

MINISTÉRIO DA EDUCAÇÃO

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Escola de Engenharia

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

PPGEM

**A DESCORRELAÇÃO DE VARIÁVEIS COM FATORIZAÇÃO MAF EM
ESTIMATIVA DE TEORES**

Camilla Zacché da Silva

Dissertação para obtenção do título de Mestre em Engenharia

Porto Alegre – RS

2013

MINISTÉRIO DA EDUCAÇÃO

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Escola de Engenharia

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

PPGEM

**A DESCORRELAÇÃO DE VARIÁVEIS COM FATORIZAÇÃO MAF EM
ESTIMATIVAS DE TEORES**

Camilla Zacché da Silva

Bacharel em Física

Trabalho realizado no laboratório de pesquisa mineral e planejamento mineiro (LPM) da escola de engenharia da Universidade Federal do Rio Grande do Sul, dentro do programa de Pós-graduação em Engenharia de Minas, Metalúrgica e de Materiais (PPGEM), como parte dos requisitos para obtenção do título de Mestre em Engenharia.

Área de concentração: Tecnologia Mineral, Ambiental e Metalurgia Extrativa

Porto Alegre – RS

2013

Esta Dissertação foi julgada adequada para obtenção do título de Mestre em Engenharia, área de concentração de Metalurgia Extrativa e Tecnologia Mineral, e aprovada em sua forma final pelo Orientador e pela Banca Examinadora do Curso de Pós-Graduação.

Orientador:

Prof. Dr. João Felipe Coimbra Leite Costa

Banca Examinadora:

Dra. Maria Noel Morales Boezio

Prof. Dr. Jair Carlos Koppe

Dra. Vanessa Koppe

Prof. Dr. Telmo Roberto Strohaecker

Coordenador do PPGEM

“... Felizes são aqueles que buscam a proteção de Deus!”

Sl 2:12

À minha querida mãe.

AGRADECIMENTOS

Ao Prof. Dr. João Felipe, por me receber no LPM, pela orientação e incentivo, sempre depositando sua confiança em minha capacidade.

Ao Conselho Nacional de Pesquisa e Desenvolvimento tecnológico, pelo apoio financeiro.

À Dra. María Boezio, pelos materiais concedidos que enriqueceram este trabalho e fundamentaram meu conhecimento no tema abordado. Também pelo seu tempo dedicado e amizade.

À Dra. Vanessa Koppe, que, ao longo destes dois anos, sempre esteve disposta a ajudar e a compartilhar conhecimento.

Aos colegas Henrique Shibata, pelo dedicação e empenho no desenvolvimento do algoritmo apresentado nesta dissertação, e Diego Marques, pelas discussões esclarecedoras que tanto ajudaram no desenvolvimento deste trabalho.

Aos colegas do LPM, pelo companheirismo e pelas amizades que fiz. Em especial, à Sarah Guazzelli, Carla De Carli, Taís Câmara e Antonio Alves.

Ao Guilherme Cezaro, pelo apoio constante nesta caminhada.

À minha mãe, pelo amor incondicional e apoio desde o início de uma longa jornada.

A Deus, pela vida.

SUMÁRIO

LISTA DE FIGURAS	III
LISTA DE TABELAS	X
RESUMO	XII
ABSTRACT	XII
CAPÍTULO 1.....	1
INTRODUÇÃO	1
1.1 PROBLEMA	2
1.2 OBJETIVO	2
1.3 METAS	3
1.4 METODOLOGIA	3
1.5 ORGANIZAÇÃO DA DISSERTAÇÃO.....	4
CAPÍTULO 2.....	5
FUNDAMENTOS TEÓRICOS.....	5
2.1 ESTIMATIVAS DE DEPÓSITOS MULTIVARIADOS	5
2.1.1 COKRIGAGEM COLOCADA.....	6
2.1.2 MODELOS DE MARKOV	6
2.1.3 ASPECTOS PRÁTICOS.....	7
2.2 FUNDAMENTOS DE ÁLGEBRA LINEAR	8
2.2.1 TRANSFORMAÇÕES LINEARES	8
2.2.2 AUTOVETORES E AUTOVALORES.....	8
2.2.3 DECOMPOSIÇÃO ESPECTRAL.....	9
2.2.4 VETOR DE MÉDIAS E MATRIZ DE COVARIÂNCIA	10
2.3 ANÁLISE DE COMPONENTES PRINCIPAIS.....	12
2.4 MÉTODO DE MÍNIMOS-MÁXIMOS FATORES DE AUTOCORRELAÇÃO (MAF).....	17
CAPÍTULO 3.....	23
METODOLOGIA APLICADA	23

3.1	ALGORITMO.....	23
	CAPÍTULO 4.....	31
	ESTUDOS DE CASO.....	31
4.1	ESTUDO DE CASO 1 – BANCO DE DADOS JURA.....	31
4.1.1	ANÁLISE EXPLORATÓRIA DOS DADOS.....	31
4.1.2	SOLUÇÃO VIA COKRIGAGEM.....	40
4.1.3	SOLUÇÃO VIA KMAF.....	46
4.1.4	ANÁLISE DOS RESULTADOS – ESTUDO DE CASO JURA.....	60
4.2	ESTUDO DE CASO 2 - BANCO DE DADOS SAPECADO.....	80
4.2.1	APRESENTAÇÕES DO BANCO DE DADOS.....	80
4.2.2	ANÁLISE EXPLORATÓRIA DOS DADOS.....	83
4.2.3	SOLUÇÃO VIA COKRIGAGEM.....	88
4.2.4	SOLUÇÃO VIA KMAF.....	93
4.2.5	ANÁLISES DOS RESULTADOS - ESTUDO DE CASO SAPECADO.....	101
4.4	ANÁLISES COMPARATIVAS ENTRE AS METODOLOGIAS.....	114
	CAPÍTULO 5.....	117
5.1	CONCLUSÃO GERAL.....	117
5.1.1	CONCLUSÕES REFERENTES ÀS METAS PROPOSTAS.....	117
5.1.2	RECOMENDAÇÕES.....	118
	REFERÊNCIAS BIBLIOGRÁFICAS.....	119
	GLOSSÁRIO.....	122

LISTA DE FIGURAS

Figura 2. 1 - Gráfico de dispersão entre as componentes principais Z_1 e Z_2	16
Figura 4. 1 Mapa amostral do banco de dados Jura (Goovaerts, 1997) com os agrupamentos amostrais destacados.....	32
Figura 4. 2 - Histograma da variável Cádmio agrupado. Valores de concentração do metal em (ppm)	32
Figura 4. 3 - Histograma da variável Cádmio desagrupado. Concentração do metal em (ppm). Desagrupamento realizado por polígonos de Vornoi (Dirichlet, 1850)	33
Figura 4. 4 - Histograma da variável Cobalto agrupado. Concentração do metal em (ppm) ...	33
Figura 4. 5 - Histograma da variável Cobalto desagrupado. Concentração do metal em (ppm). Desagrupamento realizado por polígonos de Vornoi (Dirichlet, 1850).	34
Figura 4. 6 - Histograma agrupado da variável Cromo. Concentração do metal em (ppm)	34
Figura 4. 7 - Histograma da variável Cromo desagrupado. Concentração do metal em (ppm). Desagrupamento realizado por polígonos de Vornoi (Dirichlet, 1850)	35
Figura 4. 8 - Histograma agrupado da variável Níquel. Concentração do metal em (ppm).....	35
Figura 4. 9 - Histograma da variável Níquel desagrupado. Concentração do metal em (ppm). Desagrupamento realizado por polígonos de Vornoi (Dirichlet, 1850)	36
Figura 4. 10 - Gráfico de dispersão entre a variável Cádmio (ppm) e a variável Cobalto (ppm)	37
Figura 4. 11 - Gráfico de dispersão entre a variável Cádmio (ppm) e a variável Cromo (ppm)	38
Figura 4. 12 - Gráfico de dispersão entre a variável Cádmio (ppm) e Níquel (ppm).....	38
Figura 4. 13 - Gráfico de dispersão entre as variáveis, Cobalto (ppm) e Cromo (ppm)	39
Figura 4. 14 - Gráfico de dispersão entre as variáveis Cobalto (ppm) e Níquel (ppm).....	39
Figura 4. 15 - Gráfico de dispersão entre as variáveis, Cromo (ppm) e Níquel (ppm)	40
Figura 4. 16 - Variograma experimental do atributo Cobalto na direção N45; dip = 0. Alcance da primeira estrutura esférica é 700m. Alcance da segunda estrutura esférica é 3000m	41

Figura 4. 17 - Variograma experimental do atributo Cobalto na direção N135; dip = 0. Alcance da primeira estrutura esférica é 700m. Alcance da segunda estrutura esférica é 1200m	42
Figura 4. 18 - Variograma experimental cruzado das variáveis Cádmiio e Cobalto na direção N45; dip = 0. Alcance da primeira estrutura esférica é 700m. Alcance da segunda estrutura esférica é 3000m.....	43
Figura 4. 19 - Variograma experimental cruzado das variáveis Cádmiio e Cobalto na direção N135; dip = 0. Alcance da primeira estrutura esférica é 700m. Alcance da segunda estrutura esférica é 1200m.....	43
Figura 4. 20 - Gráfico de dispersão entre os fatores MAF1 e MAF2, obtidos por meio da transformação linear, realizada com o algoritmo exposto no capítulo 3	47
Figura 4. 21 – Correlograma experimental omnidirecional cruzado dos fatores MAF1 e MAF2	47
Figura 4. 22 - Gráfico de dispersão entre os fatores MAF1 e MAF3, obtidos por meio da transformação linear, realizada com o algoritmo exposto no capítulo 3	48
Figura 4. 23 - Correlograma experimental omnidirecional cruzado entre os fatores MAF1 e MAF3	48
Figura 4. 24 - Gráfico de dispersão entre os fatores MAF1 e MAF4, obtidos por meio da transformação linear, realizada com o algoritmo exposto no capítulo 3	49
Figura 4. 25 - Correlograma experimental cruzado omnidirecional entre os fatores MAF1 e MAF4	49
Figura 4. 26 - Gráfico de dispersão entre os fatores MAF2 e MAF3, obtidos por meio da transformação linear, realizada com o algoritmo exposto no capítulo 3	50
Figura 4. 27 - Correlograma experimental cruzado entre os fatores MAF2 e MAF3	50
Figura 4. 28 - Gráfico de dispersão entre os fatores MAF2 e MAF4, obtidos por meio da transformação linear realizada com o algoritmo exposto no capítulo 3	51
Figura 4. 29 - Correlograma experimental cruzado ominidirecional dos fatores MAF2 e MAF4	51
Figura 4. 30 - Gráfico de dispersão entre os fatores MAF3 e MAF4, obtidos por meio da transformação linear, realizada com o algoritmo exposto no capítulo 3	52
Figura 4. 31 - Correlograma experimental cruzando omnidirecional dos fatores MAF3 e MAF4	52
Figura 4. 32 - Gráfico de dispersão entre a variável Cádmiio (ppm) obtida por meio da retrotransformação dos fatores MAF ao espaço dos dados originais e à variável Cádmiio (ppm) original.....	53

Figura 4. 33 - Gráfico de dispersão entre a variável Cobalto (ppm) obtida por meio da retrotransformação dos fatores MAF ao espaço dos dados originais e à variável Cobalto (ppm) original.....	53
Figura 4. 34 - Gráfico de dispersão entre a variável Cromo (ppm) obtida por meio da retrotransformação dos fatores MAF ao espaço dos dados originais e à variável Cromo (ppm) original.....	54
Figura 4. 35 - Gráfico de dispersão entre a variável Níquel (ppm) obtida por meio da retrotransformação dos fatores MAF ao espaço dos dados originais e à variável Níquel (ppm) original.....	54
Figura 4. 36 - Variogramas direcionais do fator independente MAF1 (A) Direção de maior continuidade, N135 dip = 0; (B) direção de menor continuidade, N45 dip = 0.....	56
Figura 4. 37 - Variogramas direcionais do MAF2 (A) direção de maior continuidade, N22 dip = 0; (B) direção de menor continuidade, N112 dip = 0.....	57
Figura 4. 38 - Variogramas direcionais do MAF3 (a) direção de maior continuidade, N45 dip = 0; (b) direção de menor continuidade N135 dip = 0.....	58
Figura 4. 39 - Variogramas direcionais do MAF4 (a) direção de maior continuidade, N45 dip = 0; (b) direção de menor continuidade N135 dip = 0.....	59
Figura 4. 40 - Análise de deriva na direção X (Leste) do atributo Cádmiu (ppm). As faixas estão divididas em intervalos de 100m.....	63
Figura 4. 41 - Análise de deriva na direção Y (Norte) do atributo Cádmiu (ppm). As faixas estão divididas em intervalos de 100m.....	63
Figura 4. 42 - Análise de deriva na direção X (Leste) do atributo Cobalto (ppm). As faixas estão divididas em intervalos de 100m.....	64
Figura 4. 43 - Análise de deriva na direção Y (Norte) do atributo Cobalto (ppm). As faixas estão divididas em intervalos de 100m.....	64
Figura 4. 44 - Análise de deriva na direção X (Leste) do atributo Cromo (ppm). As faixas estão divididas em intervalos de 100m.....	65
Figura 4. 45 - Análise de deriva na direção Y (Norte) do atributo Cromo (ppm). As faixas estão divididas em intervalos de 100m.....	65
Figura 4. 46 – Análise de deriva na direção X (Leste) do atributo Níquel (ppm). As faixas estão divididas em intervalos 100m	66
Figura 4. 47 - Análise de deriva na direção Y (Norte) do atributo Níquel. As faixas estão divididas a cada 100m	66
Figura 4. 48- Gráfico de dispersão das variáveis cokrigadas Cádmiu (ppm) e Cobalto (ppm)	67

- Figura 4. 49 - Gráfico de dispersão das variáveis cokrigadas Cádmio (ppm) e Cromo (ppm) 68
- Figura 4. 50 - Gráfico de dispersão das variáveis cokrigadas Cádmio (ppm) e Níquel (ppm) 68
- Figura 4. 51 - Gráfico de dispersão das variáveis cokrigadas Cobalto (ppm) e Cromo (ppm) 69
- Figura 4. 52 - Gráfico de dispersão das variáveis cokrigadas Cobalto (ppm) e Níquel (ppm) 70
- Figura 4. 53 - Gráfico de dispersão das variáveis cokrigadas Cromo (ppm) e Níquel (ppm)..70
- Figura 4. 54 - Gráfico de dispersão entre as variáveis krigadas Cádmio (ppm) e Cobalto (ppm) utilizando a KMAF.....71
- Figura 4. 55 - Gráfico de dispersão entre as variáveis krigadas Cádmio (ppm) e Cromo (ppm) utilizando a KMAF.....72
- Figura 4. 56 - Gráfico de dispersão entre as variáveis krigadas Cádmio (ppm) e Níquel (ppm) utilizando a KMAF.....72
- Figura 4. 57 - Gráfico de dispersão entre as variáveis krigadas Cobalto (ppm) e Cromo (ppm) utilizando a KMAF.....73
- Figura 4. 58 - Gráfico de dispersão entre as variáveis krigadas Cobalto (ppm) e Níquel (ppm) utilizando a KMAF.....74
- Figura 4. 59 - Gráfico de dispersão entre as variáveis krigadas Cromo (ppm) e Níquel (ppm) utilizando a KMAF.....74
- Figura 4. 60 – Gráficos de dispersão entre as variáveis cokrigadas e as variáveis estimadas por meio de KMAF. (A) Cádmio (ppm) cokrigado *versus* Cádmio (ppm) estimado por meio de KMAF; (B) Cromo (ppm) cokrigado *versus* Cromo (ppm) estimado por meio de KMAF75
- Figura 4. 61 - Gráficos de dispersão entre as variáveis cokrigadas e as variáveis estimadas por meio de KMAF. (A) Cobalto (ppm) cokrigado *versus* Cobalto (ppm) estimado por meio de KMAF; (B) Níquel (ppm) cokrigado *versus* Níquel (ppm) estimado por meio de KMAF.....76
- Figura 4. 62 – Mapas para validação visual dos processos KMAF e Cokrigagem. A escala de cores na figura corresponde ao valor dos teores de Cádmio em (ppm), tanto no mapa de localização quanto nos mapas de krigagem e cokrigagem77
- Figura 4. 63 - Mapas para validação visual dos processos KMAF e Cokrigagem. A escala de cores na figura corresponde ao valor dos teores de Cobalto em (ppm), tanto no mapa de localização quanto nos mapas de krigagem e cokrigagem78
- Figura 4. 64 - Mapas para validação visual dos processos KMAF e Cokrigagem. A escala de cores na figura corresponde ao valor dos teores de Cromo em (ppm), tanto no mapa de localização quanto nos mapas de krigagem e cokrigagem79

- Figura 4. 65 - Mapas para validação visual dos processos KMAF e Cokrigagem. A escala de cores na figura corresponde ao valor dos teores de Níquel em (ppm), tanto no mapa de localização quanto nos mapas de krigagem e cokrigagem 80
- Figura 4. 66 - Vista no plano XY do modelo do Complexo Pico na direção principal do corpo de minério (Spier *et al.*, 2004)..... 82
- Figura 4. 67 - Mapas amostrais da variável Ferro global (FET). Mapa superior: configuração amostral no plano XY (Norte e Leste); mapa central configuração das amostras no plano ZY (Vertical e Norte); mapa inferior: configuração amostral no plano ZX (Vertical e Leste). Escala de cores representa teor de Ferro em (%)..... 83
- Figura 4. 68 - Mapas amostrais da variável Sílica global (SIT). Mapa superior: configuração amostral no plano XY (Norte e Leste); mapa central configuração das amostras no plano ZY (Vertical e Norte); mapa inferior: configuração amostral no plano ZX (Vertical e Leste). Escala de cores representa teor de Sílica em (%) 84
- Figura 4. 69 – Histograma da variável Ferro (FET) agrupado. Teor de Ferro expresso em (%) 85
- Figura 4. 70 - Histograma da variável Ferro (FET) desagrupado. Teor de Ferro expresso em (%) 85
- Figura 4. 71 - Histograma da variável Sílica (SIT) agrupado. Teor de Sílica expresso em (%) 86
- Figura 4. 72 - Histograma da variável Sílica (SIT) desagrupado. Teor de Sílica expresso em (%) 86
- Figura 4. 73 - Gráfico de dispersão entre Ferro global e Sílica global, coeficiente de correlação -0.98 87
- Figura 4. 74 – Variogramas experimentais direcionais da variável Ferro (FET). (A) Variograma na direção de maior continuidade espacial, N112, **dip** = 0; (B) Variograma na direção de continuidade espacial intermediária, N22, **dip** = 0; (C) Variograma na direção de menor continuidade espacial, Vertical..... 89
- Figura 4. 75 - Gráfico de dispersão entre os fatores de autocorrelação obtidos por meio da transformação linear dos dados originais, MAF1 e MAF2 94
- Figura 4. 76 - Correlograma experimental omnidirecional dos fatores MAF1 e MAF2 obtidos por meio da transformação MAF aplicada ao banco de dados Sapecado 95
- Figura 4. 77 - Correlograma experimental omnidirecional vertical dos fatores MAF1 e MAF2 obtidos por meio da transformação MAF aplicada ao banco de dados Sapecado..... 95
- Figura 4. 78 - Gráfico de dispersão de teste entre a variável FET (%) original, e a retrotransformação do fator MAF ao FET (%)..... 96

Figura 4. 79 - Gráfico de dispersão de teste entre a variável SIT (%) original, e a retrotransformação do fator MAF à SIT (%).....	96
Figura 4. 80 – Variogramas experimentais direcionais diretos do fator MAF1. (A) Variograma na direção de maior continuidade espacial, N112 dip = 0; (B) Variograma na direção de continuidade espacial intermediária, N22 dip = 0; (C) Variograma na direção de menor continuidade espacial, vertical.....	98
Figura 4. 81 – Variogramas experimentais direcionais diretos do fator MAF2. (A) Variograma na direção de maior continuidade espacial, N112 dip = 0; (B) Variograma na direção de continuidade espacial intermediária, N22 dip = 0; (C) Variograma na direção de menor continuidade espacial, vertical.....	99
Figura 4. 82 - Análise de deriva na direção X (Leste) para a variável FET (%). Faixas divididas em intervalos de 50m.....	102
Figura 4. 83 - Análise de deriva na direção Y (Norte) para a variável FET (%). Faixas divididas em intervalos de 50m.....	103
Figura 4. 84 - Análise de deriva na direção Z(vertical) para a variável FET (%). Faixas divididas em intervalos de 50m.....	103
Figura 4. 85 - Análise de deriva na direção X (Leste) para a variável SIT (%). Faixas divididas em intervalos de 50m.....	104
Figura 4. 86 - Análise de deriva na direção Y (Norte) para a variável SIT (%). Faixas divididas em intervalos de 50m.....	104
Figura 4. 87 - Análise de deriva na direção Y (Norte) para a variável SIT (%). Faixas divididas em intervalos de 50m.....	105
Figura 4. 88 – Gráfico de dispersão entre as variáveis cokrigadas FET (%) e SIT (%).....	105
Figura 4. 89 - Gráfico de dispersão entre as variáveis estimadas por meio de KMAF, FET (%) e SIT (%)	106
Figura 4. 90 – Gráfico de dispersão entre a variável cokrigada FET (%) e a variável estimada por meio de KMAF FET(%)	107
Figura 4. 91 – Gráfico de dispersão entre a variável cokrigada SIT (%) e a variável estimada por meio de KMAF SIT(%)	107
Figura 4. 92 – Vista no plano XY (leste e norte), nível 25 dos mapas de coestimativa do teor de Ferro por meio de cokrigagem (mapa superior); estimativa do teor de Ferro via KMAF (mapa central) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Ferro em (%)	109
Figura 4. 93 - Vista no plano XZ (leste e vertical), nível 33 dos mapas de coestimativa do teor de Ferro por meio de cokrigagem (mapa superior); estimativa do teor de Ferro via KMAF	

(mapa central) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Ferro em (%) 110

Figura 4. 94 - Vista no plano YZ (norte e vertical), nível 44 dos mapas de estimativa do teor de Ferro KMAF (mapa superior esquerdo); mapas de coestimativa do teor de Ferro por meio de cokrigagem (mapa superior direito) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Ferro em (%)..... 111

Figura 4. 95 - Vista no plano XY (leste e norte), nível 25 dos mapas de coestimativa do teor de Sílica por meio de cokrigagem (mapa superior); estimativa do teor de Sílica via KMAF (mapa central) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Sílica em (%)..... 112

Figura 4. 96 - Vista no plano XZ (leste e vertical), nível 33 dos mapas de coestimativa do teor de Sílica por meio de cokrigagem (mapa superior); estimativa do teor de Sílica via KMAF (mapa central) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Sílica em (%)..... 113

Figura 4. 97 - Vista no plano YZ (norte e vertical), nível 44 dos mapas de estimativa do teor de Sílica via KMAF (mapa superior esquerdo); de coestimativa do teor de Sílica por meio de cokrigagem (mapa superior direito) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Sílica em (%)..... 114

LISTA DE TABELAS

Tabela 4. 1 - Sumário estatístico agrupado e desagrupado do conjunto de variáveis do Jura (Goovaerts, 1997). Unidades dos atributos: concentração em (ppm).....	36
Tabela 4. 2 - Parâmetros usados para construção dos variogramas experimentais direcionais do sistema de cokrigagem.	41
Tabela 4. 3 - Matriz de correionalização do efeito pepita, C_0 , para o sistema de cokrigagem. Os termos cruzados constituem o efeito pepita dos variogramas cruzados, bem como os termos diretos constituem o efeito pepita dos variogramas diretos.	44
Tabela 4. 4 - Matriz de correionalização da primeira estrutura esférica, C_1 , para o sistema de cokrigagem. Os termos cruzados constituem a contribuição da primeira estrutura esférica dos variogramas cruzados, bem como os termos diretos constituem a contribuição da primeira estrutura esférica dos variogramas diretos.	44
Tabela 4. 5 - Matriz de correionalização da segunda estrutura esférica, C_2 , para o sistema de cokrigagem. Os termos cruzados constituem a contribuição da segunda estrutura esférica dos variogramas cruzados, bem como os termos diretos constituem a contribuição da segunda estrutura esférica dos variogramas diretos.	44
Tabela 4. 6 - Parâmetros de coestimativa utilizados no Software Isatis®.	45
Tabela 4. 7 – Parâmetros usados para a construção dos variogramas direcionais experimentais do sistema de KMAF.	55
Tabela 4. 8 - Tabela dos parâmetros de krigagem para os fatores MAF. Cada fator foi interpolado separadamente com seu modelo de continuidade e somente o raio do elipsoide de busca na estratégia de krigagem permanece igual.	60
Tabela 4. 9 - Comparativo entre as médias originais desagrupadas das variáveis do banco de dados Jura e as médias obtidas por meio de coestimativa (cokrigagem).	61
Tabela 4. 10 - Comparativo entre as médias originais desagrupadas das variáveis do banco de dados Jura e as médias obtidas por meio de estimativa (KMAF).	61
Tabela 4. 11 - Tabela indicando as variáveis que serão analisadas ao longo do estudo de caso. A letra "T" na denominação de cada variável indica que se trata da quantidade global.	83
Tabela 4. 12 - Sumário estatístico da variável Ferro global (FET), e Sílica (SIT), agrupados e desagrupados. Valores expressos em (%).	87
Tabela 4. 13 - Parâmetros utilizados na construção dos variogramas experimentais direcionais dos atributos Ferro e Sílica.	88
Tabela 4. 14 - Parâmetros do modelo linear de correionalização, nas direções N112, N22 e Vertical.	90

Tabela 4. 15 - Matriz de correionalização para o efeito pepita C_0 . Os termos cruzados representam os valores atribuídos ao efeito pepita do modelo cruzado de correionalização.	90
Tabela 4. 16 - Matriz de correionalização para primeira estrutura C_1 . Os termos cruzados representam os valores atribuídos à contribuição da primeira estrutura do modelo cruzado de correionalização.	90
Tabela 4. 17 - Matriz de correionalização para a segunda estrutura C_2 . Os termos cruzados representam os valores atribuídos à contribuição da segunda estrutura do modelo cruzado de correionalização.	91
Tabela 4. 18 - Tabela de parâmetros de cokrigagem utilizados na coestimativa da variável Ferro em que este é a variável primária.	91
Tabela 4. 19 - Tabela de parâmetros de cokrigagem utilizados na coestimativa da variável Sílica em que esta é a variável primária.	92
Tabela 4. 20 – Parâmetros utilizados na construção dos variogramas experimentais direcionais dos fatores independentes MAF1 e MAF2.	97
Tabela 4. 21 - Tabela com os parâmetros do modelo de regionalização do fator independente MAF1 nas direções N112 $dip = 0$, N22 $dip = 0$ e vertical.	99
Tabela 4. 22 - Tabela com os parâmetros do modelo de regionalização para fator independente MAF2 nas direções N112 $dip = 0$, N22 $dip = 0$ e vertical.	100
Tabela 4. 23 - Parâmetros de krigagem utilizados no processo de estimativa do fator MAF1.	100
Tabela 4. 24 - Parâmetros de krigagem utilizados no processo de estimativa do fator MAF2.	101
Tabela 4. 25 - Sumário da comparação entre as médias dos dados originais desagrupadas e as médias obtidas por meio de cada método aplicado. Além disso, cálculo do erro relativo percentual das médias globais de cada variável do estudo de caso 2, banco de dados Sapecado.	102

RESUMO

Há, na mineração, no que se refere à estimativa de teores, inúmeras situações em que as variáveis de interesse exibem correlação entre si. Neste caso, devem ser consideradas as correlações no processo a fim de se obter um resultado consistente com a estatística de correlação entre as variáveis presentes no depósito. A maneira clássica de abordar problemas desta natureza é por meio do uso da cokrigagem, que exige a modelagem das covariâncias diretas e cruzadas das variáveis. Esta metodologia é muito trabalhosa quando envolve mais de duas variáveis. Alternativas aparecem na forma de métodos de fatoração, em que as variáveis originais correlacionadas são transformadas em fatores descorrelacionados. Nesta linha, se tem o método de Mínimos-Máximos Fatores de Autocorrelação (*Minimum-Maximum Autocorrelation Factors*, MAF), no qual, por meio de uma transformação linear nos atributos correlacionados, leva-os a um espaço em que estejam descorrelacionados, possibilitando que a estimativa seja realizada individualmente pela krigagem, desta forma, evitando o uso do modelo linear de correionalização. Nesta dissertação, é analisada a aplicabilidade do método de Mínimos-Máximos Fatores de Autocorrelação à estimativa de teores em depósitos multivariados, bem como analisadas suas vantagens e limitações. A técnica alternativa é comparada com o método de coestimativa clássica, cokrigagem, com as análises em diferenças de estimativas obtidas por meio de cada metodologia. Os resultados obtidos mostraram, nos estudos de caso realizados, que a solução MAF proveu resultados de maneira mais rápida e simples sem perda de precisão em relação à cokrigagem.

ABSTRACT

There are in mining, as regards the grade estimation, numerous situations in which the variables of interest exhibit correlation. In this case, the correlation must be considered in the process to obtain a result consistent with the statistical correlation between the variables present in the deposit. The classic way of addressing such problems is through the use of cokriging, which requires the modeling of direct and cross-covariances of the variables. This methodology is very troublesome when it involves more than two variables. Alternative appear as factoring methods, where the original correlated variables are transformed into uncorrelated factors. In this line, you have the method Minimum - Maximum Autocorrelation

Factors (Minimum - Maximum autocorrelation Factors, MAF) where through a linear transformation on attributes correlated, leads them to a space where they are uncorrelated, allowing that the estimation be performed individually through kriging, thus avoiding the use of the linear model of coregionalization. In this dissertation is analyzed the applicability of the method of Minimum - Maximum Autocorrelation Factors for grade estimation in multivariate deposits, as well as its advantages and limitations. The alternative technique is compared to the classic coestimation method, cokriging, by analyzing differences in the estimates obtained by each method. The results showed, in the case studies that the MAF solution provided results more quickly and easily without loss of accuracy regarding to the cokriging.

CAPÍTULO 1

INTRODUÇÃO

Na avaliação de depósitos minerais, busca-se descobrir a quantidade de minério presente na área de estudo, a qualidade deste e onde estão as zonas mais ricas para posterior lavra da região.

Tais informações podem ser obtidas por meio de sondagem e análise laboratorial das amostras, porém este é um procedimento de custo elevado, razão pela qual tem seu uso limitado. A geoestatística torna-se indispensável quando há o interesse em compreender as características de um depósito mineral, pois ela permite que se infiram informações sobre área desconhecida.

Krigagem (Matheron, 1963) é a ferramenta geoestatística mais conhecida e utilizada atualmente, porém ela trabalha com um modelo em que é considerado somente um atributo, e, por vezes, o objeto de estudo não possui somente uma variável, mas, sim, uma ou mais variáveis secundárias. Estas variáveis podem estar correlacionadas ao atributo de interesse. Portanto, devem-se levar em consideração as informações secundárias as quais se têm disponíveis. É preciso um método que considere estas correlações a fim de construir um mapa de estimativas próximo à distribuição de teores observada no depósito (Isaaks & Srivastava, 1989).

A maneira clássica de lidar com informações secundárias é a cokrigagem (Marechal, 1970), contudo, esta é extremamente laboriosa quando se lida com mais de três atributos, devido à modelagem das covariâncias diretas e cruzadas entre as variáveis. Alguns métodos de incorporação da variável secundária já vêm sendo utilizados, como a cokrigagem colocada (Xu et. al, 1992) utilizando modelos de Markov (Almeida, 1993; Journel, 1999), métodos estes mais simples que a cokrigagem. Todavia, eles apresentam uma limitação, pois necessitam da amostragem exaustiva do atributo secundário, ou que seja feita a sua complementação via krigagem ordinária ou outras metodologias, o que, inevitavelmente, acrescentará erros à estimativa final (Isaaks & Srivastava, 1989).

Outra maneira de analisar bancos de dados multivariados, em que as variáveis apresentam algum tipo de correlação entre si, seja espacial, seja temporal, são os métodos de fatoração, que consistem em decomposições espectrais dos vetores de observação, a fim de descorrelacioná-los e, desta forma, estudar cada fator independentemente. Nesta linha, tem-se

o método de Análise das Componentes Principais (*Principal Component Analysis – PCA*) (Pearson, 1901; Hoettling, 1933), que consiste em realizar uma rotação no sistema das observações correlacionadas, levando-as a um novo sistema de referências em que as variáveis correlacionadas são ortogonais entre si (correlação igual à zero). Procede-se as estimativas de cada componente de maneira independente utilizando a krigagem ordinária, e, depois de estimadas, a retrotransformação as leva ao sistema de referência original sem perda de informação (Davis & Greenes, 1983). Na limitação do PCA, a descorrelação acontece somente para vetores de separação nulos (*lag* igual à zero) (Davis & Greenes, 1983).

Outro método conhecido é o Mínimos-Máximos Fatores de Autocorrelação (*Minimum-Maximum Autocorrelation Factors – MAF*) (Switzer & Green, 1984). Este método é derivado do PCA, e, como tal, também leva as variáveis originais a fatores descorrelacionados. Entretanto, o MAF obtém a descorrelação das variáveis para vetores de separação não nulos. Desta forma, é possível trabalhar cada variável do sistema de maneira independente.

Na presente dissertação, será analisada a aplicabilidade do último método, suas vantagens e limitações.

1.1 PROBLEMA

Para estimar teores em depósitos multivariados, nos quais há atributos correlacionados, é necessário considerar as correlações existentes a fim de obter estimativas adequadas.

Apesar da geoestatística clássica prover, com a cokrigagem, o meio de considerar as correlações existentes, na prática, seu uso é limitado, já que a modelagem da correionalização das variáveis é um processo complexo.

1.2 OBJETIVO

O objetivo do presente estudo é avaliar a utilização da krigagem ordinária combinada com o método de Mínimos-Máximos Fatores de Autocorrelação (MAF) como alternativa à cokrigagem de múltiplas variáveis correlacionadas em depósitos minerais.

1.3 METAS

Como metas, têm-se:

- I. Desenvolver uma rotina que realize a transformação MAF de forma automática;
- II. Avaliar os resultados de MAF aplicado a dois bancos de dados multivariados, utilizando os resultados obtidos por meio de cokrigagem como referência.

1.4 METODOLOGIA

Serão realizados estudos de caso no banco de dados Jura (Goovaerts, 1997), e Sapecado (fornecido pela empresa Vale).

Os estudos de caso serão divididos em dois segmentos, o primeiro, cokrigagem (Marechal, 1970) das variáveis correlacionadas, que, neste caso, será tomada como referência para a análise de resultados por ser o método clássico de estimativa de depósitos multivariados.

A cokrigagem (Marechal, 1970), baseada no mesmo princípio que a krigagem (Matheron, 1963), procura realizar estimativas com erro médio zero e variância do erro mínima, porém considerando covariâncias diretas e cruzadas dos atributos.

A consideração de atributos correlacionados no processo de estimativa leva à construção de modelos de covariâncias que devem satisfazer à condição de definição positiva do Modelo Linear de Corregionalização (Goovaerts, 1997).

O processo de cokrigagem pode se tornar extremamente complexo sem o auxílio de *softwares* adequados.

O segundo segmento consiste na aplicação do método de Mínimos-Máximos Fatores de Autocorrelação (MAF), proposto, neste trabalho, como alternativa à cokrigagem, em que por meio de transformações lineares aplicadas nos dados originais, estes se tornam novos fatores que são linearmente independentes.

Sem a correlação, os atributos podem ser estimados independentemente, por meio de krigagem ordinária e retrotransformados ao seu espaço original, de modo que a construção de modelos de covariâncias cruzadas é evitada.

Baseado nos dois estudos de caso, serão analisados os resultados obtidos por meio do método MAF em relação aos obtidos pela cokrigagem a fim de avaliar a aplicabilidade do método como alternativa para estimativas de depósitos multivariados.

1.5 ORGANIZAÇÃO DA DISSERTAÇÃO

A dissertação a ser exposta está subdividida nos seguintes capítulos, e desta forma organizada: o capítulo 2 apresenta uma breve revisão dos fundamentos teóricos necessários para a compreensão do tema abordado: uma abordagem concisa da geoestatística multivariada. Serão revisados também conceitos básicos de álgebra linear e os fundamentos teóricos que envolvem PCA e MAF.

O capítulo 3 apresenta o algoritmo desenvolvido na linguagem *Python*® para a realização da transformação MAF. Serão apresentados detalhes de como utilizar a rotina MAF.

O capítulo 4 consiste na aplicação do método MAF e cokrigagem a dois estudos de caso. Serão utilizados o banco de dados Jura (Goovaerts, 1997) e o banco de dados Sapecado, fornecido pela empresa Vale S.A. Neste capítulo, também serão realizadas as análises dos resultados obtidos em cada estudo de caso.

O capítulo 5 apresenta as conclusões obtidas com o estudo realizado nesta dissertação, bem como as recomendações a trabalhos futuros.

CAPÍTULO 2

FUNDAMENTOS TEÓRICOS

Este capítulo trata de uma breve revisão teórica dos métodos clássicos de estimativa de teores em bancos de dados multivariados e de métodos alternativos com descorrelação, que são o objetivo deste trabalho.

2.1 ESTIMATIVAS DE DEPÓSITOS MULTIVARIADOS

Em muitos depósitos minerais, não são levantadas informações apenas de uma variável de interesse, mas sim múltiplas variáveis, seja pelo modelo geológico a ser construído, seja pelo interesse econômico em mais de um atributo. O que ocorre, com frequência, é que as múltiplas variáveis, das quais se possui informações e se deseja estimar, estão correlacionadas entre si.

A geoestatística trata levantamentos amostrais de natureza multivariada por meio da cokrigagem (Marechal, 1970), que, assim como a amplamente utilizada krigagem (Matheron, 1963), busca a minimização da variância do erro de estimativa.

O uso da cokrigagem em estimativas de dados multivariados é consolidado e adequado, pois esta incorpora a variabilidade espacial conjunta dos atributos. A dificuldade, porém, encontra-se na modelagem desta variabilidade espacial conjunta. Um meio de solucionar este problema é utilizar o Modelo Linear de Corregionalização (MLC). Muitas de suas aplicações podem ser encontradas na literatura, como Goovaerts (1997). O MLC procura expressar os variogramas diretos e cruzados como uma combinação linear das estruturas básicas presentes, de modo que a variância desta combinação seja sempre positiva. Apesar de o MLC solucionar o problema da criação do modelo para a variabilidade espacial conjunta dos dados, sua aplicação não é simples. No caso de haver mais de três variáveis a serem modeladas, a aplicação do MLC pode tornar-se complexa.

Uma simplificação do MLC é o Modelo de Correlação Intrínseca (MCI) (Wackernagel, 1995). O modelo MCI assume que todas as variâncias e covariâncias são proporcionais a um mesmo modelo de corregionalização. Detalhes sobre o MLC e o MCI podem ser encontrados em Goovaerts (1997).

Frente às dificuldades apresentadas pela cokrigagem e pelo MLC, foram introduzidas diversas simplificações para estimativas de depósitos multivariados, tais como: Cokrigagem colocada; modelos de Markov (1 e 2), entre outros.

2.1.1 COKRIGAGEM COLOCADA

No sistema de cokrigagem clássico, se as variáveis secundárias possuem amostragens mais abundantes que a variável primária, o sistema de cokrigagem pode tornar-se instável na resolução da inversão das matrizes. Isto se deve ao fato dos dados secundários próximos apresentarem maior correlação do que dados primários distantes. Além disso, dados secundários próximos ao local de estimativa ou, até mesmo, no local de estimativa, filtram a influência de dados secundários mais distantes (Goovaerts, 1997). Segundo Xu *et al.* (1992), uma maneira de contornar a instabilidade gerada pelo excesso de dados secundários é utilizar somente o dado secundário no local a ser estimado o primário, este procedimento ficou conhecido como Cokrigagem Colocada.

Apesar de a cokrigagem colocada solucionar o problema de estabilidade da resolução do sistema de equações que resolvem os pesos a serem dados pelas amostras, esta ainda depende da modelagem do MLC. Uma nova aproximação do método de cokrigagem que propõe uma simplificação são os Modelos de Markov.

2.1.2 MODELOS DE MARKOV

Foram desenvolvidos dois modelos chamados Modelos de Markov. O primeiro, denotado de MM1, introduzido por Almeida (1993), reduz o problema de construção do MLC a um reescalonamento das covariâncias da variável primária e secundária.

O método assume que a covariância cruzada das variáveis é proporcional à covariância da variável primária. O MM1, entretanto, necessita que todas as covariâncias cruzadas compartilhem a mesma forma da continuidade espacial que a variável primária apresenta. Em casos em que uma das secundárias é vastamente mais amostrada, esta hipótese pode não ser a mais adequada, uma vez que seria muito mais simples obter a continuidade espacial de variáveis secundárias em função da abundância de dados disponíveis (quando comparados aos dados da primária).

O segundo Modelo de Markov desenvolvido por Journel (1999), denotado MM2, assume que o modelo de covariância cruzado é proporcional à covariância da variável secundária, e a covariância da variável primária é obtida por meio de uma nova componente introduzida no sistema, a residual ou resíduo.

O resíduo é obtido pela regressão linear entre a variável primária e secundária. Apesar de o MM2 assumir uma hipótese mais adequada à configuração espacial dos dados (abundância dos dados secundários e, portanto, mais fácil de determinar sua continuidade espacial), ainda é necessária a modelagem das covariâncias diretas e cruzadas, mesmo assumindo proporcionalidade entre elas. Deste modo, a dificuldade encontrada ao utilizar o MLC não é completamente contornada, uma vez que o MM2 deve também satisfazer as condições de definição positiva do modelo.

2.1.3 ASPECTOS PRÁTICOS

Os métodos de coestimativa apresentados acima, com exceção da cokrigagem completa, exigem amostragem exaustiva da variável secundária. Porém, esta condição não se apresenta, com frequência, em bancos de dados referentes a depósitos minerais, limitando o uso das simplificações da metodologia clássica e remetendo o usuário ao MLC.

Alternativas aos procedimentos de coestimativa recaem sobre métodos de fatorização, que buscam a desconexão das variáveis envolvidas no estudo previamente ao processo de estimativa, tais como: Análise de Componentes Principais (PCA) (Pearson, 1901; Hoettling, 1933) e Mínimos-Máximos Fatores de Autocorrelação (MAF) (Switzer & Green, 1984).

Estes métodos consistem em realizar transformações lineares nos atributos originais correlacionados, e, assim, torná-los desconexados para que possam ser estimados de maneira independente por meio de krigagem. Desta forma, os métodos de fatorização permitem tratar de múltiplas variáveis individualmente, sem ter que considerar as correlações que tenham entre si. No entanto, mesmo que as variáveis secundárias não necessitem ser amostradas de maneira exaustiva, é fundamental a configuração isotópica dos dados. A condição de isotopia das variáveis pode ser contornada, no caso em que as amostras faltantes não constituem uma fração significativa do banco de dados, mantendo somente o subconjunto isotópico dos atributos analisados. Caso a fração heterotópica dos dados seja significativa em relação à totalidade do conjunto de amostras, nos locais amostrais em que há heterotopia, as

variáveis faltantes podem ser krigadas ou simuladas, de modo que a escolha da isotopização dos dados deve ser realizada com cautela. Os métodos alternativos aos clássicos apresentados aqui serão explorados neste capítulo, em detalhes.

2.2 FUNDAMENTOS DE ÁLGEBRA LINEAR

Nesta seção, é feita uma breve revisão sobre alguns conceitos de álgebra linear, auxiliares à compreensão da Análise de Componentes Principais (PCA) e sobre os Mínimos-Máximos Fatores de Autocorrelação (MAF).

2.2.1 TRANSFORMAÇÕES LINEARES

Definição: Se v é uma função que tem seu domínio em \mathbb{R}^N , e seu contradomínio em \mathbb{R}^M , então $v: \mathbb{R}^N \rightarrow \mathbb{R}^M$ é chamada de transformação de \mathbb{R}^N em \mathbb{R}^M (Anton & Rorres, 2000). Se $V(v_1, v_2, \dots, v_M)$ compor um conjunto de equações do tipo:

$$\begin{aligned} v_1 &= a_{11}u_1 + a_{12}u_2 + \dots + a_{1N}u_N \\ &\quad \vdots \\ v_M &= a_{M1}u_1 + a_{M2}u_2 + \dots + a_{MN}u_N \end{aligned} \quad (2.1)$$

Então V é dito ser uma transformação linear de \mathbb{R}^N em \mathbb{R}^M .

Em notação matricial:

$$\mathbf{V} = \mathbf{aU} \quad (2.2)$$

sendo \mathbf{a} chamada de matriz canônica ou transformação de componentes a_{ij} que leva o vetor \mathbf{U} formado por $\{u_1, u_2, \dots, u_N\}$ ao vetor \mathbf{V} composto por $\{v_1, v_2, \dots, v_M\}$.

2.2.2 AUTOVETORES E AUTOVALORES

Definição: Seja \mathbf{B} uma matriz quadrada $N \times N$, se λ é um escalar e \mathbf{a} um vetor coluna não nulo que satisfaça a relação (Matthews, 1991):

$$\mathbf{Ba} = \lambda\mathbf{a} \quad (2.3)$$

então λ é autovalor de B e \mathbf{a} é o autovetor associado ao autovalor λ . Para calcular os autovalores associados a B , deve-se solucionar a seguinte equação:

$$\det(B - \lambda I) = 0 \quad (2.4)$$

de modo que I é a matriz identidade e \det denota o determinante da matriz resultante. A relação (2.4) é chamada de polinômio característico da matriz B , e os autovalores λ são as raízes desta equação. Portanto, explicitamente:

$$\det \left(\begin{bmatrix} B_{11} & \cdots & B_{1N} \\ \vdots & \ddots & \vdots \\ B_{N1} & \cdots & B_{NN} \end{bmatrix} - \lambda \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \right) = 0 \quad (2.5)$$

Para uma matriz de dimensão N , haverá N autovalores e N autovetores.

Com a expressão (2.3), bem como os valores de λ_N determinados, é possível obter os \mathbf{a}_N autovetores da matriz B , ou seja:

$$\begin{bmatrix} B_{11} & \cdots & B_{1N} \\ \vdots & \ddots & \vdots \\ B_{N1} & \cdots & B_{NN} \end{bmatrix} \begin{bmatrix} a_{i1} \\ a_{ij} \\ \vdots \\ a_{iN} \end{bmatrix} = \lambda_i \begin{bmatrix} a_{i1} \\ a_{ij} \\ \vdots \\ a_{iN} \end{bmatrix}, i, j = 1, \dots, N \quad (2.6)$$

sendo λ_i o autovalor que possui o autovetor \mathbf{a}_i associado e a_{ij} as componentes do autovetor \mathbf{a}_i .

2.2.3 DECOMPOSIÇÃO ESPECTRAL

Decomposição espectral é um método de fatoração de matrizes quadradas, cujo objetivo é reescrevê-las de maneira mais simples em termos dos seus autovetores e autovalores (Rodrigues, 2011).

Seja B é uma matriz quadrada $N \times N$, então:

$$B = \mathbf{a}\Lambda\mathbf{a}^{-1} \quad (2.7)$$

onde \mathbf{a} é a matriz não singular de autovetores e Λ a matriz de autovalores.

2.2.4 VETOR DE MÉDIAS E MATRIZ DE COVARIÂNCIA

Para demonstrar as propriedades do vetor de médias e da matriz de covariância, serão realizados cálculos análogos aos encontrados em Boezio (2010). Considerando o vetor randômico $\Phi[U_i]$, a média das variáveis randômicas é tal que:

$$m_i = E(U_i) \quad (2.8)$$

sendo E a esperança matemática, U_i os valores que o vetor randômico Φ pode assumir e m_i sua média, a variância das variáveis randômicas, σ_i^2 , é:

$$\sigma_i^2 = E(U_i - m_i)^2 \quad (2.9)$$

Da mesma forma, a covariância entre duas variáveis randômicas, Cov_{ij} é dada por:

$$Cov_{ij} = E[(U_i - m_i)(U_j - m_j)] \quad (2.10)$$

O vetor de médias, \mathbf{m}_Φ , é:

$$\mathbf{m}_\Phi = \begin{bmatrix} m_1 \\ \vdots \\ m_N \end{bmatrix} = \begin{bmatrix} E(U_1) \\ \vdots \\ E(U_N) \end{bmatrix} = E(\Phi) \quad (2.11)$$

E a matriz de covariâncias de Φ , \mathbf{Cov}_Φ , é:

$$\mathbf{Cov}_\Phi = Cov(\Phi) = E[(\Phi - \mathbf{m}_\Phi)(\Phi - \mathbf{m}_\Phi)^T] = \begin{bmatrix} Cov_{11} & \cdots & Cov_{1N} \\ \vdots & \ddots & \vdots \\ Cov_{N1} & \cdots & Cov_{NN} \end{bmatrix} \quad (2.12)$$

sendo que $Cov_{11} = \sigma_1^2$, ou seja, na diagonal da matriz, encontram-se as respectivas variâncias de cada variável randômica. Se U_i for multiplicada por uma constante α :

$$E(\alpha U_i) = \alpha E(U_i) = \alpha m_i \quad (2.13)$$

$$\sigma_i^2(\alpha U_i) = E(\alpha U_i - \alpha m_i)^2 = \alpha^2 \sigma_i^2(U_i) = \alpha^2 \sigma_i^2 \quad (2.14)$$

Considerando outra constante β , a relação (2.15) também se manteria válida.

$$\begin{aligned} Cov_{12}(\alpha U_1, \beta U_2) &= E(\alpha U_1 - \alpha m_1)(\beta U_2 - \beta m_2) = \alpha\beta E(U_1 - m_1)(U_2 - m_2) \\ &= \alpha\beta Cov(U_1, U_2) = \alpha\beta Cov_{12} \end{aligned} \quad (2.15)$$

sendo m_1 a média de U_1 e m_2 a U_2 .

Para a combinação linear $\alpha U_i + \beta U_j$, teríamos:

$$E(\alpha U_1 + \beta U_2) = \alpha E(U_1) + \beta E(U_2) = \alpha m_1 + \beta m_2 \quad (2.16)$$

E que a variância de $\sigma^2(\alpha U_1 + \beta U_2)$ é:

$$\begin{aligned} \sigma^2(\alpha U_1 + \beta U_2) &= E[(\alpha U_1 + \beta U_2) - (\alpha m_1 + \beta m_2)]^2 = E[\alpha(U_1 - m_1) + \beta(U_2 - m_2)]^2 \\ &= E[\alpha^2(U_1 - m_1)^2 + \beta^2(U_2 - m_2)^2 + 2\alpha\beta(U_1 - m_1)(U_2 - m_2)] \\ &= \alpha^2 Var(U_1) + \beta^2 Var(U_2) + 2\alpha\beta Cov(U_1, U_2) \\ &= \alpha^2 \sigma_1^2 + \beta^2 \sigma_2^2 + 2\alpha\beta Cov_{12} \end{aligned} \quad (2.17)$$

sendo que $Var(U_1) = \sigma_1^2$, $Var(U_2) = \sigma_2^2$ e $Cov(U_1, U_2) = Cov_{12}$.

Se α e β são componentes do vetor transposto \mathbf{a}^T , a combinação linear entre U_1 e U_2 pode ser escrito na forma matricial:

$$[\alpha \quad \beta] \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \mathbf{a}^T \Phi \quad (2.18)$$

Da mesma forma:

$$[\alpha \quad \beta] \begin{bmatrix} E(U_1) \\ E(U_2) \end{bmatrix} = \mathbf{a}^T \mathbf{m}_\Phi \quad (2.19)$$

e

$$Cov_\Phi = \begin{bmatrix} \sigma_{11}^2 & \cdots & Cov_{1N} \\ \vdots & \ddots & \vdots \\ Cov_{N1} & \cdots & \sigma_{NN}^2 \end{bmatrix}$$

De modo que:

$$\sigma^2(\alpha U_1 + \beta U_2) = \sigma^2(\mathbf{a}^T \Phi) = \mathbf{a}^T Cov_\Phi \mathbf{a} \quad (2.20)$$

2.3 ANÁLISE DE COMPONENTES PRINCIPAIS

Como mencionado na seção 2.1, tratando-se de um conjunto de dados multivariados, em que haja correlação entre os atributos, para a realização de estimativas, o método recomendado seria a cokrigagem. Essa requer a construção de modelos de correionalização permissíveis pelo MLC. Nesta seção, será abordado um método alternativo, que descorrelaciona as variáveis por meio de transformações lineares.

O método de componentes principais (Principal Component Analysis – PCA) foi desenvolvido por Pearson, em 1901, com o intuito de encontrar linhas e planos que melhor ajustam um conjunto de pontos dentro de um espaço p -dimensional (Pearson, 1901). O problema abordado era de otimização geométrica. Apesar do enfoque geométrico, Pearson alegava que o método poderia ser aplicado a problemas numéricos, ainda que, na época, houvesse mais de três variáveis, a solução poderia ser extremamente complexa de realizar, pois não havia auxílio de computadores (Jolliffe, 2002).

Hoetling (1933) trabalha com a ideia introduzida por Pearson, constituindo uma nova abordagem. Hoetling propõe que o PCA seja utilizado para realizar a redução da dimensionalidade de um sistema, encontrando um novo conjunto de variáveis (componentes principais – CP's), as quais seriam independentes e “carregariam” cada uma das novas variáveis parte da variância total do sistema original. Desta forma, o novo conjunto pode ser composto somente pelas variáveis que possuem maior parte desta informação (Hoetling, 1933).

Trabalhando com a abordagem de Hoetling, é possível transportar a ideia ao problema de estimativa de teores com N atributos correlacionados. O foco principal aqui não é a redução de dimensão, mas a descorrelação dos atributos, tornando praticável a utilização das variáveis secundárias sem a necessidade do MLC.

O PCA consiste em realizar uma rotação no sistema de coordenadas (x,y,z) de um banco de dados altamente correlacionado, levando-o a um novo sistema de coordenadas (x',y',z') , de modo que agora as variáveis não possuem mais correlação (Hoetling, 1933).

Considere um vetor randômico de N variáveis, $\Phi(U_1, U_2, \dots, U_N)$, em que as informações das variâncias e covariâncias dessas variáveis são relevantes. Desse modo, a

análise das $\frac{1}{2}N(N-1)$ covariâncias e das N variâncias deve ser realizada (Jolliffe, 2002). O PCA assume que algumas variáveis contenham mais da variância do sistema do que outras. Portanto, poderiam ser analisadas somente as que possuem maior parte desta informação. As novas componentes descorrelacionadas, Z_k , são dadas pela seguinte combinação linear:

$$Z_1 = a_1^T \Phi \quad (2.21)$$

onde a_1 é um vetor transformação e a_1^T é o transposto de a_1 e Z_1 é a primeira componente principal definida. Então:

$$a_1^T \Phi = a_{11}U_1 + a_{12}U_2 + \dots = \sum_{i=1}^N a_{1i}U_i \quad (2.22)$$

Logo,

$$Z_1 = a_{11}U_1 + a_{12}U_2 + \dots = \sum_{i=1}^N a_{1i}U_i \quad (2.23)$$

onde a função linear $a_1^T \Phi$ tem máxima variância. Após, é definida uma nova função linear $a_2^T \Phi$, que tem máxima variância em relação à variância restante do sistema e é independente da anterior, de modo que cada nova variável possua máxima variância possível de acordo com a variância remanescente.

Note que a expressão (2.23) nada mais é que uma transformação linear que leva U_i em Z_k equivalente à equação (2.1) da seção 2.2.

Se as novas componentes Z_k são combinações lineares, sua variância pode ser dada pela relação (2.20). Além disso, a variância de cada componente deve também ser máxima, o que acontece se, e somente se, a_1 é finito. Portanto, o problema está sujeito ao vínculo que $a_1^T a_1 = 1$. A fim de maximizar a variância sujeita ao vínculo, serão utilizados parâmetros de Lagrange. Seja a variância de Z_1 chamada de $\sigma_{Z_1}^2$, então:

$$\sigma_{Z_1}^2 = a_1^T Cov_{\Phi} a_1 - \lambda_1 (a_1^T a_1 - 1) \quad (2.24)$$

sendo $\sigma_{Z_1}^2$ a variância da primeira componente principal Z_1 , λ_1 o parâmetro de Lagrange introduzido na relação e a_1^T o vetor transformação transposto. Realizando a derivada parcial de (2.24) em relação à a_1^T , chega-se à expressão:

$$Cov_{\Phi} a_1 - \lambda_1 a_1 = 0 \quad \rightarrow \quad (Cov_{\Phi} - \lambda_1 I_N) a_1 = 0 \quad (2.25)$$

onde I_N é a matriz identidade e pela relação (2.4), tem-se que λ_1 é autovalor de Cov_{Φ} , e a_1 é o autovetor de Cov_{Φ} , que proporcionam a máxima variância à componente principal Z_1 calculada.

A proporção da variância que cada componente principal carrega é dada pela seguinte equação:

$$(\%)\sigma_T^2 = \frac{\lambda_i}{\sum_{i=1}^N \lambda_i} \cdot 100 \quad (2.26)$$

sendo σ_T^2 a variância total do sistema de variáveis, λ_i o i -ésimo autovalor associado à i -ésima componente principal Z_i .

Para ilustrar, apresenta-se, no que segue, um exemplo simples do cálculo das componentes principais de um sistema de duas variáveis.

Considere o banco de dados Jura (Goovaerts, 1997), no qual foram medidas as concentrações de sete metais. Dentre os sete metais, consideram-se somente dois, Cobalto e Níquel, que possuem correlação $\rho=0,74$. A matriz de variância-covariância do sistema é:

$$Cov_{\Phi} = \begin{bmatrix} 12,73 & 21,47 \\ 21,47 & 65,52 \end{bmatrix}$$

em que 12,73 é o valor da variância de Cobalto, 65, 52 a variância de Níquel e 21,47 a covariância do par Cobalto e Níquel. Para calcular os autovetores e autovalores do sistema, utilizar-se-á a equação 2.4 do item 2.2.2 da seção 2.2.

$$\det(Cov_{\Phi} - \lambda I) = 0$$

$$\lambda I = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

$$\det \begin{bmatrix} 12,73 - \lambda_1 & 21,47 \\ 21,47 & 65,52 - \lambda_2 \end{bmatrix} = 0$$

Foram obtidos os autovalores $\lambda_1 = 73,15$ e $\lambda_2 = 5,1$. Os autovetores são calculados utilizando a equação (2.3) do item 2.2.2 da seção 2.2, ou seja,

$$Cov_{\Phi} \mathbf{a} = \lambda \mathbf{a}$$

explicitamente:

$$\begin{bmatrix} 12,73 & 21,47 \\ 21,47 & 65,52 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix} = \lambda_1 \begin{bmatrix} a_{11} \\ a_{12} \end{bmatrix}$$

sendo λ_1 o autovalor associado ao autovetor a_1 e a_{11}, a_{12} as componentes do autovetor a_1 . O mesmo se faz para a_2 , utilizando λ_2 . Assim: $a_1 = \begin{bmatrix} -0,33 \\ -0,94 \end{bmatrix}$ e $a_2 = \begin{bmatrix} -0,94 \\ 0,33 \end{bmatrix}$ são os autovetores associados à matriz de variância-covariância do sistema.

Antes de calcular as componentes principais é importante verificar a independência linear dos autovetores. Vetores independentes linearmente têm seu produto interno igual à zero.

O produto interno entre vetores é calculado de acordo com a seguinte equação:

$$(a_{11} \cdot a_{21}) + (a_{12} \cdot a_{22}) = 0$$

Desta forma:

$$(-0,94 \cdot -0,33) + (-0,94 \cdot 0,33) = 0,3102 - 0,3102 = 0$$

As componentes principais associadas são dadas pela seguinte relação:

$$\begin{cases} Z_1 = a_{11}U_1 + a_{12}U_2 \\ Z_2 = a_{21}U_1 + a_{22}U_2 \end{cases}$$

Se U_1 for a variável Cobalto e U_2 for a variável Níquel, a_{11} e a_{12} as componentes do autovetor a_1 , e a_{21} e a_{22} as componentes do autovetor a_2 , de acordo com a equação (2.21), as componentes principais serão, então:

$$\begin{cases} Z_1 = -0,33Co - 0,94Ni \\ Z_2 = -0,94Co + 0,33Ni \end{cases}$$

A partir das equações acima, pode calcular em cada ponto amostral um valor para cada componente principal e verificar a sua descorrelação por meio de um gráfico de dispersão, como observado na figura 2.1.

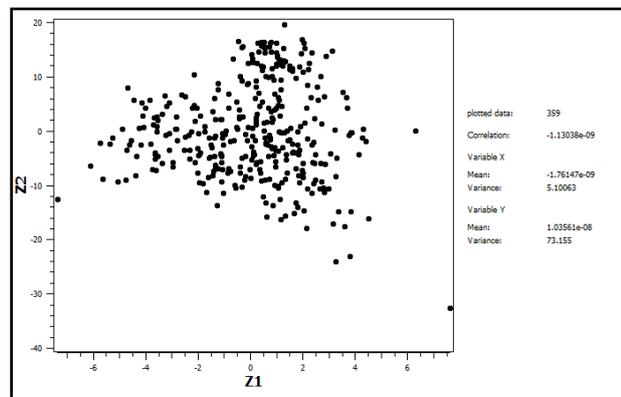


Figura 2.1 - Gráfico de dispersão entre as componentes principais Z_1 e Z_2

A proporção da variância total do sistema que cada componente principal possui, neste exemplo, é:

$$(\%)\sigma_{TZ_1}^2 = \frac{73,15}{78,25} \cdot 100 = 93,4(\%)$$

sendo $(\%)\sigma_{TZ_1}^2$ a proporção da variância total do sistema associada a componente principal Z_1 .

$$(\%)\sigma_{TZ_2}^2 = \frac{5,1}{78,25} \cdot 100 = 6,51(\%)$$

sendo $(\%)\sigma_{TZ_2}^2$ a proporção da variância total do sistema associada a componente principal Z_2 .

Para o processo de estimativas de teores em depósitos multivariados, o procedimento deve ser aplicado nos seguintes passos:

- i. Determinar a matriz de variância-covariância Cov_{Φ} dos dados;

- ii. Procurar os autovetores e autovalores associados à matriz Cov_{Φ} ;
- iii. Realizar a transformação dos dados em componentes principais Z_k de acordo com a relação (2.21);
- iv. Calcular e construir modelar os variogramas para as N componentes principais diretos e realizar a krigagem nas componentes principais individualmente;
- v. Levar as componentes principais krigadas novamente ao espaço original dos dados pela retrotransformação das componentes principais (CP's), a partir da seguinte relação:

$$U_{oki}(u) = (a_i^T)^{-1} Z_{ok} \quad (2.27)$$

onde $(a_i^T)^{-1}$ compõe a inversa da matriz transformação que leva às componentes principais aos dados originais, Z_{ok} é a componente principal estimada via krigagem ordinária e U_{oki} será o valor da variável U_i estimado no local u .

O método exposto acima parece bastante atraente, afinal, aplica-se uma transformação linear nos dados, a correlação entre eles é quebrada, realiza-se a estimativa de maneira independente nas CP's, retrotransforma-se e chega-se ao resultado. Infelizmente, na prática, o método não é tão eficiente, a limitação aqui é que a descorrelação não é garantida para amostras separada por $h \neq 0$. Somente no caso de correlação intrínseca é que a descorrelação acontece para todo e qualquer h . Em alternativa ao uso de componentes principais, surge o método de Mínimos-Máximos Fatores de Autocorrelação, introduzido por Switzer e Green (1984), no qual a descorrelação acontece também para amostras separadas por $h \neq 0$. Este método é exposto na seção seguinte.

2.4 MÉTODO DE MÍNIMOS-MÁXIMOS FATORES DE AUTOCORRELAÇÃO (MAF)

O método de Mínimos-Máximos Fatores de Autocorrelação (Minimum-Maximum Autocorrelation Factors – MAF) surgiu com a proposta de separar as componentes de sinal e ruídos presentes em processamentos de imagens multiespectrais provenientes do sensoriamento remoto (Switzer & Green, 1984). Segundo Switzer e Green (1984), em Desbarats e Dimitrakopoulos (2000), MAF consiste em transformar o vetor randômico multivariado $\Phi(u) = [U_1(u), U_2(u), \dots, U_N(u)]$ em um conjunto de combinações lineares que são independentes entre si.

$$M_i(u) = \mathbf{a}_i^T \Phi(u) \quad (2.28)$$

onde \mathbf{a}_i^T é o autovetor transposto da matriz de transformação \mathbf{a} , de modo que cada fator MAF, $M_i(u)$, apresenta uma correlação espacial crescente, ou seja, o fator MAF $M_2(u)$ apresenta correlação espacial maior que o fator MAF $M_1(u)$ (Switzer & Green, 1984).

Seja ρ_i a correlação espacial entre $M_i(u)$ e $M_i(u+h)$ para um vetor de separação $h = \delta$, as componentes \mathbf{a}_i^T são definidas de acordo com:

$$\begin{cases} \rho_1(\delta) = \text{corr}(\mathbf{a}_1^T \Phi(u), \mathbf{a}_1^T \Phi(u+\delta)) = \min \text{corr}(\mathbf{a}^T \Phi(u), \mathbf{a}^T \Phi(u+\delta)) \\ \vdots \\ \rho_N(\delta) = \text{corr}(\mathbf{a}_N^T \Phi(u), \mathbf{a}_N^T \Phi(u+\delta)) = \max \text{corr}(\mathbf{a}^T \Phi(u), \mathbf{a}^T \Phi(u+\delta)) \end{cases} \quad (2.29)$$

Onde a minimização de \mathbf{a}_i^T é sujeita ao seguinte vínculo:

$$\text{corr}(\mathbf{a}_i^T \Phi(u), \mathbf{a}_j^T \Phi(u)) = 0; \quad j < i \quad (2.30)$$

Segundo Switzer e Green (1984), os fatores MAF são obtidos por meio da decomposição espectral da matriz $\text{Cov}_{\Phi\delta} \text{Cov}_{\Phi}^{-1}$, em que Cov_{Φ} é matriz de covariância de $\Phi(u)$ para $h=0$ e Cov_{Φ}^{-1} é a sua inversa. $\text{Cov}_{\Phi\delta}$ é a matriz de covariância para as diferenças de vetores a uma separação $h = \delta$, ou seja,

$$\begin{cases} \text{Cov}_{\Phi} = \text{Cov}[\Phi(u), \Phi(u)] \\ \text{Cov}_{\Phi\delta} = \text{Cov}[(\Phi(u) - \Phi(u+\delta)), (\Phi(u) - \Phi(u+\delta))] = 2\Gamma_{\Phi}(\delta) \end{cases} \quad (2.31)$$

Se o vetor randômico $\Phi(u) = [U_1(u), \dots, U_N(u)]^T$ é estacionário de dimensão N , composto por uma componente de sinal $S(u)$ e outra de ruído $N(u)$, tal que $\Phi(u) = S(u) + N(u)$. As matrizes de variância-covariância são escritas como:

$$\begin{cases} \text{Cov}[N(u), N(u)] = \text{Cov}_N(0) = \text{Cov}_0 \\ \text{Cov}[S(u), S(u)] = \text{Cov}_S(0) = \text{Cov}_1 \\ \text{Cov}[\Phi(u), \Phi(u)] = \text{Cov}_0 + \text{Cov}_1 = \text{Cov}_{\Phi} \end{cases} \quad (2.32)$$

sendo $\text{Cov}_N(0)$, $\text{Cov}_S(0)$ e Cov_{Φ} as covariâncias para $h=0$. Então, se $\Phi(u)$ pode ser representado por um modelo linear de correlogramização de duas estruturas (Switzer & Green, 1984), a matriz de covariância espacial é dada por:

$$\begin{cases} Cov[N(u), N(u+h)] = Cov_N(h) = \rho_0(h)Cov_0 \\ Cov[S(u), S(u+h)] = Cov_S(h) = \rho_1(h)Cov_1 \\ Cov[\Phi(u), \Phi(u+h)] = Cov_{\Phi h}(h) = \rho_0(h)Cov_0 + \rho_1(h)Cov_1 \end{cases} \quad (2.33)$$

Onde $\rho_0(h)$ e $\rho_1(h)$ são as funções de correlação espacial tais que $\rho_1 > \rho_0$ para todo $h > 0$, sendo Cov_N a covariância da componente ruído, e Cov_S a covariância da componente sinal, $Cov_{\Phi h}$ a covariância do vetor composto pelo sinal e ruído com vetor de separação h . Cov_0 e Cov_1 são as duas estruturas de covariância. Assim, para $h = \delta$, teremos:

$$\begin{aligned} Cov[(\Phi(u) - \Phi(u + \delta)), (\Phi(u) - \Phi(u + \delta))] \\ &= Cov[\Phi(u)^2 + \Phi(u + \delta)^2 - 2\Phi(u)\Phi(u + \delta)] \\ &= \sigma^2[\Phi(u)] + \sigma^2[\Phi(u + \delta)] - 2Cov[\Phi(u), \Phi(u + \delta)] \\ &= Cov_{\Phi} + Cov_{\Phi} - 2(Cov_0\rho_0(\delta) + Cov_1\rho_1(\delta)) \\ &= 2Cov_{\Phi} - 2Cov_0\rho_0(\delta) - 2(Cov_{\Phi} - Cov_0)\rho_1(\delta) \\ &= 2Cov_{\Phi} - 2Cov_0\rho_0(\delta) - 2Cov_{\Phi}\rho_1(\delta) + 2Cov_0\rho_1(\delta) \\ &= 2Cov_{\Phi}(1 - \rho_1(\delta)) + 2Cov_0(\rho_1(\delta) - \rho_0(\delta)) \end{aligned}$$

Logo,

$$\begin{aligned} Cov[(\Phi(u) - \Phi(u + \delta)), (\Phi(u) - \Phi(u + \delta))] \\ &= 2Cov_{\Phi}(1 - \rho_1(\delta)) + 2Cov_0(\rho_1(\delta) - \rho_0(\delta)) \\ &= 2\Gamma_{\Phi}(\delta) \end{aligned} \quad (2.34)$$

Onde $\Gamma_{\Phi}(\delta)$ é a matriz de variogramas para um $h = \delta$. Multiplicando (2.34) por Cov_{Φ}^{-1} , a equação pode ser expressa em termos dos autovetores esquerdos, \mathbf{a}_i^T que compõe a matriz transformação \mathbf{a} , e dos seus autovalores Λ .

$$\begin{aligned} \mathbf{a}^T(\Gamma_{\Phi}(\delta)Cov_{\Phi}^{-1}) &= (1 - \rho_1(\delta))\mathbf{a}^T + (\rho_1(\delta) - \rho_0(\delta))\mathbf{a}^T Cov_0 Cov_{\Phi}^{-1} \\ &= \left(\frac{\Lambda}{2}\right)\mathbf{a}^T \end{aligned} \quad (2.35)$$

Vetores ortogonais e normalizados formam a matriz \mathbf{a}^T . Multiplicando (2.35) por $Cov_{\Phi}\mathbf{a}$, tem-se a seguinte equação:

$$(1 - \rho_1(\delta))I + (\rho_1(\delta) - \rho_0(\delta))\mathbf{a}^T Cov_0 \mathbf{a} = \frac{\Lambda}{2} \quad (2.36)$$

Lembrando que: \mathbf{a} é a matriz transformação, I a matriz identidade e Λ a matriz de autovalores.

Em Switzer e Green (1984), verifica-se que os autovetores de $2\Gamma_{\Phi}(\delta)\text{Cov}_{\Phi}^{-1}$ são os mesmos que de $\text{Cov}_0\text{Cov}_{\Phi}^{-1}$. Uma vez realizada a transformação dos dados aos fatores MAF, se constata que a matriz de variância-covariância dos fatores obtidos é uma matriz identidade, pois todos os fatores terão variância unitária e covariância nula, devido a sua independência linear (Desbarats & Dimitrakopoulos, 2000).

Como $2\Gamma_{\Phi}(\delta)\text{Cov}_{\Phi}^{-1}$ não é uma matriz simétrica, Switzer e Green (1984) sugerem que a obtenção de seus autovetores seja realizada em três passos:

- i. Realizar uma transformação linear nos dados originais $\Phi(u)$ em $V(u)$, o que pode ser feito por meio de componentes principais realizando a decomposição espectral da matriz Cov_{Φ} ;
- ii. Construir dois conjuntos de dados $[V(u) - V(u + \delta_x)]$ e $[V(u) - V(u + \delta_z)]$, por exemplo, onde δ_x e δ_z são as distancias horizontal e vertical, respectivamente. Calcular as matrizes de variância-covariância para cada um dos conjuntos e calcular a média das duas matrizes, construindo uma nova matriz $\text{Cov}_{V_m}(\delta)$;
- iii. Realizar uma decomposição PCA na matriz $\text{Cov}_{V_m}(\delta)$, sendo esta a solução MAF.

Se as componentes principais $V(u)$ são dadas pela relação $V(u) = W^T \Phi(u)$, sendo $W^T = HD^{-1/2}$, onde H é a matriz de autovetores obtidas na decomposição espectral de Cov_{Φ} e D a matriz de autovalores associados a Cov_{Φ} .

A solução MAF, em termos da matriz de variogramas, é:

$$\begin{aligned} \Gamma_V(\delta) &= \text{Cov}[V(u) - V(u + \delta), V(u) - V(u + \delta)] \\ &= 2(1 - \rho_1(\delta))I + 2(\rho_1(\delta) - \rho_0(\delta))W^T \text{Cov}_0 W \end{aligned} \quad (2.37)$$

de modo que $\Gamma_V(\delta)$ é a matriz de variogramas das diferenças das componentes principais $V(u)$ separadas por $h = \delta$. E a decomposição espectral de $\Gamma_V(\delta)$ é:

$$\Gamma_V(\delta) = C^T \frac{\Lambda}{2} C \quad (2.38)$$

Sendo Λ a matriz de autovalores e C a matriz de autovetores. Inserindo a relação (2.38) na relação (2.37), se tem:

$$\Gamma_V(\delta) = 2(1 - \rho_1(\delta))I + 2(\rho_1(\delta) - \rho_0(\delta))W^T Cov_0 W = C^T \frac{\Lambda}{2} C \quad (2.39)$$

Rearranjando os termos:

$$(1 - \rho_1(\delta))I + 2(\rho_1(\delta) - \rho_0(\delta))C^T W^T Cov_0 W C = \frac{\Lambda}{2} \quad (2.40)$$

A equação (2.40) é equivalente à equação (2.36) se $\mathbf{a} = WC$, assim, a matriz transformação \mathbf{a}^T é dada por: $C^T D^{-1/2} H^T$.

Em Desbarats e Dimitrakopoulos (2000), se indica a normalização dos dados prévia à transformação MAF para, posteriormente, utilizá-los em processos de simulação, porém, a normalização dos dados é desnecessária no que se refere a estimativas, pois, se a transformação é baseada no método de componentes principais e esta não tem a multinormalidade do vetor randômico como exigência, essa é, portanto, dispensável na transformação MAF (Bandarian et. al, 2008).

Em Desbarats e Dimitrakopoulos (2000), é apresentada uma rotina para a obtenção de fatores MAF, dividida nos seguintes passos:

- i. Calcular a matriz de variância-covariância de $\Phi[U_i(u)]$ para $h = 0$, Cov_Φ , realizar sua decomposição espectral em uma matriz de autovetores H com a correspondente matriz de autovalores D , tal que: $Cov_\Phi = HDH^T$;
- ii. Calcular as transformadas PCA $V(u) = W^T \Phi(u)$, sendo $W = HD^{-1/2}$ e $W^T Cov_\Phi W = I$;
- iii. Calcular a matriz do variograma para lag $h \neq 0$, das componentes $V(u)$, ou seja, $\Gamma_V(h \neq 0)$;
- iv. Realizar a decomposição espectral da matriz do variograma $\Gamma_V(h \neq 0)$, com a matriz de autovalores correspondente Λ , tal que: $\Gamma_V(h \neq 0) = C \frac{\Lambda}{2} C^T$;

- v. Finalmente, transforma $\Phi(u)$ em fatores MAF $M(u)$ com a matriz de autovetores \mathbf{a}^T . Sendo $\mathbf{a} = WC$, tal que $M(u) = \mathbf{a}^T \Phi(u)$, onde $M(u)$ são as transformadas MAF.

Uma vez realizada a transformação aos fatores MAF, a krigagem pode ser realizada de maneira independente. Após a krigagem dos fatores $M(u)$, deve-se realizar a retrotransformação ao espaço dos dados originais, que se dá pela relação: $\Phi(u) = (\mathbf{a}^T)^{-1} M(u)$, onde $(\mathbf{a}^T)^{-1}$ é a inversa da matriz transformação dos fatores MAF.

A metodologia pode ser derivada para sistemas mais complexos, com mais de duas estruturas de covariância. Essa situação não será realizada aqui, pois sua derivação pode ser encontrada em Boucher (2003).

Note que, em descorrelação com fatores MAF, se o $\Phi(u)$ puder ser representado pelo modelo de correlação intrínseca, a derivação em componentes principais é suficiente para a descorrelação das variáveis para vetores de separação não nulos (Goovaerts, 1993; Wackernagel, 1998).

Já no caso em que a primeira estrutura de covariância é descrita por efeito pepita puro, a decomposição em fatores MAF irá isolar esta estrutura em transformadas de baixo *rank* (índice i) para pequenos vetores de separação δ . As componentes são derivadas em ordem crescente de autocorrelação espacial, podendo eliminar a estrutura do problema (Desbarats & Dimitrakopoulos, 2000).

Pode-se observar, a partir da prática do método, que a transformação a fatores independentes se dá por meio da decomposição espectral da matriz $\Gamma_V(h \neq 0)$, de modo que seus autovetores igualmente dependem do vetor de separação h (Rondon, 2011). A escolha do vetor de separação h é delicada e ainda não há um procedimento que leve a sua escolha ótima. Em Desbarats (2001), o vetor de separação é selecionado como sendo igual à separação horizontal das amostras, não considerando o espaçamento vertical. Outros autores sugerem diferentes metodologias para tal escolha, porém esta última é a adotada nesta dissertação.

CAPÍTULO 3

METODOLOGIA APLICADA

Neste capítulo, é apresentado o algoritmo desenvolvido para decomposição em fatores independentes, MAF. O algoritmo foi utilizado nos estudos de caso apresentados no capítulo 4.

3.1 ALGORITMO

Para a realização da decomposição MAF, foi construído um algoritmo em *python*®, que pode ser executado dentro do ambiente do *software Sgems*®. O *script* foi desenvolvido em conjunto com Henrique Y. Shibata, doutorando do laboratório LPM.

O *script* importa os dados dos projetos abertos no *Sgems*®, realiza as transformações e envia os dados novamente para o ambiente de trabalho do *Sgems*®, de modo que o usuário pode realizar toda a análise exploratória dos fatores bem como estimativas ou simulações.

O *script* funciona em 17 passos:

- i. Importa as variáveis, $U_i(u)$, determinadas pelo usuário do projeto aberto no *Sgems*®;
- ii. Centraliza os dados nas suas médias;
- iii. Calcula a matriz de variâncias-covariâncias dos dados para $h=0$;
- iv. Realiza a primeira decomposição espectral da matriz de variância-covariâncias dos dados para $h=0$;
- v. Calcula a matriz W a partir da decomposição espectral da matriz de variâncias-covariâncias e sua transposta;
- vi. Obtém as componentes $V(u)$ que são as primeiras transformadas dos dados;
- vii. Calcula distâncias euclidianas entre as amostras nas três direções (X, Y e Z);
- viii. Solicita que determine um valor de h para que a matriz de variâncias-covariância seja calculada para amostras a esta distância;
- ix. Verifica quais amostras estão separadas por esta distância h ;
- x. Monta um novo vetor de amostras que contém somente aquelas que satisfazem a condição da distância;
- xi. Calcula a matriz do variograma $\Gamma_V(h)$;
- xii. Realiza a decomposição espectral da matriz de variogramas para h ;

- xiii. Calcula a matriz A e sua transposta;
- xiv. Obtém as componentes M(u) transformadas, que são as componentes MAF;
- xv. Calcula a matriz inversa da matriz A;
- xvi. Calcula a retrotransformação;
- xvii. Exporta os valores MAF ao ambiente *Sgems*®.

O *script* será exposto a fim de fornecer mais clareza do seu funcionamento em cada um dos passos.

```

import sgems
import numpy as np
import math
import scipy
from scipy import stats

coordx = sgems.get_property("sap", "_X_")
coordx1 = sgems.get_property("sap", "_X_")
coordy = sgems.get_property("sap", "_Y_")
coordy1 = sgems.get_property("sap", "_Y_")
coordz = sgems.get_property("sap", "_Z_")
coordz1 = sgems.get_property("sap", "_Z_")

t1 = sgems.get_property("sap", "FET")
t2 = sgems.get_property("sap", "SIT")

#k1 = sgems.get_property("grid", "MAF1")
#k2 = sgems.get_property("grid", "MAF1_0")

n = len(t1)-1

dist_min = 0

dist_max = 200

z = []

for i in range(len(t1)):
    a1 = (t1 - np.average(t1)) # d e f = a1, a2
    i += 1

print "\n", np.average(t1)

for i in range(len(t2)):
    a2 = (t2 - np.average(t2))
    i += 1

```

```

print "\n", np.average(t2)
# faz a covariancia dos valores
cov = np.cov([a1], [a2])

print cov

# calcule os autovetores e autovalores da covariancia
ev = np.linalg.eig(cov)
evals, evecs = ev

print "\n Eigenvalues: \n", evals

print "\n Eigenvectors: \n", evecs

diag_evals = np.diag(evals,k=0)

from scipy import linalg
inv = linalg.inv(diag_evals)

vec = inv.flatten(1)

print vec

for i in vec:
    raiz = math.sqrt(i)
    z.append(raiz)

print z

ma_inv = np.reshape(z, (2,-1))

print ma_inv

W = np.dot(evecs,ma_inv)
# mult = W
print W

mt1 = np.transpose(W)

print mt1
#Calcula a primeira fatorização: componentes principais
#-----
--

valor1 = mt1[0][0]
valor2 = mt1[0][1]

for i in range(len(t2)):
    v1 = ((np.dot(valor1, a1))+(np.dot(valor2, a2)))

valor3 = mt1[1][0]
valor4 = mt1[1][1]

for i in range(len(t1)):
    v2 = ((np.dot(valor3, a1))+(np.dot(valor4, a2)))

#vetores para o cálculo da distância da diferença
#-----
b = []

```

```

h = []
p = []
o = []
q = []
dif1 = []
dif2 = []
dif12 = []
u = []
r = []

#Cálculo da distância
#-----
for k,l in enumerate(range(0,n,1)):
    for i,j in enumerate(range(0,n,1)):

        zx = (coordx[k] - coordx1[i+k])**2

        b.append(zx)

        zy = (coordy[l] - coordy1[j+1])**2

        h.append(zy)

        zz = (coordz[l] - coordz1[j+1])**2

        u.append(zz)

        soma = zx + zy + zz
        dist = math.sqrt(soma)

        if dist_min<= dist <= dist_max:

            dif1 = v1[k] - v1[i+k]
            dif2 = v2[k] - v2[i+k]
            dif12 = v1[k] - v2[i+k]
            p.append(dif1)
            o.append(dif2)
            q.append(dif12)

        if coordx1[i+k] == coordx1[n]:break

#Cálculo da matriz do semivariograma das amostras separadas por dist
#-----
for i in range(len(p)):
    quad = p[i]**2
    r.append(quad)

soma_dif1 = np.sum(r)

d11 = soma_dif1/(2*len(p))
s11 = 2*d11
s = []

for i in range(len(o)):
    quad2 = o[i]**2
    s.append(quad2)

```

```

soma_dif2 = np.sum(s)

d22 = soma_dif2/(2*len(o))
s22 = 2*d22

t = []
for i in range(len(q)):
    quad12 = q[i]**2
    t.append(quad12)

soma_dif12 = np.sum(t)

d12 = soma_dif12/(2*len(q))
s12 = 2*d12

print d11
mtcov = np.array([[s11], [s12], [s12],[s22]])

semi = np.reshape(mtcov,(2,-1))

print semi

#-----
#Diagonalização da matriz dos semivariogramas separados por dist
#-----
evsemi = np.linalg.eig(semi)
evals_semi,  evecs_semi = evsemi

A = np.dot(W,evecs_semi)

#print a

mt2 = np.transpose(A)
#Cálculo da transposta de A: matriz transformação
trans = np.reshape(mt2,(2,-1))

print "Matriz transformacao: \n",trans
#Cálculo dos fatores MAF
#-----
valor1 = trans[0][0]
valor2 = trans[0][1]

for i in range(len(t2)):
    m1 = ((np.dot(valor1, t1))+np.dot(valor2, t2)))

valor3 = trans[1][0]
valor4 = trans[1][1]

for i in range(len(t1)):
    m2 = ((np.dot(valor3, t1))+np.dot(valor4, t2)))

#Cálculo da inversa da matriz transformação: realiza a retrotransformação
trans_inv = linalg.inv(trans)

#Cálculo dos valores retrotransformados

```

```

#-----
-----
valor1 = trans_inv[0][0]
valor2 = trans_inv[0][1]

for i in range(len(t2)):
    r1 = ((np.dot(valor1, m1))+(np.dot(valor2, m2)))

valor3 = trans_inv[1][0]
valor4 = trans_inv[1][1]

for i in range(len(t1)):
    r2 = ((np.dot(valor3, m1))+(np.dot(valor4, m2)))

#Envia os valores como propriedades no projeto do Sgems
#-----
-----

sgems.set_property("sap","MAF1", m1.tolist())
sgems.set_property("sap","MAF2", m2.tolist())

sgems.set_property("grid","retroFET", r1.tolist())
sgems.set_property("grid","retroSIT", r2.tolist())

```

No início da rotina (parte em destaque vermelho), há uma observação a ser feita: as entradas k1 e k2 são utilizadas para importar dados que já tenham sido trabalhados com os fatores MAF, ou seja, estas variáveis estão em algum modelo de blocos após terem sido krigadas ou simuladas. É necessário que o algoritmo já tenha sido aplicado, ao menos, uma vez nos dados. Caso seja a primeira vez que o algoritmo esteja sendo utilizado e o usuário ainda não possua os fatores MAF, e ainda não tenha os resultados da krigagem ou da simulação, basta comentar as entradas que o algoritmo realizará as transformações normalmente.

No parte final da rotina (destacada em vermelho), caso no início da rotina não tenha as variáveis em um modelo de blocos (a linha tenha sido comentada pelo usuário), na retrotransformação, em que há a entrada dos valores k1 e k2, basta o usuário substituir por m1 e m2, que são os componentes MAF, que será realizada a retrotransformação direta dos fatores aos dados originais. Este é um teste que servirá de verificação ao funcionamento do algoritmo.

De acordo com o funcionamento do algoritmo, o usuário terá, ao fim da primeira aplicação, os fatores MAF e a retrotransformação direta aos dados originais. Na segunda

aplicação do algoritmo, o usuário irá importar, além dos dados originais, o modelo de blocos com as estimativas dos fatores MAF ou de suas simulações, e obterá, ao fim, a retrotransformação das estimativas ou simulações ao espaço dos dados originais.

No capítulo 4, quando utilizado o algoritmo no banco de dados, todas as saídas possíveis estão apresentadas em cada estudo de caso.

É importante ressaltar que, pelo fato do método se basear em multiplicações de matrizes e decomposições espectrais, os vetores, tanto para transformação quanto para retrotransformação, devem ter a mesma dimensão, caso contrário, a multiplicação matricial não é executada. Esta condição faz com que o usuário deva ter cuidado em que os dados a serem importados ao algoritmo, para a realização das transformações lineares do método MAF, tenham configuração isotópica, ou seja, todas as variáveis serão amostradas nos mesmos locais. Na configuração isotópica, todas as variáveis são representadas por vetores de mesma dimensão.

No caso de banco de dados não isotópicos, o usuário pode contornar o problema de diversas maneiras, eliminando os locais amostrais que não contenham todas as variáveis, por exemplo. Também pode se realizar a krigagem ordinária das variáveis para gerar a isotopia. De qualquer modo, esta é uma decisão que deve ser tomada levando em conta os ganhos e as perdas que cada opção irá gerar.

Na retrotransformação dos dados, a mesma condição de isotopia se faz necessária, desta maneira, em casos em que foi realizada uma estimativa, por exemplo, estas devem ser realizadas utilizando a mesma estratégia de busca. Deste modo, os locais estimados serão os mesmos e a retrotransformação poderá ser realizada.

É também necessário que o usuário determine o tamanho do *lag* utilizado para determinação da matriz dos variogramas $\Gamma_V(h)$ e informe no início da rotina. O critério adotado nos estudos de caso do capítulo 4 é baseado no espaçamento amostral. De modo que são realizados diversos testes com *lags* múltiplos ao espaçamento amostral e observada a matriz transformação resultante. Quando a matriz transformação apresenta estabilidade, ou seja, a partir de um determinado *lag*, não há mudança significativa nos seus termos, este *lag* é o definido como a separação para a qual será realizada a decomposição MAF. Outros métodos de determinação podem ser utilizados. Recomenda-se Fonseca e Dimitrakopoulos (2003),

porém não é fornecida ainda uma metodologia que determine o vetor de separação demonstrada matematicamente.

CAPÍTULO 4

ESTUDOS DE CASO

Este capítulo apresenta um estudo de caso realizado no banco de dados bidimensional extraído de Goovaerts (1997), o Jura, e outro em um banco de dados tridimensional proveniente da companhia Vale S.A, Sapecado. Em ambos, é aplicada a metodologia MAF proposta nesta dissertação, juntamente com a realização da cokrigagem, método clássico multivariado utilizado para fins de referência na análise de resultados.

4.1 ESTUDO DE CASO 1 – BANCO DE DADOS JURA

Nesta seção, é apresentado um estudo de caso introdutório, utilizando o banco de dados presente em Goovaerts (1997), o Jura. O banco foi coletado pelo Instituto Suíço Federal de Tecnologia. Os dados foram amostrados em 359 pontos, em que foram medidas as concentrações de sete metais pesados: Cádmio, Cobalto, Cromo, Cobre, Níquel, Zinco e Chumbo. Este banco com 359 amostras foi dividido em dois grupos: o primeiro com 100 amostras, chamado de *Validation set*, o segundo: com 259 amostras, chamado de *Prediction set*.

No estudo de caso, foi utilizado o segundo subconjunto, *Prediction set*, levando em conta somente 4 metais: Cádmio, Cobalto, Cromo e Níquel.

Foi realizada a solução clássica ao problema: cokrigagem e o método alternativo proposto: a decorrelação das variáveis prévia à estimativa por meio do método MAF, com krigagem ordinária, chamado de KMAF.

4.1.1 ANÁLISE EXPLORATÓRIA DOS DADOS

A seguir, serão apresentadas as análises exploratórias dos dados. O banco é bidimensional, e consiste em 259 amostras dos metais Cádmio, Cromo, Cobalto e Níquel. O banco de dados estudado é isotópico. A distribuição das amostras na malha amostral não apresenta agrupamentos significantes, mas pequenas áreas com amostragem preferencial. A malha amostral tem espaçamento médio de 200m na direção X e 200m na direção Y como observado na figura 4.1.

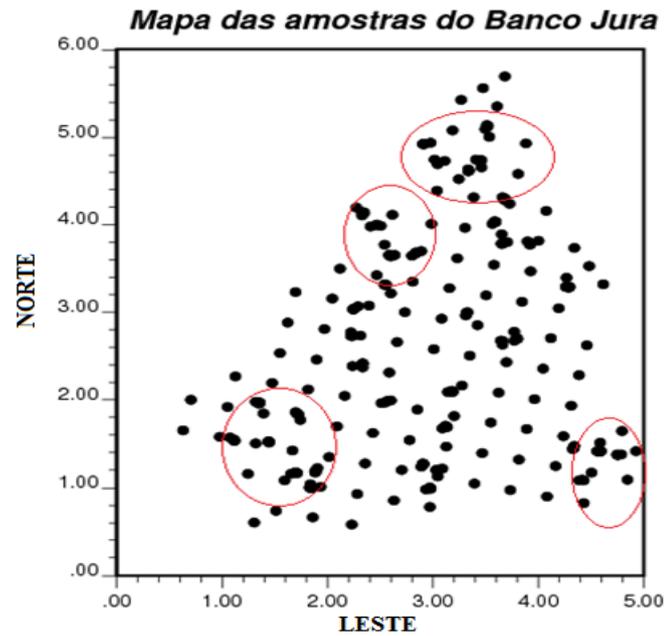


Figura 4. 1 Mapa amostral do banco de dados Jura (Goovaerts, 1997) com os agrupamentos amostrais destacados

Não obstante, o desagrupamento foi realizado por meio do método de polígonos de Voronoi (Dirichlet, 1850). Foi utilizado um modelo para realizar o desagrupamento com células de dimensões de 10m na direção X e 10m na direção Y. O modelo possui 285000 blocos. Os histogramas agrupados e desagrupados que foram obtidos a partir de cada variável serão apresentados nas figuras a seguir.

As figuras 4.2 e 4.3 apresentam os histogramas, agrupado e desagrupado, respectivamente, da variável Cádmio.

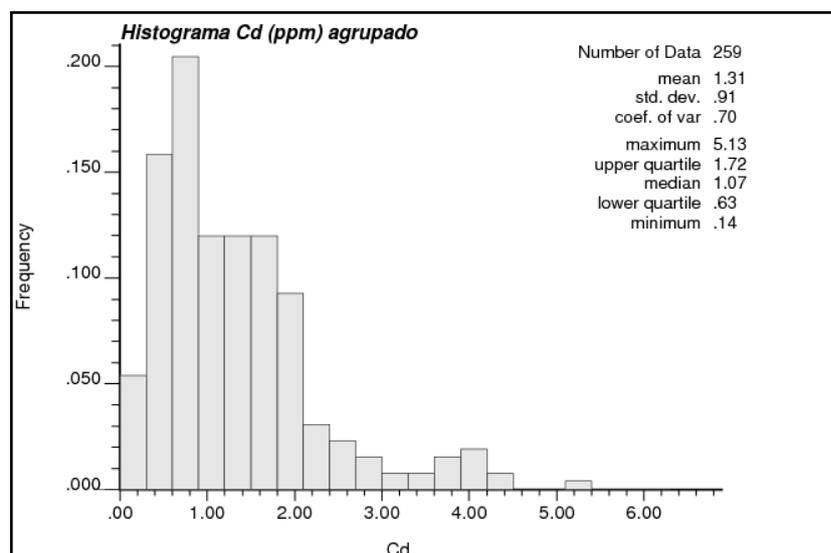


Figura 4. 2 - Histograma da variável Cádmio agrupado. Valores de concentração do metal em (ppm)

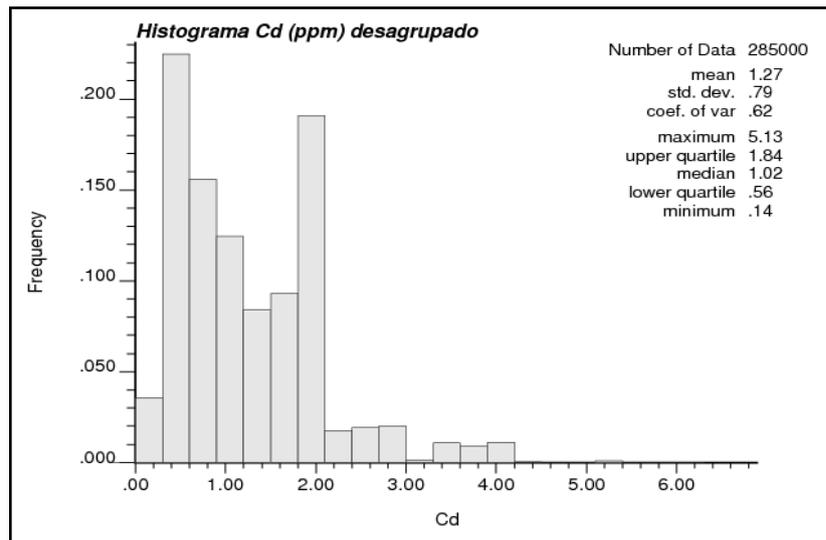


Figura 4.3 - Histograma da variável Cádmio desagrupado. Concentração do metal em (ppm). Desagrupamento realizado por polígonos de Vornoi (Dirichlet, 1850)

Nos histogramas expostos nas figuras 4.2 e 4.3, observa-se que a média global dos dados agrupados é 1,31(ppm) e diminui para 1,27(ppm) após o desagrupamento. Há uma redução de 3,05% na média global dos dados. Apesar de serem observadas duas populações no histograma da figura 4.3, o primeiro pico observado corresponde a amostras com teores em torno de 1 (ppm) de Cádmio, que se encontram em pequenos agrupamentos na malha amostral, este fato não prejudicou a validação dos resultados, como será mostrado no item 4.1.5 deste capítulo.

Nas figuras 4.4 e 4.5, são apresentados os histogramas da variável Cobalto, agrupado e desagrupado, respectivamente.

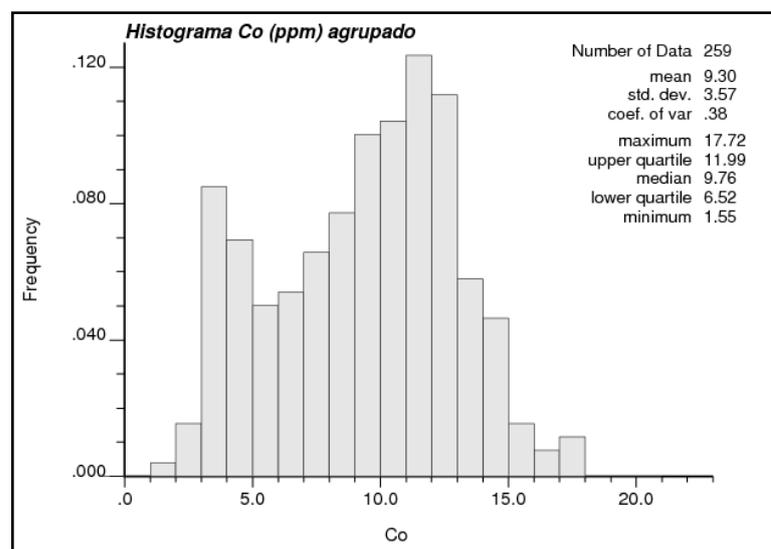


Figura 4.4 - Histograma da variável Cobalto agrupado. Concentração do metal em (ppm)

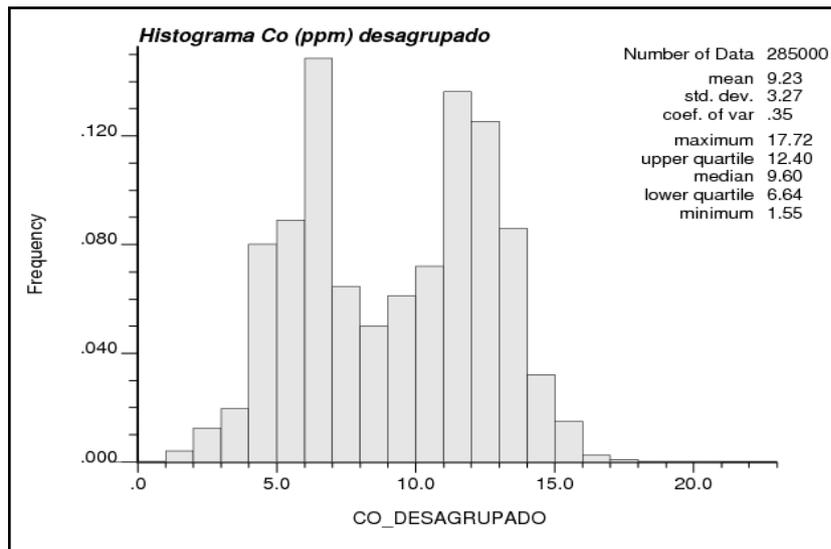


Figura 4.5 - Histograma da variável Cobalto desagrupado. Concentração do metal em (ppm). Desagrupamento realizado por polígonos de Vornoi (Dirichlet, 1850)

No gráfico da figura 4.4, vê-se que a média global dos dados agrupados é 9,30 (ppm), e, na figura 4.5, após o desagrupamento dos dados a média diminui, assim como para o Cádmio. A média desagrupada do Cobalto tem valor de 9,23 (ppm), a redução é de 0,75%.

O gráfico da figura 4.5 também aparenta duas populações que se mostram nos mapas amostrais, expostos posteriormente na análise dos resultados deste estudo de caso, com pequenos agrupamentos de amostras de baixos teores, porém, como poderá ser observado no trabalho que segue, este fato não influenciará na validação dos resultados.

Nas figuras 4.6 e 4.7, são apresentados os histogramas da variável Cromo, agrupado e desagrupado, respectivamente.

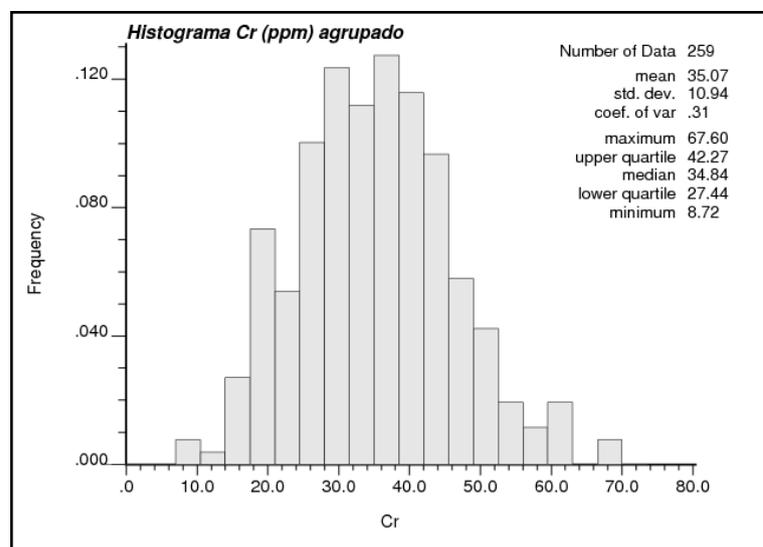


Figura 4.6 - Histograma agrupado da variável Cromo. Concentração do metal em (ppm)

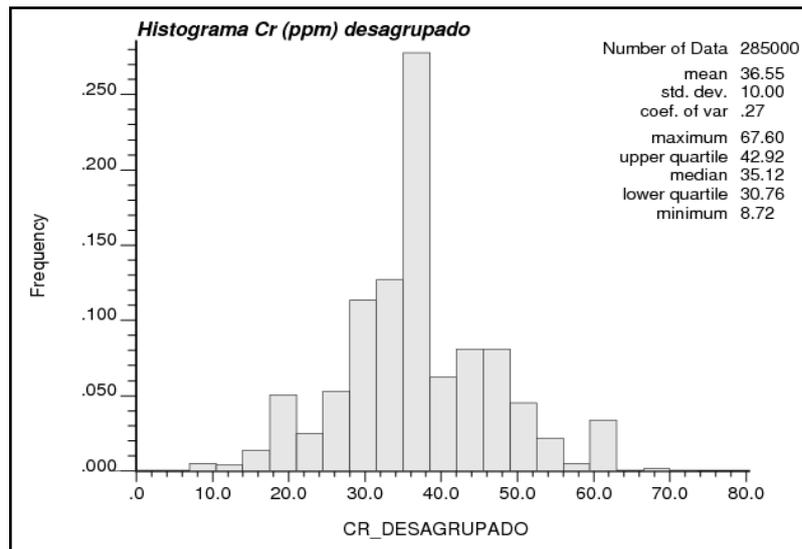


Figura 4. 7 - Histograma da variável Cromo desagrupado. Concentração do metal em (ppm). Desagrupamento realizado por polígonos de Vornoi (Dirichlet, 1850)

No histograma da figura 4.6, a média da variável Cromo agrupada tem valor de 35,07(ppm), já no histograma da figura 4.7, a média apresentada pelos dados, após o desagrupamento, é 36,55 (ppm).

Finalmente, os histogramas da variável Níquel, agrupado e desagrupado, são apresentados nas figuras 4.8 e 4.9, respectivamente.

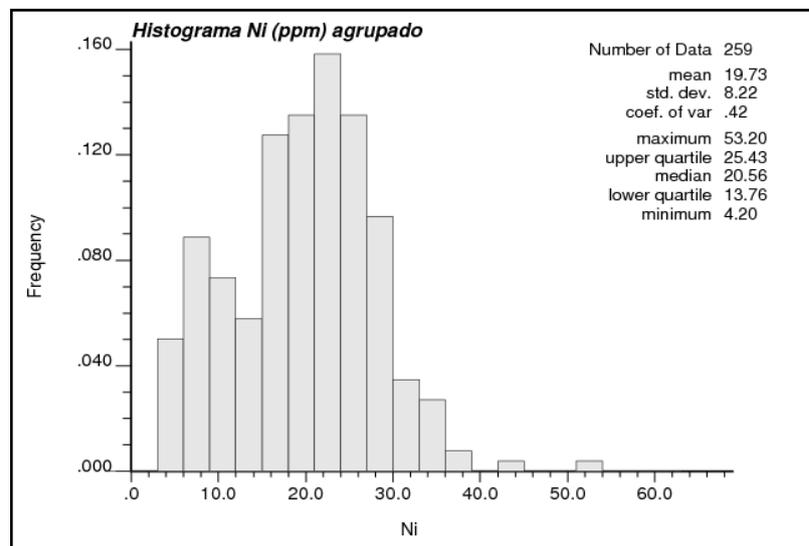


Figura 4. 8 - Histograma agrupado da variável Níquel. Concentração do metal em (ppm)

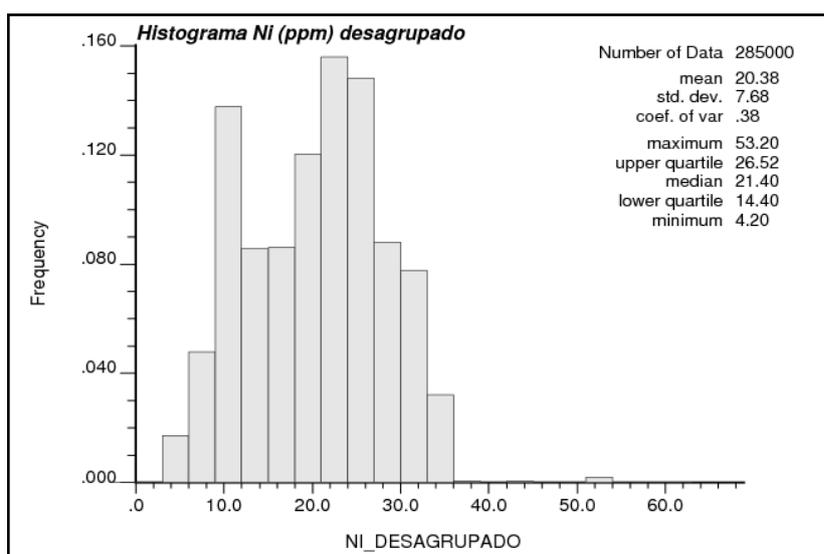


Figura 4.9 - Histograma da variável Níquel desagrupado. Concentração do metal em (ppm). Desagrupamento realizado por polígonos de Voronoi (Dirichlet, 1850)

Na figura 4.8, na qual é apresentado o histograma agrupado da variável Níquel, a média global é de 19,73 (ppm). Já no histograma desagrupado, apresentado na figura 4.9, a média global dos dados é 20,38 (ppm). Desta forma, pode-se ver que, assim como a variável Cromo, a média desagrupada é mais alta que a média agrupada. O aumento na média do Níquel é de 3,2%. O sumário estatístico das variáveis agrupadas e desagrupadas são apresentados na tabela 4.1.

Tabela 4.1 - Sumário estatístico agrupado e desagrupado do conjunto de variáveis do Jura (Goovaerts, 1997). Unidades dos atributos: concentração em (ppm).

	Cádmio	Cobalto	Cromo	Níquel
Média agr.	1,31	9,30	35,07	19,73
Média desagr.	1,27	9,23	36,55	20,38
Desvio padrão agrup.	0,91	3,57	10,94	8,22
Desvio padrão desgr.	0,79	3,27	10,0	7,68
Máximo	5,13	17,72	67,6	53,2
Mínimo	0,14	1,55	8,72	4,2
Mediana agrup.	1,07	9,76	34,84	20,56
Mediana desagr.	1,02	9,6	35,12	21,4
Quartil superior agrup.	1,72	11,99	42,27	25,43
Quartil superior desagr.	1,84	12,4	42,92	26,52

Quartil inferior agrup.	0,63	6,52	27,44	13,76
Quartil inferior desag.	0,56	6,64	30,76	14,4

O Jura (Goovaerts, 1997) é um banco de dados em que as variáveis são correlacionadas entre si. Estas correlações são visualizadas nos gráficos de dispersão dos atributos. Na figura 4.10, pode-se verificar correlação entre a variável Cádmio e a variável Cobalto.

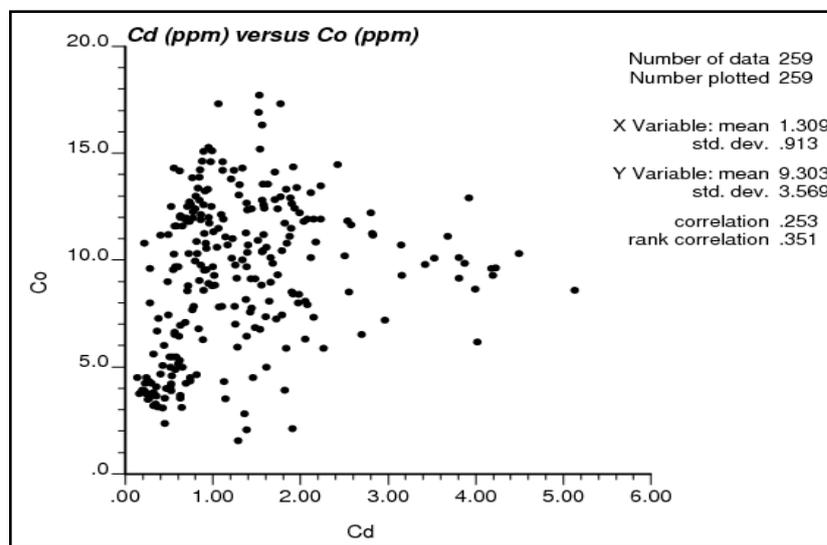


Figura 4. 10 - Gráfico de dispersão entre a variável Cádmio (ppm) e a variável Cobalto (ppm)

O coeficiente de correlação entre Cádmio e Cobalto é $\rho_{CdCo} = 0,25$, o que representa fraca correlação entre as variáveis, isso pode ser verificado na grande dispersão da nuvem de pontos apresentada no gráfico.

Na figura 4.11, é apresentado o gráfico de dispersão entre Cádmio e Cromo.

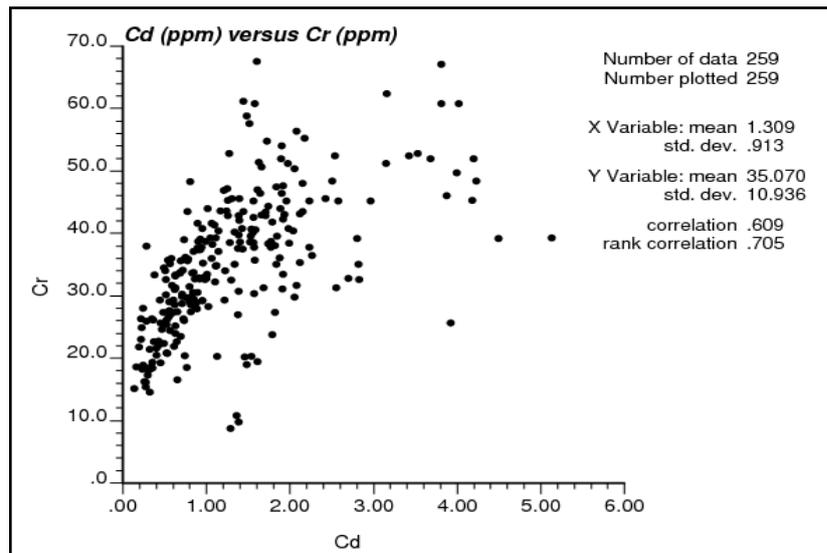


Figura 4. 11 - Gráfico de dispersão entre a variável Cádmio (ppm) e a variável Cromo (ppm)

A correlação entre Cádmio e Cromo é $\rho_{CdCr} = 0,60$, mais significativa que a verificada entre Cádmio e Cobalto. Vê-se, na figura 4.11, que a nuvem de pontos é menos dispersa do que a do gráfico apresentado na figura 4.10. Os dois atributos, neste caso, são mais correlacionados entre si.

Na figura 4.12, é apresentado o gráfico de dispersão entre o atributo Cádmio e Níquel.

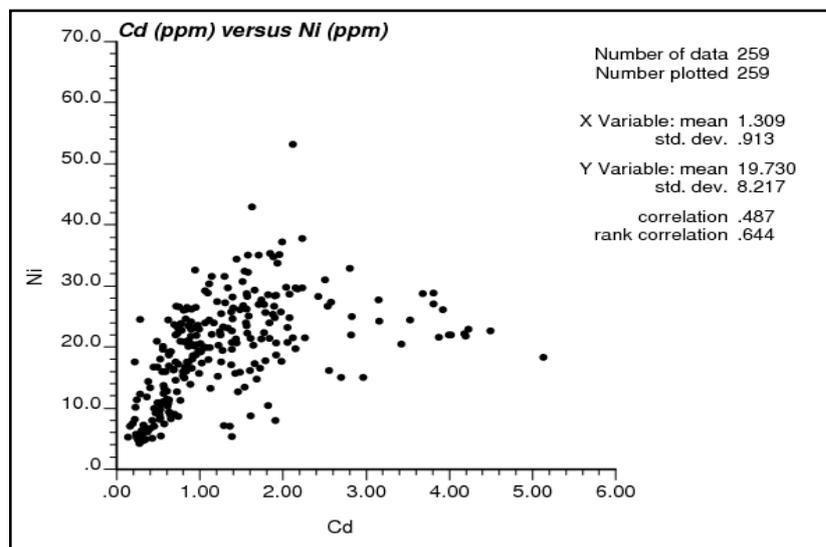


Figura 4. 12 - Gráfico de dispersão entre a variável Cádmio (ppm) e Níquel (ppm)

Quando se observa o gráfico da figura 4.12, não se vê uma nuvem de pontos tão dispersa quanto o da figura 4.10, porém é mais dispersa que o gráfico da figura 4.11. A correlação entre Cádmio e Níquel é $\rho_{CdNi} = 0,48$. Em relação às correlações restantes do banco de dados, pode-se dizer que esta é uma correlação intermediária.

Na figura 4.13, é apresentado o gráfico de dispersão entre Cobalto e Cromo.

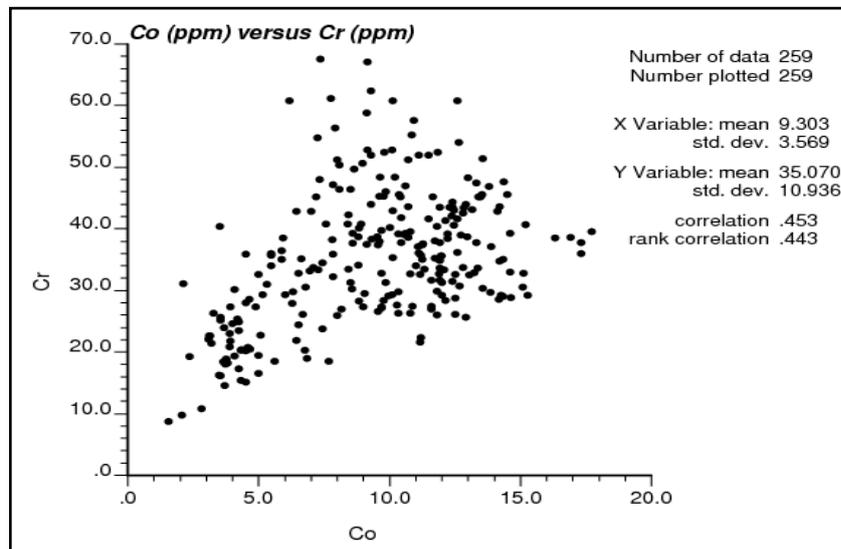


Figura 4.13 - Gráfico de dispersão entre as variáveis, Cobalto (ppm) e Cromo (ppm)

Na figura 4.13, se observa um gráfico de dispersão semelhante ao exposto na figura 4.12. Isto é devido ao valor de correlação ser próximo ao encontrado no par plotado na figura 4.12, sendo $\rho_{CoCr} = 0,45$ para o par de variáveis Cobalto e Cromo.

Na figura 4.14, é apresentado o gráfico de dispersão entre as variáveis, Cobalto e Níquel.

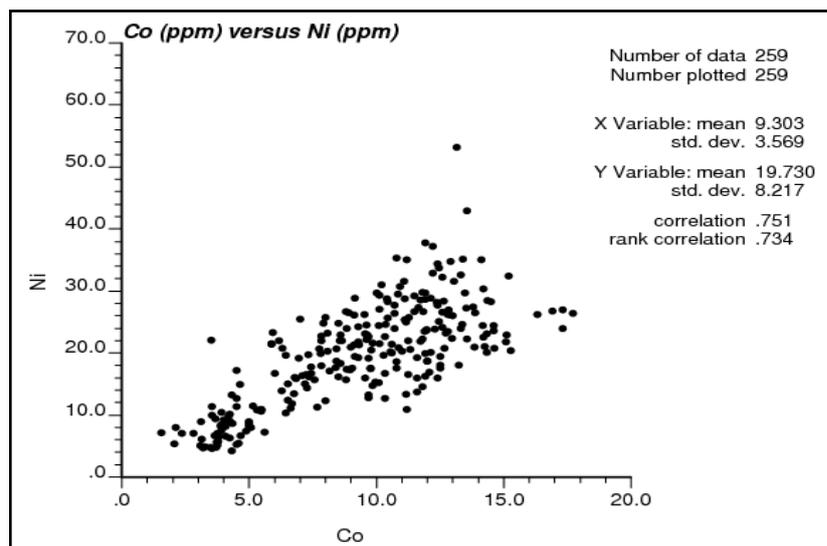


Figura 4.14 - Gráfico de dispersão entre as variáveis Cobalto (ppm) e Níquel (ppm)

Na figura 4.14, é verificada a correlação entre Cobalto e Níquel, $\rho_{CoNi} = 0,75$. Esta correlação é a mais forte apresentada entre os atributos do banco de dados. A nuvem de pontos mostra a correlação entre as duas variáveis. Finalmente, na figura 4.15, é apresentado

o gráfico de dispersão entre as variáveis, Cromo e Níquel, finalizando as análises de correlações entre as variáveis existentes no banco de dados.

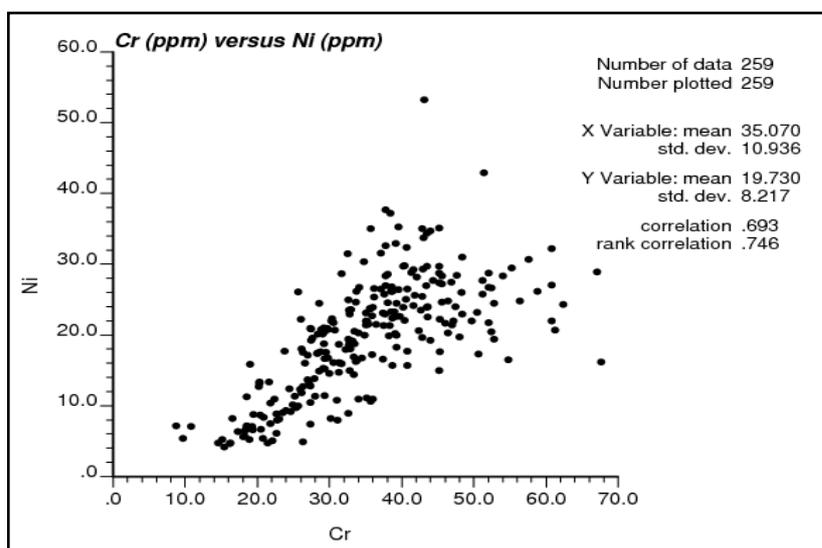


Figura 4. 15 - Gráfico de dispersão entre as variáveis, Cromo (ppm) e Níquel (ppm)

A correlação entre Cromo e Níquel é $\rho_{CrNi} = 0,69$, em relação às outras correlações existentes no banco de dados é forte.

Frente às correlações apresentadas, foram aplicadas, no banco de dados, duas soluções: a primeira: cokrigagem; a segunda: Krigagem de Mínimos-Máximos Fatores de Autocorrelação (KMAF).

4.1.2 SOLUÇÃO VIA COKRIGAGEM

A estimativa foi realizada em suporte de blocos em um modelo em que cada bloco possui dimensão de 100m na direção X por 100m na direção Y, correspondentes à metade do espaçamento amostral em cada direção. O modelo possui 50 células na direção X e 57 células na direção Y.

O sistema de cokrigagem inclui as quatro variáveis selecionadas para o estudo de caso. Devido ao grau de complexidade envolvido na modelagem dos variogramas, este foi realizado com o auxílio de *software Isatis®*, com o qual os variogramas diretos e cruzados podem ser modelados de maneira semiautomática, satisfazendo o modelo linear de correionalização (MLC). Os parâmetros para a construção dos variogramas direcionais do sistema de cokrigagem estão listados na tabela 4.2.

Tabela 4. 2 - Parâmetros usados para construção dos variogramas experimentais direcionais do sistema de cokrigagem.

Variogramas direcionais no plano XY	
Número de lags	20
Tamanho de lag	100m
Tolerância de lag	50m
Largura de banda	100m
Número de direções	8
Tolerância angular	22,5°

Como a variável Cobalto apresentou um variograma experimental bem definido, esta serviu de base para a construção de todos os modelos variográficos. O variograma do Cobalto apresentou maior continuidade espacial na direção N45, sendo $dip = 0$, e continuidade mínima espacial na direção perpendicular, N135 com $dip = 0$. Sendo assim, todos os variogramas diretos e cruzados foram modelados nestas direções. O alcance máximo encontrado foi de 3000m e o mínimo de 1200m.

No modelo utilizado, nas figuras 4.16 e 4.17, foi ajustado um *sill* maior que a variância a priori dos dados para satisfazer a condição de definição positiva.

Os variogramas construídos para o atributo Cobalto nas direções N45 e N135 podem ser vistos nas figuras 4.16 e 4.17.

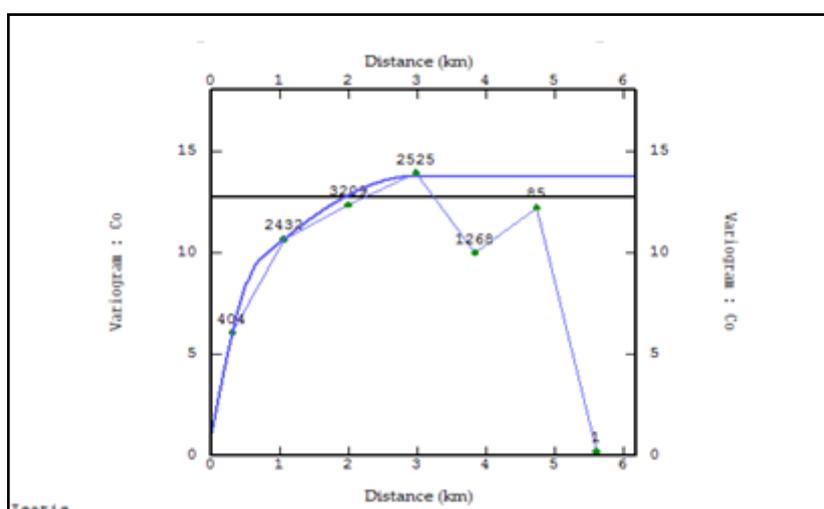


Figura 4. 16 - Variograma experimental do atributo Cobalto na direção N45; $dip = 0$. Alcance da primeira estrutura esférica é 700m. Alcance da segunda estrutura esférica é 3000m

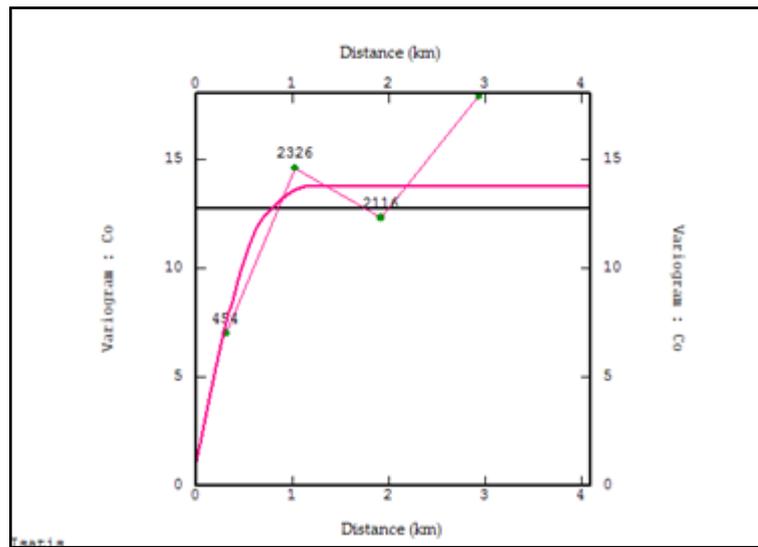


Figura 4.17 - Variograma experimental do atributo Cobalto na direção N135; $dip = 0$. Alcance da primeira estrutura esférica é 700m. Alcance da segunda estrutura esférica é 1200m

Os variogramas expostos nas figuras 4.16 e 4.17 foram os utilizados como base na construção do modelo linear de correionalização. Cada variograma das figuras acima possui três estruturas, duas com modelo esférico, além do efeito pepita. Todos os variogramas diretos e cruzados das variáveis restantes foram construídos de acordo com os das figuras 4.16 e 4.17. Os alcances, quantidade e tipo de estruturas foram mantidos os mesmos, alterando somente a contribuição de cada estrutura ao *sill* total e o efeito pepita para cada variável. As alterações foram feitas de modo que fosse satisfeita a condição de definição positiva do MLC.

Para exemplificar os variogramas cruzados, são apresentados nas figuras 4.18 e 4.19 os variogramas cruzados entre as variáveis Cádmiio e Cobalto nas direções N45 e N135, ambas com $dip = 0$.

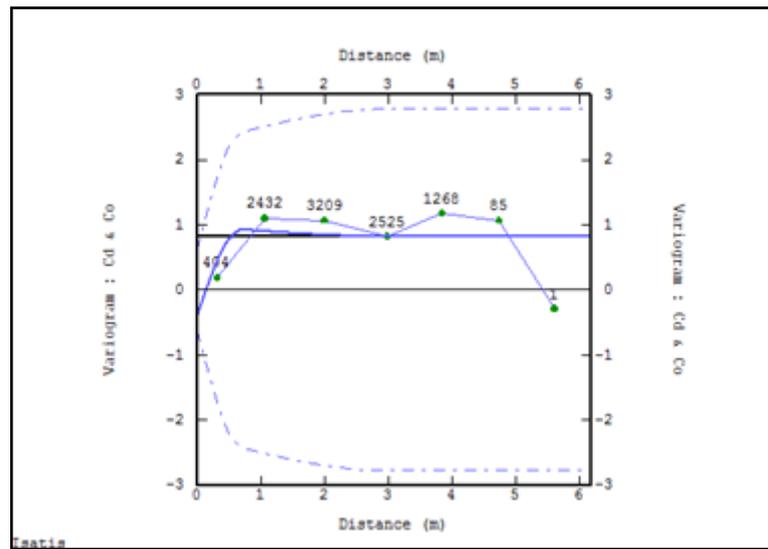


Figura 4.18 - Variograma experimental cruzado das variáveis Cádmi e Cobalto na direção N45; $dip = 0$. Alcance da primeira estrutura esférica é 700m. Alcance da segunda estrutura esférica é 3000m

Na figura 4.18, está exposto o variograma experimental cruzado entre as variáveis Cádmi e Cobalto, a fim de exemplificar o ajuste variográfico aos patamares negativos apresentados nas tabelas 4.3 a 4.5.

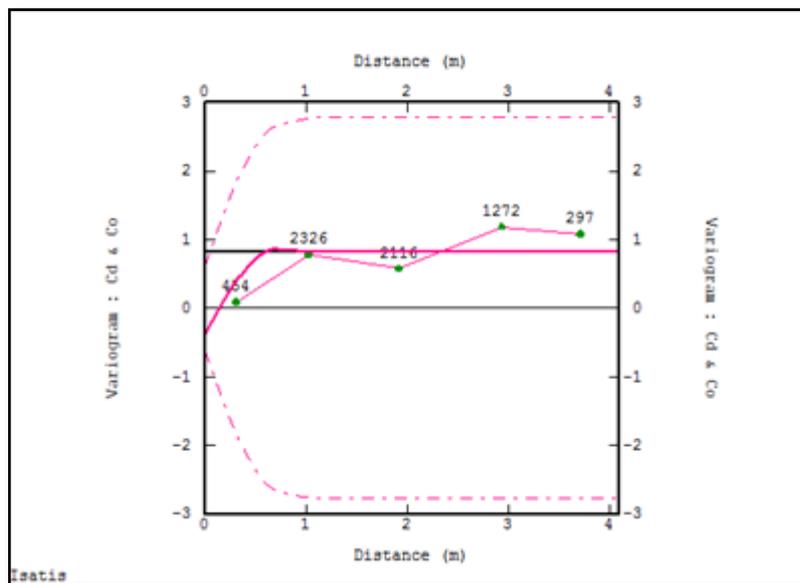


Figura 4.19 - Variograma experimental cruzado das variáveis Cádmi e Cobalto na direção N135; $dip = 0$. Alcance da primeira estrutura esférica é 700m. Alcance da segunda estrutura esférica é 1200m

No variograma apresentado na figura 4.18, assim como no da figura 4.19, o efeito pepita modelado tem valor de -0,3857 (tabela 4.3), a contribuição da primeira estrutura esférica tem valor de 1,361 (tabela 4.4) e a contribuição da segunda estrutura esférica tem valor de -0,1583 (tabela 4.5).

Sendo Cádmio, Cobalto, Cromo e Níquel as variáveis do sistema de cokrigagem, a matriz de correionalização do efeito pepita, C_0 exposta na tabela 4.3, é:

Tabela 4. 3 - Matriz de correionalização do efeito pepita, C_0 , para o sistema de cokrigagem. Os termos cruzados constituem o efeito pepita dos variogramas cruzados, bem como os termos diretos constituem o efeito pepita dos variogramas diretos.

C_0	Cádmio	Cobalto	Cromo	Níquel
Cádmio	0,3876	-0,387	1,875	-0,2461
Cobalto	-0,3857	1,017	-3,663	-1,01
Cromo	1,875	-3,663	43,99	3,121
Níquel	-0,2461	-1,01	3,121	2,661

Assim como há uma matriz de correionalização para o efeito pepita, há também para a primeira estrutura esférica do modelo variográfico. A matriz correspondente às contribuições da primeira estrutura esférica do modelo de correionalização, C_1 , é mostrada na tabela 4.4.

Tabela 4. 4 - Matriz de correionalização da primeira estrutura esférica, C_1 , para o sistema de cokrigagem. Os termos cruzados constituem a contribuição da primeira estrutura esférica dos variogramas cruzados, bem como os termos diretos constituem a contribuição da primeira estrutura esférica dos variogramas diretos.

C_1	Cd	Co	Cr	Ni
Cd	0,403	1,361	4,598	4,085
Co	1,361	6,507	21,95	17,47
Cr	4,598	21,95	75,91	63,61
Ni	4,085	17,47	63,61	60,24

Desta forma, a matriz de correionalização para as contribuições da segunda estrutura esférica do modelo, C_2 , é exposta na tabela 4.5.

Tabela 4. 5 - Matriz de correionalização da segunda estrutura esférica, C_2 , para o sistema de cokrigagem. Os termos cruzados constituem a contribuição da segunda estrutura esférica dos variogramas cruzados, bem como os termos diretos constituem a contribuição da segunda estrutura esférica dos variogramas diretos.

C_2	Cd	Co	Cr	Ni
Cd	0,04561	-0,1583	0,363	-0,03317
Co	-0,1583	6,223	0,5898	6,994

Cr	-0,363	0,5898	3,292	-1,019
Ni	-0,03317	6,994	-1,019	9,05

Os modelos variográficos finais são os apresentados nas equações de 4.1. a 4.10.

$$\gamma_{CdCd} = -0,3876 + 0,403Sph\left(\frac{700m}{N45} \frac{700m}{N135}\right) + 0,04561Sph\left(\frac{3000m}{N45} \frac{1200m}{N135}\right) \quad (4.1)$$

$$\gamma_{CdCo} = -0,3857 + 1,361Sph\left(\frac{700m}{N45} \frac{700m}{N135}\right) - 0,1583Sph\left(\frac{3000m}{N45} \frac{1200m}{N135}\right) \quad (4.2)$$

$$\gamma_{CdCr} = 1,875 + 4,598Sph\left(\frac{700m}{N45} \frac{700m}{N135}\right) - 0,363Sph\left(\frac{3000m}{N45} \frac{1200m}{N135}\right) \quad (4.3)$$

$$\gamma_{CdNi} = -0,2461 + 4,085Sph\left(\frac{700m}{N45} \frac{700m}{N135}\right) - 0,03317Sph\left(\frac{3000m}{N45} \frac{1200m}{N135}\right) \quad (4.4)$$

$$\gamma_{CoCo} = 1,017 + 6,507Sph\left(\frac{700m}{N45} \frac{700m}{N135}\right) + 6,223Sph\left(\frac{3000m}{N45} \frac{1200m}{N135}\right) \quad (4.5)$$

$$\gamma_{CoCr} = -3,663 + 21,95Sph\left(\frac{700m}{N45} \frac{700m}{N135}\right) + 0,5898Sph\left(\frac{3000m}{N45} \frac{1200m}{N135}\right) \quad (4.6)$$

$$\gamma_{CoNi} = -1,01 + 17,47Sph\left(\frac{700m}{N45} \frac{700m}{N135}\right) + 6,994Sph\left(\frac{3000m}{N45} \frac{1200m}{N135}\right) \quad (4.7)$$

$$\gamma_{CrCr} = 43,99 + 75,91Sph\left(\frac{700m}{N45} \frac{700m}{N135}\right) + 3,292Sph\left(\frac{3000m}{N45} \frac{1200m}{N135}\right) \quad (4.8)$$

$$\gamma_{CrNi} = 3,121 + 63,61Sph\left(\frac{700m}{N45} \frac{700m}{N135}\right) - 1,019Sph\left(\frac{3000m}{N45} \frac{1200m}{N135}\right) \quad (4.9)$$

$$\gamma_{NiNi} = 2,661 + 60,24Sph\left(\frac{700m}{N45} \frac{700m}{N135}\right) + 9,05Sph\left(\frac{3000m}{N45} \frac{1200m}{N135}\right) \quad (4.10)$$

A coestimativa foi realizada em blocos no *software Isatis®* utilizando os seguintes parâmetros expostos na tabela 4.6.

Tabela 4. 6 - Parâmetros de coestimativa utilizados no Software Isatis®.

Tipo de coestimativa	Cokrigagem ordinária
Número de variáveis	4
Discretização em X, Y e Z	4, 4 e 1

Elipsoide de busca	Mínimo de dados condicionantes	3
	Máximo de dados condicionais	16
	Azimute	45
	Dip	0
	Rake	0
	Raio máximo de busca	3000m
	Raio mínimo de busca	1200m

Após a realização da cokrigagem, os resultados foram validados e utilizados como referência para análise comparativa e da aplicabilidade da KMAF neste banco de dados.

4.1.3 SOLUÇÃO VIA KMAF

A descorrelação das variáveis consideradas neste estudo de caso foi realizada utilizando o método MAF, evita-se de aplicar o sistema de cokrigagem, e usa-se diretamente a krigagem ordinária em cada um dos fatores transformados.

A descorrelação se deu por meio do uso do algoritmo apresentado no capítulo 3. Após a transformação aplicada aos dados, em uma etapa prévia à estimativa, foi feita uma retrotransformação direta dos fatores aos dados para validar o algoritmo, e checar se seriam reconstituídos os dados iniciais.

Este teste foi realizado da seguinte maneira: convertem-se os dados em fatores MAF; e, após, os retrotransformam. Em seguida, plota-se o gráfico de dispersão entre a variável original e a retrotransformada. A correlação apresentada em cada gráfico deve ser máxima, ou seja, $\rho=1$.

O *lag*, que é solicitado no algoritmo para que a transformação MAF seja realizada, foi determinado como $h = 400m$. Este valor foi obtido por meio de testes com diversos valores de *lags*, variando de 50m a 600m e verificando as mudanças nos termos obtidos na matriz \mathbf{a} , da equação (2.28) da seção 2.4 do capítulo 2 utilizada na transformação dos dados em fatores

MAF. Em $h = 400m$, a matriz transformação não apresenta mudanças em seus termos. Portanto, este foi o *lag* selecionado.

O gráfico de dispersão entre os fatores MAF1 e MAF2 é apresentado na figura 4.20.

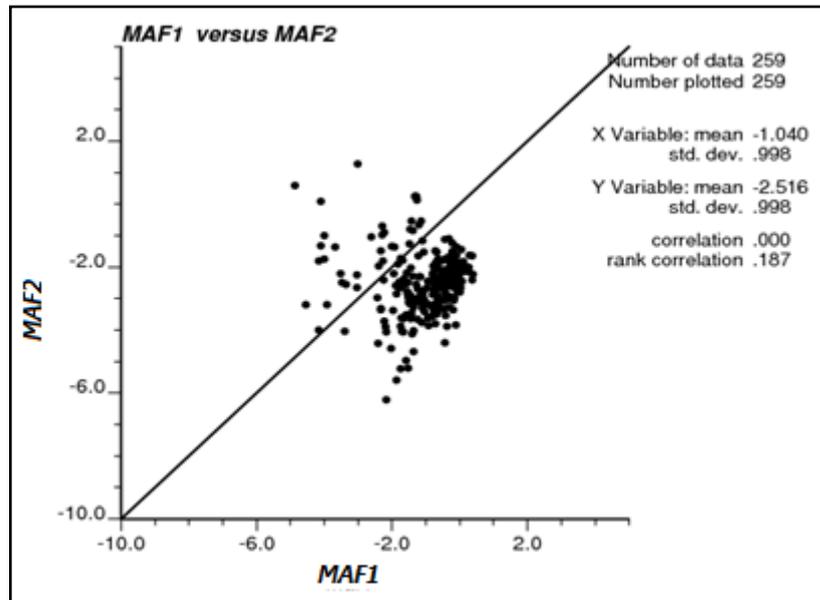


Figura 4. 20 - Gráfico de dispersão entre os fatores MAF1 e MAF2, obtidos por meio da transformação linear, realizada com o algoritmo exposto no capítulo 3

No gráfico da figura 4.20, pode-se observar a descorrelação total dos fatores MAF1 e MAF2 para $h = 0$, uma vez que o coeficiente de correlação é nulo. Na figura 4.21, é verificada a descorrelação espacial dos fatores por meio do correlograma experimental omnidirecional até $h = 400m$.

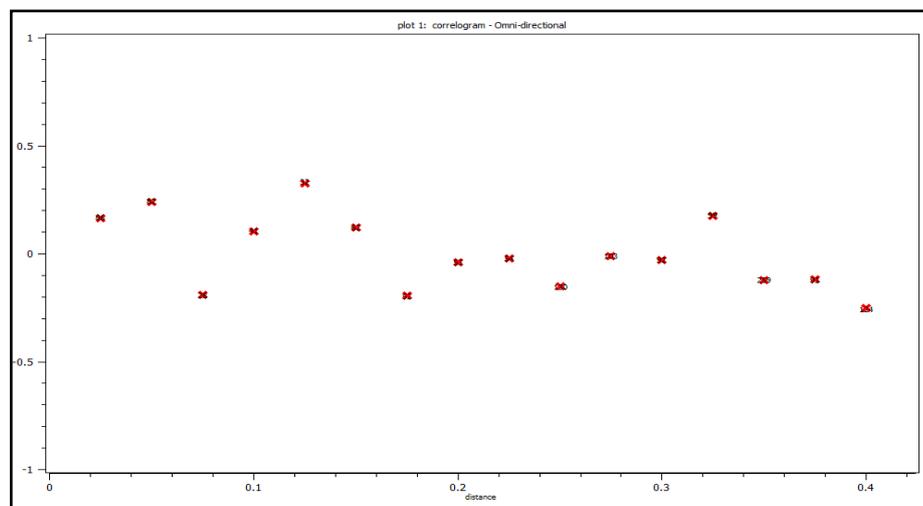


Figura 4. 21 – Correlograma experimental omnidirecional cruzado dos fatores MAF1 e MAF2

O correlograma experimental exposto na figura 4.21 foi construído utilizando 16 *lags*, tendo cada tamanho de 25m e tolerância de *lag* de 12.5m. A largura de banda utilizado foi de 25m, sendo o correlograma omnidirecional. Estes parâmetros foram utilizados para construir não só este, mas todos os correlogramas cruzados entre os fatores MAF obtidos por meio da transformação executada. Pode-se observar, na figura 4.21, o efeito pepita puro, que caracteriza a descorrelação dos fatores até a distância de 400m.

Na figura 4.22, é apresentado o gráfico de dispersão entre os fatores MAF1 e MAF3.

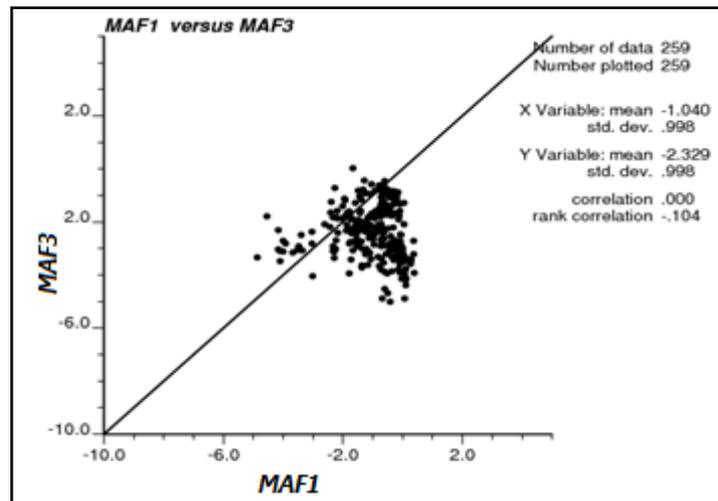


Figura 4. 22 - Gráfico de dispersão entre os fatores MAF1 e MAF3, obtidos por meio da transformação linear, realizada com o algoritmo exposto no capítulo 3

Na figura 4.22, se vê a descorrelação total dos fatores para $h = 0$. Na figura 4.23, é apresentado o correlograma experimental cruzado entre os fatores MAF1 e MAF3.

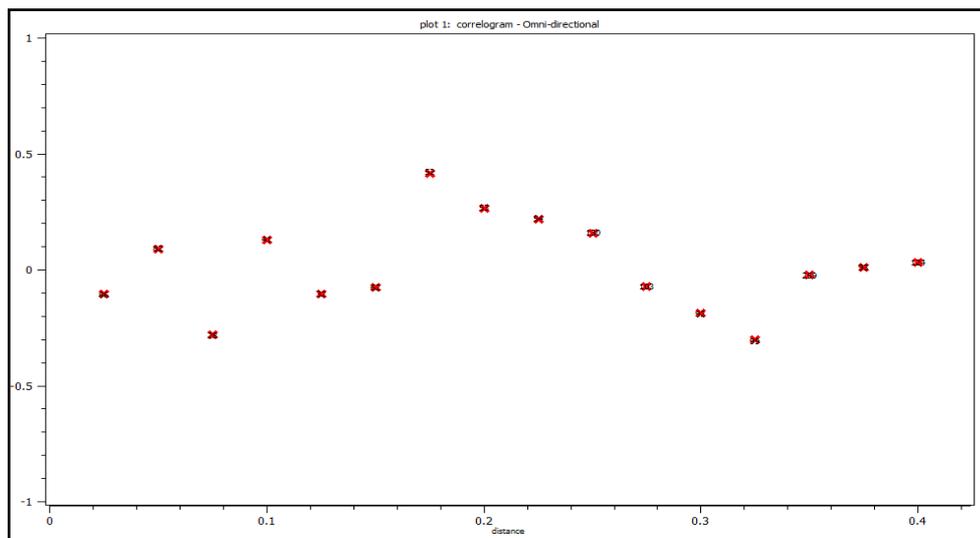


Figura 4. 23 - Correlograma experimental omnidirecional cruzado entre os fatores MAF1 e MAF3

Assim como na figura 4.21, pode ser observada na figura 4.23 a descorrelação espacial dos fatores MAF1 e MAF3 até $h = 400m$.

Na figura 4.24, é apresentada a dispersão entre os fatores MAF1 e MAF4.

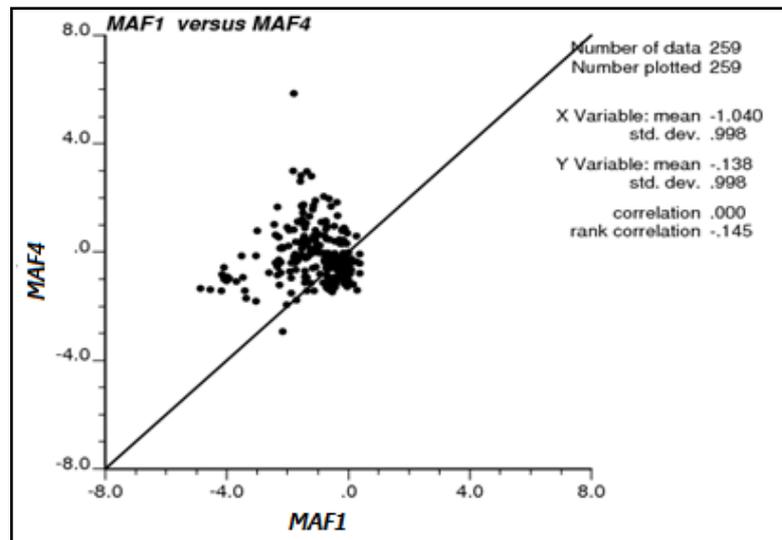


Figura 4. 24 - Gráfico de dispersão entre os fatores MAF1 e MAF4, obtidos por meio da transformação linear, realizada com o algoritmo exposto no capítulo 3

Assim como nos gráficos das figuras 4.20 e 4.22, na figura 4.24, também se pode observar a descorrelação total dos fatores MAF1 e MAF4 para $h = 0$. Na figura 4.25, é apresentado o correlograma experimental cruzado para os fatores MAF1 e MAF4.

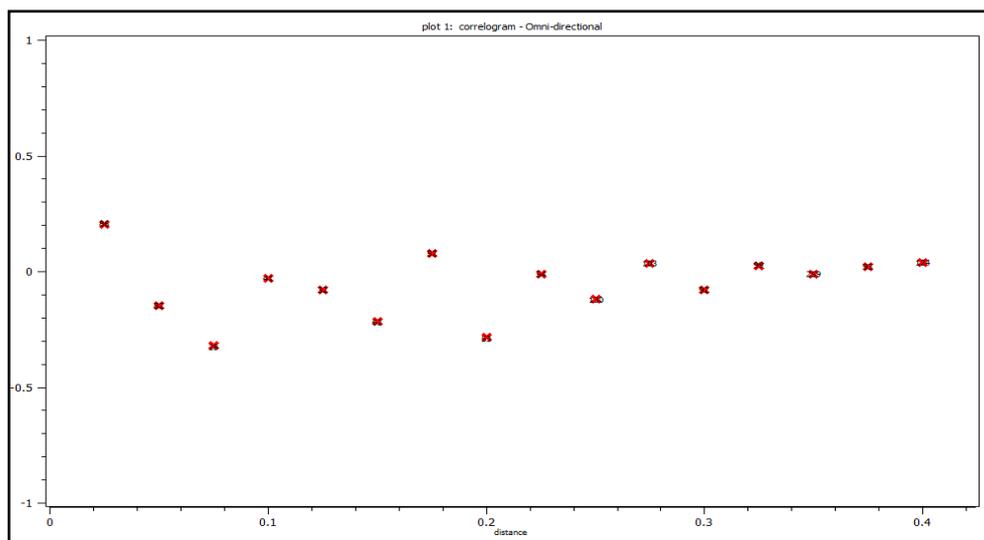


Figura 4. 25 - Correlograma experimental cruzado omnidirecional entre os fatores MAF1 e MAF4

Na figura 4.25, é vista a descorrelação dos fatores MAF1 e MAF4 até $h = 400m$, em que é possível observar o efeito pepita puro deste. Na figura 4.26, é mostrado o gráfico de

dispersão entre os fatores MAF2 e MAF3, em que se observa a descorrelação total destes para $h = 0$.

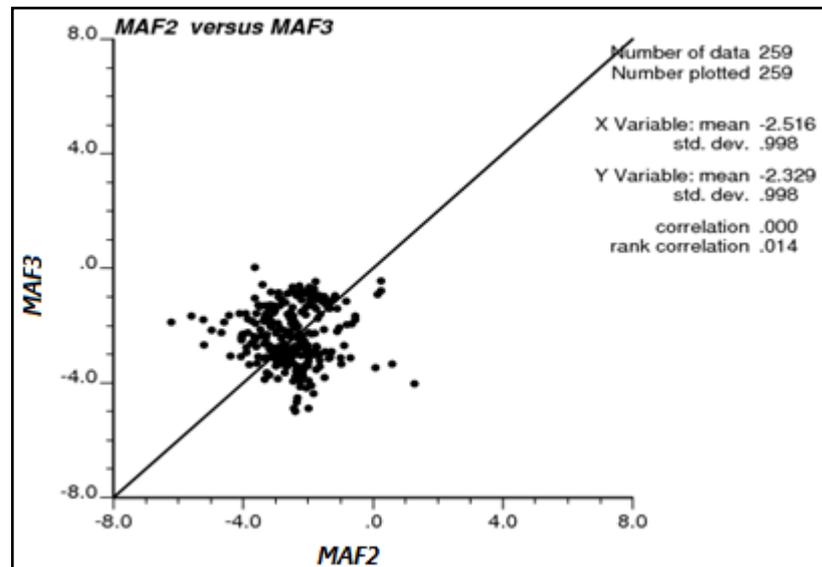


Figura 4. 26 - Gráfico de dispersão entre os fatores MAF2 e MAF3, obtidos por meio da transformação linear, realizada com o algoritmo exposto no capítulo 3

Na figura 4.26, a partir do gráfico exposto, vê-se a descorrelação total entre os fatores MAF2 e MAF3. Na figura 4.27, é apresentado o correlograma experimental cruzado entre os fatores MAF2 e MAF3.

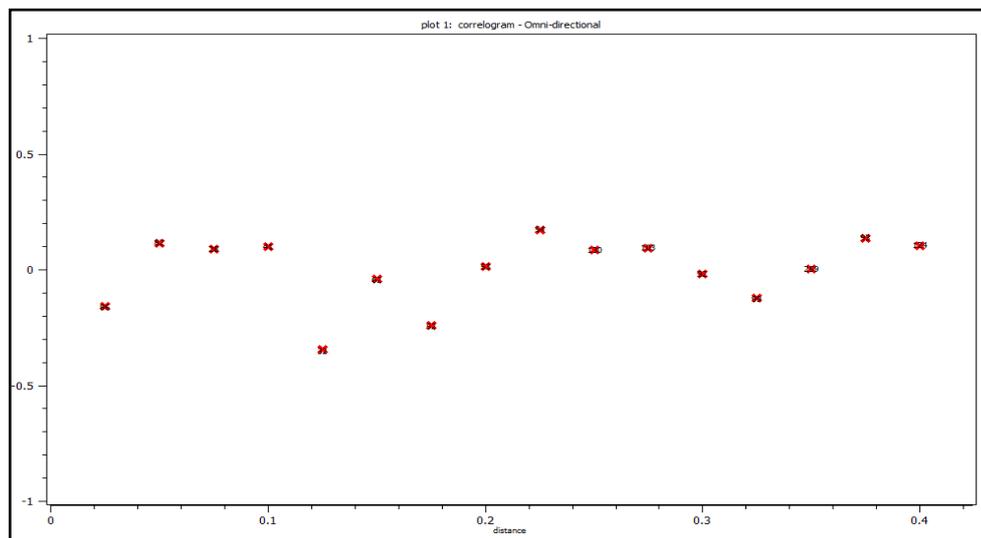


Figura 4. 27 - Correlograma experimental cruzado entre os fatores MAF2 e MAF3

Na figura 4.27, também é observada a descorrelação total até $h = 400m$ dos fatores, assim como nos correlogramas anteriores. Na figura 4.28, é mostrado o gráfico de dispersão entre os fatores MAF2 e MAF4 caracterizando a descorrelação para $h = 0$.

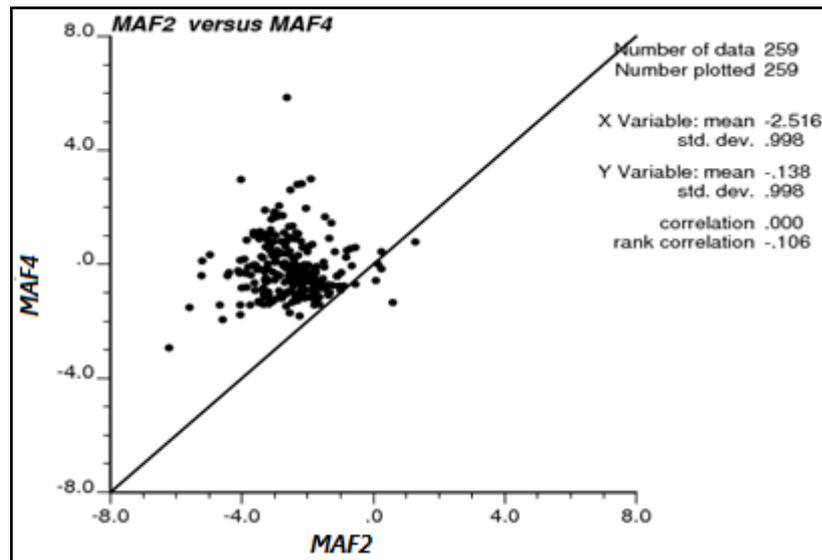


Figura 4. 28 - Gráfico de dispersão entre os fatores MAF2 e MAF4, obtidos por meio da transformação linear realizada com o algoritmo exposto no capítulo 3

Na figura 4.29, é mostrado o correlograma experimental cruzado entre os fatores MAF2 e MAF4, em que se observa a descorrelação até $h = 400m$.

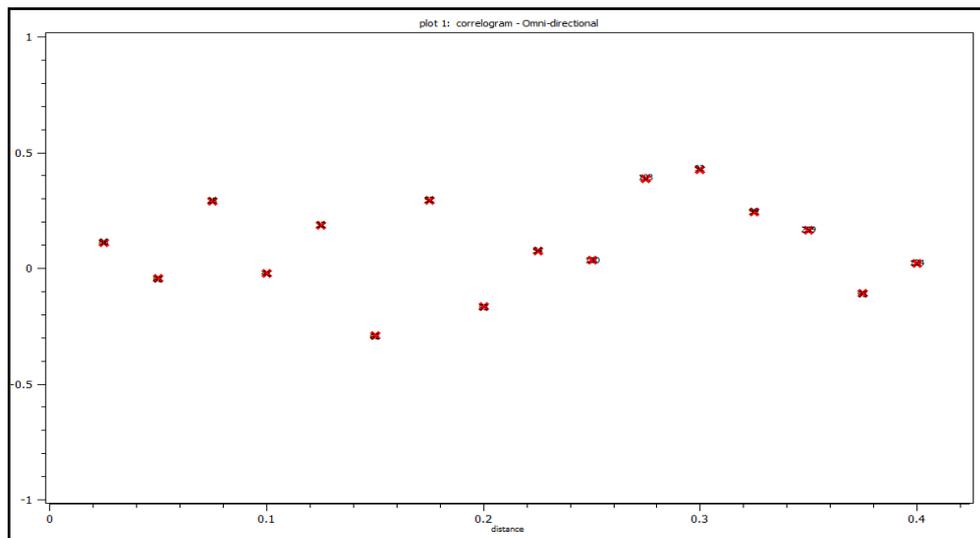


Figura 4. 29 - Correlograma experimental cruzado omnidirecional dos fatores MAF2 e MAF4

Finalmente, na figura 4.30, é apresentado o gráfico de dispersão entre os fatores MAF3 e MAF4 e, na figura 4.31, o correlograma experimental cruzado entre os fatores MAF3 e MAF4.

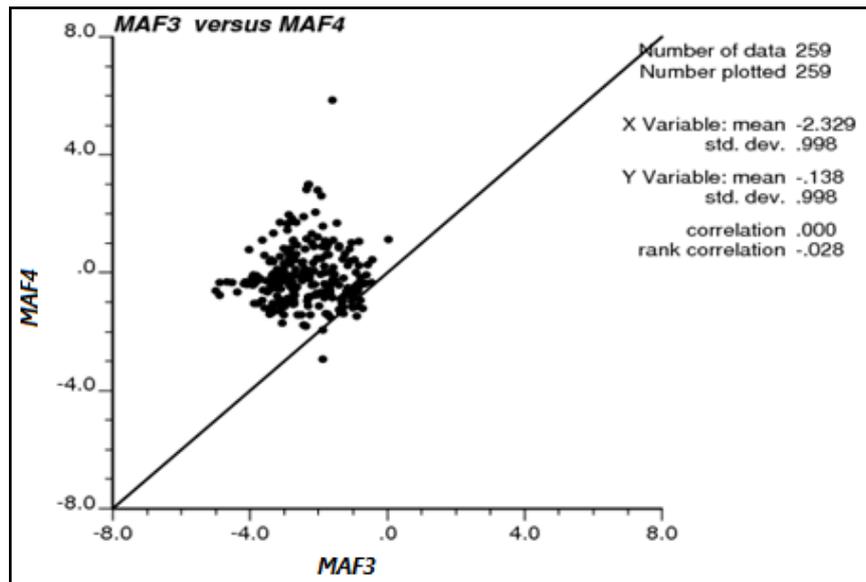


Figura 4. 30 - Gráfico de dispersão entre os fatores MAF3 e MAF4, obtidos por meio da transformação linear, realizada com o algoritmo exposto no capítulo 3

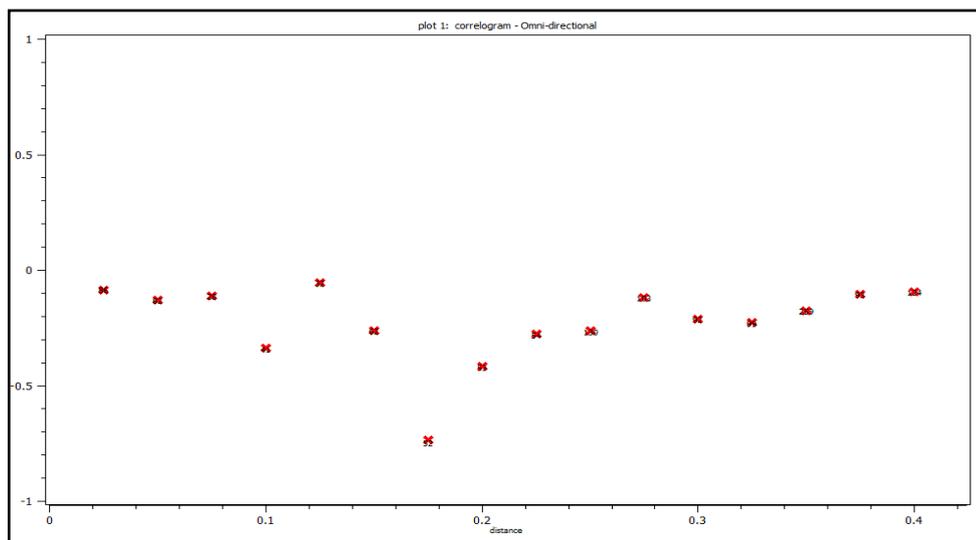


Figura 4. 31 - Correlograma experimental cruzando omnidirecional dos fatores MAF3 e MAF4

A partir das figuras 4.20 a 4.31, se vê que os fatores estão todos descorrelacionados entre si de $h = 0m$ a $h = 400m$, podendo ser tratados de maneira independente. Os diagramas de dispersão entre os fatores MAF retrotransformados e os dados originais são apresentados nas figuras 4.32 a 4.35. Note a reprodução dos dados após a retrotransformação.

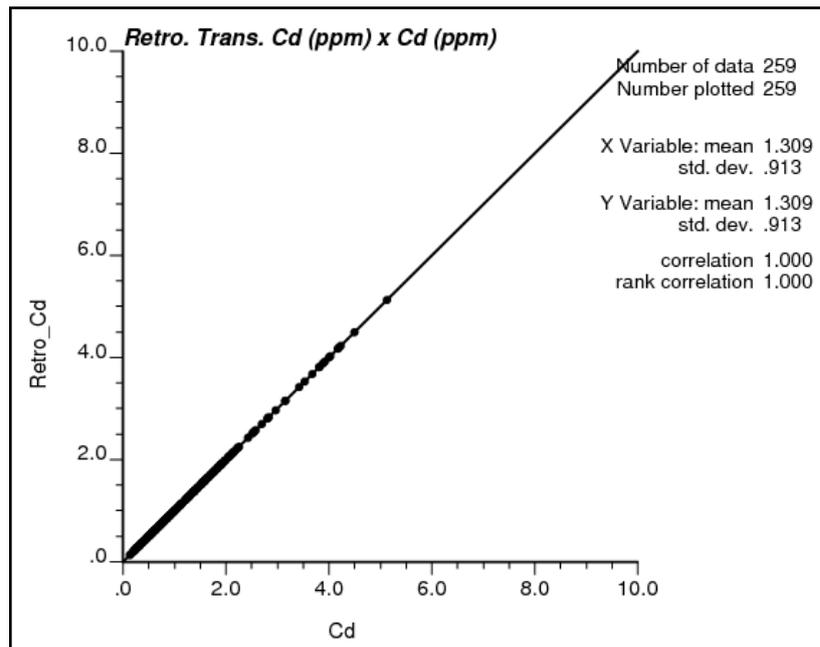


Figura 4. 32 - Gráfico de dispersão entre a variável Cádmio (ppm) obtida por meio da retrotransformação dos fatores MAF ao espaço dos dados originais e à variável Cádmio (ppm) original

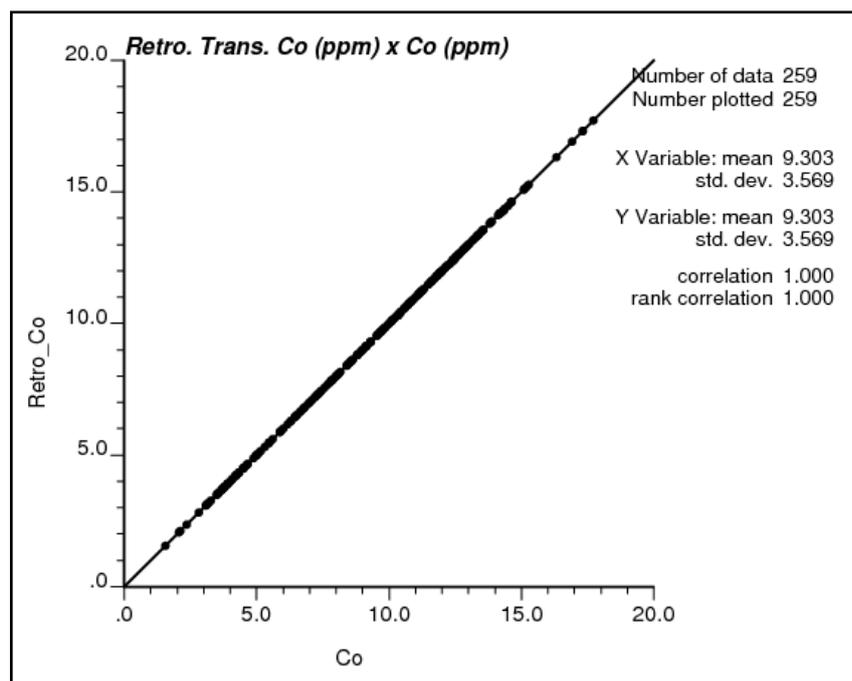


Figura 4. 33 - Gráfico de dispersão entre a variável Cobalto (ppm) obtida por meio da retrotransformação dos fatores MAF ao espaço dos dados originais e à variável Cobalto (ppm) original

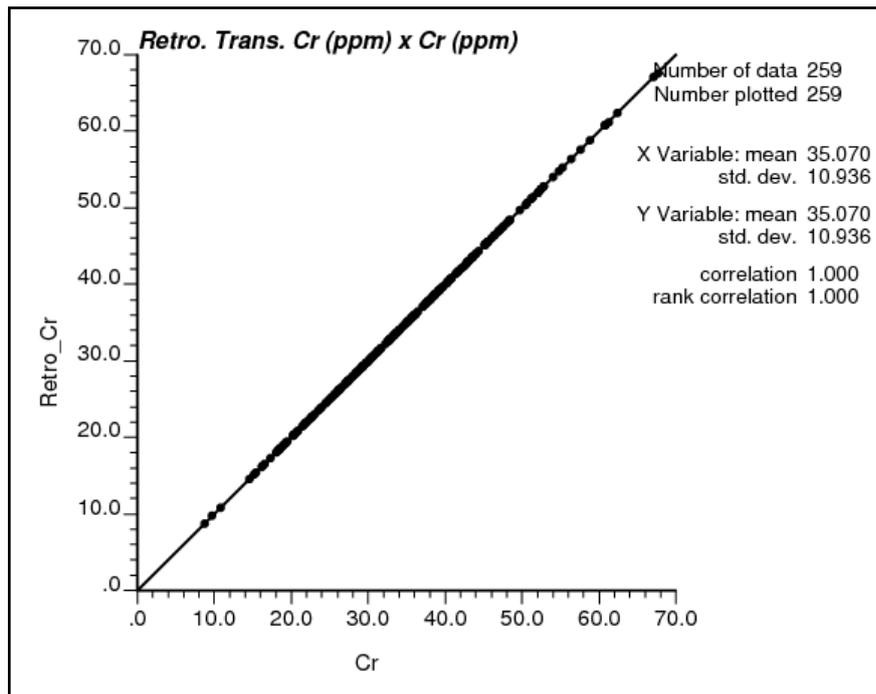


Figura 4. 34 - Gráfico de dispersão entre a variável Cromo (ppm) obtida por meio da retrotransformação dos fatores MAF ao espaço dos dados originais e à variável Cromo (ppm) original

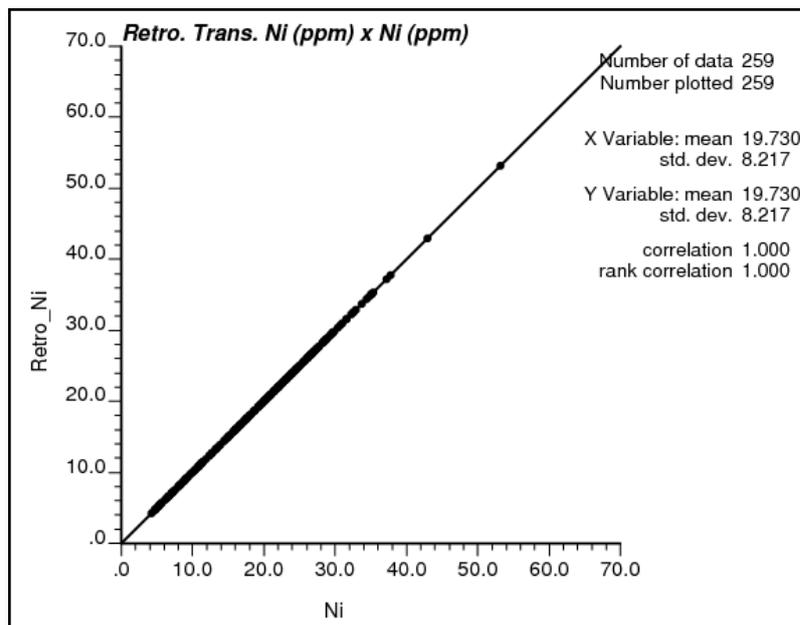


Figura 4. 35 - Gráfico de dispersão entre a variável Níquel (ppm) obtida por meio da retrotransformação dos fatores MAF ao espaço dos dados originais e à variável Níquel (ppm) original

Nos gráficos de dispersão expostos nas figuras 4.32 a 4.35, das variáveis Cádmio, Cobalto, Cromo e Níquel retrotransformadas a partir da transformação MAF, se verifica a correlação máxima com as variáveis originais. O gráfico demonstra que o algoritmo executou a transformação e retrotransformação linear adequadamente.

Uma vez que foi verificado que a transformação linear descorrelaciona os atributos e não gera erros no banco de dados, foi realizada a krigagem ordinária dos fatores independentes. Os parâmetros da variografia estão expostos na tabela 4.7.

Tabela 4.7 – Parâmetros usados para a construção dos variogramas direcionais experimentais do sistema de KMAF.

Variogramas direcionais no plano XY	
Número de lags	25
Tamanho de lag	500m
Tolerância de lag	250m
Largura de banda	500m
Número de direções	8
Tolerância angular	22,5°

Os modelos variográficos foram construídos de maneira independente para cada fator. Neste caso, são necessários somente 4 modelos, enquanto que, na cokrigagem, foram construídos dez modelos. Para todos os fatores, o efeito pepita foi obtido a partir dos variogramas omnidirecionais. Cada variograma omnidirecional foi calculado utilizando 40 *lags* de 100m cada e tolerância de *lag* de 50m, a largura de banda para cada variograma omnidirecional foi de 100m. Na figura 4.36, é apresentado o variograma experimental do fator MAF1.

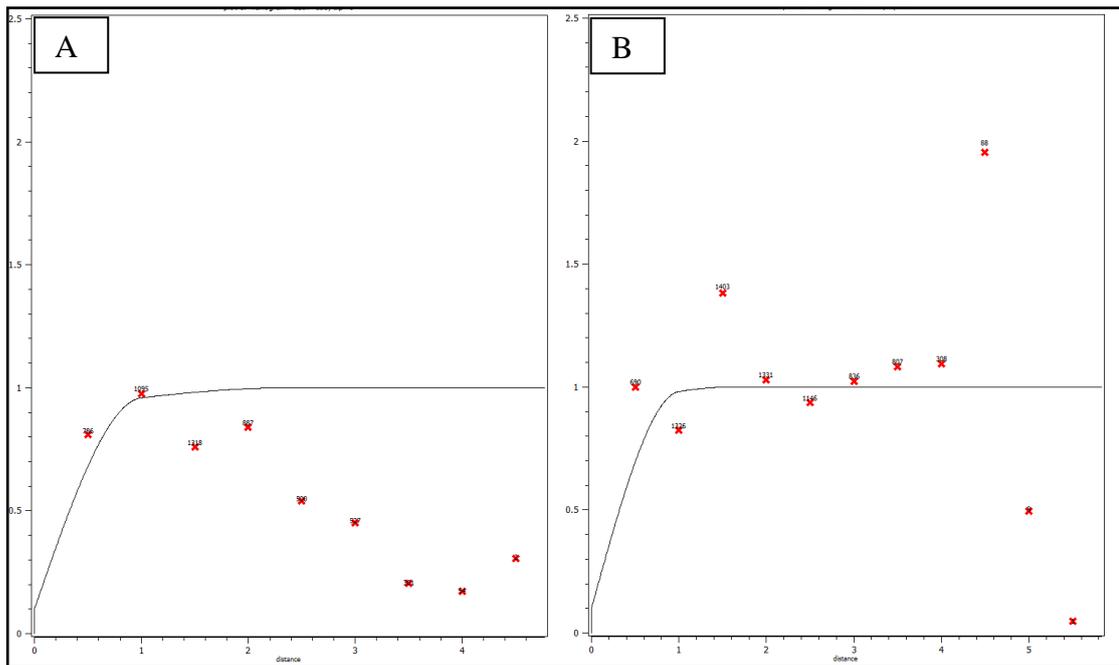


Figura 4. 36 - Variogramas direcionais do fator independente MAF1 (A) Direção de maior continuidade, N135 $dip = 0$; (B) direção de menor continuidade, N45 $dip = 0$

O variograma da figura 4.36, do fator MAF1, é constituído por três estruturas, duas com modelos esféricos, além do efeito pepita. A variância do fator MAF1 (e também dos outros fatores) é 1. O efeito pepita corresponde a 10% da variância total do fator, ou seja, $C_0 = 0,1$. A contribuição da primeira estrutura esférica, C_1 , é 80% da variância fator, ou seja, $C_1 = 0,8$. Foi atribuído à contribuição da segunda estrutura esférica, C_2 , os 10% restantes da variância total, ou seja, $C_2 = 0,1$. A direção de maior continuidade do primeiro fator, MAF1, é na N135 com $dip = 0$, com alcance de 2,4 km. A direção de menor continuidade é na N45 com $dip = 0$ e alcance de 1,6 km.

Assim como para o fator MAF1, foi construído um modelo para o fator MAF2. Na figura 4.37, é apresentando o variograma do fator MAF2.

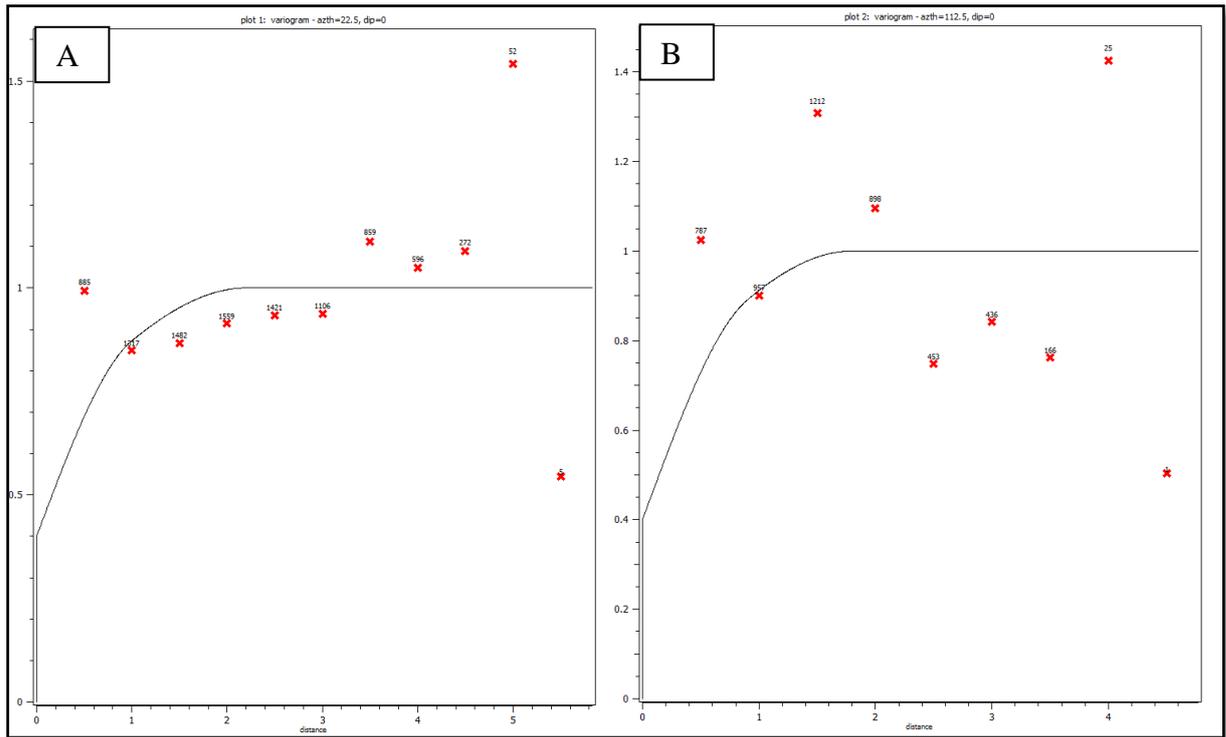


Figura 4. 37 - Variogramas direcionais do MAF2 (A) direção de maior continuidade, N22 $dip = 0$; (B) direção de menor continuidade, N112 $dip = 0$

Já para o fator MAF2, a direção de maior continuidade é N22, $dip = 0$ sendo a menor N112 $dip = 0$. Os alcances obtidos nas direções de maior e menor continuidade foram 2,2 km e 1,6 km, respectivamente. O variograma do fator MAF2 possui três estruturas, duas com modelo esférico, além do efeito pepita. A primeira estrutura esférica possui 25% da variância total do fator, $C_1 = 0,25$, a segunda estrutura esférica possui 35% do valor, ou seja, $C_2 = 0,35$. Os 40% remanescentes da variância é atribuído ao efeito pepita.

O variograma experimental do fator MAF3 apresenta a direção de continuidade espacial na N45 com $dip = 0$. O modelo de regionalização realizado também é composto por três estruturas, duas esféricas e efeito pepita. Na figura 4.38, é mostrado o variograma experimental do fator MAF3.

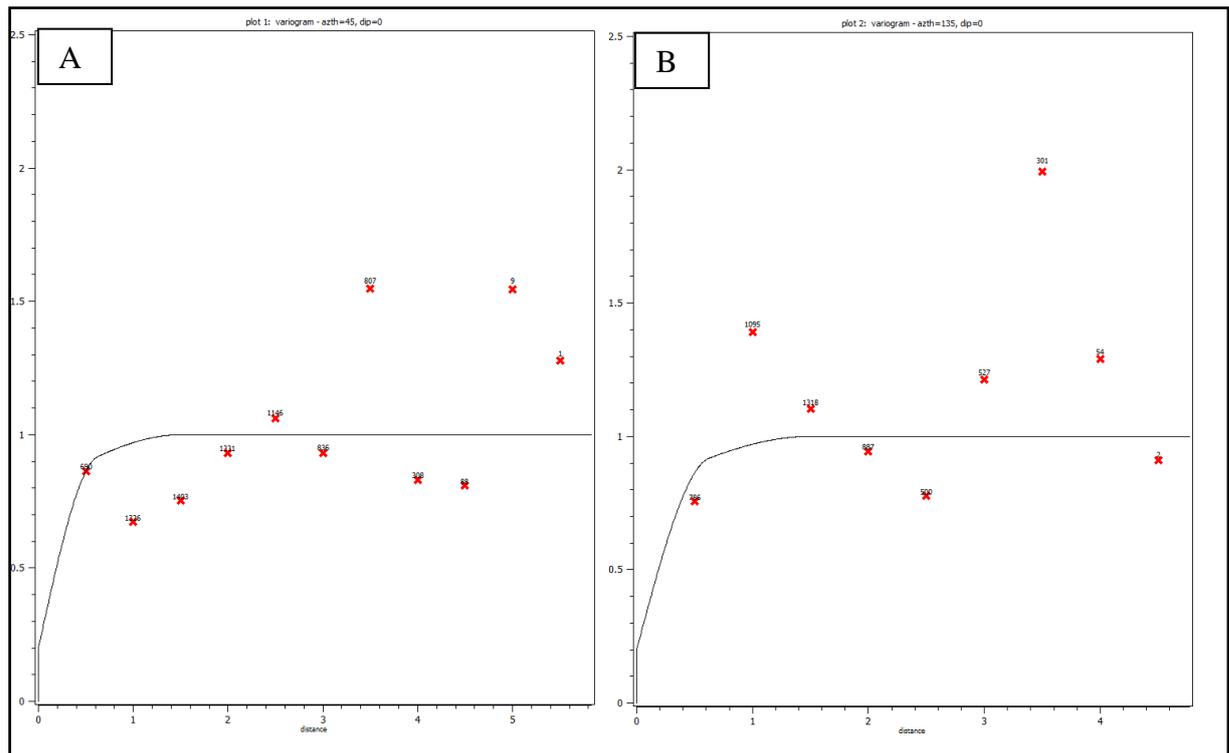


Figura 4. 38 - Variogramas direcionais do MAF3 (a) direção de maior continuidade, N45 $dip = 0$; (b) direção de menor continuidade N135 $dip = 0$

Cada estrutura que compõe o modelo de regionalização do fator MAF3 possui, assim como o fator MAF2, uma proporção da variância total do fator. A primeira estrutura esférica possui 60% variância total do fator, ou seja, $C_1 = 0,6$, a segunda estrutura esférica possui 20% da variância, $C_2 = 0,2$. Ao efeito pepita também foi atribuído com 20% da variância total, $C_0 = 0,2$. O alcance obtido no modelo variográfico do fator MAF3 na direção de maior continuidade foi de 1,5 km e na direção de menor continuidade de 1 km.

Já o modelo ajustado ao variograma experimental do fator MAF4 obteve a direção de maior continuidade na N45 com $dip = 0$, com alcance de 1,2 km. Na direção de menor continuidade, N135 e $dip = 0$, o alcance obtido foi 1 km. O modelo de regionalização também é caracterizado pela presença de três estruturas, duas com modelo esférico, além do efeito pepita. A primeira estrutura esférica possui contribuição de 70% da variância total do fator, $C_1 = 0,7$. A segunda estrutura esférica contribui com 20% da variância total do fator, $C_2 = 0,2$, enquanto que o efeito pepita possui 10% da variância total do fator. O variograma experimental do fator MAF4 é visto na figura 4.39.

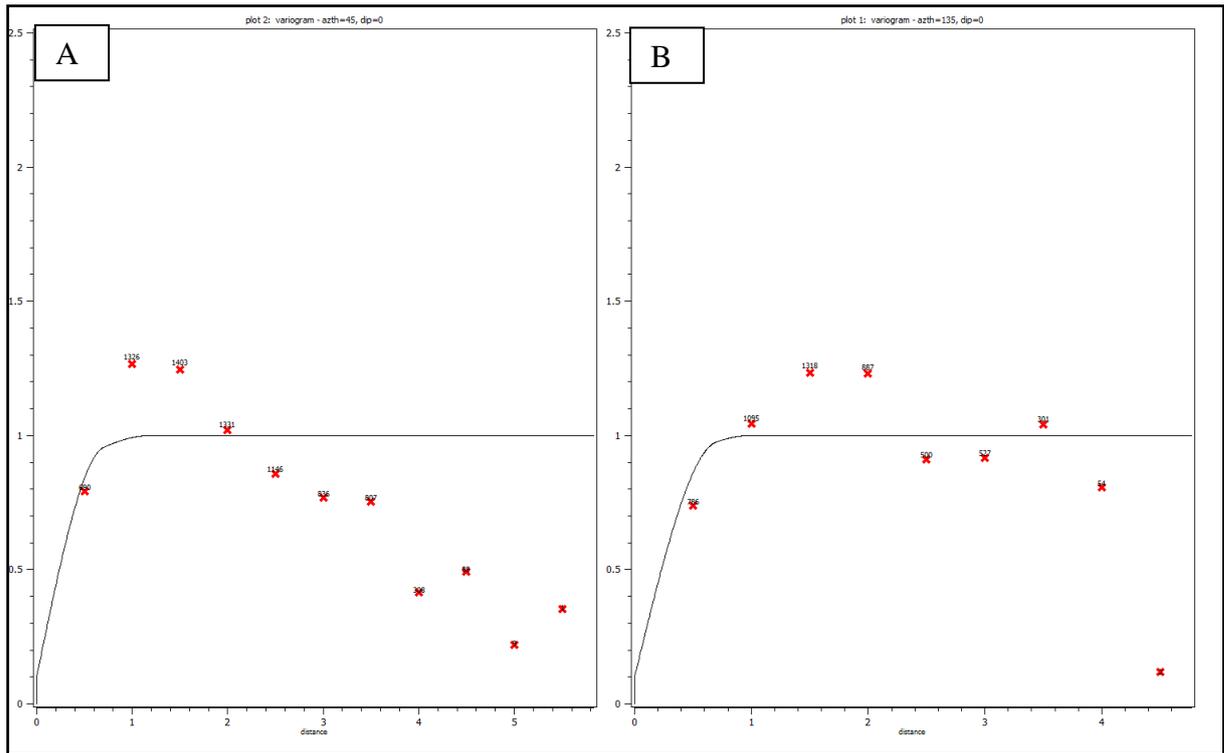


Figura 4. 39 - Variogramas direcionais do MAF4 (a) direção de maior continuidade, N45 $dip = 0$; (b) direção de menor continuidade N135 $dip = 0$

Note que somente os variogramas experimentais diretos são necessários para as estimativas com transformação MAF.

Os modelos variográficos dos fatores MAF estão expostos nas equações 4.11 a 4.14.

$$\gamma_{MAF1} = 0,1 + 0,8Sph\left(\frac{1km}{N135} \frac{0,8km}{N45}\right) + 0,1Sph\left(\frac{2,4km}{N135} \frac{1,6km}{N45}\right) \quad (4.11)$$

$$\gamma_{MAF2} = 0,4 + 0,25Sph\left(\frac{1km}{N22} \frac{0,8km}{N112}\right) + 0,35Sph\left(\frac{2,2km}{N22} \frac{1,6km}{N112}\right) \quad (4.12)$$

$$\gamma_{MAF3} = 0,2 + 0,6Sph\left(\frac{0,7km}{N45} \frac{0,8km}{N135}\right) + 0,2Sph\left(\frac{1,5km}{N45} \frac{1km}{N135}\right) \quad (4.13)$$

$$\gamma_{MAF4} = 0,1 + 0,7Sph\left(\frac{0,7km}{N45} \frac{0,7km}{N135}\right) + 0,2Sph\left(\frac{1,2km}{N45} \frac{1km}{N135}\right) \quad (4.14)$$

O modelo utilizado para estimativa possui blocos de 100m na direção X e 100m na direção Y. Cada fator MAF foi estimado por meio de krigagem ordinária, utilizando o mesmo raio de busca no elipsoide, para garantir o mesmo número dos blocos estimados. Apesar da

diferença na direção de anisotropia, as estimativas foram realizadas nos mesmos blocos, resultando na isotopia necessária para realizar a retrotransformação. Os parâmetros de krigagem estão expostos na tabela 4.8.

Tabela 4. 8 - Tabela dos parâmetros de krigagem para os fatores MAF. Cada fator foi interpolado separadamente com seu modelo de continuidade e somente o raio do elipsoide de busca na estratégia de krigagem permanece igual.

Tipo de estimativa	Krigagem ordinária	
Variável primária	MAF	
Discretização em X,Y e Z	4, 4 e 1	
Elipsoide de busca	Mínimo de dados condicionantes	2
	Máximo de dados condicionantes	16
	Azimute	Dos modelos de cada MAF
	Dip	0
	Rake	0
	Raio máximo de busca	2,2km
	Raio médio de busca	1km
	Raio mínimo de busca	---

Após as estimativas realizadas independentemente nos fatores MAF, estas são retrotransformadas ao espaço dos dados originais, no qual as análises podem ser realizadas.

4.1.4 ANÁLISE DOS RESULTADOS – ESTUDO DE CASO JURA

Como parte da validação dos resultados, sabe-se que a forma do histograma não tende a ser reproduzido devido à suavização da interpolação; porém, a média global desagrupada dos dados originais deve ser. A partir dos resultados obtidos pela cokrigagem e KMAF, pode-se comparar as médias desagrupadas com as obtidas por meio dos métodos aqui empregados e calcular o erro associado ao resultado obtido. O erro relativo percentual é calculado da seguinte maneira:

$$err_{\%} = \frac{m_{od} - m_{est}}{m_{od}} \cdot 100 \quad (4.15)$$

sendo m_{od} a média original dos dados desagrupada, m_{est} a média da estimativa, $err_{\%}$ o erro relativo percentual. Na tabela 4.9, são apresentadas as médias originais e coestimadas por meio da cokrigagem, juntamente com o erro relativo percentual associado a cada variável coestimada. É importante ressaltar que, tanto na KMAF quanto na cokrigagem, foram realizadas estimativas utilizando diferentes parâmetros nas estratégias de busca das amostras, tais como: tamanho do raio de busca do elipsoide; número de amostras condicionantes utilizadas. Foram mantidas as melhores estimativas obtidas.

Tabela 4. 9 - Comparativo entre as médias originais desagrupadas das variáveis do banco de dados Jura e as médias obtidas por meio de coestimativa (cokrigagem).

	Média original desgrupada (ppm)	Média estimada (ppm)	Erro relativo percentual (%)
Cádmio	1,27	1,32	3,93
Cobalto	9,23	9,26	0,32
Cromo	36,55	36,24	-0,84
Níquel	20,38	20,34	-0,19

A partir dos valores da tabela 4.9, pode-se observar que o erro relativo obtido na coestimativa do atributo Cádmio foi de 3,93%. O erro associado ao processo é devido ao fato de que todas as variáveis são coestimadas assumindo um modelo base que descreve o comportamento espacial conjunto delas. Porém, dentre as quatro variáveis do sistema, o atributo Cádmio possui fraca correlação ($\rho = 0,25$) com o Cobalto, variável na qual se baseou a construção do modelo de correção regionalização, levando, assim, a um erro na coestimativa da média do Cádmio mais alto que os obtidos para as outras variáveis.

Já no caso da KMAF, não se assume um modelo para o comportamento espacial de um fator que não seja o calculado e modelado para este fator. Na tabela 4.10, é feito o comparativo das médias originais desagrupadas e as médias obtidas após a KMAF.

Tabela 4. 10 - Comparativo entre as médias originais desagrupadas das variáveis do banco de dados Jura e as médias obtidas por meio de estimativa (KMAF).

	Média original desagrupada (ppm)	Média estimada via KMAF (ppm)	Erro relativo percentual (%)
--	---	--	---

Cádmio	1,27	1,29	1,57
Cobalto	9,23	9,39	1,4
Cromo	36,55	36,40	-0,41
Níquel	20,38	20,69	1,52

Note que as estimativas obtidas por meio do método KMAF apresentaram erro relativo satisfatório, não apresentando diferenças significativas em relação às coestimativas. Apesar de a cokrigagem apresentar erros relativos associados à reprodução das médias de cada variável, a exceção do Cádmio e do Cromo, mais baixos que os apresentados pelo método KMAF, a operacionalização da KMAF é mais simples e rápida que a do método clássico.

Para verificar em mais detalhe o comportamento das médias estimadas por meio dos métodos multivariados utilizados nesta dissertação, foi realizada a análise de deriva, que consiste em comparar as médias dos blocos com a média dos dados em intervalos espaciais determinados. Quando realizada na direção X, os intervalos têm dimensão de 100m nesta direção, e dimensão infinita da direção perpendicular. Do mesmo modo, quando analisado na direção Y, os intervalos têm dimensão de 100m na direção Y e infinita na direção X.

Para realizar tal validação, foi criado um delimitador do corpo de minério. Lembrando que o desagrupamento dos dados foi realizado por polígonos de Voronoi (Dirichlet, 1850), em um modelo de blocos com células com dimensão de 10m na direção X e 10m na direção Y, foi criado um filtro no modelo de blocos que mantivesse somente os blocos no entorno do corpo de minério. Assim, a análise de deriva é feita somente na área delimitada, pois regiões nas quais a estimativa é realizada utilizando poucos dados (escassez de dados na região) podem levar a suavizações exageradas dos valores. Os gráficos da análise de deriva na direção X e Y do atributo Cádmio são apresentados nas figuras 4.40 e 4.41.

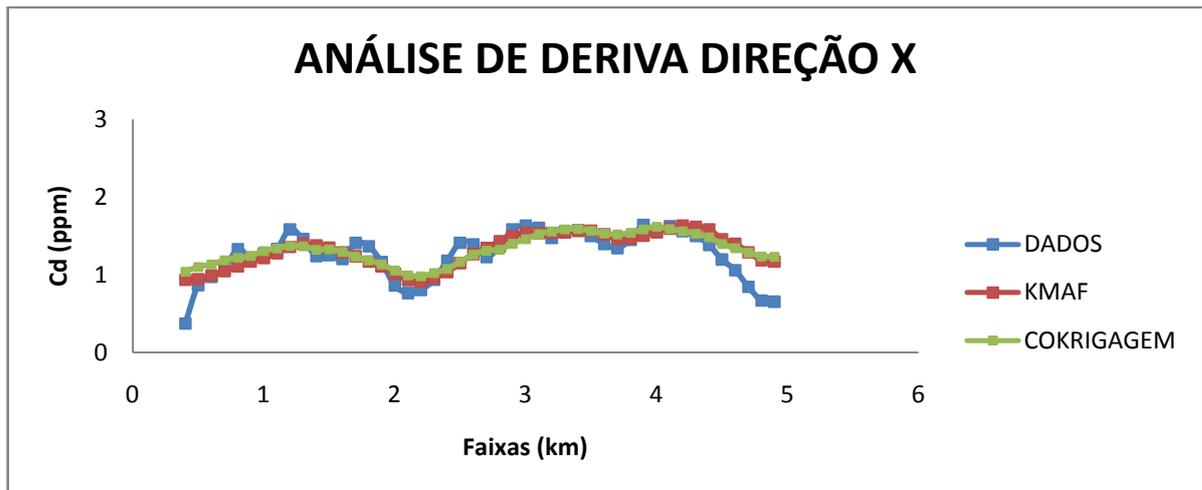


Figura 4. 40 - Análise de deriva na direção X (Leste) do atributo Cádmio (ppm). As faixas estão divididas em intervalos de 100m

Note, na figura 4.40, que a tendência local da média desagrupada dos dados é honrada tanto pela cokrigagem quanto pela KMAF. O comportamento é suavizado pelas duas metodologias. Já na figura 4.41, em que é mostrada a análise de deriva na direção Y (Norte), a cokrigagem, em torno do valor 2 km, apresenta um comportamento contrário à tendência dos dados mais acentuado que o apresentado pela KMAF.

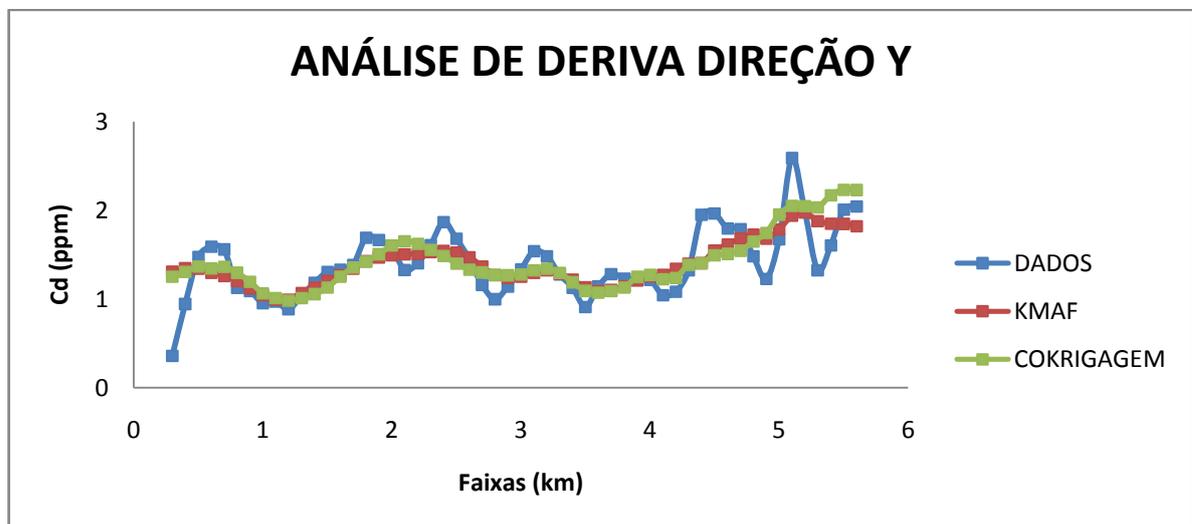


Figura 4. 41 - Análise de deriva na direção Y (Norte) do atributo Cádmio (ppm). As faixas estão divididas em intervalos de 100m

Na figura 4.42 e 4.43, são mostradas as análises de deriva para a variável Cobalto nas direções X e Y, respectivamente.

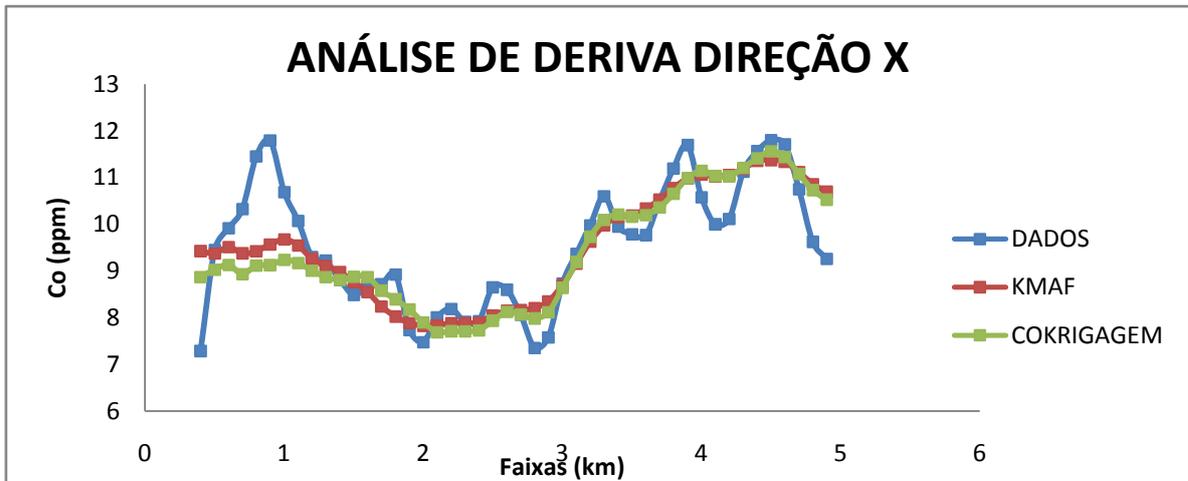


Figura 4. 42 - Análise de deriva na direção X (Leste) do atributo Cobalto (ppm). As faixas estão divididas em intervalos de 100m

Note, na figura 4.42, que, tanto a cokrigagem quanto a KMAF, suavizam o comportamento da média, principalmente nos extremos do gráfico. Até 1 km, as curvas da KMAF e da cokrigagem passam por baixo da curva dos dados suavizando a tendência apresentada pelos dados. Ambas as metodologias apresentam comportamento bastante próximo. A figura 4.43 apresenta a análise realizada na direção Y para a variável Cobalto.

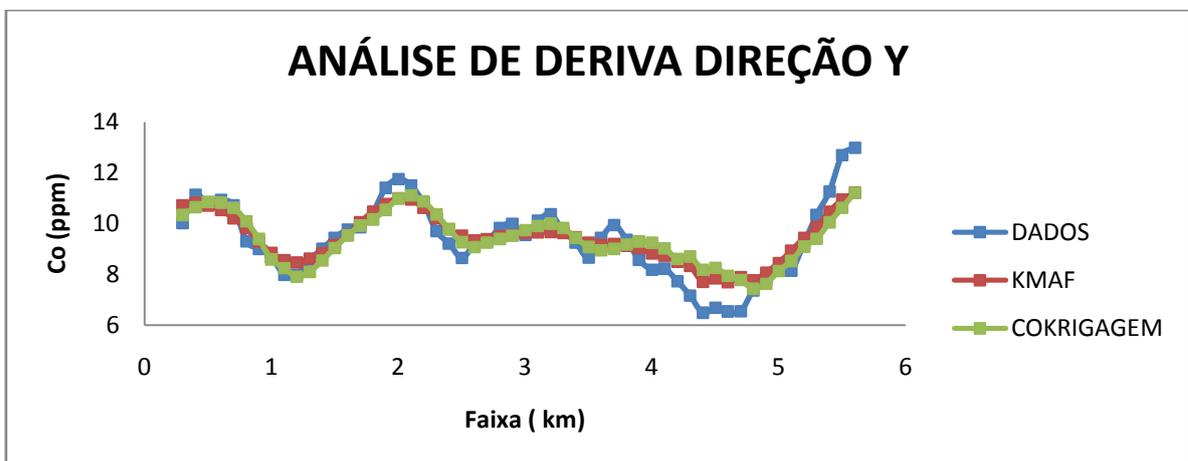


Figura 4. 43 - Análise de deriva na direção Y (Norte) do atributo Cobalto (ppm). As faixas estão divididas em intervalos de 100m

Percebe-se, a partir da figura 4.43, que, na direção Y, até 4 km, as metodologias acompanham, de forma muito próxima, a tendência local dos dados. Entre 4 km e 5 km, nota-se uma superestimativa da média tanto pela cokrigagem quanto pela KMAF.

Foi realizada também análise de deriva para o atributo Cromo. E esta é apresentada nas Figuras 4.44 e 4.45, nas direções X e Y, respectivamente.

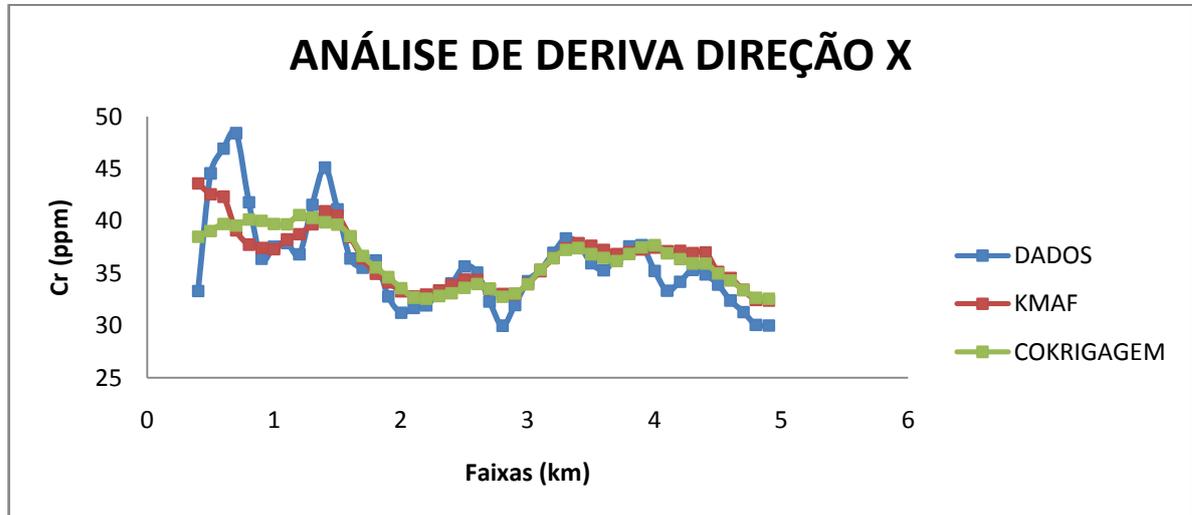


Figura 4. 44 - Análise de deriva na direção X (Leste) do atributo Cromo (ppm). As faixas estão divididas em intervalos de 100m

Note, na figura 4.44, que, em torno do valor 1 km, a curva da cokrigagem apresenta um aumento na média das estimativas, enquanto os dados apresentam um comportamento decrescente. Já a KMAF acompanha o comportamento dos dados desagrupados nesta região. A figura 4.45 apresenta a análise para a variável Cromo na direção Y.

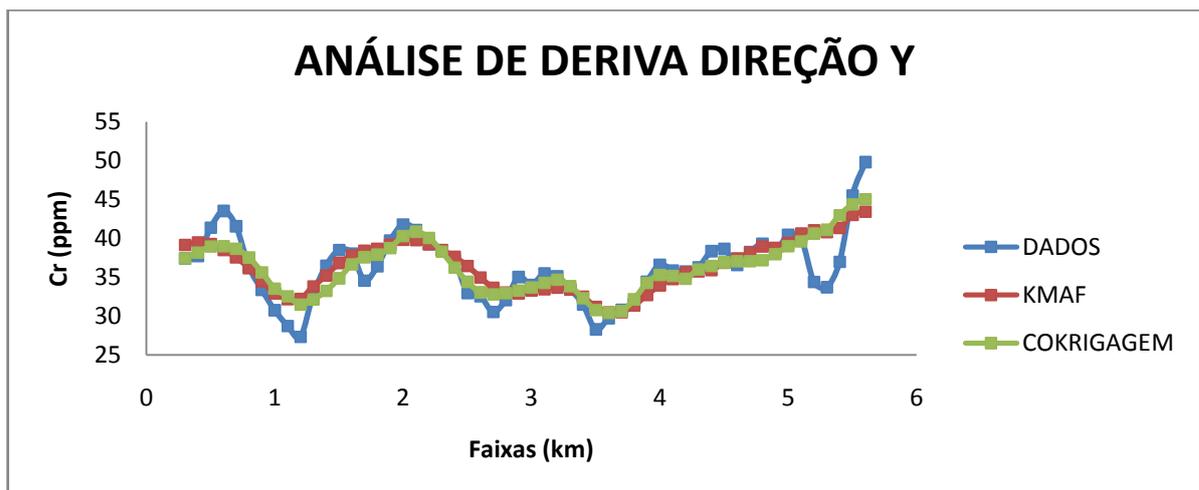


Figura 4. 45 - Análise de deriva na direção Y (Norte) do atributo Cromo (ppm). As faixas estão divididas em intervalos de 100m

Observa-se, no gráfico da figura 4.45, que, na direção Y, as curvas das duas metodologias se sobrepõem. Finalmente, a análise de deriva realizada para a variável Níquel é apresentada nas figuras 4.46 e 4.47.

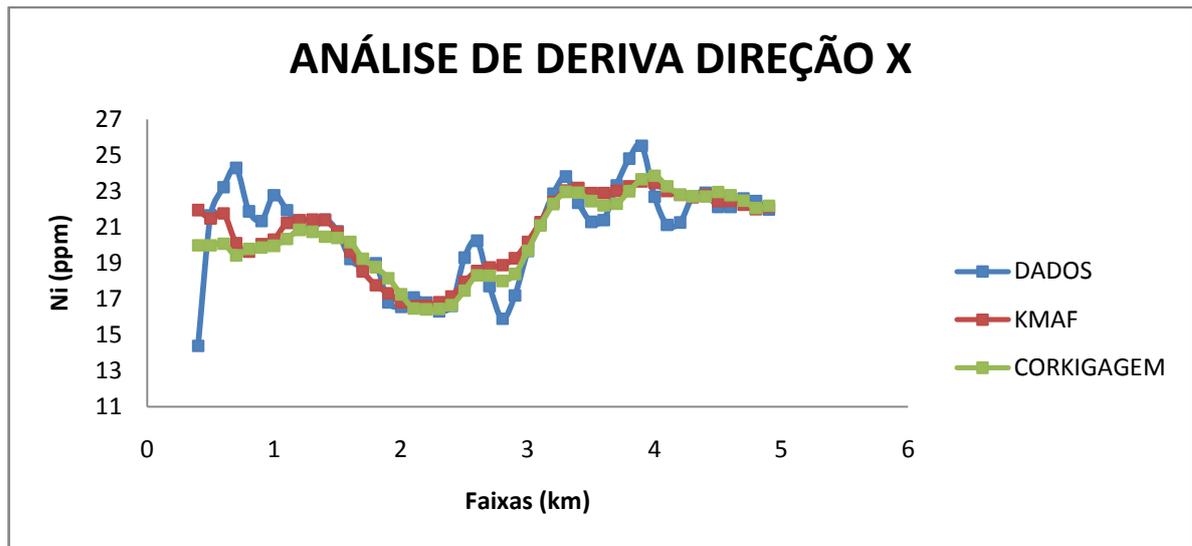


Figura 4. 46 – Análise de deriva na direção X (Leste) do atributo Níquel (ppm). As faixas estão divididas em intervalos 100m

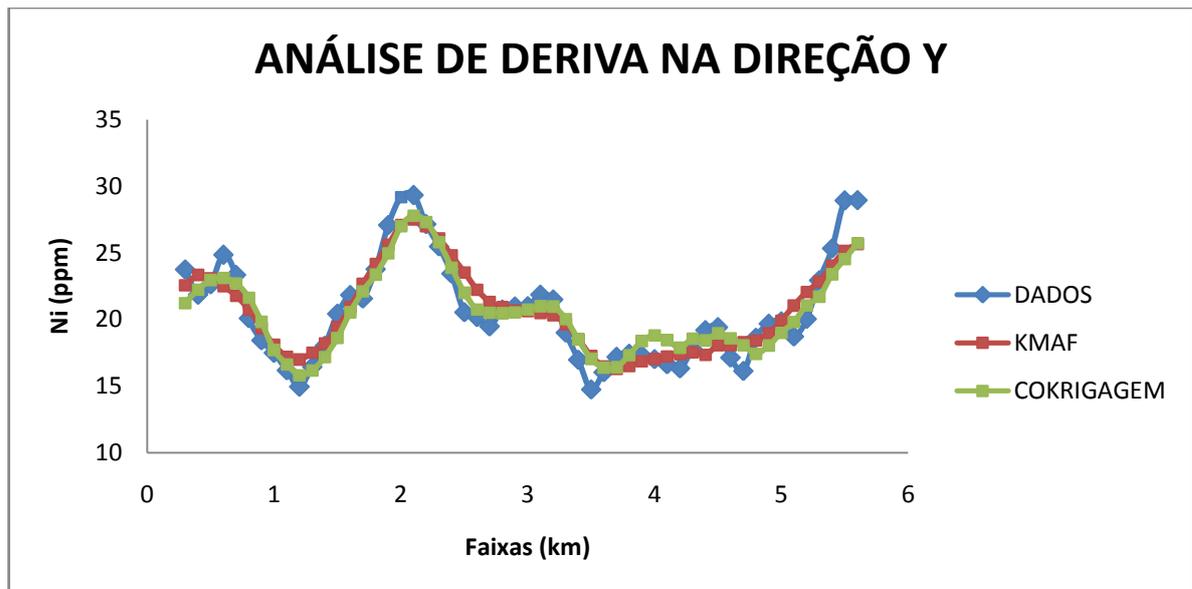


Figura 4. 47 - Análise de deriva na direção Y (Norte) do atributo Níquel. As faixas estão divididas a cada 100m

Note que, nas Figuras 4.46 e 4.47, o comportamento local da média dos dados desagrupados é reproduzido pelas duas metodologias. Ambas apresentam comportamento estatístico próximo entre os modelos e os dados.

Outro ponto a ser analisado são as correlações entre os atributos obtidas por meio de estimativas e coestimativas.

Para verificar as correlações entre as variáveis após a cokrigagem, foram feitos os gráficos de dispersão entre as saídas das coestimativas realizadas. O gráfico entre as coestimativas de Cádmio e Cobalto é apresentado na figura 4.48.

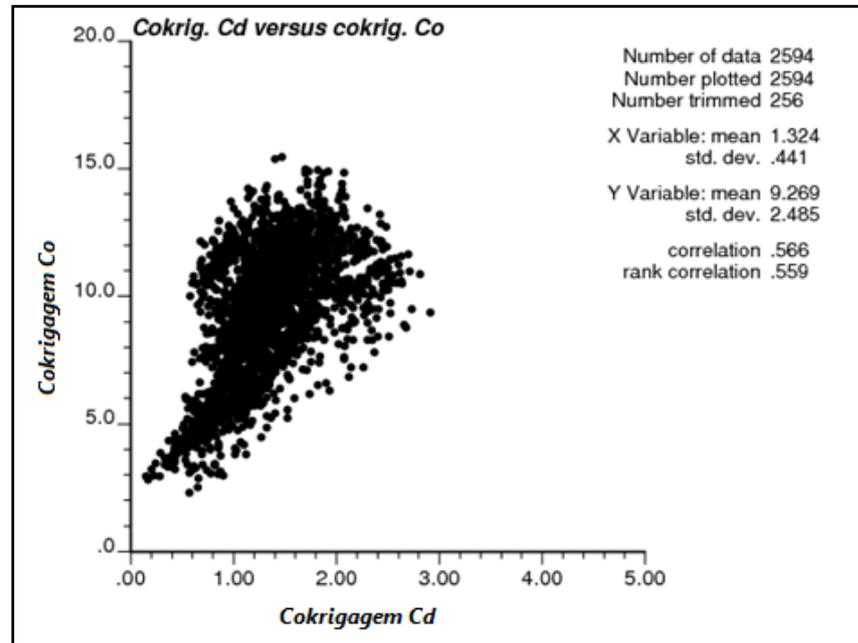


Figura 4. 48- Gráfico de dispersão das variáveis cokrigadas Cádmio (ppm) e Cobalto (ppm)

Note que a correlação apresentada pelas variáveis após a coestimativa, na figura 4.48, é $\rho_{estCdCo} = 0,56$. A nuvem de pontos é menos dispersa que a apresentada no gráfico original. O valor de correlação entre Cádmio e Cobalto original é $\rho_{CdCo} = 0,25$. A saída da coestimativa apresenta mais que duas vezes a correlação original.

O atributo Cobalto apresenta fraca correlação com a variável Cádmio. Desta forma, o modelo linear de correogionalização, que, neste estudo caso, foi baseado na variabilidade espacial do Cobalto, tornou-se inadequado à coestimativa do Cádmio, gerando, ao fim, uma correlação espúria, ou seja, sem significado físico entre o par de variáveis.

Já correlação original entre Cádmio e Cromo é $\rho_{CdCr} = 0,60$. Após a coestimativa, o valor aumenta em 20% da correlação original, como pode ser visto na figura 4.49.

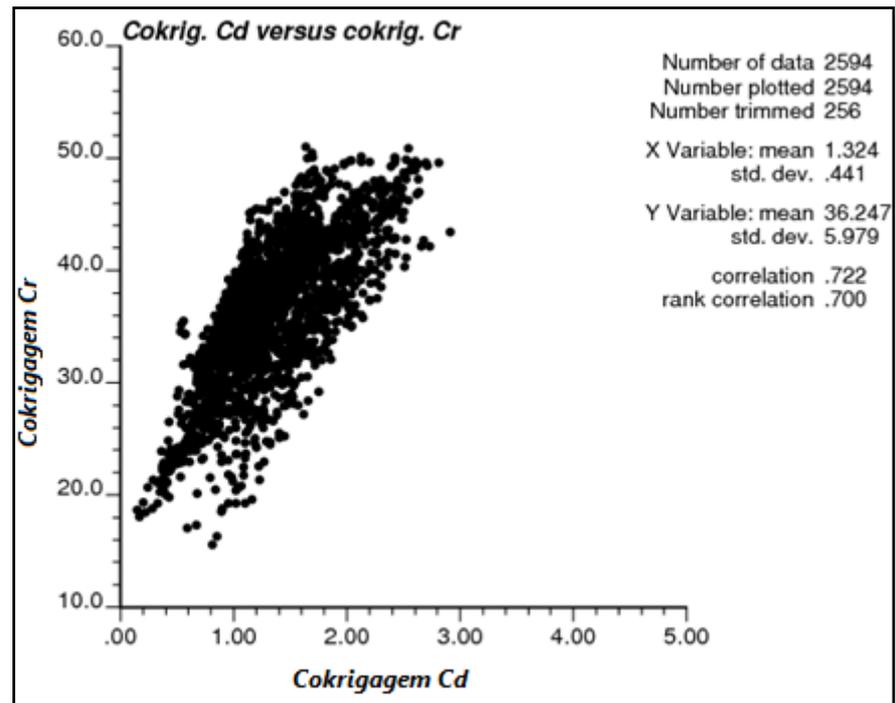


Figura 4. 49 - Gráfico de dispersão das variáveis cokrigadas Cádmio (ppm) e Cromo (ppm)

Note que, como os atributos do gráfico da figura 4.49 apresentavam uma forte correlação originalmente, a coestimativa não a aumenta tanto como fez no caso de Cádmio e Cobalto. No caso de Cádmio e Níquel, também é estabelecida pela coestimativa uma correlação que não existe previamente nos dados, como se pode verificar na figura 4.50.

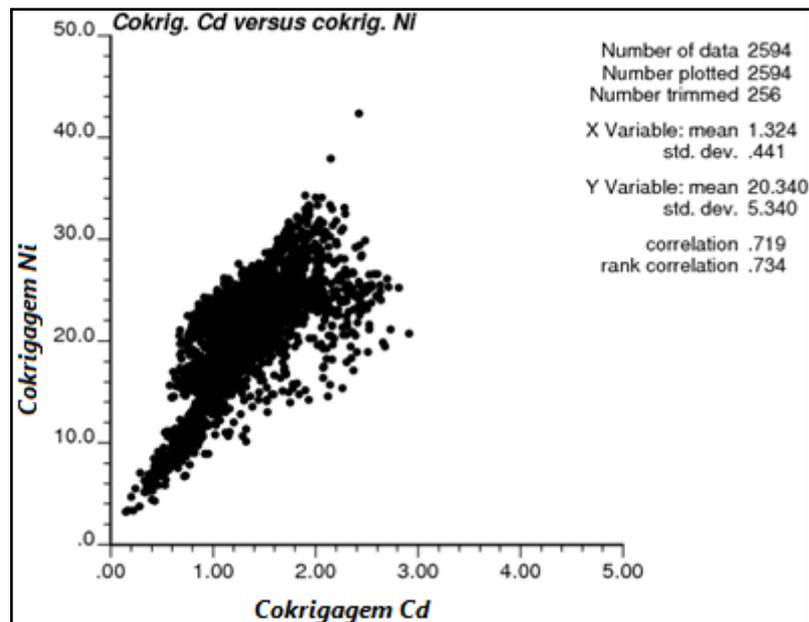


Figura 4. 50 - Gráfico de dispersão das variáveis cokrigadas Cádmio (ppm) e Níquel (ppm)

A correlação original entre Cádmio e Níquel é $\rho_{CdNi} = 0,48$, a observada no gráfico da figura 4.50 é $\rho_{estCdNi} = 0,71$. O aumento na correlação é 47% do valor original.

Para a correlação entre Cobalto e Cromo, o aumento após a cokrigagem é de 31% do valor original. Na figura 4.51, é mostrado o gráfico de dispersão entre Cobalto e Cromo.

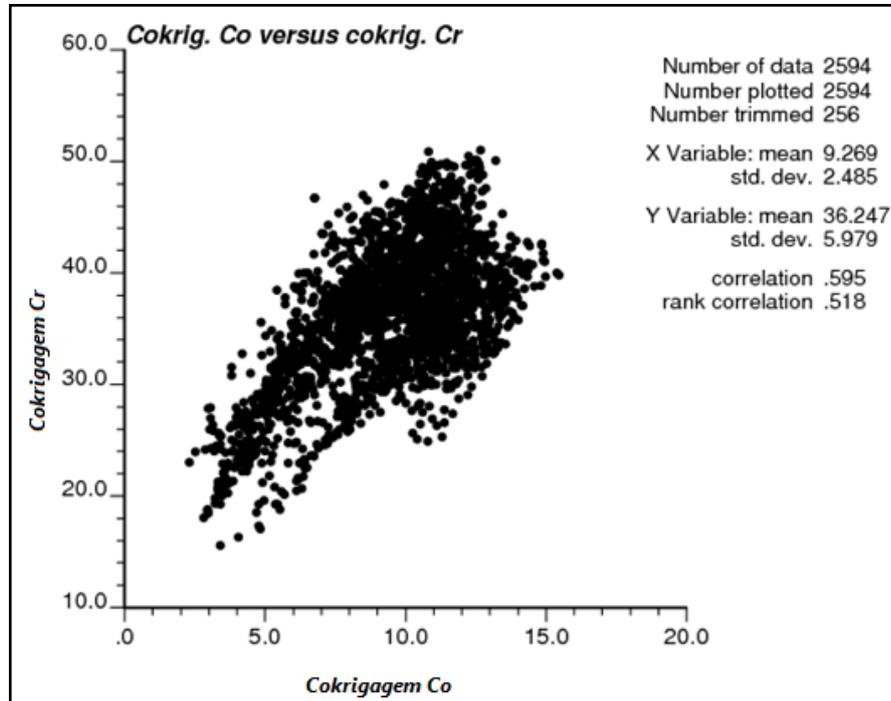


Figura 4. 51 - Gráfico de dispersão das variáveis cokrigadas Cobalto (ppm) e Cromo (ppm)

Lembrando que a correlação entre Cobalto e Cromo originalmente era $\rho_{CoCr} = 0,45$. Após a coestimativa, seu valor se tornou $\rho_{estCoCr} = 0,59$. Já o par Cobalto e Níquel apresentava a correlação mais forte presente no banco de dados, $\rho_{CoNi} = 0,75$. Após a cokrigagem, a correlação entre o par de variáveis aumentou 13% do seu valor original. A figura 4.52 mostra o gráfico de dispersão.

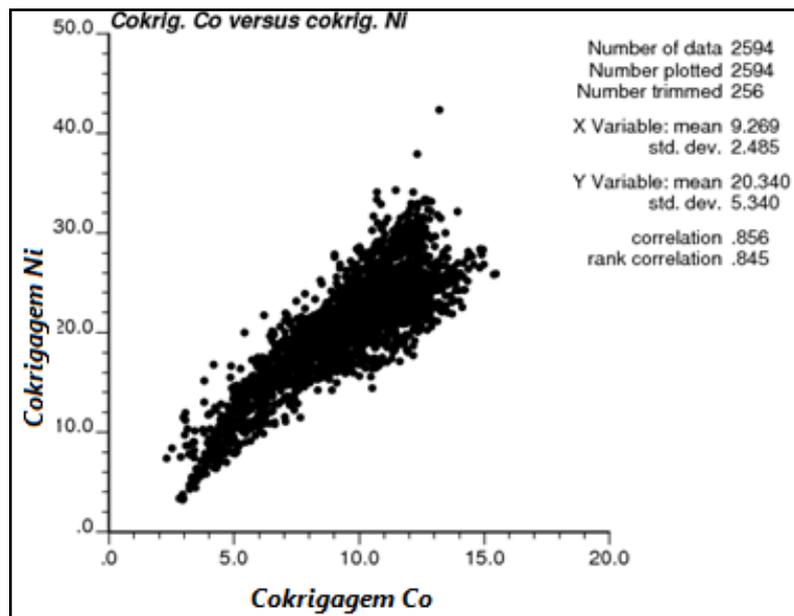


Figura 4. 52 - Gráfico de dispersão das variáveis cokrigadas Cobalto (ppm) e Níquel (ppm)

Note, na figura 4.52, que a correlação apresentada pelo par Cobalto e Níquel passou a $\rho_{estCoNi} = 0,85$.

No caso do par de variáveis Cromo e Níquel, que a correlação original era de $\rho_{CrNi} = 0,69$; após a cokrigagem, o par não apresentou aumento significativo como dos outros pares de variáveis. O gráfico de dispersão entre Cromo e Níquel cokrigados é mostrado na figura 4.53.

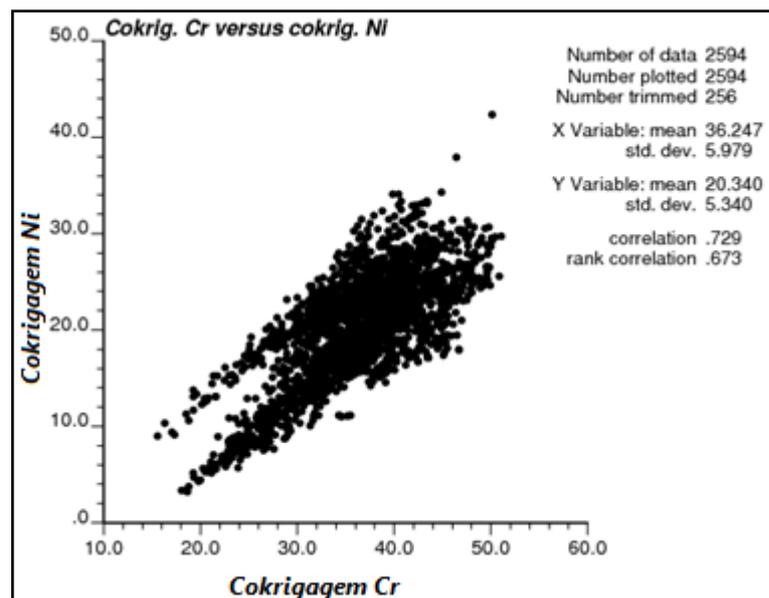


Figura 4. 53 - Gráfico de dispersão das variáveis cokrigadas Cromo (ppm) e Níquel (ppm)

A correlação apresentada no gráfico de dispersão da figura 4.53, entre Cromo e Níquel, é $\rho_{estCrNi} = 0,72$. Houve um aumento de 4.3% no valor original da correlação.

O modelo linear de corregeionalização, utilizado na cokrigagem, impõe que o comportamento espacial de cada variável envolvida no sistema seja caracterizado pelo mesmo modelo base. Portanto, as correlações exageradas entre pares que não apresentavam correlações significativas, originalmente, surgem após a coestimativa.

Na KMAF, foi realizada a mesma análise. O gráfico de dispersão entre o primeiro par, Cádmiio e Cobalto, é apresentado na figura 4.54.

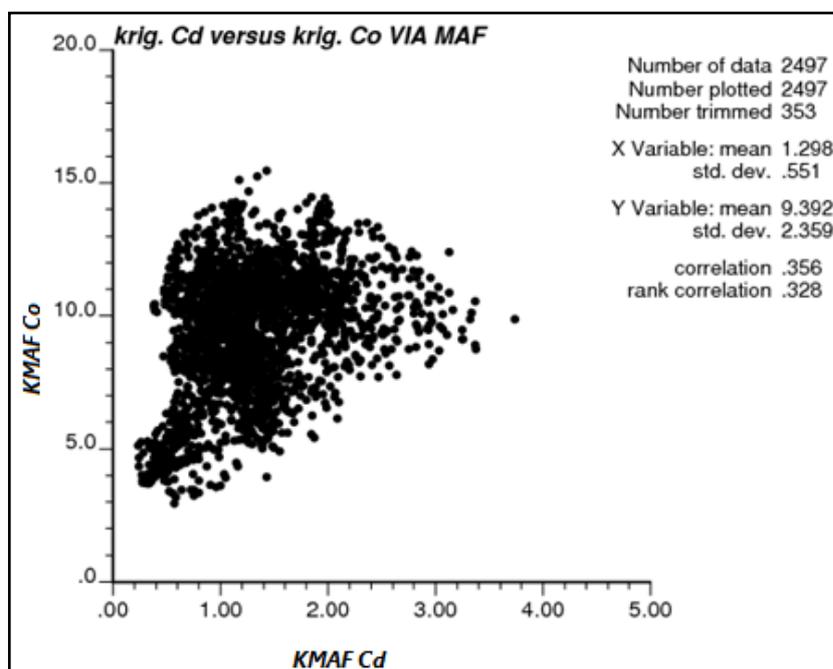


Figura 4. 54 - Gráfico de dispersão entre as variáveis krigadas Cádmiio (ppm) e Cobalto (ppm) utilizando a KMAF

Na figura 4.54, verifica-se que a correlação estabelecida após a estimativa das duas variáveis por meio de KMAF é $\rho_{estCdCo} = 0,35$, aumento de 40% no valor da correlação original. Ainda assim, menor que o valor de correlação apresentado pela cokrigagem.

Já para o par de variáveis, Cádmiio e Cromo, o valor da correlação apresentado após a estimativa é $\rho_{estCdCo} = 0,57$, sendo menor que o valor de correlação original dos dados, $\rho_{CdCo} = 0,60$. A dispersão entre as variáveis krigadas pode ser vista na Figura 4.55.

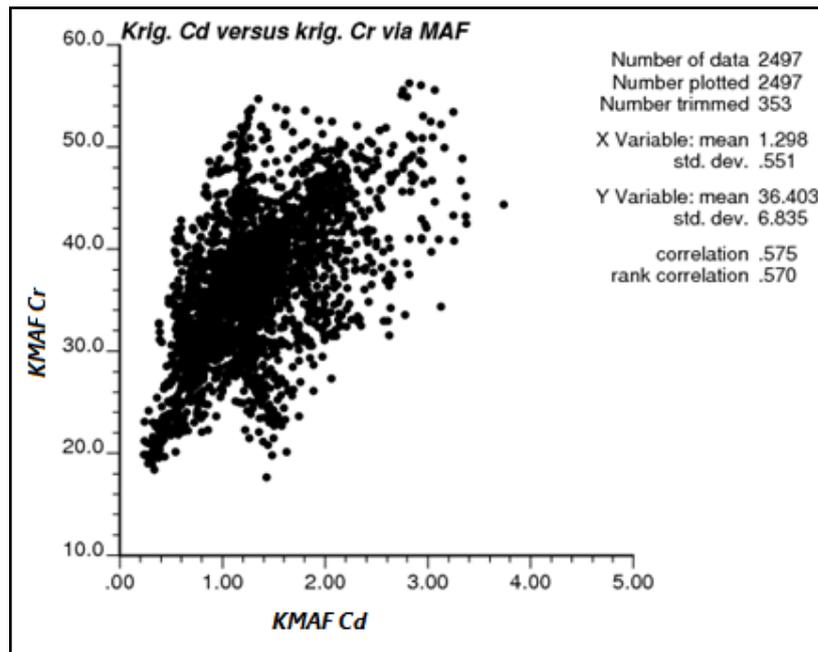


Figura 4. 55 - Gráfico de dispersão entre as variáveis krigadas Cádmio (ppm) e Cromo (ppm) utilizando a KMAF

Observa-se, na figura 4.55, que há uma diferença de 5% do valor original de correlação do par Cádmio e Cromo, $\rho_{cdCr} = 0,60$. Já a dispersão entre Cádmio e Níquel apresenta uma correlação de $\rho_{estCdNi} = 0,56$, como pode ser visto na figura 4.56.

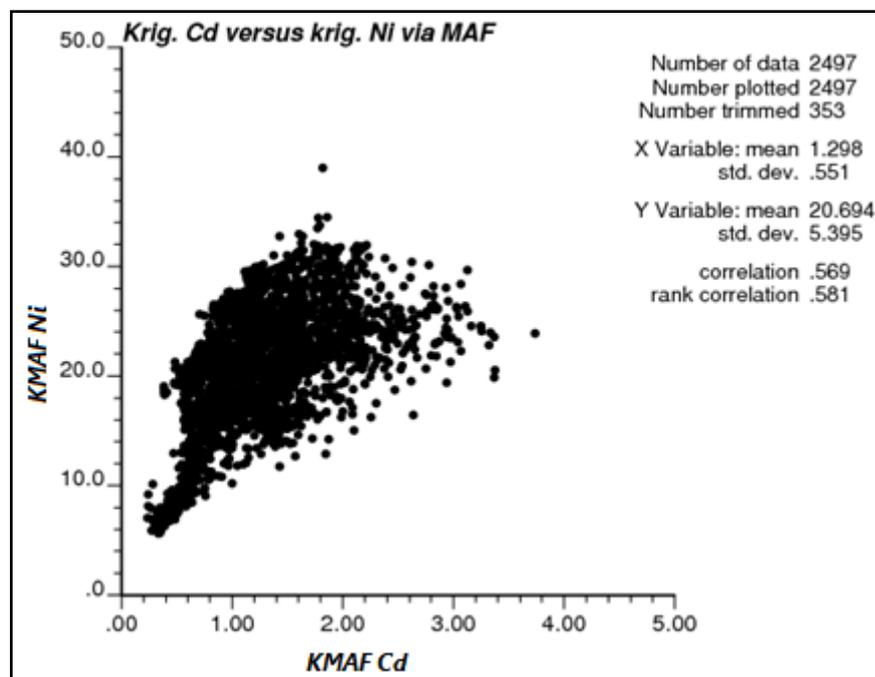


Figura 4. 56 - Gráfico de dispersão entre as variáveis krigadas Cádmio (ppm) e Níquel (ppm) utilizando a KMAF

O aumento da correlação, neste caso, é de 16,6% em relação ao valor original.

Para o caso da correlação estabelecida entre Cobalto e Cromo, após a KMAF, esta apresenta o valor de $\rho_{estCoCr} = 0,46$, como pode ser visto na figura 4.57.

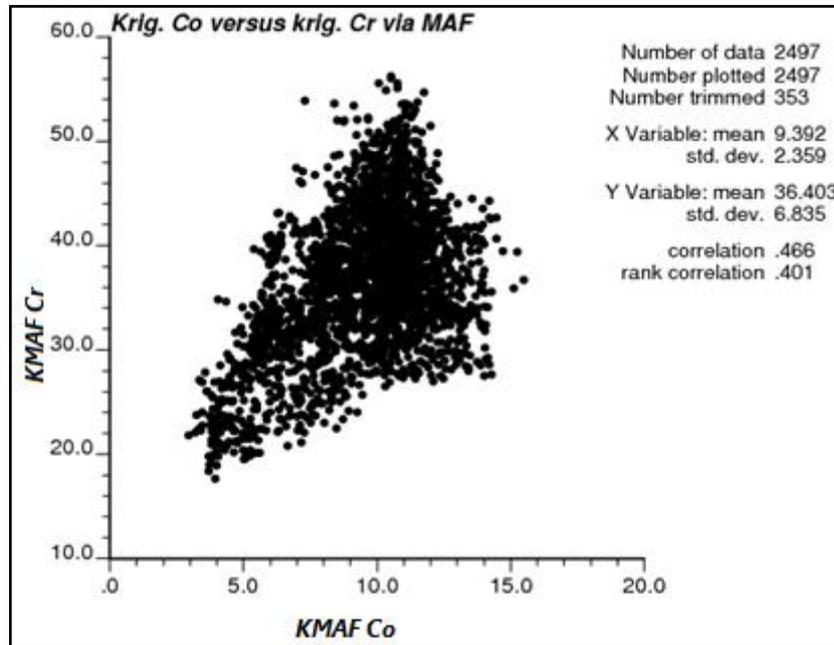


Figura 4. 57 - Gráfico de dispersão entre as variáveis krigadas Cobalto (ppm) e Cromo (ppm) utilizando a KMAF

Note, na figura 4.57, que a correlação apresentada pelo par de variáveis após a KMAF é muito próxima do valor original da correlação, $\rho_{CoCr} = 0,45$, sendo a diferença entre os dois valores de apenas 2%.

Já a correlação entre Cobalto e Níquel após a KMAF é $\rho_{estCoNi} = 0,74$, lembrando que a correlação original entre o par era $\rho_{CoNi} = 0,75$. A diferença do valor das correlações, depois de realizada, a estimativa é de 1.3%.

As correlações das variáveis, obtidas com a KMAF, é mais próxima das correlações originais que as obtidas por meio de cokrigagem. Para este mesmo par, Cobalto e Níquel, por meio de cokrigagem, a diferença obtida foi de 13%. O gráfico de dispersão entre Cobalto e Níquel estimados por meio de KMAF é apresentado na figura 4.58.

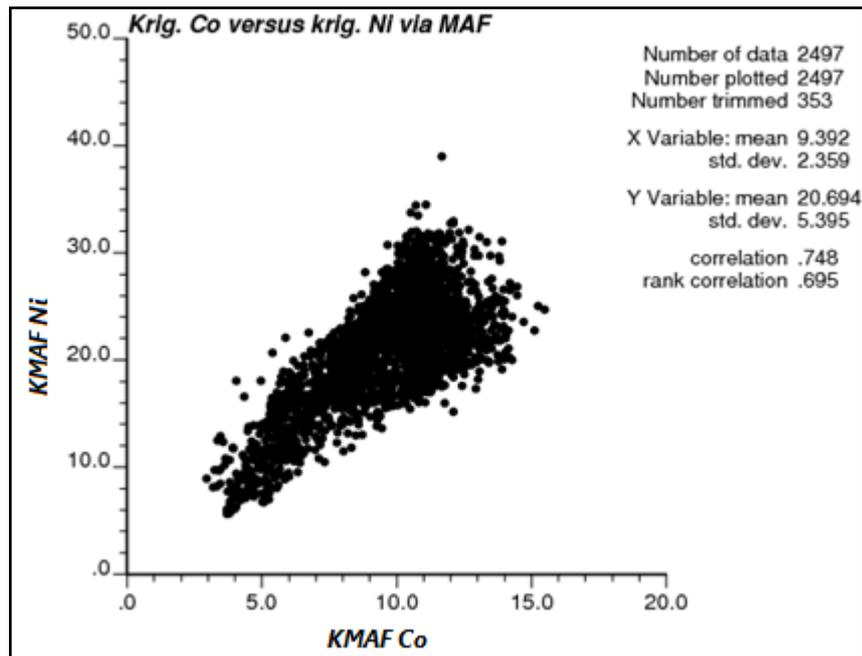


Figura 4. 58 - Gráfico de dispersão entre as variáveis krigadas Cobalto (ppm) e Níquel (ppm) utilizando a KMAF

Finalmente, para o par de variáveis Cromo e Níquel, o gráfico de dispersão é apresentado na figura 4.59.

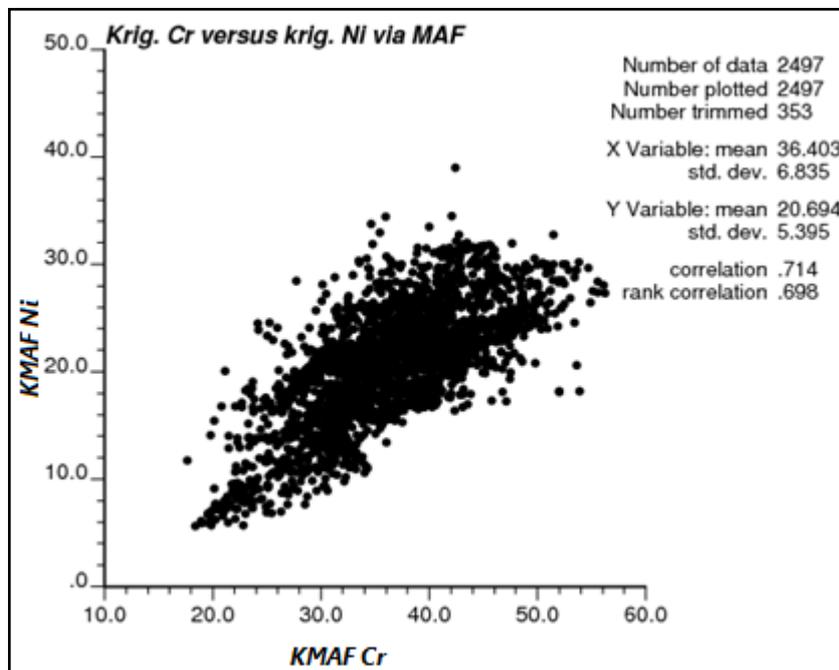


Figura 4. 59 - Gráfico de dispersão entre as variáveis krigadas Cromo (ppm) e Níquel (ppm) utilizando a KMAF

Pode-se observar, na figura 4.59, que a correlação estabelecida entre o par, após a estimativa por meio da KMAF, é $\rho_{estCoNi} = 0,71$. A correlação original do par era $\rho_{CrNi} = 0,69$, a diferença percentual entre os dois valores é de 2.8%.

Em resumo, se observa que as correlações obtidas entre as variáveis por meio de KMAF reproduzem melhor as correlações existentes originais no banco de dados do que a cokrigagem neste estudo de caso.

São construídos, também, gráficos de dispersão entre as coestimativas realizadas por meio de cokrigagem e as estimativas obtidas por meio de KMAF a fim de verificar a semelhança entre os métodos. Na figura 4.60, são apresentados os gráficos para Cádmiio e Cromo.

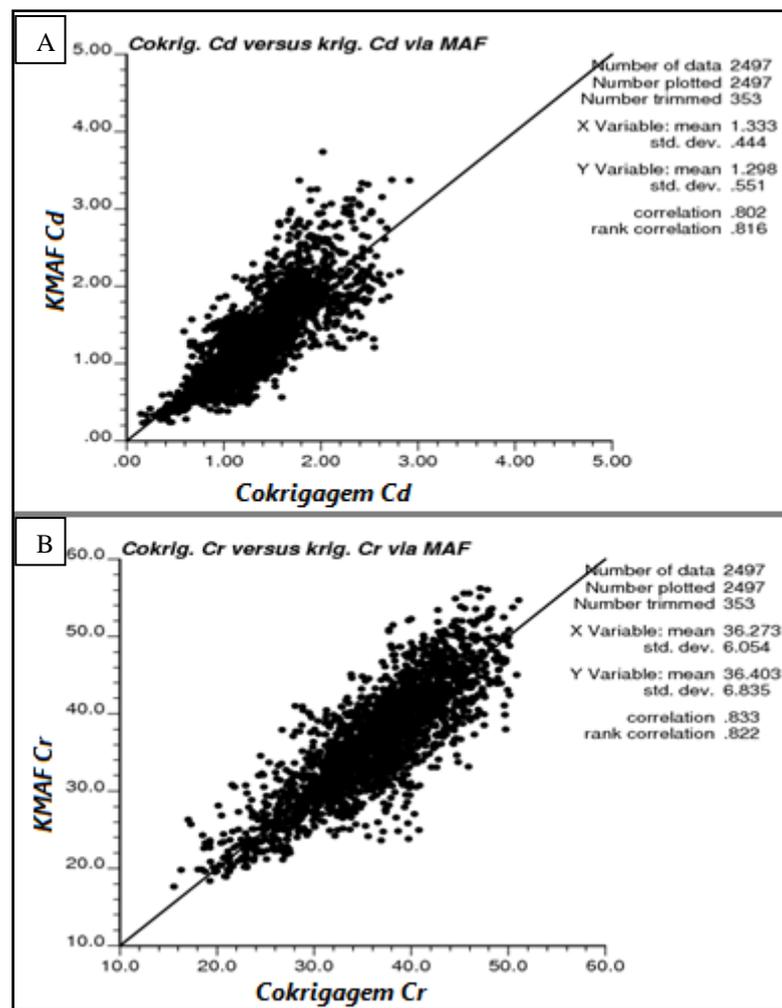


Figura 4. 60 – Gráficos de dispersão entre as variáveis cokrigadas e as variáveis estimadas por meio de KMAF. (A) Cádmiio (ppm) cokrigado versus Cádmiio (ppm) estimado por meio de KMAF; (B) Cromo (ppm) cokrigado versus Cromo (ppm) estimado por meio de KMAF

Note que as correlações apresentadas nos gráficos da figura 4.60 são altas em ambos os casos, mostrando uma grande similaridade entre os métodos. No gráfico (A) da figura 4.60 percebe-se um leve desvio da nuvem em relação à bissetriz do gráfico, evidenciando maior variabilidade entre os valores obtidos por meio de cada metodologia. A parte superior direita

da nuvem apresenta a reprodução dos teores mais altos, no eixo Y por KMAF, no eixo X por cokrigagem, é possível verificar que a metodologia KMAF aproxima-se mais do valor máximo dos dados originais do que a cokrigagem, que suavizou este resultado. Já na figura 4.61, há uma alta semelhança nas estimativas dos dois métodos para os casos Cobalto e Níquel.

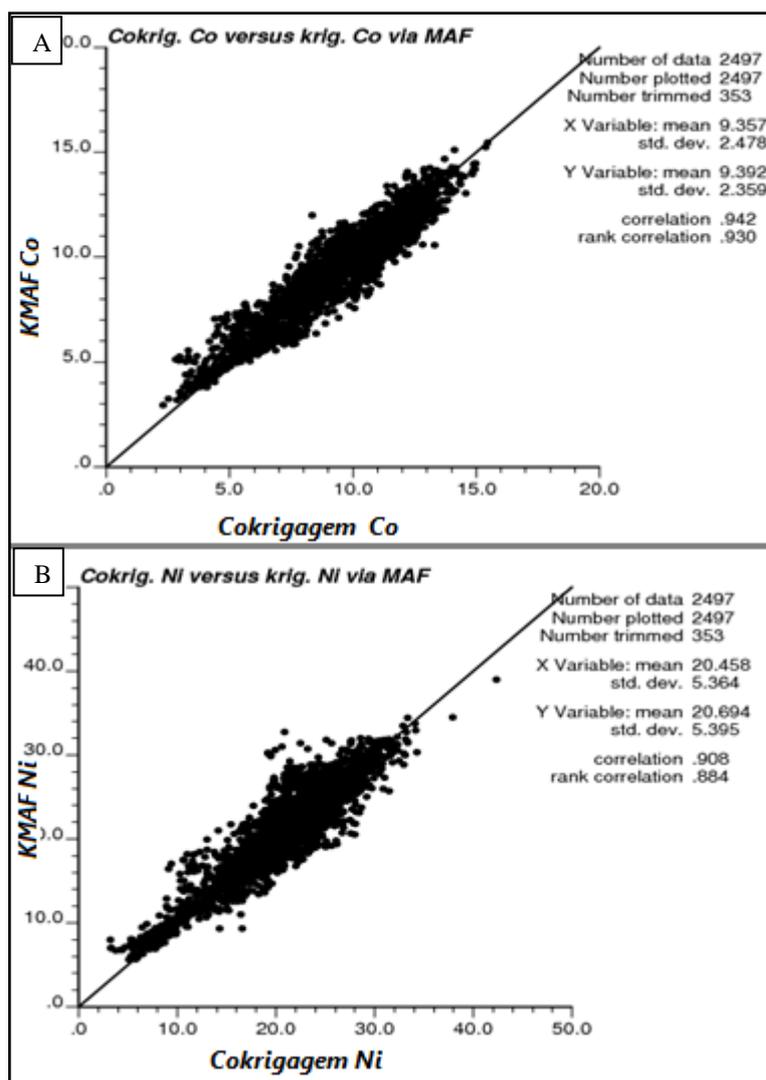


Figura 4. 61 - Gráficos de dispersão entre as variáveis cokrigadas e as variáveis estimadas por meio de KMAF. (A) Cobalto (ppm) cokrigado *versus* Cobalto (ppm) estimado por meio de KMAF; (B) Níquel (ppm) cokrigado *versus* Níquel (ppm) estimado por meio de KMAF

Pode-se observar, na figura 4.61, que há simetria em torno da bissetriz da nuvem de dispersão nos dois casos. Nas figuras 4.60 e 4.61, existem correlações entre $\rho=0,80$ e $\rho=0,94$. A correlação mais baixa apresenta-se para a estimativa do Cádmiio. A diferença entre os valores estimados entre métodos é acentuada pela estratégia de busca das amostras para estimativa ou coestimativa, pois esta não é a mesma. Caso o mesmo elipsoide de busca fosse utilizado para a cokrigagem e para a KMAF, a correlação entre os dois métodos teria se

apresentado mais alta, pois, em cada bloco estimado, teriam sido utilizadas as mesmas amostras, criando uma maior similaridade.

Por fim, é realizada uma validação visual entre os procedimentos aqui realizados. Na figura 4.62, são apresentados os mapas de localização das amostras, de estimativa realizado por meio de KMAF e o mapa de coestimativa obtido a partir da cokrigagem, da variável Cádmio.

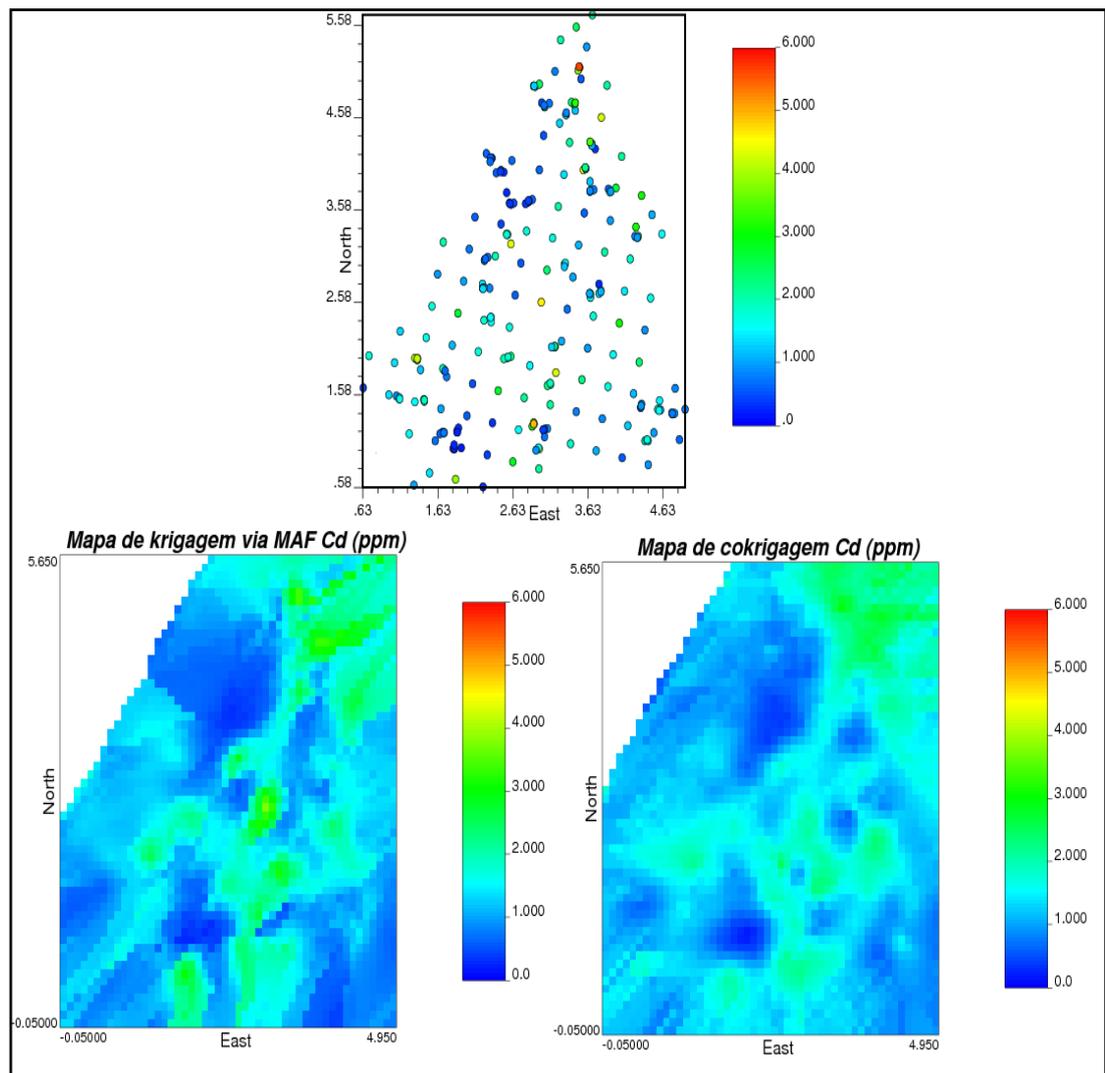


Figura 4. 62 – Mapas para validação visual dos processos KMAF e Cokrigagem. A escala de cores na figura corresponde ao valor dos teores de Cádmio em (ppm), tanto no mapa de localização quanto nos mapas de krigagem e cokrigagem

Note que, na parte superior do mapa de localização amostral, entre as coordenadas $Y = 3,58km$ e $Y = 4,58km$, há uma zona de baixos teores. Esta região é reproduzida nos mapas de KMAF e cokrigagem presentes na figura. Assim como foi realizada a análise para a variável Cádmio, também foi para outras variáveis. Nota-se uma maior suavização por parte

da cokrigagem no mapa de estimativas, isto porque o modelo ajustado não é adequado à variável Cádmiio, porém a necessidade de satisfação do MLC impediu, neste caso, que uma modelagem mais precisa fosse realizada para esta variável. Na figura 4.63, é apresentada a verificação visual das estimativas do elemento Cobalto.

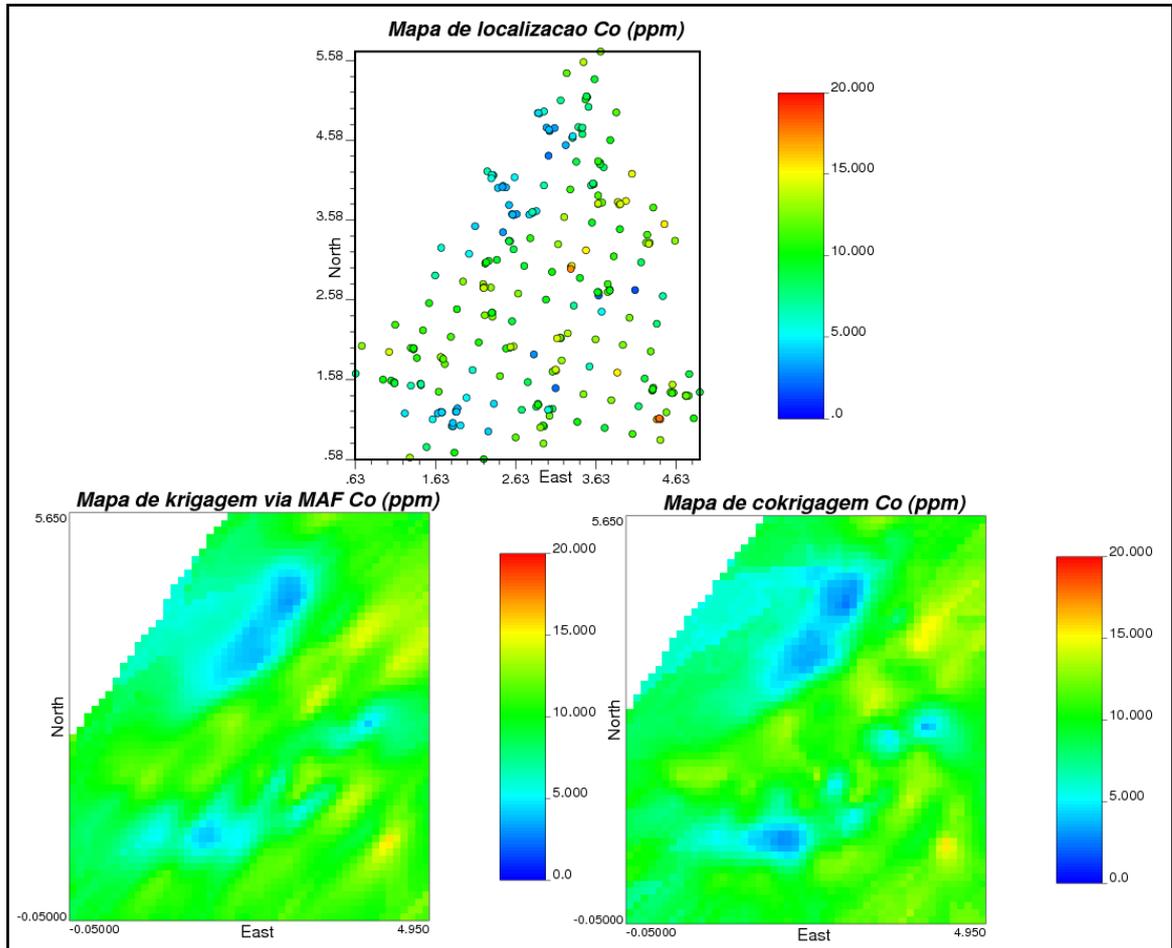


Figura 4. 63 - Mapas para validação visual dos processos KMAF e Cokrigagem. A escala de cores na figura corresponde ao valor dos teores de Cobalto em (ppm), tanto no mapa de localização quanto nos mapas de krigagem e cokrigagem

Nos mapas da figura 4.63, pode-se observar, no mapa amostral, entre as coordenadas $Y = 3,58km$ e $Y = 4,58km$, uma região de baixos teores da variável Cobalto, estas amostras são pertencentes à população de baixos teores observado no histograma da figura 4.5. Como se pode observar, este setor foi reproduzido nos mapas obtidos por meio de KMAF e cokrigagem. Nota-se, também, uma zona de teores entre 10 (ppm) e 15 (ppm) entre as coordenadas $Y = 3,58km$ e $Y = 4,58km$, e $X = 3,63km$ e $X = 4,63km$ do mapa amostral. Esta também foi reproduzida pelas estimativas e coestimativas da variável Cobalto. Na figura 4.64, são apresentados os mapas da variável Cromo.

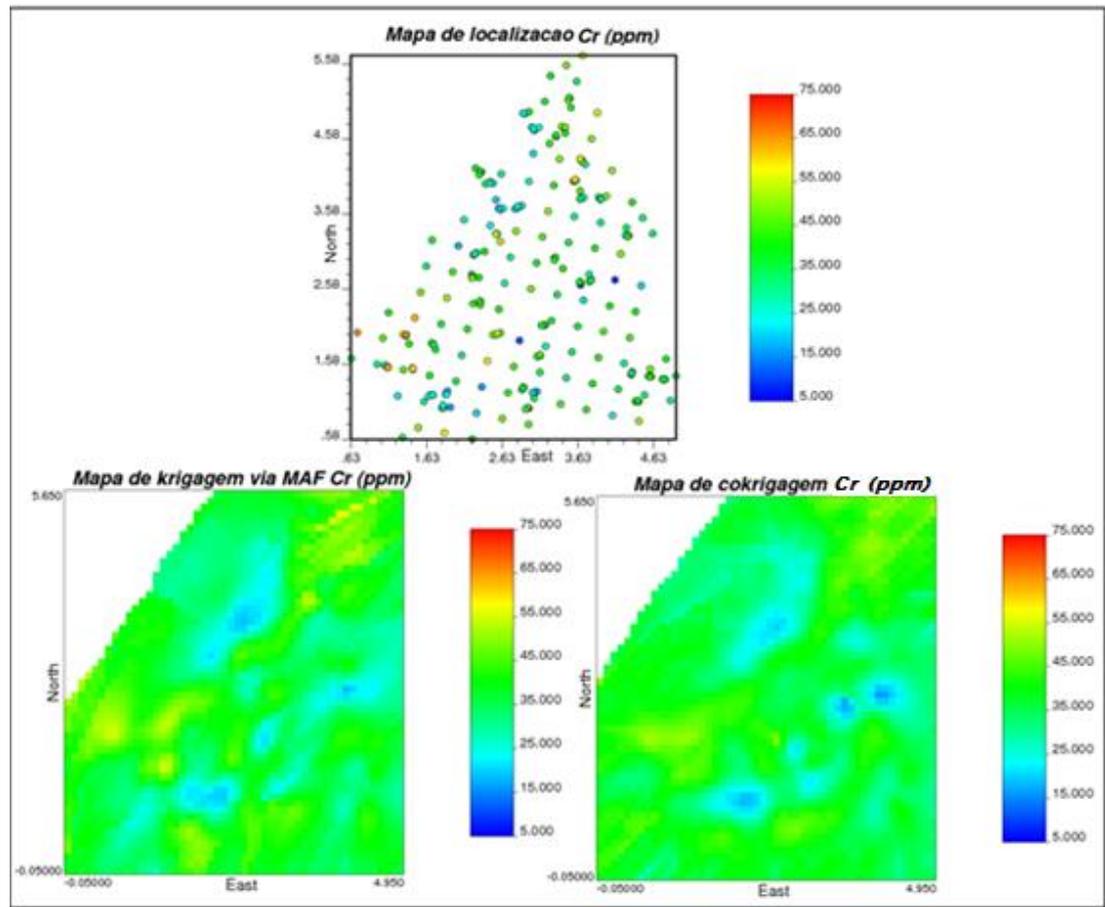


Figura 4. 64 - Mapas para validação visual dos processos KMAF e Cokrigagem. A escala de cores na figura corresponde ao valor dos teores de Cromo em (ppm), tanto no mapa de localização quanto nos mapas de krigagem e cokrigagem

Na figura 4.64, pode-se observar, entre as coordenadas $Y=1,58\text{km}$ e $Y=2,58\text{km}$ e as coordenadas $X=0,63\text{km}$ e $X=1,63\text{km}$ do mapa amostral, a presença de teores altos de Cromo e esta zona é reproduzida em ambos os mapas obtidos por meio das metodologias aplicadas. Contudo, verifica-se uma maior suavização no mapa obtido por meio de cokrigagem. Isto é devido ao ajuste do MLC, pois, para que fosse respeitada a condição de definição positiva, não foi possível ajustar o modelo a patamares mais baixos, resultando numa maior suavização das estimativas. Finalmente, na figura 4.65, são apresentados os mapas obtidos para a variável Níquel.

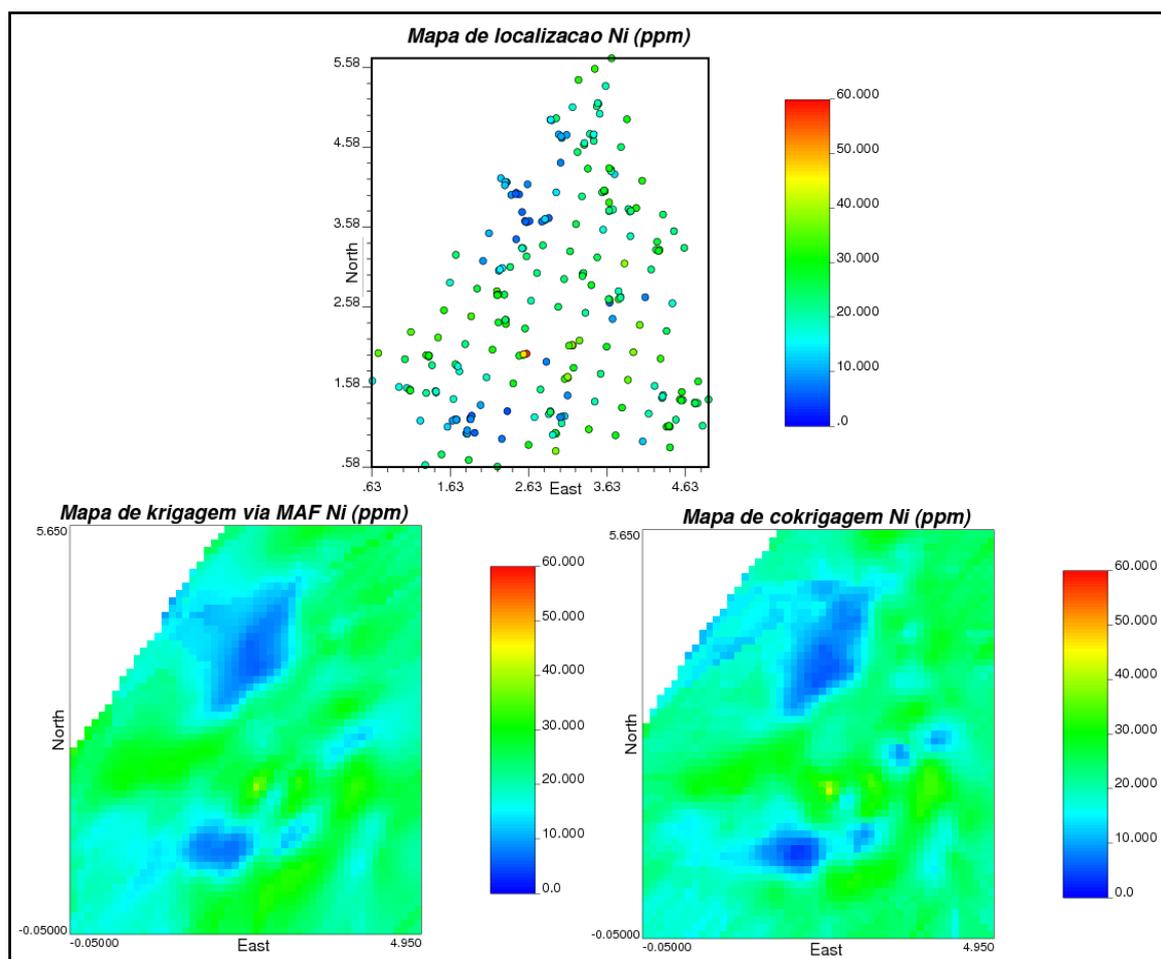


Figura 4. 65 - Mapas para validação visual dos processos KMAF e Cokrigagem. A escala de cores na figura corresponde ao valor dos teores de Níquel em (ppm), tanto no mapa de localização quanto nos mapas de krigagem e cokrigagem

No mapa amostral da figura 4.65, na coordenada X= 2,63km e entre as coordenadas Y=1,58km e Y=2,58km, se vê uma amostra teor alto do atributo Níquel. É possível verificar a reprodução deste teor nos mapas de cokrigagem e KMAF, porém mais suavizado no mapa de estimativas por meio da krigagem de fatores independentes, neste caso, o modelo ajustado para o sistema de cokrigagem mostrou-se adequado à variável Níquel. O modelo adotado utilizou como base o modelo variográfico da variável Cobalto, que é fortemente correlacionada à variável Níquel.

4.2 ESTUDO DE CASO 2 - BANCO DE DADOS SAPECADO

4.2.1 APRESENTAÇÕES DO BANCO DE DADOS

O banco de dados utilizado no presente estudo é referente às informações coletadas pela empresa Vale S.A no complexo Pico, composto pelas seguintes minas: Sapecado,

Galinheiro e Pico. O complexo está localizado entre as coordenadas -2500E a -12400E e entre 750N a 1200N no Quadrilátero Ferrífero (Spier *et al.*, 2004).

A geologia e a estratigrafia do Quadrilátero Ferrífero foram completamente descritas pela equipe da USGS-DNPM entre 1946 e 1963 (Dorr, 1969; Spier *et al.*, 2004). Os depósitos de ferro desta região estão hospedados nos Itabiritos do Supergrupo Minas, que consiste em três sequências do Paleoproterozoico. A unidade mais baixa é o grupo Caraça, a unidade intermediária é o Grupo Itabira, em que ocorre uma extensa formação ferrífera de composição variável, abrangendo a formação Cauê e Gandarela, acima está o Grupo Piracicaba e o Grupo Sabará.

No grupo Itabira, o itabirito está presente em formações ferríferas bandeadas (Banded Iron Formations – BIFs), com descontinuidades distribuídas ao longo do corpo de minério, variando de alguns centímetros a alguns metros (Dorr, 1969).

O minério de ferro no grupo Itabira é classificado como de alto teor ($Fe > 64\%$) e de baixo teor ($32\% < Fe < 64\%$), contendo poucos contaminantes, como: SiO_2 , Al_2O_3 , CaO , MgO , etc. O minério de alto teor poder ser subclassificado em: compactos (HO), friáveis (SO) e médios (MO). O minério de baixo teor poder ser subclassificado como: itabiritos ricos (RSIB), itabiritos (SIB), itabiritos limoníticos (IBLI) e itabiritos manganésiferos (IBMN). Essas subdivisões são baseadas nas características físicas e químicas do minério (Spier *et al.*, 2004).

Nesta dissertação, serão consideradas as minas de Sapecado e Sapecado Sul, utilizando os itabiritos friáveis como objeto de estudo, sendo estes os que representam a porção de interesse econômico, uma vez que o minério para lavra remanescente, após 2008, é essencialmente de baixo teor (Boezio, 2010). O modelo construído para o complexo Pico (Spier *et al.*, 2004) está exposto na figura 4.66.

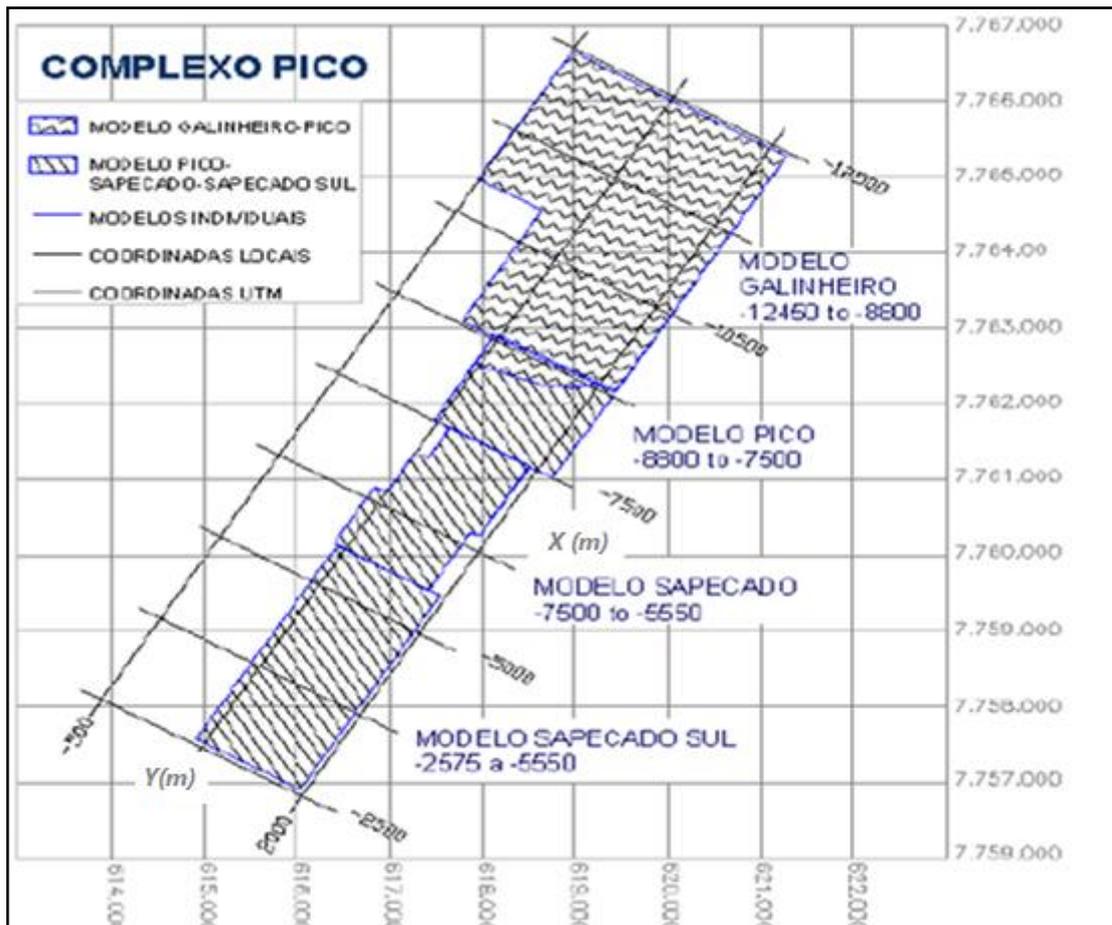


Figura 4. 66 - Vista no plano XY do modelo do Complexo Pico na direção principal do corpo de minério (Spier *et al.*, 2004)

Pode se ver, na figura 4.58, que o depósito é subdividido em três regiões: a primeira das coordenadas $X = -12450m$ a $X = -8800m$, compreendendo o modelo para a mina de Galinheiro. A partir da coordenada $X = -8800m$ até $X = -7500m$, constitui o modelo para a mina Pico e, por fim, das coordenada $X = -7500m$ a $X = -2500m$ estão compreendidos os modelos para as minas Sapecado e Sapecado Sul. Neste capítulo, serão apresentados os resultados para estimativas dos teores das espécies químicas de interesse nas minas de Sapecado e Sapecado Sul.

Nesta área, as amostras têm espaçamento de, aproximadamente, 200m na direção X e 100m na direção Y, porém, em algumas regiões, a malha amostral é mais densa, apresentando espaçamento de 50m em ambas as direções, na direção Z, o espaçamento padrão é de 8m, sendo o espaçamento mínimo de 4m (Spier *et al.*, 2004).

Dentre os elementos analisados em cada amostra, estão presentes a alumina (Al_2O_3), a sílica (SiO_2), o ferro (Fe), o fósforo (P), o manganês (Mn), perdas por calcinação (PPC),

óxido de magnésio (MgO) e dióxido de titânio (TiO₂). Os elementos que apresentaram forte correlação foram Fe e SiO₂. Para este estudo de caso, foram selecionados somente o Ferro e a Sílica como objeto de estudo a fim de analisar a aplicabilidade do método MAF como alternativa à cokrigagem.

4.2.2 ANÁLISE EXPLORATÓRIA DOS DADOS

O banco de dados que constitui a área das minas Sapecado e Sapecados Sul é composto por 1180 amostras, contendo somente os itabiritos friáveis. As variáveis analisadas estão dispostas na tabela 4.11.

Tabela 4.11 - Tabela indicando as variáveis que serão analisadas ao longo do estudo de caso. A letra "T" na denominação de cada variável indica que se trata da quantidade global.

Variável analisada	Denominação
Ferro	FET
Sílica	SIT

O mapa amostral da variável FET é apresentado na figura 4.67, com vistas nos planos XY, ZY e ZX, respectivamente.

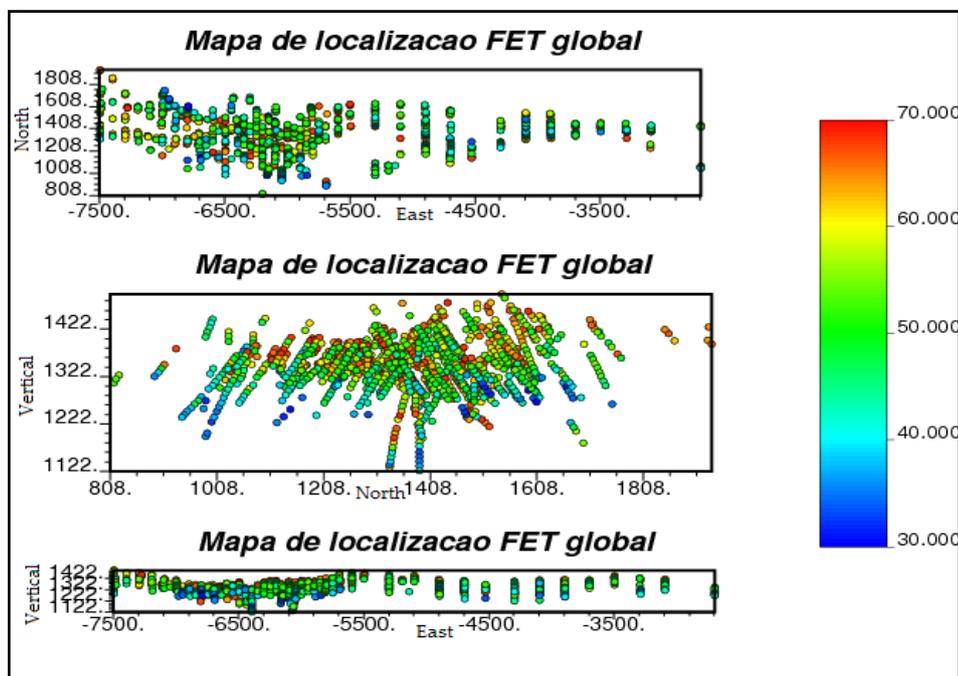


Figura 4. 67 - Mapas amostrais da variável Ferro global (FET). Mapa superior: configuração amostral no plano XY (Norte e Leste); mapa central configuração das amostras no plano ZY (Vertical e Norte); mapa inferior: configuração amostral no plano ZX (Vertical e Leste). Escala de cores representa teor de Ferro em (%)

Na figura 4.67, se vê, a partir do plano ZY, que amostras com teores altos se encontram na parte superior dos furos, e baixos teores, na parte inferior.

Os mapas de localização amostral para a variável SIT são apresentados na figura 4.68.

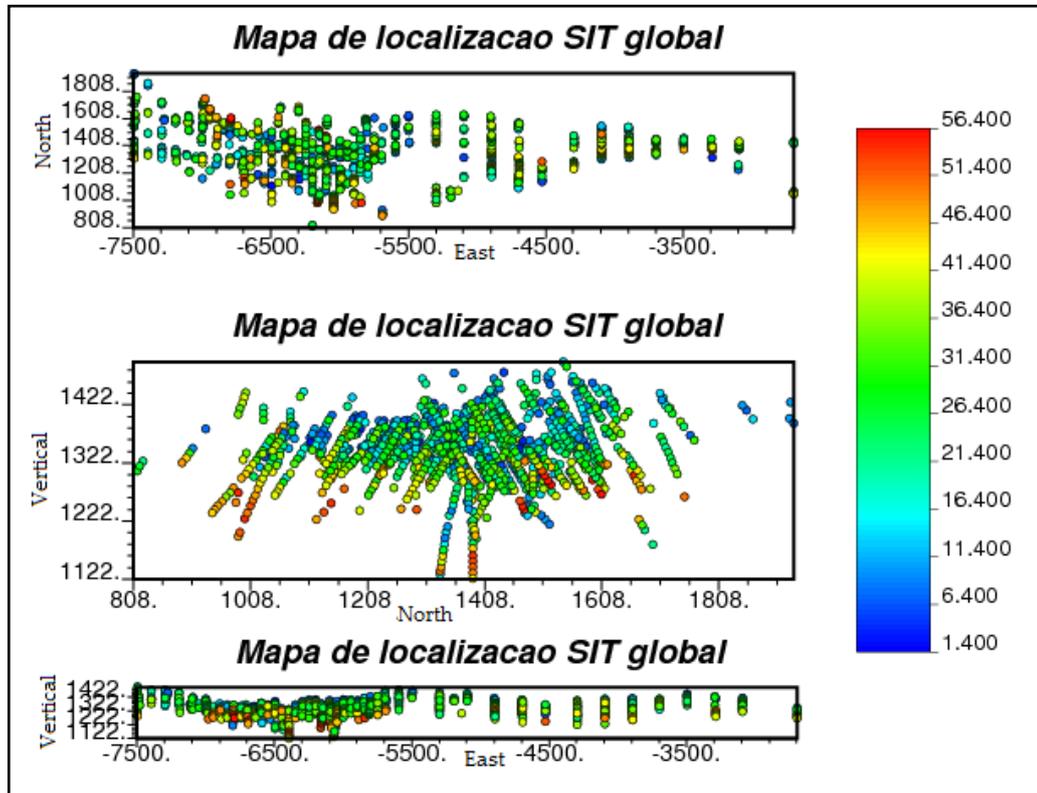


Figura 4. 68 - Mapas amostrais da variável Sílica global (SIT). Mapa superior: configuração amostral no plano XY (Norte e Leste); mapa central configuração das amostras no plano ZY (Vertical e Norte); mapa inferior: configuração amostral no plano ZX (Vertical e Leste). Escala de cores representa teor de Sílica em (%)

A partir da figura 4.68, se observa, para a variável Sílica, que a distribuição dos teores na malha é inversa à apresentada pelo Ferro na figura 4.67. Para a Sílica, Figura 4.68, os teores baixos encontram-se na parte superior dos furos.

Nas figuras 4.67 e 4.68, pode-se observar que, entre as coordenadas $X = -6500$ e $X = -5500$, há adensamento amostral. Visto isso, neste estudo de caso, como no anterior, foi utilizado o método de polígonos de Voronoi (Dirichlet, 1850) para a realização do desagrupamento e uso dos pesos obtidos para cálculo das estatísticas das espécies químicas.

Este procedimento foi realizado utilizando um modelo de blocos com dimensões $X=10m$, $Y=10m$ e $Z=5m$. Na figura 4.69, é apresentado o histograma agrupado da variável FET.

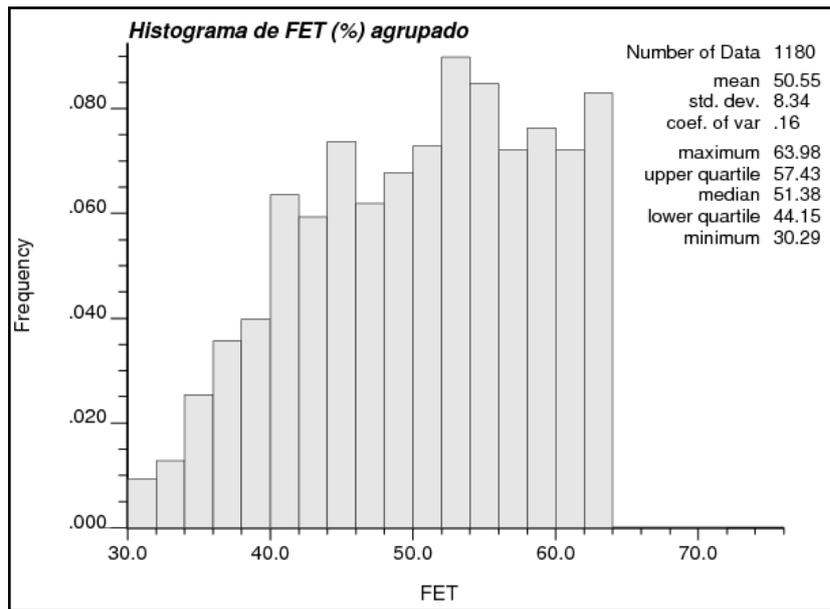


Figura 4. 69 – Histograma da variável Ferro (FET) agrupado. Teor de Ferro expresso em (%)

Note que a média de Ferro no gráfico da figura 4.69 é $m_{FET} = 50,55$. Os dados possuem uma distribuição assimétrica dos valores. O valor máximo de teor encontrado é 63,98 e mínimo é 30,29. A variável não apresenta valores extremos em seu conjunto. Na figura 4.70, é mostrado o histograma da variável FET após a realização do desagrupamento dos dados.

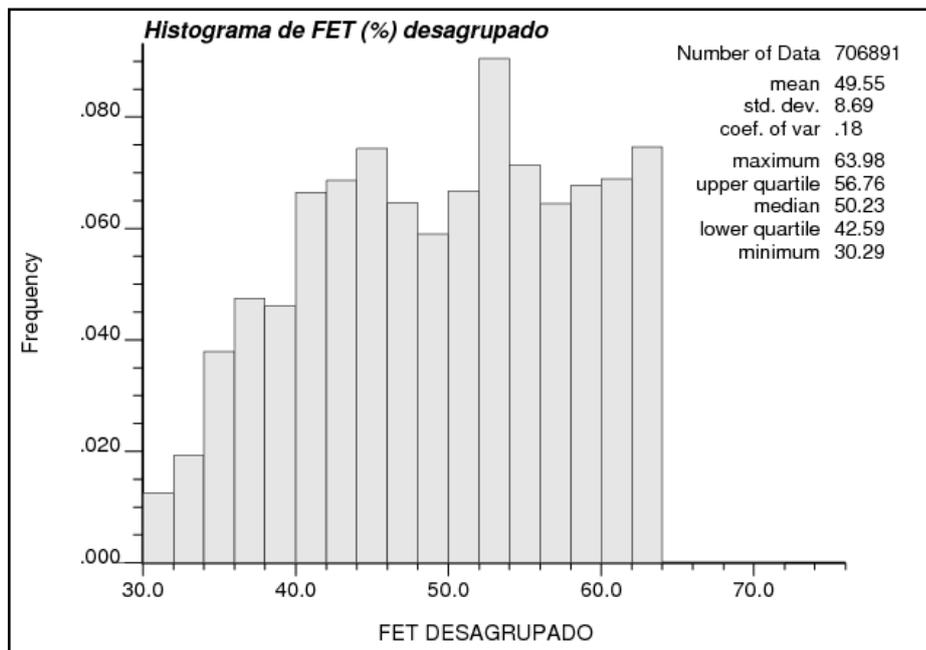


Figura 4. 70 - Histograma da variável Ferro (FET) desagrupado. Teor de Ferro expresso em (%)

A partir da figura 4.70, pode-se verificar que a média dos dados, após o processo de desagrupamento, é $m_{FETdes.} = 49,55$. A média reduz 1,97% do seu valor original. Assim como para o Ferro, foi realizado o desagrupamento para a variável Sílica. O histograma agrupado de SIT é mostrado na figura 4.71.

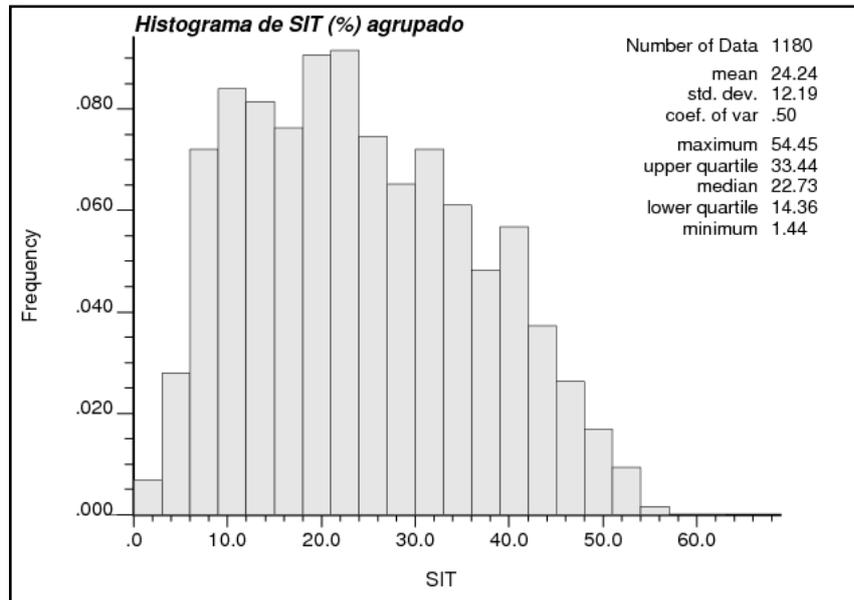


Figura 4. 71 - Histograma da variável Sílica (SIT) agrupado. Teor de Sílica expresso em (%)

Vê-se, na figura 4.71, que a média agrupada de SIT é $m_{SIT} = 24,24$, com valor máximo de 54,45 e mínimo de 1,44. A distribuição é não simétrica bem como a do Ferro. O histograma das amostras de Sílica desagrupado é mostrado na figura 4.72.

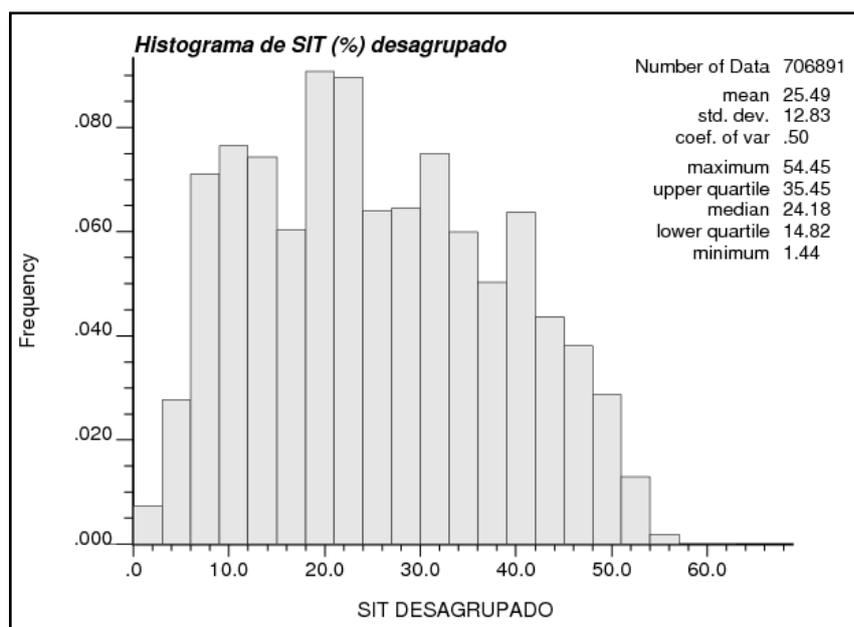


Figura 4. 72 - Histograma da variável Sílica (SIT) desagrupado. Teor de Sílica expresso em (%)

Na figura 4.72, vê-se que a média obtida para a variável Sílica, após o desagrupamento, aumenta 5,1% do valor original, para $m_{SITdes} = 25,49$. O sumário estatístico do banco de dados é apresentado na tabela 4.12.

Tabela 4. 12 - Sumário estatístico da variável Ferro global (FET), e Sílica (SIT), agrupados e desagrupados. Valores expressos em (%).

	FET	SIT
Média agrup.	50,55	24,24
Média desagrup.	49,55	25,49
Desvio padrão agrup.	8,34	12,19
Desvio padrão desagrup.	8,69	12,83
Máximo	63,98	54,45
Mínimo	30,29	1,44
Mediana agrup.	51,38	22,72
Mediana desagrup.	50,23	24,18
Quartil superior agrup.	57,4	34,44
Quartil superior desagrup.	56,76	35,45
Quartil inferior agrup.	44,15	14,36
Quartil inferior desagrup.	42,59	14,82

A dispersão entre Sílica e Ferro apresentou uma correlação $\rho = -0,98$, motivo pelo qual o estudo de caso irá focar nestas duas variáveis. O gráfico é apresentado na figura 4.73.

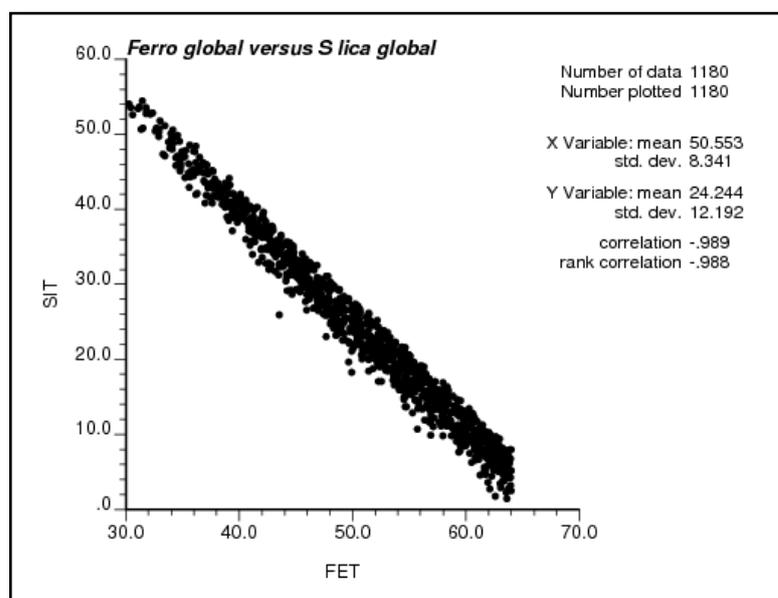


Figura 4. 73 - Gráfico de dispersão entre Ferro global e Sílica global, coeficiente de correlação -0.98

Nota-se, a partir do gráfico da figura 4.73, que a correlação entre os atributos é inversa, característica também observada nos mapas amostrais expostos nas figuras 4.67 e 4.68.

Como os atributos são fortemente correlacionados, a cokrigagem é a abordagem clássica adequada ao problema, que será apresentada na próxima seção. Posteriormente, será realizada a aplicação do método MAF e será feita a sua comparação com os resultados obtidos via cokrigagem.

4.2.3 SOLUÇÃO VIA COKRIGAGEM

A estimativa dos dados será realizada em um modelo de blocos de 50 x 20 x 10, em concordância com o modelo geoestatístico proposto no relatório interno da empresa MBR (Spier *et al.*, 2004). O sistema de cokrigagem incluirá as duas variáveis expostas no item anterior, levando à construção do modelo linear de correção (MLC) que deve ser definido positivamente.

O modelo variográfico base será o da variável Ferro, a partir do qual serão ajustados os variogramas restantes: o direto da variável Sílica e o cruzado, Ferro e Sílica.

Na construção do MLC, os modelos de cada uma das estruturas adotados e os alcances dessas são iguais para todos os variogramas de todas as variáveis diretas e cruzadas. Os parâmetros utilizados na variografia dos dados originais são apresentados na tabela 4.13.

Tabela 4. 13 - Parâmetros utilizados na construção dos variogramas experimentais direcionais dos atributos Ferro e Sílica.

Variogramas direcionais no plano XY	
Número de lags	15
Tamanho de lag	100m
Tolerância de lag	50m
Largura de banda	100m
Número de direções	8
Tolerância angular	22,5°
Variogramas direcionais na vertical	
Número de lags	15

Tamanho de lag	8m
Tolerância de lag	4m
Largura de banda	8m
Número de direções	8
Tolerância angular	22,5°

Com os parâmetros da tabela 4.13, foram construídos os variogramas experimentais, diretos e cruzados, para Ferro e Sílica. Na figura 4.74, está exposto o variograma direcional da variável Ferro nas três direções do elipsoide de anisotropia.

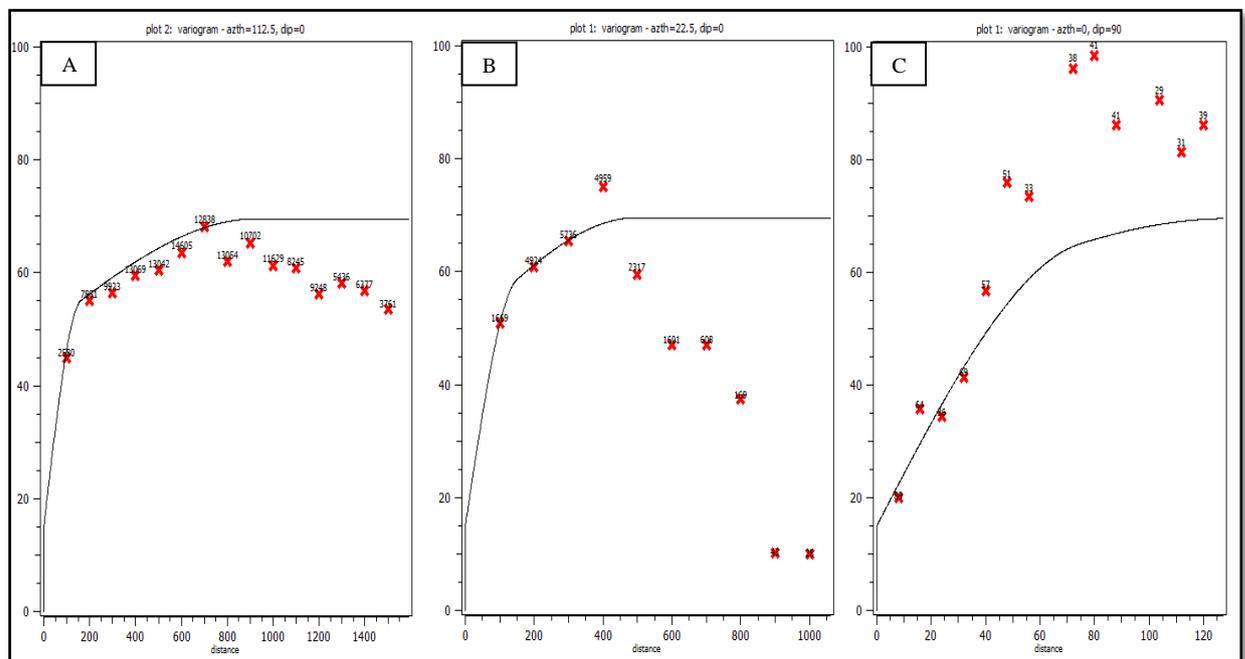


Figura 4. 74 – Variogramas experimentais direcionais da variável Ferro (FET). (A) Variograma na direção de maior continuidade espacial, N112, $dip = 0$; (B) Variograma na direção de continuidade espacial intermediária, N22, $dip = 0$; (C) Variograma na direção de menor continuidade espacial, Vertical

Os variogramas da figura 4.74 foram ajustados utilizando três estruturas, duas com modelos esféricos, além de efeito pepita. O modelo utilizado para a variável Ferro foi tomado como base na modelagem do variograma direto da Sílica e cruzado entre Ferro e Sílica, de maneira a satisfazer a condição de definição positiva do MLC. Desta forma, os dois variogramas restantes têm mesmo alcance e estruturas básicas, variando somente os valores das contribuições de *sill* em cada uma.

O modelo linear de correionalização proposto para o sistema de cokrigagem tem direção de maior alcance na N112, dip nulo, a direção de continuidade intermediária na N22,

dip nulo, a direção de menor continuidade vertical. Os parâmetros do MLC estão dispostos na tabela 4.14.

Tabela 4. 14 - Parâmetros do modelo linear de correionalização, nas direções N112, N22 e Vertical.

	N112, dip=0	N22, dip=0	Vertical
Tipo da primeira estrutura	Esférica	Esférica	Esférica
Tipo da segunda estrutura	Esférica	Esférica	Esférica
Alcance da primeira estrutura	175m	150m	72m
Alcance da segunda estrutura	1225m	500m	120m

Seja C_0 o efeito pepita atribuído aos modelos, C_1 a contribuição da primeira estrutura esférica e C_2 a contribuição para a segunda estrutura esférica; as matrizes de correionalização são apresentadas nas tabelas 4.15, 4.16 e 4.17.

Tabela 4. 15 - Matriz de correionalização para o efeito pepita C_0 . Os termos cruzados representam os valores atribuídos ao efeito pepita do modelo cruzado de correionalização.

C_0	Fe	SiO ₂
Fe	15	-24
SiO ₂	-24	40

Tabela 4. 16 - Matriz de correionalização para primeira estrutura C_1 . Os termos cruzados representam os valores atribuídos à contribuição da primeira estrutura do modelo cruzado de correionalização.

C_1	Fe	SiO ₂
Fe	35	-30.81
SiO ₂	-30.81	74.38

Tabela 4. 17 - Matriz de correionalização para a segunda estrutura C₂. Os termos cruzados representam os valores atribuídos à contribuição da segunda estrutura do modelo cruzado de correionalização.

C ₂	Fe	SiO ₂
Fe	19.55	-25.76
SiO ₂	-25.76	34.34

Os modelos variográficos, diretos e cruzados utilizados no processo de estimativa são os das equações 4.16 a 4.18.

$$\gamma_{FETFET} = 15 + 35Sph\left(\frac{175m}{N112} \frac{150m}{N22} \frac{72m}{D90}\right) + 19,55Sph\left(\frac{1225m}{N112} \frac{500m}{N22} \frac{120m}{D90}\right) \quad (4.16)$$

$$\gamma_{FETSIT} = -24 - 30,81Sph\left(\frac{175m}{N112} \frac{150m}{N22} \frac{72m}{D90}\right) - 25,76Sph\left(\frac{1225m}{N112} \frac{500m}{N22} \frac{120m}{D90}\right) \quad (4.17)$$

$$\gamma_{SITSIT} = 40 + 74,38Sph\left(\frac{175m}{N112} \frac{150m}{N22} \frac{72m}{D90}\right) + 34,34Sph\left(\frac{1225m}{N112} \frac{500m}{N22} \frac{120m}{D90}\right) \quad (4.18)$$

sendo γ_{FETFET} o variograma direto da variável Ferro, γ_{FETSIT} o variograma cruzado entre as variáveis Ferro e Sílica, e γ_{SITSIT} o variograma direto da variável Sílica.

Foi realizada cokrigagem ordinária em blocos. Os parâmetros de coestimativa estão dispostos na Tabela 4.18.

Tabela 4. 18 - Tabela de parâmetros de cokrigagem utilizados na coestimativa da variável Ferro em que este é a variável primária.

Tipo de coestimativa	Cokrigagem ordinária	
Variável primária	FET	
Variável secundária	SIT	
Discretização em X,Y e Z	4, 4 e 4	
Elipsoide de busca (variável Primária)	Mínimo de dados condicionantes	3
	Máximo de dados condicionantes	20
	Azimute	112,5
	Dip	0

	Rake	0
	Raio máximo de busca	980m
	Raio médio de busca	420m
	Raio mínimo de busca	96m
	Mínimo de dados condicionantes	3
	Máximo de dados condicionantes	20
Elipsoide de busca (variável secundária)	Azimute	112,5
	Dip	0
	Rake	0
	Raio máximo de busca	980m
	Raio médio de busca	420m
	Raio mínimo de busca	96m

A variável Sílica foi estimada também por meio de cokrigagem, utilizando os mesmos variogramas calculados na estimativa do ferro, uma vez que é o mesmo sistema de variáveis. Porém, na estimativa da Sílica, esta foi usada como variável primária e o Ferro como variável secundária. Os parâmetros de cokrigagem são expostos na tabela 4.19.

Tabela 4. 19 - Tabela de parâmetros de cokrigagem utilizados na coestimativa da variável Sílica em que esta é a variável primária.

Tipo de coestimativa	Cokrigagem ordinária	
Variável primária	SIT	
Variável secundária	FET	
Discretização em X,Y e Z	4, 4 e 4	
	Mínimo de dados condicionantes	3
Elipsoide de busca (variável Primária)	Máximo de dados condicionantes	20
	Azimute	112,5
	Dip	0
	Rake	0

Elipsoide de busca (variável secundária)	Raio máximo de busca	980m
	Raio médio de busca	420m
	Raio mínimo de busca	96m
	Mínimo de dados condicionantes	3
	Máximo de dados condicionantes	20
	Azimute	112,5
	Dip	0
	Rake	0
	Raio máximo de busca	980m
	Raio médio de busca	420m
	Raio mínimo de busca	96m

Com os parâmetros expostos na tabela 4.18 e 4.19, foi realizada a cokrigagem no banco de dados.

4.2.4 SOLUÇÃO VIA KMAF

Assim como no estudo de caso 1, é aplicada a solução KMAF ao banco de dados. Primeiramente, os dados foram submetidos à transformação linear apresentada na seção 2.5 do segundo capítulo desta dissertação. A transformação linear foi aplicada utilizando o algoritmo apresentado no capítulo 3 do presente trabalho.

Da mesma maneira que para o estudo de caso 1, definiu-se o *lag* até onde a decorrelação deve ocorrer. Foi realizado o mesmo teste ao executado no estudo de caso anterior (Jura), verificando as matrizes \mathbf{a} (equação 2.28, seção 2.5), resultantes para múltiplos valores de *lag* entre 50 e 600m. O *lag* encontrado onde não há mudanças nas entradas da matriz \mathbf{a} foi de 200m.

A decorrelação calculada no algoritmo usando as variâncias e covariâncias é omnidirecional, ou seja, é calculada ignorando uma direção preferencial, portanto, não é necessário determinar o ângulo da direção no qual a decorrelação deve ocorrer. Os fatores MAF1 e MAF2 são os obtidos depois de realizada a transformação MAF nos dados originais.

O gráfico de dispersão dos fatores MAF obtidos na transformação é apresentado na figura 4.75.

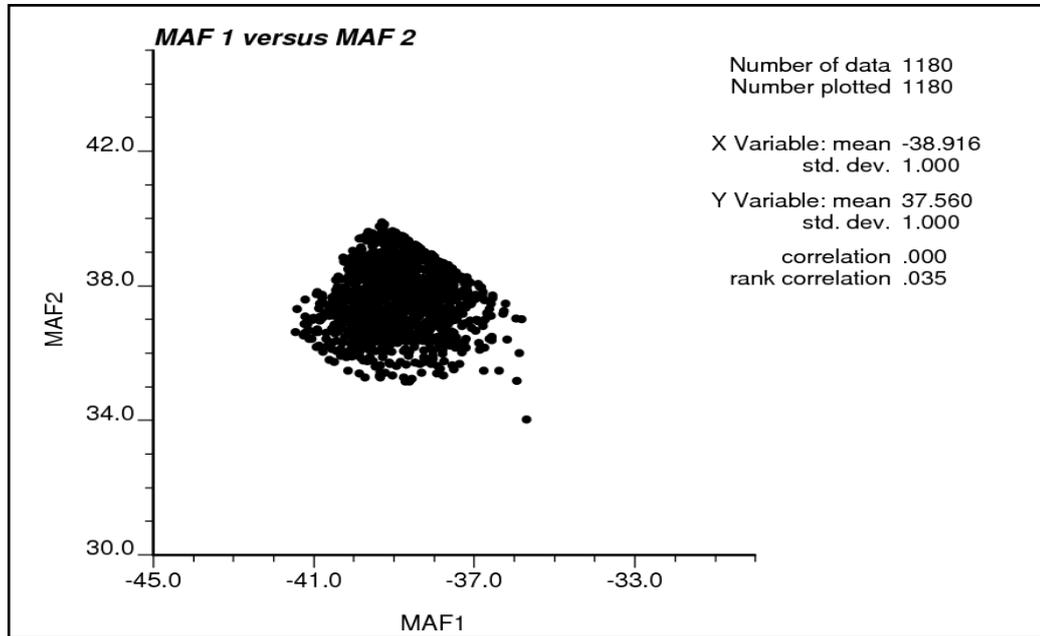


Figura 4.75 - Gráfico de dispersão entre os fatores de autocorrelação obtidos por meio da transformação linear dos dados originais, MAF1 e MAF2

Como se pode observar no gráfico da figura 4.75, os fatores não apresentam correlação entre si, de modo que a transformação linear foi operada de maneira correta. Os fatores também devem apresentar variância (unitária). Esta propriedade pode ser observada na figura 4.75. Nas figura 4.76 e 4.77, é observada a descorrelação espacial dos fatores MAF1 e MAF2. A figura 4.76 corresponde ao correlograma experimental omnidirecional dos fatores MAF1 e MAF2 no plano XY. A figura 4.77 representa o correlograma experimental omnidirecional vertical dos fatores MAF1 e MAF2, em ambas as figuras a descorrelação é observada.

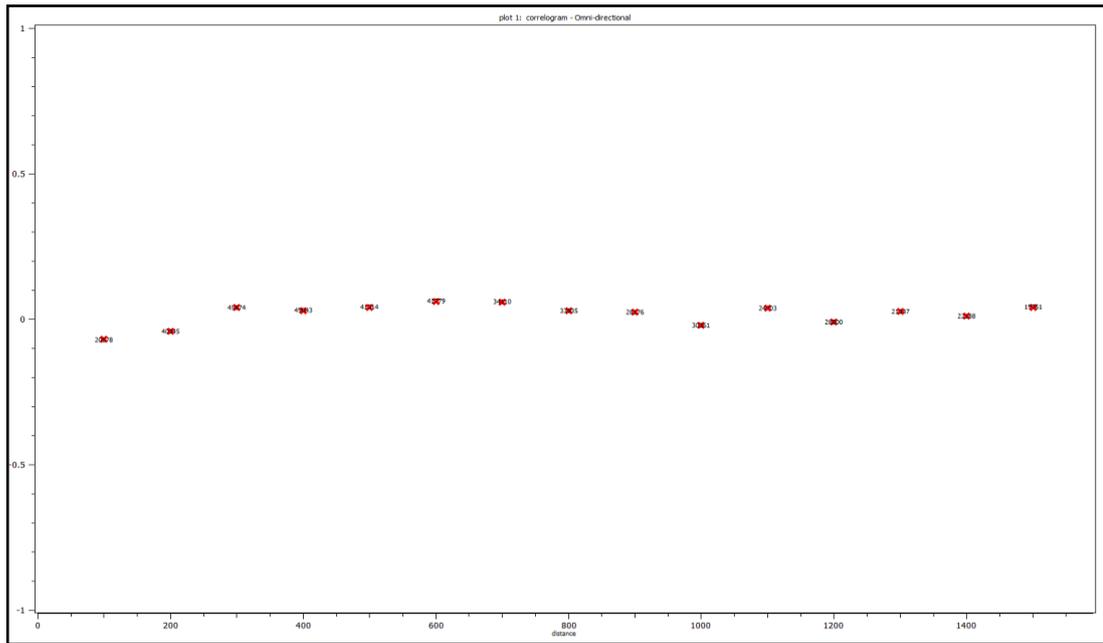


Figura 4. 76 - Correlograma experimental omnidirecional dos fatores MAF1 e MAF2 obtidos por meio da transformação MAF aplicada ao banco de dados Sapecado

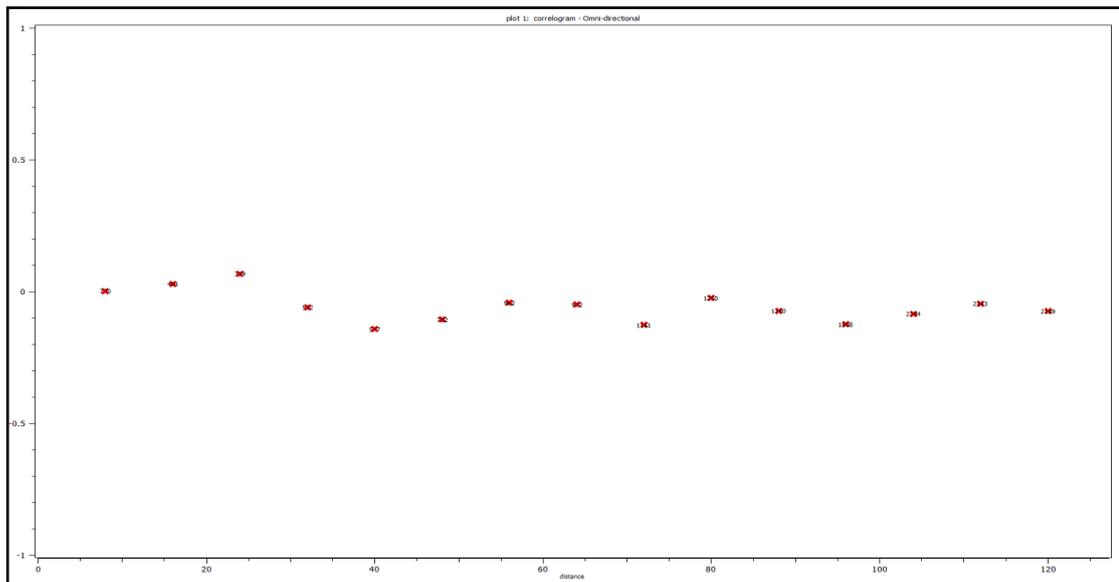


Figura 4. 77 - Correlograma experimental omnidirecional vertical dos fatores MAF1 e MAF2 obtidos por meio da transformação MAF aplicada ao banco de dados Sapecado

Antes de realizar a krigagem ordinária nos fatores independentes, foi realizada a retrotransformação nos fatores para verificar se estes não estavam gerando valores distintos aos originais quando retornando o espaço dos dados Ferro e Sílica. O gráfico de dispersão entre a retrotransformação ao Ferro e a variável Ferro é apresentado na figura 4.78. O gráfico de dispersão entre a retrotransformação à Sílica e a variável Sílica é apresentado na figura 4.79.

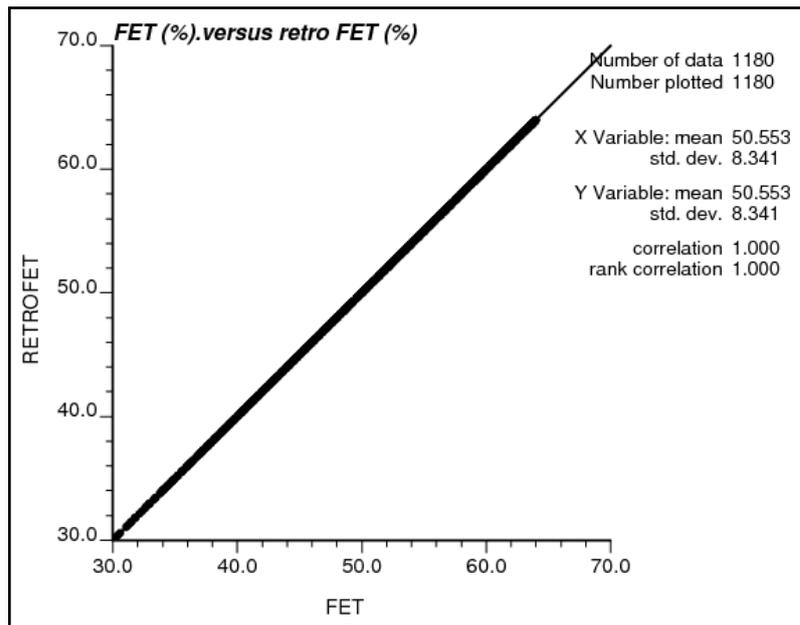


Figura 4. 78 - Gráfico de dispersão de teste entre a variável FET (%) original, e a retrotransformação do fator MAF ao FET (%)

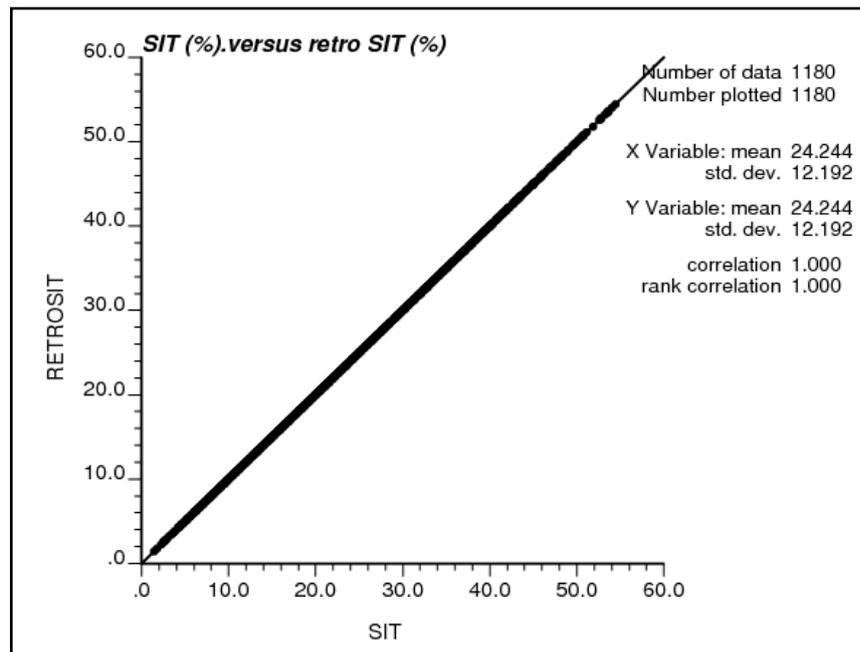


Figura 4. 79 - Gráfico de dispersão de teste entre a variável SIT (%) original, e a retrotransformação do fator MAF à SIT (%)

Percebe-se a correlação máxima entre as retrotransformações e os dados originais, nas Figuras 4.78 e 4.79. Uma vez que foi verificado que a transformação não gera anomalias no banco de dados, os fatores MAF foram variografados de maneira independente e estimados. Os variogramas dos fatores MAF1 e MAF2 diretos foram desenvolvidos com os parâmetros apresentados na tabela 4.20.

Tabela 4. 20 – Parâmetros utilizados na construção dos variogramas experimentais direcionais dos fatores independentes MAF1 e MAF2.

Variogramas direcionais no plano XY	
Número de lags	15
Tamanho de lag	100m
Tolerância de lag	50m
Largura de banda	100m
Número de direções	8
Tolerância angular	22,5°
Variogramas direcionais na vertical	
Número de lags	15
Tamanho de lag	8m
Tolerância de lag	4m
Largura de banda	8m
Número de direções	8
Tolerância angular	22,5°

Serão construídos somente modelos de regionalização utilizando os variogramas diretos experimentais uma vez que não há correlação entre os fatores. O variograma experimental do fator MAF1 é apresentado na figura 4.80.

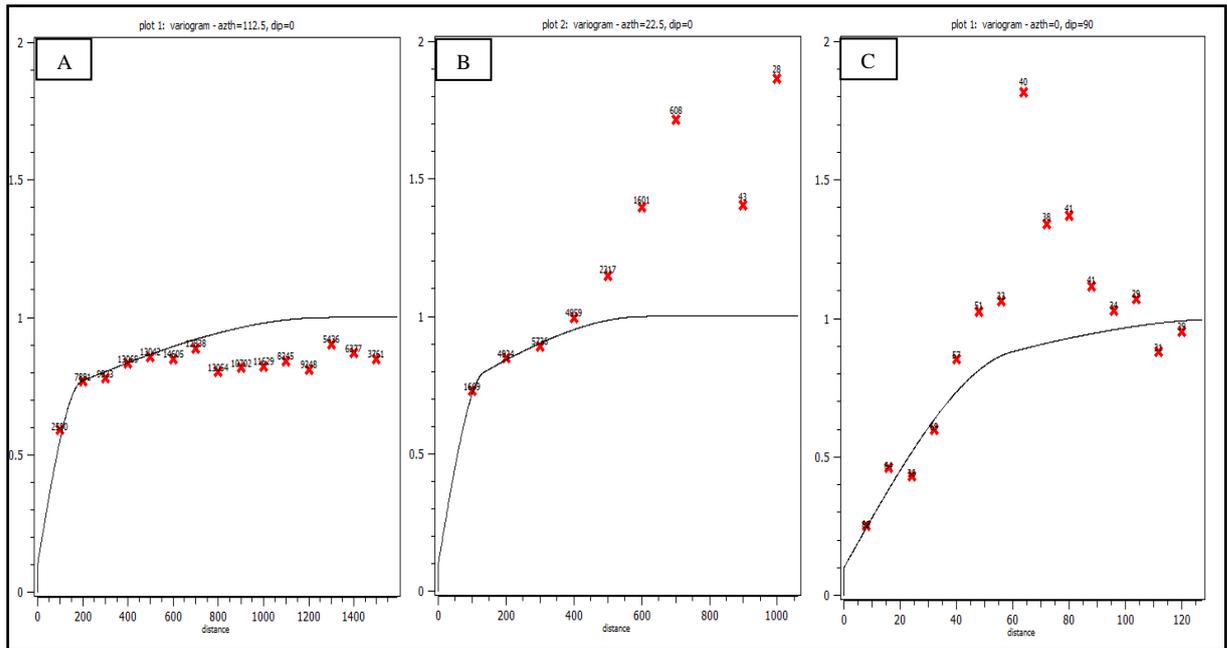


Figura 4.80 – Variogramas experimentais direcionais diretos do fator MAF1. (A) Variograma na direção de maior continuidade espacial, N112 $dip = 0$; (B) Variograma na direção de continuidade espacial intermediária, N22 $dip = 0$; (C) Variograma na direção de menor continuidade espacial, vertical

O variograma experimental direto do fator MAF1, mostrado na figura 4.80, foi modelado utilizando duas estruturas esféricas, além do efeito pepita. Ao efeito pepita foram atribuídos 10% da variância total do fator, ou seja, $C_0 = 0,1$. À primeira estrutura esférica, foi determinada uma contribuição de 60% da variância, $C_1 = 0,6$, à segunda estrutura esférica, os 30% restantes da variância, $C_2 = 0,3$. O alcance máximo obtido foi de 1300m na direção N112, dip nulo. O alcance da direção intermediária é de 625m na N22, $dip = 0$, e o alcance da direção de menor continuidade foi de 130m na vertical.

No variograma experimental do fator MAF2, as direções de continuidade obtidas foram as mesmas que as obtidas para o fator MAF1, como pode ser visto na figura 4.81.

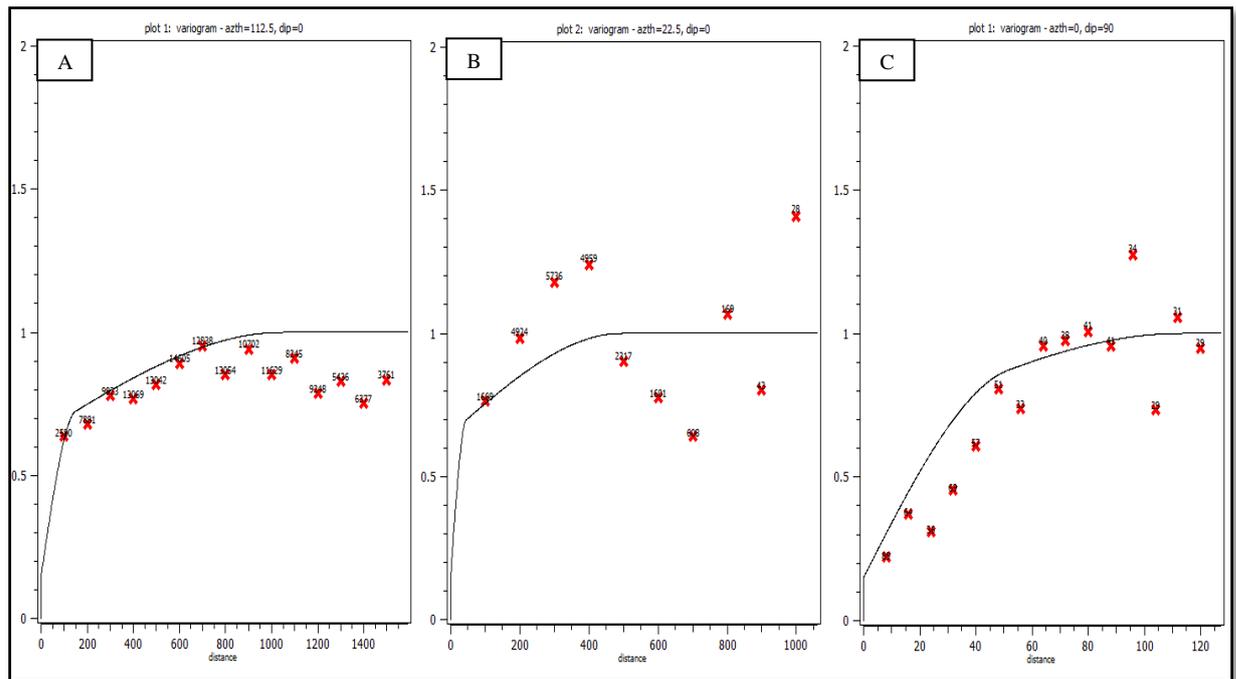


Figura 4.81 – Variogramas experimentais direcionais diretos do fator MAF2. (A) Variograma na direção de maior continuidade espacial, N112 $dip = 0$; (B) Variograma na direção de continuidade espacial intermediária, N22 $dip = 0$; (C) Variograma na direção de menor continuidade espacial, vertical

No variogramas experimentais da figura 4.81, o efeito pepita é de 15% da variância, $C_0 = 0,15$. À contribuição da primeira estrutura esférica 50% da variância, $C_1 = 0,50$, e à contribuição da segunda estrutura esférica 35% da variância, $C_2 = 0,35$.

Os modelos variográficos são apresentados nas tabelas 4.21 e 4.22 para os fatores MAF1 e MAF2, respectivamente. Diferentemente do que foi observado em alguns fatores independentes do estudo de caso realizado na seção 4.1, os agora apresentados tiveram suas direções iguais às modeladas para os dados originais.

Tabela 4.21 - Tabela com os parâmetros do modelo de regionalização do fator independente MAF1 nas direções N112 $dip = 0$, N22 $dip = 0$ e vertical.

	N112, $dip=0$	N22, $dip=0$	Vertical
Tipo da primeira estrutura	Esférica	Esférica	Esférica
Tipo da segunda estrutura	Esférica	Esférica	Esférica
Alcance da primeira estrutura	200m	135m	56m
Alcance da segunda estrutura	1300m	625m	130m

Tabela 4. 22 - Tabela com os parâmetros do modelo de regionalização para fator independente MAF2 nas direções N112 dip = 0, N22 dip = 0 e vertical.

	N112, dip=0	N22, dip=0	Vertical
Tipo da primeira estrutura	Esférica	Esférica	Esférica
Tipo da segunda estrutura	Esférica	Esférica	Esférica
Alcance da primeira estrutura	150m	50m	48m
Alcance da segunda estrutura	1050m	500m	105m

Os modelos variográficos utilizados no processo de estimativa são:

$$\gamma_{MAF1} = 0,1 + 0,6Sph\left(\frac{200m}{N112} \frac{135m}{N22} \frac{56m}{D90}\right) + 0,3Sph\left(\frac{1300m}{N112} \frac{625m}{N22} \frac{130m}{D90}\right) \quad (4.19)$$

$$\gamma_{MAF2} = 0,15 + 0,5ph\left(\frac{150m}{N112} \frac{50m}{N22} \frac{48m}{D90}\right) + 0,35Sph\left(\frac{1050m}{N112} \frac{500m}{N22} \frac{105m}{D90}\right) \quad (4.20)$$

Foi realizada a krigagem ordinária em blocos e nas tabelas 4.23 e 4.24 estão apresentados os parâmetros de krigagem para os fatores MAF1 e MAF2, respectivamente.

Tabela 4. 23 - Parâmetros de krigagem utilizados no processo de estimativa do fator MAF1.

Tipo de estimativa	Krigagem ordinária	
Variável primária	MAF1	
Discretização em X, Y e Z	4, 4 e 4	
Elipsoide de busca	Mínimo de dados condicionantes	3
	Máximo de dados condicionantes	20
	Azimute	112,5
	Dip	0
	Rake	0
	Raio máximo de busca	980m
	Raio médio de busca	420m
	Raio mínimo de busca	96m

Tabela 4. 24 - Parâmetros de krigagem utilizados no processo de estimativa do fator MAF2.

Tipo de estimativa	Krigagem ordinária	
Variável primária	MAF2	
Discretização em X,Y e Z	4, 4 e 4	
Elipsoide de busca	Mínimo de dados condicionantes	3
	Máximo de dados condicionantes	20
	Azimute	112,5
	Dip	0
	Rake	0
	Raio máximo de busca	980m
	Raio médio de busca	420m
	Raio mínimo de busca	96m.

Após a krigagem de cada fator, é realizada a retrotransformação dos dados, para permitir que os resultados possam ser analisados em termos FET(%) e SIT(%) obtidas por meio da KMAF.

4.2.5 ANÁLISES DOS RESULTADOS - ESTUDO DE CASO SAPECADO

Nesta seção, serão apresentados os resultados obtidos para o estudo de caso 2, realizando as mesmas validações feitas na seção 4.1.4, do estudo de caso 1.

Primeiramente, será verificada a reprodução das médias globais obtidas por meio dos métodos de coestimativa e estimativa, aplicados ao banco de dados. Será calculado também o erro percentual relativo ($err_{\%}$) entre as médias. Assim como para o estudo de caso 1, foram testados exaustivamente diferentes parâmetros nas estratégias de krigagem e cokrigagem, sendo mantidos os que obtiveram as melhores estimativas. Os valores são mostrados na tabela 4.25.

Tabela 4. 25 - Sumário da comparação entre as médias dos dados originais desagrupadas e as médias obtidas por meio de cada método aplicado. Além disso, cálculo do erro relativo percentual das médias globais de cada variável do estudo de caso 2, banco de dados Sapecado.

Variável	m_{od}	m_{est} Cokrigagem	m_{est} KMAF	$err_{\%}$ Cokrigagem	$err_{\%}$ KMAF
FET(%)	49,55	49,15	49,23	-0,80%	-0,64%
SIT(%)	25,49	25,78	25,61	1,13%	0,47%

Pode ser visto na tabela 4.25 que tanto a cokrigagem quanto a KMAF reproduziram as médias globais dos dados com o erro relativo percentual baixo, sendo que o erro relativo mais alto foi obtido na coestimativa da Sílica. As médias globais das variáveis foram reproduzidas com sucesso, por ambos os métodos, tanto o clássico quanto o proposto nesta dissertação.

Para verificar os resultados obtidos com respeito à tendência local das médias com os procedimentos aplicados, foi realizada uma análise de deriva. Assim como para o estudo de caso 1 (Jura), para o estudo de caso 2 (Sapecado) foi construído um delimitador do corpo de minério a partir do método de polígonos de Voronoi a fim de filtrar as áreas de extrapolação das estimativas e coestimativas. Na figura 4.82, é apresentada a análise de deriva para a variável FET (%).

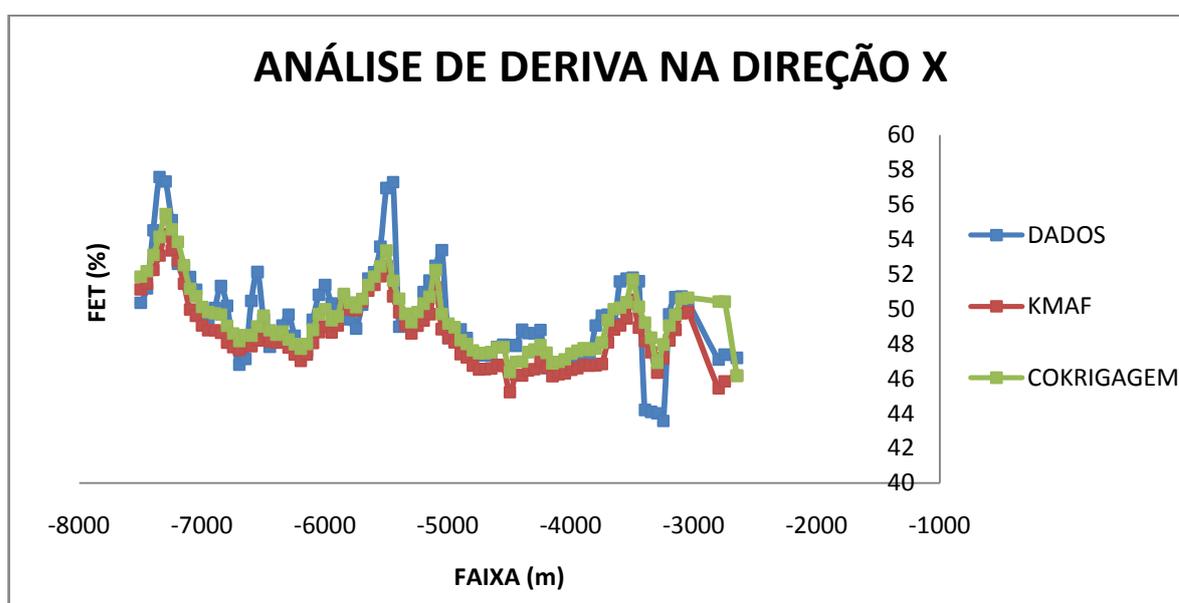


Figura 4. 82 - Análise de deriva na direção X (Leste) para a variável FET (%). Faixas divididas em intervalos de 50m

Pode-se observar, na figura 4.82, que ambos os métodos acompanham a tendência da média dos dados desagrupados. Em torno de $X = -3000m$, há uma superestimativa da média

tanto pela cokrigagem quanto pela KMAF. Na figura 4.83, é apresentada a análise de deriva na direção Y (Norte).

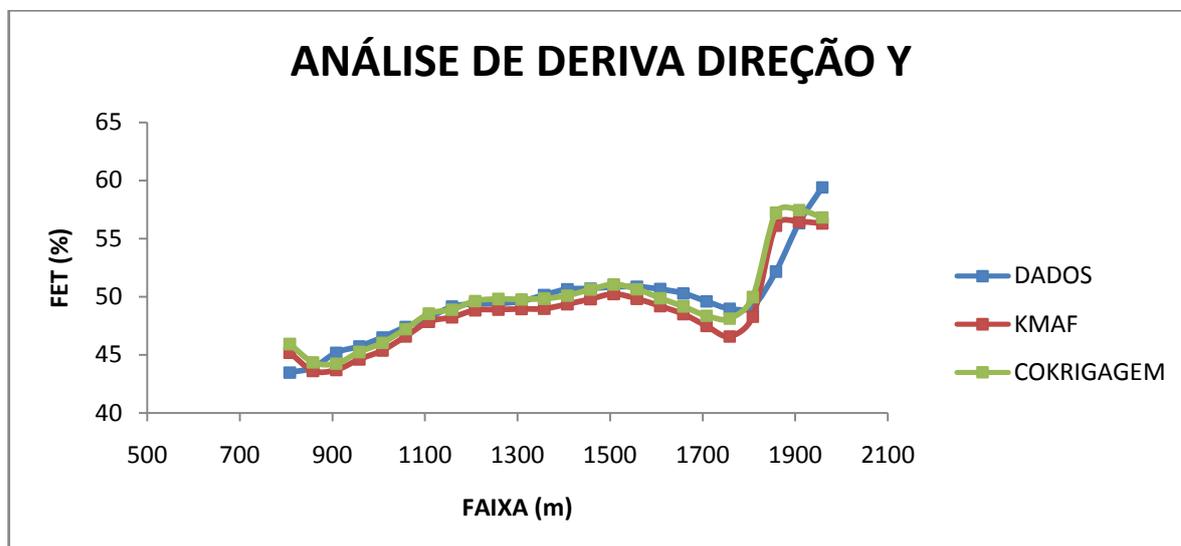


Figura 4. 83 - Análise de deriva na direção Y (Norte) para a variável FET (%). Faixas divididas em intervalos de 50m

Vê-se que a curva da média de KMAF passa por baixo das curvas dos dados e da cokrigagem, indicando a subestimativa recorrente de valores. Entretanto, o comportamento da estimativa por meio de KMAF é muito próximo ao obtido por meio do método clássico. Na figura 4.84, é apresentada análise de deriva na direção vertical.

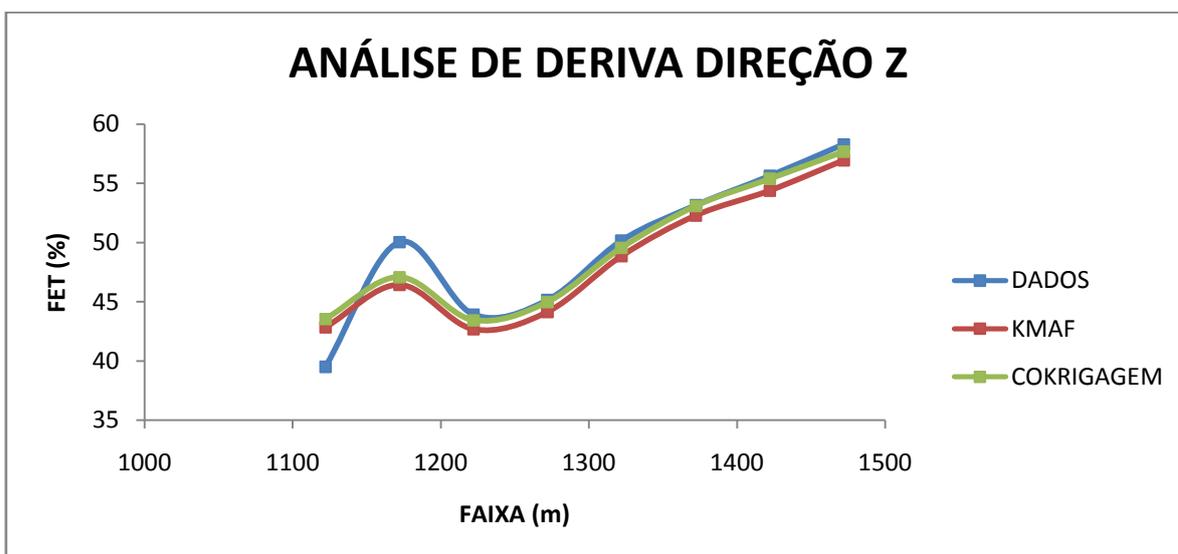


Figura 4. 84 - Análise de deriva na direção Z(vertical) para a variável FET (%). Faixas divididas em intervalos de 50m

Assim como na análise de deriva na direção Norte, a análise na direção vertical revela uma subestimativa sistemática do processo KMAF, porém acompanhando a tendência dos dados originais e próximo à curva da cokrigagem.

O mesmo procedimento foi realizado para variável Sílica. O gráfico da análise de deriva na direção Leste é mostrado na figura 4.85.

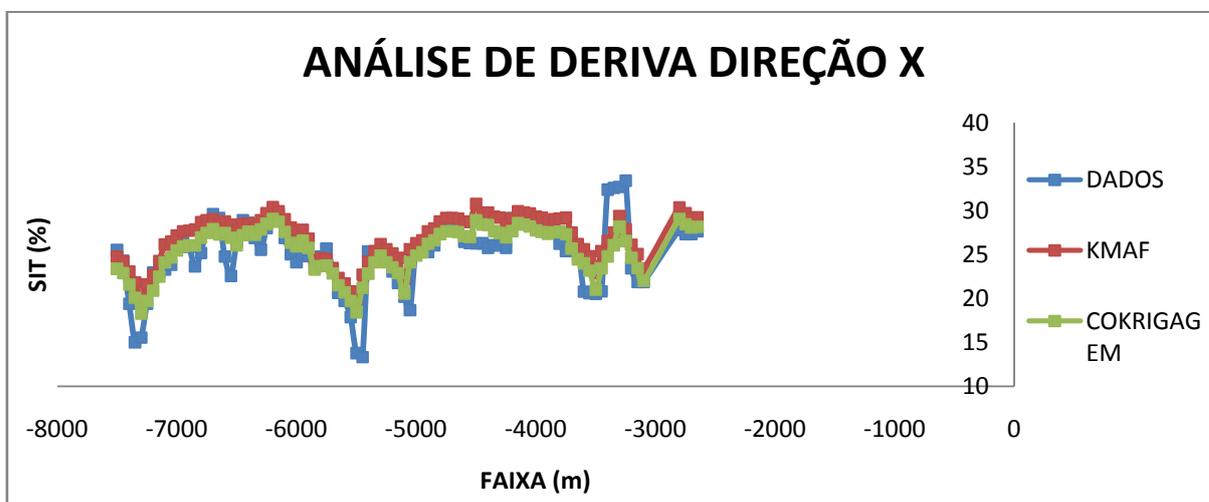


Figura 4. 85 - Análise de deriva na direção X (Leste) para a variável SIT (%). Faixas divididas em intervalos de 50m

Pode-se ver, a partir da figura 4.85, que, no caso da variável Sílica, há uma superestimativa dos valores tanto da cokrigagem quanto na KMAF. Na figura 4.86, é apresentada a análise realizada na direção norte.

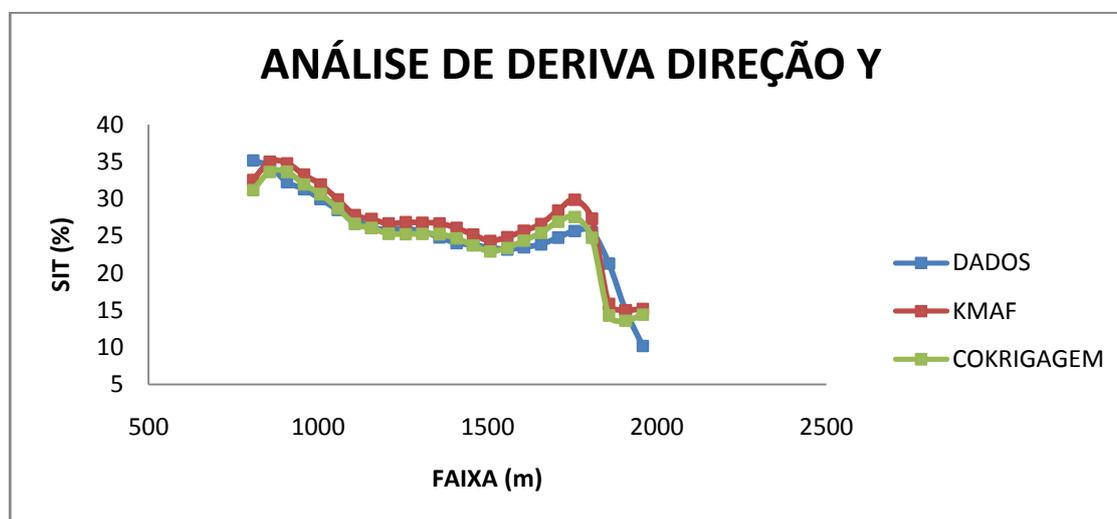


Figura 4. 86 - Análise de deriva na direção Y (Norte) para a variável SIT (%). Faixas divididas em intervalos de 50m

Note, na figura 4.86, que a cokrigagem acompanha a tendência dos dados, na região central do gráfico, a curva de cokrigagem se sobrepõe à dos dados. No caso de KMAF, ao longo do gráfico, a curva passa acima da curva dos dados indicando estimativas mais altas do que os valores apresentados pelos dados. Na figura 4.87, é mostrada a análise de deriva na direção vertical.

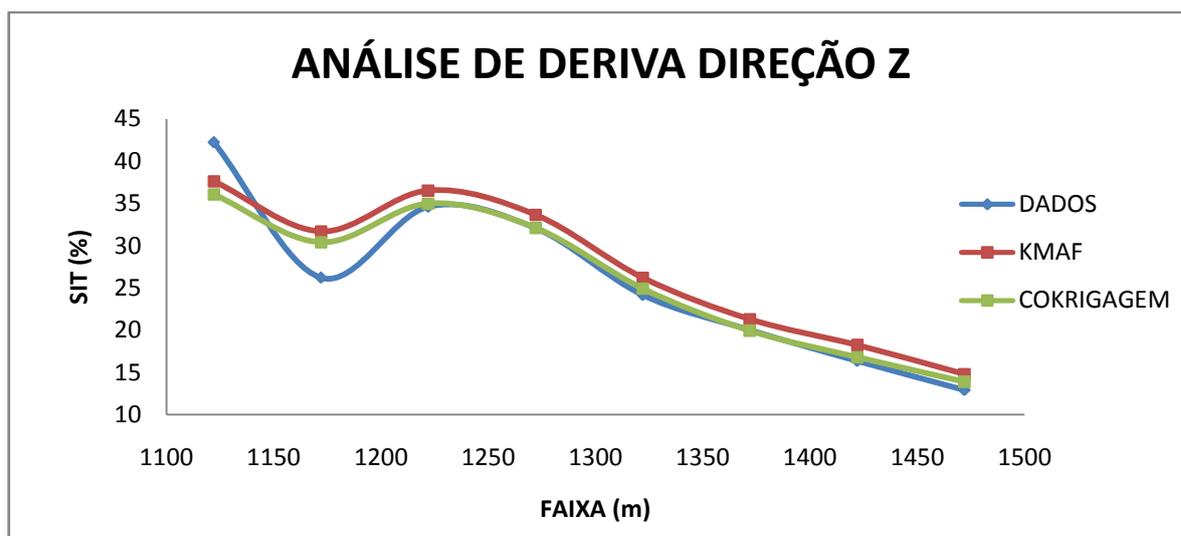


Figura 4. 87 - Análise de deriva na direção Y (Norte) para a variável SIT (%). Faixas divididas em intervalos de 50m

Observa-se, na figura 4.86, que o comportamento apresentado por ambos os métodos é análogo ao que pode ser visto na análise da direção Norte. A curva da cokrigagem se sobrepõe à curva dos dados desagrupados, bem como a KMAF passa acima da curva dos dados, apresentando o mesmo comportamento verificado na figura 4.87.

Assim como analisado no estudo de caso 1, são verificados os gráficos de dispersão entre as variáveis estimadas a fim de analisar a reconstituição das correlações prévias dos dados. Tanto o método clássico, quanto a KMAF satisfazem esta premissa. A figura 4.88 apresenta o gráfico obtido por meio de cokrigagem.

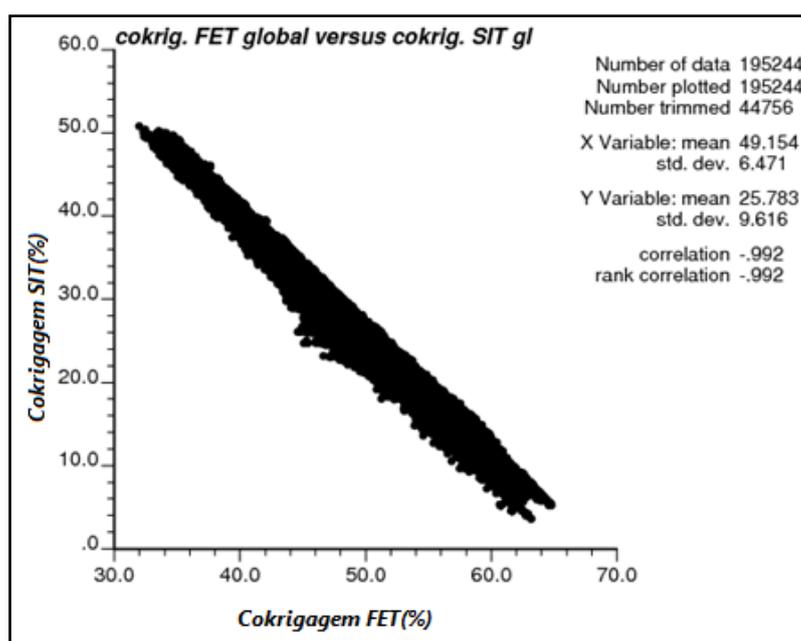


Figura 4. 88 – Gráfico de dispersão entre as variáveis cokrigadas FET (%) e SIT (%)

Na figura 4.88, pode-se verificar a reconstituição da correlação prévia dos dados. Como entre o par havia uma correlação forte, $\rho_{FETSIT} = -0,98$, após o processo de coestimativa, há um aumento de 1% no valor original da correlação. Com uma correlação tão alta entre os dados, o modelo de correção regionalização torna-se adequado qualquer que seja a variável escolhida para criar o modelo base, pois as restantes são regidas pelo mesmo comportamento.

Na figura 4.89, é apresentado o gráfico de dispersão entre as variáveis estimadas por meio do método KMAF.

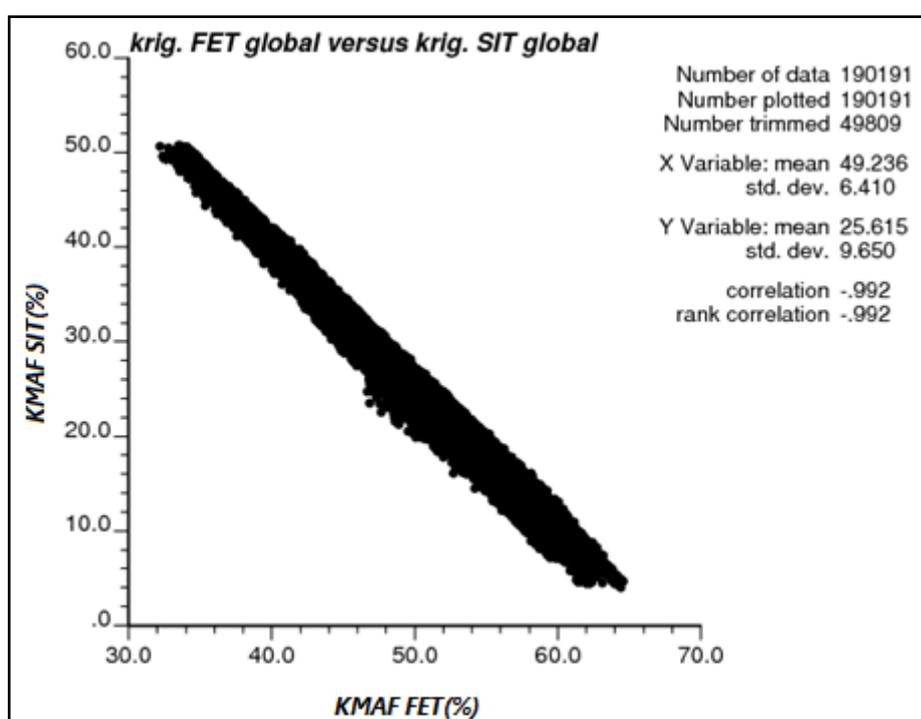


Figura 4. 89 - Gráfico de dispersão entre as variáveis estimadas por meio de KMAF, FET (%) e SIT (%)

Pode-se observar, na figura 4.89, que, como no caso da cokrigagem, a KMAF reproduz a correlação original dos dados. No segundo estudo de caso, porém, a construção dos modelos de regionalização é mais simples, pois os variogramas são ajustados de maneira independente.

Também foi analisada a similaridade entre os procedimentos utilizados. Para tanto, foram gerados gráficos de dispersão entre as saídas da cokrigagem e da KMAF, analisando a correlação entre elas. Na figura 4.90, é mostrado gráfico entre a variável FET(%) cokrigada e a variável FET(%) estimada por meio de KMAF.

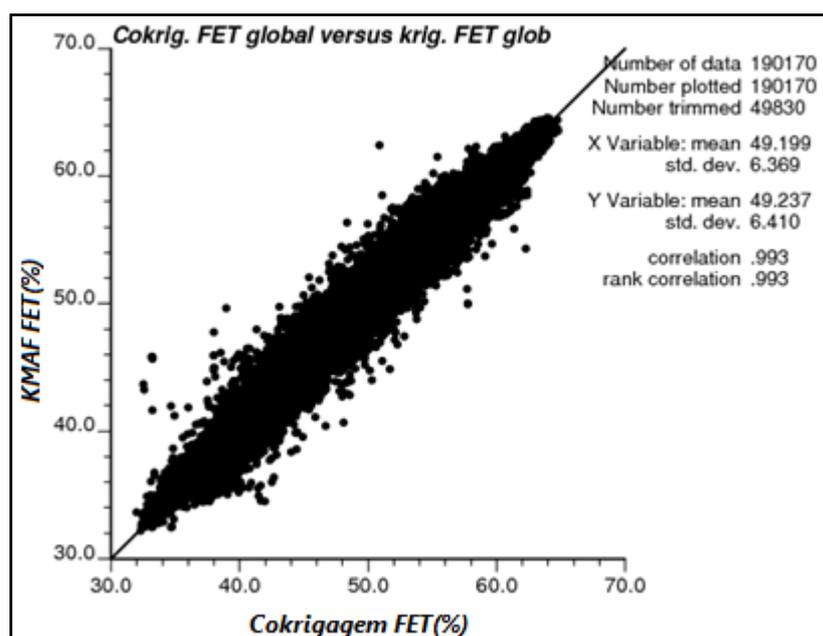


Figura 4. 90 – Gráfico de dispersão entre a variável cokrigada FET (%) e a variável estimada por meio de KMAF FET(%)

Na Figura 4.90, vê-se que a correlação entre os valores obtidos por meio dos dois métodos é de $\rho_{COK_{FET}KMAF_{FET}} = 0,99$, mostrando grande semelhança entre eles. O mesmo é realizado para a variável Sílica, apresentado na figura 4.91.

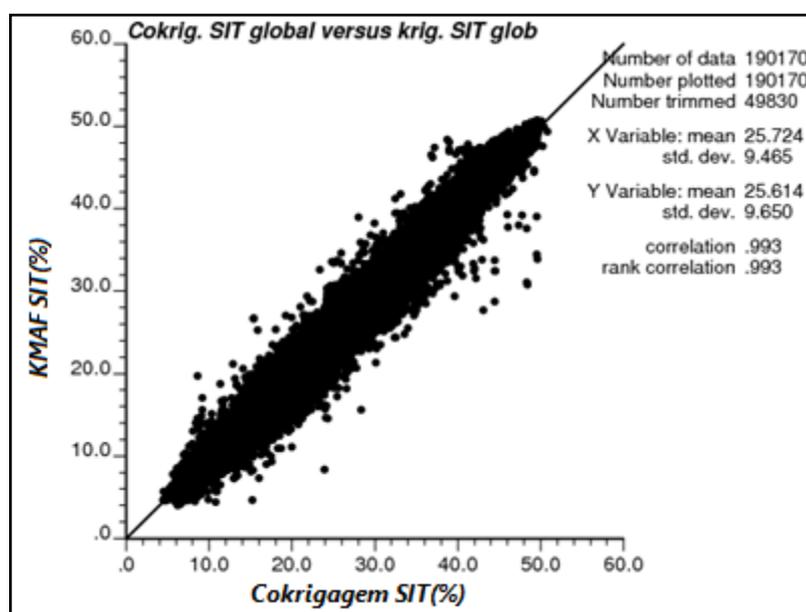


Figura 4. 91 – Gráfico de dispersão entre a variável cokrigada SIT (%) e a variável estimada por meio de KMAF SIT(%)

Na figura 4.91, as estimativas para a Sílica produzidas por meio dos dois métodos apresentaram uma correlação $\rho_{COK_{SIT}KMAF_{SIT}} = -0,99$, ou seja, os métodos apresentam grande similaridade nas estimativas. Também é observada, nas figuras 4.90 e 4.91, a simetria

da nuvem de dispersão em torno da bissetriz, indicando que não há discrepância significativa entre os valores estimados por meio de ambas as metodologias.

Por fim, é realizada a validação visual das estimativas. Neste item, serão realizados mapas de estimativa de teores, nos planos XY, XZ e YZ, pois o modelo de blocos é tridimensional. Foram escolhidas algumas seções transversais em cada plano.

No plano XY, foi selecionada a seção transversal no nível 25, que corresponde à coordenada $Z = 1372m$; no plano YZ, foi selecionado o nível 44 que corresponde à coordenada $X = -5300m$ e, no plano XZ, foi selecionado o nível 33 que corresponde à coordenada $Y = 1268m$.

Para cada nível selecionado, foram buscadas as amostras mais próximas ao bloco estimado.

Primeiramente, serão mostradas as análises no plano XY, XZ e YZ para os resultados obtidos para as duas variáveis FET (%) por meio das duas metodologias empregadas. Após, serão feitas as mesmas análises para a variável SIT (%).

Na figura 4.92, são mostrados os mapas de cokrigagem e KMAF para o atributo Ferro juntamente com o mapa de localização das amostras mais próximas ao nível selecionado.

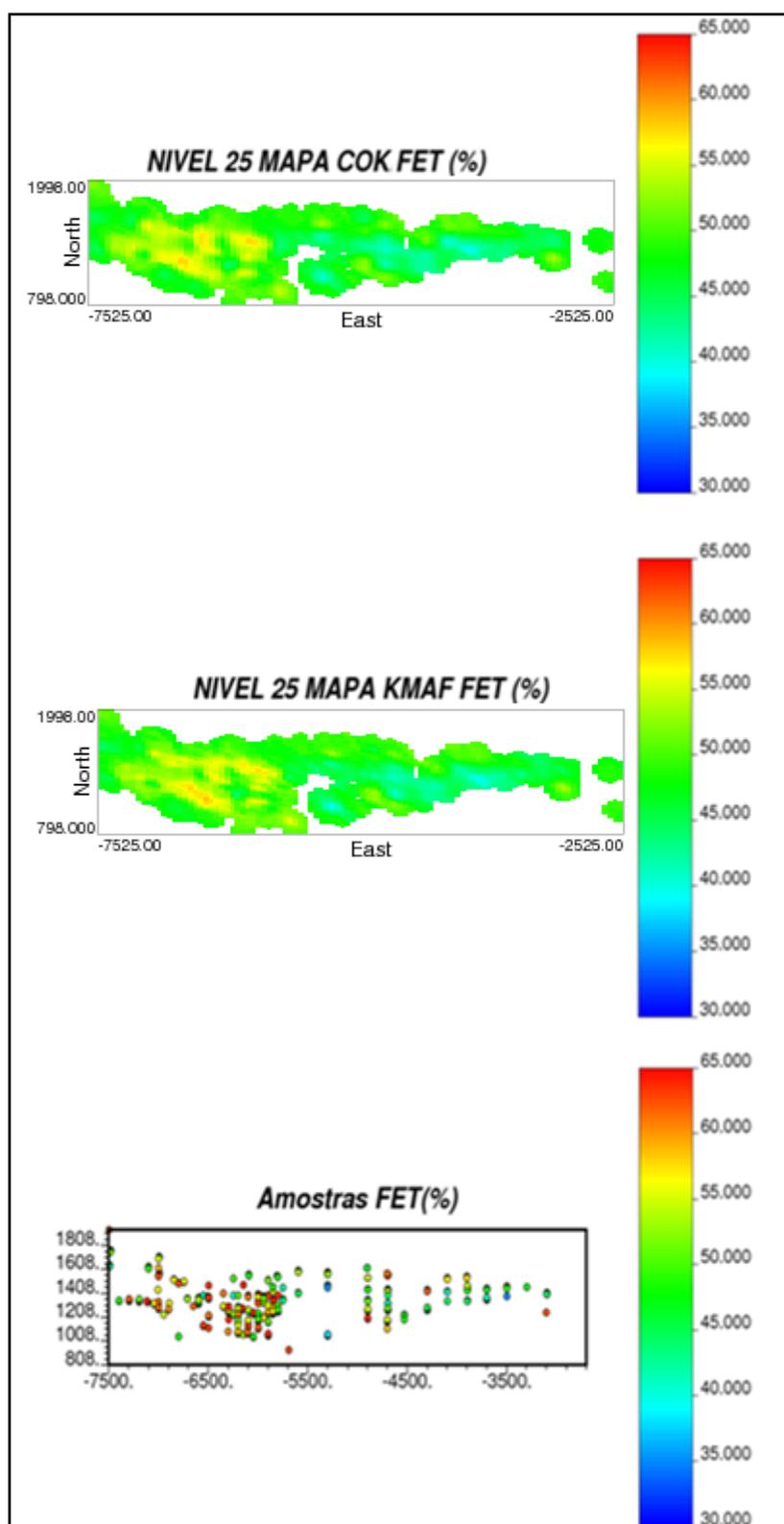


Figura 4. 92 – Vista no plano XY (leste e norte), nível 25 dos mapas de coestimativa do teor de Ferro por meio de cokrigagem (mapa superior); estimativa do teor de Ferro via KMAF (mapa central) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Ferro em (%)

Pode-se ver, no mapa da figura 4.92, que representa a posição amostral entre as coordenadas $X = -5500$ e $X = -7500$, que há uma região com presença de altos teores.

Esta zona é reproduzida, tanto no mapa de coestimativa, mapa superior, quanto no mapa da variável krigada, por meio da transformação MAF, mapa central. Também entre as coordenadas $X = -4500$ e $X = -3500$ e entre $Y = 1408$ e $Y = 1608$ do mapa obtido por meio de KMAF, observa-se também uma região de valores do Ferro entre 50% e 60% mais acentuada que a observada no mapa de cokrigagem. De acordo com a análise de deriva realizada, na direção X, as médias estimadas por meio de KMAF são mais altas que as obtidas por meio de cokrigagem e se aproximam melhor da curva dos dados originais.

Os mapas de estimativa e coestimativa realizados para o nível 33 do plano XZ são apresentados na figura 4.93.

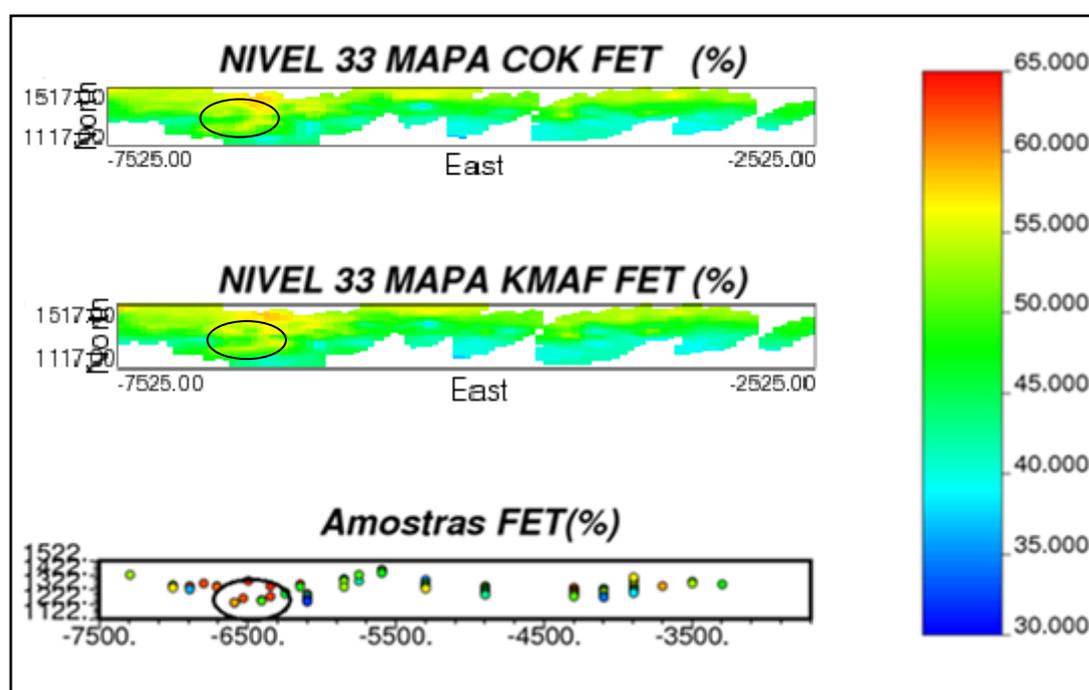


Figura 4.93 - Vista no plano XZ (leste e vertical), nível 33 dos mapas de coestimativa do teor de Ferro por meio de cokrigagem (mapa superior); estimativa do teor de Ferro via KMAF (mapa central) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Ferro em (%)

Na figura 4.93, na região destacada dos mapas, nota-se que a cokrigagem atribui valores mais altos que a KMAF. Na figura 4.94, é apresentado o mapa no plano YZ, nível 44.

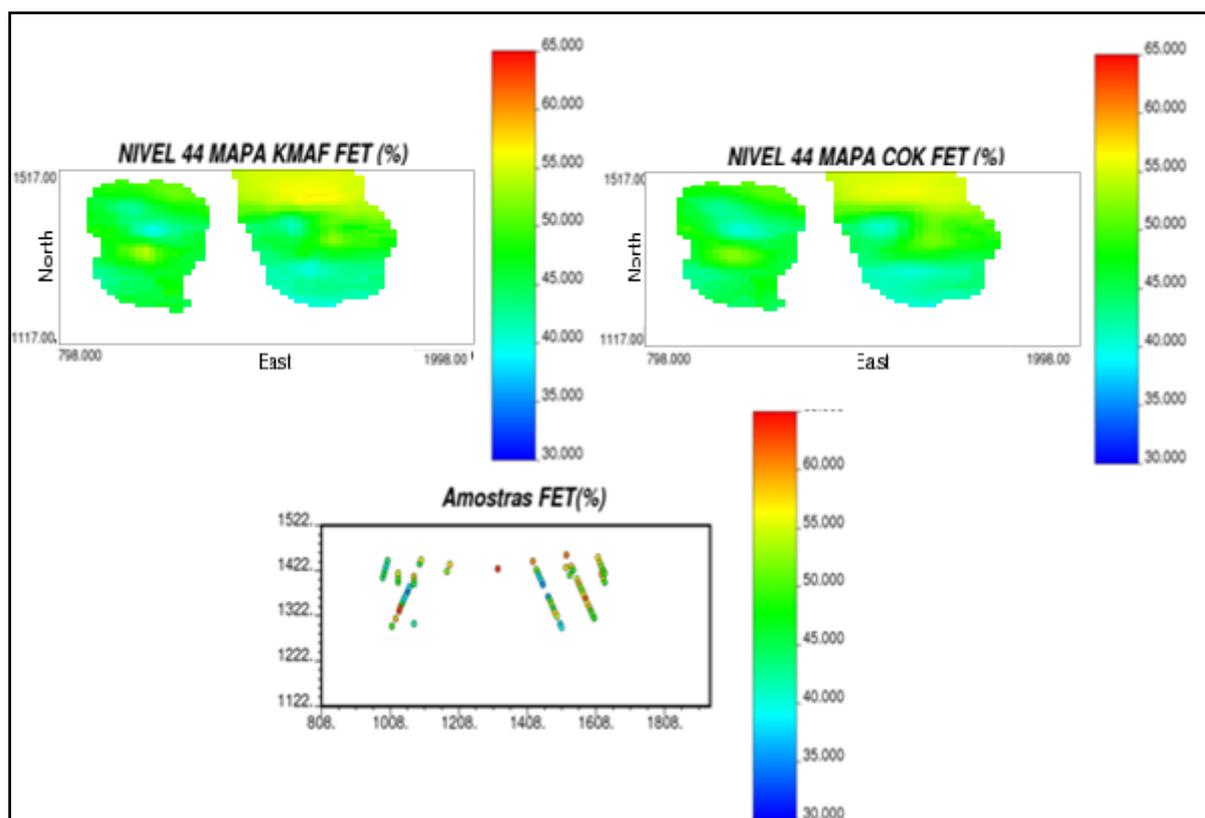


Figura 4. 94 - Vista no plano YZ (norte e vertical), nível 44 dos mapas de estimativa do teor de Ferro KMAF (mapa superior esquerdo); mapas de coestimativa do teor de Ferro por meio de cokrigagem (mapa superior direito) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Ferro em (%)

Observa-se, na figura 4.94, que a estimativa obtida por meio da KMAF apresenta menor suavização do que a cokrigagem. Na figura 4.95, é apresentada a validação para a variável SIT (%) no plano XY.

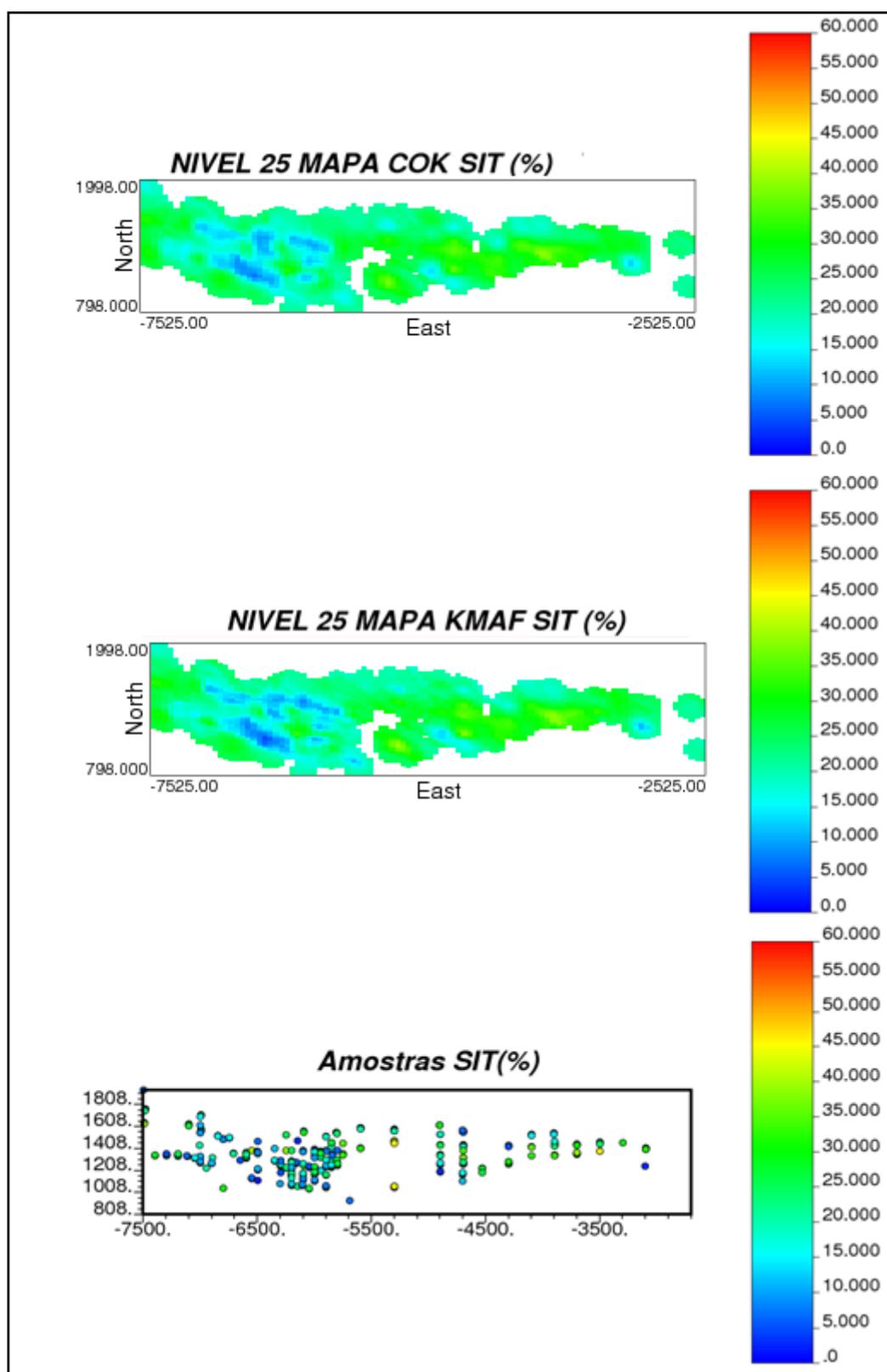


Figura 4. 95 - Vista no plano XY (leste e norte), nível 25 dos mapas de coestimativa do teor de Sílica por meio de cokrigagem (mapa superior); estimativa do teor de Sílica via KMAF (mapa central) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Sílica em (%)

A partir da figura 4.95, vê-se que os baixos teores presentes são reproduzidos pelos dois métodos aplicados, cokrigagem e KMAF. Não se notam discrepâncias nos mapas, do

mesmo modo que se verificou nos mapas da variável Ferro no plano XY. Na figura 4.96, é apresentado o mapa no plano XZ para o nível 33 da variável SIT (%).

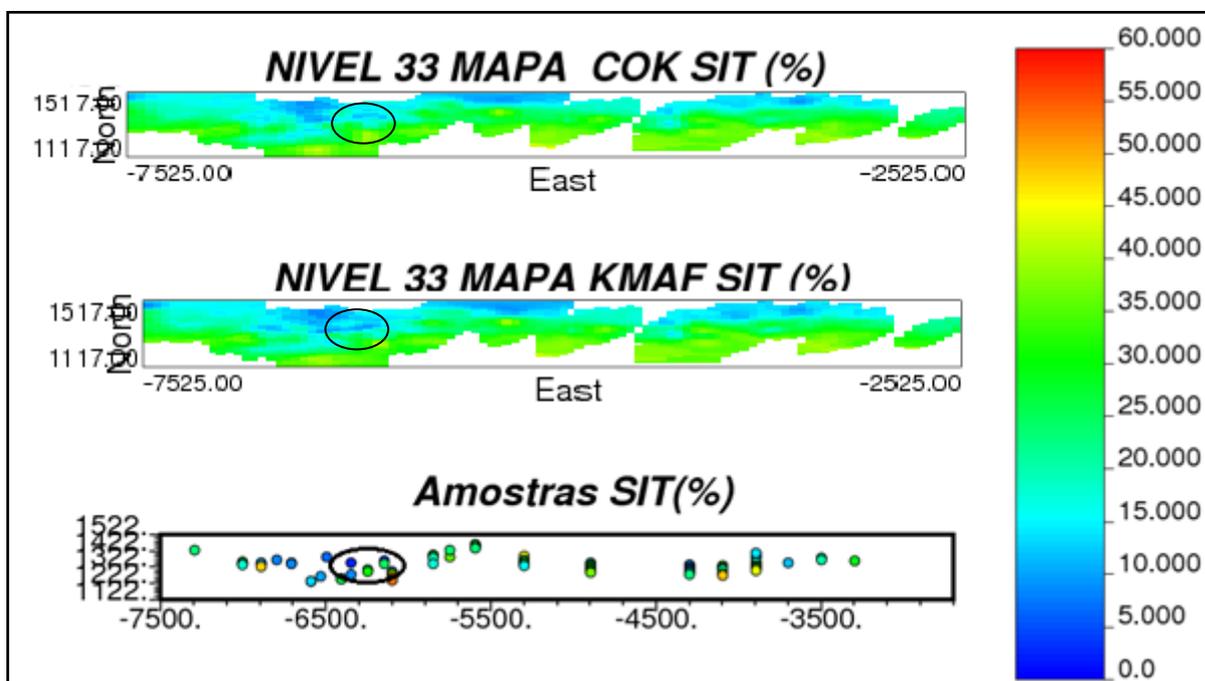


Figura 4.96 - Vista no plano XZ (leste e vertical), nível 33 dos mapas de coestimativa do teor de Sílica por meio de cokrigagem (mapa superior); estimativa do teor de Sílica via KMAF (mapa central) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Sílica em (%)

Observa-se, na figura 4.96, no mapa de estimativas, que, nas regiões destacadas, a cokrigagem suaviza as regiões mais do que é percebido no mapa obtido por KMAF. Por fim, são apresentados os mapas no plano YZ, nível 44 para a variável SIT (%), na figura 4.97.

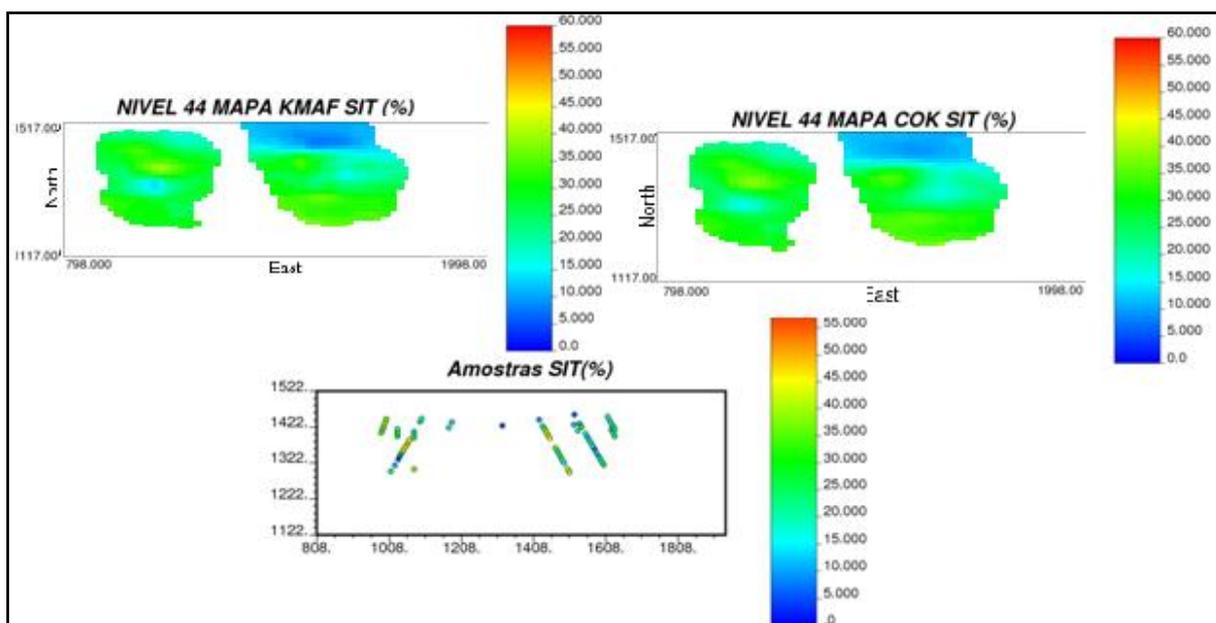


Figura 4. 97 - Vista no plano YZ (norte e vertical), nível 44 dos mapas de estimativa do teor de Sílica via KMAF (mapa superior esquerdo); de coestimativa do teor de Sílica por meio de cokrigagem (mapa superior direito) e mapa de localização das amostras próximas aos blocos estimados (mapa inferior). Escala de cores representa teor de Sílica em (%)

Assim como para a análise no plano YZ realizada para a variável FET (%) (na figura 4.94), se vê, na figura 4.97, uma menor suavização dos valores no mapa gerado pela KMAF.

4.4 ANÁLISES COMPARATIVAS ENTRE AS METODOLOGIAS

A cokrigagem foi utilizada como referência por ser o método classicamente empregado para problemas de estimativa de depósitos com multielementos.

É cediço que a cokrigagem, dentre os métodos de coestimativa, apresenta a vantagem de lidar adequadamente com bancos de dados heterotópicos, provendo estimativas precisas e acuradas. A sua limitação está na dificuldade na modelagem do MLC. No caso de mais de duas variáveis, o usuário se encontra impelido a utilizar *softwares* com algoritmos semiautomáticos para chegar a sua construção. No estudo de caso 1 realizado, observou-se que, ao serem construídos modelos simultâneos a todas as variáveis, para considerar as correlações existentes entre elas, chega-se a variogramas que, individualmente, não seriam os mais adequados. A continuidade espacial assumida a uma variável não implica na mesma continuidade espacial assumida por outra variável. Além disso, também se verificou a geração de uma pseudocorrelação entre variáveis que, originalmente, não apresentavam tal característica. A correlação apresentada após a cokrigagem é dita espúria, ou seja, não possui sentido físico associado a ela.

No caso da krigagem dos fatores independentes, KMAF, a construção do MLC é contornada quando se propõe a descorrelação das variáveis. A continuidade espacial é modelada para cada fator analisado. A descorrelação é obtida para vetores de separação não nulos, o que não se obtém se esta for realizada por meio de PCA, em bancos de dados que não apresentem correlação intrínseca. A limitação do método, que também insere erros no processo de estimativa, é a determinação do vetor h , até onde ocorrerá a descorrelação dos dados; ainda não há uma metodologia que explicita a maneira de obtê-los. Observou-se, também, nos estudos de caso analisados, que os fatores MAF podem ou não apresentar a mesma direção preferencial dos dados. Isto será definido pelo grau de correlação entre as variáveis do banco, uma vez que cada fator MAF é composto por uma porção de cada variável presente (combinação linear). Se as variáveis forem fortemente correlacionadas como no estudo de caso 2, é provável que os fatores apresentem a mesma direção de anisotropia que as variáveis originais.

O método apresenta também a necessidade de dados serem isotópicos tanto no espaço original quanto nos blocos com os fatores estimados para a realização da retrotransformação. A isotopia dos dados pode ser atingida inicialmente ou filtrando os locais amostrais em que não há análise de todos os elementos estudados, ou por meio da krigagem ordinária de cada uma das variáveis a fim de preencher os locais nos quais não há amostras de todas as espécies químicas estudadas. No caso da isotopia nas estimativas, estas podem ser realizadas utilizando a mesma estratégia de busca para todos os fatores, de maneira que se tenham os mesmos blocos estimados com todos os fatores estimados. Ou, ainda, pode ser gerado um filtro após a estimativa que mantenha somente os blocos estimados que tenham em comum a estimativa de todos os fatores.

Os resultados analisados nos estudos de caso mostram que a KMAF reproduziu as correlações originais dos dados de maneira mais adequada do que a cokrigagem, isto pela descorrelação entre eles.

Os gráficos de dispersão entre as duas metodologias, mostrados nas análises de resultados em cada estudo realizado, mostram também que o emprego da transformação linear MAF não implica a perda de precisão das estimativas.

Mesmo o método proposto tendo limitações, em ambos os estudos de caso realizados, a krigagem ordinária de fatores independentes, KMAF, realiza as estimativas dos atributos

com um grande ganho operacional. O procedimento é simples, reduz o tempo empregado na modelagem da continuidade espacial e pode ser realizado em qualquer *software*.

CAPÍTULO 5

Neste capítulo, são apresentadas as conclusões e recomendações obtidas por meio de estudo realizado nesta dissertação.

5.1 CONCLUSÃO GERAL

Em relação ao objetivo deste trabalho, avaliar a utilização do método de Mínimos-Máximos Fatores de Autocorrelação (MAF) combinada com krigagem como alternativa à cokrigagem, aquele foi atingido. Uma vez que, nos estudos de caso realizados, a metodologia MAF proveu resultados satisfatórios, sem perda de precisão em relação à metodologia clássica.

O problema que motivou o estudo aqui realizado é a limitação do uso da cokrigagem em estimativas de depósitos minerais multivariados, devido à necessidade do uso do MLC. A metodologia proposta se mostrou capaz, nos estudos de caso analisados, de contornar o problema exposto por meio da decorrelação das variáveis. Além disso, foi observado, ao longo do desenvolvimento deste trabalho, que a cokrigagem não só tem a complexidade de sua aplicação como limitadora de seu uso, mas também a inadequação muitas vezes do modelo de correionalização adotado para modelar o comportamento espacial de todas as variáveis presentes no sistema. Como visto nos estudos de caso, isso gera correlações entre atributos que, na realidade, são inexistentes. O método MAF, por sua vez, obteve maior aderência na reprodução das correlações originais, além de contornar a problemática do MLC.

5.1.1 CONCLUSÕES REFERENTES ÀS METAS PROPOSTAS

- I. A meta proposta de desenvolver uma rotina que execute a transformação MAF de maneira automática foi atingida. A rotina apresentada no capítulo 3, e utilizada nos estudos de caso, desempenhou a transformação MAF corretamente, satisfazendo os testes realizados;
- II. A meta de aplicar a metodologia a dois bancos de dados multivariados distintos foi executada, provendo resultados positivos na avaliação do uso da metodologia proposta. Esta se mostrou eficiente e com facilidade operacional ao usuário, uma vez que se faz desnecessária a modelagem do MLC no sistema. Os gráficos de dispersão entre as estimativas e coestimativas obtidas

por meio de cada método, apresentados na discussão dos resultados, mostram que: há forte correlação entre os resultados obtidos para todas as variáveis estimadas. Desta forma, se vê que a metodologia foi aplicada aos casos analisados sem perda de precisão nas estimativas de teores, em relação à cokrigagem.

5.1.2 RECOMENDAÇÕES

Para trabalhos futuros, recomenda-se que sejam realizados mais estudos de caso utilizando o método MAF combinado com krigagem a fim de avaliar o seu uso em outros tipos de depósitos. Além disso, que seja realizado um estudo amplo sobre as implicações que têm os métodos de isotopização dos dados sobre as estimativas em diferentes cenários de configuração das amostras, como a perda, se houver, na precisão da estimativa por meio de KMAF para cada método de isotopização empregado.

Recomenda-se, também, que seja realizado um estudo sobre a escolha do vetor de separação ao qual ocorrerá a descorrelação dos fatores MAF. Não existe metodologia que demonstre matematicamente como definir este valor.

REFERÊNCIAS BIBLIOGRÁFICAS

Almeida, A. S., 1993, Joint simulation of multiple variables with a Markov-type coregionalization model, tese de doutorado, Universidade de Stanford, 199p.

Anton, H. & Rorres, C., 2000, Álgebra linear com aplicações. Tradução de Claus Ivo Doering, editora Bookman, 572p.

Bandarian, E. M., Bloom, L. M. & Mueller, U. A., 2008, Direct minimum-maximum autocorrelation factors within the framework of two structure linear model of corregeionalisation, *Computers&Geosciences*, 34(3), pp. 190-200.

Boezio, M. N., 2010, Estudo das metodologias alternativas da geoestatística multivariada aplicada a estimativas de teores de depósitos de Ferro, tese de doutorado, Universidade Federal do Rio Grande do Sul, 465p.

Boucher, A., 2003, Conditional joint simulation of random fields on block support, Dissertação de Mestrado, Universidade de Queensland, Brisbane, 1610p.

Davis, B. & Greenes, K. A., 1983, Estimation using spatially distributed multivariate data: an exemple with coal quality, *Mathematical Geology*, 15(2), pp. 287-300.

Desbarats, A., 2001, Geostatistical modeling of regionalized grain size distributions using min/max autocorrelation factors, *GeoEnvIII – geostatistics for environmental applications*, Proceedings of the third European conference on geostatistics for environmental applications, pp. 441-452.

Desbarats, A. & Dimitrakopoulos, R., 2000, Geostatistical simulation of regionalized pore-size distributions using min/max autocorrelation factors, *Mathematical Geology*, 32(8), pp. 919-942.

Dirichlet, G. L., 1850, Über die Reduktion der positive quadratischen formem mi drei unbestimmten ganzen Zahlen, *Journal für die reine und angewandte Mathematik*, pp. 209-227.

Dorr, J. V. N., 1969, Physiographic, stratigraphic and structural development of the Quadrilátero Ferrífero, *USGS Professional Paper*, N° 641-A, 110p.

Goovaerts, P., 1993, Spatial Orthogonality of the principal components computed from regionalized variables, *Mathematical Geology*, 25(3), pp. 281-302.

Goovaerts, P., 1997, *Geostatistics for natural resource evaluation*, Oxford University Press, 483p.

Hoettling, H., 1933, Analysis of a complex statistical variables of principal components, *Journal of educational psychology*, 24(7), pp. 417-441.

Jolliffe, I. T., 2002, *Principal component analysis*, Springer, 518p.

Jounel, A., 1999, Markov model for cross-covariances, *Mathematical Geology*, 31(8), pp. 955-964.

Marechal, A., 1970, Cokriegeage et regression em correlation intrisique, *Centre de Geostatistique de Fontainebleau*, 40p.

Matheron, G., 1963, Principles of geostatistics, *Economic Geology*, 58(8), pp. 1246-1266.

Mathews, K. R., 2008, *Elementary Linear Algebra*, Springer, 301p.

Pearson, K., 1901, On lines and planes closest fit to systems of points in space, *Philosophical Magazine*, 2(6), pp. 559-572.

Rodrigues, C. F., 2011, Análise comparativa entre os métodos de decomposição em valores singulares e análise de componentes principais envolvendo matrizes de grande porte, dissertação de mestrado, Universidade Federal de Minas Gerais, 187p.

Rondon, O., 2011, Teaching aid: minimum/maximum autocorrelation factors for joint simulations of attributes, *Mathematical Geosciences*, 44 (4), pp. 469-504.

Spier, C. A., Moura, M. M., Ribeiro, D. & Reis da Cruz, F. A., 2004, Appraisal of geological resources of the Pico complex – Quadrilátero Ferrífero – Minas Gerais, Relatório Interno MBR (Minerações Brasileiras Reunidas), 168p.

Switzer, P. & Green, A. A., 1984, Min/max autocorrelation factors for multivariate spatial imagery, Reporte Técnico Número 6, Universidade de Stanford, 14p.

Wackernagel, H., 1998, Principal component analysis for autocorrelated data: a geostatistical perspective, Reporte Técnico Número 22, Escola de Mines de Paris, 41p.

Xu, W., Tran, T. T., Srivastava, R. M. & Journel, A. G., 1992, Integrating seismic data in reservoir modeling: the collocated cokriging alternative, Anais do 67th Annual Technical Conference of the Society of Petroleum Engineers, 24742, pp. 833-842.

GLOSSÁRIO

Expressão em inglês	Expressão em português	Definição
<i>Coefficient of variation (coef. of var.)</i>	Coeficiente de variação	Medida de assimetria da distribuição
<i>Correlation</i>	Correlação	Sumariza a relação existente entre duas variáveis
<i>Dip</i>	Mergulho	Deslocamento em ângulos na vertical em relação ao plano horizontal
<i>Lag</i>	Espaçamento	Intervalo métrico pré-definido utilizado no cálculo do variograma experimental
<i>Lower quartile</i>	Quartil inferior	Indica o valor no qual o histograma cumulativo é igual a 75%.
<i>Maximum</i>	Máximo	Indica o valor no qual o histograma cumulativo é 0%.
<i>Mean</i>	Média	Centro de gravidade da distribuição da variável
<i>Median</i>	Mediana	Indica o valor no qual a distribuição é dividida na metade.
<i>Minimum</i>	Mínimo	Indica o valor no qual o histograma cumulativo é 100%
<i>Number of data</i>	Número de dados	Consiste na quantidade de amostras presentes nos histogramas

<i>Number plotted</i>	Número plotado	Quantidade de pares dispostos em um gráfico de dispersão
<i>Rank correlation</i>	Correlação de ranking	Sumariza a relação existente entre diferentes rankings de duas variáveis
<i>Sill</i>	Patamar	Valor do variograma correspondente ao seu alcance. A partir deste ponto não se observa dependência espacial entre as amostras
<i>Standard deviation (std. dev.)</i>	Desvio padrão	Raiz quadrada da variância
<i>Upper quartile</i>	Quartil superior	Valor no qual se tem o histograma cumulativo igual a 25%
<i>Variable</i>	Variável	Variável considerada na análise estatística