro da variância sob a restrição de que eles somem um, portanto, satisfazem as seguintes  $n + 1$  equações:

$$\sum_{j=1}^n w_j \tilde{C}_{ij} + \mu = \tilde{C}_{i0} \quad \forall i = 1, \dots, n \quad (2.17)$$

$$\sum_{i=1}^n w_j = 1 \quad (2.18)$$

Esse sistema de equações geralmente é referido como sistema de krigagem ordinária, e pode ser escrito em notação matricial, conforme apresentado na equação abaixo (ISAAKS; SRIVASTAVA, 1989).

$$C \cdot w = D \quad (2.19)$$

$$\begin{bmatrix} \tilde{C}_{11} & \dots & \tilde{C}_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \tilde{C}_{n1} & \dots & \tilde{C}_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \tilde{C}_{10} \\ \vdots \\ \tilde{C}_{n0} \\ 1 \end{bmatrix}$$

Para calcular os pesos, é necessário multiplicar os dois lados da equação por  $C^{-1}$ , o inverso da matriz de covariância do lado esquerdo, resultando nas equações abaixo.

$$C \cdot w = D$$

$$C^{-1} \cdot C \cdot w = C^{-1} \cdot D$$

$$I \cdot w = C^{-1} \cdot D$$

$$w = C^{-1} \cdot D \quad (2.20)$$

É possível substituir os pesos encontrados anteriormente para obter o valor da variância da krigagem ordinária, expressando como:

$$\tilde{\sigma}_R^2 = \tilde{\sigma}^2 - \sum_{i=1}^n w_i \tilde{C}_{i0} + \mu \quad (2.21)$$

Em síntese, a matriz  $D$  fornece um esquema de pesos semelhante ao método de estimativa denominado inverso da distância. A covariância entre cada amostra e o ponto a ser estimado geralmente diminui com a distância, até receber o valor zero. Assim, a matriz  $D$  contém uma distância, que não é geométrica e sim uma distância estatística. A matriz  $C$  além de transformar os valores da matriz  $D$  para que a soma seja igual a um, também utiliza as distâncias entre cada amostra para fornecer para o sistema de krigagem informações sobre o agrupamento dos dados. Dessa forma, se duas amostras estiverem próximas umas às outras,

será atribuído um valor alto, caso contrário, será atribuído um valor baixo. A multiplicação de  $D$  por  $C^{-1}$  ajusta o inverso das distâncias estatísticas dos pesos em  $D$  para levar em consideração as possíveis redundâncias entre as amostras. Assim como na matriz  $D$ , a matriz  $C$  registra a distância em termos de distância estatística. Dessa forma, a possível redundância entre amostras depende não simplesmente da distância geométrica, mas sim da continuidade espacial. O uso da continuidade espacial permite customizar o procedimento de estimativa com o auxílio de informações qualitativas. Assim, a escolha de um modelo de covariância (semi variograma ou correlograma) é um pré-requisito para a krigagem ordinária.

### 2.2.2 Semi Variograma

O semi variograma é a ferramenta geoestatística mais utilizada na mineração para definir a continuidade espacial de vários atributos geológicos, como teor da mineralização e espessura de zonas mineralizadas (JOURNAL; HUIJBREGTS, 1978; GOOVAERTS, 1997). O semi variograma experimental  $\gamma(h)$  de uma variável regionalizada  $Z(x)$  é uma função que descreve a continuidade espacial desta variável em função da direção e da distância entre pontos amostrais. Em outras palavras, representa uma medida da dissimilitude espacial entre os dados em função do vetor crescente de separação entre eles ( $h$ ), e é calculado como metade da diferença quadrática separados pelo vetor ( $h$ ).

$$\gamma(h) = \frac{1}{2N} \sum_{i=1}^n [Z(x_i + h) - Z(x_i)]^2 \quad (2.22)$$

Mudando o vetor ( $h$ ), seu tamanho e sua direção, será possível quantificar o padrão de continuidade espacial de uma determinada variável regionalizada. Como a diferença calculada utilizando a equação (2.22) é uma grandeza quadrática, o variograma é independente do sinal do vetor, portanto é simétrico em relação a ( $h$ ).

Os valores de um variograma calculados ao longo de uma direção escolhida  $\gamma(h)$  podem ser plotados em relação à distância ( $h$ ), conforme mostra a Figura 2.3. Esta é uma maneira tradicional de apresentar os valores do variograma quando toda a informação é sintetizada em um único ponto por classe de distância. Na prática, o variograma experimental é calculado sobre incrementos regulares ( $h$ ), denominados de *lags* do variograma.

A análise variográfica requer o ajuste de uma função matemática ao variograma experimental. Existem vários modelos admissíveis que podem ser utilizados para se aproximar ao modelo experimental. Os modelos mais comumente utilizados para representar a

continuidade espacial das características geológicas e dos teores do depósito são: esférico, exponencial, potência, Gaussiano, linear e efeito pepita puro. Cada um desses modelos é utilizado para descrever os diferentes comportamentos das variáveis estudadas. Na prática, dois ou mais modelos de variograma geralmente precisam ser combinados para se adaptar aos variogramas experimentais.

Os variogramas geralmente apresentam algumas características principais, conforme mostra a Figura 2.3. Na origem, em  $h = 0$ , o valor de  $\gamma(h)$  seria de zero, porém ele pode ser descontínuo logo após a origem, e essa descontinuidade é denominada de “efeito pepita”. O gráfico, geralmente, aumenta continuamente com a distância ( $h$ ), podendo estabilizar em um determinado nível, chamado de “patamar” ou “sill”. A distância entre a origem do variograma até o patamar é denominado de “alcance” ou “range”. O alcance transmite um significado quantitativo ao conceito geológico de “zona de influência” de uma amostra. Quando o variograma atinge o patamar, isso significa que naquela distância de separação não há mais correlação entre as amostras. O patamar e o alcance do variograma podem variar nas diferentes direções do depósito refletindo a anisotropia, situação muito comum quando uma variável regionalizada é mais contínua ao longo de uma direção específica do que em outra.

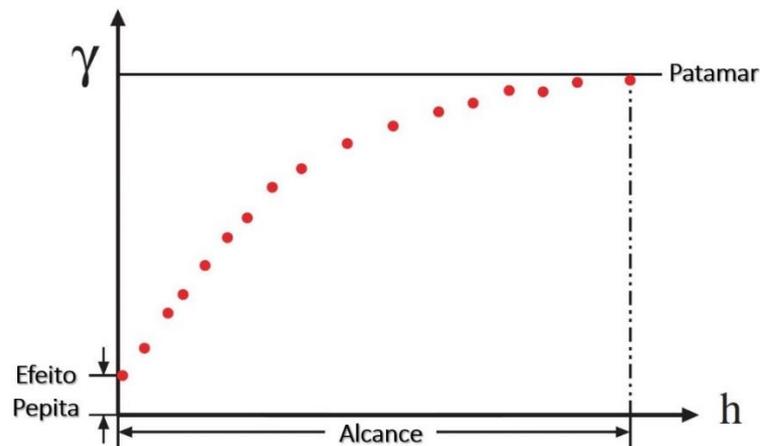


Figura 2.3 – Características de um variograma. Fonte: modificado de Rossi e Deutsch (2014).

### 2.2.3 Efeito Pepita e a Krigagem

Os variogramas que são descontínuos na origem são o tipo mais comum encontrado nos estudos de depósitos minerais (JOURNEL; HUIJBREGTS, 1978). Essa descontinuidade na origem, quando o variograma  $\gamma(h)$  não tende a zero com  $h$  aproximando-se de zero, indica um comportamento irregular (alta variabilidade) da variável em questão a uma curta

distância. Essa descontinuidade é chamada de efeito pepita, pois remete a heterogeneidade em pequena escala associada à presença de pepitas de ouro.

Utilizando os mesmos dados e técnica de krigagem, a estimativa obtida utilizando um modelo de variograma com efeito pepita é mais suavizada, isto é, possui menor definição em pequena escala do que a previsão obtida quando não se adiciona nenhum efeito pepita. Isso acontece, porque a adição de um efeito de pepita resulta em uma maior distribuição dos pesos de krigagem para dados mais distantes ao lugar estimado. Isto é, com o aumento do efeito pepita, o processo de estimativa torna-se mais parecido como uma simples média aritmética dos dados disponíveis. Assim, um modelo de continuidade espacial de efeito pepita puro implica uma completa falta de correlação espacial. Matematicamente, o efeito pepita subtrai um valor constante em todos os valores fora da diagonal principal da matriz de krigagem, equação (2.19), diminuindo a covariância entre as amostras e aumentando a variância do erro.

Alguns autores (CHILES; DELFINER, 1999; CRESSIE, 1991) argumentam que o efeito pepita compreende uma combinação de um “verdadeiro efeito pepita”, o qual é causado por uma continuidade menor que a menor distância entre as amostras, e um erro de medida ou posicionamento, advindos do erro de amostragem.

$$c_0 = c_{MS} + c_{ME} \quad (2.23)$$

Na equação (2.23),  $c_0$  é o valor do efeito pepita;  $c_{MS}$  é o “verdadeiro efeito pepita”, ou seja, a variância em microescala; e  $c_{ME}$  é a variância do erro de medida.

A decomposição do efeito pepita não é realmente significativa quando se trata de modelar o variograma, entretanto é importante quando se trata da krigagem. A krigagem na ausência de erros de medida é um interpolador exato, sempre retornando o valor da amostra como a estimativa em qualquer um dos locais amostrados. Contudo, assumindo que exista um erro de medida, a propriedade da krigagem de ser um interpolador exato invalida-se. Isto é, a krigagem pode honrar os dados se assumida como um interpolador exato,  $c_{ME} = 0$ ; ou pode suavizar os dados nos locais amostrados se  $c_{ME} > 0$ , a partir de uma modificação no sistema de krigagem.

#### 2.2.4 Krigagem com Variância do Erro de Medida

Assumir  $c_{ME} = 0$  é uma suposição irrealista quando se utiliza dados advindos do processo de amostragem (ver Erros de Amostragem). Ao ignorar esses erros, subestima-se

consideravelmente a incerteza e conseqüentemente o risco econômico associado a um projeto de mineração (CUBA; LEUANGTHONG; ORTIZ, 2012).

Em vista disso, vários métodos geoestatísticos estão disponíveis para extrair o máximo de informação de dados que sejam imprecisos, enviesados ou ambos. Uma abordagem para realizar estimativas integrando dados com fontes variadas de precisão é a Krigagem com Variância do Erro de Medida (KVME ou OKVME – *Ordinary Kriging with Variance of Measurement Error*). Essa metodologia foi primeiro descrita por Delhomme (1976), porém pouca bibliografia foi gerada sobre o assunto desde então. O método é abordado brevemente em alguns livros (CHILES; DELFINER, 1999; WACKERNAGEL, 2003), e citado em poucos artigos científicos (CHAUTRU, 2015; CORNAH; MACHAKA, 2015; DERAISME; STRYDOM, 2009).

Na Krigagem com Variância do Erro de Medida, a precisão de cada dado amostrado  $z(x_i)$  deve ser conhecida ou estimada de maneira confiável, para que então possa ser incorporada na estimativa. A incerteza sobre esse dado pode ser assumido como uma variância ( $\sigma_i^2$ ), a qual pode ser obtida a partir de um cálculo estatístico, ou estimada aproximadamente pela equação do intervalo de incerteza (DELHOMME, 1976).

Assumindo que os erros  $\varepsilon(x_i)$  são:

- |  |   |                        |
|--|---|------------------------|
| a. Não sistemáticos (ou seja, randômicos)          | $E[\varepsilon(x_i)] = 0$                     | $\forall i$            |
| b. Não correlacionáveis entre eles                 | $cov[\varepsilon(x_i), \varepsilon(x_j)] = 0$ | $\forall i \neq j$     |
| c. Não correlacionáveis com as variáveis estudadas | $cov[\varepsilon(x_i), Z(x)] = 0$             | $\forall i, \forall x$ |

Essa incerteza nos dados introduz uma pequena mudança no sistema de krigagem. Os dados não são mais considerados valores  $z(x_i)$ , mas uma soma de “valor verdadeiro + erro”, isto é,  $z(x_i) + \varepsilon(x_i)$ . O estimador gerado para a KVME, portanto, assume a forma (DELHOMME, 1976):

$$z_0^* = \sum_{i=1}^n w_i (z(x_i) + \varepsilon(x_i)) \quad (2.24)$$

Como os erros são randômicos,  $E[\varepsilon(x_i)] = 0$ , as condições de não-viés não serão modificadas. O sistema de krigagem pode ser escrito então:

$$\sum_{j=1}^n w_j \tilde{C}_{ij} + w_i \sigma_i^2 + \mu = \tilde{C}_{i0} \quad i = 1, \dots, n \quad (2.25)$$

$$\sum_{i=1}^n w_i = 1 \quad (2.26)$$

Esse sistema de equações pode ser escrito em notação matricial, conforme apresentado na equação (2.27). A diferença para o sistema de krigagem ordinária, equações (2.17), (2.18), (2.19), está nos termos da diagonal da matriz de covariância  $C$ , no qual se adiciona  $\sigma_i^2$ . Nenhuma mudança ocorre na matriz  $D$ , conforme mostrado abaixo.

$$\begin{bmatrix} C_{11} + \sigma_1^2 & \cdots & C_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C_{n1} & \cdots & C_{nn} + \sigma_n^2 & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} C_{10} \\ \vdots \\ C_{n0} \\ 1 \end{bmatrix} \quad (2.27)$$

Substituindo na fórmula (2.28) os pesos encontrados anteriormente, obtém-se o valor da variância de krigagem.

$$\sigma_R^2 = \sigma^2 - \sum_{i=1}^n w_i C_{io} + \mu \quad (2.28)$$

Dessa forma, através da adição na diagonal da matriz de covariância do valor correspondente ao erro, a KVME penaliza com menor peso as amostras com maior incerteza associada. Delhomme (1976) apresenta um exemplo para visualizar os efeitos da presença de uma incerteza, caracterizada por  $\sigma_i^2$ , em um dado  $z(x_i)$ . No exemplo, a variável estudada possui variograma  $\gamma(h) = |h|$ , e os dados estão localizados nos seis vértices de um hexágono regular, conforme mostra a Figura 2.4 (a). Cinco desses dados são considerados livres de erros ( $\sigma_i^2 = 0$  para  $i = 1, \dots, 5$ ), e o sexto dado  $z(x_6)$  possui erro representado por  $\sigma_6^2$ . Uma maneira de demonstrar a influência da incerteza na KVME é observar o peso  $w_6$  atribuído a  $z(x_6)$  na estimativa do valor no centro do hexágono, apresentada na Figura 2.4 (b). Quando  $\sigma_6^2 = 0$ , o peso atribuído ao sexto dado é o mesmo que os outros cinco, ou seja,  $w_6 = 1/6$  (0,16). Conforme  $\sigma_6^2$  cresce o peso  $w_6$  diminui, e o limite tende a zero quando  $\sigma_6^2$  tende ao infinito. Nessa lógica, dados totalmente incertos seriam equivalentes a falta de dados (DELHOMME, 1976).

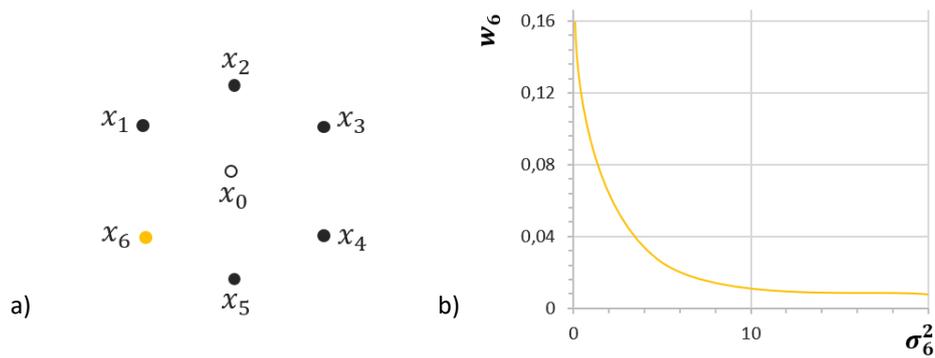


Figura 2.4 – a) Dados sem erro ( $i = 1, \dots, 5$ ), dado com erro associado ( $i = 6$ ), e valor a ser estimado  $z(x_0)$ . b) Influência da incerteza sobre os pesos. Fonte: modificado de Delhomme (1976).

Como citado na seção 2.2.3, a krigagem em sua versão original é considerada um método de interpolação exato; porém, a partir de uma modificação no sistema essa propriedade é rescindida. Na KVME, a função não passará exatamente sobre os dados de entrada, pois a técnica permite que os erros de medida sejam especificados e, assim, suavizados para obter um preditor estatisticamente mais adequado. Desta maneira, a KVME não irá honrar os dados de entrada que possuam erros, tornando-se assim um método de ajuste (*fitting*). A Figura 2.5 ajuda a elucidar a diferença do comportamento da KVME em relação a interpolação por krigagem sem erro amostral. Em (a) foram gerados dados sintéticos a partir de uma função (linha tracejada), e nesses dados foram inseridos erros produzidos através de uma distribuição gaussiana. Em (b) e (c) é possível ver os resultados da interpolação por krigagem, respectivamente sem e com efeito pepita. Mesmo com efeito pepita, a curva passa exatamente pelos dados, o que seria pouco provável de ocorrer quando os erros de medida são significativos. Por último, em (d), o estimador se ajusta aos dados com erro, levando em consideração a suas variâncias de medida.

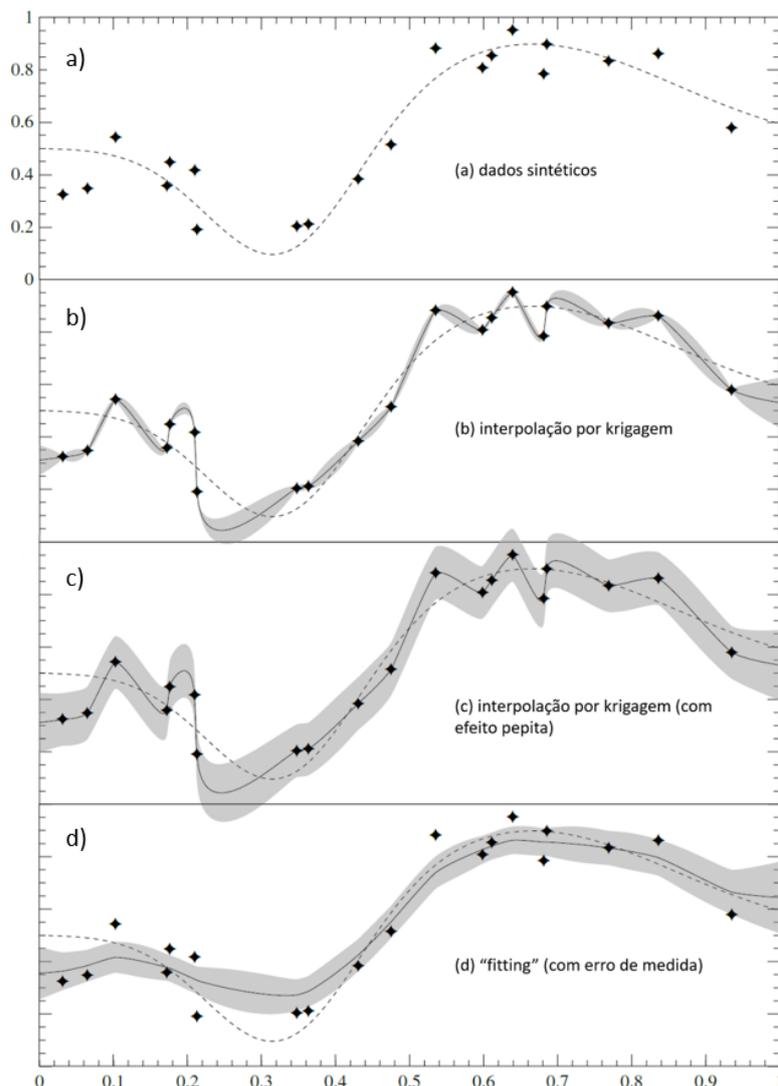


Figura 2.5 – Exemplos unidimensionais de interpolação e *fitting*. a) Dados sintéticos gerados a partir de uma função (linha tracejada). b) Resultado da interpolação por krigagem. A equação de interpolação é plotada como a linha sólida, enquanto a variância do erro de krigagem é mostrada como a faixa sombreada. c) O mesmo que em (b) porém com efeito pepita. d) Resultado da estimativa ao se ajustar o estimador aos dados com variância do erro de medida. Fonte: modificado de Press et al. (2008).

Como citado anteriormente, alguns trabalhos investigam a KVME como uma alternativa aos métodos clássicos, capazes de trabalhar com dados de diferentes confiabilidades. Deraisme & Strydom (2009) valeram-se de amostras obtidas através de distintas técnicas, as quais apresentam diferenças estatísticas significativas. Foram realizadas estimativas utilizando krigagem com variância do erro de medida, levando em consideração a precisão de cada amostra; krigagem ordinária utilizando apenas as amostras com maior confiabilidade; e krigagem ordinária combinando todos as amostras, ignorando a heterogeneidade entre os dados. A partir da análise dos resultados dos três métodos de

estimativa, os autores concluíram que a KVME apresentou os melhores resultados, garantindo estimativas sem viés.

Cornah & Machaka (2015) também empregaram diferentes métodos geoestatísticos, utilizando informações secundárias imprecisas e/ou enviesadas. As técnicas utilizadas foram a krigagem com variância do erro de medida, krigagem ordinária, cokrigagem e cokrigagem multicolocada. Os autores recomendam que, na presença de dados secundários imprecisos, porém sem viés, deve-se integrá-los utilizando a KVME, pois essa fornece uma melhoria na precisão das estimativas em comparação com os demais métodos.

Chautru (2015) utilizou dados auxiliares de menor precisão, porém bem distribuídos na área de interesse, para contribuir na definição de limites em diferentes aplicações na modelagem geológica. Para isso, foram combinados dois métodos geoestatísticos: krigagem com variância do erro de medida e “*conditional expectation with inequalities*”. Chautru (2015) afirma que essas técnicas possibilitam a inclusão do conhecimento empírico, geológico ou geofísico, aprimorando os resultados da modelagem.

#### 2.2.4.1 Obtenção da Variância do Erro de Medida

Nos artigos citados anteriormente (CHAUTRU, 2015; CORNAH; MACHAKA, 2015; DERAISME; STRYDOM, 2009), foram desenvolvidas diferentes técnicas para atribuir valores de variância do erro de medida para as amostras. Independente da forma de obtenção, a imprecisão da amostragem e dos métodos analíticos precisa ser analisada e calculada por métodos estatísticos. Somente a partir destas informações, apropriadas para a utilização na KVME, é possível proceder com as estimativas.

A imprecisão dos dados poderá ser avaliada para cada campanha de amostragem com base nos procedimentos de controle de qualidade. Uma outra abordagem para obter a variância do erro da medida, apresentada com mais detalhe nos estudos de caso deste trabalho, é dada pela modelagem dos variogramas das distintas populações. Nesse método, dados de melhor qualidade permitirão a identificação no variograma de uma estrutura referente a variância em pequena escala ( $c_{MS}$ ); já o variograma modelado aos dados imprecisos mostrará a mesma estrutura anterior, incrementada por uma estrutura que corresponde à variação específica do erro de medida ( $c_{ME}$ ).

### 2.3 *Outliers e Capping*

Um pequeno número de valores muito altos, denominados *outliers*, pode ter grande impacto nas estatísticas dos dados, como a média, a variância e as medidas de continuidade espacial. Além disso, com a presença de *outliers*, poderá ocorrer a superestimativa dos recursos. Considerando estas questões, muitos autores concordam que validar os valores extremos deve ser a primeira tarefa a ser feita em um banco de dados (LEUANGTHONG; NOWAK, 2015). Se a amostra for considerada livre de erros, como por exemplo, erros derivados da manipulação das amostras, erros de digitação, entre outros, e excluindo o erro inerente ao processo de amostragem, esta deve permanecer no banco de dados. Posteriormente, para evitar a superestimativa dos recursos, será necessário encontrar possíveis soluções para limitar a influência dos valores extremos (COSTA, 2003).

*Capping* de valores *outliers* é uma prática bastante comum na indústria de mineração. A técnica envolve o truncamento de valores extremos para um nível inferior, para fins de estimativa dos recursos. Existem várias ferramentas para auxiliar na escolha de um valor apropriado para o truncamento dos teores. Leuangthong & Nowak (2015) apresentam uma revisão de algumas dessas ferramentas, como gráficos de probabilidade, análise de decisão (PARRISH, 1997), análise “*metal-at-risk*”, *cutting curves* (ROSCOE, 1996), e *cutting statistics plots*.

O gráfico de probabilidade é provavelmente a ferramenta mais comumente utilizada, devido a sua simplicidade e a disponibilidade para ser utilizada em *softwares* de geoestatística. Pontos de inflexão e/ou lacunas na distribuição são frequentemente apontados como possíveis valores de truncamento, geralmente entre os percentis 95 a 99 da distribuição. Em alguns casos, com a prática e o conhecimento de um determinado depósito, é possível estabelecer qual o valor de *capping* mais adequado a ser aplicado. Essa escolha deve ser revisada toda vez que o banco de dados e/ou os domínios forem atualizados.

Apesar de a prática de *capping* ser bastante usual, o tratamento de valores *outliers* é um problema complexo que geralmente não possui uma solução padrão. A questão principal é não atribuir pesos muito elevados para os valores extremos, o qual poderia contribuir para uma tonelagem aparente de minério com alto teor que provavelmente não irá existir. Além disso, deve-se garantir que os teores altos contido no depósito não sejam perdidos, o que poderia chegar a inviabilizar o projeto. Logo, *outliers* devem ser considerados caso a caso, com

estudos de sensibilidade, e considerando seu impacto nas estimativas locais e globais. Em geral, a avaliação de recursos será beneficiada com a compreensão da continuidade dos teores altos dentro das estruturas geológicas do depósito (LEUANGTHONG; NOWAK, 2015).

### 3. ESTUDO DE CASO – 1

Este capítulo apresenta um estudo de caso realizado a partir do banco de dados bidimensional Walker Lake (ISAAKS; SRIVASTAVA, 1989). O banco de dados inicial consiste em três variáveis medidas em 78.000 pontos com espaçamento de 1 x 1 metro, em um grid de 260 x 300 metros. Neste trabalho, será utilizado apenas a variável “V” medida em parte por milhão (ppm). O histograma do banco de dados exaustivo é apresentado na Figura 3.1 (a). Os dados compreendem valores de 0 a 1631,16 ppm, com uma acentuada assimetria positiva, refletida pela diferença entre a mediana de 221,25 ppm e a média 277,98 ppm.

A fim de obter o modelo de blocos da variável “V”, o conjunto de pontos foi transformado em 780 blocos de 10 x 10 metros, no qual o teor do bloco representa a média de pontos dentro dos blocos. Esse modelo de blocos é utilizado como cenário de referência, para análise e comparação das estimativas. O histograma e o sumário estatístico do modelo são mostrados na Figura 3.1(b).

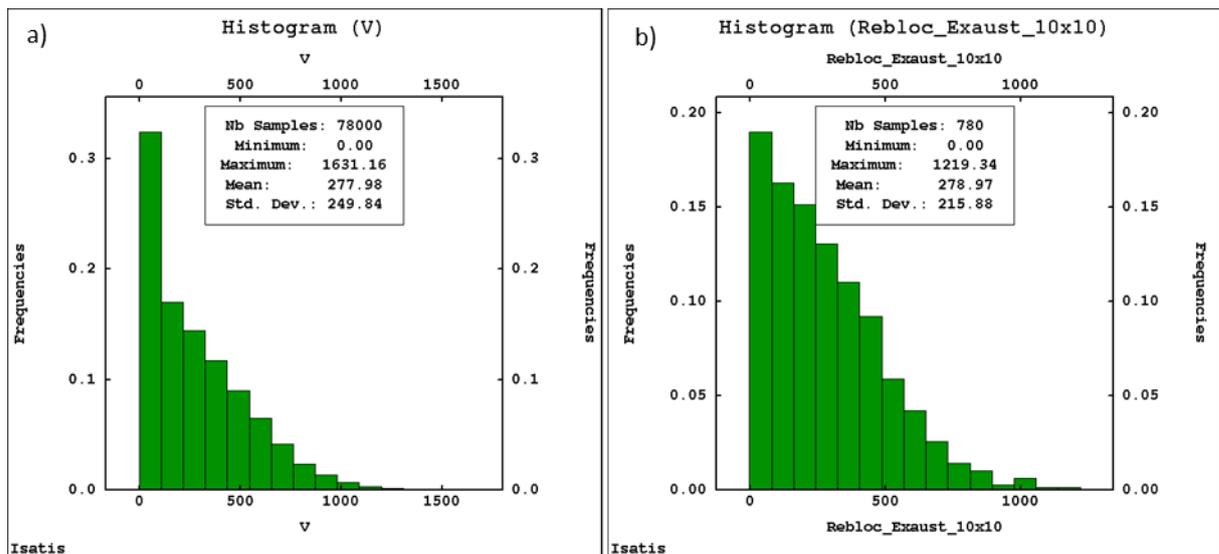


Figura 3.1 – a) Histograma do banco de dados exaustivo da variável V. b) Histograma da variável V após a modificação para blocos de 10 x 10 metros.

Comparando os histogramas, é possível observar que após a transformação de V de pontos para blocos, a média se manteve constante, e o desvio padrão diminuiu, o que era esperado devido ao incremento do suporte. A Figura 3.2 (a) mostra a distribuição espacial da variável V exaustiva em suporte de pontos, e na Figura 3.2 (b) a variável em suporte de blocos. Verifica-se que os teores mais elevados ocorrem na porção oeste do mapa, com orientação aproximadamente norte-sul, e os teores mais baixos encontram-se dispersos no mapa.

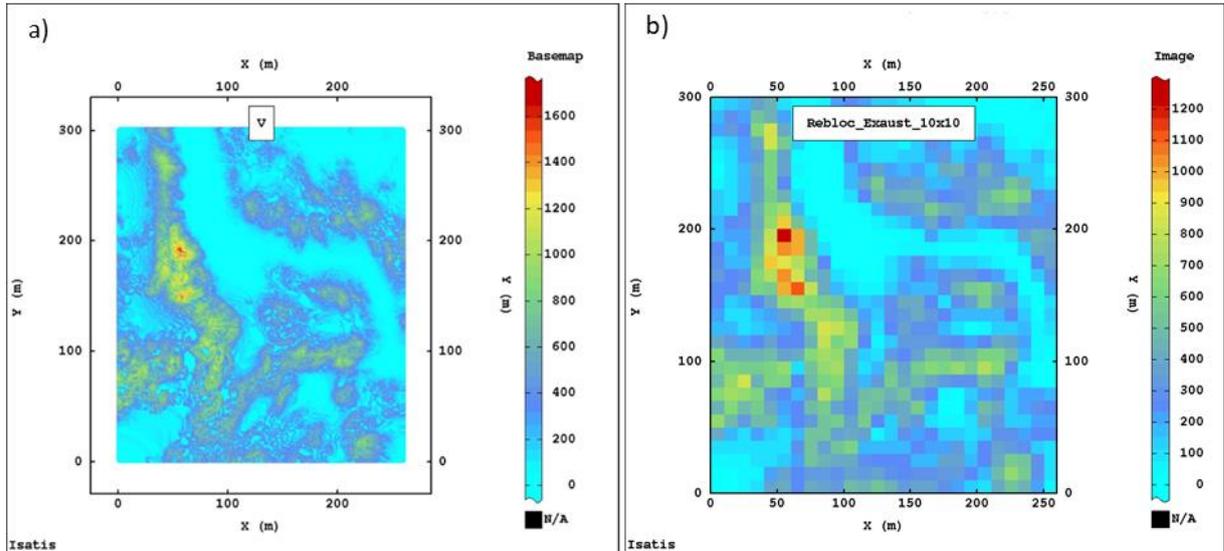


Figura 3.2 – a) Distribuição espacial da variável exaustiva V. b) Distribuição espacial de V após transformação para suporte bloco. Escala de cores em ppm.

A partir do banco de dados exaustivo foram escolhidos 470 pontos para representar um conjunto de dados típico de amostragem, denominado de “banco de dados amostral”. Desses 470 pontos, 195 estão localizados em uma malha aproximadamente regular com espaçamento de 20 x 20 metros. Regiões com teores altos (>500 ppm) foram detalhadas por uma malha de aproximadamente 10 x 10 metros. Essa segunda campanha soma ao conjunto mais 275 amostras.

O histograma do banco de dados amostral da variável V é apresentado na Figura 3.3 (a). Embora o coeficiente de variação de 0,68 seja moderado, existem valores com concentração bastante elevada e que se apresentam agrupados, o que pode vir a trazer complicações nos processos de estimativa. A distribuição espacial das amostras é representada na Figura 3.3 (b), onde observa-se na porção oeste a amostragem preferencial de teores altos.

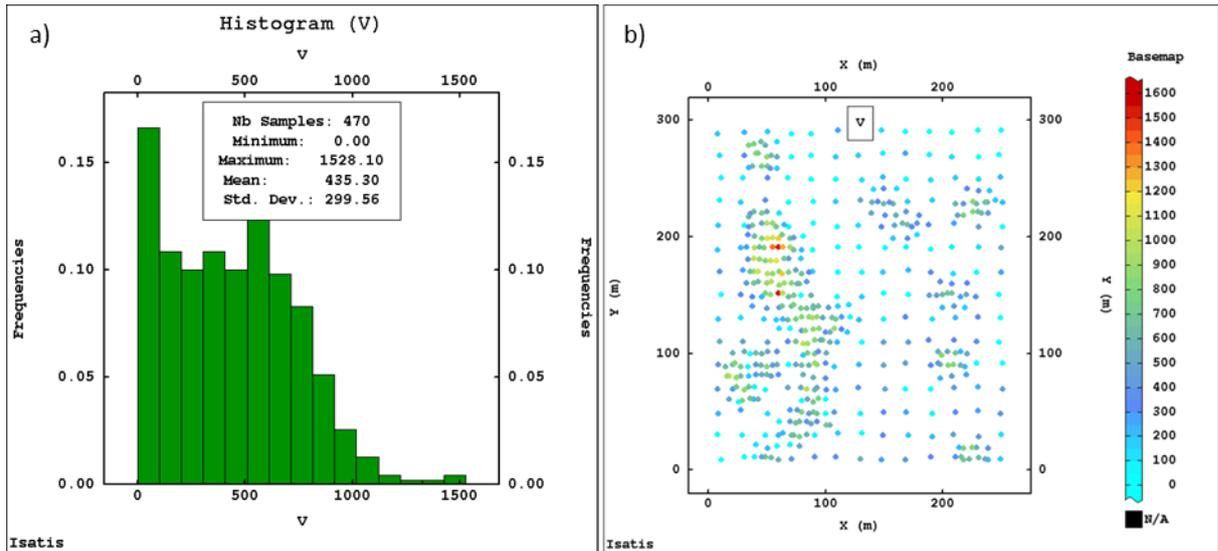


Figura 3.3 – a) Histograma do banco de dados amostral da variável V. b) Distribuição espacial das amostras de V.

A partir do banco de dados amostral da variável V, serão simulados diferentes cenários, incorporando erros com variadas magnitudes. Esses erros representam a imprecisão adicionada durante a etapa de coleta, preparação e análise química das amostras. Não será adicionado viés nos valores das amostras, pois a metodologia aplicada nesse trabalho foi formulada para considerar apenas variâncias nos valores medidos (imprecisão), não corrigindo possíveis vieses presentes em bancos de dados.

Cada cenário será apresentado junto com a descrição do procedimento utilizado para adição dos erros nas amostras. Estes cenários serão então estimados por Krigagem Ordinária (OK) e Krigagem com Variância do Erro de Medida (KVME), sendo necessário definir a variância do erro da medida para cada cenário elaborado. Por fim, os resultados das estimativas serão comparados e discutidos.

### 3.1 Cenários Hard – Soft

Os cenários deste estudo de caso simulam um depósito que possui dados com diferentes qualidades. Será estabelecido como *hard data* os 195 pontos amostrais localizados na malha de 20 x 20 metros, e sua variância do erro da medida será constante de 0 ppm<sup>2</sup>. Os dados *hard* representam uma malha regular de exploração, obtida por sondagem diamantada. Já as outras 275 amostras, localizadas em uma malha de aproximadamente 10 x 10 metros, serão designadas como *soft data*, e erros serão adicionados em seus teores medidos. Os dados *soft* representam uma típica malha de adensamento ou sondagem de produção. A variância do erro de medida do *soft data* será determinada através de interpretações dos variogramas.

A Figura 3.4 exibe a localização das amostras que foram definidas como *hard data*, e aquelas que foram definidas como *soft data*.

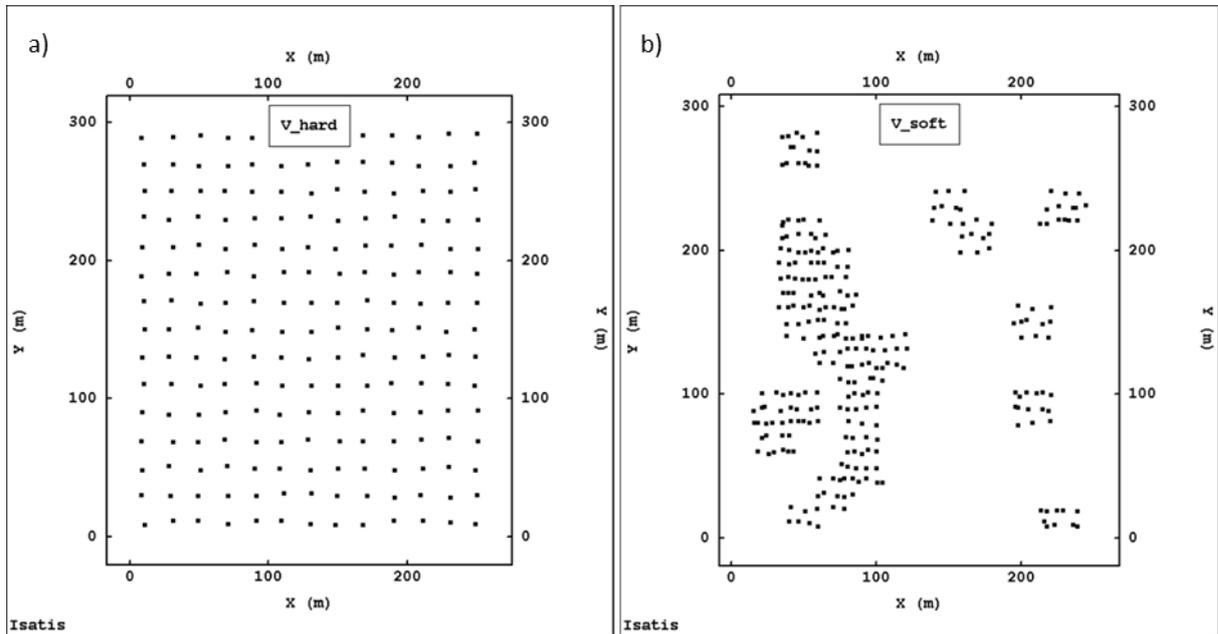


Figura 3.4 – a) Localização das amostras “hard data”. b) Localização das amostras “soft data”.

### 3.2 Adicionando imprecisão aos dados

Os valores lidos nos 275 *soft data* foram perturbadas pela adição de um erro aleatório, assumindo uma distribuição gaussiana com média zero, e desvio padrão determinado como uma porcentagem do desvio padrão do banco de dados original. Foram adotados três diferentes valores para esse desvio padrão, denominados de *std.dev\_1*, *std.dev\_2* e *std.dev\_3*, e para cada um deles foram simulados 10 cenários para minimizar o efeito de anomalias na simulação e garantir a consistência nos dados. Os distintos valores adotados como desvio padrão tiveram o objetivo de simular diferentes magnitudes de incerteza, dessa forma adotou-se  $std.dev_1 < std.dev_2 < std.dev_3$ . Em relação ao banco de dados original, em *std.dev\_1* o desvio padrão aumentou em 4%, em *std.dev\_2* em 9% e em *std.dev\_3* em 18%.

Para ilustrar o ruído gerado nos teores, as Figuras 3.5 (a), 3.6 (a) e 3.7 (a) apresentam diagramas de dispersão, onde o eixo  $x$  representa as 275 amostras do banco de dados original de  $V$ , e o eixo  $y$  representa as 275 amostras de três diferentes cenários, referentes aos distintos níveis de incerteza (*std.dev\_1*, *std.dev\_2* e *std.dev\_3*). Com o aumento da intensidade do erro, mais descorrelacionado ficam os dados simulados dos dados originais. Também se verifica que quanto maior o nível de incerteza, mais elevado é a quantidade de amostras com teor zero. Isso acontece, porque todos aqueles valores que após a adição do

erro tornaram-se negativos, foram transformados em zero, assegurando a coerência nos dados. As Figuras 3.5 (b), 3.6 (b) e 3.7 (b) mostram a distribuição espacial de três cenários, que são o conjunto das 275 amostras, *soft data*, obtidas a partir da adição de erros aos dados originais, junto dos 195 *hard data*, que não sofreram modificações.

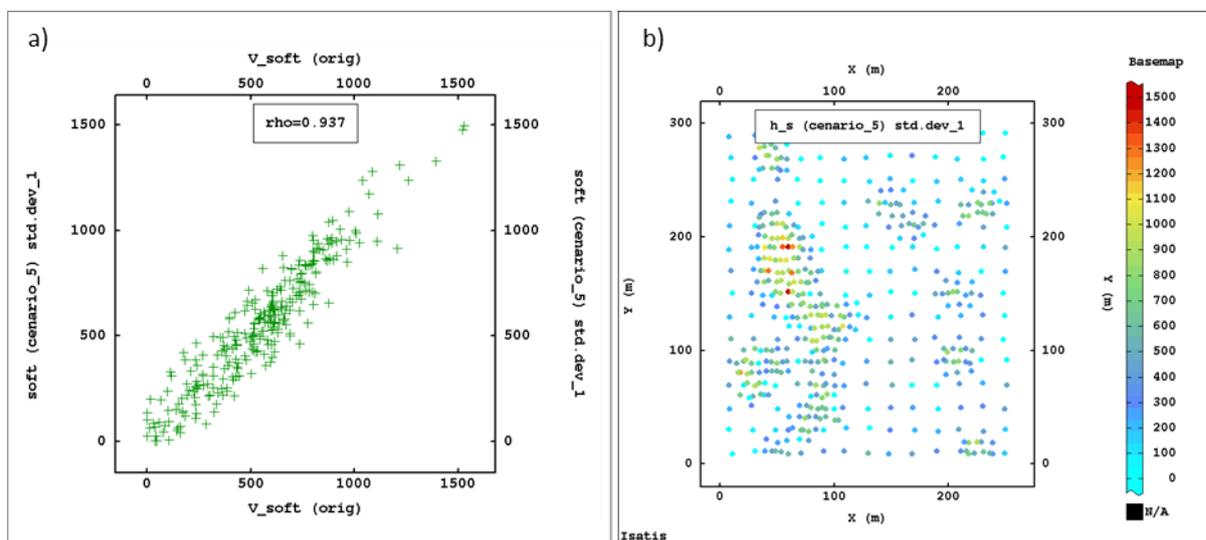


Figura 3.5 – a) Diagrama de dispersão dos 275 dados originais de V versus 275 dados do cenário 5 (*std.dev\_1*). b) Distribuição espacial do cenário\_5 (*std.dev\_1*).

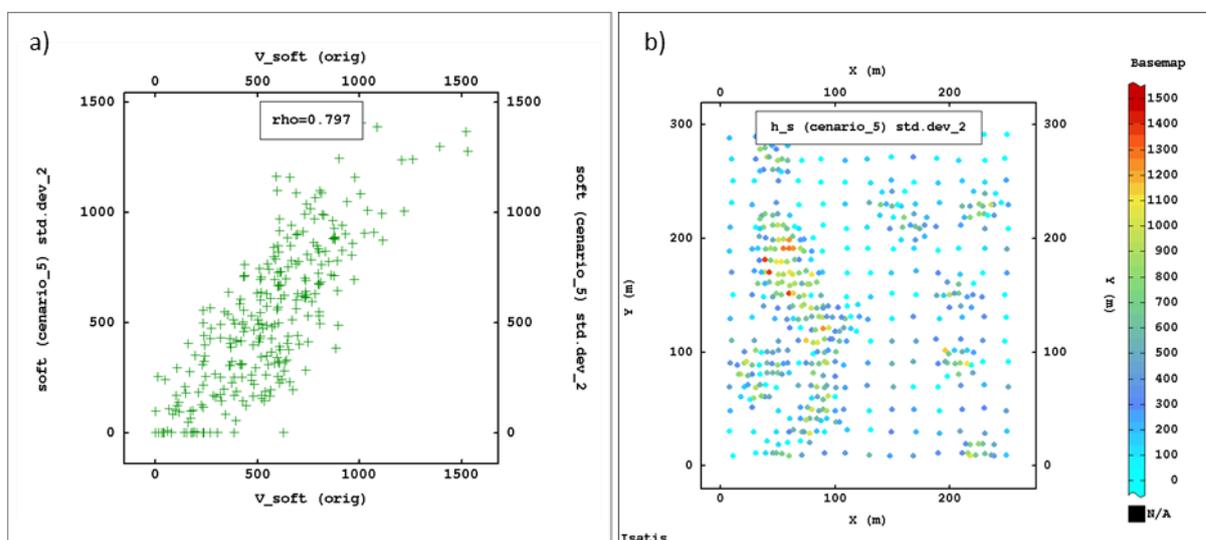


Figura 3.6 – a) Diagrama de dispersão dos 275 dados originais de V versus 275 dados do cenário 5 (*std.dev\_2*). b) Distribuição espacial do cenário\_5 (*std.dev\_2*).

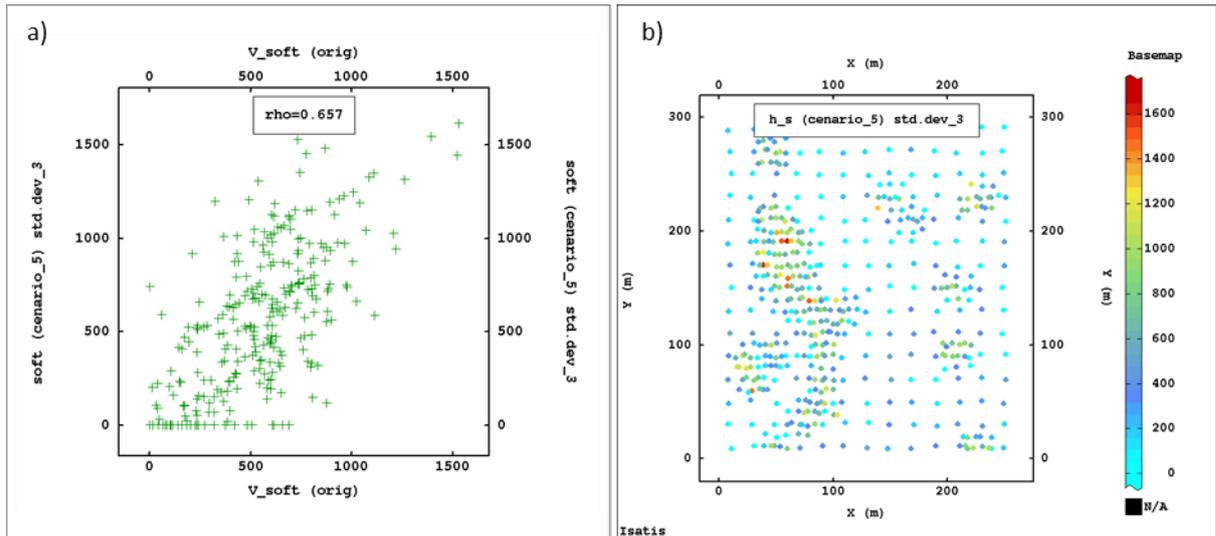


Figura 3.7 – a) Diagrama de dispersão dos 275 dados originais de V versus 275 dados do cenário 5 (*std.dev\_3*). b) Distribuição espacial do cenário\_5 (*std.dev\_3*).

A Tabela 3.1 apresenta um resumo estatístico dos bancos de dados resultantes dos cenários simulados, mostrando também a média estatística destes. As médias de todos os cenários se mantiveram próximas ao valor original de V, o que era esperado, já que não houve adição de viés. Além disso, o truncamento dos valores negativos em zero não gerou viés significativo nos dados. O desvio padrão aumentou até 20% devido a adição da componente de erro aos valores do *dataset* original. A Tabela 3.1 mostra também que com o incremento do erro, mais desconhecidos os dados de cada banco de dados simulado estão dos seus valores originais.

Tabela 3.1 – Estatística descritiva dos cenários simulados em função do desvio padrão adicionado na geração dos erros.

Cenário	Coefficiente de Correlação	Mínimo	Máximo	Média	Desvio Padrão
V (original)	1	0	1528,1	435,3	299,56
Cen_1 (std.dev_1)	0,971	0	1694,9	434,58	309,06
Cen_5 (std.dev_1)	0,967	0	1492,06	436,64	305,34
Cen_10 (std.dev_1)	0,968	0	1511,9	437,45	307,3
Média Cenários (std.dev_1)	0,969	0	1552,88	435,52	307,43
Cen_1 (std.dev_2)	0,894	0	1306,98	438,73	315,48
Cen_5 (std.dev_2)	0,880	0	1403,1	417,27	322,14
Cen_10 (std.dev_2)	0,891	0	1623,91	450,64	334,24
Média Cenários (std.dev_2)	0,891	0	1601,66	435,98	326,48
Cen_1 (std.dev_3)	0,777	0	1737,87	437,39	357,73
Cen_5 (std.dev_3)	0,791	0	1641,07	448,57	368,81
Cen_10 (std.dev_3)	0,786	0	1584,56	447,2	352,67
Média Cenários (std.dev_3)	0,799	0	1754,99	442,69	355,77

### 3.3 Continuidade Espacial e Variância do Erro de Medida

Para utilizar a KVME, cada teor do banco de dados precisa estar associado a um valor que represente a variância do erro em sua medida. Para encontrar esse valor, primeiramente, foi modelado um variograma omnidirecional dos 195 *hard data* (Figura 3.8). Como foi assumido que esses dados não possuem erro, ou seja, sua variância do erro de medida é de 0 ppm<sup>2</sup>, o variograma apresentará apenas o efeito pepita referente a variância em microescala ( $c_{MS}$ ).

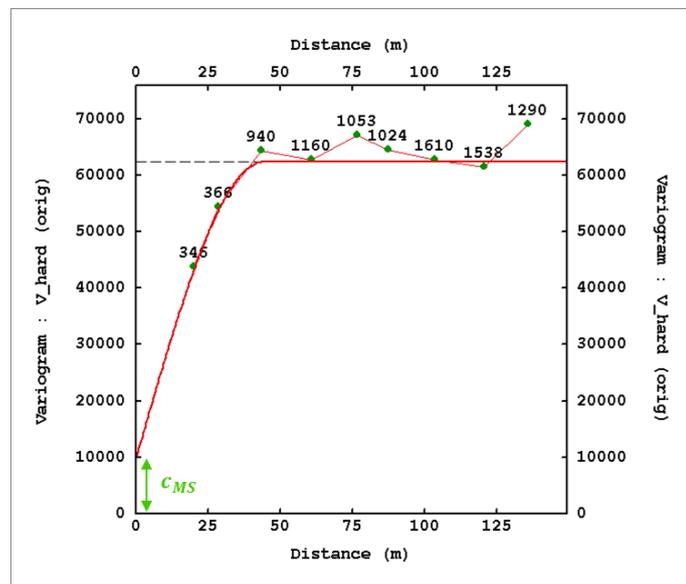


Figura 3.8 – Variograma omnidirecional *hard data* (195 amostras), mostrando o efeito pepita relativo a variância em microescala ( $c_{MS}$ ).

Posteriormente, o variograma omnidirecional de cada cenário foi modelado (*hard + soft data*), os quais mostram a mesma estrutura anterior,  $c_{MS}$ , incrementada por uma segunda componente do efeito pepita, esta correspondente à variância do erro de medida,  $c_{ME}$ . O valor encontrado de  $c_{ME}$  para cada cenário será utilizado como variância do erro de medida para os 195 *soft data*. Para definir o valor de  $c_{ME}$  a partir desta metodologia, foi assumido que os dados estão no mesmo domínio estacionário. As Figuras 3.9, 3.10 e 3.11 apresentam os variogramas omnidirecionais de cenários referentes aos distintos níveis de incerteza, respectivamente *std.dev\_1*, *std.dev\_2* e *std.dev\_3*. As setas indicam os componentes do efeito pepita. Estes variogramas foram escolhidos aleatoriamente para exemplificação da metodologia.

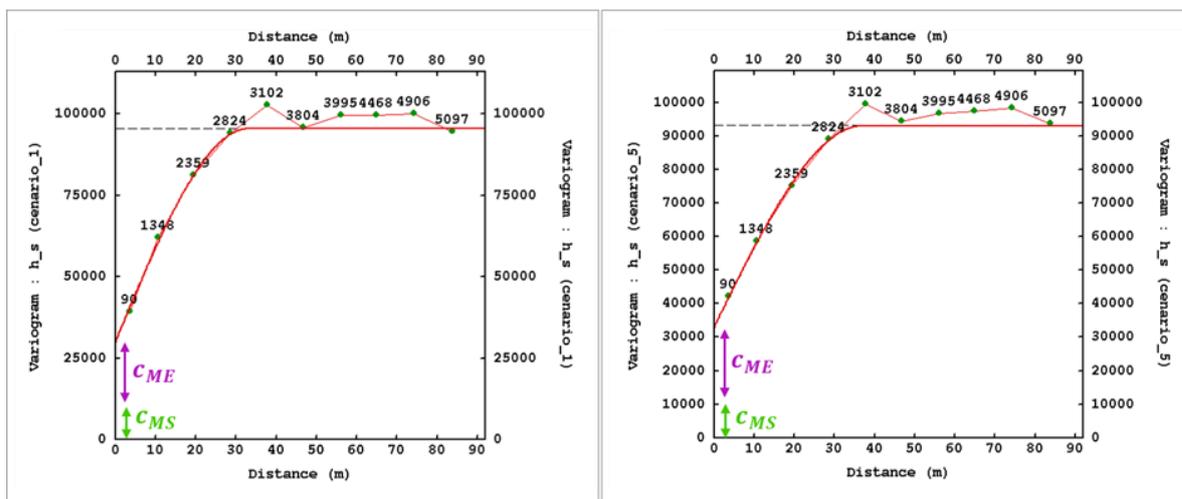


Figura 3.9 – Variogramas omnidirecionais de cenários referentes ao erro  $std.dev_1$ , indicando o efeito pepita relativo a variância em microescala ( $c_{MS}$ ) e relativo a variância do erro de medida ( $c_{ME}$ ).

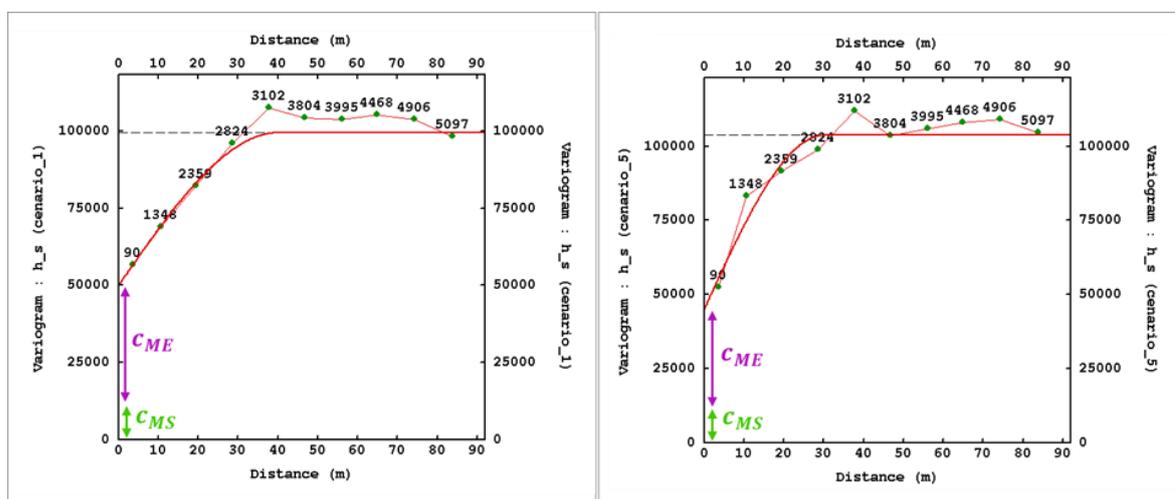


Figura 3.10 – Variogramas omnidirecionais de cenários referentes ao erro  $std.dev_2$ , indicando o efeito pepita relativo a variância em microescala ( $c_{MS}$ ) e relativo a variância do erro de medida ( $c_{ME}$ ).

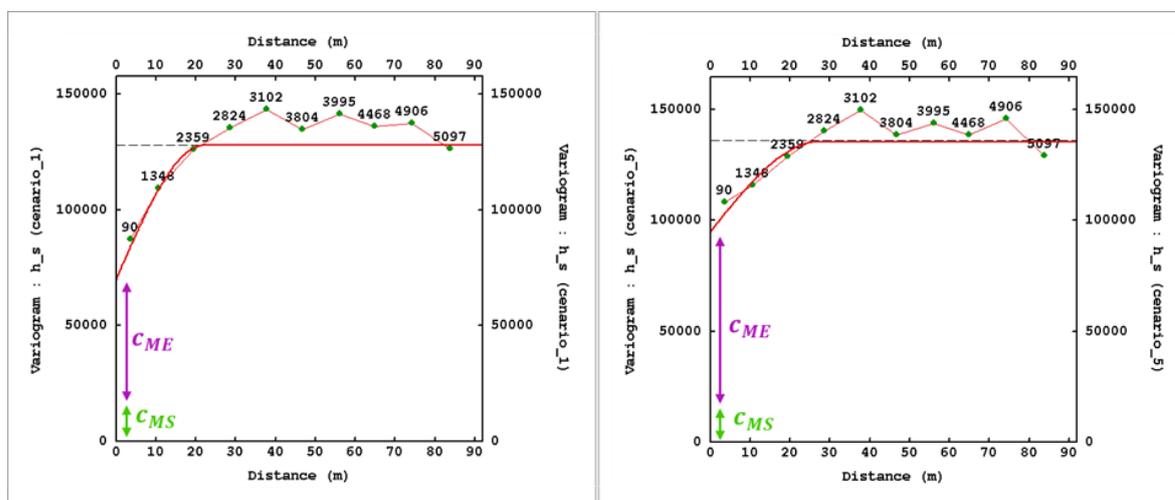


Figura 3.11 – Variogramas omnidirecionais de cenários referentes ao erro  $std.dev_3$ , indicando o efeito pepita relativo a variância em microescala ( $c_{MS}$ ) e relativo a variância do erro de medida ( $c_{ME}$ ).

As Tabelas 3.2, 3.3 e 3.4 apresentam os resultados obtidos com a modelagem do variograma omnidirecional utilizando apenas *hard data*, e dos variogramas omnidirecionais de todos os cenários (*hard data + soft data*). Com base nessa modelagem, foram determinados valores para a variância do erro de medida, o qual será o efeito pepita total menos o efeito pepita relativo a variância de microescala ( $c_{ME} = c_0 - c_{MS}$ ). Para realizar a krigagem ordinária, foi utilizado o variograma direcional médio dos dados *hard data + soft data*, apresentado no final das Tabelas 3.2, 3.3 e 3.4, sem distinção entre  $c_{ME}$  e  $c_{MS}$ . Já na KVME, foi mantido o valor de efeito pepita encontrado para os valores *hard data*, assumido como  $c_{MS}$ , e foi adicionado o valor médio calculado de  $c_{ME}$  como variância do erro de medida no *input* dos dados.

Tabela 3.2 – Variogramas referentes aos cenários com erro *std.dev\_1*. (\*Valor adicionado como variância do erro de medida).

<b>std.dev_1</b>	<b><math>c_0</math></b>	<b><math>c_{MS}</math></b>	<b><math>c_{ME}</math></b>	<b>Contribuição (Sph)</b>	<b>Alcance</b>
variog omni hard data	10.000	10.000	0	52.000	44m
variog omni cenário_1	30.000	10.000	20.000	65.500	33m
variog omni cenário_2	25.000	10.000	15.000	66.000	34m
variog omni cenário_3	26.000	10.000	16.000	72.500	32m
variog omni cenário_4	28.000	10.000	18.000	67.400	35m
variog omni cenário_5	33.000	10.000	23.000	60.000	37m
variog omni cenário_6	34.000	10.000	24.000	61.200	35m
variog omni cenário_7	34.000	10.000	24.000	60.500	34m
variog omni cenário_8	37.000	10.000	27.000	55.500	36m
variog omni cenário_9	32.000	10.000	22.000	62.400	35m
variog omni cenário_10	22.000	10.000	12.000	72.500	32m
variog dir médio cenários	30.100	10.000	(20.100)*	$31.000 \left( \frac{N157^\circ}{35m}; \frac{N67^\circ}{16m} \right) + 34.000 \left( \frac{N157^\circ}{82m}; \frac{N67^\circ}{20m} \right)$	

Tabela 3.3 – Variogramas referentes aos cenários com erro *std.dev\_2*. (\*Valor adicionado como variância do erro de medida).

<b>std.dev_2</b>	<b><math>c_0</math></b>	<b><math>c_{MS}</math></b>	<b><math>c_{ME}</math></b>	<b>Contribuição (Sph)</b>	<b>Alcance</b>
variog omni hard data	10.000	10.000	0	52.000	40m
variog omni cenário_1	50.000	10.000	40.000	49.000	40m
variog omni cenário_2	55.000	10.000	45.000	53.500	35m
variog omni cenário_3	57.000	10.000	47.000	59.000	35m
variog omni cenário_4	52.000	10.000	42.000	57.000	28m
variog omni cenário_5	45.000	10.000	35.000	58.700	30m
variog omni cenário_6	50.000	10.000	40.000	61.500	28m
variog omni cenário_7	50.000	10.000	40.000	59.000	32m
variog omni cenário_8	45.000	10.000	35.000	54.500	33m
variog omni cenário_9	45.000	10.000	35.000	52.000	30m
variog omni cenário_10	55.000	10.000	45.000	56.500	33m
variograma dir médio	50.400	10.000	(40.400)*	$1.900 \left( \frac{N157^\circ}{15m}; \frac{N67^\circ}{15m} \right) + 40.000 \left( \frac{N157^\circ}{70m}; \frac{N67^\circ}{20m} \right)$	

Tabela 3.4 – Variogramas referentes aos cenários com erro *std.dev\_3*. (\*Valor adicionado como variância do erro de medida).

<b>std.dev_3</b>	<b><math>c_0</math></b>	<b><math>c_{MS}</math></b>	<b><math>c_{ME}</math></b>	<b>Contribuição (Sph)</b>	<b>Alcance</b>
variog omni hard data	10.000	10.000	0	52.000	44m
variog omni cenário_1	70.000	10.000	60.000	58.000	22m
variog omni cenário_2	80.000	10.000	70.000	45.000	18m
variog omni cenário_3	80.000	10.000	70.000	53.000	25m
variog omni cenário_4	80.000	10.000	70.000	41.000	25m
variog omni cenário_5	95.000	10.000	85.000	40.500	27m
variog omni cenário_6	85.000	10.000	75.000	42.700	28m
variog omni cenário_7	90.000	10.000	80.000	41.000	25m
variog omni cenário_8	70.000	10.000	60.000	47.000	35m
variog omni cenário_9	75.000	10.000	65.000	46.000	25m
variog omni cenário_10	65.000	10.000	55.000	59.000	28m
variograma dir médio	79.000	10.000	(69.000)*	$25000 \left( \frac{N157^\circ}{20m}; \frac{N67^\circ}{5m} \right) + 29500 \left( \frac{N157^\circ}{35m}; \frac{N67^\circ}{10m} \right)$	

A utilização de valores médios se deve ao fato de os variogramas serem bastantes parecidos, e que seria dispendioso modelar manualmente os variogramas direcionais de todos os cenários. Esse variograma médio foi obtido levando em consideração todos os cenários, escolhendo parâmetros que se ajustassem a grande maioria deles. Este recurso foi testado e, dessa forma, foi assegurado não haver modificações significativas no resultado das estimativas a partir da utilização de valores médios.

### 3.4 Estimativas

A estimativa foi realizada em suporte de blocos, de dimensões 10 x 10 metros. O raio de busca aplicado foi anisotrópico, com a direção e distâncias obtidas na etapa de variografia.

Foi utilizada a opção de restrição da busca por octantes, com um máximo de 2 amostras por setor. A estimativa dos blocos foi discretizada em 5 x 5 x 1 pontos internos. Esses parâmetros apresentaram os melhores resultados considerando os critérios de validação das estimativas. Os mesmos parâmetros de busca por amostras e modelo de blocos foram utilizados nos 10 cenários de cada erro, assim a comparação das estimativas pode ser feita de maneira adequada.

As Figuras 3.12 (a) e 3.13 (a) apresentam a elipse de busca sobre as amostras do mesmo cenário, demonstrando os pesos fornecidos para estas nos diferentes tipos de estimativa. As diferenças de pesos e valor final do bloco estimado podem ser conferidas ao comparar os resultados dos cálculos nas Figuras 3.12 (b) e 3.13 (b). Na Figura 3.13 (b), verifica-se a existência de uma coluna a mais (“variance”) para o cálculo dos pesos (“lambda”). Tendo em conta essa variância do erro de medida, aquelas amostras com erro associado contribuirão de menor forma nas estimativas, como no exemplo abaixo, as amostras 3, 7, 8 e 9 são aquelas que tem seu peso diminuído devido ao erro de medida.

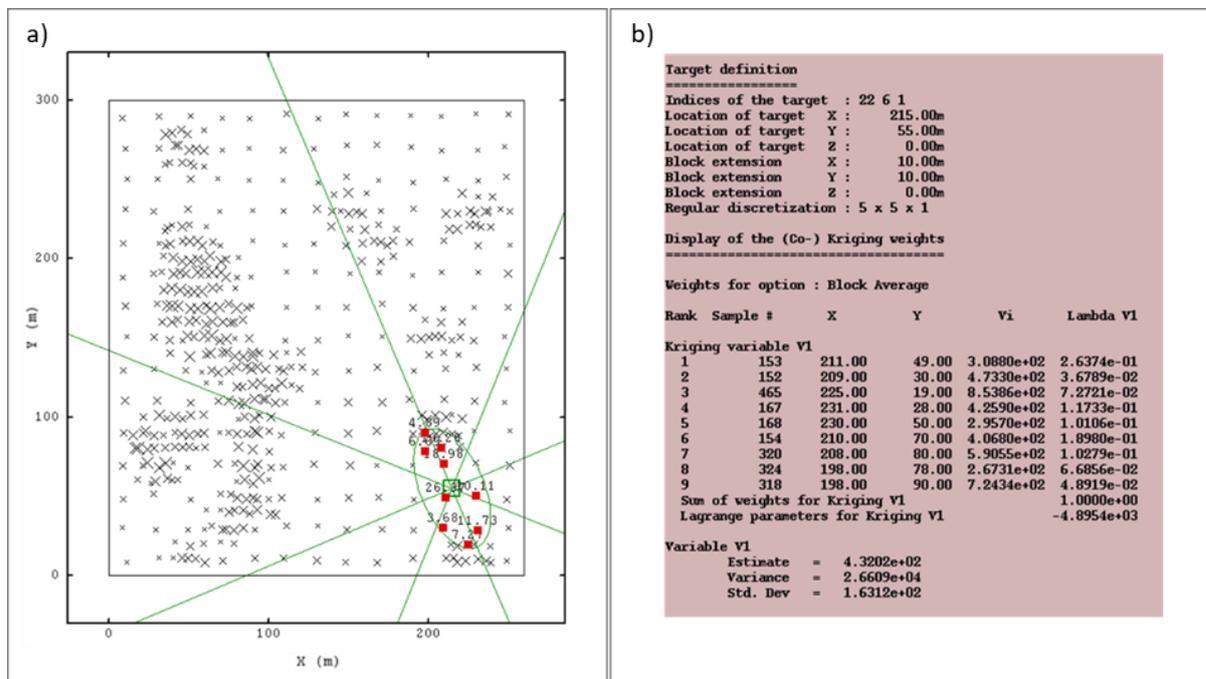


Figura 3.12 – a) Elipse de busca sobre as amostras de um cenário estimado por krigagem ordinária. As amostras, representadas por quadrados em vermelho, indicam seus pesos. b) Informações sobre a estimativa por krigagem ordinária do bloco com índice  $x = 22$ ,  $y = 6$  e  $z = 1$ .

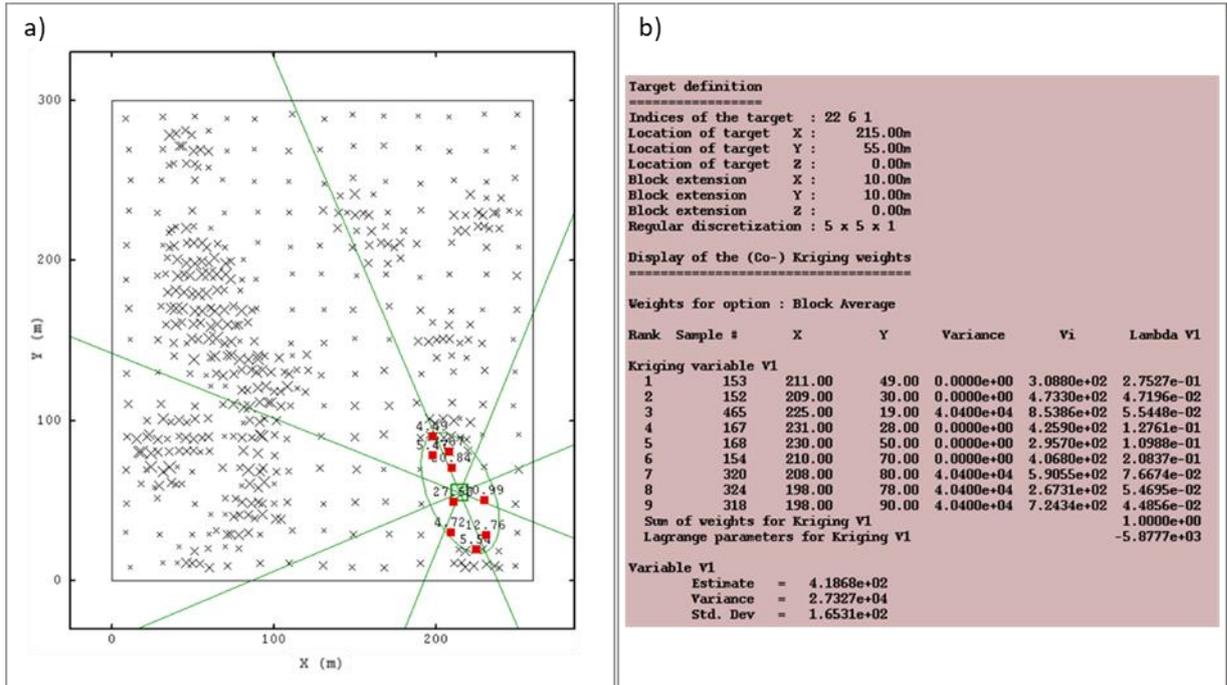


Figura 3.13 – a) Elipse de busca sobre as amostras de um cenário estimado por krigagem com variância do erro de medida. As amostras, representadas por quadrados em vermelho, indicam seus pesos. b) Informações sobre a estimativa por krigagem com variância do erro de medida do bloco com índice  $x = 22$ ,  $y = 6$  e  $z = 1$ .

As Figuras 3.14, 3.15 e 3.16 ilustram, cada uma com diferente nível de incerteza, o mesmo cenário estimado por KVME e OK, junto com a referência “*Rebloc\_Exhaust\_10x10*”. Todos os cenários estão com a mesma legenda de cores para auxiliar na comparação dos mapas. Visualmente, as estimativas apresentam ser bastantes semelhantes, com diferenças apenas nos teores mais altos, na porção oeste do mapa, onde estão concentrados a maioria dos *soft data*. Observa-se, que a OK e a KVME estão suavizadas em comparação com o modelo de referência, e quanto maior o erro adicionado nos cenários, maior a suavização.

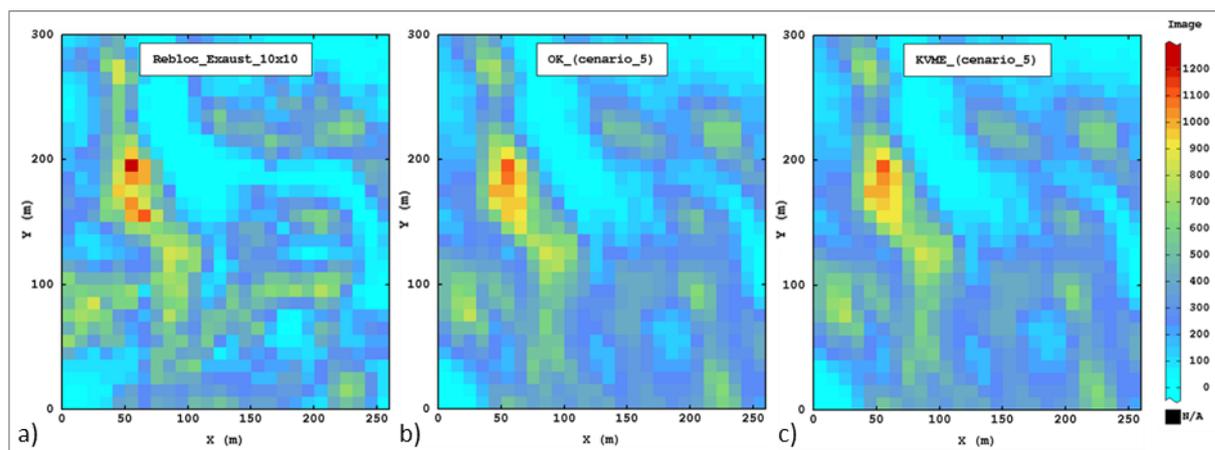


Figura 3.14 – a) Modelo de referência. b) Krigagem ordinária do cenário\_5 referente ao erro  $std.dev_1$ . c) Krigagem com variância do erro de medida do cenário\_5 referente ao erro  $std.dev_1$ .

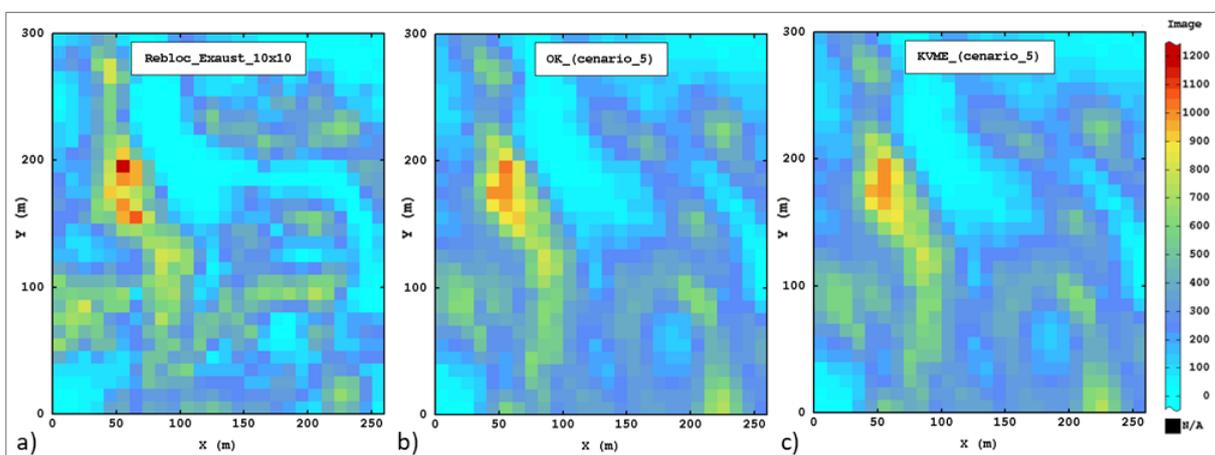


Figura 3.15 – a) Modelo de referência. b) Krigagem ordinária do cenário\_5 referente ao erro  $std.dev_2$ . c) Krigagem com variância do erro de medida do cenário\_5 referente ao erro  $std.dev_2$ .

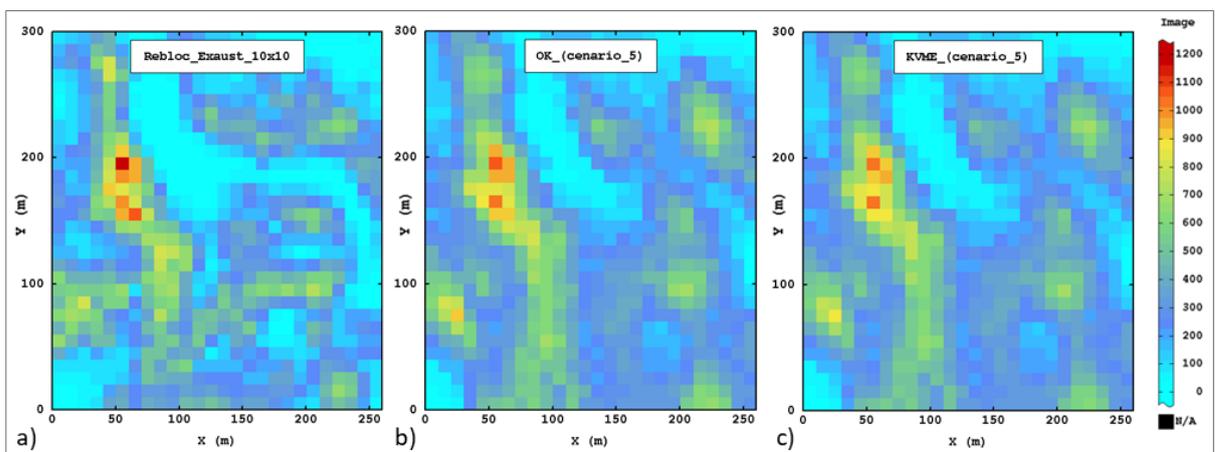


Figura 3.16 – a) Modelo de referência. b) Krigagem ordinária do cenário\_5 referente ao erro  $std.dev_3$ . c) Krigagem com variância do erro de medida do cenário\_5 referente ao erro  $std.dev_3$ .

Os resultados das estimativas são sumarizados na Tabela 3.5. Foi obtida a média dos 10 cenários referentes ao mesmo nível de erro para uma comparação global entre as krigagens. Com o aumento do erro de medida nos teores, uma gradual redução do coeficiente de correlação entre os cenários e o modelo de referência acontece. Apesar de ser pequena a diferença, a correlação entre a KVME e a referência é sempre superior se comparado com a OK. O desvio quadrático médio, o qual indica quanto os valores desviam-se do valor de referência, mostra uma diferença menor da KVME e a variável V reblocada, em comparação com a OK. Verifica-se nos resultados da Tabela 3.5, que quanto maior o erro associado aos dados ( $std.dev_3 > std.dev_2 > std.dev_1$ ), mais as estimativas por KVME se sobressaem a OK.

Tabela 3.5 – Valores de referência e média dos cenários estimados por OK e KVME. Mínimo de todos os cenários estimados de 0 ppm.

Cenários	Máximo (ppm)	Média (ppm)	Desvio Padrão (ppm)	Coefficiente de Correlação	Erro Quadrático Médio (ppm)
V (reblockado)	1219,34	278,97	215,88	1	-
Cenários OK ( $std.dev_1$ )	1163,19	288,92	193,02	0,883	101,54
Cenários KVME ( $std.dev_1$ )	1138,42	287,95	190,85	0,889	98,68
Cenários OK ( $std.dev_2$ )	1115,02	294,38	188,94	0,863	110,35
Cenários KVME ( $std.dev_2$ )	1078,52	292,32	186,05	0,875	105,50
Cenários OK ( $std.dev_3$ )	1086,97	311,04	186,36	0,823	129,09
Cenários KVME ( $std.dev_3$ )	1057,22	305,39	183,42	0,852	117,42

Para representar visualmente a correlação dos cenários estimados com o modelo de referência, as Figuras 3.17, 3.18 e 3.19 apresentam diagramas de dispersão, no qual o eixo  $x$  representa o banco de dados exaustivo reblocado, e o eixo  $y$  as estimativas. Os triângulos em vermelho simbolizam os blocos dos cenários estimados por krigagem ordinária, e os círculos em azul por krigagem com variância do erro de medida. A linha de referência  $y = x$  representa a correlação perfeita entre o valor estimado e a referência. Percebe-se, que quanto maior o nível de incerteza ( $std.dev_1$ ,  $std.dev_2$  e  $std.dev_3$ ), maior a diferença entre o valor estimado e a referência, e maior a dispersão dos pontos em torno da reta.

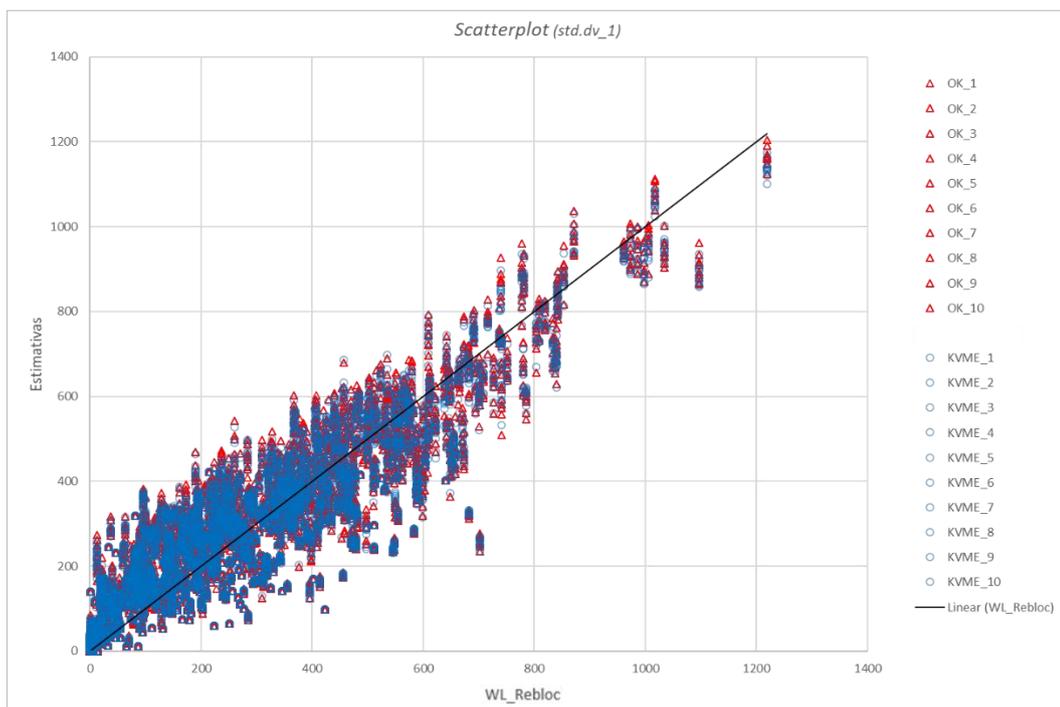


Figura 3.17 – Diagrama de dispersão: OK\_cenários versus referência (triângulos em vermelho), e KVME\_cenários versus referência (círculos em azul). Nível de incerteza  $std.dev_1$ .

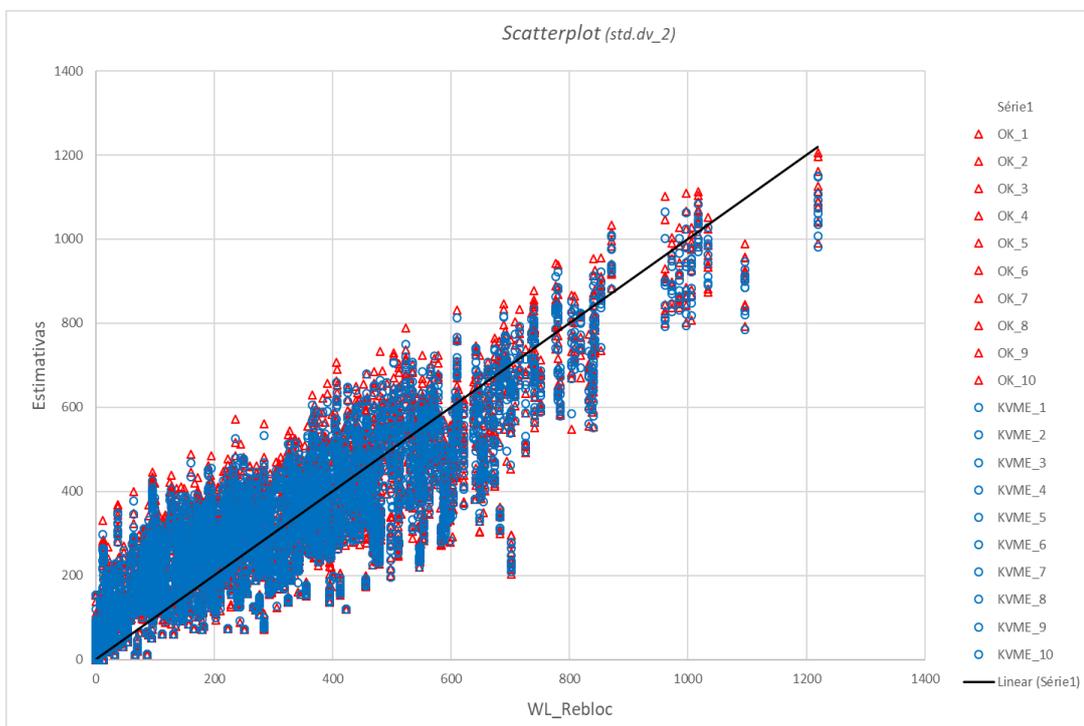


Figura 3.18 – Diagrama de dispersão: OK\_cenários versus referência (triângulos em vermelho), e KVME\_cenários versus referência (círculos em azul). Nível de incerteza  $std.dev_2$ .

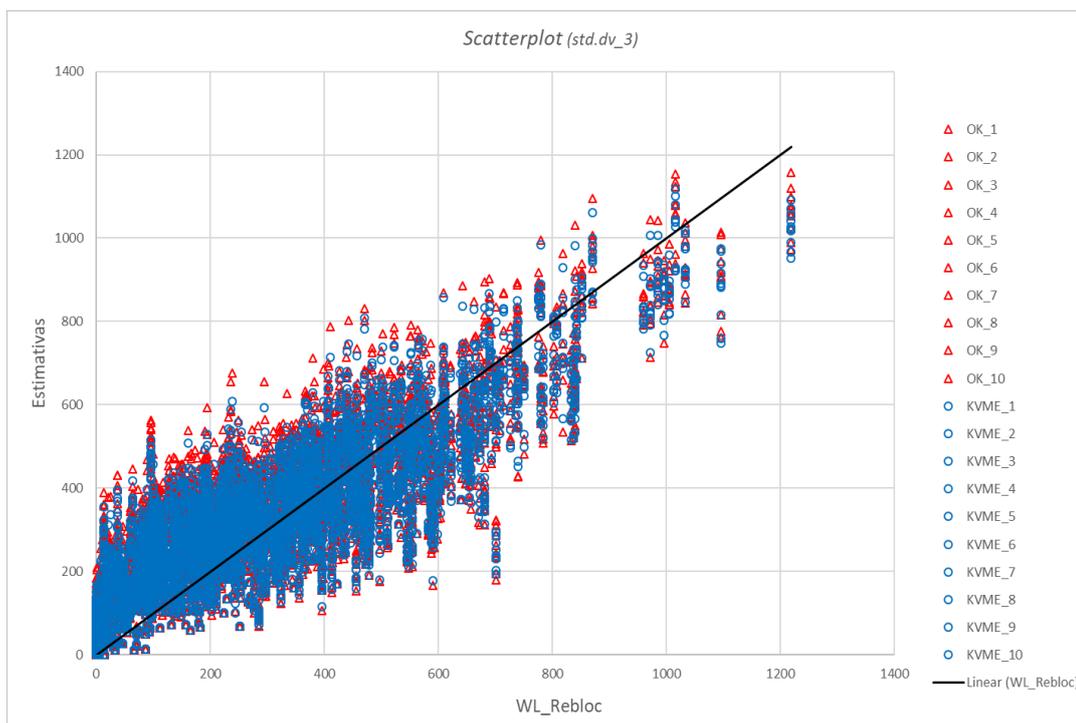


Figura 3.19 – Diagrama de dispersão: OK\_cenários *versus* referência (triângulos em vermelho), e KVME\_cenários *versus* referência (círculos em azul). Nível de incerteza *std.dev\_3*.

Observa-se ainda nos diagramas de dispersão, principalmente na Figura 3.19, que os blocos estimados por krigagem ordinária tendem a ter seus teores superestimados, enquanto os blocos estimados por krigagem com variância do erro de medida mantem seus teores mais próximos a referência. Esse comportamento permanece em quase toda a faixa de teores, porém, nos teores mais altos, a partir de 900 ppm, os cenários estimados por OK se aproximam mais da referência, enquanto a KVME subestima os teores dos blocos. Para compreender melhor esse comportamento, a Tabela 3.6 e a Figura 3.20 apresentam o erro quadrático médio por faixa de teor.

Tabela 3.6 – Erro quadrático médio dos cenários por faixa de teor. Células na cor verde representam erro quadrático médio da KVME < OK; em amarelo desvio KVME = OK; em vermelho desvio KVME > OK.

Faixa de Teores (ppm)	Erro quadrático médio (média cenários - <i>std.dv_1</i> )		Erro quadrático médio (média cenários - <i>std.dv_2</i> )		Erro quadrático médio (média cenários - <i>std.dv_3</i> )	
	OK	KVME	OK	KVME	OK	KVME
<100	86,27	84,52	99,10	94,00	126,62	113,82
>100; <200	96,07	94,89	109,70	103,32	138,97	124,10
>200; <300	93,86	92,57	94,64	90,58	111,28	101,85
>300; <400	92,43	89,91	95,17	90,18	98,45	90,85
>400; <500	116,95	116,47	119,26	116,93	124,60	121,51
>500; <600	133,15	134,78	147,53	146,87	150,75	150,73
>600; <700	135,69	136,71	148,47	148,22	145,54	145,74
>700; <800	153,19	152,29	153,05	153,07	165,36	162,56
>800; <900	87,08	88,89	116,99	117,89	147,64	147,72
>900; <1000	56,33	59,73	95,45	95,21	117,95	123,28
>1000; <1100	116,31	118,82	120,14	127,47	137,15	144,28
>1100	56,15	79,04	110,39	147,19	157,04	193,51

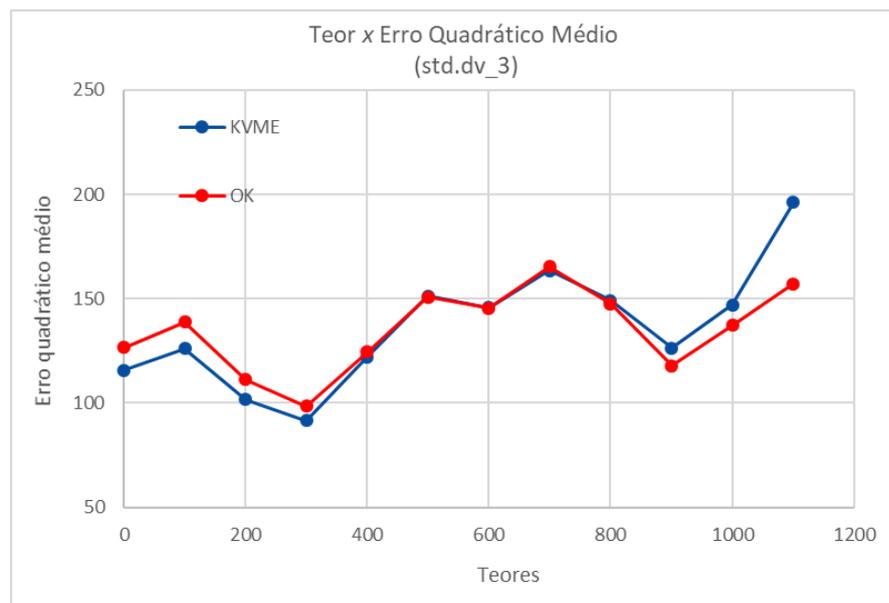


Figura 3.20 – Erro quadrático médio dos cenários (*std.dv\_3*) por faixa de teor. Linha em azul representa os desvios da KVME, e linha em vermelho desvios da OK.

Conforme a Tabela 3.6 e Figura 3.20, até 500 ppm o erro quadrático médio da KVME é menor que a OK; entre 500 e 800 ppm os desvios são muito semelhantes; e depois de 800 ppm o erro quadrático médio da OK é menor que a KVME. Esse comportamento da KVME, de subestimar os blocos com teores acima de 800 ppm, acontece porque as amostras de teores mais elevado tem maior probabilidade de serem encontrados nos dados assumidos como *soft data*, e especialmente fazem parte da campanha de adensamento nas áreas de alto teor.

Portanto, esses dados receberam menor peso nas estimativas pois possuem maior erro de medida associado. Já os blocos abaixo de 500 ppm, estimados por KVME se aproximaram mais da referência justamente por conceder maior peso para as amostras com menor teor, *hard data*, as quais não possuem erro vinculado. Os pesos atribuídos para as amostras com teores intermediários foram similares, logo os blocos estimados por OK e KVME possuem concentrações semelhantes.

Para conferir a tendência dos teores estimados em relação ao modelo de referência, os *swath plots* das Figuras 3.21, 3.21 e 3.22 mostram o comportamento nas direções *X* e *Y* das médias estimadas por meio da KVME e da OK, e referência “*Rebloc\_Exhaust\_10x10*”. A análise foi realizada com faixas de 10 metros para o eixo *X* e para o eixo *Y*. Observa-se, que tanto a OK quanto a KVME suavizam o comportamento da média. Porém, verifica-se, principalmente na Figura 3.23, a qual apresenta a análise de deriva das estimativas referente aos dados com maior erro associado (*std.dv\_3*), que nos teores altos a OK se aproxima mais do teor verdadeiro, e superestima os teores baixos. Já a KVME subestima teores altos e se aproxima mais da referência nos teores baixos.

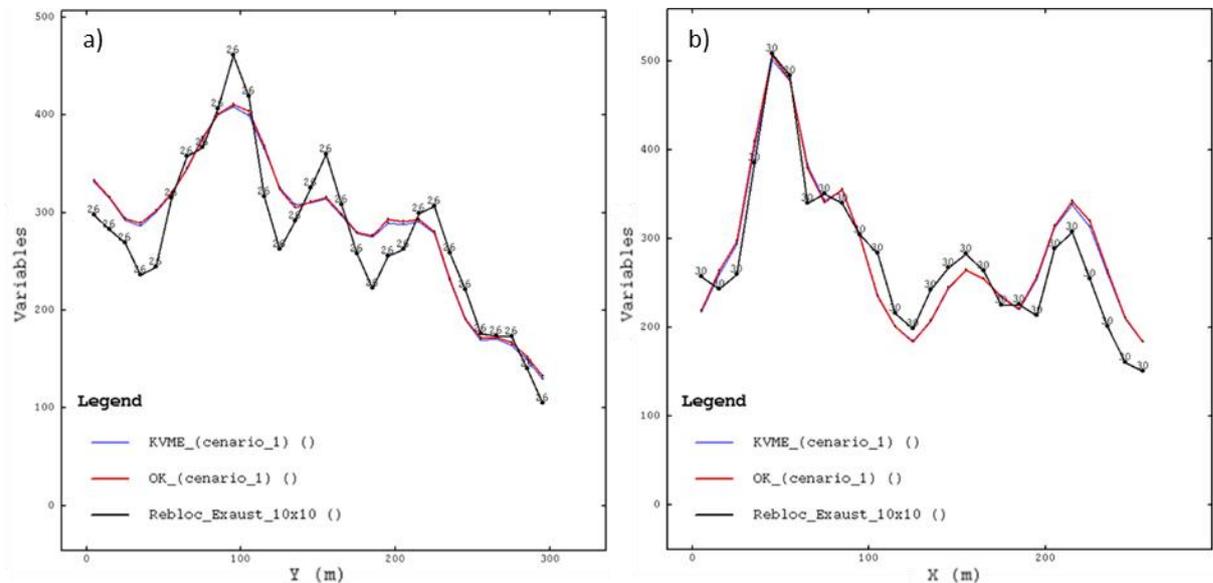


Figura 3.21 – Análise de deriva de um cenário pertencente ao erro *std.dev\_1*. a) Análise na direção *Y*. b) Análise na direção *X*.

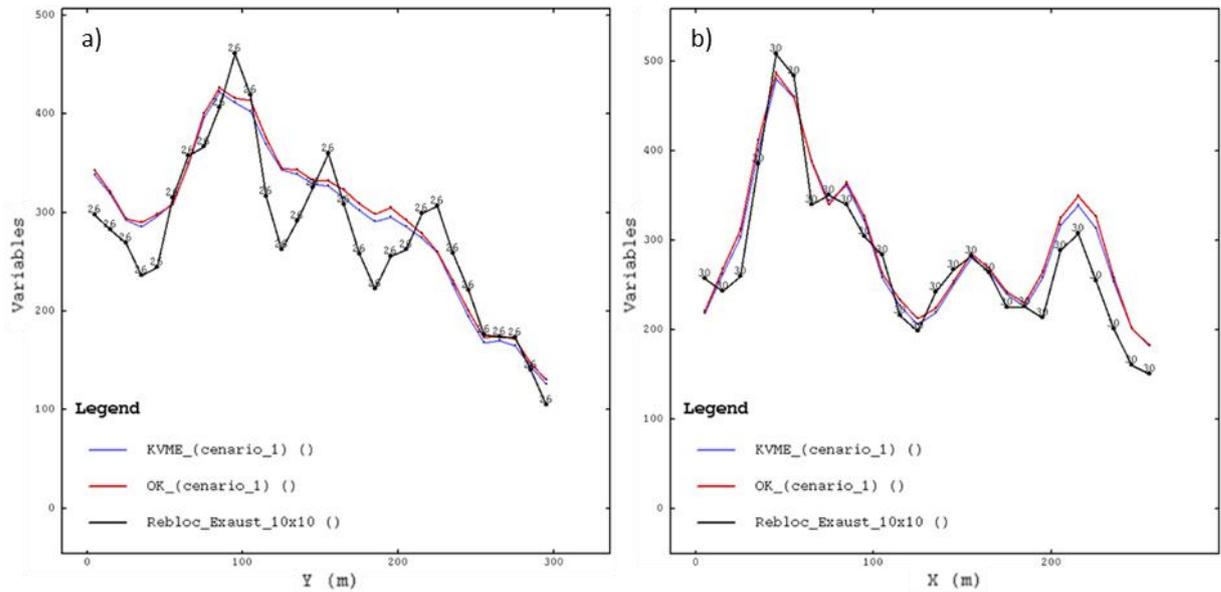


Figura 3.22 – Análise de deriva de um cenário pertencente ao erro  $std.dev_2$ . a) Análise na direção Y. b) Análise na direção X.

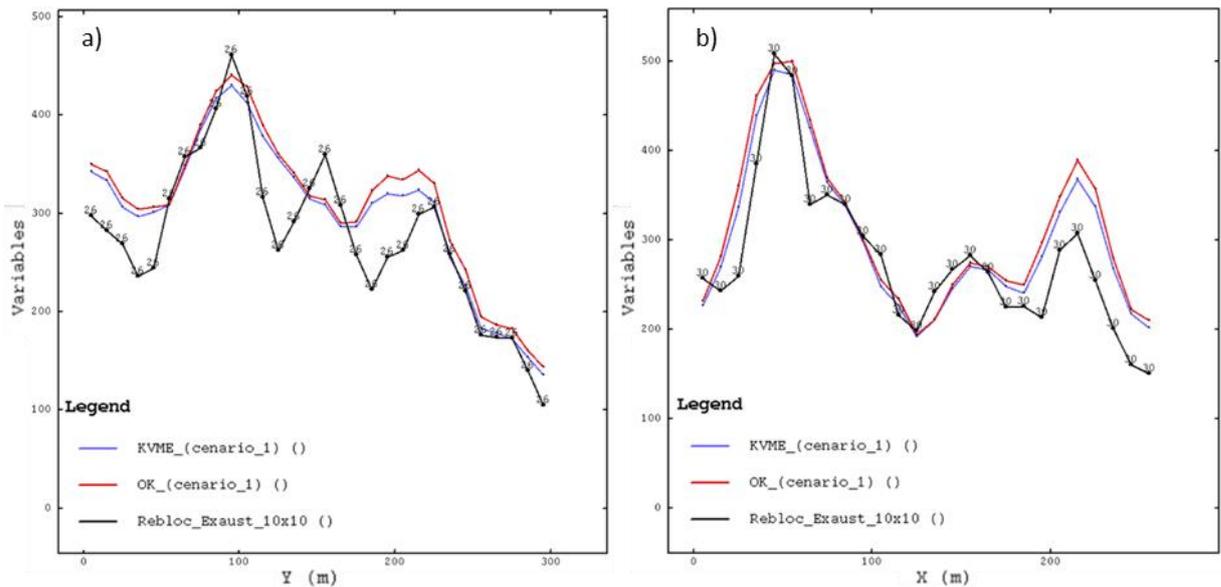


Figura 3.23 – Análise de deriva de um cenário pertencente ao erro  $std.dev_3$ . a) Análise na direção Y. b) Análise na direção X.

### 3.5 Discussões

Como visto nos resultados apresentados, as estimativas obtidas por OK e KVME são distintas devido a distribuição de pesos serem realizadas de diferentes formas. Ao simular os cenários nesse estudo de caso, foi adicionado erro apenas nos dados que tipicamente seriam sondados por métodos de adensamento ou *grade control*. Nessa disposição, existe uma relação entre dados e erros nos teores mais altos, fazendo com que a KVME adicionasse

menor peso nesses teores. Isso causou uma subestimativa nas zonas de teores altos, maior que na OK quando comparadas.

A decisão de empregar a OK ou KVME está relacionado a dois fatores. O primeiro é o grau de confiança em relação aos teores altos. Se as amostras com teores acima de 800 ppm tem respaldo geológico, como uma assembleia mineral condizente com teores extremamente altos, tais faixas de teores devem ser honradas. Por outro lado, é necessário cautela caso haja suspeita de que esses valores sejam simplesmente variações, causadas pelo erro associado ao processo de medição. Nesse caso, a estimativa por KVME oferece conservadorismo, pois atribui menor peso as amostras de alto teor, o que minimiza a probabilidade de erros amostrais superestimarem blocos no entorno dessas amostras.

O segundo fator é a precisão do método nos valores próximos ao teor de corte do depósito. Em um cenário onde a faixa de teores subestimados é distante do teor de corte e, portanto, os blocos continuam sendo considerados minério, isso não afetará a quantificação de reservas. Por outro lado, a maior acurácia das estimativas próximos ao teor de corte tem impacto positivo na classificação dos blocos como minério ou estéril.

Para ilustrar o segundo fator, a Figura 3.24 apresenta a curva de parametrização teor-tonelagem dos blocos de referência e das estimativas dos cenários com o erro *std.dev\_3*. Os maiores desvios da KVME em relação a distribuição real dos blocos ocorrem nos teores de corte acima de 380 ppm. A OK tem um comportamento inverso, com maiores desvios abaixo de 380 ppm. Logo, uma metodologia será mais adequada que a outra dependendo do *cutoff* escolhido.

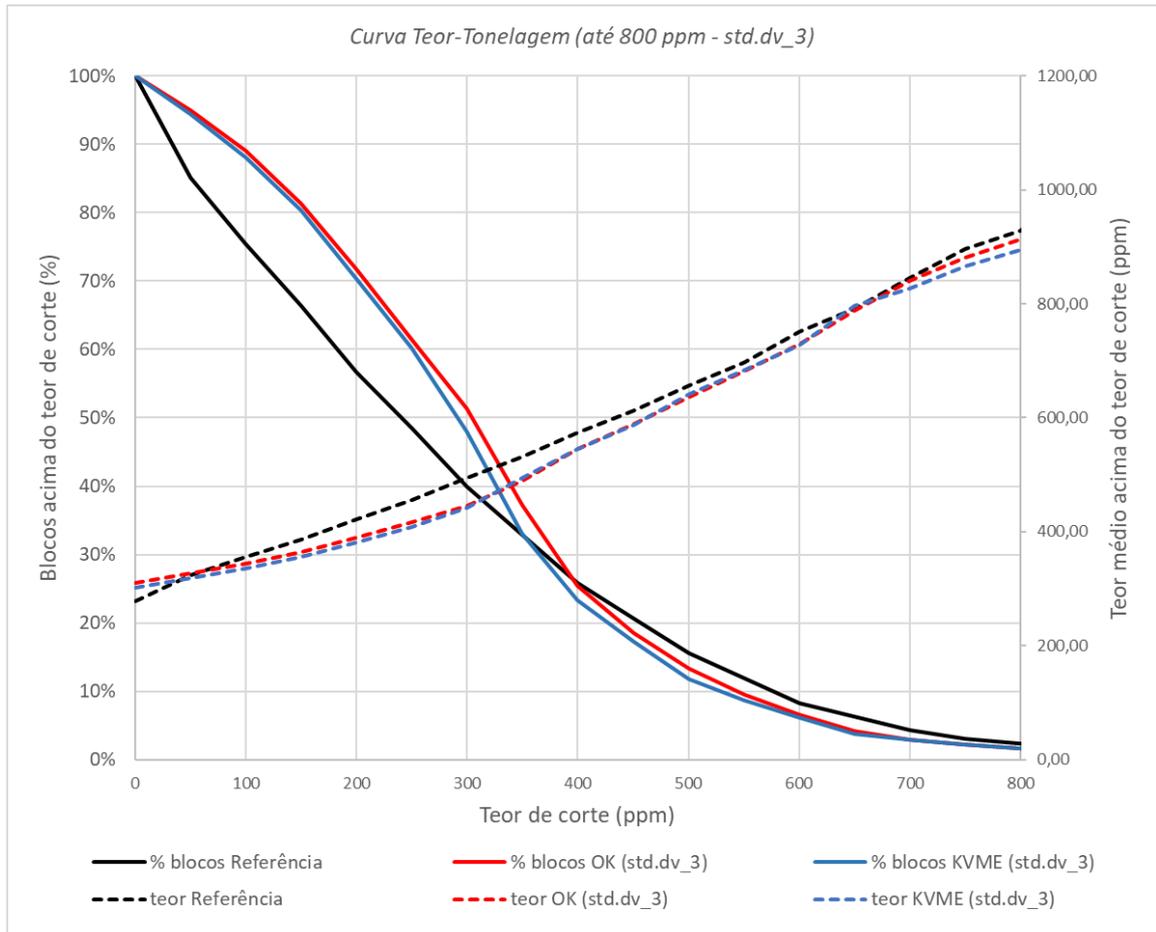


Figura 3.24 – Curva de parametrização teor x tonelagem. Curva obtida através da média dos cenários pertencente ao erro *std.dev\_3*.

### 3.6 Capping e KVME

Nesta seção, será analisada a utilização da KVME para limitar a influência dos *outliers*, como uma alternativa ao *capping*. Um valor *outlier* é caracterizado por uma concentração anormalmente alta da variável em questão, indicando ser um valor inconsistente com a grande maioria dos dados. *Capping* geralmente envolve a redução dos valores que são considerados *outliers* a um nível inferior, geralmente no percentil 95 a 99 da distribuição.

Ao perturbar os valores originais dos 275 *soft data*, vários *outliers* foram gerados devido à adição do erro aleatório. Valores que originalmente correspondiam a teores altos, transformaram-se em teores ainda mais elevados. Como visto anteriormente, a KVME distribui pesos de acordo com o erro associado a amostra. Dessa forma, considerando que nesse estudo de caso os altos teores estão relacionados ao maior erro, a KVME atribuiu menor peso para as amostras localizadas na cauda da distribuição. Ao avaliar não apenas o quão anômalo é um valor, mas o contexto onde ele se localiza, a KVME dialoga com as boas práticas

de analisar a vizinhança de teores de um valor inicialmente classificado como um *outlier*. Essa propriedade da KVME pode ser considerada uma opção mais coerente ao se trabalhar com valores extremos, pois um cluster de *outliers* pode indicar uma região de altos teores, no qual o corte dos valores extremos não será a melhor solução.

Para comparar as duas metodologias, primeiramente será determinado valores para o truncamento superior dos teores, e assim, proceder com o *capping*. Esses dados serão estimados por krigagem ordinária e, posteriormente, os resultados da OK serão comparados com as estimativas por KVME (sem *capping*). Assim, foram gerados gráficos de probabilidade dos cenários referentes ao erro *std.dev\_3*, o qual apresenta a maior quantidade de valores extremos. A Figura 3.26 expõe os gráficos gerados para dois cenários, onde o eixo *x* representa os valores das amostras, e em *y* as frequências cumulativas correspondentes. Pontos de inflexão e/ou lacunas na distribuição são geralmente direcionados como possíveis valores de *capping*. Nos gráficos abaixo, a distribuição segue uma forma linear de 15% a até 93-96%, e após esse ponto há uma quebra na linearidade. Assim, foram selecionados dois percentis para o *capping*: P95 e P99.

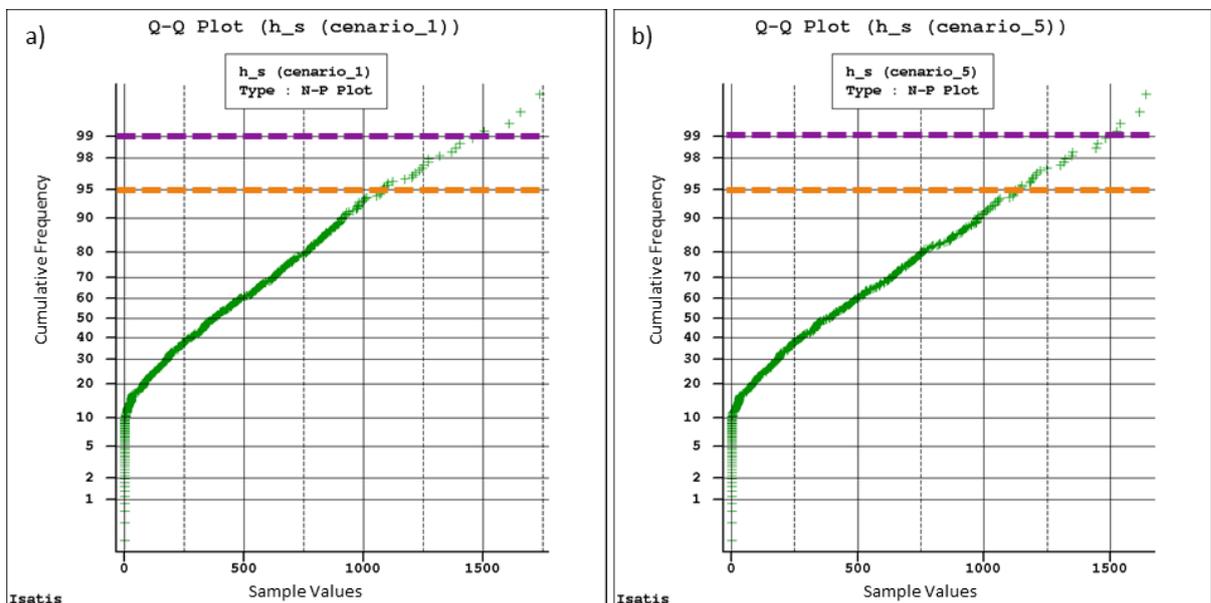


Figura 3.25 – Gráficos de probabilidade de dois cenários referentes ao erro *std.dev\_3*. Linha tracejada em laranja representa *capping* no P95, e em lilás P99.

A Tabela 3.7 apresenta os teores máximos de todos os cenários após a realização do *capping* P95 e P99. As médias, após a transformação dos valores extremos, se mantiveram próximas aos valores originais dos cenários, diminuindo 4% no *capping* P95, e 1% no *capping*

P99. O desvio padrão diminuiu 7% no *capping* P95 e 2% no *capping* P99. Dessa forma, como era esperado, o truncamento dos valores extremos para o P95 modificou mais intensamente o conjunto de dados, em comparação com a transformação para o P99.

Tabela 3.7 – Teores originais máximos nos cenários simulados (h\_s cenários), na referência (V original), e após a realização do *capping* em P99 e P95.

Cenário	Máximo (original)	Máximo (P99)	Máximo (P95)
V (original)	1528,1	-	-
h_s (cenário_1)	1737,87	1455,21	1080,97
h_s (cenário_2)	1803,26	1384,47	1108,91
h_s (cenário_3)	1748,26	1463,54	1119,51
h_s (cenário_4)	1940,34	1363,85	1064,76
h_s (cenário_5)	1641,07	1479,88	1123,68
h_s (cenário_6)	1803,75	1551,20	1080,69
h_s (cenário_7)	1695,93	1544,54	1088,09
h_s (cenário_8)	1757,81	1480,56	1070,06
h_s (cenário_9)	1837,10	1442,51	1021,91
h_s (cenário_10)	1584,56	1287,08	1056,21

O variograma utilizado para realizar a krigagem ordinária dos cenários após o *capping* será o mesmo calculado anteriormente, exposto na Tabela 3.4. O variograma foi mantido, pois apesar da transformação dos valores extremos, os dados resultantes apresentaram as mesmas estruturas. Os parâmetros para proceder com a krigagem ordinária também foram os mesmos utilizados anteriormente. Os resultados das estimativas são apresentados na Tabela 3.8, na qual foi obtida a média dos 10 cenários para uma análise global das krigagens. Como pode ser observado, a correlação entre a KVME e a referência é superior a OK onde foi aplicado o *capping*, tanto em P95 quanto em P99. O desvio quadrático médio mostra uma diferença menor da KVME e a variável V reblocada, em comparação com a OK. Verifica-se também que, quanto maior o *capping* aplicado aos dados (P95 > P99), mais as estimativas por OK se distanciam da referência.

Tabela 3.8 – Valores de referência e média dos cenários estimados por OK (P95 e P99) e KVME. Mínimo de todos os cenários estimados de 0 ppm.

Cenários	Máximo (ppm)	Média (ppm)	Desvio Padrão (ppm)	Coefficiente de Correlação	Erro Quadrático Médio (ppm)
V (reblockado)	1219,34	278,97	215,88	1	-
Cenários KVME ( <i>std.dev_3</i> )	1057,22	305,39	183,42	0,852	117,42
Cenários OK <i>capping</i> 99	1115,02	294,38	188,94	0,822	127,11
Cenários OK <i>capping</i> 95	1073,31	290,10	185,08	0,820	126,76

As Figuras 3.26 e 3.27 apresentam diagramas de dispersão, correlacionando os cenários estimados com o modelo de referência. O eixo  $x$  representa o banco de dados exaustivo reblockado, o eixo  $y$  as estimativas, e a linha de referência  $y = x$  representa a correlação perfeita entre o valor estimado e a referência. Os triângulos em amarelo simbolizam os blocos dos cenários estimados por krigagem ordinária após *capping*, e os círculos em azul os blocos estimados por krigagem com variância do erro de medida. Observe nos dois diagramas que, em geral, os blocos estimados por KVME apresentam uma maior aderência à linha de referência. A Figura 3.26 (*capping* P99) mostra, nos teores mais elevados, um comportamento muito semelhante entre as duas estimativas. Entretanto, a Figura 3.27 (*capping* P95) apresenta uma subestimativa muito mais intensa dos blocos estimados por OK, se comparados com a KVME.

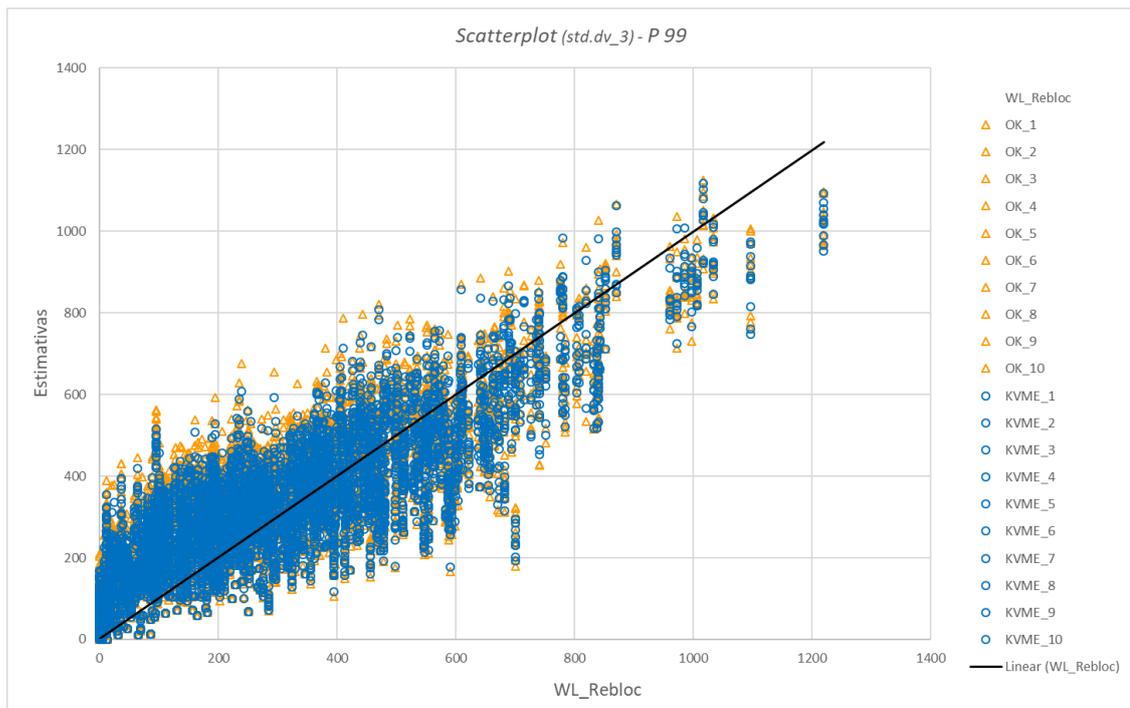


Figura 3.26 – Diagrama de dispersão: OK\_cenários versus referência (triângulos em amarelo), e KVME\_cenários versus referência (círculos em azul). *Capping* P99.



Tabela 3.9 – Erro quadrático médio dos cenários por faixa de teor. Células na cor verde representam desvio quadrático médio da KVME < OK; em amarelo desvio KVME = OK; em vermelho desvio KVME > OK.

Faixa de Teores (ppm)	Erro quadrático médio (média cenários - <i>std.dv_3</i> )		
	KVME	OK_Capping 99	OK_Capping 95
<100	113,82	126,59	125,79
>100; <200	124,10	138,87	137,28
>200; <300	101,85	111,17	109,63
>300; <400	90,85	98,34	96,74
>400; <500	121,51	124,18	120,62
>500; <600	150,73	150,45	148,89
>600; <700	145,74	145,48	145,71
>700; <800	162,56	164,35	163,90
>800; <900	147,72	147,24	151,44
>900; <1000	123,28	132,73	187,81
>1000; <1100	144,28	145,25	197,57
>1100	193,51	197,05	307,80

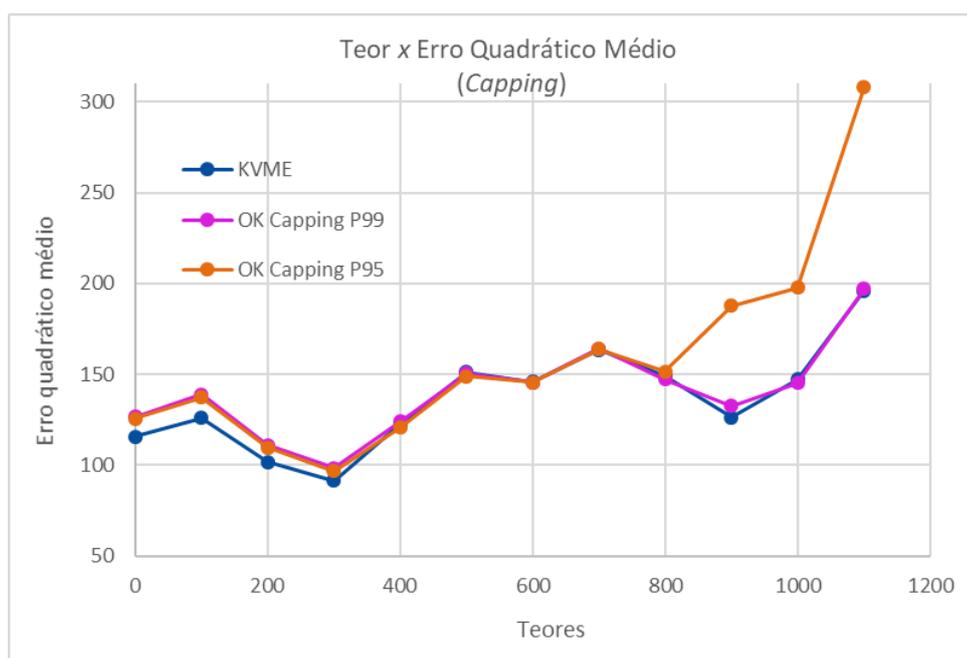


Figura 3.28 – Erro quadrático médio do cenários (*std.dv\_3*) por faixa de teor. Linha em azul representa os desvios da KVME, e linha em vermelho desvios da OK.

As diferenças nos resultados acontecem porque a KVME mantém o teor de todas as amostras, e distribui peso a estas conforme sua confiabilidade. Assim, dados *outliers* recebem menor peso nas estimativas pois, como esses teores estão associados a uma alta imprecisão, possuem maior variância do erro de medida. Já o truncamento dos dados, estabelece novos

valores aos *outliers*, muitas vezes a partir de uma decisão arbitrária, a qual influenciará nas estimativas. Isso pode ser constatado pelas diferenças nos resultados da OK dos dados em P95 e P99. Além disso, ao se aplicar *capping*, não é levado em consideração a distribuição local de probabilidades, e os valores que estão além do P95 ou P99 da distribuição, podem espacialmente estar relacionados a uma região de altos teores. Ao se utilizar a variância do erro de medida, obtida a partir do variograma, a decisão sobre a importância dada aos *outliers* é melhor embasada e consistente com as boas práticas de considerar a localização dos dados.

#### 4. ESTUDO DE CASO – 2

Este capítulo apresenta um segundo estudo de caso também realizado a partir do banco de dados Walker Lake (ISAAKS; SRIVASTAVA, 1989). A descrição detalhada da variável “V” (obtenção, localização das amostras, sumário estatístico), e a elaboração de um modelo de blocos, utilizado como cenário de referência para análise e comparação das estimativas, é apresentada na primeira seção do Capítulo 3 (Figura 3.1, 3.2 e 3.3).

##### 4.1 Cenários Hard – Soft

Os cenários deste estudo de caso simulam um depósito que possui dados obtidos em diferentes campanhas de sondagem. Será estabelecido como *hard data* os 275 pontos amostrais localizados na malha de aproximadamente 10 x 10 metros, e sua variância do erro de medida será constante de 0 ppm<sup>2</sup>. Esses dados representam amostras coletadas em uma campanha de sondagem mais recente, associada a uma maior precisão durante a etapa de coleta, preparação e análise das amostras. As outras 195 amostras, localizados na malha de 20 x 20 metros, serão designadas como *soft data*, e erros serão adicionados em seus teores medidos. Os dados *soft* mimetizam uma campanha de sondagem mais antiga, com menor controle na qualidade dos dados. A Figura 4.1 exibe a localização das amostras que foram definidas como *hard*, e aquelas que foram definidas como *soft*.

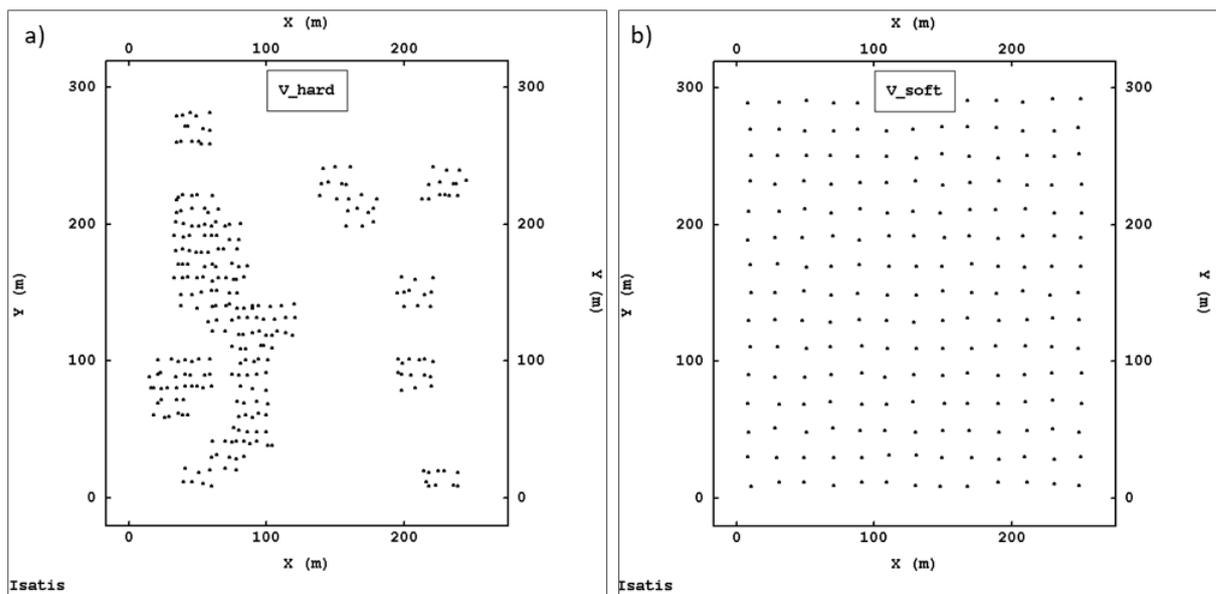


Figura 4.1 – a) V – *hard data*. b) V – *soft data*.

## 4.2 Adicionando imprecisão aos dados

Os valores lidos nos 195 *soft data* foram perturbadas por adição de um erro aleatório, assumindo uma distribuição gaussiana com média zero e desvio padrão determinado como uma porcentagem do desvio padrão do banco de dados original. Foram adotados três diferentes valores para esse desvio padrão, denominados de *std.dev\_1*, *std.dev\_2* e *std.dev\_3*, e para cada um deles foram simulados 10 cenários para minimizar o efeito de anomalias na simulação e garantir a consistência nos dados. Os distintos valores adotados como desvio padrão tiveram o objetivo de simular diferentes níveis de incerteza, dessa forma adotou-se  $std.dev_1 < std.dev_2 < std.dev_3$ .

Para ilustrar o ruído gerado nos teores, as Figuras 4.2 (a), 4.3 (a) e 4.4 (a) apresentam diagramas de dispersão, onde o eixo *x* representa as 195 amostras do banco de dados original de *V*, e o eixo *y* representa as 195 amostras de três diferentes cenários, referentes aos distintos níveis de incerteza (*std.dev\_1*, *std.dev\_2* e *std.dev\_3*). Com o aumento do erro, mais descorrelacionado ficam os dados simulados dos dados originais. Também se verifica que quanto maior o nível de incerteza, mais elevado é a quantidade de amostras com teor zero. Isso ocorre, porque todos aqueles valores que após a adição do erro tornaram-se negativos, foram transformados em zero, assegurando a coerência nos dados. As Figuras 4.2 (b), 4.3 (b) e 4.4 (b) mostram a distribuição espacial de três cenários, que são o conjunto das 195 amostras, *soft data*, obtidas a partir da adição de erros aos dados originais, junto dos 275 *hard data*, que não sofreram modificações.

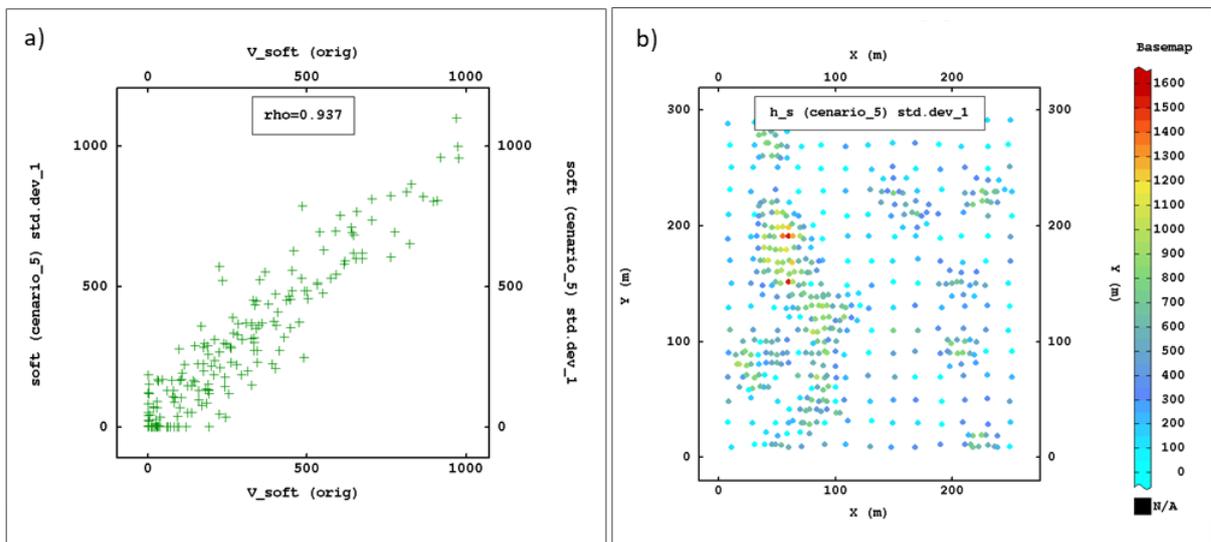


Figura 4.2 – a) Diagrama de dispersão dos 195 dados originais de *V* versus 195 dados do cenário 5 (*std.dev\_1*). b) Distribuição espacial do cenário\_5 (*std.dev\_1*).

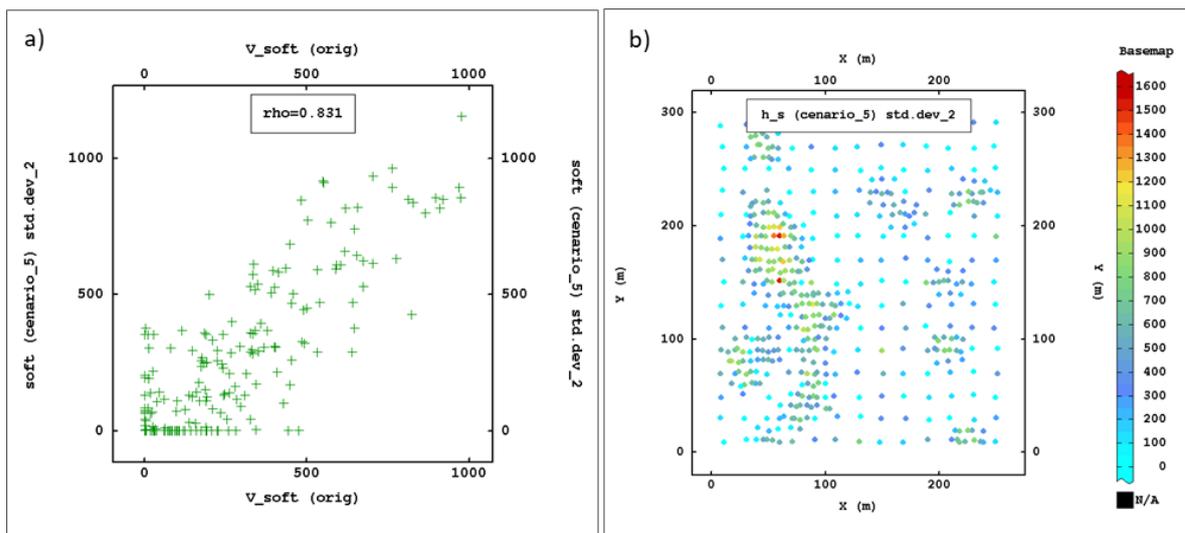


Figura 4.3 – a) Diagrama de dispersão dos 195 dados originais de V versus 195 dados do cenário 5 (*std.dev\_2*). b) Distribuição espacial do cenário\_5 (*std.dev\_2*).

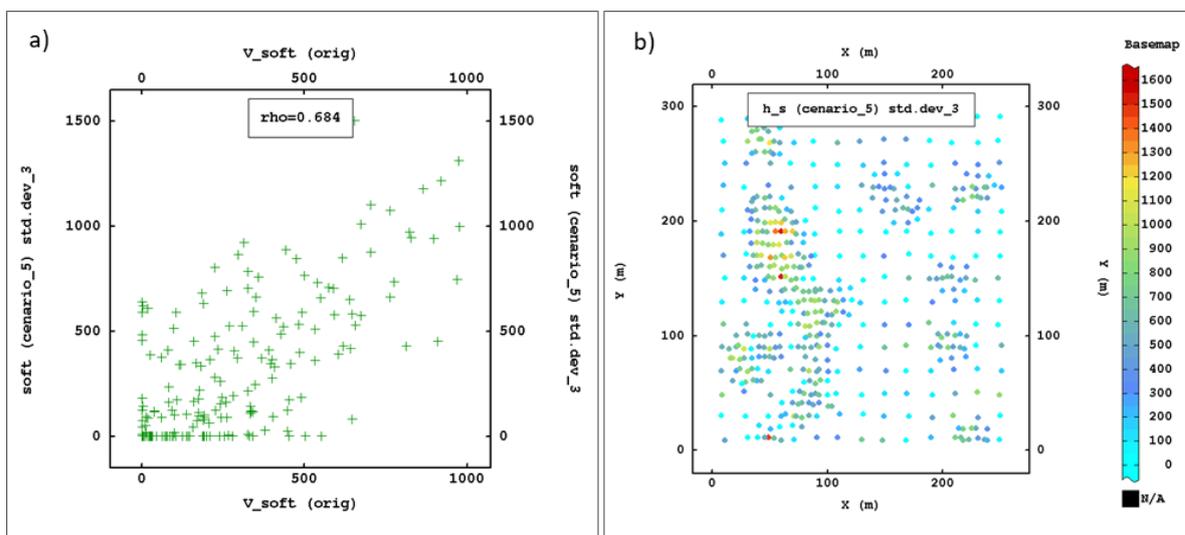


Figura 4.4 – a) Diagrama de dispersão dos 195 dados originais de V versus 195 dados do cenário 5 (*std.dev\_3*). b) Distribuição espacial do cenário\_5 (*std.dev\_3*).

A Tabela 4.1 apresenta um resumo estatístico dos bancos de dados resultantes dos cenários simulados. O valor mínimo de todos os cenários é igual a zero, e o valor máximo encontrado nos cenários é igual ao valor máximo do banco de dados original de 1528,1 ppm. A Tabela 4.1 mostra também que com o incremento do erro, mais descorrelacionado os dados simulados estão dos seus valores originais. Em geral, a média dos cenários se manteve próxima ao valor original de V. Entretanto, nos cenários com o erro *std.dev\_3*, a média aumentou em torno de 4% em relação ao valor original de V, pois a transformação de valores negativos em zero gerou um pequeno viés nos dados. Se o erro adicionado aos dados fosse maior, o viés em relação à média original de V aumentaria, e a utilização da KVME não seria

adequada em tal situação. Além disso, foi mantido a mesma proporção de erros (*std.dev\_1*, *std.dev\_2* e *std.dev\_3*) nos dois estudos de caso, para posterior comparação do comportamento da KVME nos diferentes tipos de erro.

Tabela 4.1 – Estatística descritiva dos cenários simulados em função do desvio padrão adicionado na geração dos erros.

Cenário	Coefficiente de Correlação	Média	Desvio Padrão
V (original)	1	435,3	299,56
Cen_1 (std.dev_1)	0,985	441,53	297,65
Cen_5 (std.dev_1)	0,981	438,38	300,80
Cen_10 (std.dev_1)	0,981	436,00	300,31
Média Cenários (std.dev_1)	0,980	439,96	300,32
Cen_1 (std.dev_2)	0,929	451,91	303,14
Cen_5 (std.dev_2)	0,947	431,66	311,67
Cen_10 (std.dev_2)	0,937	441,41	305,91
Média Cenários (std.dev_2)	0,936	444,17	306,88
Cen_1 (std.dev_3)	0,852	462,79	314,96
Cen_5 (std.dev_3)	0,878	445,22	325,39
Cen_10 (std.dev_3)	0,877	459,36	311,38
Média Cenários (std.dev_3)	0,869	455,76	319,34

### 4.3 Continuidade Espacial e Variância do Erro de Medida

O método adotado para encontrar a variância do erro de medida, também utiliza a diferença entre as estruturas modeladas nos variogramas dos dados *soft* e *hard*. Primeiramente, foi modelado um variograma omnidirecional dos 275 *hard data* (Figura 4.5). Como foi assumido que esses dados não possuem erro, ou seja, sua variância do erro de medida é de 0 ppm<sup>2</sup>, o variograma apresentará apenas o efeito pepita referente a variância em microescala ( $c_{MS}$ ). Posteriormente, o variograma omnidirecional de cada cenário foi modelado (*hard + soft data*), os quais mostram a mesma estrutura anterior,  $c_{MS}$ , incrementada por uma segunda componente do efeito pepita, correspondente à variância do erro de medida, ( $c_{ME}$ ). O valor encontrado de  $c_{ME}$  para cada cenário será utilizado como variância do erro de medida para os 195 *soft data*. Para definir o valor de  $c_{ME}$  a partir desta metodologia, foi assumido que os dados estão no mesmo domínio estacionário. A Figuras 4.6 apresenta os variogramas omnidirecionais de cenários referentes aos distintos níveis de incerteza (*std.dev\_1*, *std.dev\_2* e *std.dev\_3*), indicando os dois componentes do efeito pepita. Estes variogramas foram escolhidos aleatoriamente para exemplificação da metodologia.

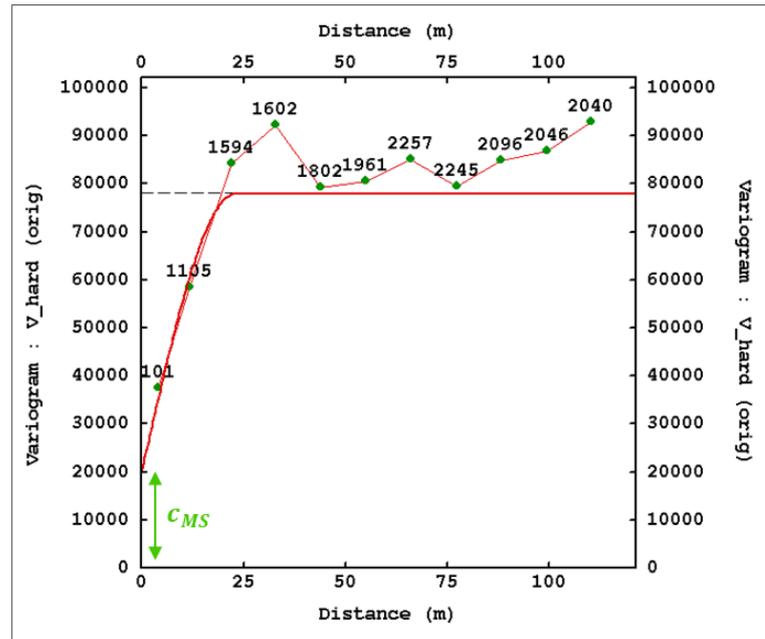


Figura 4.5 – Variograma omnidirecional *hard data* (275 amostras), mostrando o efeito pepita relativo a variância em microescala ( $c_{MS}$ ).

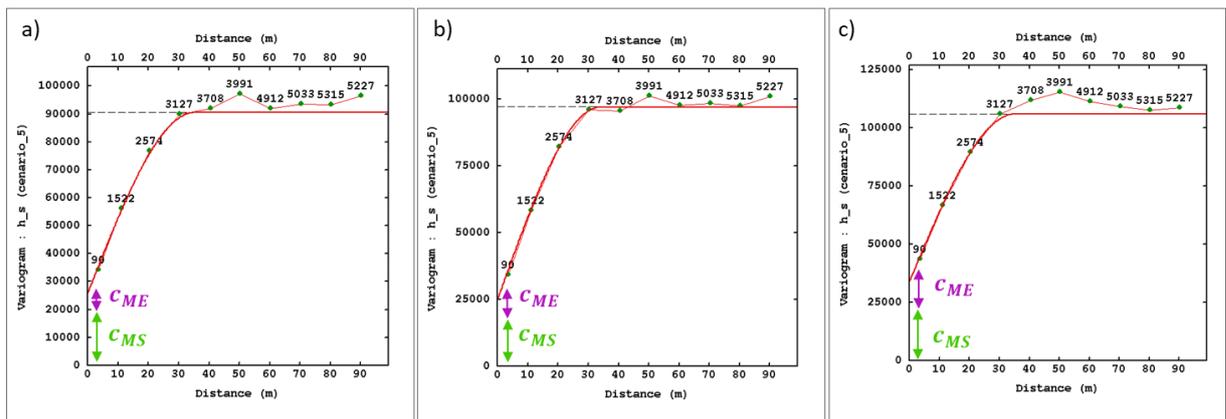


Figura 4.6 – Variogramas omnidirecionais de cenários referentes ao erro *std.dev\_1* (a), erro *std.dev\_2* (b) e ao erro *std.dev\_3* (c), indicando o efeito pepita relativo a variância em microescala ( $c_{MS}$ ) e relativo a variância do erro de medida ( $c_{ME}$ ).

As Tabelas 4.2, 4.3 e 4.4 apresentam os resultados obtidos com a modelagem do variograma utilizando apenas *hard data*, e dos variogramas omnidirecionais de todos os cenários (*hard data* + *soft data*). Com base nessa modelagem, foram determinados valores para a variância do erro de medida. Para realizar a krigagem ordinária, foi utilizado o variograma direcional médio dos dados *hard data* + *soft data*, apresentado no final das Tabelas 4.2, 4.3 e 4.4, sem distinção entre  $c_{ME}$  e  $c_{MS}$ . Já na KVME, foi mantido o valor de efeito pepita encontrado para os valores *hard data*, assumido como  $c_{MS}$ , e foi adicionado o valor médio calculado de  $c_{ME}$  como variância do erro de medida apenas no *input* dos dados.

Tabela 4.2 – Variogramas referentes aos cenários com erro *std.dev\_1*. (\*Valor adicionado como variância do erro de medida).

<b>std.dev_1</b>	<b><math>c_0</math></b>	<b><math>c_{MS}</math></b>	<b><math>c_{ME}</math></b>	<b>Contribuição (Sph)</b>	<b>Alcance</b>
variog omni hard data	20.000	20.000	0	58.000	23m
variog omni cenário_1	25.500	20.000	5.500	63.000	35m
variog omni cenário_2	26.000	20.000	6.000	63.000	33m
variog omni cenário_3	29.000	20.000	9.000	64.270	34m
variog omni cenário_4	26.000	20.000	6.000	66.000	35m
variog omni cenário_5	26.000	20.000	6.000	64.500	35m
variog omni cenário_6	26.000	20.000	6.000	64.500	34m
variog omni cenário_7	25.000	20.000	5.000	64.000	33m
variog omni cenário_8	24.000	20.000	4.000	65.500	34m
variog omni cenário_9	26.000	20.000	6.000	63.300	36m
variog omni cenário_10	25.500	20.000	5.500	64.700	36m
variograma dir médio	26.000	20.000	(6.000)*	$40.000 \left( \frac{N157^\circ}{45m}; \frac{N67^\circ}{25m} \right) + 24.300 \left( \frac{N157^\circ}{90m}; \frac{N67^\circ}{25m} \right)$	

Tabela 4.3 – Variogramas referentes aos cenários com erro *std.dev\_2*. (\*Valor adicionado como variância do erro de medida).

<b>std.dev_1</b>	<b><math>c_0</math></b>	<b><math>c_{MS}</math></b>	<b><math>c_{ME}</math></b>	<b>Contribuição (Sph)</b>	<b>Alcance</b>
variog omni hard data	20.000	20.000	0	58.000	23m
variog omni cenário_1	31.900	20.000	11.900	60.000	32m
variog omni cenário_2	29.000	20.000	9.000	65.400	35m
variog omni cenário_3	30.000	20.000	10.000	66.500	34m
variog omni cenário_4	32.000	20.000	12.000	64.400	35m
variog omni cenário_5	25.100	20.000	5.100	72.000	34m
variog omni cenário_6	25.200	20.000	5.200	69.000	33m
variog omni cenário_7	29.000	20.000	9.000	62.400	34m
variog omni cenário_8	29.200	20.000	9.200	66.000	34m
variog omni cenário_9	31.000	20.000	11.000	60.100	31m
variog omni cenário_10	30.000	20.000	10.000	63.600	35m
variograma dir médio	29.240	20.000	(9.240)*	$42.000 \left( \frac{N157^\circ}{52m}; \frac{N67^\circ}{25m} \right) + 23.000 \left( \frac{N157^\circ}{100m}; \frac{N67^\circ}{25m} \right)$	

Tabela 4.4 – Variogramas referentes aos cenários com erro *std.dev\_3*. (\*Valor adicionado como variância do erro de medida).

<b>std.dev_1</b>	<b><math>c_0</math></b>	<b><math>c_{MS}</math></b>	<b><math>c_{ME}</math></b>	<b>Contribuição (Sph)</b>	<b>Alcance</b>
variog omni hard data	20.000	20.000	0	58.000	23m
variog omni cenário_1	36.000	20.000	16.000	63.200	35m
variog omni cenário_2	35.000	20.000	15.000	64.350	34m
variog omni cenário_3	33.000	20.000	13.000	66.700	34m
variog omni cenário_4	42.000	20.000	22.000	65.760	36m
variog omni cenário_5	34.000	20.000	14.100	71.900	35m
variog omni cenário_6	44.000	20.000	24.000	64.110	34m
variog omni cenário_7	34.000	20.000	14.000	68.470	33m
variog omni cenário_8	33.000	20.000	13.000	68.200	33m
variog omni cenário_9	27.000	20.000	17.000	72.450	33m
variog omni cenário_10	42.000	20.000	22.000	54.800	33m
variograma dir médio	36.000	20.000	(16.000)*	$41.000 \left( \frac{N157^\circ}{50m}; \frac{N67^\circ}{20m} \right) + 23.000 \left( \frac{N157^\circ}{95m}; \frac{N67^\circ}{25m} \right)$	

No variograma omnidirecional dos dados *hard*, o efeito pepita modelado é de 20000 ppm<sup>2</sup> (Figura 4.5), e verifica-se que esse valor é o dobro daquele encontrado como efeito pepita no estudo de caso anterior (Figura 3.8). Isso se deve ao efeito proporcional, quando a variância local é correlacionada com a média local. Em geral, quando o efeito proporcional é combinado com o agrupamento de dados, ocorre uma superestimativa do efeito de pepita. Como os dados *hard* neste estudo de caso estão agrupados em locais onde a média é alta, o valor encontrado como efeito pepita, considerado apenas como variância de microescala ( $c_{MS}$ ), apresentou um alto valor (20000 ppm<sup>2</sup>). Logo, ao modelar os variogramas omnidirecionais dos cenários (*hard + soft*), e inferir a variância da medida do erro ( $c_{ME}$ ), foram encontrados valores relativamente menores para esse, quando comparados com o estudo de caso anterior.

#### 4.4 Estimativas

A estimativa foi realizada em suporte de blocos de dimensões 10 x 10 metros. O raio de busca aplicado foi anisotrópico, com a direção e distâncias obtidas na etapa de variografia. Foi utilizada a opção de restrição da busca por octantes, com um máximo de duas amostras por setor. A estimativa dos blocos foi discretizada em 5 x 5 x 1 pontos internos. Esses parâmetros apresentaram os melhores resultados considerando os critérios de validação das estimativas. Os mesmos parâmetros de busca por amostras e modelo de blocos foram utilizados nos 10 cenários de cada erro, assim a comparação das estimativas pode ser feita de maneira adequada.

As Figuras 4.7, 4.8 e 4.9 ilustram, cada uma com diferente nível de incerteza, o mesmo cenário estimado por KVME e OK, junto com a referência “Rebloc\_Exhaust\_10x10”. Todos os cenários estão com a mesma legenda de cores para auxiliar na comparação dos mapas. Visualmente, as estimativas são bastante semelhantes, com pequenas diferenças entre elas. Observa-se, que tanto a OK quanto a KVME estão suavizadas em comparação com o modelo de referência, e quanto maior o erro adicionado dados imputados aos cenários, maior a suavização.

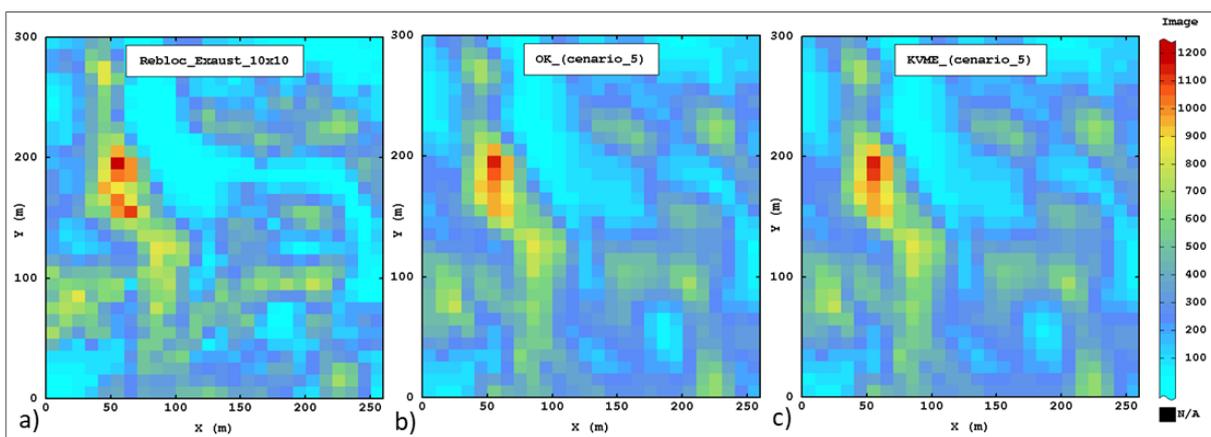


Figura 4.7 – a) Modelo de referência. b) Krigagem ordinária do cenário\_5 referente ao erro *std.dev\_1*. c) Krigagem com variância do erro de medida do cenário\_5 referente ao erro *std.dev\_1*.

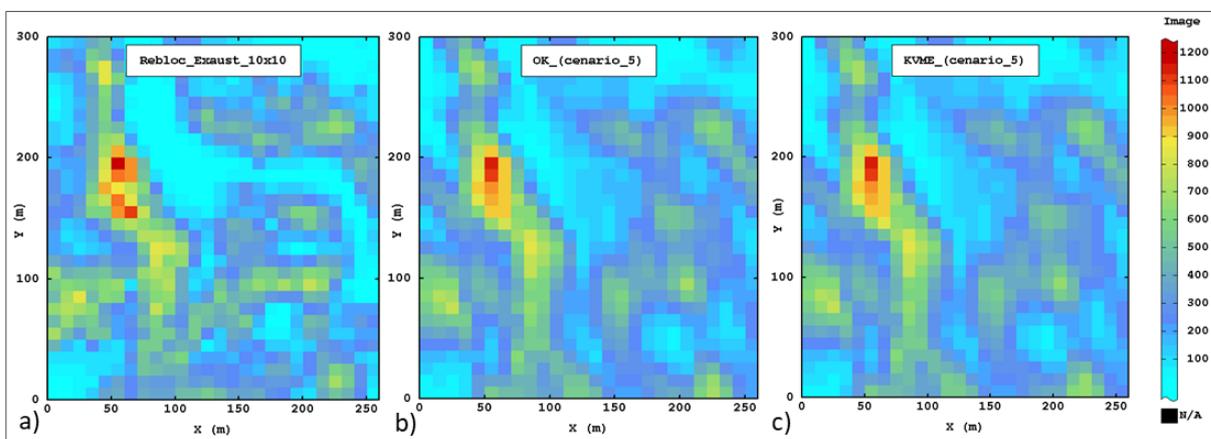


Figura 4.8 – a) Modelo de referência. b) Krigagem ordinária do cenário\_5 referente ao erro *std.dev\_2*. c) Krigagem com variância do erro de medida do cenário\_5 referente ao erro *std.dev\_2*.

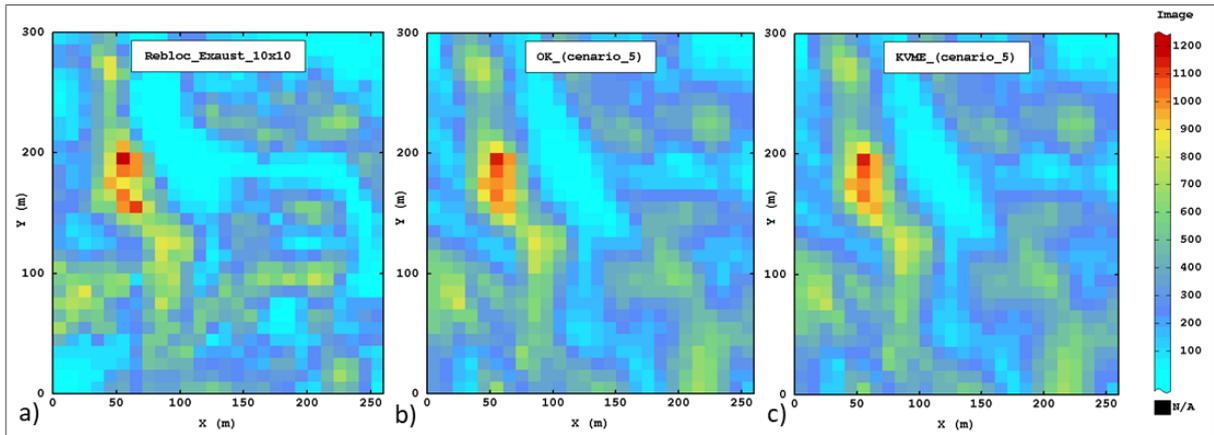


Figura 4.9 – a) Modelo de referência. b) Krigagem ordinária do cenário\_5 referente ao erro *std.dev\_3*. c) Krigagem com variância do erro de medida do cenário\_5 referente ao erro *std.dev\_3*.

Os resultados das estimativas são sumarizados na Tabela 4.5, na qual foi obtida a média dos 10 cenários referentes ao mesmo nível de erro para uma comparação global entre as krigagens. Com o aumento do erro nos teores, uma gradual redução do coeficiente de correlação entre os cenários e o modelo de referência acontece. Como o erro de medida adicionado nos cenários com erro *std.dev\_1* foi pequeno, as estimativas por OK e KVME são muito similares. Já nos cenários com os erros *std.dev\_2* e *std.dev\_3*, a correlação entre a KVME e a referência é superior a OK. O erro quadrático médio, o qual indica quanto os valores desviam-se do valor de referência, mostra uma diferença menor da KVME e a variável V reblocada, em comparação com a OK. Verifica-se também, que quanto maior o erro associado aos dados ( $std.dev_3 > std.dev_2 > std.dev_1$ ), mais as estimativas por KVME se sobressaem a OK.

Tabela 4.5 – Valores de referência e média dos cenários estimados por OK e KVME. Mínimo de todos os cenários estimados de 0 ppm.

Cenários	Máximo (ppm)	Média (ppm)	Desvio Padrão (ppm)	Coefficiente de Correlação	Erro Quadrático Médio (ppm)
V (reblockado)	1219,34	278,97	215,88	1	-
Cenários OK ( <i>std.dev_1</i> )	1156,07	293,99	189,74	0,883	102,71
Cenários KVME ( <i>std.dev_1</i> )	1156,77	294,92	188,89	0,883	103,02
Cenários OK ( <i>std.dev_2</i> )	1147,55	305,71	187,02	0,837	121,70
Cenários KVME ( <i>std.dev_2</i> )	1148,31	307,29	185,63	0,840	120,91
Cenários OK ( <i>std.dev_3</i> )	1136,43	328,46	189,59	0,761	151,10
Cenários KVME ( <i>std.dev_3</i> )	1138,63	330,22	186,72	0,775	147,12

Para representar visualmente a correlação dos cenários estimados com o modelo de referência, as Figuras 4.12, 4.13 e 4.14 apresentam diagramas de dispersão, no qual o eixo *x*

representa o banco de dados exaustivo reblockado, e o eixo y as estimativas. Os triângulos em vermelho simbolizam os blocos dos cenários estimados por krigagem ordinária, e os círculos em azul por krigagem com variância do erro de medida. A linha de referência  $y = x$  representa a correlação perfeita entre o valor estimado e a referência. Percebe-se que, quanto maior o nível de incerteza ( $std.dev_3 > std.dev_2 > std.dev_1$ ), maior a diferença entre o valor estimado e a referência, e maior a dispersão dos pontos em torno da reta.

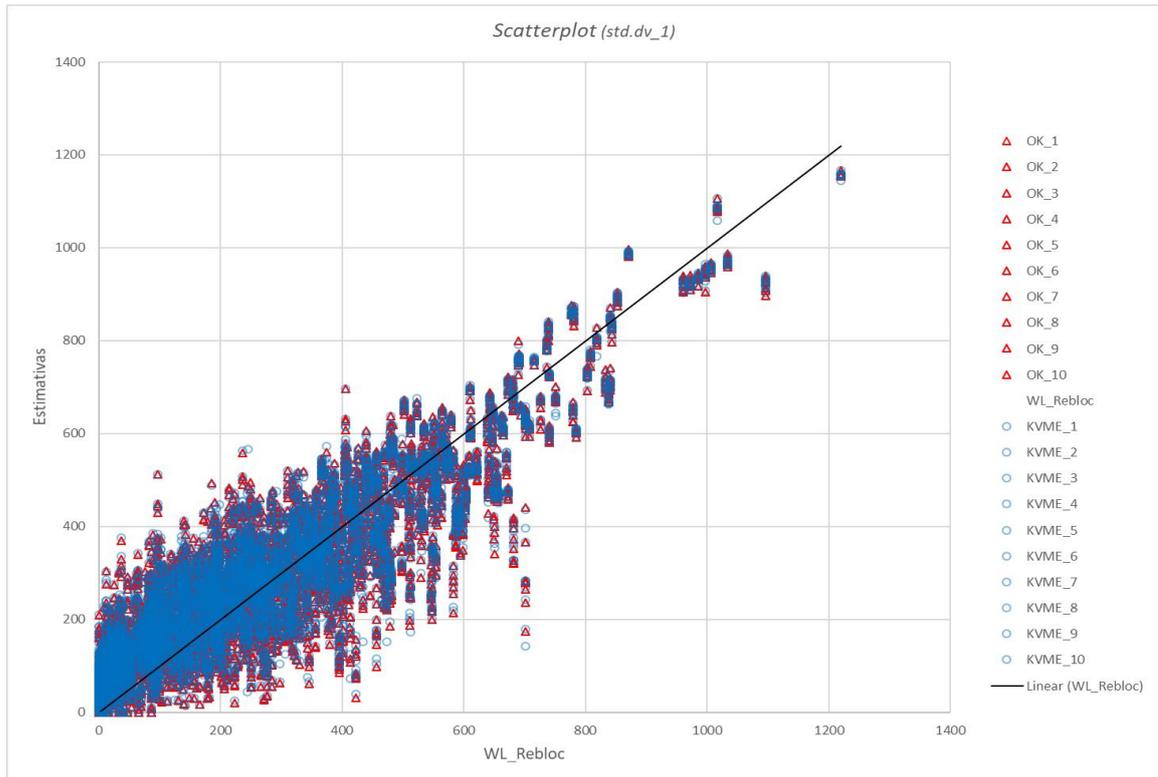


Figura 4.10 – Diagrama de dispersão: OK\_cenários versus referência (triângulos em vermelho), e KVME\_cenários versus referência (círculos em azul). Nível de incerteza  $std.dev_1$ .

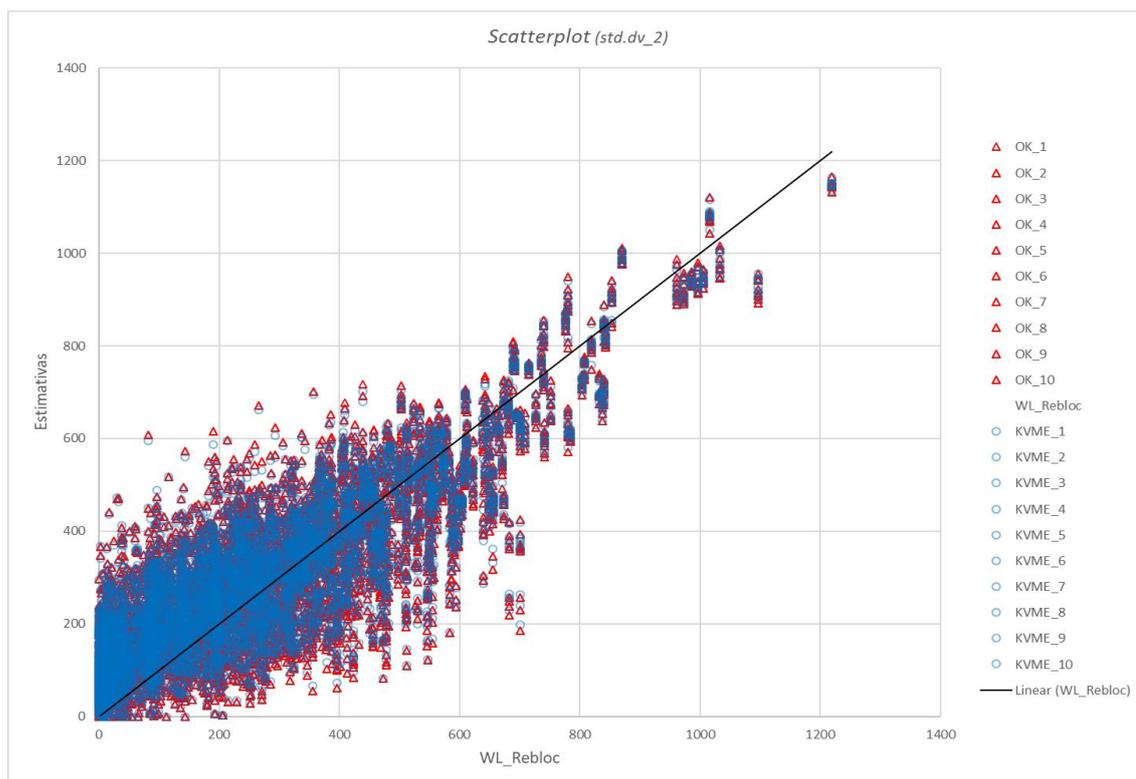


Figura 4.11 – Diagrama de dispersão: OK\_cenários *versus* referência (triângulos em vermelho), e KVME\_cenários *versus* referência (círculos em azul). Nível de incerteza *std.dev\_2*.

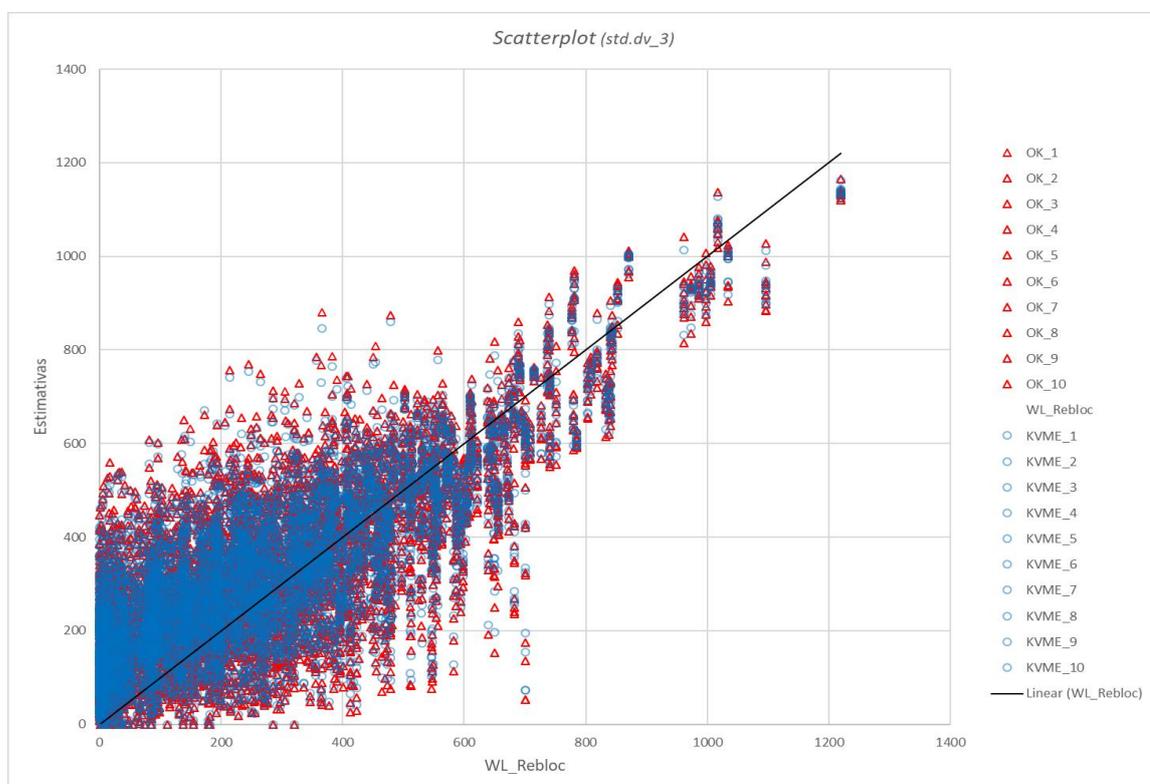


Figura 4.12 – Diagrama de dispersão: OK\_cenários *versus* referência (triângulos em vermelho), e KVME\_cenários *versus* referência (círculos em azul). Nível de incerteza *std.dev\_3*.

Verifica-se nos diagramas de dispersão, principalmente na Figura 4.12, que os blocos estimados por krigagem com variância do erro de medida mantêm seus teores mais próximos a referência, se comparadas às estimativas por krigagem ordinária. Esse comportamento permanece em quase toda a faixa de teores, com certa indefinição onde os teores são mais baixos e mais altos. Para compreender melhor esse comportamento, a Tabela 4.6 e a Figura 4.13 apresentam o erro quadrático médio por faixa de teor.

Tabela 4.6 – Erro quadrático médio dos cenários por faixa de teor. Células na cor verde representam erro quadrático médio da KVME < OK; em amarelo desvio KVME = OK; em vermelho desvio KVME > OK.

Faixa de Teores (ppm)	Erro quadrático médio (média cenários - <i>std.dv_1</i> )		Erro quadrático médio (média cenários - <i>std.dv_2</i> )		Erro quadrático médio (média cenários - <i>std.dv_3</i> )	
	OK	KVME	OK	KVME	OK	KVME
<100	94,81	97,12	129,10	130,77	179,53	179,54
>100; <200	96,85	98,66	119,91	120,85	152,02	152,12
>200; <300	93,51	92,89	110,74	108,12	138,58	133,54
>300; <400	96,76	94,73	107,95	103,74	130,58	123,87
>400; <500	117,21	115,02	122,01	117,36	136,25	128,51
>500; <600	127,66	126,78	140,09	137,25	141,84	137,00
>600; <700	123,64	122,08	133,38	130,16	137,78	132,43
>700; <800	141,73	140,82	137,25	135,01	151,94	148,02
>800; <900	90,73	91,25	96,14	97,18	100,30	100,03
>900; <1000	47,84	48,91	49,93	51,81	66,84	65,51
>1000; <1100	102,56	101,77	101,40	99,24	95,91	93,91
>1100	63,27	62,05	71,79	71,01	82,90	79,74

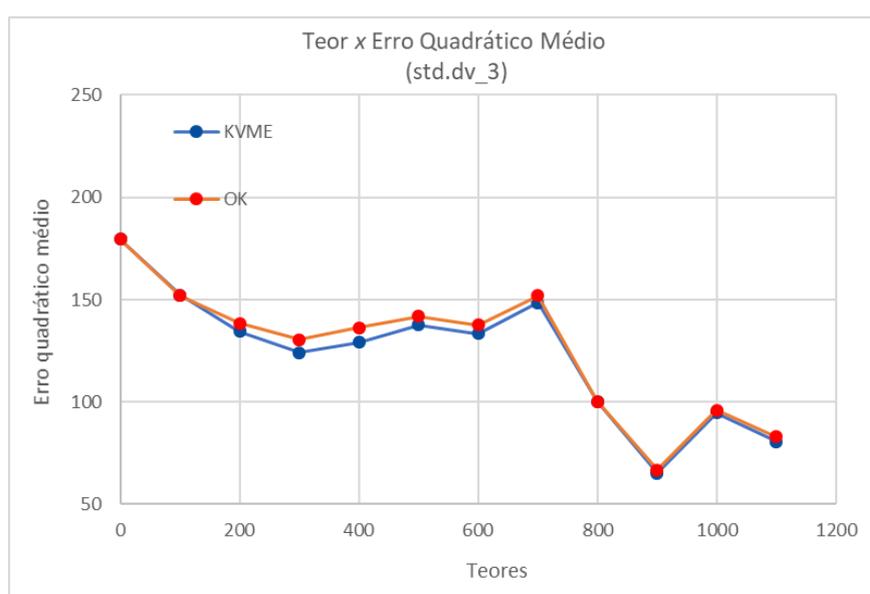


Figura 4.13 – Erro quadrático médio do cenários (*std.dv\_3*) por faixa de teor. Linha em azul representa os desvios da KVME, e linha em vermelho desvios da OK.

Conforme mostra a Tabela 4.6 e a Figura 4.13, de 0 a 200 ppm, o erro quadrático médio da OK é menor que a KVME, pois esta última superestima os teores mais baixos. Isto acontece porque a KVME dá menor peso às amostras de baixo teor, uma vez que estas estão espacialmente relacionadas com a primeira campanha de sondagem, associada a um maior erro nas análises. Sendo assim, a KVME transfere o peso às amostras de alto teor (e menor erro) que estejam próximos ao bloco estimado.

Nos teores acima de 200 ppm, a KVME apresenta menor erro quadrático médio em relação a OK, com valores de desvios muito semelhantes entre 800 e 1000 ppm. Nos teores maiores que 1000 ppm, observa-se que a OK subestima mais intensamente os blocos em relação a KVME. Isto ocorre, pois, ao estimar nas regiões com amostras agrupadas de alto teor, a OK distribui maior peso para as amostras de menor teor, pois elas têm menor redundância. Já a KVME transfere maior peso para as amostras com maior teor pois essas apresentam menor erro, o que resulta em uma melhor aproximação ao teor de referência. Comparando as estimativas, verifica-se também que, quanto maior o erro associado aos dados ( $std.dev_3 > std.dev_2 > std.dev_1$ ), menor é o erro quadrático médio da KVME em relação a referência, ou seja, mais as estimativas por KVME se sobressaem a OK.

Para conferir a tendência dos teores estimados em relação ao modelo de referência, os *swath plots* da Figura 4.14 mostram o comportamento nas direções  $X$  e  $Y$  das médias estimadas por meio da KVME e da OK, e da referência “*Rebloc\_Exhaust\_10x10*”. A análise foi realizada com faixas de 10 metros para o eixo  $X$  e para o eixo  $Y$ . É apresentado apenas a análise de deriva das estimativas referente aos dados com maior erro associado ( $std.dev_3$ ), pois nos cenários com menor erro, não foi possível observar diferenças no gráfico. Observa-se que, tanto a OK quanto a KVME estão suavizadas em relação a referência, e o comportamento das duas técnicas é muito semelhante. Entretanto, a diferença entre elas pode ser notada nas faixas de teores mais baixos, onde a OK está mais próxima a referência.

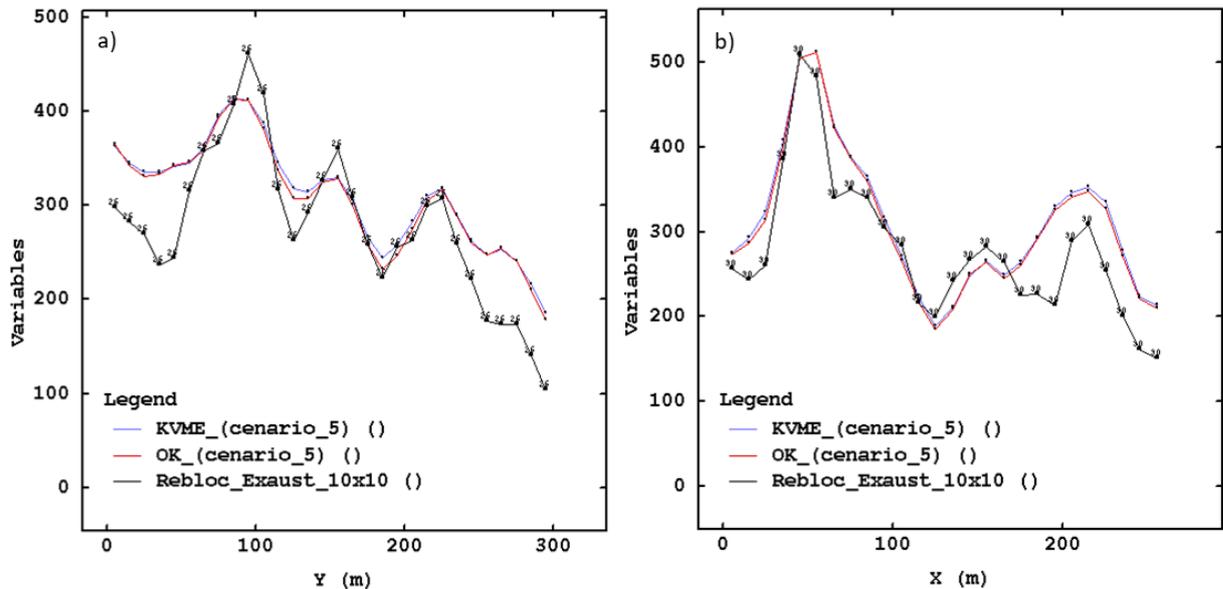


Figura 4.14 – Análise de deriva de um cenário pertencente ao erro *std.dev\_3*. a) Análise na direção Y. b) Análise na direção X.

#### 4.5 Discussões

Ao simular os cenários nesse estudo de caso, adicionou-se erro apenas nos dados localizados na malha regular de 20 x 20 metros, as quais representam uma primeira campanha de amostragem, com menor controle na qualidade dos dados. Nessa organização das amostras, existe uma relação entre dados e erros nos teores mais baixos. Dessa forma, ao realizar a KVME, foi distribuído menor peso para estes. Conforme observado nos resultados, as estimativas obtidas pela KVME estão próximas a referência nos teores altos, e apresentaram uma superestimativa nos teores mais baixos. Já as estimativas OK mostraram um comportamento contrário, aproximando-se a referência nos teores baixos, e sendo subestimadas nas regiões com teores altos.

A escolha de qual metodologia de estimativa deve ser empregada, em tal correlação das amostras com o erro, está associado ao teor de corte. A maior precisão das estimativas próximos ao teor de corte resultará na correta classificação dos blocos em minério e estéril. A Figura 4.15 apresenta a curva de parametrização teor-tonelagem dos blocos de referência e das estimativas dos cenários com o erro *std.dev\_3*. As duas metodologias apresentam resultados bastante parecidos entre si, conseqüentemente suas distribuições no gráfico são muito semelhantes. Porém, com auxílio da Tabela 4.6, observa-se que os maiores desvios da KVME em relação a distribuição real dos blocos ocorrem nos teores de corte abaixo de 250 ppm. A OK tem um comportamento inverso, com maiores desvios acima de 250 ppm.

Considerando que o valor utilizado como *cutoff* em um depósito geralmente não será em teores muito baixos, levando em consideração a lucratividade do empreendimento, em muitas situações a KVME, por se aproximar mais a realidade em teores intermediários e altos, será uma melhor alternativa à OK.

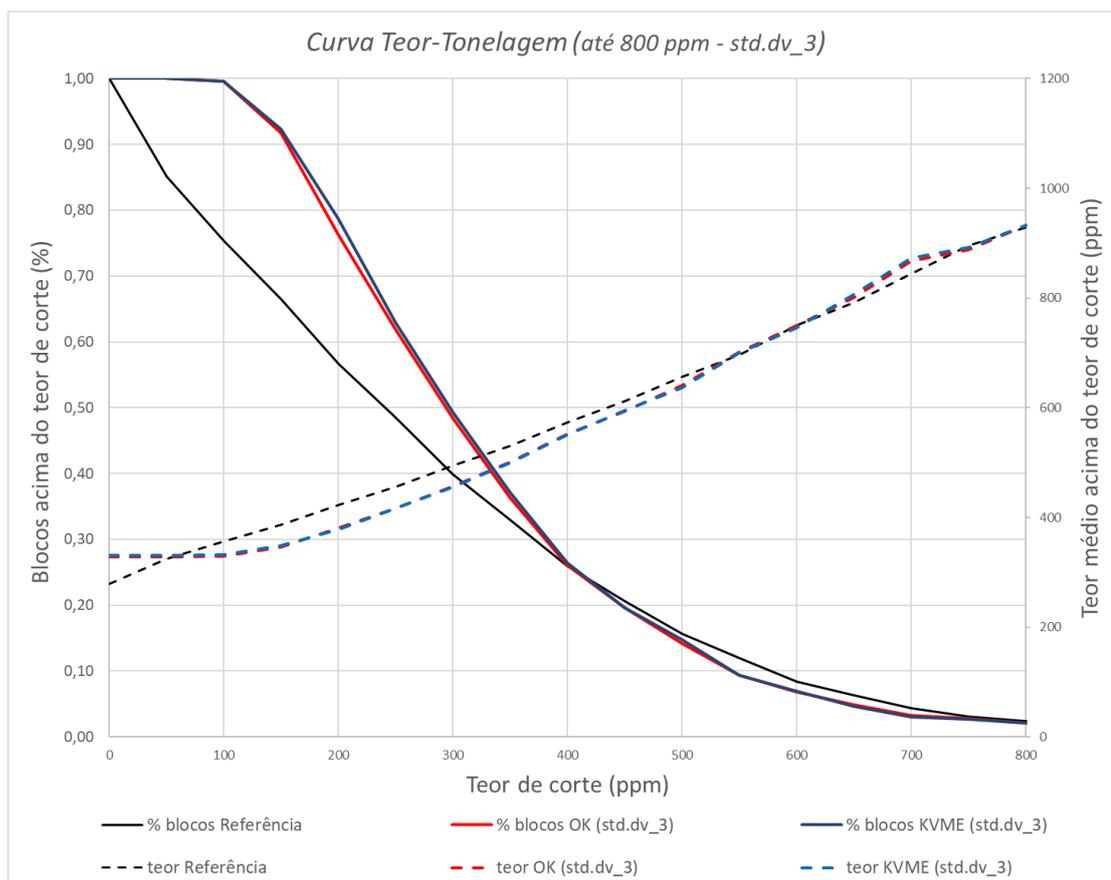


Figura 4.15 – Curva de parametrização teor x tonelagem. Curva obtida através da média dos cenários pertencente ao erro *std.dev\_3*.

## 5. CONCLUSÕES

O problema que motivou este estudo, é a limitação dos métodos geoestatísticos de estimativa para trabalhar simultaneamente com dados de diferentes graus de precisão. Como visto ao longo do trabalho, os bancos de dados podem ser compostos de dados obtidos por distintas técnicas de amostragem, e em diferentes momentos ao longo do projeto de mineração. Desta forma, é importante a pesquisa de métodos de estimativa que utilizem dados de diferentes fontes, extraindo informações destes de maneira adequada.

A KVME pondera o peso distribuído aos dados de acordo com a precisão individual de cada valor, característica essa diferente das metodologias clássicas que trabalham com diversas fontes de dados. Sendo assim, o principal objetivo deste trabalho foi investigar a metodologia “Krigagem com Variância do Erro de Medida, pois pouca bibliografia está disponível sobre o assunto, mesmo que ela tenha sido descrita pela primeira vez na década de 70. Desta forma, a meta central do trabalho foi realizada.

Fundamentado nas diferentes análises apresentadas nos resultados, conclui-se, de maneira geral, que a KVME forneceu uma melhoria na precisão das estimativas quando comparadas à OK. Essa conclusão é consistente com diversos trabalhos já publicados, os quais afirmam que dados imprecisos e/ou enviesados devem ser utilizados na estimativa dos recursos minerais, desde que integrados à estimativa de maneira adequada.

A metodologia assume que o erro deva ser independente do teor. Ao ignorar esse pressuposto, e aplicar a KVME quando existe uma relação entre erros e teores, concluiu-se que poderá ocorrer viés nas estimativas. Isto é, a KVME faz jus ao erro atribuído aos dados, dessa forma, dependendo da disposição espaciais destes, poderá acontecer enviesamento das estimativas nos teores mais altos ou nos teores mais baixos.

Ao testar a KVME como alternativa ao *capping*, essa se mostrou eficiente, uma vez que honra todas as amostras, mas distribui peso a essas conforme sua confiabilidade. Isso torna desnecessário o truncamento dos dados a partir de uma decisão arbitrária, a qual influenciará nas estimativas.

Ao longo desse trabalho também se concluiu que uma das limitações da KVME é de a obtenção da variância do erro de medida, e então proceder com a estimativa. Como na prática é improvável que a precisão de todas as medidas seja conhecida, a imprecisão dos dados precisa ser obtida de maneira indireta, não sendo uma atividade trivial. Opções consistentes

e mencionados no trabalho são através de cálculos estatísticos, baseado nos procedimentos de controle de qualidade, ou modelada pelos variogramas das distintas populações, método utilizado neste trabalho.

### 5.1 Trabalhos Futuros

Para trabalhos futuros, primeiramente recomenda-se que sejam realizados estudos de caso utilizando depósitos reais, investigando o comportamento da KVME em diferentes configurações espaciais de erro. Além disso, sugere-se buscar a integração das informações obtidas nos estudos de controle de qualidade destes depósitos reais, como opção para a obtenção da variância do erro de medida. Recomenda-se também aprofundar os estudos da performance da KVME comparado ao capping, e comparada às metodologias clássicas de estimativa com dados secundários, como a cokrigagem.

## REFERÊNCIAS

- ABZALOV, M. Sampling errors and control of assay data quality in exploration and mining geology. In: **Applications and Experiences of Quality Control**. <http://www.intechopen.com/books/applications-and-experiences-of-quality-control/sampling-errors-and-control-of-assay-data-quality-in-exploration-and-mining-geology>: IntechOpen, 2011. p. 611–644.
- ABZALOV, M. **Applied Mining Geology**. Cham: Springer, 2016.
- ALABERT, F. G. **Stochastic Imaging of Spatial Distributions Using Hard and Soft Information**. 1987. Stanford University, Stanford, 1987.
- CHAUTRU, J. M. Mapping with Auxiliary Data of Varying Accuracy. In: 2015, **Anais... : EAGE Petroleum Geostatistics**. Biarritz, 2015.
- CHILES, J.-P.; DELFINER, P. **Geostatistics: modeling spatial uncertainty**. New York: John Wiley & Sons, 1999.
- CHRISTENSEN, W. F. Filtered kriging for spatial data with heterogeneous measurement error variances. **Biometrics**, Washington, DC, v. 67, n. 3, p. 947–957, 2011.
- CORNAH, A.; MACHAKA, E. Integration of imprecise and biased data into Mineral resource estimates. **Journal of the Southern African Institute of Mining and Metallurgy**, Johannesburg, v. 115, n. 6, p. 523–530, 2015.
- COSTA, J. F. Reducing the impact of outliers in ore reserves estimation. **Mathematical Geology**, v. 35, n. 3, p. 323–345, 2003.
- CRESSIE, N. A. C. **Statistics for Spatial Data**. New York: John Wiley & Sons, 1991.
- CSA. **National Instruments 43-101. Standards of disclosure for mineral projects**. Canadian Securities Administrators., 2005.
- CUBA, M.; LEUANGTHONG, O.; ORTIZ, J. Transferring sampling errors into geostatistical modelling. **Journal of the Southern African Institute of Mining and Metallurgy**, Johannesburg, v. 112, n. 11, p. 971–983, 2012.
- DELHOMME, J. P. **Applications de la Théorie des Variables Régionalisées dans les Sciences de L'eau**. 1976. (Thèse de Docteur) - Université Pierre & Marie Curie, Paris, 1976.
- DERAISME, J.; STRYDOM, M. Estimation of iron ore resources integrating diamond and percussion drillholes. In: 37TH APPLICATION OF COMPUTERS AND OPERATIONS RESEARCH IN THE MINERAL INDUSTRY (APCOM) 2009, **Anais... : Proceedings of the 37th APCOM**. Vancouver, 2009.
- GEOVARIANCES. **Technical References - Isatis Software 2018**, Géovariances Fontainebleau, 2018.
- GOOVAERTS, P. **Geostatistics for Natural Resources Evaluation**. New York: Oxford University Press, 1997.
- GY, P. **Sampling of particulate materials, theory and practice, 2nd edition**. Amsterdam: Elsevier, 1982.

- ISAAKS, E. H.; SRIVASTAVA, M. R. **An Introduction to Applied Geostatistics**. New York: Oxford University Press, 1989.
- JORC. Australasian Code for Reporting of Exploration Results, Mineral Resources and Ore Reserves, AusIMM.2012.
- JOURNEL, A. G.; HUIJBREGTS, C. J. **Mining Geostatistics**. New York: Academic Press, 1978.
- KRIGE, D. G. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of Chemical Metallurgical & Mining Society of South Africa**, Johannesburg, v. 52, n. 6, p. 119–193, 1951.
- KRIVORUCHKO, K.; GRIBOV, A.; HOEF, J. M. A new method for handling the nugget effect in kriging. **Stochastic modeling and geostatistics: Principles, methods, and case studies, volume II**, 2006. p. 81–89.
- LEUANGTHONG, O.; NOWAK, M. Dealing with high-grade data in resource estimation. **Journal of the Southern African Institute of Mining and Metallurgy**, Johannesburg, v. 115, n. January, p. 27–36, 2015.
- MATHERON, G. Principles of Geostatistics. **Economic Geology**, Lancaster, v. 58, p. 1246–1266, 1963.
- MINNITT, R. C. A. Sampling: The impact on costs and decision making. **Journal of the Southern African Institute of Mining and Metallurgy**, Johannesburg, v. 107, n. 7, p. 451–462, 2007.
- MINNITT, R. C. A.; DEUTSCH, C. V. Cokriging for optimal mineral resource in mining operations. **Journal of the Southern African Institute of Mining and Metallurgy**, Johannesburg, v. 114, n. 3, p. 189–203, 2014.
- PARRISH, I. S. Geologist's Gordian knot: to cut or not to cut. **Mining Engineering**, New York, p. 45–49, 1997.
- PITARD, F. F. **Pierre Gy's sampling theory and sampling practice, 2nd edition**. New York: Press, 1993.
- PRESS, W. H. et al. **Numerical Recipes. The Art of Scientific Computing**. New York: Cambridge University Press, 2008.
- ROSCOE, W. E. Cutting curves for grade estimation and grade control in gold mines. In: 98TH ANNUAL GENERAL MEETING 1996, Edmonton, Alberta. **Anais...** Edmonton, Alberta: Canadian Institute of Mining, Metallurgy and Petroleum, 1996.
- ROSSI, M. E.; DEUTSCH, C. V. **Mineral Resource Estimation**. New York: Springer, 2014.
- SULLIVAN, J. Conditional Recovery Estimation Through Probability Kriging - Theory and Practice. In: VERLY, G. et al. (Eds.). **Geostatistics for Natural Resources Characterization**. Dordrecht: Springer, 1984. p. 365–384.
- THOMPSON, M.; HOWARTH, R. J. The rapid estimation and control of precision by duplicate determinations. **Analyst**, [s. l.], v. 98, n. 1164, p. 153–160, 1973.
- WACKERNAGEL, H. **Multivariate Geostatistics - An Introduction with Applications**. Berlin: Springer, 2003.