

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

*LUCIANA ARNT ABICHEQUER*

***ANÁLISE E PROPOSIÇÃO DE METODOLOGIAS PARA SIMULAÇÃO DE  
COMPOSIÇÕES EM DEPÓSITOS MULTIVARIADOS COMPLEXOS***

Porto Alegre  
2016

LUCIANA ARNT ABICHEQUER

**ANÁLISE E PROPOSIÇÃO DE METODOLOGIAS PARA SIMULAÇÃO DE  
COMPOSIÇÕES EM DEPÓSITOS MULTIVARIADOS COMPLEXOS**

*Tese submetida ao Programa de Pós Graduação em Engenharia de Minas, Metalúrgica e de Materiais, da Universidade Federal do Rio Grande do Sul, como requisito parcial à obtenção do título de Doutora em Engenharia.*

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

*Porto Alegre  
2016*

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Carlos Alexandre Neto

Vice-reitor: Rui Vicente Opperman

ESCOLA DE ENGENHARIA

Diretora: Denise Carpena Dal Molin

Vice-Diretor: Carlos Eduardo Pereira

PROGRAMA DE PÓS GRADUAÇÃO

Coordenador: Carlos Pérez Bergmann

Coordenador Substituto: Afonso Reguly

Arnt Abichequer, Luciana

ANÁLISE E PROPOSIÇÃO DE METODOLOGIAS PARA  
SIMULAÇÃO DE COMPOSIÇÕES EM DEPÓSITOS MULTIVARIADOS  
COMPLEXOS / Luciana Arnt Abichequer. -- 2016.  
213 f.

Orientador: João Felipe Coimbra Leite Costa.

Tese (Doutorado) -- Universidade Federal do Rio  
Grande do Sul, Escola de Engenharia, Programa de Pós-  
Graduação em Engenharia de Minas, Metalúrgica e de  
Materiais, Porto Alegre, BR-RS, 2016.

1. Depósitos Multivariados Complexos. 2.  
Decomposição MAF. 3. Análise de Dados Composicionais.  
4. Simulação Direta em Blocos. I. Coimbra Leite  
Costa, João Felipe, orient. II. Título.

*Luciana Arnt Abichequer*

**ANÁLISE E PROPOSIÇÃO DE METODOLOGIAS PARA SIMULAÇÃO DE  
COMPOSIÇÕES EM DEPÓSITOS MULTIVARIADOS COMPLEXOS**

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

---

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

---

Prof. Dr. Carlos Pérez Bergmann  
Coordenador do PPGEM

Aprovado em \_\_/\_\_/2016

BANCA EXAMINADORA:

Prof. Dra. Vanessa Cerqueira Koppe – DEMIN/UFRGS

Dr. Diniz Tamantini Ribeiro – Vale S.A.

Dra. Maria Noel Morales Boezio – DINAMIGE, Uruguai

Aos meus pais

## AGRADECIMENTOS

À Unipampa, pelo afastamento.

Aos colegas Raul e Eduardo, por me substituírem e em especial ao Eduardo, pelas contribuições.

Ao professor João Felipe pela orientação.

Aos colegas do LPM pelo companheirismo e troca de ideias.

Ao grupo LPM pela estrutura necessária para a realização do trabalho.

Aos meus pais, pelo amor e tudo o mais que for incondicional.

Às minhas amigas tão queridas.

À MRN pelo banco de dados e dados de campo.

“Open mind for a different view.”  
Nothing Else Matters – The Black Album  
Mettalica

## RESUMO

A cossimulação de teores em depósitos multivariados complexos com mais de três variáveis envolvidas, quando realizada por métodos clássicos, é extremamente trabalhosa e acaba por gerar resultados que precisam de ajuste posterior, principalmente com relação ao fechamento da soma dos teores nos blocos ou faixas granulométricas de interesse. A necessidade de ajustes posteriores, aliada à falta de praticidade destes métodos, motiva a busca por soluções alternativas que gerem resultados tão ou mais precisos e, sejam mais facilmente implementáveis no dia a dia da indústria mineira. Nesta tese foi analisada a viabilidade de aplicação de cinco combinações de métodos como alternativa aos métodos clássicos: a simulação individual pelo método de bandas rotativas de cada uma das variáveis, em suporte de pontos posteriormente reblocados (combinação 1); a simulação das transformações ilrs, também pelo métodos de bandas rotativas, de forma individual e em suporte de pontos posteriormente reblocados (combinação 2); a simulação diretamente em suporte de blocos dos fatores MAFs (combinação 3); a simulação diretamente em suporte de blocos dos fatores MAFs das ilrs, com a utilização da variável artificial Resto, também utilizada nas combinações 1,2 e 3 (combinação 4) e; a simulação diretamente em suporte de blocos dos fatores MAFs das ilrs, com a operação de fechamento (combinação 5). Para aplicação da metodologia proposta foram simulados os teores das variáveis Alap, Fe, Si e Ti, retidas na faixa 14#, para um depósito de bauxita do norte do Brasil. Os resultados obtidos em todas as combinações foram analisados com relação ao grau de reprodução das características estatísticas e de continuidade espacial, reprodução das somas dos teores dos dados originais também na faixa simulada e a reprodução das correlações, ou reprodução da ausência de correlações, também nos cenários simulados. Após a análise dos resultados obtidos, se pode concluir que todas as combinações testadas são passíveis de utilização, no entanto, somente as combinações quatro e cinco proveem uma solução acabada para o problema em questão. Isto se deve ao fato de estas duas combinações de métodos serem as únicas que combinaram: a decomposição MAF, para descorrelacionar as variáveis, simular cada uma de modo independente das demais e mesmo assim

garantir a reprodução das correlações no final do processo; a transformação ilr, que, por considerar as amostras como composições pertencentes ao espaço de soma restrita e constante, Simplex, garante que a soma dos teores esteja garantida nos resultados das simulações; e, a simulação diretamente em suporte de blocos, que foi incluída por evitar que um grande número de pontos precise ser simulado e posteriormente ajustado para o suporte de blocos, dando agilidade ao processo. A combinação cinco se mostrou ainda mais vantajosa por não fazer uso da variável artificial Resto, o que diminui o número de variáveis a serem efetivamente simuladas. Apesar das vantagens apontadas, é inegável que o usuário precisa conhecer e estar atento à sequência em que as transformações devem ser aplicadas sobre os dados originais.

Palavras chave: Métodos de Cossimulação, Transformações logarítmicas isométricas, Fatores Mínimos e Máximos de Correlação, Depósitos Multivariados Complexos.

## ABSTRACT

Cosimulation of complex multivariate deposits with more than three variables involved, when performed by classical methods, is extremely labor intensive and ultimately generate results that need further adjustment, especially with respect to the closing of the sum of the attributes of interest. The need for further adjustments, together with the difficulties in practical implementation of these methods, motivates the search for alternative solutions that generate results as or more accurate and which can be more easily implementable at the mining industry. This thesis analyzed the feasibility of five combinations of methods as an alternative to the classical ones: individual simulation by turning bands method, each variable at a time, in subsequent block support adjustment (combination 1); the simulation of ILRs transformations, also using the turning bands method, individually and followed by the block support correction (combination 2); the directly block simulation of MAFs factors (combination 3); direct block simulation of MAFs factors of ILRs with the use of artificial variable "Resto" also used in combinations 1,2 and 3 (combination 4) and direct block simulation of MAFs factors of ILRs with the closing operation (combination 5). To apply the proposed methodology, the following variables from a large bauxite deposit were chosen: mass retained at 14 # sieve, Alap, Fe, Si and Ti were simulated. The results obtained in all combinations were analyzed with respect to the level of reproduction of statistical characteristics and spatial continuity, reproduction of the sums of the grades of the original data also in the simulated scenarios, allied to variables cross-correlation reproduction in the simulated scenarios. The results lead to conclude that all tested combinations are amenable to use, however, only the 4 and 5 combinations provide a finished solution to the problem in question. The MAF decomposition was used to decorrelate the variables and so each one of those could be independently simulated, and the maintenance of correlations after the process was assured, and the ilr transformation was included in both combinations 4 and 5 to provide results with constant sums of the grades in the sieve of interest. These two combinations guarantee the maintenance of the grades closed sum and correlations reproduction in simulated scenarios, and, still are computationally efficient as the

block simulation was used to avoid the simulation of a large amount of points that need to be post processed. The last workflow proved to be even more advantageous for not using an artificial variable Resto, which reduces the number of variables to be effectively simulated. Despite of these advantages, it is undeniable that the user needs to know and to be aware of the sequence in which transformation should be applied to the original data.

Keywords: Cossimulation methods, Complex Deposits, Isometric Log Ratios, Minimum/Maximum Correlation Factors, Direct Block Simulation.

## LISTA DE FIGURAS

Figura 1.1: MÉTODO TRADICIONAL DE SIMULAÇÃO EM BLOCOS (modificado de Boucher (2003)).	42
Figura 1.2: fluxograma da metodologia da combinação i.	55
Figura 1.3: fluxograma da metodologia para a combinação ii.	55
Figura 1.4: fluxograma da metodologia para a combinação iii.	56
Figura 1.5: fluxograma da metodologia para as combinação IV.	56
Figura 1.6: fluxograma da metodologia para as combinação V.	57
Figura 2.1: Representação gráfica da perturbação de um segmento no $S^3$ (esquerda) e em coordenadas (direita) (Pawlowsky-Glahn <i>et al.</i> , 2013).	72
Figura 2.2: Representação gráfica da potência de um vetor no $S^3$ (esquerda) e em coordenadas (direita) (Pawlowsky-Glahn <i>et al.</i> , 2013).	72
Figura 2.3: Representação gráfica de círculos e elipses no $S^3$ (esquerda) e em coordenadas (direita) (Pawlowsky-Glahn <i>et al.</i> , 2013).	73
Figura 2.4: Representação gráfica de pares de linhas paralelas no $S^3$ (esquerda) e em coordenadas (direita) (Pawlowsky-Glahn <i>et al.</i> , 2013).	73
Figura 3.1: O perfil geológico da bauxita no depósito (modificado de Carvalho <i>et al.</i> , 1997).	88
Figura 3.2: Amostras de diferentes horizontes do perfil de bauxita plotadas em um diagrama ternário de $SiO_2$ - $Fe_2O_3$ - $Al_2O_3$ (modificado de Carvalho <i>et al.</i> 1997).	91
Figura 3.3: Mapa de localização das amostras para a variável Alap.	93
Figura 3.4: Exemplo ilustrativo de duas amostras analisadas em duas faixas granulométricas. A recuperação mássica das amostras é diferente entre si na mesma faixa e entre as faixas (Modificado de Chaves, A.P., Perez, A.E.C., 1999).	96
Figura 3.5: esquema simplificado com a transformação de coordenadas, onde a seria a camada ferruginosa granular, b seria o horizonte de bauxita e c a camada caulinita inferior. Na parte esquerda da figura os horizontes estão representados em coordenadas cartesianas e na parte direita somente a camada de bauxita está representada em coordenadas estratigráficas. (Modificado de Ecole des Mines de Paris & Geovariances, 2012).	97
Figura 3.6: histograma da variável rec14#, percebe-se o espalhamento dos valores de recuperação nas diferentes amostras.	98
Figura 3.7: correlações lineares entre as variáveis, acompanhadas do respectivo coeficiente de correlação. Exibem correlação significativa: AlapxFe, AlapxSi, RestoxFe e RestoxSi. Não exibem correlação nenhuma, ou exibem correlação insignificante: AlapxResto, AlapxTi, FexSi, FexTi, RestoxTi e SixTi.	101
Figura 3.8: Correlograma ao longo do furo para a variável Alap. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado.	102
Figura 3.9: Correlograma ao longo do furo para a variável Fe. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado.	103
Figura 3.10: Correlograma ao longo do furo para a variável artificial Resto. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado.	103

Figura 3.11: Correlograma ao longo do furo para a variável Si. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado. ....	104
Figura 3.12: Correlograma ao longo do furo para a variável Ti. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado. ....	104
Figura 3.13: Correlograma ao longo do furo para a variável Rec14#. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado. ....	105
Figura 3.14: Correlogramas direcionais para a variável Rec 14#. ....	106
Figura 3.15: Correlogramas direcionais para a variável Alap. ....	106
Figura 3.16: Correlogramas direcionais para a variável Fe. ....	107
Figura 3.17: Correlogramas direcionais para a variável Resto. ....	107
Figura 3.18: Correlogramas direcionais para a variável Si. ....	108
Figura 3.19: Correlogramas direcionais para a variável Ti. ....	108
Figura 3.20: Correlograma omnidirecional para a variável Rec14#. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura. ....	109
Figura 3.21: Correlograma omnidirecional para a variável Alap. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura. ....	110
Figura 3.22: Correlograma omnidirecional para a variável Fe. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura. ....	110
Figura 3.23: Correlograma omnidirecional para a variável Resto. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura. ....	111
Figura 3.24: Correlograma omnidirecional para a variável Si. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura. ....	111
Figura 3.25: Correlograma omnidirecional para a variável Ti. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura. ....	112
Figura 4.1: Correlograma omnidirecional dos dados normalizados da variável Rec14#. ....	115
Figura 4.2: correlogramas experimentais dos modelos simulados (linhas pretas); correlograma experimental (vermelho, linha fina) e correlograma modelado para os dados originais (vermelho, linha grossa). Os correlogramas experimentais dos valores simulados acompanham o correlograma experimental e modelado dos dados originais. ....	117
Figura 4.3: correlogramas experimentais dos modelos simulados (linhas pretas); correlograma experimental (vermelho, linha fina) e correlograma modelado para os dados originais (vermelho, linha grossa). Os correlogramas experimentais dos valores simulados acompanham o correlograma experimental e modelado dos dados originais. ....	118
Figura 4.4: Análise de deriva ao longo do eixo "x". As médias simuladas da variável Rec14# (linhas pretas) seguem a mesma tendência das médias originais da variável (linha contínua vermelha). As linhas vermelhas pontilhadas representam a envoltória do desvio padrão das amostras. ....	120
Figura 4.5: Análise de deriva ao longo do eixo "y". As médias simuladas da variável Rec14# (linhas pretas) seguem a mesma tendência das médias originais da variável	

(linha contínua vermelha). As linhas vermelhas pontilhadas representam a envoltória do desvio padrão das amostras. ....	120
Figura 4.6: Análise de deriva ao longo do eixo "z". As médias simuladas da variável Rec14# (linhas pretas) seguem a mesma tendência das médias originais da variável (linha contínua vermelha) com uma leve discrepância para os valores maiores do que 3m de cota em Z por falta de dados nessa cota. As linhas vermelhas pontilhadas representam a envoltória do desvio padrão das amostras. ....	121
Figura 4.7: Correlogramas simulados (linhas pretas), experimental (linha fina vermelha) e modelado (linha grossa vermelha) para a variável Rec14#.....	123
Figura 4.8: Correlogramas simulados (linhas pretas), experimental (linha fina vermelha) e modelado (linha grossa vermelha) para a variável Rec14#.....	123
Figura 4.9: Análise de deriva em "x". As linhas pretas representam os blocos simulados e a linha vermelha os dados originais. Os blocos simulados (linhas pretas) apresentam médias locais bastante semelhantes às dos dados originais (linha vermelha), acompanhando inclusive as oscilações mais bruscas das mesmas. As linhas vermelhas pontilhadas representam a envoltória do desvio padrão das amostras.....	124
Figura 4.10: Análise de deriva em "y". As linhas pretas representam os blocos simulados e a linha vermelha os dados originais. Os blocos simulados (linhas pretas) apresentam médias locais bastante semelhantes às dos dados originais (linha vermelha), acompanhando inclusive as oscilações mais bruscas das mesmas. As linhas vermelhas pontilhadas representam a envoltória do desvio padrão das amostras.....	125
Figura 4.11: Análise de deriva em "z". As linhas pretas representam os blocos simulados e a linha vermelha os dados originais. Os blocos simulados (linhas pretas) apresentam médias locais bastante semelhantes às dos dados originais (linha vermelha), acompanhando inclusive as oscilações mais bruscas das mesmas. ....	125
Figura 4.12: Correlações lineares entre as variáveis originais (linhas vermelhas) e cenários simulados (linhas pretas). As correlações AlapxResto, SixResto e TixResto foram perfeitamente reproduzidas nos cenários simulados. As correlações AlapxSi, AlapxTi, FexResto, FexSi e SixTi, apesar de não apresentarem uma reprodução perfeita apresentam um comportamento próximo ao das correlações entre os dados originais. Já AlapxFE e FexTi possuem correlações bastante distintas sob o olhar dos dados originais e dos cenários obtidos.....	129
Figura 4.13: Correlações lineares entre as variáveis originais (linhas vermelhas) e cenários simulados (linhas pretas). As correlações AlapxResto, SixResto e TixResto foram perfeitamente reproduzidas nos cenários simulados. As correlações AlapxSi, AlapxTi, FexResto, FexSi e SixTi, apesar de não apresentarem uma reprodução perfeita apresentam um comportamento próximo ao das correlações entre os dados originais. Já AlapxFE e FexTi possuem correlações bastante distintas sob o olhar dos dados originais e dos cenários obtidos.....	135
Figura 4.14: Correlações lineares entre as variáveis originais (linhas vermelhas) e cenários simulados (linhas pretas). Note que, a grande maioria das correlações foi quase que perfeitamente reproduzida. Somente quando olhamos RestoxTi percebemos resultados simulados um pouco destoantes do original.....	140
Figura 4.15: Correlações lineares entre as variáveis originais (linhas vermelhas) e cenários simulados (linhas pretas). AlapxFE, AlapxResto, AlapxSi, AlapxTi, FexSi, FexTi, RestoxTi e SixTi, exibem uma aderência quase que completa dos valores simulados aos valores originais. E, FexResto e SixResto, apesar de não exibirem este comportamento praticamente perfeito, mostram correlações suficientes. ....	145

Figura 4.16: Correlações lineares entre as variáveis originais (linhas vermelhas) e cenários simulados (linhas pretas). AlapxFe, AlapxTi, FexSi, FexTi, e SixTi, exibem uma aderência quase que completa dos valores simulados aos valores originais. E, AlapxSi, apesar de não exibirem este comportamento praticamente perfeito, mostram correlações suficientes. ....150

Figura A.0.1: médias locais ao longo dos eixos x, y e z, para a variável Alap (%) na combinação 1. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada..... 195

Figura A.0.2: médias locais ao longo dos eixos x, y e z, para a variável Fe (%) na combinação 1. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada..... 196

Figura A.0.3: médias locais ao longo dos eixos x, y e z, para a variável Resto (%) na combinação 1. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada..... 196

Figura A.0.4: médias locais ao longo dos eixos x, y e z, para a variável Si (%) na combinação 1. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada..... 197

Figura A.0.5: médias locais ao longo dos eixos x, y e z, para a variável Ti (%) na combinação 1. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada..... 197

Figura A.0.6: médias locais ao longo dos eixos x, y e z, para a variável Alap (%) na combinação 2. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada..... 198

Figura A.0.7: médias locais ao longo dos eixos x, y e z, para a variável Fe (%) na combinação 2. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada..... 198

Figura A.0.8: médias locais ao longo dos eixos x, y e z, para a variável Resto (%) na combinação 2. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada..... 199

Figura A.0.9: médias locais ao longo dos eixos x, y e z, para a variável Si (%) na combinação 2. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados



as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada.....	204
Figura A.0.20: médias locais ao longo dos eixos x, y e z, para a variável Ti (%) na combinação 4. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada.....	205
Figura A.0.21: médias locais ao longo dos eixos x, y e z, para a variável Alap (%) na combinação 5. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada.....	205
Figura A.0.22: médias locais ao longo dos eixos x, y e z, para a variável Fe (%) na combinação 5. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada.....	206
Figura A.0.23: médias locais ao longo dos eixos x, y e z, para a variável Si (%) na combinação 5. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada.....	206
Figura A.0.24: médias locais ao longo dos eixos x, y e z, para a variável Ti (%) na combinação 5. A linha vermelha espessa representa as médias locais dos dados originais, as linhas vermelhas tracejadas representam o desvio padrão das médias locais dos dados originais e as linhas pretas representam as médias locais dos vinte cenários simulados na referida combinação testada.....	207
Figura A.0.25: correlogramas experimentais (linha vermelha fina), correlogramas modelados (linha vermelha espessa) e correlogramas obtidos nas simulações (linhas pretas) para as variáveis Alap, Fe, Resto, Si e Ti.....	208
Figura A.0.26: correlogramas experimentais (linha vermelha fina), correlogramas modelados (linha vermelha espessa) e correlogramas obtidos nas simulações (linhas pretas) para as variáveis Alap, Fe, Resto, Si e Ti.....	209
Figura A.0.27: correlogramas experimentais (linha vermelha fina), correlogramas modelados (linha vermelha espessa) e correlogramas obtidos nas simulações (linhas pretas) para as variáveis Alap, Fe, Resto, Si e Ti.....	210
Figura A.0.28: correlogramas experimentais (linha vermelha fina), correlogramas modelados (linha vermelha espessa) e correlogramas obtidos nas simulações (linhas pretas) para as variáveis Alap, Fe, Resto, Si e Ti.....	211
Figura A.0.29: correlogramas experimentais (linha vermelha fina), correlogramas modelados (linha vermelha espessa) e correlogramas obtidos nas simulações (linhas pretas) para as variáveis Alap, Fe, Si e Ti.....	212

## LISTA DE TABELAS

Tabela 2.1: Exemplo de uma matriz de sinais, utilizada para codificar uma sequência de partição binária e construção de uma base ortonormal. A parte inferior da tabela mostra a matriz $\psi$ , da base equivalente. (Modificado de Pawlowsky-Glahn et al. (2013)).....	67
Tabela 2.2: Exemplo de uma matriz de sinais para $D=5$ , utilizada para codificar a partição binária sequencial de modo padrão. A parte inferior da tabela mostra a matriz $\psi$ da base.....	68
Tabela 3.1: Estatística básica univariada das variáveis estudadas. ....	99
Tabela 4.1: Parâmetros de busca utilizados nas simulações da variável Rec14# e em todas as demais simulações realizadas. ....	116
Tabela 4.2: Estatística básica dos dados originais (topo da tabela) e dos vinte cenários simulados para a variável Rec14#. Os valores simulados apresentam os resultados esperados, com valores mínimos e máximos preservados, assim como a média oscilando em torno dos valores originais.....	119
Tabela 4.3: Estatística básica dos dados originais (topo da tabela) e dos vinte cenários simulados para a variável Rec14#. ....	122
Tabela 4.4: Informações sobre o número de blocos estimados dentro do intervalo permitido (entre 95% e 100%) e a média da soma dos teores simulados para a faixa 14#. É possível perceber que apesar de a soma dos teores simulados não ser exatamente igual à soma dos teores dos dados originais (100%), os valores obtidos são bastante próximos e suficientes para uma análise preliminar.....	130
Tabela 4.5: Informações sobre o número de blocos estimados dentro do intervalo permitido (entre 95% e 100%) e a média da soma dos teores simulados para a faixa 14#. É possível perceber que a soma dos teores simulados é exatamente igual à soma dos teores dos dados originais (100%). ....	136
Tabela 4.6: Informações sobre o número de blocos estimados dentro do intervalo permitido (entre 95% e 100%), após o ajuste da variável Resto, e, a média da soma dos teores simulados para a faixa 14#. É possível perceber que apesar de a soma dos teores simulados não ser exatamente igual à soma dos teores dos dados originais (100%), os valores obtidos são bastante próximos e suficientes para uma análise preliminar.....	141
Tabela 4.7: Análise do fechamento da soma dos teores simulados para a faixa 14#. ....	146
Tabela 4.8: Análise do fechamento da soma das proporções simuladas para a faixa 14#. ....	151
Tabela 4.9: Médias globais dos dados originais (no topo) e dos 20 cenários simulados para Alap (abaixo). ....	152
Tabela 4.10: Médias globais dos dados originais (no topo) e dos 20 cenários simulados para Fe (abaixo). ....	152
Tabela 4.11: Médias globais dos dados originais (no topo) e dos 20 cenários simulados para Si (abaixo). ....	153
Tabela 4.12: Médias globais dos dados originais (no topo) e dos 20 cenários simulados para Ti (abaixo). ....	153
Tabela 5.1: quadro resumo com dados sobre a qualidade de reprodução dos resultados simulados quando comparados com os dados originais e sobre o grau de facilidade de implementação da metodologia.....	162
Tabela A.2: média dos dados originais para a variável Alap (%) e médias simuladas para os vinte cenários gerados para a mesma variável. ....	179



# SUMÁRIO

Resumo.....	8
Abstract .....	10
Lista de figuras .....	12
Lista de tabelas .....	18
Sumário .....	20
Capítulo 1.....	22
Introdução.....	22
1.1. A incerteza geológica.....	23
1.2. Métodos multivariados para explorar a incerteza geológica .....	25
1.3. Simulação direta em blocos.....	41
1.4. Análise de dados composicionais.....	44
1.5. Problema .....	51
1.6. Meta .....	53
1.7. Objetivos .....	53
1.8. Metodologia.....	54
1.9. Organização da tese .....	57
Capítulo 2.....	59
Revisão dos métodos aplicados.....	59
2.1. Análise dos dados composicionais (CODA).....	59
2.2. Decomposição em fatores de autocorrelação mínimos e máximos (MAF) .....	74
2.3. Simulação direta em blocos .....	78
2.4. Simulação por bandas rotativas .....	82
2.5. Comentários .....	85
Capítulo 3.....	86
O banco de dados .....	86
3.1. Generalidades .....	86
3.2. Mineralogia e geologia.....	87
3.3. A estatística básica .....	93
3.4. Comentários .....	112
Capítulo 4.....	113
Resultados e discussão.....	113
4.1. Simulação da variável Rec14# .....	114

4.2. Simulação dos dados originais .....	126
4.3. Simulação das ilrs .....	131
4.4. Simulação direta em blocos dos fatores MAF .....	137
4.5. Simulação direta em blocos dos fatores MAFs das ilrs .....	142
4.6. Simulação direta em blocos dos MAFs das ilrs com a operação de fechamento.	147
4.7. Comparação entre os métodos .....	155
4.8. Comentários .....	159
Capítulo 5.....	160
1. Conclusões .....	160
2. Recomendações para trabalhos futuros .....	163
Referências bibliográficas .....	166
Anexo A – Verificação da reprodução das características estatísticas de continuidade espacial.....	179
1. Médias globais .....	179
1.1. Médias globais – Combinação 1 .....	179
1.2. Médias globais – Combinação 2.....	182
1.3. Médias globais – Combinação 3.....	186
1.4. Médias globais – Combinação 4.....	189
1.5. Médias globais – Combinação 5.....	192
2. Médias locais .....	195
2.1. Médias locais – Combinação 1.....	195
2.2. Médias locais – Combinação 2.....	198
2.3. Médias locais – Combinação 3.....	200
2.4. Médias locais – Combinação 4.....	203
2.5. Médias locais – Combinação 5.....	205
3. Continuidade espacial.....	208
3.1. Correlogramas para a combinação 1 .....	208
3.2. Correlogramas para a combinação 2 .....	209
3.3. Correlogramas para a combinação 3 .....	210
3.4. Correlogramas para a combinação 4.....	211
3.5. Correlogramas para a combinação 5.....	212

# Capítulo 1

## INTRODUÇÃO

Os depósitos minerais são constituídos por diversas espécies químicas e propriedades físicas que devem ser propriamente identificadas e quantificadas para o correto planejamento das etapas posteriores de lavra e beneficiamento. As primeiras investigações geológico-mineiras da história se restringiam ao estudo da variável de interesse econômico existente no depósito, ignorando a existência de múltiplas variáveis, a correlação entre elas ou até mesmo sua influência no processo de beneficiamento.

A evolução das técnicas de pesquisa, caracterização tecnológica e de beneficiamento dos minérios, além de alertar para a importância, tornou praticamente imprescindível a determinação das características e a quantificação das múltiplas variáveis que influenciam os processos e a lucratividade do empreendimento mineiro.

A geoestatística é a técnica que permite realizar estimativas de variáveis regionalizadas por meio de diferentes variantes dos métodos de krigagem (Matheron, 1965), e, cuja aplicabilidade depende do objetivo final que se quer alcançar. Existem métodos mais adequados para a estimativa de uma só variável (Isaaks & Srivastava, 1989), para a estimativa de múltiplas variáveis de forma conjunta (Wackernagel, 1994), e, ainda, métodos capazes de avaliar a variabilidade *in situ* dos teores de um ou mais atributos correlacionados ou não (Chliés & Delfiner, 1999).

Depósitos de ferro, manganês, níquel e bauxita são exemplos de ocorrências em que a correta avaliação da distribuição dos teores de múltiplas espécies químicas e a possível correlação entre as mesmas devem ser garantidas. Além disto, as estimativas das múltiplas espécies químicas nesses depósitos devem garantir o fechamento dos balanços de massa e de teores nas diferentes faixas granulométricas. E, por serem sistemas que envolvem a

consideração conjunta de diversos fatores correlacionados, a utilização das técnicas da geoestatística multivariada é recomendável.

Métodos como a cokrigagem (Marechal, 1970) e suas variações fornecem estimativas bastante próximas à realidade na quantificação das variáveis presentes em um depósito mineral multivariado. Como uma extensão natural da krigagem (Matheron, 1965), a cokrigagem fornece estimativas não tendenciosas, que minimizam a variância do erro e ainda, como vantagem, consideram a correlação conjunta existente entre as variáveis.

No entanto, quando o objetivo é a previsão da variabilidade *in situ* dos teores, por exemplo, para fins de planejamento de mina, a cokrigagem não se apresenta como a melhor alternativa. Isto porque o método só é capaz de fornecer uma única estimativa sobre os teores das variáveis em um dado local. Mesmo que este valor estimado seja o mais próximo da realidade que se pode obter, a cokrigagem não considera adequadamente a variação que estes teores podem sofrer, já que por construção, os valores extremos são suavizados.

Surge então a motivação para a busca de metodologias alternativas para a estimativa da variabilidade dos teores em depósitos multivariados complexos.

### 1.1. A INCERTEZA GEOLÓGICA

O conceito de que os resultados obtidos na cokrigagem ou em qualquer outro método linear de estimativa não podem ser utilizados para alimentar processos de decisão que executem cálculos com base em funções não lineares é bastante consolidado (ex.: Chliés & Delfiner, 1999; Goovaerts, 1997; Lantuejoul, 2002). Como consequência deste fato, o conhecimento do intervalo de possíveis soluções se torna necessário para a avaliação da incerteza sobre a resposta associada ao atributo de interesse. A simulação estocástica é o processo que viabiliza a obtenção de inúmeros cenários equiprováveis, ou seja, fornece várias soluções utilizadas para a estimativa do intervalo de incerteza sobre as variáveis estudadas.

Os métodos de simulação estocástica têm sido largamente utilizados em diversas aplicações que vão desde a análise de risco para contaminações no solo até a análise econômica de projetos. No caso de um depósito mineral, a

estatística básica dos teores: média, variância e covariância; deverão ser respeitadas nos cenários simulados. Se, além destas características, os modelos reproduzirem nos locais amostrados os valores dos dados originais, estes modelos são chamados de condicionalmente simulados (Journel, 1994).

Os primeiros estudos envolvendo simulação geoestatística de teores em depósitos minerais ocorreram nas décadas de 70 e 80 (David, 1973; David *et al.*, 1974; Clark & White, 1976; Journel & Huijbregts, 1978; Dagbert, 1981; Chilès, 1984 e Deraisme *et al.*, 1984). Depois destes, muitos outros trabalhos foram desenvolvidos, principalmente envolvendo aplicações diretas na indústria mineira (Blackwell & Sinclair, 2002).

Soluções para locação amostral em exploração mineral foram propostas por Pilger (2000) e Koppe *et al.* (2011). Já Souza *et al.* (2004) fez uso de cenários simulados para melhorar a qualidade da estimativa e classificação de recursos em um depósito mineral. Bonato *et al.* (2000) e Blackwell *et al.* (1999) estudaram a simulação geoestatística como ferramenta auxiliar no planejamento da lavra. Ainda, Gambin *et al.* (2005), Marques *et al.* (2010), Beretta *et al.* (2010) e Abichequer *et al.* (2011) avaliaram os benefícios do método como auxiliar no planejamento de pilhas de homogeneização e na diminuição da variabilidade dos teores alimentados na planta de processamento.

Os métodos que quantificam e exploram a incerteza geológica são bastante eficientes quando se trata apenas da análise do comportamento univariado. No entanto, são pouco adaptados para a modelagem de fenômenos multivariados. Por exemplo, um modelo realista de um depósito de ferro deve conter informações sobre a sílica, o fósforo, a alumina e o manganês, além do ferro, para caracterizar adequadamente a distribuição dos teores em um bloco. Como o impacto da correlação existente entre as variáveis não é analisado nas etapas de planejamento da mina, otimização do *pit* ou sequenciamento, o modelo de blocos deve garantir que a dependência ou correlação entre os atributos seja previamente considerada e reproduzida. Um modelo realista irá implicar em *designs* de cava mais coerentes com a realidade, melhores indicadores econômicos e facilitar a análise do risco envolvido na exploração do depósito.

Preservar a variabilidade é importante, mas não o suficiente quando se trata de uma função de transferência multivariada. Marcotte *et al.* (2005)

demonstraram isto em um estudo de caso na área do cimento. Nove variáveis foram alimentadas em um *software* para homogeneização da mistura do cimento. Na primeira tentativa cada uma das variáveis foi inserida isoladamente, na segunda, a correlação entre elas foi considerada. Os resultados obtidos demonstraram que os valores simulados considerando a correlação eram muito mais precisos e próximos da realidade do que os que a omitiram. Quando o objeto de estudo for multivariado-correlacionado esta característica deve ser preservada para que os resultados gerados tenham certo grau de confiabilidade e reproduzam as características dos teores no depósito mineral.

Dimitrakopoulos & Ramazan (2004) demonstraram a importância de incorporar a correlação entre as variáveis nos modelos simulados por meio de um estudo de caso em um depósito laterítico de níquel-cobalto. O sequenciamento da lavra foi estudado segundo o comportamento de sete variáveis diferentes. Os modelos não correlacionados se mostraram muito otimistas e com variações de teor muito mais suavizadas do que os modelos correlacionados.

Também com o intuito de comprovar a importância da cossimulação Dimitrakopoulos & Fonseca (2003) analisaram as curvas de teor/tonelagem em um depósito de cobre no norte do Brasil. Como o teor de corte é dependente dos teores de cobre, ferro e potássio, é imprescindível que a correlação entre as variáveis seja reproduzida e incluída na análise da variação das curvas.

Outro aspecto relevante da incerteza geológica é a mudança de suporte amostral. Na mineração, o suporte amostral mais comum e praticamente único, são os furos de sondagem, enquanto que as estimativas são feitas em um suporte muito maior, os blocos. A transição de pontos para blocos é fundamental (Krige, 1951; Matheron, 1976; David, 1977; Journel & Huijbregts, 1978; David, 1988) para a correta avaliação dos depósitos, uma vez que a maioria das aplicações são função do teor dos blocos e da relação existente entre eles.

## 1.2. MÉTODOS MULTIVARIADOS PARA EXPLORAR A INCERTEZA GEOLÓGICA

A grande maioria dos métodos de simulação multivariada é bastante complexa, pouco prática e, para evitar complicações ainda maiores, considera só dois momentos principais, a média e a covariância (simples e cruzadas). O

conjunto de metodologias de cossimulação pode ser dividido em duas grandes categorias: no domínio da distribuição espacial e no domínio da distribuição espectral ou de frequências.

Independentemente do domínio escolhido, existem também dois grandes tipos de simulação: simulação condicional e não condicional. Simulações não condicionais geram cenários com valores reproduzindo adequadamente a covariância e a distribuição de frequências de um dado vetor de valores  $\mathbf{Z}(u)$ . Já uma simulação condicional, além de reproduzir a covariância e o histograma, também honra os valores nos locais onde  $\mathbf{Z}(u)$  é conhecido.

Aplicações na indústria mineira como o planejamento ou sequenciamento de mina requerem simulações condicionais. Se a simulação não for condicional, os resultados não serão restringidos pelos dados originais e a comparação entre os valores reais e simulados não será possível.

O pós-condicionamento das variáveis simuladas pode ser obtido por meio da adição do resíduo da cokrigagem dos dados inicialmente não condicionados (Carr & Meyers, 1985). O procedimento está demonstrado na equação 1.1, onde  $\mathbf{Z}_{cs}(u')$  é o vetor condicionalmente simulado para os  $n$  atributos,  $\mathbf{Z}_{nc}(u')$  são os dados cossimulados sem condicionamento,  $\mathbf{Z}_{nc}^*(u')$  é o vetor cokrigado, utilizando os dados não condicionados nos  $u$  locais, e,  $\mathbf{Z}_{ob}^*(u')$  são os dados originais cokrigados.

$$Z_{cs}(u') = Z_{nc}(u') - Z_{nc}^*(u') + Z_{ob}^*(u') \quad (1.1)$$

O método é muito laborioso e demanda alta capacidade computacional. Vários sistemas de cokrigagem precisam ser resolvidos, além da aplicação da equação de condicionamento em todos os locais estimados. Por exemplo, se existirem cinco variáveis amostradas em dez locais diferentes, o condicionamento irá requerer a inversão de uma matriz 50x50 termos para a obtenção de  $\mathbf{Z}_{ob}^*(u')$ , além da de uma muito maior para  $\mathbf{Z}_{nc}^*(u')$ , já que o *malha de estimativa (grid)* simulado possui muito mais locais com informações (previamente simuladas) do que o espaço amostrado. Isto faz com que a aplicação do método seja quase impossível na realidade da operação na indústria de mineração, já que a avaliação de depósitos minerais inclui a estimativa de

teores em grandes áreas, ou seja, muitos pontos ou blocos em um *malha de estimativa (grid)* de igual tamanho.

#### 1.2.1. O DOMÍNIO DE FREQUÊNCIAS OU ESPECTRAL

Os métodos de decomposição espectral admitem as amostras como "sinais" que podem ser decompostos em "ondas" de diferentes frequências por meio da transformada de Fourier. Muitos métodos realizam a decomposição das amostras para o domínio das frequências, descorrelacionando-as. Assim, estas podem ser simuladas de maneira independente e a correlação entre elas será mantida quando os valores simulados retornarem ao espaço original. Embora estes métodos sejam computacionalmente eficientes, a dificuldade de gerar cenários condicionalmente simulados e de simular vetores com muitas variáveis ou extensos, faz com que sua aplicação na vida real seja prejudicada.

Borgman *et al.* (1984) propuseram um método para cossimular as variáveis dentro do domínio de frequências, baseando-se na propriedade que um conjunto  $\mathbf{Z}_n$  de variáveis, onde  $n=0, \dots, N-1$ , é periódico, correlacionado, estacionário, com valores reais e comprimento máximo  $N$ . Sua transformada de Fourier pode ser descrita por  $\mathbf{A}_m$ ,  $m=1, \dots, N/2$ , um conjunto de dados não correlacionados de valores complexos. A ideia é simular os  $\mathbf{A}_m$  e depois retornar ao espaço original invertendo os  $\mathbf{A}_m$  simulados por meio de um algoritmo rápido para inversão de transformadas de Fourier (FFT). O resultado será um conjunto de valores simulados com a correlação cruzada assegurada.

O método é rápido, mas a grande desvantagem é que a simulação não é condicional e, assim como os outros métodos aplicados no domínio das frequências, exige uma série periódica de valores. Motivo que torna sua aplicação só viável quando a malha de amostragem for regular, prejudicando o pós-condicionamento dos atributos. Como solução os autores propuseram três maneiras de pós condicionamento, sendo que todas elas exigem que a simulação seja realizada em grandes malhas e com poucos dados condicionantes. Estas condições são muito restritivas para a aplicação no meio mineral, onde os dados não obedecem a uma distribuição espacial regular e onde existe grande quantidade de informação amostrada. Assim, a cokrigagem do resíduo volta a ser a melhor opção para o pós-condicionamento dos dados simulados.

Fazendo uso da representação espectral da covariância Robin *et al.* (1993) propuseram um método não condicional para geração de duas funções randômicas correlacionadas. Utilizaram o teorema da decomposição espectral para substituir estruturas correlacionadas no espaço original, por estruturas descorrelacionadas no domínio das frequências. A retro transformação dos dados descorrelacionados no domínio das frequências para o espaço original onde devem estar correlacionadas é feita por uma transformada direta de Fourier. É uma adaptação melhorada da metodologia proposta por Borgman *et al.* (1984), já que não precisa de um modelo de covariâncias, a covariância é extraída diretamente dos dados. O método é muito eficiente computacionalmente, mas apresenta uma série de desvantagens. As simulações não são condicionais e ainda, a malha de estimativa (*malha de estimativa (grid)*) não pode ser definido de acordo com a distribuição espacial dos dados, mas sim, deve ser função da covariância para evitar a periodicidade e minimizar os erros de aproximação. Como na mineração as malhas de amostragem não são necessariamente regulares e a discretização dos blocos depende dos parâmetros da lavra, o método também não se aplica para estas situações.

Pardo-Iguzquiza & Chica-Olmo (1994) propuseram uma extensão do método da integral de Fourier (Pardo-Iguzquiza & Chica-Olmo, 1993) para obtenção de cenários não condicionalmente simulados. É um método de simulação no espaço espectral baseado em uma função de densidade cruzada obtida pela transformada de Fourier das funções de covariância cruzadas do espaço original. Basicamente, é uma adaptação do método desenvolvido por Borgman *et al.* (1984). A grande vantagem da proposta é que o modelo linear de correogionalização (MLC), indispensável para a cokrigagem, não precisa ser ajustado, no entanto, as condições de definição positiva devem ser checadas para todos os vetores de separação  $h$  necessários. Além disto, é um método bastante eficiente computacionalmente. Mas, assim como os demais, produz resultados não condicionados e por isso, não apresenta vantagem para utilização na indústria mineira.

Gutjahr *et al.* (1997) utilizou o método da simulação espectral para cossimular funções randômicas correlacionadas e propôs um método para o pós-condicionamento ainda no domínio das frequências. Assim como as metodologias abordadas até o momento, a idéia é decompor  $\mathbf{Z}(u)$  em dados

descorrelacionados no domínio espectral. Estes são simulados e voltam ao espaço original por meio da transformação de Fourier. Ao invés de pós-condicionar os cenários com a cokrigagem do resíduo, os autores propõem uma alternativa ainda dentro do espaço espectral, já que a metodologia permite o condicionamento de uma das duas variáveis envolvidas. Considerando  $dY_1(f)$  uma função randômica complexa obtida a partir da decomposição espectral de  $Z_1(u)$ , a idéia é obter uma  $d^*Y_1(f)$  krigada e condicionada pelos dados. Para isto, a covariância entre  $dY_1(f)$  e os dados originais deve ser analisada, considerando as partes reais e imaginárias. A cossimulação condicional de  $Z_1(u)$  é feita no domínio espectral, a partir dos dados krigados,  $d^*Y_1(f)$ . O método é bastante complexo e somente uma das variáveis é condicionada, sendo que embora os autores garantam que a aplicação possa ser estendida para mais atributos, isto ainda não foi demonstrado.

O método apresenta a vantagem de não precisar do ajuste do Modelo Linear de Corregionalização (MLC) inerente à cokrigagem, no entanto sua aplicação é limitada a malhas regulares e pouco extensas. Gutjahr *et al.* (1997) utilizaram uma malha de estimativa (*grid*) 2D regular de 128x128metros. Estas dimensões são muito pequenas para estudos na área da mineração, o que mais uma vez, desfavorece a proposta.

Em geral, os métodos de cossimulação no domínio das frequências, ou espectral, não são eficientes para aplicação no meio mineral por todos os motivos já abordados, por exemplo, periodicidade, restrições na malha amostrada e incapacidade de simular cenários extensos. A maioria foi desenvolvida para aplicações no meio ambiente e na hidrogeologia, onde o condicionamento não é tão importante quanto na mineração. A grande vantagem destes métodos é a velocidade de processamento, mas a desvantagem é que a maioria precisa do pós condicionamento dos modelos aos dados.

No entanto, o princípio da decomposição de um vetor randômico se mostrou muito eficiente. Isto é, decompor um dado vetor  $\mathbf{Z}(u)$ , em dados não correlacionados, simular cada um independentemente e por fim, retornar ao espaço original.

### 1.2.2. O DOMÍNIO ESPACIAL

As primeiras aplicações de métodos de cossimulação no domínio espacial foram propostas por Dowd (1983) e Chilès (1984), utilizando variáveis auxiliares obtidas pela análise da correlação.

Menos específico, mas ainda não completamente aplicável em depósitos multivariados complexos, Carr & Meyers (1985) introduziram a utilização da correlação cruzada para simular independentemente cada atributo e o pós-condicionamento dos cenários por meio da cokrigagem. Resumidamente, o método utiliza a simulação em bandas rotativas (Matheron, 1973; Journel e Huijbregts, 1978) para simular independentemente e não condicionalmente, cada função randômica, utilizando a cokrigagem para o pós condicionamento dos cenários.

O método proposto não é uma cossimulação propriamente dita, já que a correlação cruzada entre as variáveis só é considerada na etapa de pós condicionamento via cokrigagem. No entanto, os autores concluíram que os resultados obtidos mesmo com a incorporação tardia da correlação cruzada, são mais próximos da realidade que os gerados na simulação individual dos atributos. Seguindo esta mesma linha de raciocínio, qualquer conjunto de simulações univariadas não condicionais, poderia ser pós processado via cokrigagem. Mas, isto não garantiria teoricamente a reprodução da correlação existente entre as variáveis.

Myers (1988), estendeu a decomposição LU da matriz condicional de covariância univariada desenvolvida por Davis (1987), para aplicação em depósitos multivariados. Com as matrizes triangulares inferior e superior obtidas pela decomposição de Cholesky das matrizes de covariância, variância e covariância cruzada (**C**), a cossimulação é realizada multiplicando-se a matriz triangular inferior (**L**) por um número randômico **w**.

$$\begin{aligned} \mathbf{C} &= \mathbf{LU} \\ \mathbf{Z}(u) &= \mathbf{Lw} \end{aligned} \tag{1.2}$$

O ajuste do MLC é necessário para garantir as condições de definição positiva das matrizes de variância, covariância e covariância cruzada. O método compartilha as mesmas vantagens do caso univariado, como por exemplo, a velocidade de processamento para geração de vários cenários simulados e a necessidade de somente ser necessária a decomposição de uma matriz. A metodologia apresenta algumas desvantagens, sendo que a incapacidade de lidar com grande número de dados é a mais significativa para a aplicação em depósitos minerais.

Verly (1993) combinou a simulação LU dos vetores com a sequência de simulação nos nós do malha de estimativa (*grid*), para tornar viável a geração de imagens envolvendo grandes áreas. Resumidamente o método funciona da seguinte maneira: a simulação LU é realizada em cada nó do malha de estimativa (*grid*), para cossimular todas as variáveis neste nó, antes de seguir para o próximo. O processo é repetido até que todos os locais sejam visitados. Mas, mais uma vez, o ajuste do MLC é necessário.

Uma abordagem diferente do problema foi proposta por Gomez-Hernandez & Journel (1993). Em vez de utilizar a decomposição LU, os autores estimaram via cokrigagem simples, a média e a variância condicionadas em cada nó da malha de estimativa (*grid*). O resultado da simulação é amostrado randomicamente da distribuição condicional gaussiana resultante de cada nó simulado. A sequência de simulação da malha de estimativa (*grid*) está baseada na decomposição das variáveis em termos da distribuição condicional (Rosenblatt, 1952; Johnson, 1987; Ripley, 1987), onde uma função de distribuição de probabilidade condicional (pdf) de  $\mathbf{Z}(u_i)$  com N pontos pode ser expressa em termos da pdf de um único ponto N.

A esperança matemática e a variância de cada variável são obtidas por cokrigagem simples, no método proposto por Gomez-Hernandez & Journel, (1993) e pela decomposição LU no método de Verly (1993). As duas abordagens garantem a reprodução dos histogramas, dos variogramas diretos e cruzados, para vizinhanças suficientemente grandes.

O grande problema é que a cokrigagem simples das variáveis é bastante trabalhosa, dependendo do número de dados condicionantes e da estimativa em cada um desses nós. Como alternativa, os autores propõem a diminuição de

dados condicionantes, mas de modo que a reprodução dos variogramas não seja afetada. Mais uma vez, o ajuste do MLC está presente para garantir as condições de definição positiva das matrizes de covariância.

Como extensão da correspondência da decomposição LU com a krigagem simples (Alabert, 1987), a cossimulação sequencial dos vetores por decomposição LU ou por cokrigagem simples são praticamente a mesma coisa.

Uma alternativa bastante interessante para acessar a incerteza geológica sobre os teores em depósitos multivariados é a utilização de métodos baseados na decomposição das matrizes correionalizadas, obtidas pela modelagem da variabilidade espacial dos atributos (Wackernagel, 1994). A maioria dos métodos tenta fatorizar as variáveis com o objetivo de gerar vetores ortogonais, assumidos como independentes para todos os vetores de separação  $h$ , e por isso, permitindo que cada variável seja simulada independentemente. Depois da retro transformação, os momentos de primeira e segunda ordem (média e variância) são reproduzidos para todas as variáveis.

Sendo  $\mathbf{A}$  um conjunto de  $p$  vetores ortogonais que fazem a transformação linear do vetor  $\mathbf{Z}(u)$  em  $\mathbf{Y}(u)$ , tal que  $\mathbf{Y}(u)=\mathbf{Z}(u)\mathbf{A}$ . Se a matriz de covariância  $\mathbf{C}_z(h)$  de  $\mathbf{Z}(u)$  satisfizer determinadas condições, a covariância de  $\mathbf{Y}(u)$  será diagonalizável.

Sendo ortogonal,  $\mathbf{Y}(u)$  pode ser simulado de maneira independente, com qualquer método de simulação univariada. A retro transformação do vetor  $\mathbf{Z}(u)$  é obtida da seguinte maneira (1.3):

$$\mathbf{Z}(u) = \mathbf{Y}(u)\mathbf{A}^{-1} \quad (1.3)$$

onde:

$\mathbf{Z}(u)$  será o conjunto de valores estimados correlacionados

$\mathbf{Y}(u)$  será o conjunto de valores estimados de forma independente

$\mathbf{A}^{(-1)}$  é a matriz inversa de  $\mathbf{A}^{(1)}$ , utilizada para a retrotransformação

A decomposição em componentes principais (PCA), introduzida na geoestatística por David *et al.* (1984), é o método mais utilizado para simulação de depósitos multivariados (David *et al.* 1984; Suro-Perez & Journel, 1990;

Suro-Perez & Journel 1991). Nela, os vetores ortogonais são obtidos pela decomposição espectral da matriz de variância/covariância. Muitos autores atentaram para o fato de que o método só ortogonaliza  $\mathbf{Z}(u)$  quando a correlação entre as variáveis for intrínseca ou quando a matriz de covariâncias de  $\mathbf{Z}(u)$  for diagonalizável (Wackernagel *et al.* 1989; Petitgas & Touffait, 1989; Lajaunie, 1992; Goovaerts, 1993). Caso contrário, poderão aparecer falsas correlações entre os fatores, dificultando a reprodução da correlação original depois da retro-transformação. Um modelo de correlação intrínseca é uma séria restrição para os ajustes dos variogramas diretos e cruzados, fazendo com que a metodologia possa ser aplicada em muito poucos casos na prática.

Como alternativa ao PCA, Desbarats & Dimitrakopoulos (2000) propuseram a decomposição em fatores de auto correlação mínimos e máximos (MAF) para ortogonalizar  $\mathbf{Z}(u)$ . A decomposição MAF foi primeiramente desenvolvida por Switzer & Green (1984) para aplicações na área do sensoriamento remoto e tem como grande vantagem, o fato de descorrelacionar as variáveis para qualquer distância de separação, desde que os variogramas de  $\mathbf{Z}(u)$  atendam a um modelo linear de correogionalização com no máximo duas estruturas.

A possibilidade de utilização de duas estruturas para a modelagem da covariância cruzada faz com que a decomposição MAF possa ser aplicada em muitos mais casos do que a por PCA. A descrição completa da decomposição em fatores MAF será detalhada no capítulo 2.

Existem diversos outros métodos capazes de ortogonalizar  $\mathbf{Z}(u)$ . Tercan (1999) comparou cinco métodos, incluindo a decomposição espectral e a decomposição de Choleski, para concluir que o melhor entre eles era a decomposição pelo método de Choleski (a decomposição MAF não fez parte do estudo). Os vetores finais da aplicação deste método, assim como os da decomposição MAF, são resultado de duas decomposições da matriz de variância/covariância em vetores de separação para diferentes  $h$ s. Uma decomposição para o vetor de separação nulo, e, outra para o *lag* correspondente ao alcance segunda estrutura do variograma (considerando o efeito pepita como sendo a primeira estrutura).

A grande vantagem dos métodos baseados na fatorização das variáveis correlacionadas é que, na prática, os fatores que podem ser estimados de modo independente, também podem ser modelados de modo independente, não necessitando do ajuste do MLC. No entanto, quando os fatores são ajustados isoladamente é necessário que se verifique a reprodução dos variogramas diretos e cruzados, além da reprodução da correlação existente entre os dados, no modelo final retro transformado ao espaço original.

Bandarian *et al.* (2008) estudaram uma alternativa para a decomposição MAF para o caso de simulações com abordagem multi gaussiana, com o objetivo de eliminar o primeiro passo, a normalização dos dados de cada atributo. Os autores propuseram um método para obtenção dos fatores MAF partindo diretamente dos dados originais, o que chamaram de DMAF.

O cálculo dos coeficientes de transformação DMAF é realizado assumindo-se um MLC com duas estruturas e duas transformações do tipo PCA seguidas. Estes coeficientes serão posteriormente utilizados para a transformação dos dados originais em fatores DMAF.

Sendo a matriz de variância-covariância  $\mathbf{C}(\mathbf{h}) = \mathbf{B}_0 c_0(\mathbf{h}) + \mathbf{B}_1 c_1(\mathbf{h})$ , onde as matrizes de correionalização definidas positivas,  $\mathbf{B}_0$  e  $\mathbf{B}_1$  contêm os patamares dos modelos de covariância definidos positivos  $c_0$  e  $c_1$ , respectivamente. E,  $\mathbf{B} = cov\{\mathbf{Z}(\mathbf{u}), \mathbf{Z}(\mathbf{u})\}$ , como a matriz de variância-covariância definida positiva  $\mathbf{C}(0)$  para  $\mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1$ , o primeiro passo é ortogonalizar  $\mathbf{B}$  de acordo com  $\mathbf{B} = \mathbf{Q}_1 \mathbf{\Lambda}_1 \mathbf{Q}_1^T$ , onde  $\mathbf{Q}_1$  é a matriz de autovetores de  $\mathbf{B}$ ,  $\mathbf{Q}_1^T \mathbf{Q}_1 = \mathbf{I}$  e  $\mathbf{\Lambda}_1$  é a matriz diagonal dos autovalores correspondentes arranjados em ordem decrescente de magnitude.

Definir  $\mathbf{W}$ , como  $\mathbf{W} = \mathbf{Q}_1 \mathbf{\Lambda}_1^{-1/2}$  de modo que  $\mathbf{W}^T \mathbf{B} \mathbf{W} = \mathbf{I}$ , então a primeira transformação será:  $\mathbf{C}_w(\mathbf{h}) = \mathbf{W}^T \mathbf{B}_0 \mathbf{W} c_0(\mathbf{h}) + \mathbf{I} c_1(\mathbf{h}) - \mathbf{W}^T \mathbf{B}_1 \mathbf{W} c_1(\mathbf{h})$ . No entanto, esta primeira transformação só ortogonaliza os fatores transformados  $\mathbf{W}(\mathbf{u}) = (\mathbf{Q}_1 \mathbf{\Lambda}_1^{-1/2})^T \mathbf{Z}(\mathbf{u})$ , para o vetor de separação nulo. Para os casos em que a correlação não for intrínseca, os fatores ainda poderão estar correlacionados para separações maiores do que zero.

Então, uma segunda transformação por PCA é realizada na matriz simétrica  $\mathbf{C}_w(\mathbf{h})$ , e, a matriz resultante dos autovetores  $\mathbf{Q}_2$ , será utilizada para ortogonalizar os fatores transformados  $\mathbf{M}(\mathbf{u}) = \mathbf{Q}_2^T \mathbf{W}(\mathbf{u})$  para todos

os  $h$ . Assim a nova matriz será:  $C_W(\mathbf{h}) = \mathbf{Q}_2[\Lambda_2\mathbf{c}_0(\mathbf{h}) + (\mathbf{I} - \Lambda_2)\mathbf{c}_1(\mathbf{h})]\mathbf{Q}_2^T$ , onde  $\mathbf{Q}_2$  é a matriz ortogonal de autovetores e  $[\Lambda_2\mathbf{c}_0(\mathbf{h}) + (\mathbf{I} - \Lambda_2)\mathbf{c}_1(\mathbf{h})]$  é a matriz diagonal dos autovalores correspondentes.

A matriz de transformação  $\mathbf{A}^T$  é obtida pela combinação das duas decomposições, sendo  $\mathbf{A}^T = \mathbf{Q}_2^T\Lambda_1^{-1/2}\mathbf{Q}_1^T$ . Os atributos são transformados em fatores seguindo  $\mathbf{M}(\mathbf{u}) = \mathbf{A}^T\mathbf{Z}(\mathbf{u})$ .

Os fatores DMAF são simulados de maneira independente e os resultados são retrotransformados para representarem os valores simulados dos dados originais, utilizando-se a inversa da matriz de transformação. Então,  $\mathbf{Z}(\mathbf{u}) = (\mathbf{A}^T)^{-1}\mathbf{M}(\mathbf{u})$ .

Além das decomposições MAF e DMAF, existem diversos métodos que podem ser utilizados para descorrelacionar uma família de atributos até a distância necessária, sendo que todos eles estão baseados na diagonalização das matrizes de variância/covariância e de variogramas. Mueller & Ferreira (2012), propuseram a diagonalização por meio de iterações gaussianas uniformemente distribuídas (U-Wedge). As autoras testaram a aplicabilidade da proposta em um banco de dados de ferro e provaram que a grande vantagem do método é que este pode ser aplicado aos casos onde os variogramas seriam melhor ajustados com três estruturas, já que as decomposições MAF e DMAF restringem a descorrelação a variogramas que poderiam ser modelados com no máximo duas estruturas. No entanto, a grande maioria dos depósitos minerais têm a continuidade espacial das variáveis suficientemente caracterizadas até o alcance da segunda estrutura do variograma.

Todos os métodos citados anteriormente foram desenvolvidos para funções estacionárias. Embora sejam apropriados para a grande maioria dos casos, modelos não estacionários devem ser considerados. Matheron (1973) assumiu funções randômicas intrínsecas de ordem  $k$ , ou IRF- $k$ , como modelos não estacionários descritos como incrementos de funções randômicas estacionárias. Apesar da elegância teórica do método para a modelagem de fenômenos complexos, o IRF- $k$  é muito pouco utilizado por sua complexidade e principalmente pela dificuldade na inferência dos parâmetros de ajuste (Definer, 1976; Matheron, 1976; Dimitrakopoulos, 1990; Chliés & Delfiner, 1999).

Dowd (1988) cossimulou a profundidade e a espessura de uma formação geológica, para testar a IRF-2 ( $k=2$ ), como exercício para avaliar o grau de dificuldade de inferir as covariâncias cruzadas. Em seu trabalho, o autor ressaltou a dificuldade de obtenção dos coeficientes dos polinômios dos incrementos e a sensibilidade do resultado final, como dependente destes ajustes. Em geral, os usuários preferem fazer pequenos ajustes nos métodos estacionários, como alterações na vizinhança de busca por exemplo, a utilizar o IRF-k.

### 1.2.3. APLICAÇÕES DOS MÉTODOS DE COSSIMULAÇÃO EM DEPÓSITOS MULTIVARIADOS COMPLEXOS.

Antes de Carr & Meyers (1985), alguns métodos com propostas diferentes eram utilizados para a cossimulação de variáveis correlacionadas, principalmente em aplicações na mineração e na indústria do petróleo. No entanto, todos eles estavam baseados em metodologias adaptadas para resolver problemas específicos, além de não serem métodos de simulação multivariada propriamente ditos.

Chilès (1984) simulou um depósito laterítico 2D utilizando a espessura como variável principal e removendo sua influência sobre as demais por meio de regressões lineares. O objetivo do estudo era a reprodução do gráfico de dispersão das variáveis correlacionadas. Consequentemente, todas as correlações cruzadas entre as variáveis foram obtidas nos *scatterplots* e não em variogramas cruzados. Cada variável foi simulada independentemente e sem condicionamento, retro-transformada para o espaço original e krigada para o pós-condicionamento (com a utilização da correlação intrínseca, a cokrigagem é equivalente à krigagem).

Dowd (1983) também propôs uma abordagem para solucionar um problema específico quando simulou a espessura de quatro camadas sedimentares em um reservatório de petróleo. Com o objetivo de manter uma representação realista da bacia, algumas correlações entre os atributos foram consideradas para a elaboração do modelo. Por exemplo, um aumento na espessura provocado por dobramentos ou uma redução drástica causada por um aumento de pressão sobre as camadas, devem ser refletidos em todas elas. O

autor propôs que a correlação entre as espessuras das camadas fosse obtida por meio de variáveis auxiliares obtidas na normalização cruzada entre as espessuras das camadas. Utilizando métodos de simulação condicional e transferindo os resultados para o espaço original, o método não pode ser definido como multivariado propriamente dito, já que somente as correlações forçadas são utilizadas.

Com o desenvolvimento de métodos verdadeiramente multivariados de simulação, tanto no domínio espectral quanto no domínio espacial, simulações genuinamente multivariadas em 2D e 3D começaram a surgir na literatura. Sendo que, aplicações no domínio espectral se restringem a aplicações na hidrologia e no meio ambiente, onde se pode trabalhar com áreas menores e o condicionamento não é essencial. Já os métodos como a simulação por bandas rotativas, ou por métodos sequenciais são largamente utilizadas na mineração e no petróleo exatamente pelo motivo contrário: são capazes de simular grandes áreas, além de o condicionamento nos métodos sequenciais ser bastante simples.

Shive *et al.* (1990) analisou a performance dos métodos geofísicos: gravimetria, sísmica, GPR e resistividade, para localizar cavidades de diferentes tamanhos. Para isto, a velocidade de propagação na sísmica, densidade, resistividade e a constante dielétrica foram cossimuladas em um malha de estimativa (grid) 3D e as cavidades reais também foram consideradas para comparação com os resultados da simulação.

Esta é uma aplicação da simulação multivariada no domínio das frequências, assim como a proposta por Borgman *et al.* (1984) onde as realizações foram posteriormente correlacionadas com o método geofísico a que estavam atreladas. As imagens simuladas com as respostas dos métodos geofísicos serviram como base para a análise da performance de cada método.

Gutjahr *et al.* (1994) cossimularam a permeabilidade e a topografia no domínio espectral. A correlação entre as variáveis foi reproduzida por um modelo linear, posteriormente utilizado para a cossimulação por meio de um algoritmo rápido para aplicação da transformada de Fourier. A covariância foi obtida pela relação física existente entre os atributos e o pós condicionamento foi feito para dois casos: utilizando somente dados de transmissividade e utilizando dados de

transmissividade e topografia. O algoritmo provou ser bastante rápido e eficiente para um malha de estimativa (grid) 2D de 64x64m, com dimensões muito pequenas para aplicação em depósitos minerais ou na indústria do petróleo. O mais interessante do trabalho é que a correlação entre as variáveis foi extraída de uma correlação física representada por uma equação diferencial parcial, em vez da utilização do comportamento de distribuição espacial.

Guibal *et al.* (1996) utilizou a simulação multivariada por bandas rotativas para acessar a incerteza sobre os teores em um depósito de ferro do oeste da Austrália. Seis elementos foram considerados: Fe, SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, P, Mn e PPC, além da covariância cruzada modelada pelo MLC. O método apresentou bons resultados, tendo em vista que os histogramas e variogramas foram reproduzidos. No entanto, os autores mencionaram que sua aplicação é bastante complexa e que 21 variogramas tiveram que ser modelados, sem falar no pós condicionamento dos resultados.

Leuangthong *et al.* (2002) utilizou o modelo de Markov para cossimular um depósito laterítico de níquel, atentando para as tendências e restrições mineralógicas. Para simular quatro variáveis (Ni, Fe, MgO e SiO<sub>2</sub>), os autores utilizaram uma variável secundária para indentificar alguma possível tendência vertical e removê-la dos dados originais. Para isto, as tendências de primeira e segunda ordem foram obtidas plotando-se as médias e desvios padrão em função da profundidade. As tendências sobre as médias foram removidas pela utilização do resíduo obtido por regressão linear e as tendências na variância foram consideradas nos variogramas verticais, em função do aumento da profundidade. Os resíduos foram simulados utilizando uma classificação hierárquica (Almeida & Journel 1994), onde somente o modelo de correionalização de Markov é necessário. Logo, a correlação entre as variáveis para o vetor de separação nulo é suficiente.

A classificação hierárquica funciona da seguinte maneira: primeiro as variáveis devem ser classificadas por ordem de importância; depois a primeira delas é simulada de modo independente das demais e por fim as demais variáveis são simuladas utilizando os dados da primária simulada nos pontos amostrados, como informação secundária. Embora o método não seja tão completo quanto o pós condicionamento via cokrigagem, os resultados são mais

estáveis porque as variáveis secundárias são utilizadas em menos locais, além disto, é bastante satisfatório para aplicações onde há abundância de informação secundária. No entanto, é importante salientar que justamente a estimativa do comportamento da variável mais importante hierarquicamente, acaba não incluindo nenhuma informação sobre as demais.

Soares (2001) também utilizou o modelo de correionalização de Markov, mas elaborou um algoritmo de simulação direta que não requer que se assuma a multigaussianidade dos dados. No modelo, a pdf é construída a partir do histograma dos dados krigados ou cokrigados, quando condicionada aos dados colocados, à média e à variância das estimativas. A grande vantagem é a não necessidade de normalização dos dados, além de não ser necessário assumir um comportamento multigaussiano, uma hipótese bastante forte. No entanto, o método ainda peca pela utilização do modelo de correionalização de Markov.

Para resolver este problema, Horta & Soares (2010) propuseram um método de cossimulação (Co-DSS) baseado em funções de distribuição de probabilidade conjunta. Considerando duas variáveis correlacionadas,  $Z_1(u)$  e  $Z_2(u)$ , os valores de  $Z_2(u)$  utilizados na simulação são retirados da função de distribuição de probabilidade conjunta entre  $Z_1(u)$  e  $Z_2(u)$ . Ou, mais precisamente, da distribuição de  $Z_2(u_1)$ , no local  $u_1$ , condicionada ao valor de  $Z_1(u_1)$  previamente simulado. A metodologia garante a reprodução das funções de distribuição de probabilidades e a correlação nos modelos simulados. A única desvantagem do método está no fato de que o Modelo de Markov é um tanto quanto simplista na modelagem da covariância cruzada.

David *et al.* (1984) simularam por decomposição em componentes principais (PCA) as variáveis  $U_3O_8$  e As presentes em um depósito de urânio cujos limites dos domínios haviam sido previamente simulados com o uso de indicadores. Para isto, os atributos  $U_3O_8$  e As foram rotacionados com a decomposição espectral da matriz de covariância, para que se tornassem duas componentes principais independentes. A simulação independente de cada um dos fatores, agora combinações lineares de  $U_3O_8$  e As, reproduziram muito bem a correlação existente entre as variáveis, quando retro-transformados.

Seguindo David (1988), Suro-Perez & Journel (1990) utilizaram a decomposição PCA para simular seis fácies fortemente correlacionadas presentes

em um reservatório de petróleo. As fácies foram representadas por indicadores, de acordo com:

$$I_k(u; k) = 1 \text{ se } Z_k(u) = \text{categoria } k$$

$$I_k(u; k) = 0 \text{ para os casos contrários}$$

Depois, a matriz de covariância foi decomposta em componentes principais, que foram independentemente simulados e por fim, trazidos para a base original.

A grande vantagem sobre o método proposto por Guibal *et al.*(1996) é que os autores conseguiram simular as seis variáveis com somente seis variogramas, ao invés de precisar de vinte e um modelos ajustados. Após retornar para o espaço original, a proporção e as covariâncias diretas e cruzadas foram bem reproduzidas. Em seu trabalho os autores concluíram que a utilização de uma única matriz diagonal de covariâncias para todos os vetores de separação  $h$  foi suficiente, assim como para a grande maioria dos casos em que aplicaram o método. No entanto, recomendaram que a ortogonalização dos vetores deve sempre ser verificada para maiores separações.

Desbarats & Dimitrakopoulos (2000) utilizaram a decomposição em fatores de auto correlação mínimos e máximos (MAF) para simular a distribuição dos tamanhos dos poros em um depósito de tufos no Arizona. Os autores subdividiram a distribuição global em quinze novas sub classes. Como uma classe continha informações relevantes sobre as demais, a idéia foi utilizar a correlação cruzada entre elas, para que toda informação disponível fosse aproveitada na simulação de determinada classe. Depois da decomposição, os autores perceberam que aproximadamente 50% das classe correspondiam a ruídos nos dados e por isto reduziram o número de classes para seis. Estas foram simuladas independentemente com um algoritmo de simulação sequencial gaussiana e quando de volta ao espaço original reproduziram muito bem as correlações cruzadas entre as classes. No entanto, nesta metodologia os cenários foram simulados em pontos e precisavam ser reblocados após todo o procedimento, o que aumentava bastante o volume de trabalho dependendo do número de modelos gerados.

Para resolver este problema, Boucher & Dimitrakopoulos (2009;2012) desenvolveram um algoritmo para a simulação direta em blocos de múltiplas

variáveis correlacionadas (DBMAFSIM). A metodologia propõe que os dados sejam primeiramente ortogonalizados com o uso da decomposição em fatores de autocorrelação mínimos e máximos (MAF) e posteriormente simulados em pontos equivalentes aos da discretização que se quer para os blocos. Por fim, o algoritmo desenvolvido calcula a média destes pontos e a assume como sendo a média do bloco. A metodologia foi testada em estudos na área do petróleo e em um banco de dados de ferro com cinco variáveis, Fe, SiO<sub>2</sub>, Al<sub>2</sub>O<sub>3</sub>, P e PPC, variáveis críticas na definição da qualidade deste minério. Os modelos simulados reproduziram os variogramas diretos e cruzados, mesmo que a simulação tenha sido feita de maneira descorrelacionada, considerando somente os variogramas diretos. A metodologia é, sem dúvida, bastante eficiente na caracterização da variabilidade *in situ* de teores de depósitos multivariados complexos.

A desvantagem da proposta está na não adequada reprodução dos variogramas cruzados para os casos em que o coeficiente de correlação linear é significativamente diferente do coeficiente de ranqueamento, ou seja, a correlação entre as variáveis é não linear. Como alternativa, os autores propuseram a utilização dos métodos propostos por Soares (2001) e Bandarian (2008), que não requerem a prévia normalização dos dados, já que este é o fator que mais influencia para que os variogramas cruzados não sejam adequadamente reproduzidos.

Tanto o método proposto por Bandarian (2008), quanto o proposto por Soares (2001) geram cenários com simulações de teores em pontos, não em blocos. Assim, retorna o problema da reblockagem após todo o procedimento.

### 1.3. SIMULAÇÃO DIRETA EM BLOCOS

O segundo aspecto essencial da representação numérica dos teores de um depósito mineral é o suporte em que a informação está disponível. Em aplicações mineiras, o suporte de blocos deve ser utilizado, sendo este, função da seletividade que pode ser obtida pelos equipamentos usados e pelo método de lavra. Conseqüentemente, as funções de transferência que possuem relação não linear com os teores (como planejamento ou *design* de cava) requerem seu cálculo usando suporte de blocos. A grande maioria dos métodos multivariados

existentes é capaz apenas de simular pontos. Assim, em seguida estes devem ser reblockados para que possam ser utilizados para os mais diversos fins.

A mudança de suporte pode ser realizada com um método bastante seguro, mas pouco eficiente (Figura 1.1), que consiste na simulação em pontos seguida da reblockagem por meio da média aritmética dos pontos que pertençam a determinado bloco. Esta abordagem possui duas grandes desvantagens do ponto de vista computacional: o primeiro é que o algoritmo precisa resolver um grande sistema de equações (matrizes), dificultando o manejo dos dados e tornando o processo lento como um todo; e, o segundo é o cálculo da média aritmética dos pontos em si, que pode ser bastante prejudicado, dependendo do número de pontos e blocos que se quer estimar.

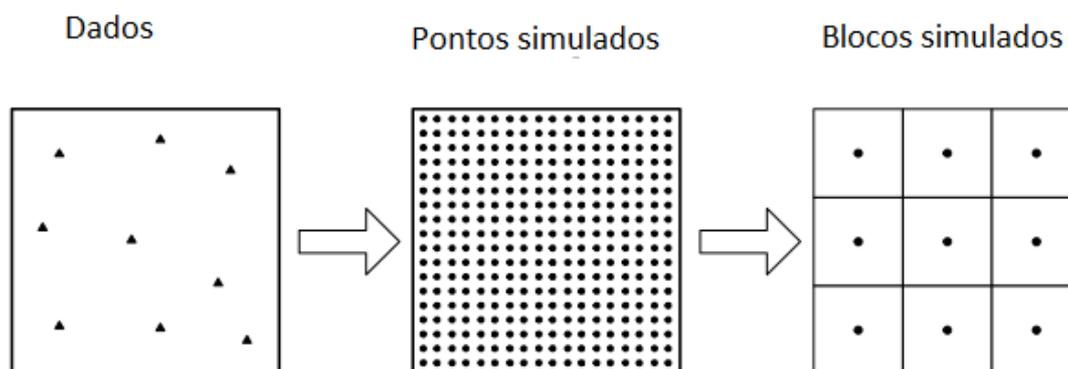


FIGURA 1.1: MÉTODO TRADICIONAL DE SIMULAÇÃO EM BLOCOS (MODIFICADO DE BOUCHER (2003)).

Como alternativa para reduzir a capacidade e tempo computacional necessários no método clássico, Godoy (2002) propôs um algoritmo para simulação direta em blocos por meio da simulação simultânea dos nós pertencentes a um mesmo bloco. Uma vez que a média dos nós tenha sido calculada, os pontos são deletados da memória e somente o valor do teor médio do bloco é mantido. O teor calculado do bloco é então utilizado para condicionar o cálculo do próximo e assim por diante. As grandes vantagens do método são: o cálculo simples para mudança de suporte e a velocidade de processamento dos dados.

Boucher (2003) estendeu o algoritmo de Godoy (2002) para a simulação multivariada de teores estimados via decomposição MAF, o que chamou de DBMAFSIM. O princípio de funcionamento do método é praticamente o mesmo,

com a particularidade do cálculo do teor de diversas variáveis em vez de uma só. Em ambos os casos, os autores puderam concluir que a simulação diretamente em suporte de blocos diminui em cerca de 90% o tempo de processamento computacional.

O método proposto por Boucher (2003) resolve dois principais problemas da simulação multivariada de teores:

- i. A decomposição MAF ortogonaliza os vetores para separações maiores que zero, garantindo que cada variável possa ser estimada independentemente, mantendo a correlação entre elas no retorno ao espaço original e eliminando a utilização de variogramas cruzados;
- ii. A simulação direta em blocos, que garante mais agilidade e menor necessidade de capacidade computacional para o procedimento.

Além do método proposto por Godoy (2002), também existem métodos que fornecem respostas bastante aproximadas por meio dos chamados modelos de mudança de suporte. A grande maioria destes modelos foi proposta por Georges Matheron, e, um resumo delas pode ser encontrado em Chilès & Delfiner (2012). Dentre os diversos métodos existentes, a abordagem via modelo gaussiano discreto é a mais utilizada.

Matheron (1967) foi quem propôs o primeiro método de simulação em blocos com a utilização de um modelo gaussiano discreto. Em 2005, Emery e Ortiz apresentaram uma nova abordagem para a proposta de Matheron, na qual simplificaram o cálculo do coeficiente de correlação entre pontos e blocos, acelerando de forma significativa o processo de simulação de atributos diretamente em suporte de blocos. Ambos os métodos serão abordados em detalhe no capítulo 2.

Contudo, nenhum dos métodos citados anteriormente garante que as estimativas das variáveis tenham soma correta (100% por exemplo), gerando em cada bloco um correto balanço de massa e de teores entre as variáveis estimadas.

#### 1.4. ANÁLISE DE DADOS COMPOSICIONAIS

Dados composicionais são aqueles que carregam informações relativas sobre um todo. Na grande maioria das vezes, os dados composicionais estão restringidos a uma soma fechada constante, por exemplo: a soma dos teores de todos os elementos químicos presentes em uma amostra mineral deverá ser 100%.

As condições de soma fechada fazem com que os dados composicionais apresentem algumas restrições: serão sempre positivos em um intervalo de 0 a 100, ou outra constante; e, por se tratarem de proporções, uma vez que o valor de um dado aumente o do outro irá diminuir obrigatoriamente, mesmo que a correlação entre eles não exista na origem.

Os métodos da estatística clássica foram desenvolvidos para a análise de dados reais, que podem estar em intervalo infinito positivo e negativo, não sendo adequados, portanto, para a interpretação de dados composicionais, rodeados por correlações espúrias. Muitos autores têm apontado problemas na utilização da estatística clássica em dados que fazem parte de uma composição. Pearson (1897) foi o primeiro, seguido por Sarmanov & Vistelius (1959), Krumbein (1962), Chayes(1971), Butler(1979), Aitchison(1986), Davis (1986), Rock(1988), Rollinson(1995) e Aitchison & Egozcue (2005), entre outros.

A análise dos dados composicionais passou por basicamente quatro fases principais. A primeira é pré anos 60 e se desenvolveu justamente na época em que muito se pensava sobre a estatística clássica multivariada, claramente eficiente para a análise e solução dos problemas no espaço real. No entanto, como citado anteriormente, o comportamento de vetores composicionais, onde as partes são as proporções de uma soma constante, são completamente diferentes de vetores não restritos. Mas até 1960, as composições eram abordadas por métodos clássicos sem nenhuma restrição.

Foi somente com o estudo de Chayes (1960), que apontou problemas de fechamento com a utilização de métodos clássicos da geoestatística multivariada para a análise de dados composicionais, que a segunda fase iniciou. Como solução, o autor propôs algumas alterações nos métodos clássicos, sem propor

uma metodologia de abordagem direta às composições, mesmo caso de Sarmanov & Vistelius (1959) e Mosimann (1962).

A terceira fase começa na década de 80, quando Aitchison apresenta a idéia de que as composições fornecem informações relativas, não absolutas, sobre as componentes. Portanto, qualquer análise pode ser feita considerando a razão entre as componentes (Aitchison 1981,1982,1983 e 1984) em vez de seus valores puros, independentes. O fato de as razões logarítmicas serem mais fáceis de lidar matematicamente em relação às razões em si, aliado à possibilidade de utilização da mesma análise estatística multivariada no espaço real e no transformado, favorece a utilização de transformações logarítmicas das relações para o estudo das composições.

A quarta fase inicia quando alguns autores perceberam que as operações de perturbação simples (interna), potência simples (externa) e a métrica simples, definiam um espaço vetorial métrico, ou seja, um espaço de Hilbert (Billheimer *et al.*, 1997; 2001; Pawlowisky-Glan & Egozcue, 2001). Em se tratando de composições, muitos problemas podem ser resolvidos dentro deste espaço vetorial (o Simplex), desde que tratados com sua álgebra e geometria específica. Esta abordagem é conhecida como *permacer-no-Simplex* (Mateu-Figueras, 2003; Pawlowsky-Glahn, 2003). A idéia básica é representar as composições em coordenadas, assim como no espaço Euclidiano, e, interpretar as variáveis e suas relações por meio de suas representações no Simplex.

A terceira fase, que propõe a transformação das relações em razões logarítmicas aditivas, merece atenção especial, já que é uma técnica utilizada há mais de um século. A grande vantagem da proposta está na remoção do espaço vetorial restrito, o Simplex, e, adoção de outro irrestrito, multivariado real, possibilitando a utilização das técnicas clássicas de abordagem aos problemas da estatística multivariada, seguida da retro transformação para o Simplex. As principais transformações utilizadas nesta fase são: transformação em razões logarítmicas aditivas (alr) e transformação em razões logarítmicas centradas (clr) (Aitchison, 1986).

A razão logarítmica isométrica (ilr) (Egozcue *et al.* 2003; Egozcue & Pawlowsky-Glahn, 2005) foi desenvolvida já na quarta fase, para corrigir um problema das metodologias anteriores alr e clr, a não preservação da métrica

dos sistemas quando das operações de mudança de base. Basicamente, a proposta da ilr é a mudança de base dos dados ( $S^D$ ) por meio da utilização de uma matriz de transformação aleatória, para obtenção das razões logarítmicas isométricas no espaço real ( $R^{D-1}$ ). Estando as ilr representadas no espaço real, as técnicas clássicas de ajuste de variograma e estimativa podem ser utilizadas, sendo seguidas da retro transformação das ilr interpoladas através da inversa da matriz de transformação. Esta abordagem é uma aplicação do princípio de “trabalhar em coordenadas” (Pawlowsky-Glahn, 2003), onde interpolações válidas podem ser obtidas. Ou seja, os vetores estimados são positivos, mantêm a condição de soma restrita e a continuidade espacial original.

#### 1.4.1. A ANÁLISE EXPLORATÓRIA DO SIMPLEX

A estatística descritiva padrão não é muito informativa quando nos referimos a composições pertencentes ao Simplex. Em particular, a média aritmética e o desvio padrão ou a variância das componentes individuais não se encaixam na geometria Aitchison como medidas de tendência central e de dispersão. Estas estatísticas foram definidas no âmbito da geometria euclidiana no espaço Real, que não é uma geometria sensível para dados composicionais. Portanto, é necessário introduzir alternativas como os conceitos de centro (Aitchison, 1997), matriz de variação e variância total (Aitchison, 1986).

##### 1.4.1.1. CENTRO

A medida de tendência central para dados composicionais é a média geométrica fechada (g). Para um determinado número n de dados esta medida é chamada de centro e é definida por:

$$g = C [g_1, g_2, \dots, g_D] , \quad (1.4)$$

Sendo que:

$$g_i = \left( \prod_{j=1}^n x_{ij} \right)^{1/n}, \quad i = 1, 2, \dots, D. \quad (1.5)$$

Onde:

C é a constante de fechamento e  $x_{ij}$  são as amostras.

#### 1.4.1.2. MATRIZ DE VARIAÇÃO

A dispersão, em um conjunto de dados composicionais, pode ser descrita tanto pela matriz de variação, originalmente definida por Aitchison (1986):

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1D} \\ t_{21} & t_{22} & \cdots & t_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ t_{D1} & t_{D2} & \cdots & t_{DD} \end{pmatrix}, \quad t_{ij} = \text{var} \left( \ln \frac{x_i}{x_j} \right), \quad (1.6)$$

Ou, pela matriz de variância normalizada:

$$\mathbf{T}^* = \begin{pmatrix} t_{11}^* & t_{12}^* & \cdots & t_{1D}^* \\ t_{21}^* & t_{22}^* & \cdots & t_{2D}^* \\ \vdots & \vdots & \ddots & \vdots \\ t_{D1}^* & t_{D2}^* & \cdots & t_{DD}^* \end{pmatrix}, \quad t_{ij}^* = \text{var} \left( \frac{1}{\sqrt{2}} \ln \frac{x_i}{x_j} \right). \quad (1.7)$$

Onde:

$t_{ij}$  é a razão logarítmica entre as partes  $i$  e  $j$ , enquanto  $t_{ij}^*$  é a razão logarítmica normalizada entre as partes  $i$  e  $j$ .

#### 1.4.1.3. VARIÂNCIA TOTAL

Outra medida importante de dispersão é a variância total, dada por:

$$\text{totvar}[\mathbf{X}] = \frac{1}{2D} \sum_{i=1}^D \sum_{j=1}^D \text{var} \left( \ln \frac{x_i}{x_j} \right) = \frac{1}{2D} \sum_{i=1}^D \sum_{j=1}^D t_{ij} = \frac{1}{D} \sum_{i=1}^D \sum_{j=1}^D t_{ij}^*. \quad (1.8)$$

Por definição  $\mathbf{T}$  e  $\mathbf{T}^*$  são simétricas e sua diagonal contém apenas zeros. Além disso, nem a variância total, nem nenhuma das entradas das matrizes de variância, normalizada ou não, dependem da constante de fechamento associada à dimensão  $D$  do Simplex, já que as constantes são eliminadas nas razões. Assim sendo, pode-se afirmar que o reescalonamento não influencia nestas medidas de dispersão. Além disso, a variância total pode ser definida como a matriz de

variância em uma única posição, já que todas compartilham o denominador, ou seja, a matriz de variância explica como a variância total se distribui entre as partes.

As medidas estatísticas do Simplex são muito relevantes e podem ser observadas em estudos de composições nas mais diversas áreas de conhecimento (e.g. Pawlowski-Glahn & Buccianti, 2011). No entanto, neste estudo, ainda se manteve a análise da qualidade dos resultados em termos da geometria euclidiana.

#### 1.4.2. APLICAÇÕES DAS RAZÕES LOGARÍTMICAS

Uma das críticas para a transformação em razões logarítmicas é que esta abordagem não admite a presença de zeros no banco de dados. No entanto Martín-Fernandez, Barceló-Vidal & Pawlowsky-Glahn (2003) e Pawlowsky-Glahn & Olea (2004) propuseram algumas alternativas para se lidar com estes valores, como por exemplo, a utilização do limite de detecção dos equipamentos utilizados para a análise química das amostras.

Outro problema inicialmente apontado era que as estimativas realizadas produziram resultados inviesados, mas Pawlowsky-Glahn & Egozcue (2002) provaram que em uma escala relativa as interpolações de dados não regionalizados não eram tendenciosas. Já Tolosana-Delgado (2006), estendeu a teoria para o caso de variáveis regionalizadas por meio de um estudo de caso na área de contaminação ambiental, comprovando sua eficiência.

Martín-Fernandez (2001) comparou os resultados da krigagem ordinária dos dados composicionais com a estimativa por cokrigagem das composições em um reservatório de petróleo. Já Tozzo-Martins *et al.*(2009) compararam os resultados da cokrigagem dos dados originais com os da cokrigagem dos dados composicionais em amostras de solo.

Buccianti *et al.*(2006), por meio da *Geological Society*, elaboraram uma publicação especial contendo um conjunto de artigos com aplicações práticas em dados reais, além de uma abordagem teórica sobre os dados composicionais. O objetivo deste trabalho era divulgar e alertar grupos de pesquisa que trabalham com composições, mas abordam os problemas via métodos da estatística

clássica. Neste livro, foram abordados principalmente problemas ligados à importância da manutenção de somas fechadas e correlações entre variáveis geoquímicas, mas um estudo sobre a distribuição de espécies na paleontologia também foi apresentado.

Já em 2011, Pawlowski-Glahn & Buccianti, em uma espécie de homenagem a John Aitchson que completava 85 anos, publicaram uma coletânea de artigos contendo o estado da arte da metodologia até o momento, junto com aplicações da análise de dados composicionais nos mais variados meios de estudo: ecologia, biologia, geoquímica, astronomia, química, economia e geoestatística. O objetivo era novamente disseminar a idéia de que os dados restritos por uma soma constante são partes de uma composição em um espaço restrito (Simplex) e por este motivo, seu comportamento deve ser analisado com a geometria apropriada.

Em se tratando especificamente de aplicações da teoria em depósitos minerais multivariados complexos, Boezio *et al.* (2012) compararam os resultados obtidos na cokrigagem direta dos dados originais de um depósito de ferro, com aqueles resultantes da cokrigagem das razões logarítmicas aditivas (alr). Como conclusão, perceberam que os resultados da cokrigagem dos alr eram melhores do que os da cokrigagem direta dos dados originais, uma vez que não produziam valores negativos e mantinham a condição de soma fechada, sem nenhum tipo de pós processamento para distribuição do erro.

Mueller & Ward (2012) também testaram a utilização de dados composicionais para a estimativa de teores de diversas variáveis em um depósito de ferro, comparando os resultados com aqueles obtidos na cokrigagem clássica. No entanto, analisaram dois tipos de retro transformação das estimativas, a agl (retro-transformação padrão) e uma aproximação por polinômios de Hermite, obtendo resultados muito semelhantes. Além disto, a análise estatística dos resultados foi realizada com a geometria euclidiana (correspondente ao espaço Real) e a geometria de Aitchson (correspondente ao Simplex). Com isto, puderam perceber que a geometria de Aitchson se mostrou mais adequada para a reprodução das restrições (soma constante e resultados positivos) para ambos os casos. Já no espaço Real, as estimativas resultantes da cokrigagem direta foram menos enviesadas do que as obtidas pela cokrigagem dos alrs. Por outro

lado, a distribuição do erro da estimativa dos alrs se mostrou mais próxima à realidade do depósito.

Boezio *et al.* (2012) aplicaram a decomposição MAF às composições de um depósito de ferro com cinco variáveis. A grande vantagem da proposta está no fato de a transformação em razões logarítmicas aditivas implicar na diminuição do número de variáveis a serem estimadas, já que  $S^D \rightarrow R^{D-1}$ . E, a decomposição em fatores de autocorrelação mínimos e máximos (MAF), por ortogonalizar as variáveis, permitiu que cada uma delas fosse estimada de modo independente, evitando o ajuste do MLC. Os resultados foram comparados com os obtidos pela cokrigagem clássica e pela cokrigagem dos alrs e o método proposto se mostrou bastante eficiente e menos laborioso que os demais.

Tolosana Delgado *et al.* (2013) utilizaram a transformação em razões logarítmicas aditivas para a estimativa dos teores em um depósito de ferro. Para evitar que os resultados fossem tendenciosos, propuseram a utilização de simulações de Monte Carlo como alternativa à cokrigagem e puderam concluir que os resultados obtidos foram satisfatórios.

Rubio *et al.* (2016) também utilizaram métodos de simulação geoestatística para a estimativa de teores em um depósito de bauxita do norte do Brasil. Assim como Tolosana Delgado *et al.* (2014) fizeram para um depósito de ferro, os autores propuseram a que a retrotransformação dos valores simulados para os nós da malha de estimativa anteceda o procedimento de reblocagem, evitando portanto o surgimento de resultados enviesados.

Tolosana Delgado *et al.* (2015) mediram a influência da consideração das amostras de um depósito mineral de ferro como composições pertencentes ao Simplex no processo de beneficiamento. Aplicaram para tanto a simulação LU, sobre quatro elementos de interesse e puderam perceber que a utilização das composições se aproxima do comportamento real exibido na rota de beneficiamento de modo mais adequado do que a consideração individual das variáveis.

Tolosana Delgado *et al.* (2016) construíram diagramas de contato entre os possíveis pares de composições para a análise de contato entre diferentes domínios geoestatísticos. Neste estudo puderam mapear a remobilização

ocorrida entre os elementos principais e deletérios em um depósito laterítico de níquel cobalto.

Barnett & Deutsch (2012) apresentaram uma espécie de guia para utilização em sequência das diversas transformações não lineares que tornam as técnicas práticas da geoestatística possíveis em estudo de depósitos multivariados complexos. No estudo, propuseram a utilização em sequência da transformação alr, seguida da decomposição em fatores de autocorrelação mínimos e máximos, para a simulação conjunta dos teores em um depósito de níquel. Mas novamente foram utilizadas técnicas da estatística clássica euclidiana na análise dos resultados.

Mueller *et al.* (2014) propuseram a utilização em sequência da transformação alr e da decomposição MAF para a estimativa dos teores em um depósito laterítico de níquel. Com a utilização da metodologia, os autores puderam perceber que a reprodução das correlações entre as variáveis, assim como a manutenção do fechamento da soma constante de teores estiveram garantidas no processo. No entanto, as simulações foram realizadas em suporte de pontos posteriormente reblocados.

Até o presente momento, não foram encontradas aplicações da combinação MAF+CODA via ilr(análise de composições) para a simulação dos teores em blocos de depósitos multivariados complexos.

### 1.5. PROBLEMA

Depósitos multivariados complexos como os de ferro, bauxita, fosfato e manganês, onde os fechamentos são dados tanto pelos balanços de massa quanto pela estequiometria em cada faixa granulométrica são adequados para o tratamento por técnicas da geoestatística multivariada. No entanto, existem alguns problemas relacionados às técnicas clássicas de estimativa, seja para o cálculo do melhor teor para um dado bloco, seja para a análise do espaço de incerteza sobre este valor, são eles:

- i. As técnicas clássicas não garantem o fechamento das somas e balanços de massas e teores nas estimativas.

- ii. As técnicas clássicas podem gerar resultados negativos para as estimativas dos teores nos locais de interesse.
- iii. Conforme aumenta o número de atributos a serem estimados, cresce também a dificuldade para o ajuste do MLC, motivo pelo qual o modelo utilizado nem sempre representa adequadamente a continuidade espacial direta e cruzada das variáveis.
- iv. A grande maioria dos métodos utilizados não garante a reprodução das correlações cruzadas nos modelos gerados.
- v. A cossimulação clássica dos teores em depósitos multivariados complexos é realizada em pontos em vez de blocos. Sendo o procedimento computacional necessário para a reblockagem, extremamente laborioso.

A análise composicional dos dados (CODA) surgiu como alternativa para solucionar os problemas (i.), (ii.) e (iii.). Considerando que as amostras estejam restritas a um espaço onde as somas sejam sempre constantes e os valores sejam todos positivos (Simplex) com sua álgebra e geometria específica, o método propõe uma mudança de espaço do tipo  $S^D \rightarrow R^{D-1}$  para a realização das estimativas no espaço Real, seguida do retorno ao Simplex. Assim, os valores calculados também teriam que obedecer às condições de restrição e, além disto, o número de variáveis a serem estimadas diminuiria em uma unidade.

Outro benefício do método é a análise do comportamento das relações entre as variáveis, em vez delas propriamente ditas. O que contribui para a manutenção das correlações originais nos modelos estimados.

Já a ortogonalização dos vetores, como por exemplo, a decomposição em fatores de autocorrelação mínimos e máximos (MAF) resolve o problema (iii.). A ideia principal do método é a realização de uma mudança de base dos vetores, para que se tornem ortogonais entre si. Deste modo, estes não estarão mais correlacionados e cada um poderá ser estudado de modo independente. O processo termina com o retorno à base original, onde as correlações devem novamente ser reproduzidas.

A simulação direta em blocos (DBMAFSIM) se apresenta como melhor alternativa para solucionar (item v.). Por produzir resultados diretamente em

suporte de blocos, evita a última etapa da cossimulação, tornando o procedimento mais rápido e eficiente. No entanto, o algoritmo proposto não incorpora a transformação ilr.

Pensando nos fatores apontados foi que Boucher e Dimitrakopoulos (2009; 2012) propuseram a simulação direta em blocos dos fatores MAF provenientes de um depósito de ferro e Deraisme *et al.* (2008) aplicaram a proposta de Emery e Ortiz (2005) no depósito de Chuquicamata, no Chile. Já Boezio *et al.* (2012) propuseram a utilização conjunto das duas metodologias (CODAVia alr+MAF) para a estimativa dos teores em um depósito de ferro. Ward & Muller (2012) e Boezio *et al.* (2012) cokrígaram os alrs também em um depósito multivariado complexo de ferro. Sendo que os primeiros realizaram a análise dos resultados dentro e fora do espaço restrito.

Ainda não foram estudadas metodologias que proponham uma solução acabada aos problemas da simulação de teores de diversas variáveis em depósitos multivariados complexos.

#### 1.6. META

Com base nos problemas apontados, a meta da presente tese é estudar e desenvolver uma solução combinando a simulação direta em blocos dos fatores MAF originados dos ilr, de modo simples e computacionalmente eficiente.

E também, analisar as vantagens e desvantagens do método proposto sobre algumas, também possíveis, combinações de métodos para simulação de teores em depósitos multivariados complexos.

#### 1.7. OBJETIVOS

Para atender à meta proposta será implementada a combinação ILR+MAF+DB (transformação ilr, transformação MAF e diretamente em suporte de blocos). E, será realizada uma análise comparativa dos resultados obtidos por meio da avaliação da coerência e reprodução das características originais do depósito, do fechamento de balanços de teores e de massa, além do grau de simplicidade na utilização para as seguintes combinações de métodos:

- i. Simulação individual dos dados originais pelo método de bandas rotativas, primeiramente em suporte de pontos, posteriormente reblocados. (Combinação I)
- ii. Simulação individual dos ilrs pelo método de bandas rotativas, primeiramente em suporte de pontos, posteriormente reblocados. (Combinação II)
- iii. Simulação direta em blocos dos fatores MAF. (Combinação III)
- iv. ilr+MAF+DB, proposta desta tese, com cinco variáveis. (Combinação IV)
- v. ilr+MAF+DB, proposta desta tese, com quatro variáveis. (Combinação V, sem a variável Resto, ver explicação a seguir.)

Como a transformação ilr requer que os dados de entrada possuam soma igual e constante, para as combinações (i.ii.iii. e iv) utilizou-se uma variável artificial chamada de Resto, criada como complemento para garantir que, em todas as amostras, a soma dos teores fosse 100%.

Mas, como esta variável artificial leva a correlações espúrias e dificulta a reprodução das mesmas nos cenários simulados, além de aumentar o número e variáveis a serem simuladas, a quinta combinação foi realizada com as proporções relativas de cada elemento, onde a soma constante passou a não ser mais 100% mas 1 (um).

A utilização da variável artificial e a das proporções relativas está explicada detalhadamente no capítulo 3, onde o banco de dados utilizado é apresentado.

## 1.8. METODOLOGIA

O roteiro proposto será aplicado na faixa granulométrica 14# em um domínio geoestatístico de um depósito multivariado complexo de bauxita. Serão simulados os teores das variáveis mais importantes:  $Al_2O_3$ ,  $SiO_2$ ,  $Fe_2O_3$ ,  $TiO_2$  e (Resto quando for o caso) dentro desta faixa. A metodologia será executada de acordo com o fluxograma proposto nas figuras 1.2, 1.3, 1.4, 1.5 e 1.6 seguida da análise das vantagens e desvantagens de cada proposta sobre as demais.

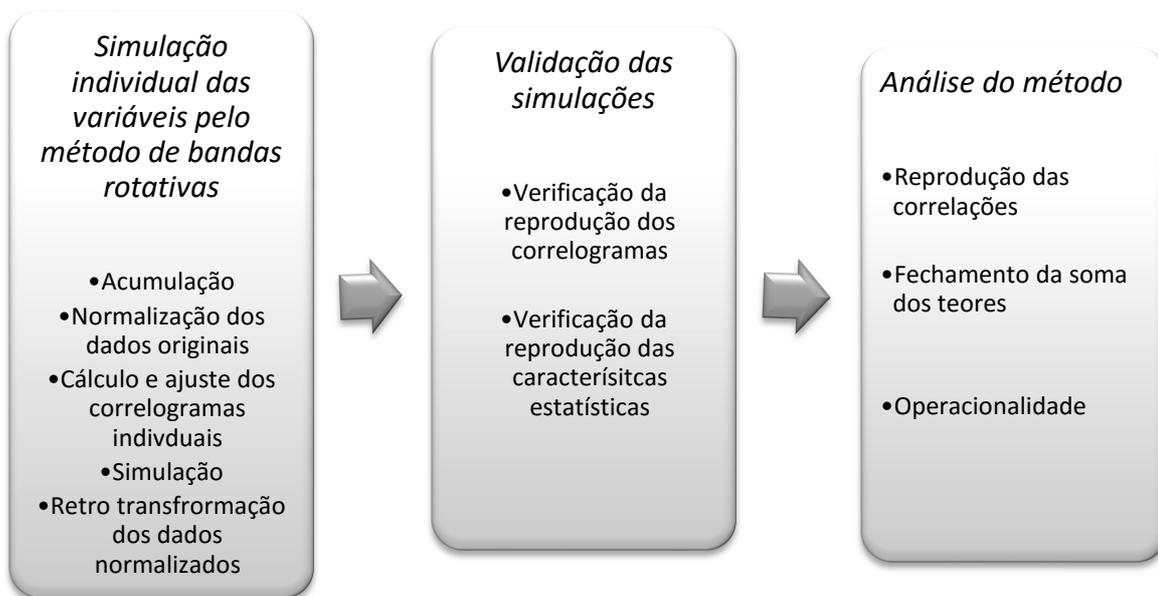


FIGURA 1.2:fluxograma da metodologia da combinação I.

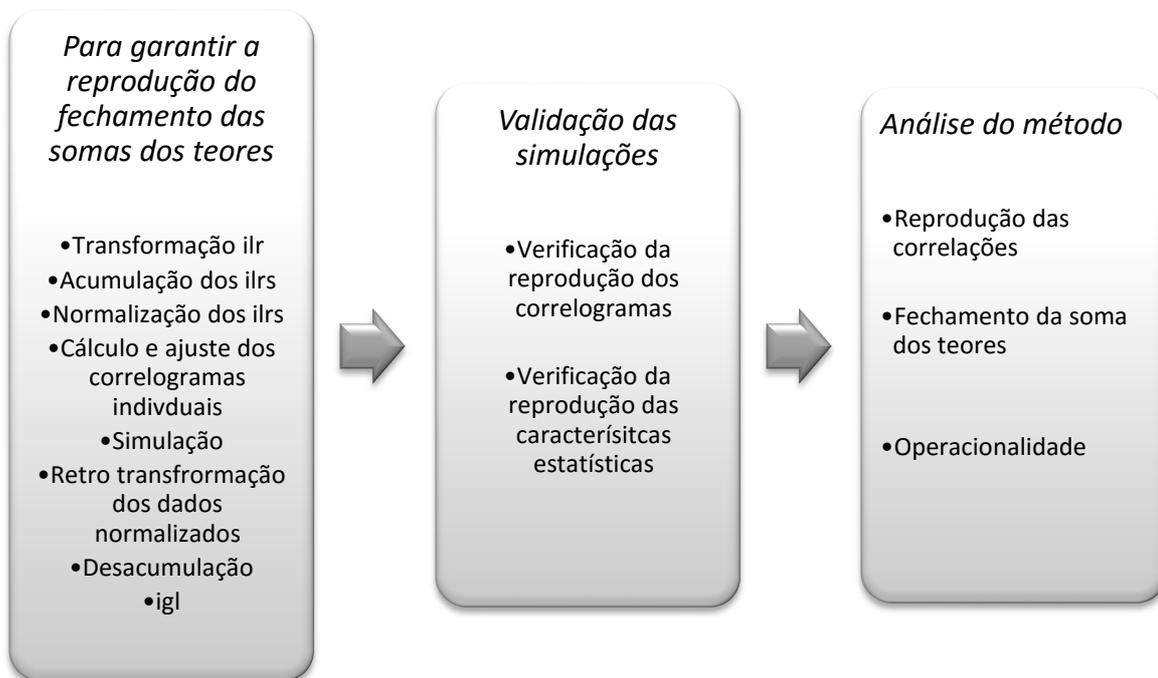


FIGURA 1.3:fluxograma da metodologia para a combinação II.

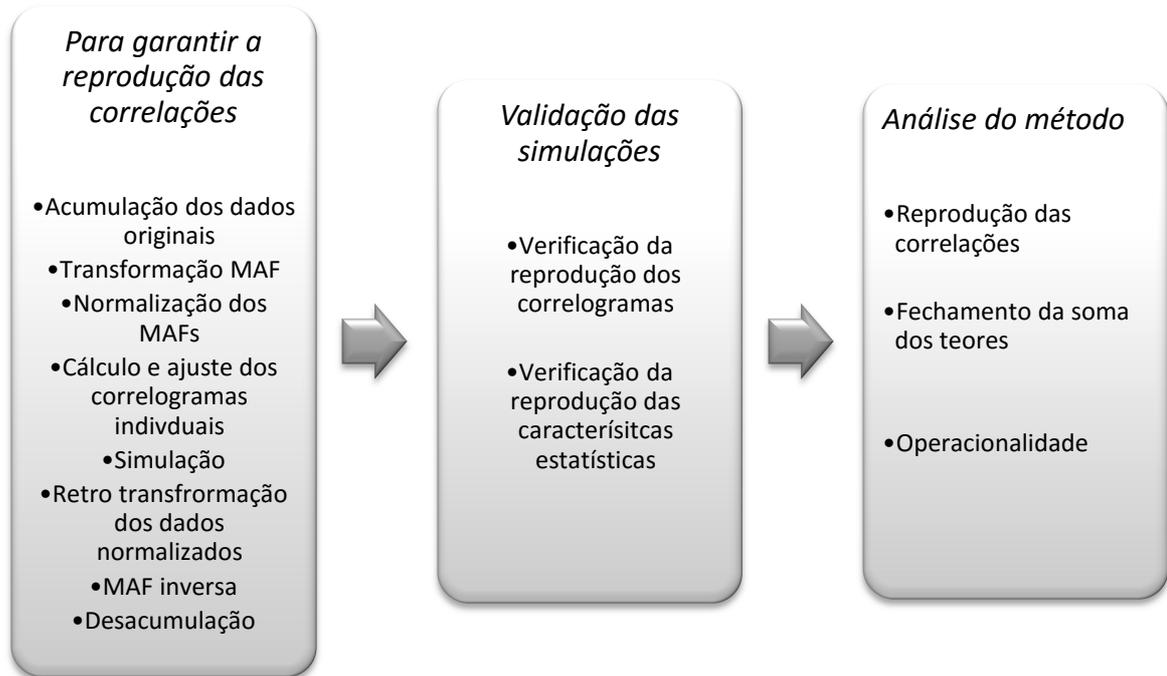


FIGURA 1.4:fluxograma da metodologia para a combinação III.

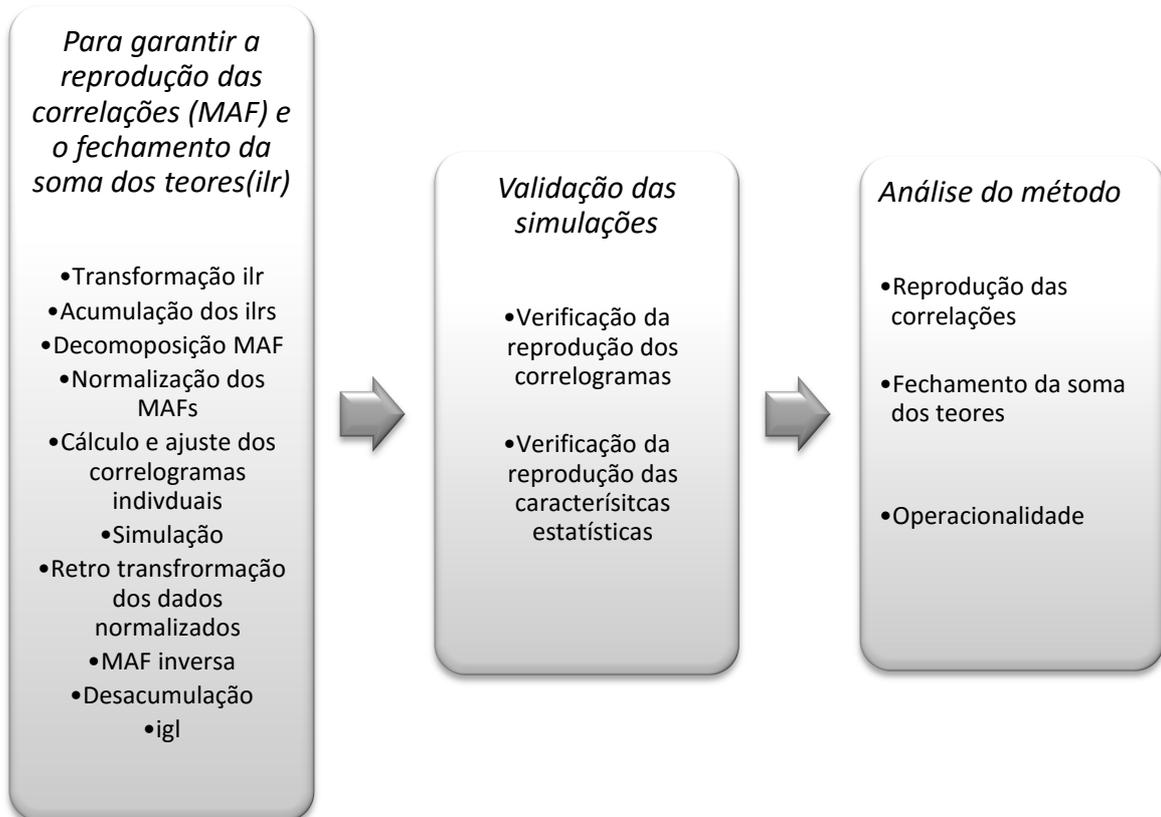


FIGURA 1.5:fluxograma da metodologia para as combinação IV.

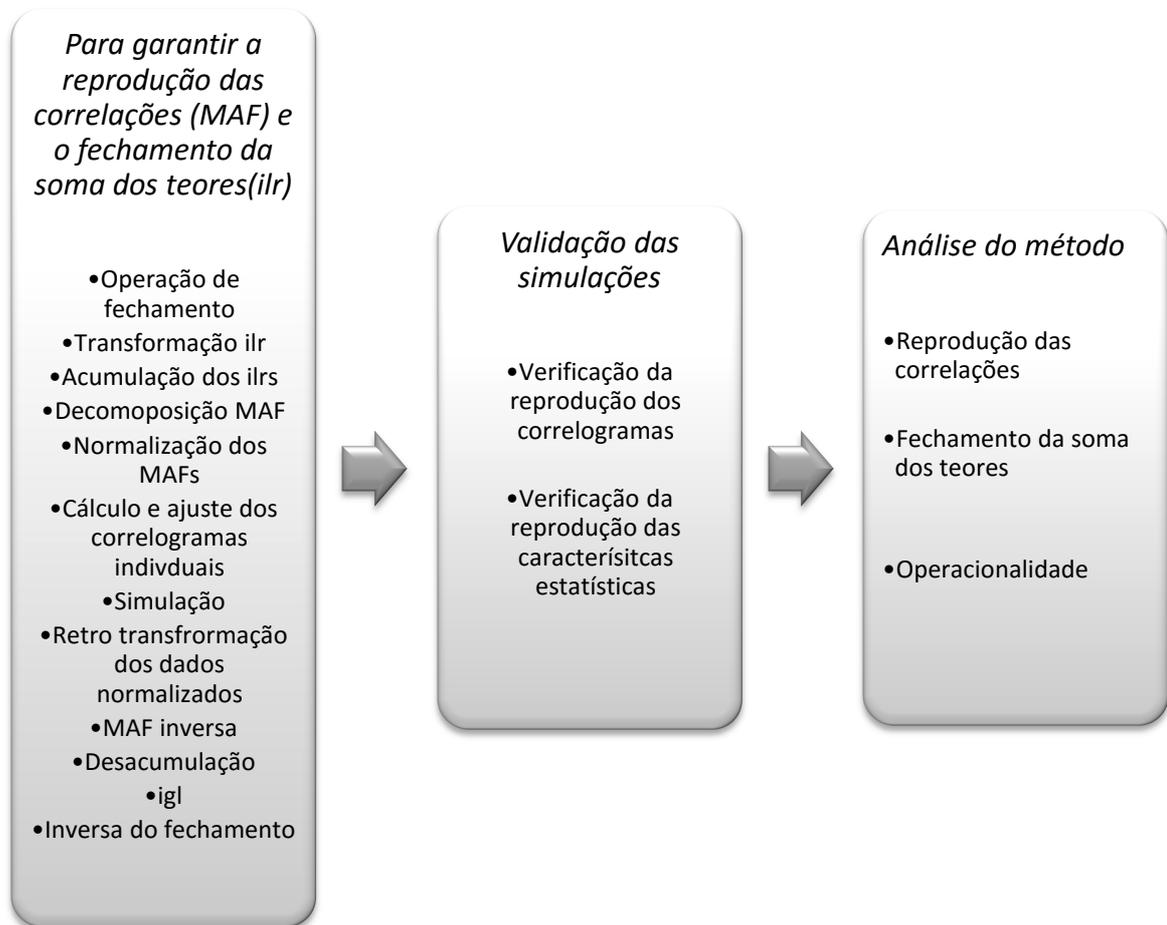


FIGURA 1.6: fluxograma da metodologia para as combinação V.

## 1.9. ORGANIZAÇÃO DA TESE

A tese está dividida em capítulos organizados da seguinte forma:

O capítulo 2 apresenta uma breve revisão dos métodos que serão utilizados para a realização deste estudo: análise de dados composicionais, decomposição em fatores de autocorrelação mínimos e máximos, simulação direta em blocos e simulação por bandas rotativas.

O capítulo 3 deverá introduzir o estudo de caso a ser tomado como base para investigação da aplicabilidade dos métodos abordados. Os aspectos regionais e locais relevantes serão apresentados, bem como as características do

banco de dados utilizado por meio de medidas de estatística descritiva das variáveis consideradas.

No capítulo 4 serão aplicadas as metodologias e analisados os resultados obtidos nas combinações testadas. Avaliando-se as vantagens e desvantagens de cada metodologia aplicada sobre as demais.

O capítulo 5 englobará as conclusões obtidas a partir dos trabalhos realizados, bem como recomendações para trabalhos futuros.

# Capítulo 2

## REVISÃO DOS MÉTODOS APLICADOS

A estimativa de teores e tonelagens de um depósito mineral deve fornecer modelos que reproduzam adequadamente sua realidade. A variabilidade *in situ* destes teores também deve ser propriamente avaliada, partindo de cenários que respeitem as restrições físicas e químicas das variáveis analisadas.

Neste sentido, é imprescindível que os modelos gerados tanto para a estimativa do melhor teor para determinado local, quanto para o estudo de sua variabilidade local e espacial, atendam a três condições principais (em depósitos multivariados com variáveis cujo somatório atinja uma constante C): o balanço de massas e de teores e a reprodução da correlação existente entre os múltiplos elementos.

Esse capítulo apresenta uma breve revisão sobre os métodos estudados para a geração de modelos de variabilidade de teor que respeitem as restrições do depósito e ainda, sejam mais facilmente aplicados do que os métodos tradicionais. São eles: análise de dados composicionais (CODA), decomposição em fatores de autocorrelação mínimos e máximos (MAF) e o algoritmo de simulação direta em blocos desenvolvido por Emery e Ortiz (2005).

O algoritmo de simulação por bandas rotativas também será brevemente revisado, uma vez que será utilizado nas combinações i. e ii..

### 2.1. ANÁLISE DOS DADOS COMPOSICIONAIS (CODA)

John Aitchison (1986) atentou para o fato de que o valor absoluto de um atributo qualquer é irrelevante quando este faz parte de uma composição, já que o interesse está na proporção relativa entre os componentes analisados. Para viabilizar o estudo dos dados composicionais dentro do espaço real, Aitchison propôs dois métodos de transformação para um subespaço euclidiano: as razões logarítmicas aditivas (alr) e as razões logarítmicas centradas (clr). Valendo-se das técnicas clássicas da estatística para a análise dos dados transformados, utilizou a transformação alr na modelagem de fenômenos e a clr para as técnicas

baseadas na preservação da métrica. Isto porque, a transformação alr não preserva as chamadas distâncias entre as composições, enquanto a transformação clr garante a preservação da métrica, mas leva a uma matriz de covariâncias singular, o que limita sua utilização nas técnicas da geoestatística baseadas na krigagem. Em termos matemáticos, podemos dizer que a alr é um isomorfismo, mas não uma isometria, enquanto a clr é uma isometria, logo, um isomorfismo entre o espaço Simplex de D dimensões ( $S^D$ ) e um subespaço Real com D dimensões ( $R^D$ ), o que leva a perda de qualidade das distribuições depois de transformadas. Neste sentido, Aitchison desenvolveu duas metodologias muito interessantes, mas certo cuidado deve ser tomado na aplicação de ambas.

Utilizando a estrutura do espaço vetorial euclidiano uma nova abordagem algébrico-geométrica à proposta de Aitchison se torna viável. Dentro deste contexto, uma transformação de coeficientes é equivalente a expressar as observações em outro sistema de coordenadas. Estamos habituados a trabalhar em um sistema ortogonal, conhecido como sistema de coordenadas cartesianas. Neste sistema, sabemos como alterar as coordenadas e rotacionar eixos. Mas, as transformações alr e clr não podem ser diretamente associadas a um sistema ortogonal de coordenadas no Simplex, fato que levou Egozcue *et al.* (2003) a desenvolverem uma nova transformação, chamada de ilr (razões logarítmicas isométricas), garantindo a isometria entre o  $S^D$  e o  $R^{D-1}$ , evitando portanto, os problemas da alr e da clr. A operação ilr transforma as composições do  $S^D$  em coordenadas do  $R^{D-1}$  e esta será a abordagem apresentada na tese, juntamente com um tipo particular de coordenadas, os balanços, muito úteis na modelagem e interpretação dos dados.

Dentro do estudo proposto, a importância da ilr está em transformar as composições do Simplex de D dimensões em vetores de coordenadas do espaço Real de D-1 dimensões. A transformação permite a utilização das técnicas da geoestatística (que foram pensadas para o espaço euclidiano) garantindo, na retrotransformação, resultados dentro do espaço restrito onde valores negativos não são permitidos e a soma constante das composições está garantida. Como vantagem, a ilr ainda traz a diminuição em uma unidade do número de variáveis a serem estimadas.

### 2.1.1. AS COMPOSIÇÕES NO ESPAÇO REAL

Composições pertencentes ao Simplex de D dimensões são usualmente expressas em termos de uma base canônica  $\{e_1, e_2, \dots, e_D\}$  do  $R^D$ . De fato, qualquer vetor  $\mathbf{x} \in R^D$  pode ser escrito como:

$$x = x_1[1, 0, \dots, 0] + x_2[0, 1, \dots, 0] + \dots + x_D[0, 0, \dots, 1] = \sum_{i=1}^D x_i e_i \quad (2.1)$$

O problema é que o conjunto de vetores  $\{e_1, e_2, \dots, e_D\}$  não é gerador, nem uma base que respeite a estrutura do  $S^D$ . Ou seja, nem todas as combinações de coeficientes resultam em um elemento pertencente ao  $S^D$  onde valores nulos e negativos não são permitidos e por este motivo,  $e_i$  não necessariamente pertence ao Simplex.

### 2.1.2. SISTEMAS GERADORES

O primeiro passo para a obtenção de uma base ortonormal apropriada é encontrar um sistema gerador ( $w_i$ ) que possa ser utilizado para a construção da base (2.2):

$$w_i = C[1, 1, \dots, e, \dots, 1], i = 1, 2, \dots, D \quad (2.2)$$

onde em cada  $w_i$ , o número  $e$  é colocado na  $i$ -ésima coluna e a operação potência ( $\odot$ ) é aplicada em cada componente do vetor. Assim, qualquer vetor  $x$  e  $S^D$  pode ser escrito de acordo com a equação 2.3.

Obs.: A potência ( $\odot$ ) e a perturbação ( $\oplus$ ) são os operadores de multiplicação e soma respectivamente, dentro do espaço Simplex e suas propriedades são basicamente as mesmas da soma (+) e multiplicação (x) dentro do espaço real.

$$\begin{aligned} \mathbf{x} &= \bigoplus_{i=1}^D \ln x_i \odot \mathbf{w}_i = \\ &= \ln x_1 \odot [e, 1, \dots, 1] \oplus \ln x_2 \odot [1, e, \dots, 1] \oplus \dots \oplus \ln x_D \odot [1, 1, \dots, e] \end{aligned} \quad (2.3)$$

Os coeficientes de um sistema gerador não são únicos, então a equação (2.4) pode substituir a (2.3).

$$\begin{aligned} \mathbf{x} &= \bigoplus_{i=1}^D \ln \frac{x_i}{g(\mathbf{x})} \odot \mathbf{w}_i = \\ &= \ln \frac{x_1}{g(\mathbf{x})} \odot [e, 1, \dots, 1] \oplus \dots \oplus \ln \frac{x_D}{g(\mathbf{x})} \odot [1, 1, \dots, e] \end{aligned} \quad (2.4)$$

onde  $g(\mathbf{x})$  é a média geométrica da composição (2.5);

$$g(\mathbf{x}) = \left( \prod_{i=1}^D x_i \right)^{1/D} = \exp \left( \frac{1}{D} \sum_{i=1}^D \ln x_i \right) \quad (2.5)$$

Também na equação 2.4, é possível perceber que os coeficientes equivalem à transformação logarítmica centrada definida por Aitchison (1986). Note que na mesma expressão, o denominador pode ser substituído por qualquer constante, o que está de acordo com o conceito de que as composições são classes de equivalência (Barceló-Vidal *et al.*, 2001).

A transformação clr expressa uma composição em termos de coeficientes logarítmicos centrados (equação 2.6).

$$\text{clr}(\mathbf{x}) = \left[ \ln \frac{x_1}{g(\mathbf{x})}, \ln \frac{x_2}{g(\mathbf{x})}, \dots, \ln \frac{x_D}{g(\mathbf{x})} \right] = \boldsymbol{\xi} \quad (2.6)$$

A transformação inversa, que expressa os coeficientes numa base canônica do espaço Real pode ser escrita de acordo com a equação (2.7).

$$\text{clr}^{-1}(\boldsymbol{\xi}) = \mathcal{C} [\exp(\xi_1), \exp(\xi_2), \dots, \exp(\xi_D)] = \mathbf{x} \quad (2.7)$$

A clr mantém a simetria das componentes, mas para isto, estabelece mais uma restrição aos dados transformados: a soma das componentes deve ser nula. Fato que implica na singularidade da matriz de covariâncias de  $\boldsymbol{\xi}$ , ou seja, seu

determinante é zero. A definição formal da transformação clr pode ser escrita como:

Para uma composição  $x \in S^D$ , os coeficientes da clr são os componentes de  $\xi = [\xi_1, \xi_2, \dots, \xi_D] = \text{clr}(x)$ , o único vetor que satisfaz a equação 2.8:

$$x = \text{clr}^{-1}(\xi) = \mathcal{C}(\exp(\xi)) \quad , \quad \sum_{i=1}^D \xi_i = 0 \quad (2.8)$$

O  $i$ -ésimo coeficiente da clr será:

$$\xi_i = \frac{\ln x_i}{g(x)} \quad (2.9)$$

onde  $g(x)$  é a média geométrica dos componentes de  $x$ .

Embora os coeficientes da clr não sejam coordenadas referentes a uma base do Simplex, eles possuem propriedades bastante importantes. Dentre elas, a mais importante é a manutenção das operações e da métrica do Simplex no Real. Denotando a distância, a norma e o produto interno no  $R^{D-1}$  por:  $d(\cdot, \cdot)$ ,  $\|\cdot\|$ , e  $\langle \cdot, \cdot \rangle$ , respectivamente, as seguintes propriedades se mantêm:

Considerando  $x_k \in S^D$  e  $\alpha$  e  $\beta$ , constantes reais, então:

$$\begin{aligned} \text{clr}(\alpha \odot x_1 \oplus \beta \odot x_2) &= \alpha \cdot \text{clr}(x_1) + \beta \cdot \text{clr}(x_2) ; \\ \langle x_1, x_2 \rangle_a &= \langle \text{clr}(x_1), \text{clr}(x_2) \rangle ; \\ \|x_1\|_a &= \|\text{clr}(x_1)\| \quad , \quad d_a(x_1, x_2) = d(\text{clr}(x_1), \text{clr}(x_2)) . \end{aligned} \quad (2.10)$$

### 2.1.3. COORDENADAS ORTONORMAIS

Omitindo-se um vetor do sistema gerador da equação (2.2), uma base pode ser obtida. Por exemplo, se omitirmos  $w_D$  o resultado será  $\{w_1, w_2, \dots, w_{D-1}\}$ . Esta base não será ortonormal, o que pode ser verificado pelo cálculo do produto interno de qualquer um de seus pares de vetores. Mas, uma nova base ortonormal pode ser obtida pelo método de Gram-Schmidt (Egozcue *et al.*, 2003). A base calculada é apenas uma dentre as infinitas bases ortonormais

possíveis dentro do espaço euclidiano. Por este motivo é importante conhecermos suas características gerais.

Sendo  $\{e_1, e_2, \dots, e_{D-1}\}$  uma base ortonormal genérica do Simplex de  $D$  dimensões, e,  $\boldsymbol{\psi}$  uma matriz  $(D-1, D)$  cujas linhas são os  $\text{clr}(e_i)$ . Uma base ortonormal deverá satisfazer  $\langle e_i, e_j \rangle_a = \delta_{ij}$  ( $\delta_{ij}$  é o delta de Kronecker, sendo nulo para  $i \neq j$ , e unitário quando  $i=j$ ). De acordo com as propriedades descritas em 2.10, podemos escrever:

$$\langle e_i, e_j \rangle_a = \langle \text{clr}(e_i), \text{clr}(e_j) \rangle = \delta_{ij} \quad (2.11)$$

O que implica que a matriz  $\boldsymbol{\psi}$  satisfaça  $\boldsymbol{\psi} \boldsymbol{\psi}' = \mathbf{I}_{D-1}$ , sendo  $\mathbf{I}_{D-1}$  a matriz identidade de  $D-1$  dimensões. Quando o produto entre as matrizes é invertido, chegamos em  $\boldsymbol{\psi}' \boldsymbol{\psi} = \mathbf{I}_D - (\mathbf{1}/D) \mathbf{1}'_D \mathbf{1}_D$ , onde  $\mathbf{D}$  é a matriz identidade de  $D$  dimensões e  $\mathbf{1}_D$  um vetor linha com  $D$  uns (1). As composições das bases são obtidas em  $\boldsymbol{\psi}$  utilizando-se a operação inversa da  $\text{clr}$  ( $\text{clr}^{-1}$ ) em cada linha da matriz.

Uma vez que uma base ortonormal tenha sido escolhida, a composição  $x$  e  $S^D$  pode ser expressa como (2.12):

$$\mathbf{x} = \bigoplus_{i=1}^{D-1} x_i^* \odot \mathbf{e}_i, \quad x_i^* = \langle \mathbf{x}, \mathbf{e}_i \rangle_a \quad (2.12)$$

onde  $x^* = [x_1^*, x_2^*, \dots, x_{D-1}^*]$  é o vetor de coordenadas equivalentes de  $x$  para a base calculada. A função  $\text{ilr}: S^D \rightarrow R^{D-1}$ , com a mudança de coordenadas  $x^*$  para  $x$  é chamada de transformação  $\text{ilr}$  (razões logarítmicas isométricas), sendo uma isometria e um isomorfismo entre os espaços vetoriais. Para simplificação, algumas vezes esta função é chamada de  $h$ , ou seja,  $\text{ilr} = h$ , e, o  $*$  é utilizado para denotar coordenadas quando for conveniente. As propriedades descritas em (2.13) se mantêm.

Considerando  $x_k \in S^D$  e  $\alpha$  e  $\beta$ , constantes reais, então:

$$\begin{aligned} h(\alpha \odot x_1 \oplus \beta \odot x_2) &= \alpha \cdot h(x_1) + \beta \cdot h(x_2) = \alpha \cdot x_1^* + \beta \cdot x_2^* ; \\ \langle x_1, x_2 \rangle_a &= \langle h(x_1), h(x_2) \rangle = \langle x_1^*, x_2^* \rangle ; \\ \|x_1\|_a &= \|h(x_1)\| = \|x_1^*\| \quad , \quad d_a(x_1, x_2) = d(h(x_1), h(x_2)) = d(x_1^*, x_2^*) . \end{aligned} \quad (2.13)$$

A principal diferença entre as propriedades descritas em (2.10) para a clr e em (2.13) para a ilr é que a primeira se refere a vetores de coeficientes pertencentes ao  $R^D$ , enquanto a segunda trata de vetores de coordenadas do  $R^{D-1}$ , correspondente a real dimensão do  $S^D$ .

Considerando as propriedades descritas em (2.10) e (2.13) e, utilizando a matriz imagem clr da base,  $\Psi$ , as coordenadas de uma composição  $x$  podem ser escritas de forma compacta. De acordo com (2.12), uma determinada coordenada se origina do produto interno de Aitchison e pode ser expressa como um produto interno entre os coeficientes da clr. Agrupando-se todas as coordenadas em um único vetor se obtém uma matriz produto simples (2.14).

$$x^* = \text{ilr}(x) = h(x) = \text{clr}(x) \cdot \Psi' \quad (2.14)$$

A inversão da ilr, ou seja, o retorno para composições pode ser escrito de acordo com (2.12). De fato, tomando-se os coeficientes clr em ambos os lados da equação (2.12) e considerando as propriedades descritas em (2.8) é possível obter (2.15):

$$\text{clr}(x) = x^* \Psi \quad , \quad x = \mathcal{C}(\exp(x^* \Psi)) \quad (2.15)$$

Um algoritmo eficiente para a recuperação de  $x$  de suas coordenadas equivalentes  $x^*$  pode ser descrito pelos seguintes passos:

- i. Construção da matriz clr da base,  $\Psi$ ;
- ii. Cálculo da matriz produto  $x^* \Psi$ ;
- iii. Aplicação de  $\text{clr}^{-1}$  para a obtenção de  $x$ .

Existem diversos métodos que podem ser utilizados para a obtenção de uma base ortonormal no Simplex. O critério principal para a seleção de uma base

ortonormal é a garantia de qualidade na interpretação dos dados em coordenadas. Por exemplo, na análise das componentes principais uma base ortogonal é selecionada de modo que a primeira coordenada (componente principal) represente a direção de maior variabilidade. Um caso particular a ser analisado é a obtenção de uma base seguindo-se uma partição binária sequencial do vetor da composição (Egozcue & Pawlowsky-Glahn, 2005). A grande vantagem da utilização destas bases é que são facilmente interpretadas em termos de grupos de partes de uma composição. As coordenadas cartesianas em uma destas bases são chamadas de *balanços* e as composições da base são chamadas de *elementos de balanço*.

A partição binária sequencial nada mais é do que uma espécie de hierarquia entre as partes de uma composição. Na primeira posição, todas as partes são divididas em dois grupos. Nos passos seguintes, os grupos da primeira posição são novamente divididos em dois grupos, e assim por diante até que todos os grupos possuam uma parte única, como ilustrado na Tabela (2.1). Para cada ordem de partição, é possível definir o balanço entre dois subgrupos formados em um nível: sendo  $i_1, i_2, \dots, i_r$  as  $r$  partes do primeiro subgrupo (que receberam o código +1) e  $j_1, j_2, \dots, j_s$  as  $s$  partes do segundo (que receberam o código -1). O balanço pode ser definido como a razão logarítmica normalizada da média geométrica de cada grupo de partes (eq. 2.16).

TABELA 2.1: Exemplo de uma matriz de sinais, utilizada para codificar uma sequência de partição binária e construção de uma base ortonormal. A parte inferior da tabela mostra a matriz  $\psi$ , da base equivalente. (Modificado de Pawlowsky-Glahn et al. (2013)).

ordem	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	r	s
1	+1	+1	-1	-1	+1	+1	4	2
2	+1	-1	0	0	-1	-1	1	3
3	0	+1	0	0	-1	-1	1	2
4	0	0	0	0	+1	-1	1	1
5	0	0	+1	-1	0	0	1	1

ordem	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
1	$\frac{1}{4}\sqrt{\frac{4\cdot 2}{4+2}}$	$\frac{1}{4}\sqrt{\frac{4\cdot 2}{4+2}}$	$-\frac{1}{2}\sqrt{\frac{4\cdot 2}{4+2}}$	$-\frac{1}{2}\sqrt{\frac{4\cdot 2}{4+2}}$	$\frac{1}{4}\sqrt{\frac{4\cdot 2}{4+2}}$	$\frac{1}{4}\sqrt{\frac{4\cdot 2}{4+2}}$
2	$+\frac{\sqrt{3}}{2}$	$-\frac{1}{\sqrt{12}}$	0	0	$-\frac{1}{\sqrt{12}}$	$-\frac{1}{\sqrt{12}}$
3	0	$+\frac{\sqrt{2}}{\sqrt{3}}$	0	0	$-\frac{1}{\sqrt{6}}$	$-\frac{1}{\sqrt{6}}$
4	0	0	0	0	$+\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$
5	0	0	$+\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$	0	0

$$b = \sqrt{\frac{rs}{r+s}} \ln \frac{(x_{i_1} x_{i_2} \cdots x_{i_r})^{1/r}}{(x_{j_1} x_{j_2} \cdots x_{j_s})^{1/s}} = \ln \frac{(x_{i_1} x_{i_2} \cdots x_{i_r})^{a_+}}{(x_{j_1} x_{j_2} \cdots x_{j_s})^{a_-}} \quad (2.16)$$

onde:

$$a_+ = +\frac{1}{r} \sqrt{\frac{rs}{r+s}}, \quad a_- = -\frac{1}{s} \sqrt{\frac{rs}{r+s}}$$

$$a_0 = 0 \quad (2.17)$$

$a_+$  para as partes do numerador,  $a_-$  para as partes do denominador e  $a_0$  para as partes não envolvidas neste nível. O balanço será (2.18):

$$b_i = \sum_{j=1}^D a_{ij} \ln x_j \quad (2.18)$$

onde  $a_{ij}$  será substituído por  $a_+$ , se o código na  $i$ -ésima ordem de partição for +1, para a  $j$ -ésima parte; será substituído por  $a_-$  se o código for -1 e  $a_0 = 0$  se o código for nulo, utilizando-se os valores de  $r$  e  $s$  na  $i$ -ésima ordem de partição. A matriz com as entradas  $a_{ij}$  será a matriz  $\Psi$  e pode ser observada na parte inferior da Tabela 2.1.

Como a obtenção da base ortonormal de transformação,  $\Psi$ , não é trivial, um exemplo elucidativo será apresentado a seguir.

Em Egozcue *et al.* (2003) uma base ortonormal do Simplex foi obtida por meio do método de Gram-Schmidt. Ela corresponde à partição binária sequencial mostrada na Tabela 2.2.

TABELA 2.2: Exemplo de uma matriz de sinais para  $D=5$ , utilizada para codificar a partição binária sequencial de modo padrão. A parte inferior da tabela mostra a matriz  $\psi$  da base.

level	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	r	s
1	+1	+1	+1	+1	-1	4	1
2	+1	+1	+1	-1	0	3	1
3	+1	+1	-1	0	0	2	1
4	+1	-1	0	0	0	1	1

1	$+\frac{1}{\sqrt{20}}$	$+\frac{1}{\sqrt{20}}$	$+\frac{1}{\sqrt{20}}$	$+\frac{1}{\sqrt{20}}$	$-\frac{2}{\sqrt{5}}$		
2	$+\frac{1}{\sqrt{12}}$	$+\frac{1}{\sqrt{12}}$	$+\frac{1}{\sqrt{12}}$	$-\frac{\sqrt{3}}{\sqrt{4}}$	0		
3	$+\frac{1}{\sqrt{6}}$	$+\frac{1}{\sqrt{6}}$	$-\frac{\sqrt{2}}{\sqrt{3}}$	0	0		
4	$+\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$	0	0	0		

As entradas da matriz  $\Psi$  podem ser facilmente expressas por ( $D$  é o número de dimensões):

$$\Psi_{ij} = a_{ji} = +\sqrt{\frac{1}{(D-i)(D-i+1)}}, \quad j \leq D-i,$$

$$\Psi_{ij} = a_{ji} = -\sqrt{\frac{D-i}{D-i+1}}, \quad j = D-i+1;$$

$$\Psi_{ij} = 0 \tag{2.19}$$

A interpretação dos balanços é feita com base em suas próprias propriedades. A primeira propriedade vem da expressão matemática dos balanços em si, especialmente quando a média geométrica é utilizada no numerador e no denominador da fórmula (2.20).

$$b = \sqrt{\frac{rs}{r+s}} \ln \frac{(x_1 \cdots x_r)^{1/r}}{(x_{r+1} \cdots x_D)^{1/s}} \tag{2.20}$$

As médias geométricas são os valores centrais das partes em cada grupo de partes; sua razão mede o peso relativo de cada grupo; a utilização do logaritmo provê a escala adequada e a raiz quadrada dos coeficientes é uma constante normalizada que permite a comparação entre diferentes balanços. Um balanço positivo significa que, na média geométrica, o grupo de partes do numerador possui peso maior na composição do que o grupo de partes do denominador (e o contrário para os balanços negativos).

O segundo elemento interpretativo está relacionado com a ideia intuitiva de balanço. Imagine que, em uma eleição, os partidos sejam divididos em dois grupos, os de direita e os de esquerda, cada um com diversos partidos adeptos. Se após as eleições, um jornal publicasse somente as proporções de voto dentro de cada grupo, sem informar sobre as proporções entre - grupos não seria possível conhecer o resultado final das eleições. A informação de que partido ganhou só será completa com a apresentação do *balanço* entre os dois grupos. De modo mais preciso, assumo que existam seis partidos no total e as

composições dos votos sejam  $x \in S^6$ ; assumamos também que quatro partidos sejam de esquerda  $\{x_1, x_2, x_5, x_6\}$  e dois de direita  $\{x_3, x_4\}$ ; considere a partição binária sequencial da Tabela 2.1. A primeira partição separa as duas frentes e por isto os balanços nos informam sobre o equilíbrio entre as duas partes. Se este balanço for desconsiderado, os balanços seguintes irão nos informar somente sobre o grupo da esquerda (balanços 3 e 4) e somente sobre o grupo da direita (balanço 5). Então, ao reter somente a informação relativa ao quinto balanço, só teremos informações relativas ao comportamento interno da subcomposição dos partidos de direita. Do mesmo modo, os balanços 2,3 e 4 informam somente o que ocorre dentro do grupo de esquerda. A conclusão a que se chega é que o balanço 1 não nos informa sobre as relações entre os partidos pertencentes a cada grupo, nos apresenta somente a informação sobre o balanço entre os dois grupos.

Muitas questões podem ser facilmente respondidas com a utilização dos balanços. Por exemplo, suponhamos que estejamos interessados somente nas informações sobre os partidos pertencentes ao grupo da esquerda, e, conseqüentemente podemos remover as informações sobre o grupo da direita. Seguindo o modo tradicional, removeríamos as partes  $x_3$  e  $x_4$  e fecharíamos a subcomposição restante. Isto seria equivalente a projetar uma composição de seis partes ortogonalmente no subespaço associado ao grupo de esquerda, o que seria facilmente realizável assumindo-se  $b_5=0$ . Se fizéssemos isto, a projeção da composição obtida seria (2.21):

$$x_{proj} = C[x_1, x_2, g(x_3, x_4), g(x_3, x_4), x_5, x_6], \quad g(x_3, x_4) = (x_3 x_4)^{1/2} \quad (2.21)$$

onde cada parte pertencente ao grupo da direita é substituída por sua média geométrica interna. Esta composição ainda incluiria o primeiro balanço, entre os partidos de direita e esquerda,  $b_1$ . Se estivéssemos interessados em removê-la ( $b_1=0$ ), a informação restante seria somente sobre o comportamento interno dos partidos de esquerda, ou seja, seria a projeção ortogonal da subcomposição dos partidos de esquerda (2.22).

$$x_{esq} = C[x_1, x_2, g(x_1, x_2, x_5, x_6), g(x_1, x_2, x_5, x_6), x_5, x_6] \quad (2.22)$$

onde,

$$g(x_1, x_2, x_5, x_6) = (x_1, x_2, x_5, x_6)^{1/4} \quad (2.23)$$

A conclusão é que os balanços podem ser muito úteis para projetar composições em subespaços especiais, bastando para isto que alguns balanços sejam retidos e outros anulados.

#### 2.1.4. TRABALHANDO EM COORDENADAS

As coordenadas referentes a uma base ortonormal pertencente a um espaço vetorial linear no Simplex obedecem às mesmas regras das operações no espaço Real. Como consequência, a perturbação ( $\oplus$ ) no  $S^D$  é equivalente à translação ou soma (+) no espaço real, e, a operação potência ( $\odot$ ) no  $S^D$  equivale à multiplicação ( $\cdot$ ). Então, se considerarmos o vetor de coordenadas  $h(x) = x^* \in R^{D-1}$ , transformado a partir do vetor de composições  $x \in S^D$  com uma base ortonormal aleatória, temos que:

$$h(x \oplus y) = h(x) + h(y) = x^* + y^* , \quad h(\alpha \odot x) = \alpha \cdot h(x) = \alpha \cdot x^* \quad (2.24)$$

Também,

$$d_a(x, y) = d(h(x), h(y)) = d(x^*, y^*) \quad (2.25)$$

onde  $d$  é a chamada distância no espaço euclidiano. Isto significa que, na análise dos dados composicionais, os resultados que podem ser obtidos com a utilização das composições e da geometria de Aitchison são exatamente os mesmos que os obtidos com a utilização das coordenadas das composições juntamente com a tradicional geometria euclidiana. A dupla possibilidade de representação das composições, no Simplex e por coordenadas do Real, introduz uma gama de possibilidades para a análise de seu comportamento (figuras ilustrativas 2.1, 2.2, 2.3 e 2.4). O alerta é que a base de transformação deve ser cuidadosamente escolhida para uma correta interpretação do comportamento das coordenadas.

A transformação para coordenadas também pode ser feita "às cegas", método chamado de "caixa preta", dependendo do objetivo da transformação. Neste método, uma base aleatória é selecionada para a transformação e os resultados obtidos em coordenadas são retro transformados para o Simplex para interpretação. Esta estratégia cega, embora aceitável, impossibilita certas análises que podem ser relevantes para o estudo, mas novamente, a escolha entre utilizar uma base cega ou não, depende do objetivo do trabalho. Para

aplicações na área da geoestatística a utilização da base cega é permitida, desde que a mesma base seja utilizada na retro transformação dos resultados.

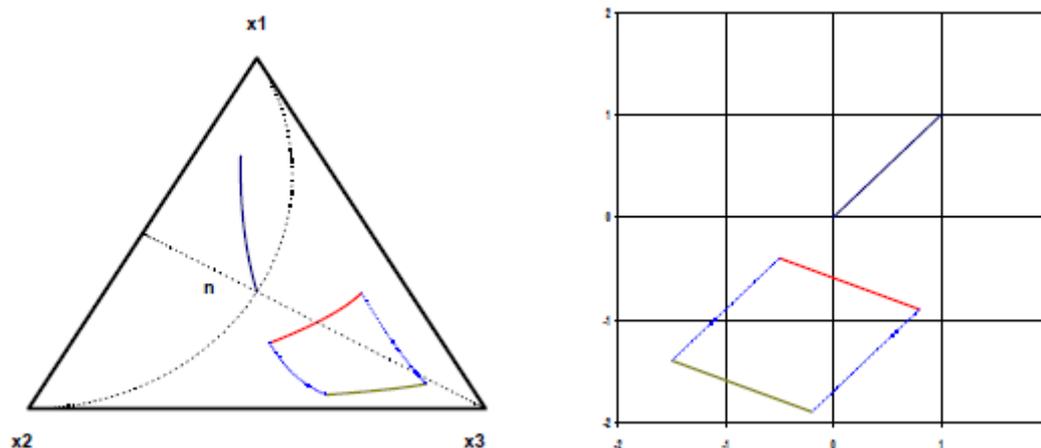


FIGURA 2.1: Representação gráfica da perturbação de um segmento no  $S^3$  (esquerda) e em coordenadas (direita) (Pawlowsky-Glahn *et al.*, 2013).

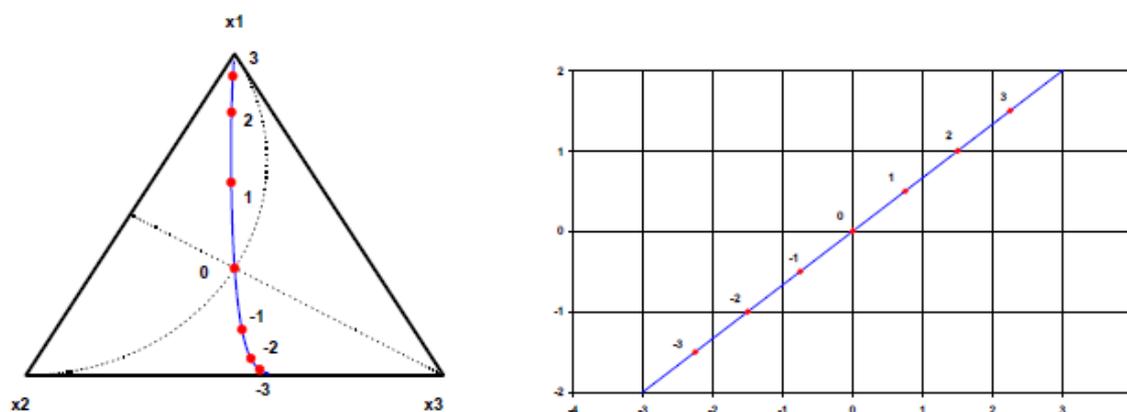


FIGURA 2.2: Representação gráfica da potência de um vetor no  $S^3$  (esquerda) e em coordenadas (direita) (Pawlowsky-Glahn *et al.*, 2013).

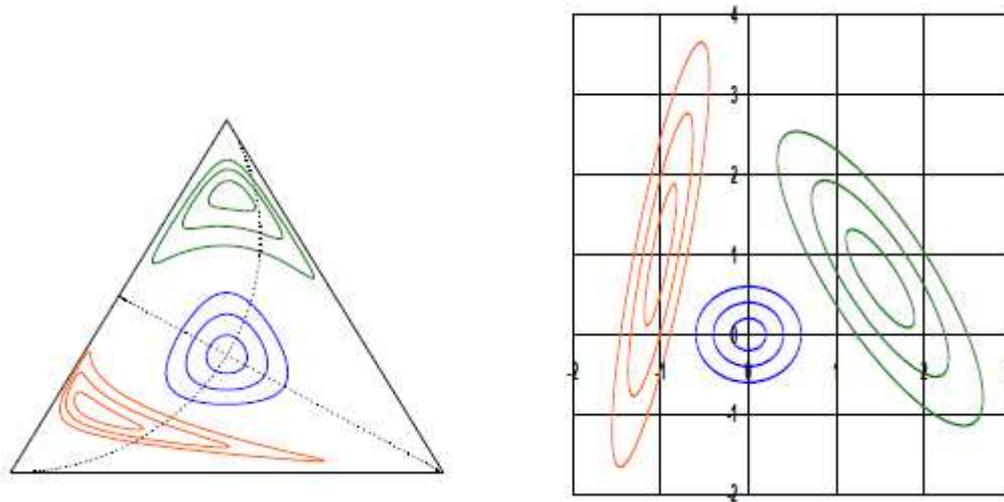


FIGURA 2.3: Representação gráfica de círculos e elipses no  $S^3$  (esquerda) e em coordenadas (direita) (Pawlowsky-Glahn *et al.*, 2013).

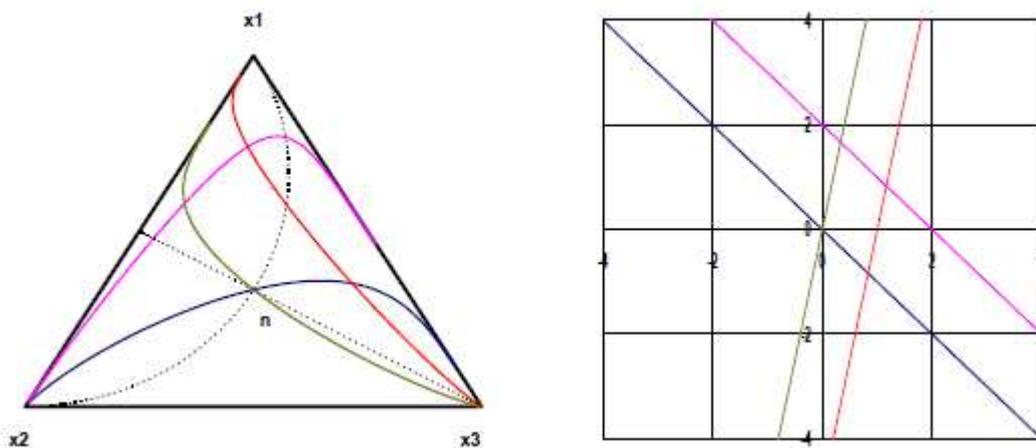


FIGURA 2.4: Representação gráfica de pares de linhas paralelas no  $S^3$  (esquerda) e em coordenadas (direita) (Pawlowsky-Glahn *et al.*, 2013).

É essencial salientar, que nesta abordagem, não são permitidos valores nulos, já que uma divisão por zero não é admissível e nem a aplicação do logaritmo é permitida. Este item não será discutido nesta tese uma vez que o banco de dados utilizado para o estudo não apresenta valores nulos. Métodos de como lidar com este problema podem ser encontrados em (Aitchison, 1986; Aitchison & Kay, 2003; Bacon-Shone, 2003; Fry *et al.*, 1996; Martín-Fernandez, 2001; Martín-Fernandez *et al.*, 2000;2003).

## 2.2. DECOMPOSIÇÃO EM FATORES DE AUTOCORRELAÇÃO MÍNIMOS E MÁXIMOS (MAF)

Os fatores de autocorrelação mínimos e máximos (MAF) foram desenvolvidos por Switzer & Green (1984) para filtrar o ruído nas imagens obtidas por sensoriamento remoto. A proposta original era separar o ruído do sinal, em imagens multi-espectrais, quando os parâmetros dos ruídos fossem desconhecidos, assumindo-se que os parâmetros dos sinais sejam mais autocorrelacionados que os do ruído. Para os casos em que ambos o ruído e o sinal podem ser modelados por meio de modelos de correionalização intrínsecos, fica demonstrado que o método produz fatores ortogonais para todos os vetores de separação  $h$  (a demonstração está no item 2.2.1).

Neste estudo, a importância da decomposição MAF está em descorrelacionar as coordenadas obtidas na transformação ilr (item 2.1). A descorrelação irá permitir que o espaço de incerteza de cada um dos atributos possa ser estimado de modo independente, o que resolve os problemas relacionados a modelagem do MLC, velocidade de processamento e reprodução das correlações entre os dados originais, existentes nos métodos clássicos de cossimulação.

Como demonstrado em Desbarats & Dimitrakopoulos (2000), a idéia é transformar vetores de observações multivariadas do tipo  $\mathbf{Z}(u) = \{Z_1(u), \dots, Z_p(u)\}^T$  utilizando um conjunto  $p$  de combinações lineares ortogonais.

$$Y_i(u) = \mathbf{a}_i^T \mathbf{Z}(u), \quad i = 1, \dots, p \quad (2.26)$$

Cada transformação  $Y_i(u)$  exibe uma correlação espacial maior do que a transformação  $Y_j(u)$  previamente calculada sob a restrição da ortogonalidade.

Sendo  $\rho_i(h)$  a correlação espacial entre  $Y_i(u)$  e  $Y_i(u+h)$  para um *lag* suficientemente pequeno, os coeficientes  $\mathbf{a} = (a_1, \dots, a_p)^T$  são tais que:

$$\begin{aligned}\rho_1 &= \text{corr}(\mathbf{a}_1^T \mathbf{Z}(u), \mathbf{a}_1^T \mathbf{Z}(u+h)) = \min_a \text{corr}(\mathbf{a}^T \mathbf{Z}(u), \mathbf{a}^T \mathbf{Z}(u+h)) \\ \rho_i &= \text{corr}(\mathbf{a}_i^T \mathbf{Z}(u), \mathbf{a}_i^T \mathbf{Z}(u+h)) = \min_a \text{corr}(\mathbf{a}^T \mathbf{Z}(u), \mathbf{a}^T \mathbf{Z}(u+h)) \\ \rho_p &= \text{corr}(\mathbf{a}_p^T \mathbf{Z}(u), \mathbf{a}_p^T \mathbf{Z}(u+h)) = \max_a \text{corr}(\mathbf{a}^T \mathbf{Z}(u), \mathbf{a}^T \mathbf{Z}(u+h))\end{aligned}\quad (2.27)$$

restritos por:

$$\text{corr}(\mathbf{a}_i^T \mathbf{Z}(u), \mathbf{a}_j^T \mathbf{Z}(u)) = 0 \text{ for } j < i \quad (2.28)$$

### 2.2.1. DESCORRELAÇÃO COM MÍNIMOS E MÁXIMOS

Switzer & Green (1984) e posteriormente Desbarats & Dimitrakopoulos (2000), mostraram como obter os fatores por meio dos autovetores da matriz  $2\Gamma_Z(h)\mathbf{B}^{-1}$ , de acordo com:

$$\begin{aligned}\mathbf{B} &= \text{cov}[\mathbf{Z}(u), \mathbf{Z}(u)] \\ 2\Gamma_Z(h) &= \text{cov}[(\mathbf{Z}(u) - \mathbf{Z}(u+h)), (\mathbf{Z}(u) - \mathbf{Z}(u+h))]\end{aligned}\quad (2.29)$$

onde  $\mathbf{B}$  é a matriz de variância/covariância de  $Z(u)$  e  $\Gamma_Z(h)$  é a matriz de variogramas para o *lag*  $h$ , e, ambas são assumidas como invertíveis.

Sendo  $\mathbf{Z}(u) = \{Z_1(u), \dots, Z_p(u)\}$  uma função estacionária vetorial aleatória de  $p$  dimensões, tal que:

$$\mathbf{Z}(u) = \mathbf{S}(u) + \mathbf{N}(u) \quad (2.30)$$

onde  $\mathbf{S}(u)$  e  $\mathbf{N}(u)$  são as componentes de sinal e ruído descorrelacionadas. Se assumirmos um modelo de correionalização intrínseco para  $\mathbf{S}(u)$  e  $\mathbf{N}(u)$ , conseqüentemente  $\mathbf{Z}(u)$  será representado por um modelo linear de correionalização onde  $\rho_0(h)$  e  $\rho_1(h)$  são funções de correlação espacial tais que  $\rho_0(h) < \rho_1(h)$  para todo  $h$  maior que zero. O que implica em:

$$\begin{aligned}
\text{cov}[\mathbf{N}(u), \mathbf{N}(u+h)] &= \mathbf{C}_N(h) = \mathbf{B}_0 \rho_0(h) \\
\text{cov}[\mathbf{S}(u), \mathbf{S}(u+h)] &= \mathbf{C}_S(h) = \mathbf{B}_1 \rho_1(h) \\
\text{cov}[\mathbf{Z}(u), \mathbf{Z}(u+h)] &= \mathbf{C}_Z(h) = \mathbf{B}_0 \rho_0 + \mathbf{B}_1 \rho_1(h)
\end{aligned} \tag{2.31}$$

Deste modo, as matrizes de variância/covariância ( $\rho(0) = 1$ ) de  $\mathbf{N}(u)$ ,  $\mathbf{S}(u)$  e  $\mathbf{Z}(u)$  são:

$$\begin{aligned}
\text{cov}[\mathbf{N}(u), \mathbf{N}(u)] &= \mathbf{C}_N(0) = \mathbf{B}_0 \\
\text{cov}[\mathbf{S}(u), \mathbf{S}(u)] &= \mathbf{C}_S(0) = \mathbf{B}_1 \\
\text{cov}[\mathbf{Z}(u), \mathbf{Z}(u)] &= \mathbf{C}_Z(0) = \mathbf{B}_0 + \mathbf{B}_1 = \mathbf{B}
\end{aligned} \tag{2.32}$$

A matriz de variogramas  $\Gamma_Z(h)$  para  $hs$  de curta distância é:

$$\begin{aligned}
2\Gamma_Z(h) &= \text{cov}[(\mathbf{Z}(u) - \mathbf{Z}(u+h), (\mathbf{Z}(u) - \mathbf{Z}(u+h),))] \\
&= 2(\mathbf{B} - \mathbf{B}_0 \rho_0(h) - \mathbf{B}_1 \rho_1(h)) = 2(1 - \rho_1(h))\mathbf{B} + 2(\rho_1(h) - \rho_0(h))\mathbf{B}_0
\end{aligned} \tag{2.33}$$

Multiplicando  $\Gamma_Z(h)$  por  $\mathbf{B}^{-1}$  chega-se a (2.34):

$$2\Gamma_Z(h)\mathbf{B}^{-1} = 2(1 - \rho_1(h))\mathbf{I} + 2(\rho_1(h) - \rho_0(h))\mathbf{B}_0\mathbf{B}^{-1} \tag{2.34}$$

O produto obtido pode ser expresso em função de  $\mathbf{A}^T$ , tal como seu conjunto de autovetores e autovalores ( $\mathbf{\Lambda}$ ) associados. A multiplicação da equação (2.34) por  $\mathbf{A}^T$  leva a:

$$\mathbf{A}^T \Gamma_Z(h) \mathbf{B}^{-1} = (1 - \rho_1(h))\mathbf{A}^T + (\rho_1(h) - \rho_0(h))\mathbf{A}^T \mathbf{B}_0 \mathbf{B}^{-1} = \frac{\Lambda_h}{2} \mathbf{A}^T \tag{2.35}$$

Rearranjando os termos:

$$\mathbf{A}^T \mathbf{B}_0 \mathbf{B}^{-1} = \frac{1}{(\rho_1(h) - \rho_0(h))} \left[ \frac{\Lambda_h}{2} - (1 - \rho_1(h))\mathbf{I} \right] \mathbf{A}^T = \mathbf{\Lambda}^* \mathbf{A}^T \tag{2.36}$$

A equação (2.36) estabelece que os autovetores de  $\Gamma_z(h)\mathbf{B}^{-1}$  são os mesmos que os da matriz independente  $\mathbf{B}^0\mathbf{B}^{-1}$ , garantindo a ortogonalização para todos os vetores de separação  $h$ . Na equação (2.36) também está demonstrado o atendimento à condição  $\rho_0(h) < \rho_1(h)$ , necessária para a obtenção de autovalores positivos.

### 2.2.2. APLICANDO EM $Z(u)$

Considerando  $\mathbf{Y}(u) = \mathbf{A}^T\mathbf{Z}(u)$  e que a esperança matemática dos valores de  $\mathbf{Z}(u)$  seja nula ( $E[\mathbf{Z}(u)] = 0$ ), então a variância de  $\mathbf{Y}(u)$  pode ser expressa por

$$E[\mathbf{Y}(u)\mathbf{Y}(u+h)] = \rho_1(h)\mathbf{I} - (\rho_1(h) - \rho_0(h))\mathbf{A}^T\mathbf{B}_0\mathbf{A} = \mathbf{I} - \frac{\Lambda_h}{2} \quad (2.37)$$

E a covariância para o lag  $h$  (2.38):

$$\text{cov}[\mathbf{Y}(u), \mathbf{Y}(u)] = E[\mathbf{A}^T\mathbf{Z}(u)^2\mathbf{A}] = \mathbf{A}^TE[\mathbf{Z}(u)^2]\mathbf{A} = \mathbf{A}^T\mathbf{B}\mathbf{A} = \mathbf{I} \quad (2.38)$$

produzindo uma matriz diagonalizável para todos os  $h$ s.

### 2.2.3. A RELAÇÃO DA DECOMPOSIÇÃO MAF COM A DECOMPOSIÇÃO EM COMPONENTES PRINCIPAIS (PCA)

Com o objetivo de evitar a decomposição espectral da matriz não simétrica, Switzer & Green(1984) definiram a transformação MAF em termos das composições principais derivadas da decomposição espectral de uma matriz simétrica. A transformação MAF é equivalente a duas transformações PCA consecutivas na matriz de variância/covariância, a primeira dos dados originais e a segunda dos dados fatorizados para  $h=0$ .

Sendo  $\mathbf{V}(u) = \mathbf{W}^T\mathbf{Z}(u)$  as componentes principais de  $Z(u)$ , onde  $\mathbf{W}^T = \mathbf{H}\mathbf{D}^{1/2}$  e  $\mathbf{H}$  é a matriz de autovetores originada na decomposição espectral da matriz de variância/covariância tal que  $\mathbf{B} = \mathbf{H}\mathbf{D}\mathbf{H}^T$ . A matriz de variogramas para um lag de separação ( $h$ ),  $\Gamma_v(h)$ , da função vetorial aleatória  $\mathbf{V}(u)$ , pode ser expressa por:

$$\begin{aligned} \Gamma_v(h) &= \text{cov}[\mathbf{V}(u) - \mathbf{V}(u+h), \mathbf{V}(u) - \mathbf{V}(u+h)] \\ &= 2(1 - \rho_1(h))\mathbf{I} + 2(\rho_1(h) - \rho_0(h))\mathbf{W}^T\mathbf{B}_0\mathbf{W} \end{aligned} \quad (2.39)$$

e, a decomposição espectral de  $\Gamma_v(h)$  é:

$$\Gamma_v(h) = \mathbf{C} \frac{\Lambda_h}{2} \mathbf{C}^T \quad (2.40)$$

Combinando as equações (2.39) e (2.40), a matriz de variogramas de  $\mathbf{Y}(u) = \mathbf{C}^T \mathbf{V}(u) = \mathbf{C}^T \mathbf{W}^T \mathbf{Z}(u)$  pode ser escrita como:

$$(1 - \rho_1(h)) \mathbf{I} + (\rho_1(h) - \rho_0(h)) \mathbf{C}^T \mathbf{W}^T \mathbf{B}_0 \mathbf{W} \mathbf{C} = \frac{\Lambda_h}{2} \quad (2.41)$$

Finalmente, substituindo  $\mathbf{A} = \mathbf{W} \mathbf{C}$ , a equação (2.41) se torna idêntica à equação (2.37) descrevendo a covariância para duas estruturas.

De forma resumida, a decomposição MAF pode ser operacionalizada em cinco passos.

- i. Realizar a decomposição espectral de  $\mathbf{B}$ , tal que  $\mathbf{B} = \mathbf{H} \mathbf{D} \mathbf{H}^T$ .
- ii. Rotacionar a variável  $\mathbf{V}(u) = \mathbf{W}^T \mathbf{Z}(u)$ , onde  $\mathbf{W}$  é tal que  $\mathbf{W}^T \mathbf{B} \mathbf{W} = \mathbf{I}$ .
- iii. Calcular a matriz de variogramas  $\Gamma_v(h)$  de  $\mathbf{V}(u)$ .
- iv. Realizar a decomposição espectral de  $\Gamma_v(h)$  de modo que ela equivalha a  $\mathbf{C} \mathbf{\Lambda} \mathbf{C}^T$ .
- v. Rotacionar  $\mathbf{Z}(u)$ , de acordo com os vetores  $\mathbf{A} = \mathbf{W} \mathbf{C}$ .

Informações adicionais podem ser encontradas em (Switzer & Green, 1984; Berman, 1985; Green *et al.*, 1988; Wackernagel *et al.*, 1989; Desbarats & Dimitrakopoulos, 2000).

### 2.3. SIMULAÇÃO DIRETA EM BLOCOS

Na indústria mineira, normalmente, as informações sobre a distribuição de teores de um depósito mineral estão disponíveis em um suporte amostral significativamente menor do que aquele em que a decisão precisa ser tomada. Ou seja, dispõe-se de um banco de dados composto basicamente por amostras de furo de sondagem ao mesmo tempo em que o planejamento e o sequenciamento de lavra precisam ser pensados em um suporte maior, os blocos, cujos tamanhos variam de acordo com, por exemplo: tamanho de equipamentos, geotécnica e produção.

Sob o olhar das teorias sobre funções randômicas, podemos afirmar que os resultados da distribuição de probabilidade no suporte de lavra são dependentes da distribuição espacial geral dos valores de interesse. Esta distribuição pode ser gerada por meio de diversos métodos de simulação, mas, a grande maioria requer alta capacidade computacional para ser executada. É este o principal motivo que leva à utilização de métodos que fornecem uma solução aproximada por meio dos chamados modelos de mudança de suporte.

### 2.3.1. O MODELO GAUSSIANO DISCRETO

Sendo  $Z(x)$  uma função randômica estacionária que pode ser expressa como a transformação de outra função randômica estacionária  $Y(x)$  com distribuição marginal normal padrão. Então, podemos afirmar que  $Z(x) = \Phi(Y(x))$ , com a função de transformação  $\Phi$  dada por:  $\Phi = F^{-1} \circ G$ , onde  $F$  é a c.d.f. marginal de  $Z$ , e,  $G$  é a c.d.f. normal padrão. Da mesma forma, podemos considerar que o teor médio  $Z(v)$ , de um bloco  $v$  qualquer, pode ser dado por  $Z(v) = \Phi_v(Y_v)$ , onde  $Y_v$  é uma variável randômica normal padrão e  $\Phi_v$  é a função de transformação que queremos conhecer.

#### 2.3.1.1. O MÉTODO CLÁSSICO

Considere um ponto qualquer  $x$  pertencente a um bloco  $v$ , e, que  $F$  é a c.d.f. de  $Z(x)$  para todos os  $x$ . Portanto,  $Z(x)$  pode ser expressa como a transformação  $\Phi(Y(x))$  da variável aleatória  $Y(x)$ .

Matheron (1976) assumiu que a distribuição bivariada  $(Y(x), Y_v)$  é gaussiana e possui um coeficiente de correlação positivo  $r$ , concluindo portanto que a função de transformação para blocos  $(\Phi_v)$  pode ser dada por:

$$\Phi_v(y_v) = \int \Phi(r y_v + \sqrt{1-r^2} u) g(u) du \quad (2.42)$$

onde  $u$  é a posição da média centrada dos blocos.

A equação 2.42 também define a distribuição de  $Z(v)$ .

Na prática,  $\Phi$  é expresso em função da expansão dos polinômios de Hermite:

$$\Phi(y) = \sum_{n=0}^{\infty} \Phi_n X_n(y) \quad (2.43)$$

onde os  $X_n$  são os polinômios de Hermite normalizados e os coeficientes  $\Phi_n$  são dados por:

$$\Phi_n = \int \Phi(y) X_n(y) G(dy) \quad (2.44)$$

Seguindo a equação 2.42 podemos expressar  $\Phi_v$  como:

$$\Phi_v(y) = \sum_{n=0}^{\infty} \Phi_n r^n X_n(y) \quad (2.45)$$

Assim, a variância de  $Z(v)$  pode ser calculada por meio dos coeficientes  $\Phi_n r^n$  originados da expansão de  $\Phi_v$  ou então da função de covariância  $C(h)$  de  $Z$ .

O coeficiente  $r$  de mudança de suporte será obtido então por:

$$\sum_{n=1}^{\infty} \Phi_n^2 r^{2n} = \frac{1}{|v|^2} \int_v \int_v C(x'-x) dx dx' \quad (2.46)$$

Onde  $x$  e  $x'$  são dois pontos aleatórios pertencentes ao bloco  $v$ .

Para isto, é necessário assumir que, em (2.46),  $r$  possui uma única solução possível pertencente ao intervalo de 0 a 1.

### 2.3.1.2. A ADAPTAÇÃO DO MÉTODO CLÁSSICO

A simplificação do método clássico proposta por Emery e Ortiz (2005), pode ser mais facilmente aplicada porque nela a distribuição bivariada entre dois pontos quaisquer,  $Y(x)$  e  $Y(x')$ , pertencentes a um bloco  $v$ , é assumida como gaussiana. Com base neste fato, o autor mostra que  $r^2$  é a variância da média de  $Y(v)$ .

$$r^2 = \frac{1}{|v^2|} \int_v \int_v C(x' - x) dx dx' \quad (2.47)$$

Onde  $C(h)$  é a covariância (aqui o correlograma) de  $Y$  e  $Y_v$  é a média  $Y(v)$  reescalada para a variância unitária pelo coeficiente de mudança de suporte:

$$Y_v = Y(v) / r \quad (2.48)$$

Isto traz uma grande simplificação para a estimativa local dos teores, principalmente para a abordagem multivariada gaussiana.

Chilès (2012) realizou uma análise de ambos os métodos (clássico e simplificado) para saber quais seriam as situações onde a utilização de cada um deles é recomendada. O autor concluiu que o modelo proposto por Matheron (1976) fornece uma boa aproximação da verdadeira distribuição de teores dos blocos, a não ser nos casos em que existem valores extremos, o que faz com que o logaritmo do desvio padrão seja muito alto. Concluiu também que, além de o método simplificado carregar consigo esta mesma limitação, os blocos a serem estimados não podem possuir, em nenhum dos eixos principais, dimensões muito próximas às do alcance do variograma.

Com base no estudo de Chilès (2012), pôde-se concluir que a aproximação do método clássico proposta por Emery e Ortiz (2005) é capaz de gerar resultados suficientemente precisos para o estudo de caso em questão.

## 2.4. SIMULAÇÃO POR BANDAS ROTATIVAS

Os cenários simulados pelo método de bandas rotativas são resultado da soma de um grande número de simulações independentes definidas em linhas de varredura no plano (Chilès e Delfiner, 1999). Segundo Journel (1974), o processo de simulação pode ser separado em duas etapas, simulação não condicional e pós-condicionamento. Na primeira, os valores atribuídos aos nós da malha simulada não são condicionados pelas amostras, apesar de reproduzirem seu modelo de covariância. Na segunda, o condicionamento é feito sobre os valores já simulados.

A simulação por bandas rotativas é um dos métodos de simulação multidimensional aleatória mais antigos, proposto por Matheron (1973) e Journel (1974). No entanto, até a década de 1990 foi pouco utilizado em aplicações geoestatísticas, devido a limitação computacional da época. As restrições são referentes ao número de bandas, aos problemas com a suavização da variância próxima ao ponto simulado e a não utilização de facilidades, como a retro-transformação dos valores simulados.

Porém, alguns algoritmos de computação mais recentes permitem a utilização do processo com considerada confiabilidade. A simulação por bandas rotativas, devido ao fato de não ser diretamente condicionada na primeira etapa, é um processo bem rápido quando comparado com outros algoritmos, o que favorece sua aplicação em problemas geológicos-mineiros, com grandes bancos de dados e modelos de blocos.

A realização da simulação em uma dimensão pode ser feita de muitas formas (Journel, 1974), desde que exista uma função aleatória com uma covariância conhecida e uma média igual a zero. Na teoria, o método para gerar a função aleatória é livre. Alguns métodos podem ser mais complexos e utilizarem transformações de Fourier. Porém, dentre as possibilidades, o mais empregado na prática é o de médias móveis, que utiliza uma função peso aleatória, com valores mais facilmente encontrados com relação aos outros métodos conhecidos. No programa Isatis<sup>®</sup> (Geovariances Inc.), utilizado para gerar os resultados deste estudo, o método de geração da função aleatória depende do modelo de covariância utilizado:

- quando o modelo de covariância for Gaussiano, Carinal-Sine, J-Bessel ou Couchy, a função aleatória é gerada pelo Método Espectral, que usa transformações de Fourier de uma distribuição positiva para gerar distribuição de covariância:

$$Y(x) = \sqrt{2} \cos (\langle \Omega, x \rangle + \Phi) \quad (2.49)$$

onde:  $\Omega$  é um vetor aleatório com a distribuição de probabilidade  $X$ ;  
 $\Phi$  é uma variável uniforme entre 0 e  $2\pi$ .

- para modelos de covariância Esféricos ou Cúbicos, é utilizado o Método da Diluição, que gera uma função numérica  $F$  e partições em intervalos com comprimento constante. Cada intervalo é aleatoriamente avaliado com  $F$  ou  $-F$ . A função aleatória é obtida como a diluição das funções primárias:

$$Y(x) = \sum_{u \in P} \varepsilon(u) g(x-u) \quad (2.50)$$

onde:  $P$  é um processo de Poisson de intensidade determinada por uma covariância geométrica;

$\varepsilon$  é uma família de variáveis aleatórias padrão;

$g$  é uma função numérica.

- o Método da Migração é utilizado para gerar a função aleatória, para o caso do modelo de covariância exponencial, onde a simulação da covariância é obtida por uma soma com projeção das simulações em um dado número de linhas da covariância.

Na primeira etapa, o valor a ser simulado em um ponto  $u_n (x,y)$  pertencente a um plano qualquer, é a soma de valores oriundos de projeções do próprio ponto em cada linha de uma simulação unidirecional. Conforme a equação 2.51, a soma das projeções ( $u_{1,n}$ ) do ponto  $u_1$ , nas  $n$  linhas que dividem o plano, multiplicadas pelo valor simulado ( $Z_{1,n}$ ) ao longo da linha  $n$ , compõe o valor simulado no ponto em questão ( $Z^*(u_1)$ ).

$$Z^*(u_1) = \frac{1}{\sqrt{n}} \cdot \sum_{n=1}^N Z_{1,n} \cdot (u_{1,n}) \quad (2.51)$$

A covariância dada na direção  $D_n$  (ao longo da linha  $n$ ) contribuirá para a obtenção do valor simulado no ponto. Este procedimento é realizado para várias bandas orientadas por vetores diferentemente distribuídos no espaço. As

diversas projeções do ponto  $Z^*(x,y,z)$  nas direções  $D_i$ , ponderadas por suas covariâncias, compõem o valor de  $Z^*(x,y,z)$ .

Segundo David (1977), a prática permite afirmar que 15 bandas são suficientes para uma boa simulação em 2D. Chilès (1977), recomenda 180 bandas, uma por cada grau do ângulo, enquanto Gneiting (1999) diz que 64 linhas garantem uma ótima reprodução da covariância. Devido ao grande avanço tecnológico dos processadores dos computadores nas últimas décadas, o número de bandas simuladas já não traz grandes problemas de tempo destinado às simulações. Sendo assim, mesmo para simulações em duas dimensões, um número elevado de bandas (da ordem de centenas) pode ser utilizado sem grandes demandas de tempo.

A Equação 2.51 pode ser usada para casos tridimensionais, com a seguinte equação, que relaciona o modelo de covariância de uma dimensão para  $n$  dimensões de forma isotrópica:

$$C(h) = \int C_1(h,u) \cdot w_s(du) \quad (2.52)$$

onde  $C_1(h,u)$  é a covariância da projeção do vetor  $h$  no vetor  $u$ ; e  $w_s$  é a distribuição de probabilidade para o vetor  $s$ .

As covariâncias são tomadas por meio dos covariogramas (ou variogramas) inseridos no processo. A equação 2.52 considera o modelo de covariâncias para  $n$  dimensões como sendo isotrópico, porém, há casos onde há anisotropia de covariâncias para as diferentes direções. Segundo Journel e Huijbregts (1978), o modelo de covariância anisotrópica é considerado uma soma das covariâncias isotrópicas de cada direção, a Equação 2.53 mostra a aplicação para o caso em três dimensões.

$$C(h_u, h_v, h_w) = K_0 C_0(r) + K_1 C_1(h_w) + K_2 C_2(\sqrt{h_u^2 + h_v^2}) + K_3 C_3(r) \quad (2.53)$$

onde  $K_n$  são constantes positivas,  $C_n$  são covariâncias isotrópicas definidas nas três direções, podendo ser, por exemplo, modelos esféricos com alcances e contribuições de covariâncias definidos.

Na segunda etapa, os dados são condicionados aos dados normalizados (Journel e Huijbregts, 1978). Para isso, é utilizada a Equação 2.51, gerando dados condicionados nos locais de interesse onde haja dados normalizados.

Segundo Lantuéjoul (2002), o condicionamento dos valores simulados é feito é feito da seguinte forma:

$$Z^*_c(u_1) = Z(u_1)^k + [Z^*(u_1) - Z^*(u_1)^k] \quad (2.54)$$

onde  $Z^*_c(u_1)$  é o valor simulado em um ponto  $u_1$ , condicionado aos dados normais;  $Z(u_1)^k$  é o valor obtido no local  $u_1$  por krigagem dos dados normalizados;  $Z^*(u_1)$  é a simulação não condicional para o local  $u_1$  proveniente da primeira etapa; e  $Z^*(u_1)^k$  é o valor obtido no local  $u_1$  por krigagem dos valores simulados (não-condicionais) nos locais dos dados normais.

## 2.5. COMENTÁRIOS

O presente capítulo abordou os três métodos que serão utilizados de forma conjunta nesta tese, a transformação ilr que garante estimativas positivas e a soma constante dos teores em um bloco; a simulação direta em blocos, cuja vantagem está na aceleração do processo de simulação para obtenção dos diferentes cenários de distribuição de teores do depósito; e, a decomposição MAF, que traz consigo o benefício de evitar o cálculo e ajuste de variogramas cruzados, por descorrelacionar as variáveis e permitir que cada uma seja simulada independentemente das demais.

A simulação por bandas rotativas também foi resumidamente explicada já que será utilizada para a simulação individual dos dados (combinação i.) e para a simulação dos ilrs (combinação ii.).

No terceiro capítulo, o banco de dados utilizado neste estudo será apresentado, juntamente com a análise da estatística básica de cada variável e entre as variáveis de interesse.

# Capítulo 3

## O BANCO DE DADOS

A metodologia proposta será aplicada em um depósito de bauxita localizado no Brasil. Informações sobre a localização do empreendimento e sobre a empresa responsável pela mina são sigilosas e por este motivo não fazem parte do presente capítulo que abordará: a geologia do depósito, a estatística básica de cada variável, possíveis correlações entre as variáveis e a análise da continuidade espacial dos atributos.

### 3.1. GENERALIDADES

A bauxita foi descoberta em 1821 por Berthier, na localidade de Les Baux, no sul da França (por isto recebeu o nome de bauxita). Normalmente, tem coloração avermelhada e é rica em alumínio, contendo mais de 40% de alumina ( $\text{Al}_2\text{O}_3$ ). Na verdade, é a proporção dos óxidos de ferro que determina a coloração da rocha: a bauxita branca contém de 2 a 4% de óxidos de ferro, enquanto que na bauxita vermelha essa proporção atinge 25%. A bauxita é a fonte natural do alumínio e mesmo com sua elevada abundância, não há notícias acerca da ocorrência de alumínio metálico na natureza. Constata-se sua maior ocorrência na forma combinada com outros elementos, principalmente o oxigênio, com o qual forma a alumina.

A rocha é composta por uma mistura impura de minerais de alumínio, sendo que os mais importantes são: gibbsita  $\text{Al}(\text{OH})_3$ , diásporo  $\text{AlO}(\text{OH})$  e boehmita  $\text{AlO}(\text{OH})$ . Estes minerais são conhecidos como oxi-hidróxidos de alumínio e suas proporções na rocha variam muito entre os depósitos, assim como o tipo e a quantidade das impurezas do minério, tais como: óxidos de ferro, argila, sílica e dióxido de titânio. A maioria dos depósitos economicamente viáveis possui um conteúdo de alumina ( $\text{Al}_2\text{O}_3$ ) entre 50 e 55% e o teor mínimo para que ela seja aproveitável é da ordem de 30%.

Até a segunda metade do século XIX, quase toda a bauxita era produzida na França e empregada basicamente para fins não metalúrgicos. Naquela época,

a produção de alumina destinava-se principalmente ao uso como mordente na indústria têxtil (utilizado para fixação da cor nos tecidos). No entanto, com o desenvolvimento do processo Hall-Héroult (1886), a alumina disponível foi, de modo crescente, sendo utilizada para a produção de alumínio metálico. Mesmo assim, com o passar dos anos surgiram diversas aplicações para a bauxita não metalúrgica: abrasivos, refratários, produtos químicos, cimentos de alta alumina e próteses humanas. Tornou-se evidente que matérias-primas com alta alumina e baixo teor de álcalis têm vantagens especiais, restando apenas a busca para uma solução do problema custo/benefício devido principalmente ao beneficiamento do minério.

### 3.2. MINERALOGIA E GEOLOGIA

O depósito estudado ocorre na forma de diversos platôs espalhados em uma área de aproximadamente 2.200km<sup>2</sup>. De origem laterítica, o conjunto de platôs é fortemente cortado, com altitudes variando de 160 a 190 metros. Seguindo um mergulho muito suave (1° a 5°), os platôs possuem encostas em formato convexo que podem ter uma inclinação de até 30°.

A região é caracterizada principalmente por uma sedimentação clástica de origem continental com espessuras superiores a 600m e constituída por: arenito argiloso e ardósia intercalados com conglomerados, argilas e silte.

#### 3.2.1. PERFIL GEOLÓGICO

O perfil do depósito pode ser considerado homogêneo ao longo de toda a sua extensão, apresentando praticamente a mesma sequência de horizontes (do topo para a base): camada caulínica superior, camada de bauxita granular, camada ferruginosa granular, camada de bauxita (com espessura variando entre 6 e 10m), camada caulínica inferior e sedimento basal (figura 3.1).

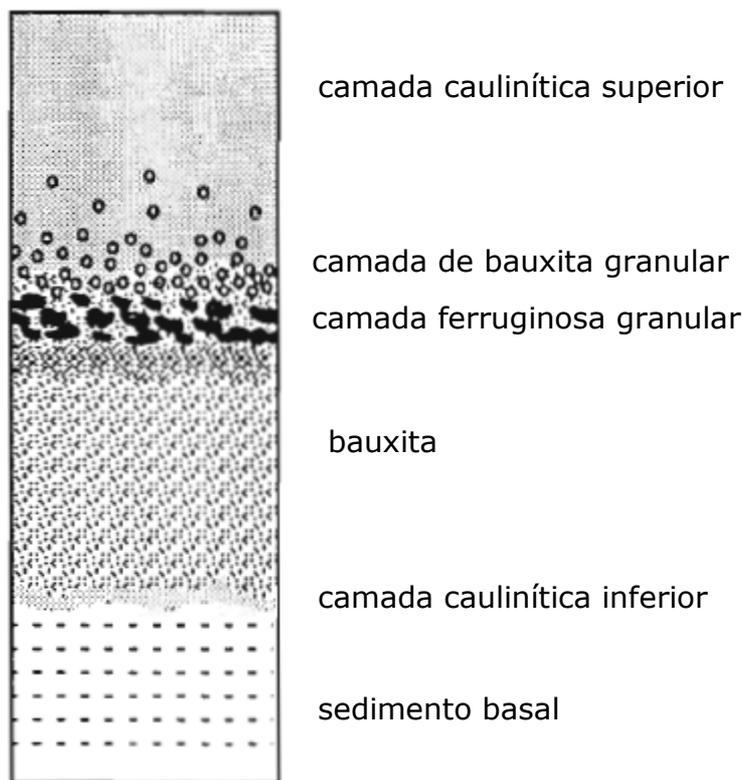


FIGURA 3.1: O perfil geológico da bauxita no depósito (modificado de Carvalho *et al.*, 1997)

### 3.2.1.1. CAMADA CAULÍNICA SUPERIOR

Bastante homogênea e sem estratificação aparente. A espessura varia de 8 a 10m no centro do platô, e, de 0 a 5m nas bordas. É constituída principalmente por caulinita (80%), gibbsita (10%) e quartzo (10%). As proporções entre a caulinita e a gibbsita tendem a se inverter na medida em que a profundidade aumenta e a camada de bauxita granular se aproxima. O ferro está presente em pequena quantidade em forma de goethita na parte superior e de forma mais significativa na camada granular, em forma de hematita.

As características do quartzo e dos grânulos de bauxita encontrados na camada indicam que estes não foram transportados e sim se originaram da transformação *in situ* de uma antiga camada de bauxita. O processo inclui a desferruginização da hematita em goethita e a dissolução da gibbsita. Parte da alumina sofre ressilificação formando a caulinita e o resto é transferido para a bauxita subjacente.

#### 3.2.1.2. CAMADA DE BAUXITA GRANULAR (1 A 3M)

Os grãos de bauxita pertencentes a esta camada são irregulares e pequenos (menores do que 5cm), mas às vezes aparecem aglomerados formando verdadeiros blocos, o que provoca a heterogeneidade dos tamanhos dos nódulos de bauxita da camada. Cada grânulo é composto por cristais muito pequenos de gibbsita. Na parte inferior da camada, começam a ocorrer grãos ferruginosos que surgem cada vez mais espessos com o aumento da profundidade. Na parte inferior, estes grãos praticamente dominam, formando uma zona de transição de aproximadamente 10 cm entre a camada de bauxita granular e a camada ferruginosa granular.

#### 3.2.1.3. CAMADA FERRUGINOSA GRANULAR (1M)

A parte superior da camada (30cm) é constituída por grãos com aproximadamente 10cm de diâmetro, envolvidos em uma matriz gibbsítica. Na parte central (60cm), a matriz passa a ser argilosa, voltando a ser gibbsítica na porção inferior (10cm).

#### 3.2.1.4. CAMADA DE BAUXITA (1 A 6M)

A porção superior da camada (1m) é alumino - ferrosa, compacta e localmente coberta por uma crosta ferruginosa com zonas bastante enriquecidas em ferro. A porção inferior (5m), efetivamente lavrada, é formada por uma bauxita friável e bastante porosa. A camada é entrecortada por grandes bolsões que contém blocos residuais e fragmentos de bauxita e argilas.

#### 3.2.1.5. CAMADA CAULINÍTICA INFERIOR

Apresenta alguns nódulos alumino - ferrosos residuais formados por gibbsita e hematita, envoltos em uma matriz argilosa. Os nódulos residuais de bauxita são formados por grandes cristais de gibbsita por vezes associados à presença de quartzo.

#### 3.2.1.6. SEDIMENTO BASAL

A base do perfil é composta por um sedimento quartzo-argiloso estratificado, com presença de argila em pequena escala.

#### 3.2.2. COMPOSIÇÃO QUÍMICA

A análise química dos elementos da camada superior de caulinita revelou que a camada como um todo pode ser considerada homogênea. A presença da  $\text{SiO}_2$  e o  $\text{Al}_2\text{O}_3$  reflete sua natureza argilo-quartzosa. O conteúdo de  $\text{Fe}_2\text{O}_3$  é baixo (cerca de 8%), enquanto que os conteúdos de  $\text{TiO}_2$  (2,8%) e Zr (1640ppm) podem ser considerados altos.

A bauxita nodular é muito rica em alumina (61%), mas possui um teor bastante baixo de ferro (2%). Já o  $\text{TiO}_2$  e o Zr podem ser encontrados em quantidades significativas na matriz da rocha.

A camada ferruginosa, como o nome já diz, é muito rica em  $\text{Fe}_2\text{O}_3$  (45%), apresentando grandes quantidades de alumina na forma de gibbsita. O  $\text{TiO}_2$  e o Zr possuem uma forte correlação positiva em todas as camadas, constituindo aproximadamente 2% da matriz da camada ferruginosa.

A bauxita possui um conteúdo baixo de sílica (menor que 3%) e uma alta concentração de  $\text{Al}_2\text{O}_3$  (50% a 60%). O conteúdo de ferro oscila entre os níveis superiores e inferiores da camada, variando de 25% na porção próxima à camada ferruginosa e 3% na parte inferior. O  $\text{TiO}_2$  e o Zr aparecem em pequena quantidade (1%).

Se considerarmos a composição química da caulinita podemos dizer que a camada de caulinita inferior apresenta uma baixa relação sílica/alumina. A grande quantidade de alumina se deve à forte presença da gibbsita e, os conteúdos de  $\text{TiO}_2$  e  $\text{Fe}_2\text{O}_3$  apresentam uma forte correlação negativa. Na medida em que as concentrações do óxido de ferro diminuem, aumenta o conteúdo do dióxido de titânio.

A correlação positiva entre  $\text{SiO}_2$  e  $\text{Al}_2\text{O}_3$  é bastante forte, o que comprova que o sedimento basal é formado essencialmente por quartzo e caulinita. A matriz da rocha é pobre em  $\text{Fe}_2\text{O}_3$ , mas existem alguns pontos de forte

concentração deste elemento (18%). O conteúdo de  $\text{TiO}_2$  é baixo (1%) aumentando em direção à camada superior.

A análise do diagrama ternário  $\text{SiO}_2\text{-Al}_2\text{O}_3\text{-TiO}_2$  (Figura 3.2) mostra que cada fácies possui uma composição química definida e que a transição entre elas não é clara. De forma geral, a correlação entre  $\text{TiO}_2$  e  $\text{Fe}_2\text{O}_3$  é baixa para todas as camadas. É por este motivo que as camadas superior e inferior de caulinita apresentam as maiores quantidades de  $\text{TiO}_2$  e as menores de  $\text{Fe}_2\text{O}_3$ .

Dentre os elementos traço, pode-se perceber que a presença de zircônio também é bastante variável entre as camadas. A correlação  $\text{Zr/TiO}_2$  também é bastante forte devido à resistência de ambos à alteração.

O vanádio e, em menor proporção, o cromo, possuem forte correlação com o ferro, indicando um comportamento semelhante dos três elementos durante as fases de transferência e mobilização.

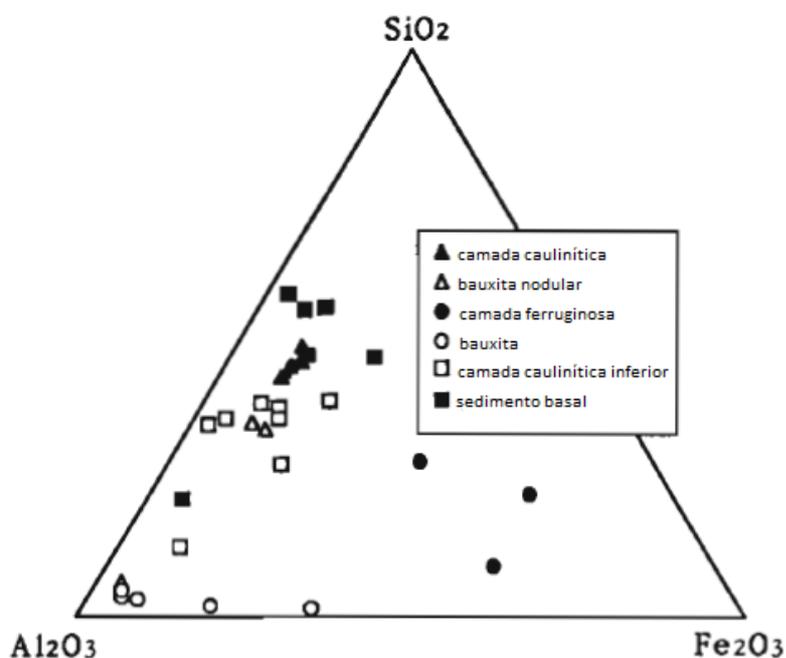


FIGURA 3.2: Amostras de diferentes horizontes do perfil de bauxita plotadas em um diagrama ternário de  $\text{SiO}_2\text{-Fe}_2\text{O}_3\text{-Al}_2\text{O}_3$  (modificado de Carvalho et al. 1997).

### 3.2.3. A EVOLUÇÃO DO PERFIL

Ao contrário do que se pensou por um longo período de tempo, o conjunto de horizontes não foi originado de um processo de deposição de sedimentos, mas sim de uma sequência de eventos geoquímicos. A espessa cobertura sedimentar, submetida a condições tropicais úmidas, sofreu uma série de transformações:

i. Evento de alteração: a espessa camada sedimentar foi submetida a alterações climáticas que provocaram a lixiviação da sílica e a acumulação do ferro e da alumina, originando a bauxita. O conteúdo de ferro ainda era baixo, uma vez que a camada sedimentar não era enriquecida neste elemento. A presença de vegetação na porção superior da camada de bauxita provocou a desferruginação e posterior ressilificação da gibbsita, originando a caulinita.

ii. O evento de alteração seguiu atingindo a camada sedimentar, aumentando a espessura do horizonte de bauxita. Ao mesmo tempo, mas mais lentamente, a alteração devida à vegetação e ao solo foi se aprofundando. O horizonte superior residual de bauxita sofreu a ressilificação, originando a caulinita e mantendo sua espessura constante ao longo do perfil. Já o ferro mobilizado se acumulou em bolsões dentro da camada de bauxita. A elevação gradual do conteúdo de ferro também se deu pela presença de água na camada de bauxita.

iii. Os eventos de alteração foram ocorrendo ao longo de todo o perfil de modo que as camadas de bauxita e superior de caulinita foram tornando-se mais e mais espessas.

iv. O ferro mobilizado se acumulou em um determinado nível do perfil originando uma verdadeira crosta que impediu o prosseguimento dos eventos de alteração nos níveis inferiores.

Não se conhece ao certo os fatores morfológicos e climáticos que ocorreram ao longo dos 90m de espessura sedimentar do platô. De qualquer forma, muito dificilmente ele tenha sido praticamente plano ao longo de toda sua existência. A hipótese mais aceita é que a estrutura ocorreu inicialmente em um sistema ondulado, com espessuras variando ao longo do comprimento.

Conteúdos residuais de alumínio, ferro e sílica são continuamente mobilizados e a crosta de ferro formada é que determina a preservação da camada de bauxita efetivamente lavrada.

### 3.3. A ESTATÍSTICA BÁSICA

O banco de dados é isotópico e possui 819 furos de sondagem dispostos em uma área de aproximadamente 225km<sup>2</sup>, no horizonte da camada de bauxita. A malha de amostragem é praticamente regular, com agrupamento preferencial em alguns locais. Os furos estão separados por uma distância de cerca de 200m nos eixos x e y, e, as amostras estão regularizadas em 0,5 metros. A figura 3.3 representa o mapa de localização das amostras para a variável Alap, onde pode-se perceber a presença dos teores mais altos nas porções superior esquerda e inferior direita da imagem, já os teores mais baixos podem ser encontrados na porção superior central.

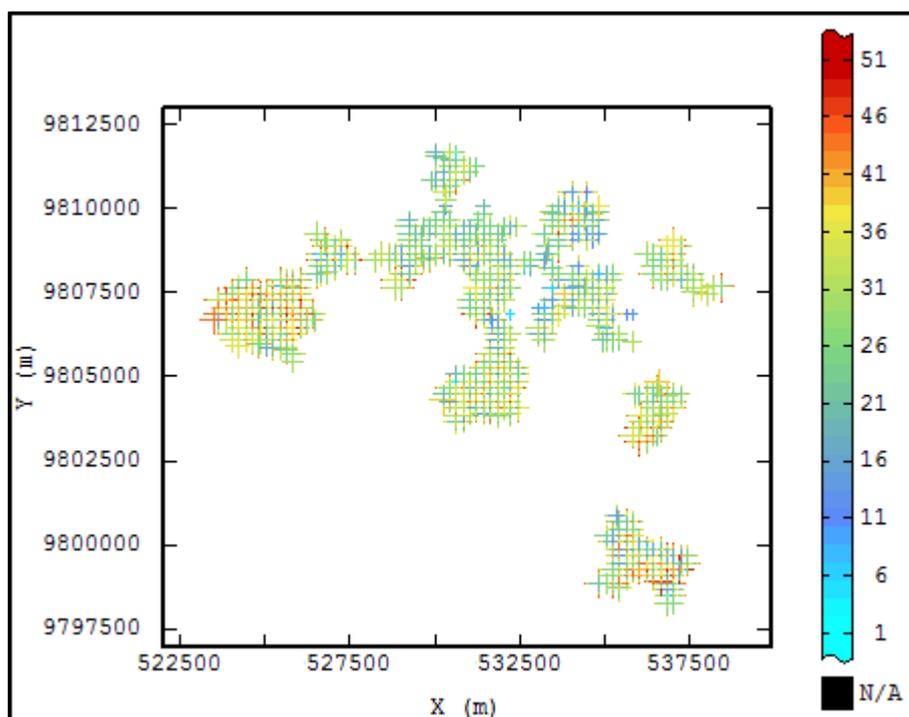


FIGURA 3.3: Mapa de Localização das amostras para a variável Alap.

Como é possível perceber no mapa de localização das amostras, o agrupamento das mesmas não é significativo, o que levou a não utilização de nenhum método de desagrupamento.

### 3.3.1. AS VARIÁVEIS ESTUDADAS

Como já afirmado nos capítulos 1 e 2, as composições carregam informações sobre as proporções pertencentes a um conjunto maior de observações. No caso de um depósito mineral, este conjunto maior é representado pela soma dos teores, ou ainda, pela soma das massas recuperadas em cada faixa granulométrica de interesse.

Os exemplos mais comuns possuem uma soma constante  $k$  e são conhecidos na literatura como dados fechados (*closed data*) (Chayes, 1971). Normalmente,  $k=1$ , o que significa que as medidas foram feitas, ou transformadas em um suporte unitário. A constante  $k$  também pode ser igual a 100%, quando os resultados das análises realizadas são dados em porcentagem. Para que a transformação  $ilr (S^D \rightarrow R^{D-1})$  possa ser aplicada, precisamos que a soma dos teores ou recuperações seja a mesma para todas as amostras analisadas ( $k$ ). E, como isto nem sempre está garantido, principalmente porque, na maioria dos casos, as análises químicas não abrangeram todo o espectro de minerais presentes nas amostras, precisamos garantir este fechamento artificialmente.

Existem duas soluções para este problema de fechamento: criar uma variável artificial para completar os valores faltantes da soma ou; realizar a operação de fechamento, dada por:

*Sendo  $\mathbf{z}$  um vetor qualquer, formado por  $D$  componentes reais positivas:*

$$\mathbf{z} = [z_1, z_2, \dots, z_D] \in \mathbb{R}_+^D \quad (3.1)$$

*( $z_i > 0$  para todos os  $i=1,2,3,\dots,D$ ), o fechamento  $C$  de  $\mathbf{z}$  é:*

$$C(\mathbf{z}) = \left[ \frac{k \cdot z_1}{\sum_{i=1}^D z_i}, \frac{k \cdot z_2}{\sum_{i=1}^D z_i}, \dots, \frac{k \cdot z_D}{\sum_{i=1}^D z_i} \right] \quad (3.2)$$

Egozcue (2003) provou que, para a maioria dos casos, ambas as operações levam a resultados idênticos ou pelo menos, compatíveis.

Nas quatro primeiras combinações de métodos propostas nesta tese, o comportamento das quatro variáveis principais, dentro da faixa 14#, será

estudado: alumínio aproveitável (Alap), sílica (Si), ferro (Fe) e titânio (Ti), juntamente com a variável artificial "Resto", adicionada ao banco de dados para garantir o fechamento da soma dos teores em 100%.

Já na quinta combinação, será aplicada a operação de fechamento (*closure*) e, somente os teores de alumínio aproveitável (Alap), sílica (Si), ferro (Fe) e titânio (Ti) serão simulados.

Como um dos objetivos do trabalho é analisar a qualidade da reprodução das correlações entre as variáveis, e, a criação de uma variável artificial implica diretamente no surgimento de correlações que não existem fisicamente, espera-se que a presença de correlações espúrias surja nas aplicações com a utilização da variável artificial (combinações i. ii. iii. e iv.). Essa espera-se seja eliminada na aplicação da combinação v (operação de fechamento).

### 3.3.2. ACUMULAÇÃO

A coleta de amostras de furos de sondagem é bastante difícil em depósitos friáveis como os de bauxita. Ao contrário do que acontece com depósitos metálicos com rochas competentes geomecanicamente, onde as amostras são retiradas praticamente intactas, em depósitos formados por rochas friáveis as amostras obtidas são compostas por pedaços de diferentes granulometrias.

A fácil fragmentação dessas rochas também contribui para que dificilmente se consiga uma recuperação de 100% da massa que compõe as amostras, sendo o percentual de recuperação mássica significativamente diferente entre elas (figura 3.4).

Com o objetivo de conhecer as proporções pertencentes a cada tamanho de grão de interesse procede-se o peneiramento da amostra em diferentes faixas granulométricas. Desta forma, os resultados obtidos permitirão que qualquer interpolação ou simulação realizada considere as proporções relativas entre a massa da amostra em cada granulometria e o teor analisado das variáveis de interesse por meio do processo de acumulação.

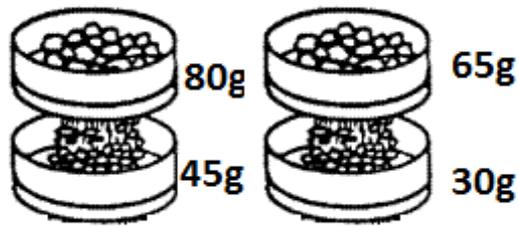


FIGURA 3.4: Exemplo ilustrativo de duas amostras analisadas em duas faixas granulométricas. A recuperação mássica das amostras é diferente entre si na mesma faixa e entre as faixas (Modificado de Chaves, A.P., Perez, A.E.C., 1999).

A simulação de um atributo pertencente a um depósito que demanda o uso da variável auxiliar acumulação dos teores (equação 3.3) envolve a posterior desacumulação dos mesmos (equação 3.4) para efeito de validação dos cenários simulados frente aos dados originais. Ou seja, as variáveis acumuladas são simuladas e posteriormente o valor obtido em cada bloco é dividido pelo valor da recuperação mássica, também simulado para aquele bloco.

$$Z_{ac}(u) = \frac{Z(u)Rec_{\#}(u)}{100} \quad (3.3)$$

onde,  $Z_{ac}(u)$  é o valor de  $Z(u)$  acumulado pela recuperação mássica da amostra na faixa de interesse ( $Rec_{\#}(u)$ ).

$$Z^*(u) = \left( \frac{Z^*_{ac}(u)}{Rec^*_{\#}(u)} \right) \quad (3.4)$$

onde,

$Z^*(u)$  é o valor de  $Z(u)$  desacumulado

$Z^*_{ac}(u)$  é o valor simulado da variável acumulada

$Rec^*_{\#}$  é o valor simulado da recuperação mássica na faixa de interesse.

Na prática, esse procedimento implica na utilização dos mesmos parâmetros e modelos variográficos para a simulação de ambos os atributos: acumulação e recuperação mássica, de modo a evitar discrepâncias nos valores finais obtidos. Por exemplo, se para determinado bloco o valor sorteado para o atributo acumulado for maior do que a média do mesmo, e, o valor da recuperação simulada for menor do que sua média, o valor final simulado poderá ser muito superior ao valor real do local.

Neste estudo, todas as variáveis, em todas as combinações, foram simuladas com os parâmetros e modelos variográficos utilizados na simulação da recuperação mássica da faixa estudada (Rec14#) em suporte de pontos ou em suporte de blocos, dependendo da combinação em questão.

### 3.3.3. AS COORDENADAS

Outro fato é que as coordenadas verticais (z) dos furos também precisaram ser ajustadas de acordo com o comportamento do corpo geológico. A camada de bauxita se encontra entre duas outras camadas (camada ferruginosa granular e camada caulínica inferior), sendo que o limite entre elas é irregular. Para viabilizar a modelagem da continuidade espacial, a camada de bauxita foi planificada, tomando como referência as coordenadas das camadas superior e inferior. Ou seja, as coordenadas cartesianas foram transformadas em coordenadas estratigráficas.

A figura 3.5 mostra um esquema simplificado com a transformação de coordenadas, onde A seria a camada ferruginosa granular, B seria o horizonte de bauxita e C a camada caulínica inferior. Na parte esquerda da figura os horizontes estão representados em coordenadas cartesianas e na parte direita somente a camada de bauxita está representada em coordenadas estratigráficas.

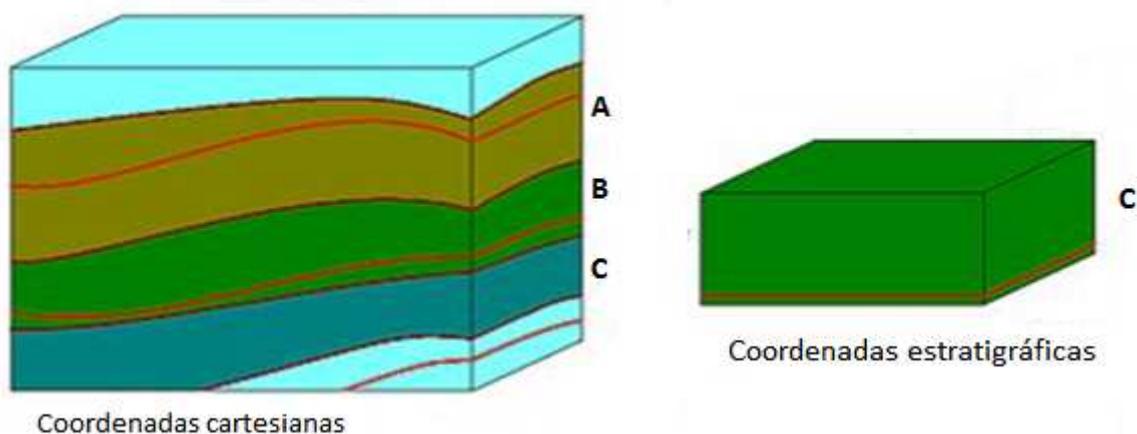


Figura 3.5: esquema simplificado com a transformação de coordenadas, onde a seria a camada ferruginosa granular, b seria o horizonte de bauxita e c a camada caulínica inferior. Na parte esquerda da figura, os horizontes estão representados em coordenadas cartesianas e na parte direita somente a camada de bauxita está representada em coordenadas estratigráficas. (Modificado de Ecole des Mines de Paris & Geovariances, 2012).

### 3.3.4. A ESTATÍSTICA BÁSICA

Nesta seção, será apresentada a estatística básica univariada das variáveis de interesse: Rec14#, Alap, Fe, Si, Ti e Resto. A tabela 3.1 contém informações importantes sobre o teor médio, o desvio padrão e os valores máximos e mínimos amostrados. Também é importante salientar, que os valores exibidos são resultado da ponderação dos valores amostrados pela respectiva recuperação amostral da faixa 14#, ou seja são estatísticas ponderadas uma vez que cada datum possui distinto suporte mássico

O desvio padrão da variável Rec14# é bastante alto (figura 3.6), o que corrobora para a utilização dos valores acumulados para a realização das simulações, uma vez que a variação entre as recuperações das diferentes amostras é significativa.

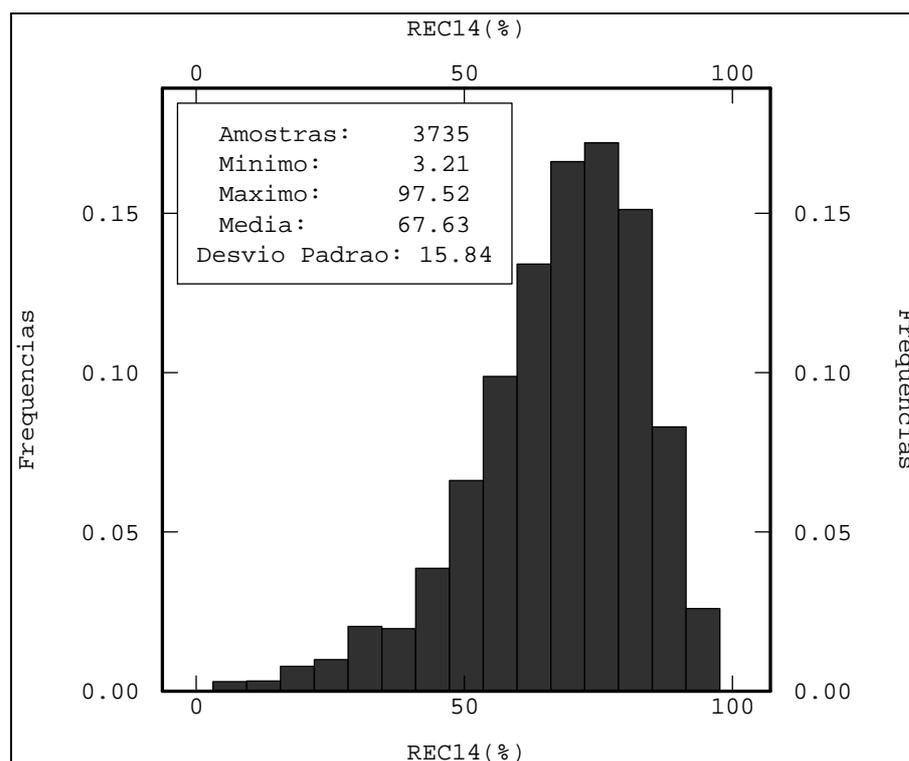


FIGURA 3.6: histograma da variável rec14#, percebe-se o espalhamento dos valores de recuperação nas diferentes amostras.

TABELA 3.1: Estatística básica univariada das variáveis estudadas.

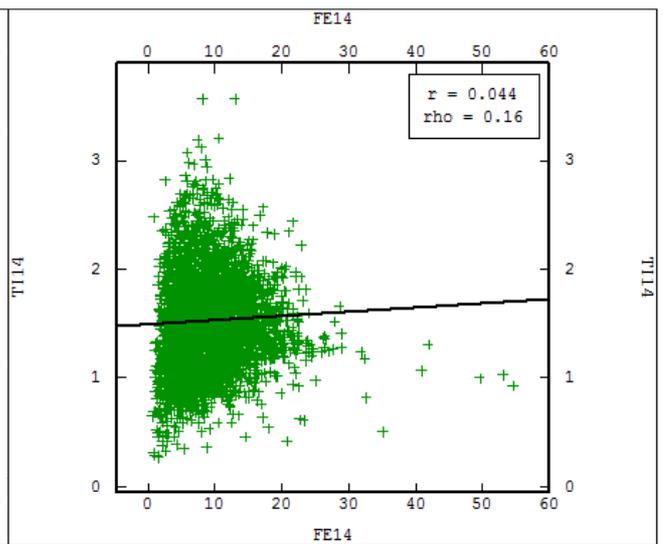
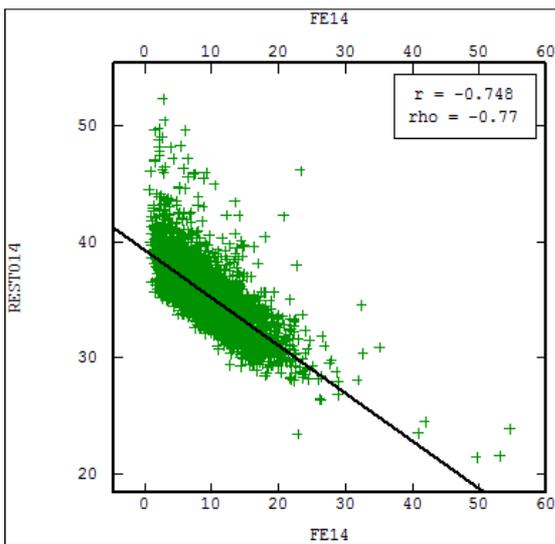
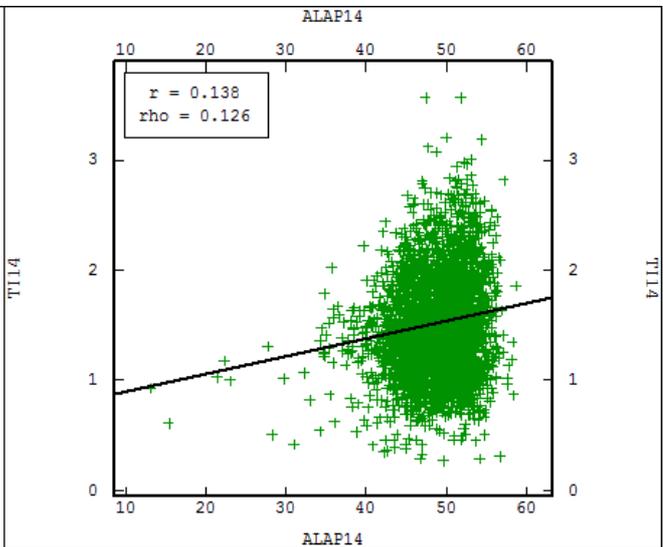
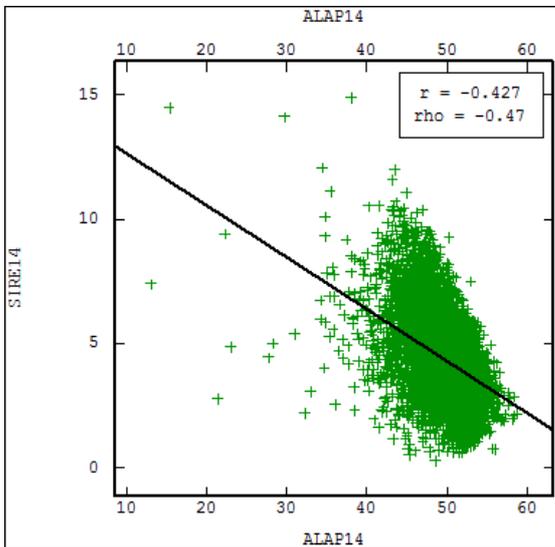
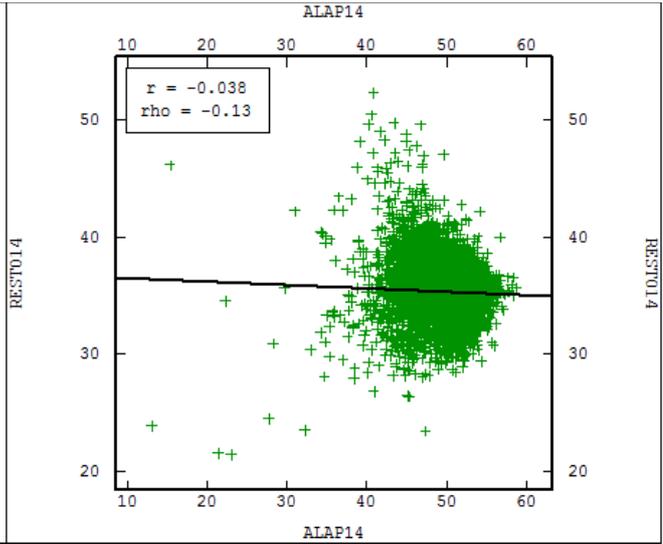
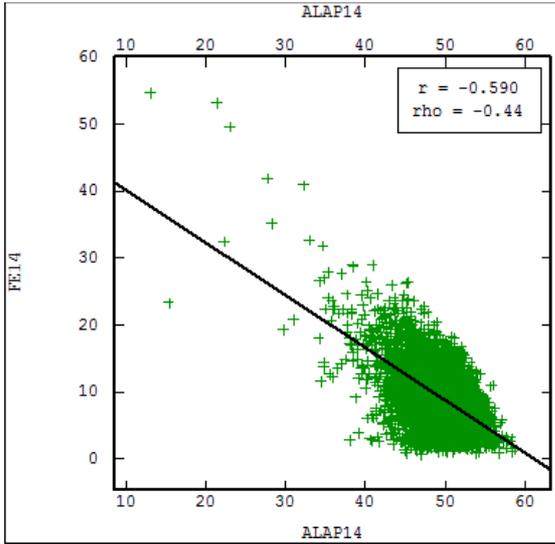
<i>Dados</i>	<i>Mínimo (%)</i>	<i>Máximo (%)</i>	<i>Média ponderada (%)</i>	<i>Desvio Padrão (%)</i>	
<i>Rec14#</i>	3735	3,21	97,52	67,63	15,84
<i>Alap 14#</i>	3735	13,07	58,78	49,17	3,83
<i>Fe14#</i>	3735	0,69	54,65	9,49	4,98
<i>Resto14#</i>	3735	21,47	52,32	35,76	2,88
<i>Si14#</i>	3735	0,27	14,9	4,44	1,97
<i>Ti14#</i>	3735	0,27	3,57	1,53	0,45

### 3.3.5. CORRELAÇÕES

Na figura 3.7, estão exibidas as correlações lineares entre as variáveis (ponderadas pela recuperação da faixa de interesse) e acompanhadas dos respectivos coeficientes de correlação linear ( $r$ ) e de Spearman ( $\rho$ ). Exibem correlação linear significativa:  $Alap \times Fe$ ,  $Alap \times Si$ ,  $Fex \times Resto$  e  $Resto \times Si$ . Não exibem correlação nenhuma, ou exibem correlação linear insignificante:  $Alap \times Resto$ ,  $Alap \times Ti$ ,  $Fex \times Si$ ,  $Fex \times Ti$ ,  $Resto \times Ti$  e  $Si \times Ti$ .

Analisando os resultados de correlação não paramétrica, é possível afirmar que somente  $Fex \times Resto$  e  $Resto \times Si$  possuem um coeficiente que representa uma correlação significativa, mas que já será verificada quando da análise da correlação linear, já que esta também é representativa. Assim, a reprodução do coeficiente de Spearman não será analisada nesta tese.

A variável artificial Resto, como já salientado, não existe no banco de dados original, muito menos representa um único elemento pertencente ao depósito. Por surgir no estudo como forma de fechamento da soma dos teores em 100%, esta variável representa sim um grupo de elementos não analisados e por este motivo qualquer correlação que ela apresente com as demais é artificial, dita espúria, e, deve ser desconsiderada.



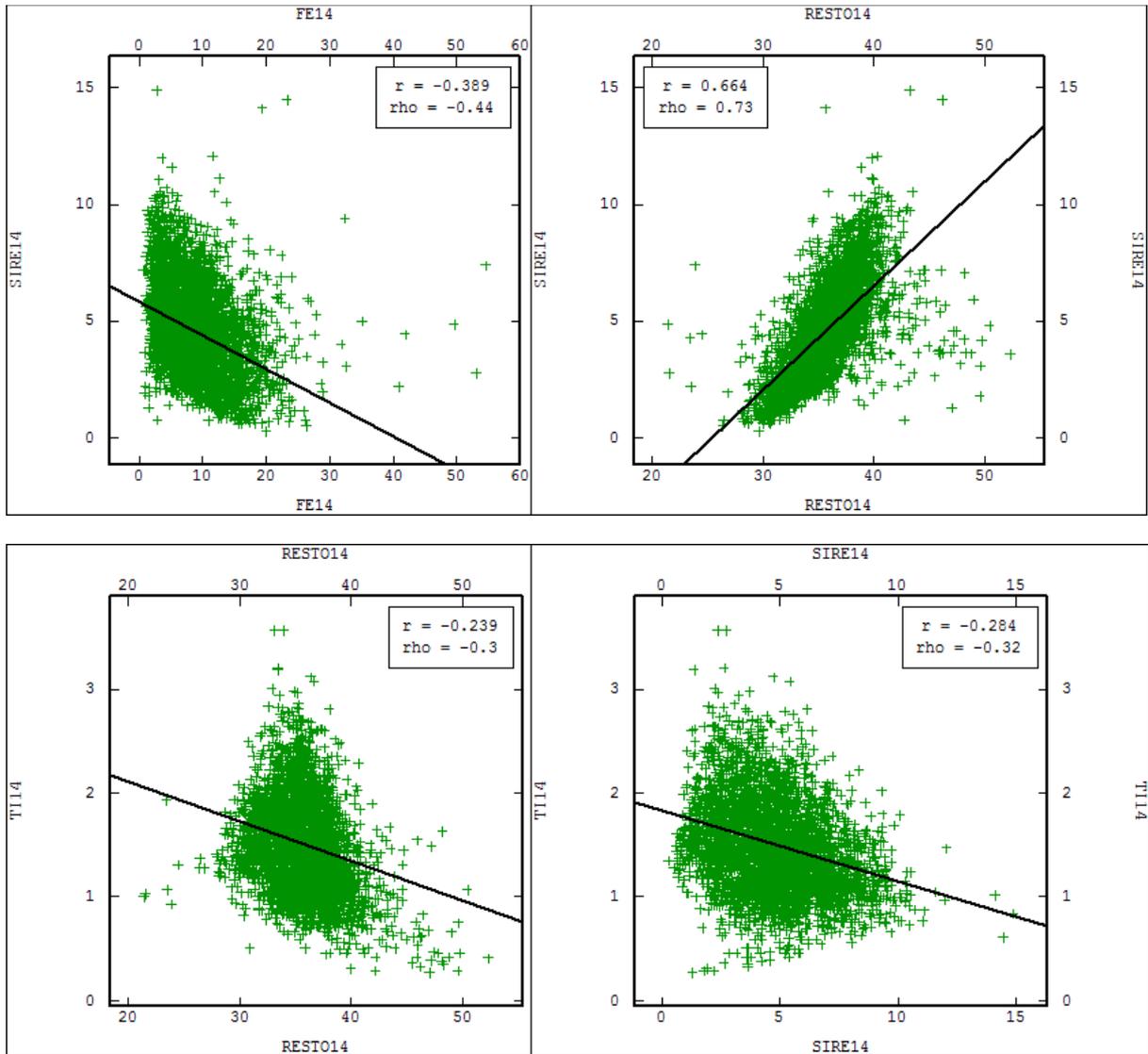


FIGURA 3.7: correlações lineares entre as variáveis, acompanhadas do respectivo coeficiente de correlação. Exibem correlação significativa: AlapxFe, AlapxSi, RestoxFe e RestoxSi. Não exibem correlação nenhuma, ou exibem correlação insignificante: AlapxResto, AlapxTi, FexSi, FexTi, RestoxTi e SixTi.

### 3.3.6. CONTINUIDADE ESPACIAL

A grande variabilidade das amostras disponíveis favoreceu a utilização de correlogramas, em vez de variogramas, para o mapeamento das possíveis direções de anisotropia do depósito. Uma vez que consideram a covariância entre os vetores de separação  $h$ , os correlogramas (equação 3.5) são mais adequados para a análise de distribuições erráticas que os variogramas (equação 3.6).

Assim como para as correlações, os correlogramas foram calculados para as variáveis de interesse ponderadas pela recuperação da faixa 14#, da respectiva amostra.

$$\frac{1}{n} \sum_n (Z_\alpha - m_z) (Z_\beta - m_z) / \sigma_z^2 \quad (3.5)$$

$$\frac{1}{2n} \sum_n (Z_\alpha - Z_\beta)^2 \quad (3.6)$$

onde,

$n$  é o número de pares de dados separados por uma determinada distância;

$Z_\alpha$  e  $Z_\beta$  representam os valores das duas amostras que formam um par, separadas por uma determinada distância;

$\sigma_z^2$  é a variância total do banco de dados;

$m_z$  é a média dos dados.

### 3.3.6.1. CORRELOGRAMAS AO LONGO DO FURO (DTH)

Inicialmente, os correlogramas “ao longo do furo” foram ajustados para possibilitar a inferência do efeito pepita associado a cada atributo (figuras 3.8 a 3.13). Percebe-se, que o comportamento das variáveis ao longo da vertical é bastante semelhante. Grande parte dos correlogramas foi modelada com duas estruturas esféricas de alcance aproximado de 2,5m e 5m respectivamente.

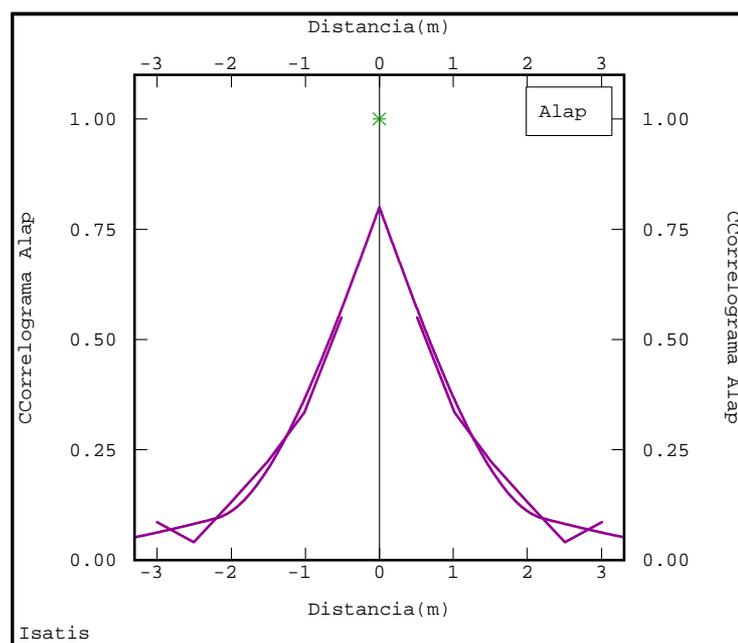


FIGURA 3.8: Correlograma ao longo do furo para a variável Alap. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado.

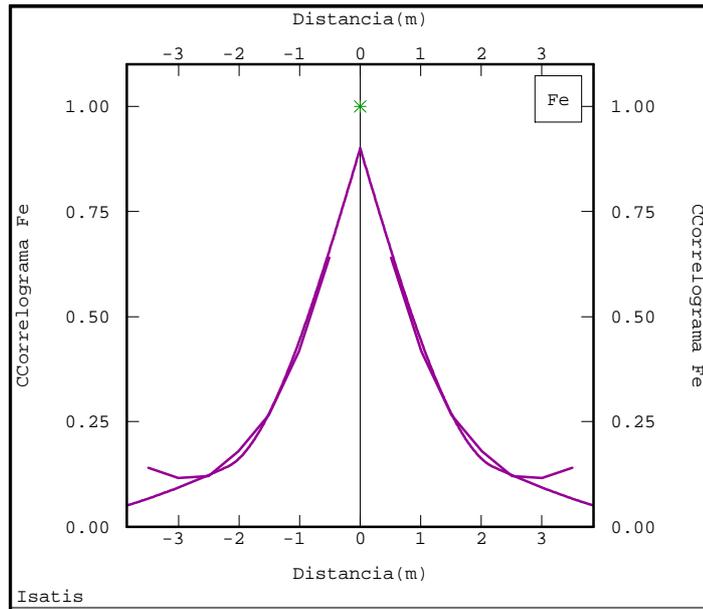


FIGURA 3.9: Correlograma ao longo do furo para a variável Fe. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado.

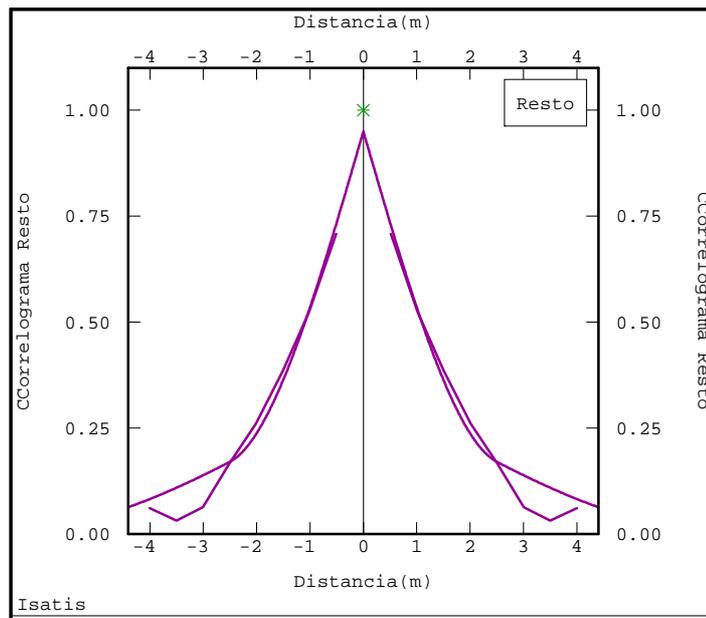


FIGURA 3.10: Correlograma ao longo do furo para a variável artificial Resto. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado.

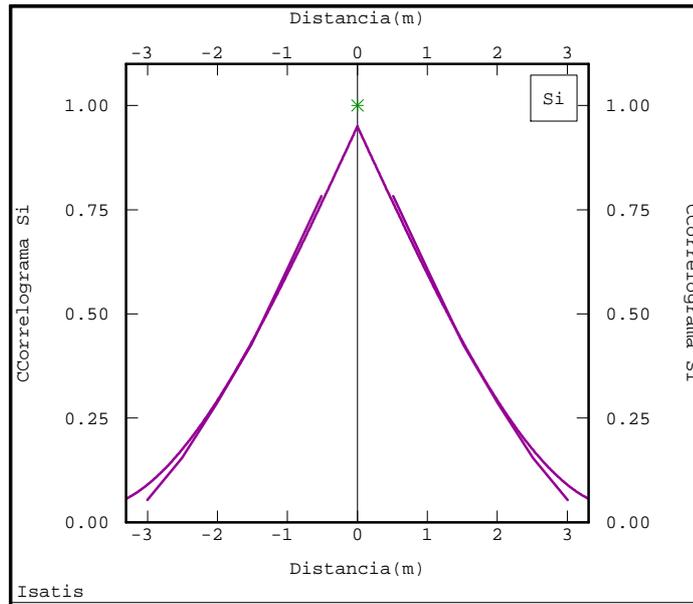


FIGURA 3.11: Correlograma ao longo do furo para a variável Si. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado.

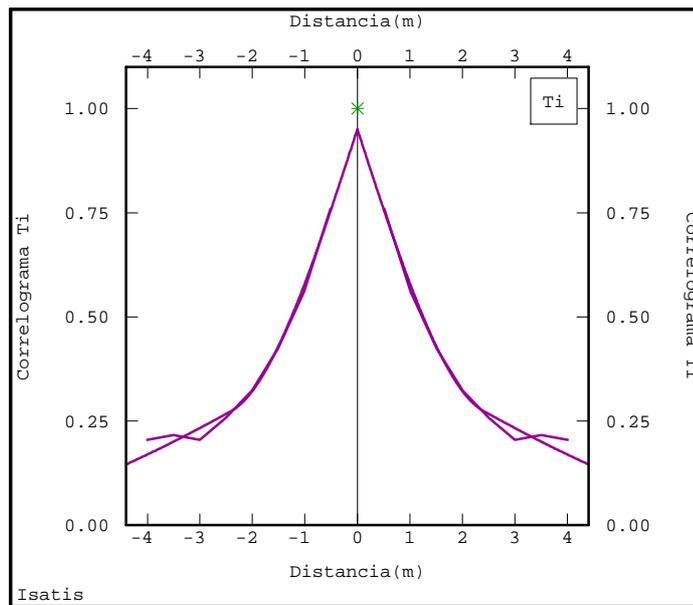


FIGURA 3.12: Correlograma ao longo do furo para a variável Ti. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado.

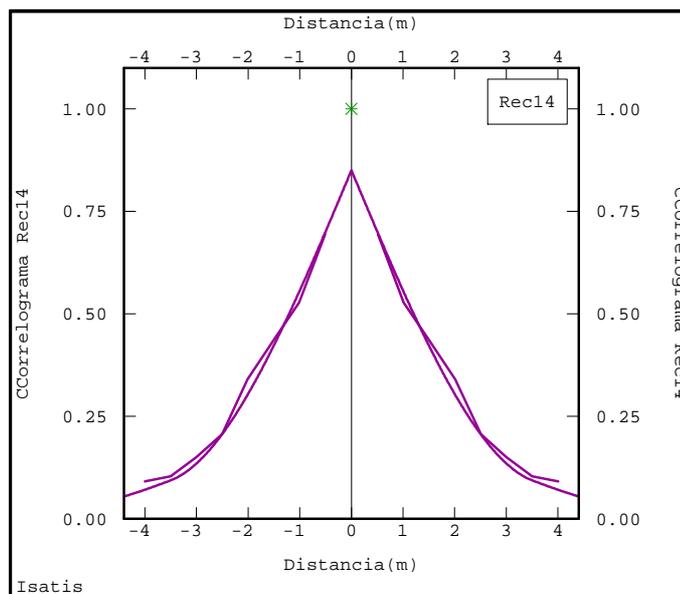


FIGURA 3.13: Correlograma ao longo do furo para a variável Rec14#. A linha roxa pontilhada representa o correlograma calculado e a linha roxa contínua representa o modelo ajustado.

### 3.3.6.2. VERIFICAÇÃO DAS DIREÇÕES DE CONTINUIDADE ESPACIAL

O primeiro passo para a modelagem da continuidade espacial das variáveis é o cálculo dos correlogramas experimentais direcionais ao longo do plano xy (figuras 3.14 a 3.19). A análise dos correlogramas direcionais servirá para a identificação das possíveis direções de anisotropia de cada variável, quando for o caso.

Os cálculos foram executados com os seguintes parâmetros:

- i. Lag: 200m, correspondente ao espaçamento da malha amostral;
- ii. Número de vetores de separação h: 10, para analisar a continuidade espacial em um raio de até 2 km
- iii. Tolerância: 50% do lag, para captar amostras com espaçamento de no mínimo 100m;
- iii. Tolerância angular:  $22,5^\circ$  para garantir que somente as amostras correspondentes à direção que se quer analisar sejam captadas;
- iv. Espessura da fatia: 0,5m, evitando a influência da variabilidade vertical.

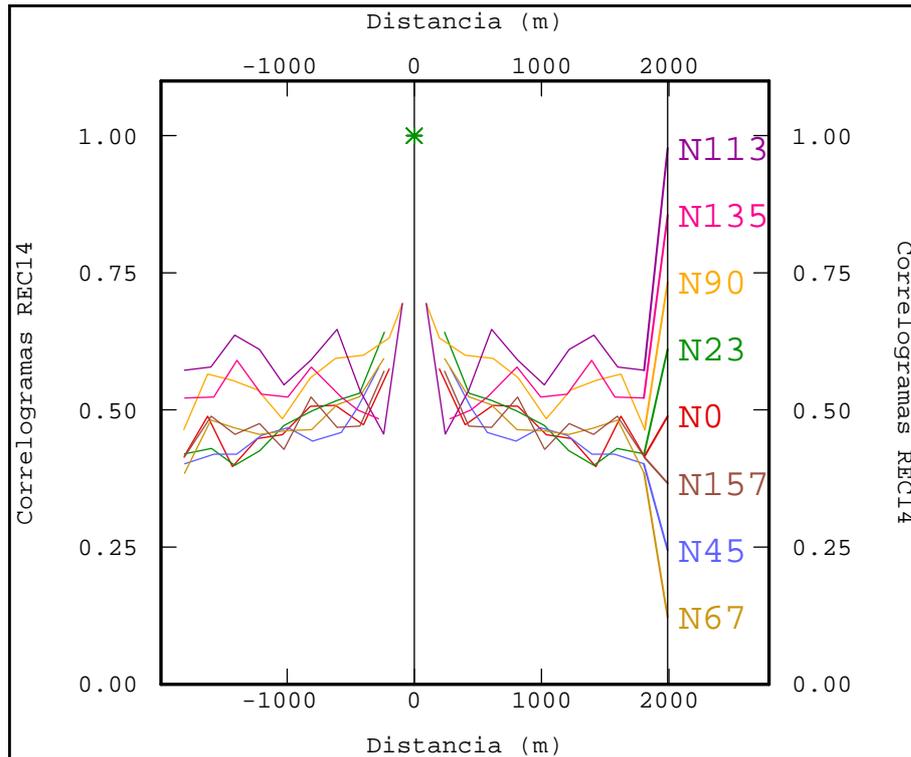


FIGURA 3.14: Correlogramas direcionais para a variável REC 14#.

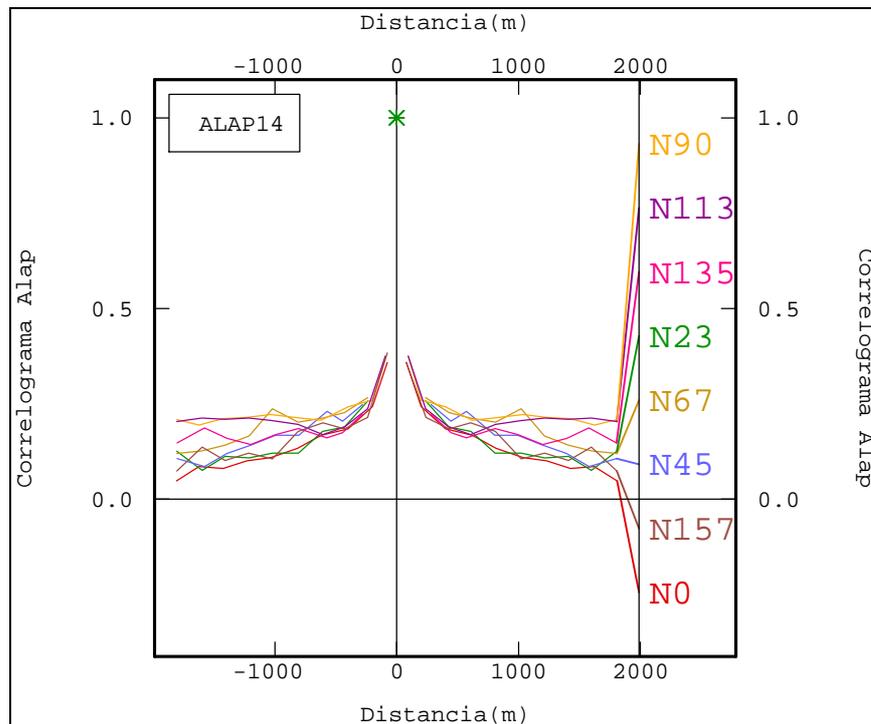


FIGURA 3.15: Correlogramas direcionais para a variável Alap.

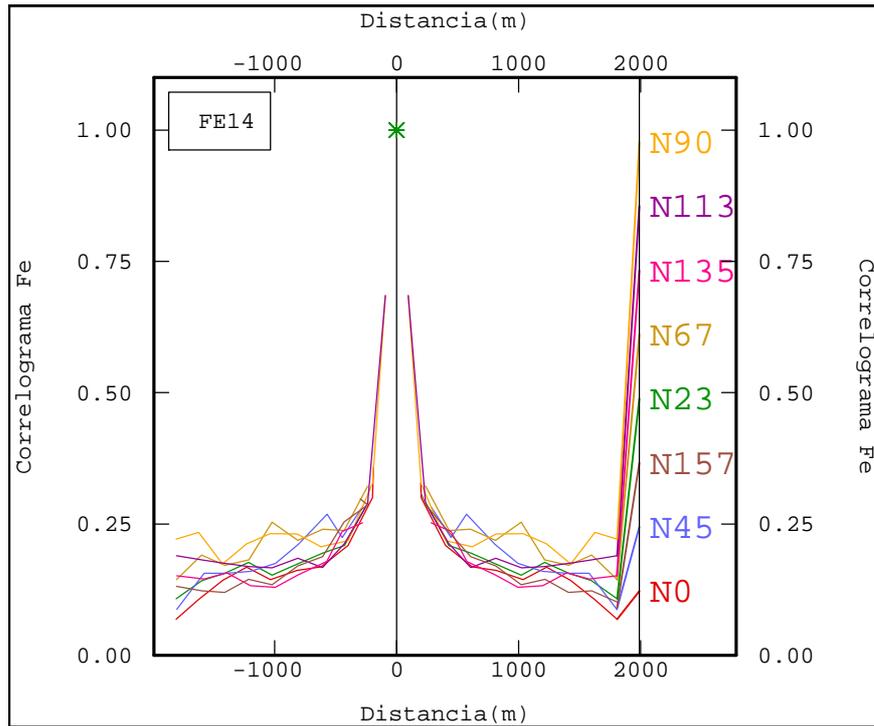


FIGURA 3.16: Correlogramas direcionais para a variável Fe.

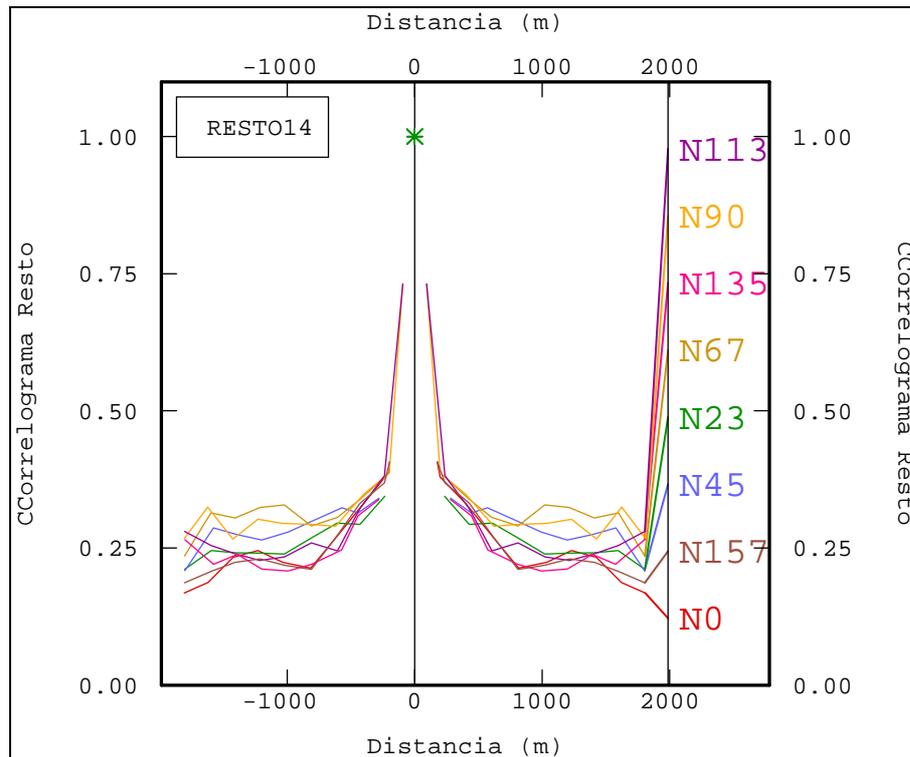


FIGURA 3.17: Correlogramas direcionais para a variável Resto.

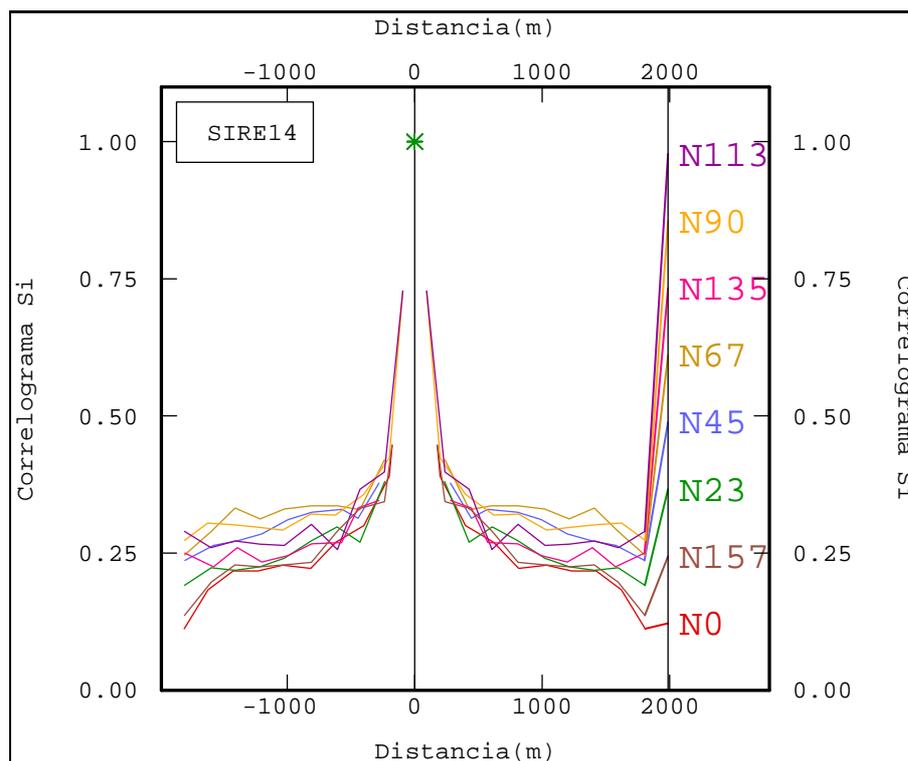


FIGURA 3.18: Correlogramas direcionais para a variável Si.

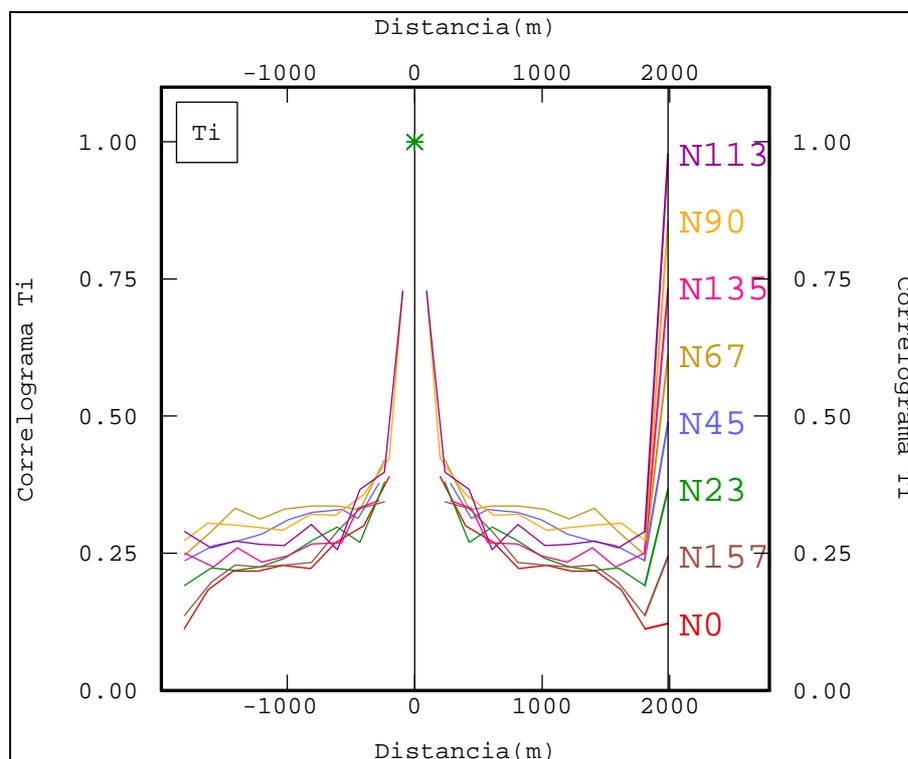


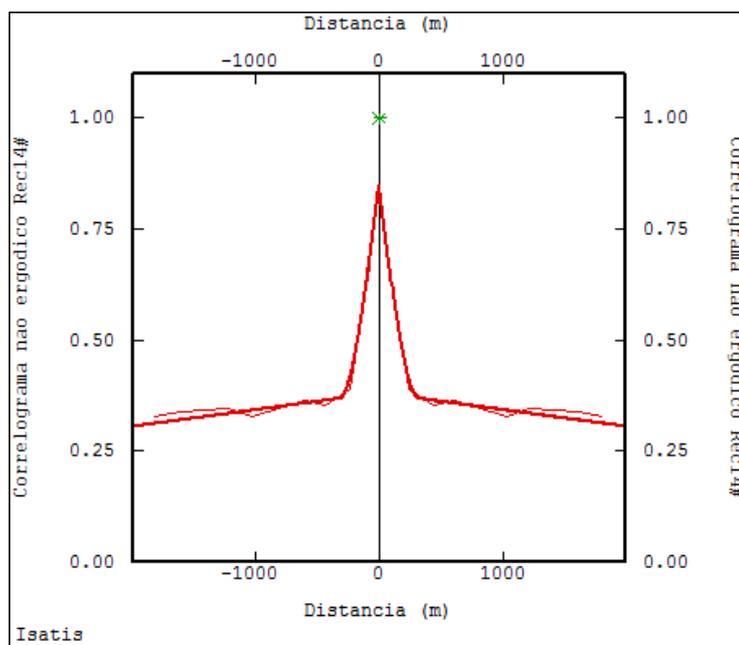
FIGURA 3.19: Correlogramas direcionais para a variável Ti.

Note que, em todas as direções analisadas, para todas as variáveis ocorre o fenômeno chamado de "anisotropia zonal", ou seja, a distribuição dos teores no depósito é tão contínua horizontalmente que a variância máxima dos

dados nunca é atingida (ou em termos teóricos a uma distância infinita). Também, é possível perceber que a primeira estrutura é responsável por aproximadamente 90% da variância total.

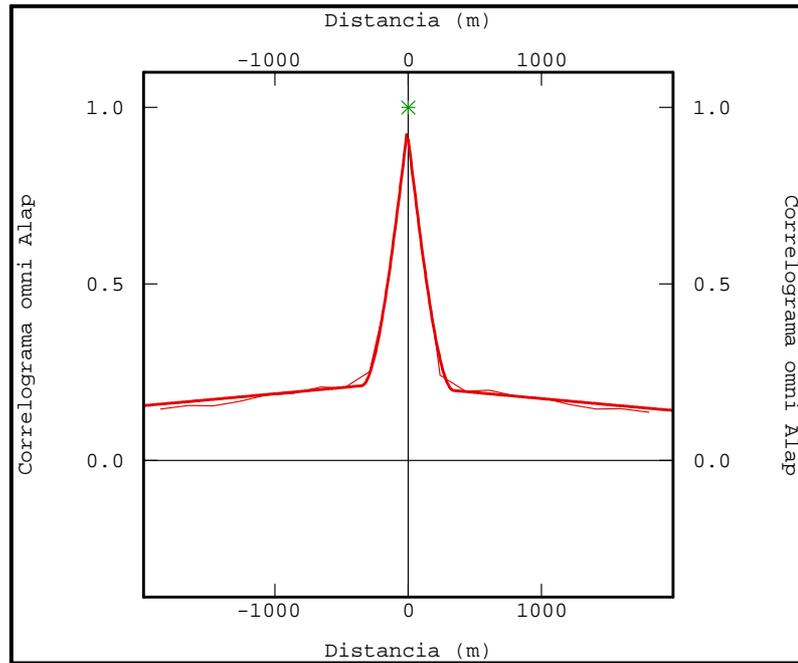
Outro fato é que, para todos os atributos, nenhuma das direções analisadas possui um alcance significativamente superior às demais. Neste sentido, optou-se pela modelagem e utilização dos correlogramas omnidirecionais no plano horizontal (figuras 3.19 a 3.24).

Como o depósito é tabular, onde a espessura é muito menor que a extensão no plano xy, a verificação das direções de mergulho e inclinação de uma provável anisotropia não é possível, nem necessária. Neste tipo ocorrência mineral, não são encontrados pares de amostras para mergulhos e inclinações diferentes ou muito próximos ao grau zero.



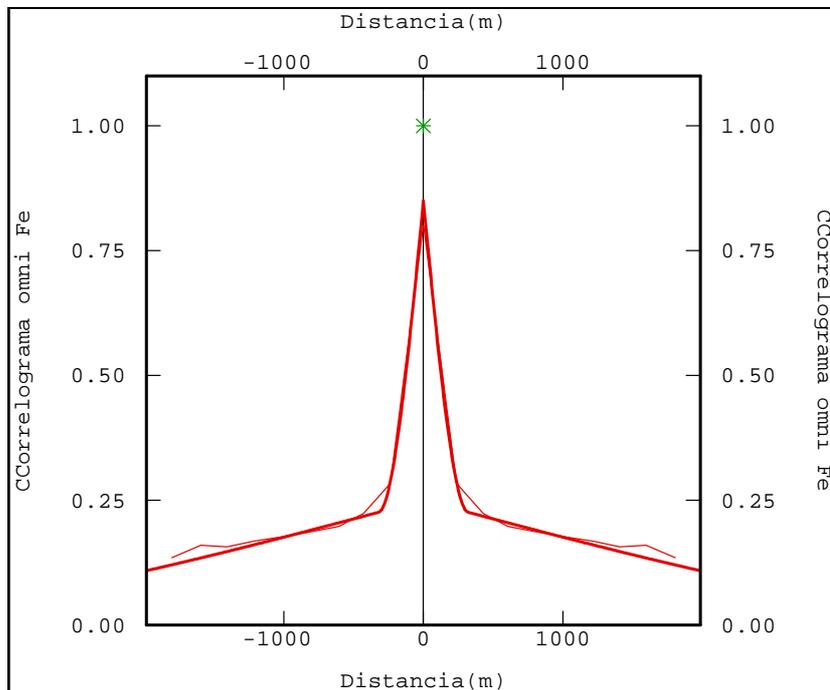
$$\gamma(h) = 0,15 + 0,47Sph(320) + 0,38Sph(15000)$$

FIGURA 3.20: Correlograma omnidirecional para a variável Rec14#. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura.



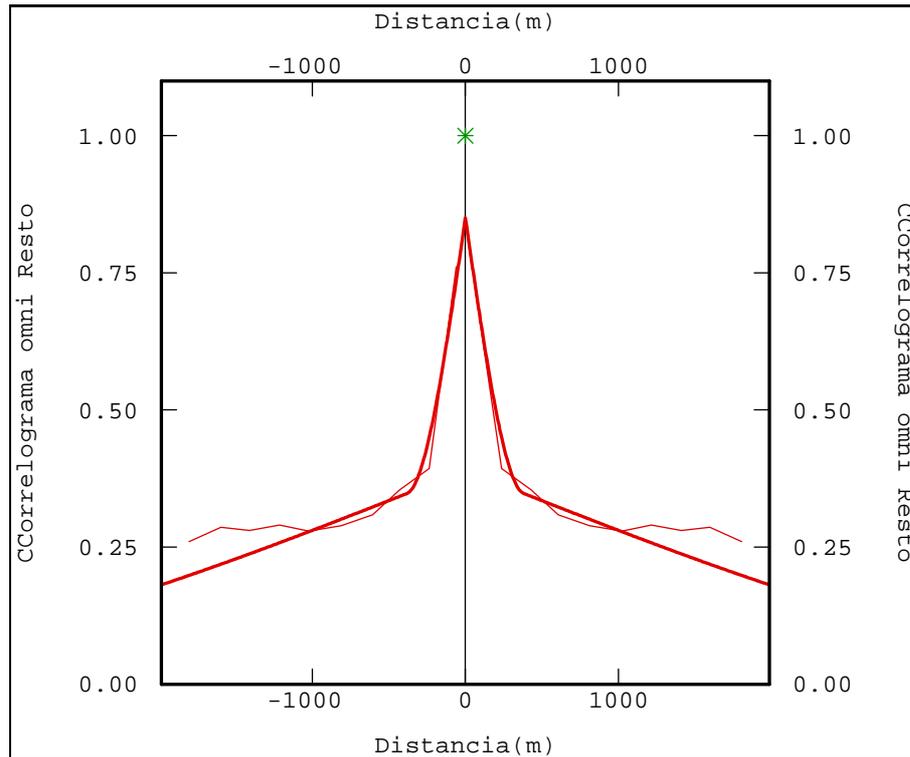
$$\gamma(h) = 0,2 + 0,7Sph(340) + 0,1Sph(9000)$$

FIGURA 3.21: Correlograma omnidirecional para a variável Alap. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura.



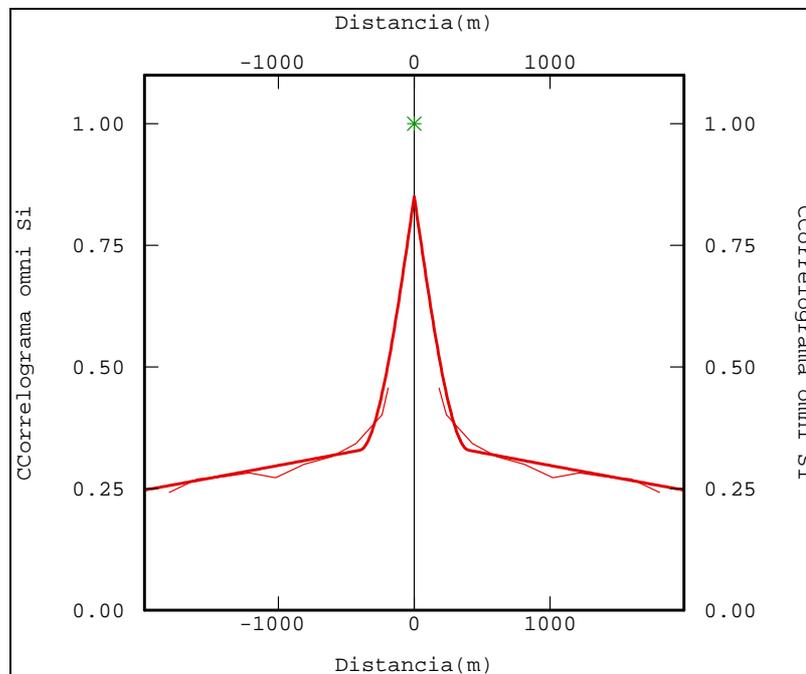
$$\gamma(h) = 0,1 + 0,6Sph(320) + 0,30Sph(5000)$$

FIGURA 3.22: Correlograma omnidirecional para a variável Fe. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura.



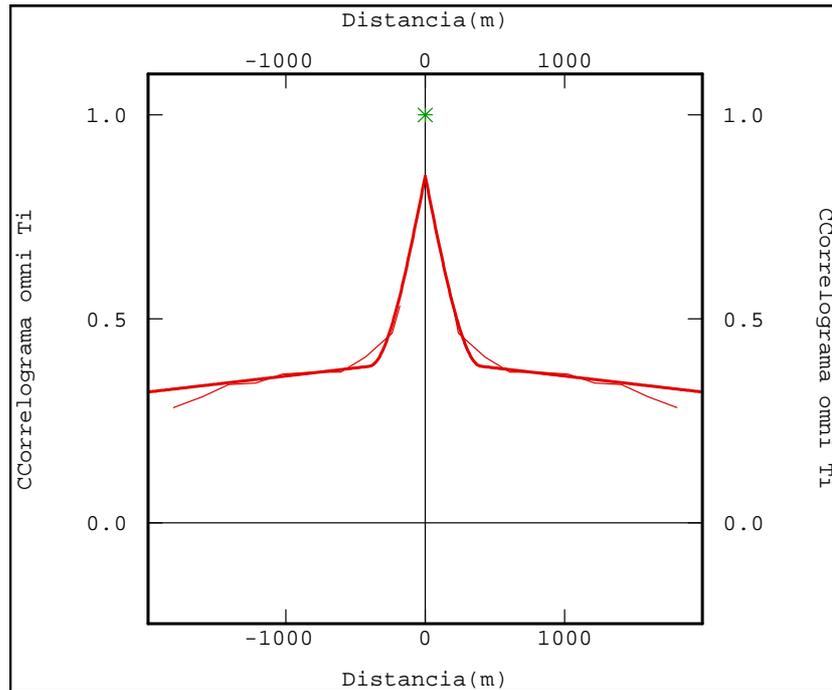
$$\gamma(h) = 0,05 + 0,46Sph(380) + 0,49Sph(5000)$$

FIGURA 3.23: Correlograma omnidirecional para a variável Resto. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura.



$$\gamma(h) = 0,05 + 0,5Sph(400) + 0,45Sph(10000)$$

FIGURA 3.24: Correlograma omnidirecional para a variável Si. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura.



$$\gamma(h) = 0,05 + 0,45Sph(400) + 0,50Sph(15000)$$

FIGURA 3.25: Correlograma omnidirecional para a variável Ti. A linha vermelha estreita representa o correlograma experimental e a linha espessa o modelo ajustado. A equação do modelo ajustado está abaixo da figura.

### 3.4. COMENTÁRIOS

O presente capítulo abordou alguns detalhes da geologia do depósito, juntamente com uma análise da estatística básica das variáveis. A continuidade espacial individual dos atributos também foi calculada e ajustada, sendo que os parâmetros calculados para a Rec14# serão utilizados nas simulações realizadas futuramente, e, os correlogramas calculados e modelados para as demais variáveis serão utilizados para a validação das simulações realizadas.

As respostas provenientes da aplicação das cinco combinações propostas neste estudo serão apresentadas no quarto capítulo, juntamente com uma análise das questões positivas e negativas da aplicação de cada uma delas.

# Capítulo 4

## RESULTADOS E DISCUSSÃO

A apresentação e discussão dos resultados obtidos na: simulação individual dos teores originais (combinação 1); simulação dos teores originais após a transformação em razões logarítmicas isométricas (combinação 2), a partir de agora chamada de simulação das ilrs; simulação direta em blocos dos fatores de autocorrelação mínimos e máximos (MAFs) chamada de simulação direta dos MAFs (combinação 3); a simulação dos MAFs das razões logarítmicas isométricas (simulação dos MAFs das ilrs – combinação 4) e; a simulação dos MAFs das ilrs com a utilização da operação de fechamento (combinação 5), fazem parte deste quarto capítulo.

O mapeamento do espaço de incerteza foi considerado como parte secundária das análises aqui realizadas, assim, vinte cenários possíveis e equiprováveis foram gerados para cada variável, dentro de cada combinação proposta, quantidade que demonstrou ser suficiente para o atendimento do objetivo estabelecido. É importante salientar que apesar de não ser o foco desse estudo, o mapeamento da incerteza é o objetivo final da simulação e em casos multivariáveis, quando não se leva em conta as correlações entre as variáveis na simulação, arrisca-se tomar decisões erradas quando se faz parametrizações cruzadas a partir das realizações. A análise dos resultados consiste basicamente da:

i. Verificação da reprodução da estatística básica e das características originais de continuidade espacial (anexo A). Para que um modelo simulado seja considerado adequado para utilizações futuras, ele deve reproduzir as principais características estatísticas do depósito estudado.

ii. Verificação da reprodução das correlações entre as variáveis. Se determinadas variáveis possuem uma correlação significativa ou, nenhuma correlação verificada nos dados originais, os cenários simulados também devem respeitar e garantir que estas características sejam reproduzidas.

iii. Análise do fechamento de teores. A soma dos teores dos elementos presentes em um depósito mineral, global ou para determinada faixa granulométrica, nos dados originais, deve também ser a soma dos teores nos blocos ou faixas simulados. Ou, para o caso da simulação dos MAFs das ilrs com a operação de fechamento, a soma das proporções dos teores nos blocos deve ser uma constante (unitária neste caso), a mesma da soma das proporções de teores das amostras.

Os resultados das combinações testadas foram avaliados de forma individual e entre grupos, de acordo com os critérios i., ii. e iii. Por fim, são apontadas as vantagens e desvantagens de cada combinação sobre as demais.

#### 4.1. SIMULAÇÃO DA VARIÁVEL REC14#

Como já salientado no capítulo 3, a simulação da variável Rec14# se faz necessária para viabilizar a desacumulação dos teores simulados para as variáveis de interesse. Assim sendo, a variável Rec14# foi simulada por bandas rotativas em suporte de pontos posteriormente reblocados, para utilização nas combinações 1 e 2, e, simulada diretamente em suporte de blocos, para utilização nas combinações 3, 4 e 5.

##### 4.1.1. PARÂMETROS UTILIZADOS

No primeiro método, a simulação de cada nó da malha de estimativa (*grid*) foi realizada com a utilização de mil bandas e com a mesma semente inicial para todas as variáveis, em todas as combinações.

Para a utilização do método de bandas rotativas, é recomendado que o conjunto de dados de entrada siga uma distribuição de probabilidade gaussiana, o que nem sempre é uma realidade para as variáveis medidas em depósitos minerais. É por este motivo que o primeiro passo para sua aplicação é a normalização dos dados, seguida do ajuste do modelo variográfico dos dados normalizados utilizados nas simulações futuras (figura 3.18).

A simulação direta em blocos foi realizada também pelo método de bandas rotativas seguindo a variante da abordagem via modelo gaussiano

discreto Emery e Ortiz (2005), com mil bandas e a mesma semente inicial para todas as variáveis.

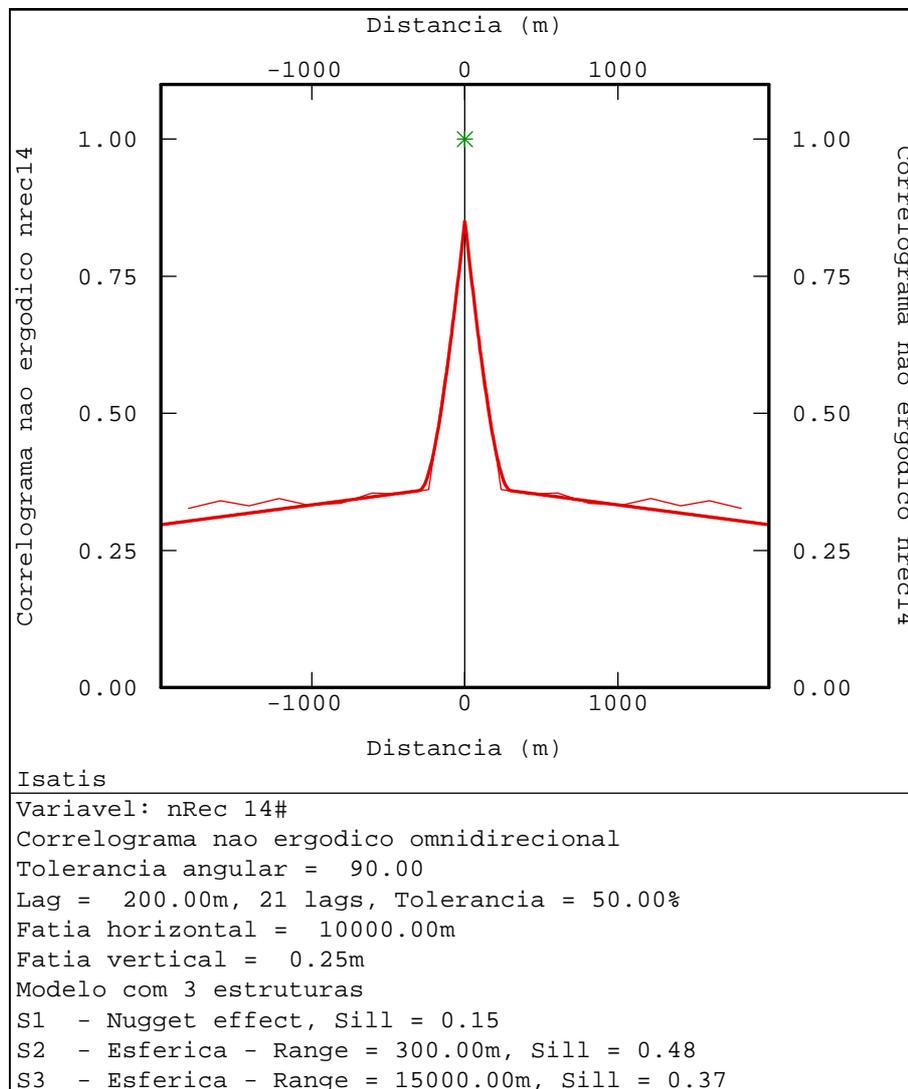


FIGURA 4.1: Correlograma omnidirecional dos dados normalizados da variável Rec14#.

Partindo dos modelos definidos, é possível que se estabeleçam alguns parâmetros utilizados na vizinhança de busca das simulações. Na verdade, as informações obtidas nos modelos variográficos serão utilizadas principalmente na elaboração do elipsoide de busca, sendo os demais parâmetros definidos pelas características do depósito simulado e após diversas tentativas de ajuste. A tabela 4.1 mostra as escolhas realizadas para este estudo.

TABELA 4.1: Parâmetros de busca utilizados nas simulações da variável Rec14# e em todas as demais simulações realizadas.

Parâmetro	Valor
Elipsóide de busca (x,y,z)	1000x1000x0,5m
Número mínimo de amostras	12
Número máximo de amostras	24
Número de nós previamente simulados	12
Número de setores angulares	4

As distâncias de 1000 metros no plano horizontal e de 0,5 metros no plano vertical foram pensadas de modo a garantir que um número maior de dados pertencentes ao plano horizontal fosse utilizado, já que este é o plano que representa a maior continuidade do depósito. O valor 1000m também foi escolhido por ser suficientemente grande para a busca por amostras. Dificilmente, amostras separadas por esta distância serão consideradas, uma vez que o número de dados disponíveis a curtas distâncias é significativo.

O elipsoide de busca foi dividido em quatro setores angulares para garantir o desagrupamento na busca. Então, para um número mínimo de doze e máximo de vinte e quatro dados condicionantes, a simulação de cada ponto ou bloco foi realizada a partir de três dados originais e de até seis valores previamente simulados por setor, respectivamente.

#### 4.1.2. VALIDAÇÃO DA SIMULAÇÃO EM PONTOS

Para a verificação da qualidade dos resultados obtidos, foram analisadas: a reprodução das características de continuidade espacial e a estatística básica da variável simulada Rec14#.

##### 4.1.2.1. REPRODUÇÃO DOS CORRELOGRAMAS

As figuras 4.2, para a análise omnidirecional, e, 4.3, para a análise vertical, mostram a adequada aderência dos correlogramas dos cenários

simulados (em preto) que flutuam em torno do correlograma experimental dos dados originais (em vermelho, linha fina) e do modelo ajustado para o mesmo (vermelho, linha grossa). Isto significa que os modelos simulados possuem as mesmas características de continuidade espacial que os dados originais amostrados do depósito.

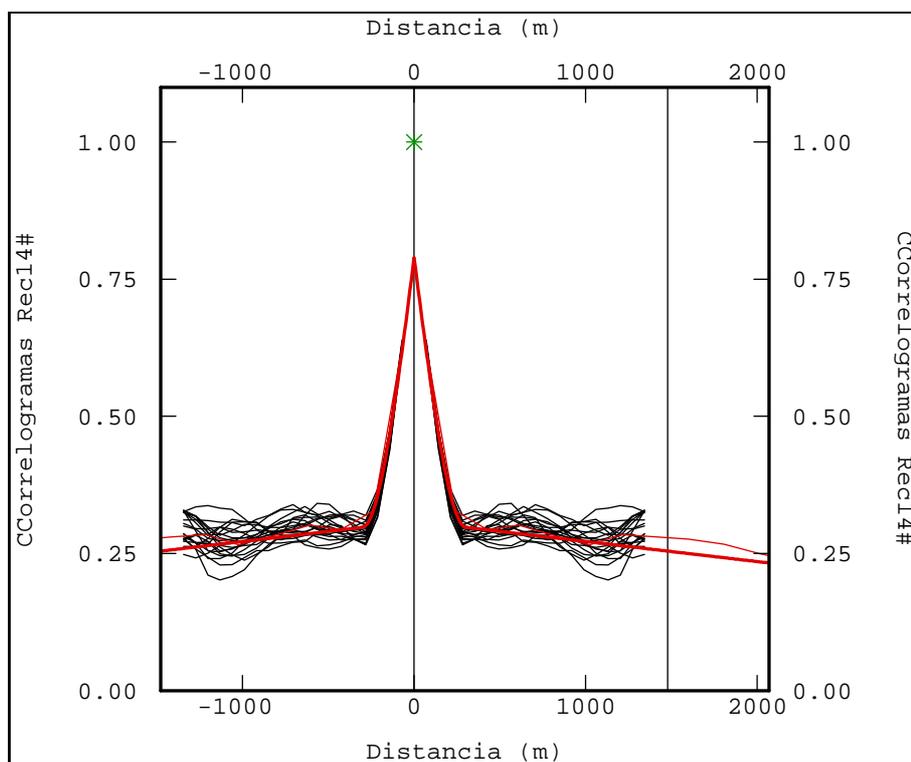


FIGURA 4.2: Correlogramas experimentais dos modelos simulados (linhas pretas); correlograma experimental (vermelho, linha fina) e correlograma modelado para os dados originais (vermelho, linha grossa). Os correlogramas experimentais dos valores simulados acompanham o correlograma experimental e modelado dos dados originais.

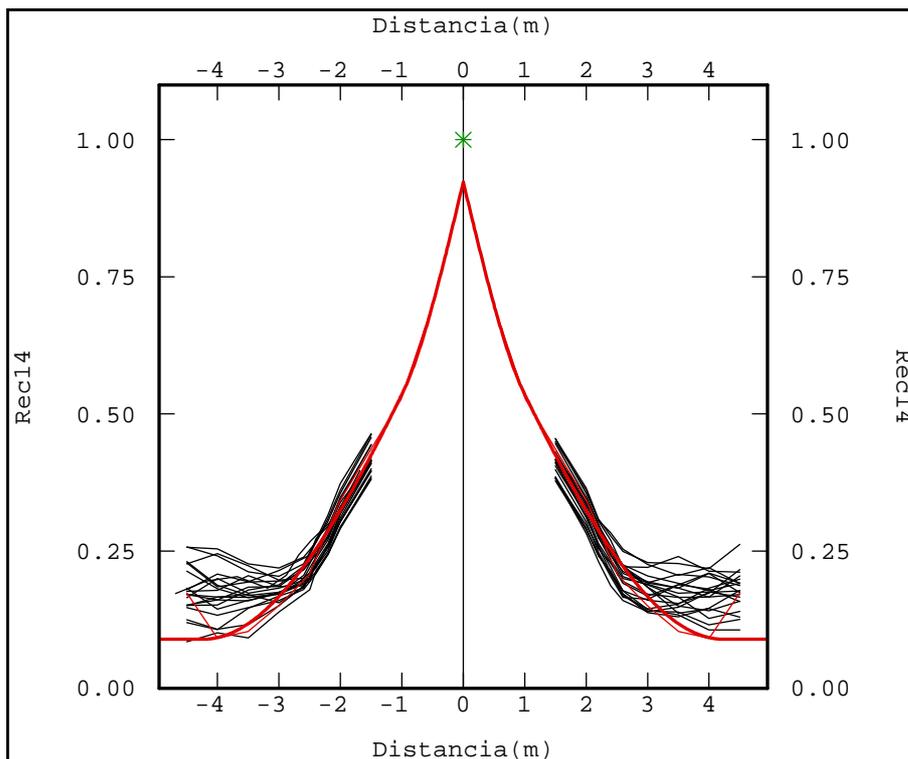


FIGURA 4.3: Correlogramas experimentais ao longo do furos para os modelos simulados (linhas pretas); correlograma experimental (vermelho, linha fina) e correlograma modelado para os dados originais (vermelho, linha grossa). Os correlogramas experimentais dos valores simulados acompanham o correlograma experimental e modelado dos dados originais.

#### 4.1.2.2. REPRODUÇÃO DAS CARACTERÍSTICAS ESTATÍSTICAS

A análise da reprodução da estatística básica das variáveis foi realizada considerando-se os seguintes parâmetros: valores mínimo e máximo e média. Na tabela 4.2, é possível verificar que os valores simulados apresentam os resultados esperados, com valores mínimos e máximos preservados, assim como a média oscilando em torno dos valores originais.

TABELA 4.2: Estatística básica dos dados originais (topo da tabela) e dos vinte cenários simulados para a variável Rec14#. Os valores simulados apresentam os resultados esperados, com valores mínimos e máximos preservados, assim como a média oscilando em torno dos valores originais.

VARIÁVEL	PONTOS/ BLOCOS	MÍNIMO(%)	MÁXIMO(%)	MÉDIA(%)	DESVIO PADRÃO (%)
<b>DADOS ORIGINAIS</b>	<b>3735</b>	<b>3,21</b>	<b>97,52</b>	<b>67,62</b>	<b>15,84</b>
REC[00001]	71420	3,21	97,52	67,62	15,24
REC[00002]	71379	3,21	97,52	67,85	15,36
REC[00003]	71453	3,21	97,52	67,48	15,24
REC[00004]	71387	3,21	97,52	67,98	15,69
REC[00005]	71401	3,21	97,52	68	15,01
REC[00006]	71338	3,21	97,52	67,85	15,02
REC[00007]	71402	3,21	97,52	67,54	15,98
REC[00008]	71397	3,21	97,52	67,39	15,42
REC[00009]	71431	3,21	97,52	67,17	16
REC[00010]	71406	3,21	97,52	67,64	14,98
REC[00011]	71361	3,21	97,52	68,05	15,54
REC[00012]	71354	3,21	97,52	67,19	15,25
REC[00013]	71354	3,21	97,52	67,95	15,69
REC[00014]	71335	3,21	97,52	67,17	15,87
REC[00015]	71371	3,21	97,52	67,65	15,23
REC[00016]	71348	3,21	97,52	68,05	15,01
REC[00017]	71315	3,21	97,52	67,25	15,03
REC[00018]	71300	3,21	97,52	67,96	15,64
REC[00019]	71339	3,21	97,52	67,75	15,81
REC[00020]	71397	3,21	97,52	67,41	15,24

#### 4.1.2.3. ANÁLISE DE DERIVA

A análise de deriva corresponde à verificação da reprodução de médias locais, sejam elas de teores ou de recuperações. Um modelo simulado adequado possuirá não somente uma média global semelhante a dos dados originais, mas também médias locais parecidas. As figuras 4.4, 4.5 e 4.6 mostram a análise de deriva para nos eixos x, y e z da variável Rec14#. Note que a aderência dos valores simulados aos dados originais é bastante satisfatória nos três eixos analisados.

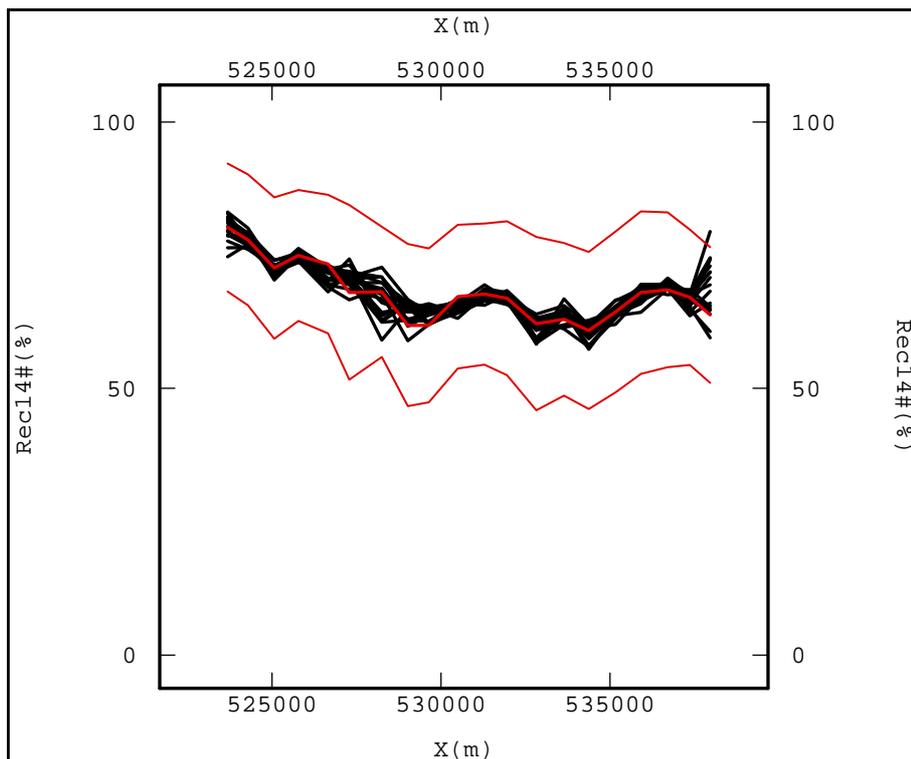


FIGURA 4.4: Análise de deriva ao longo do eixo "x". As médias simuladas da variável Rec14# (linhas pretas) seguem a mesma tendência das médias originais da variável (linha contínua vermelha). As linhas vermelhas pontilhadas representam a envoltória do desvio padrão das amostras.

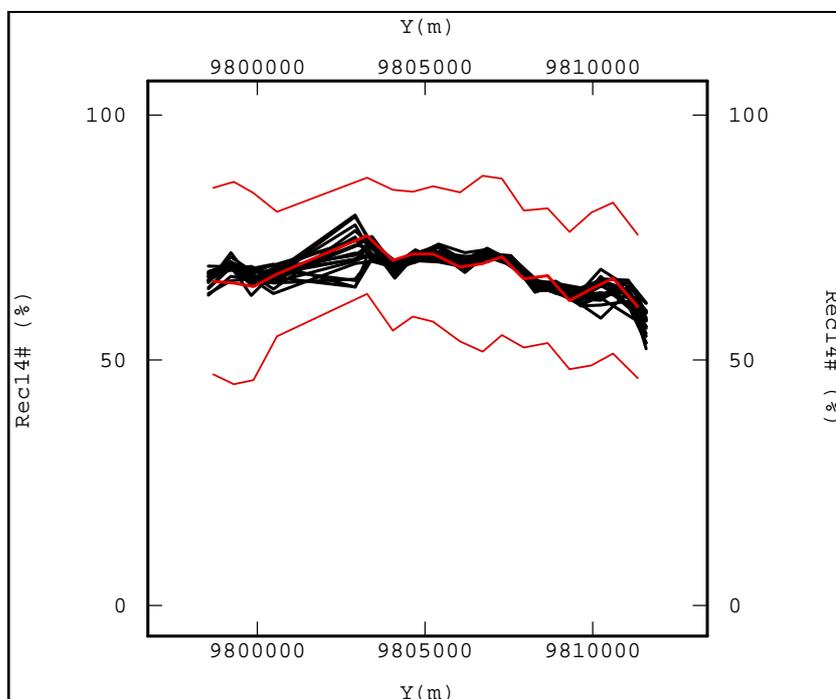


FIGURA 4.5: Análise de deriva ao longo do eixo "y". As médias simuladas da variável Rec14# (linhas pretas) seguem a mesma tendência das médias originais da variável (linha contínua vermelha). As linhas vermelhas pontilhadas representam a envoltória do desvio padrão das amostras.

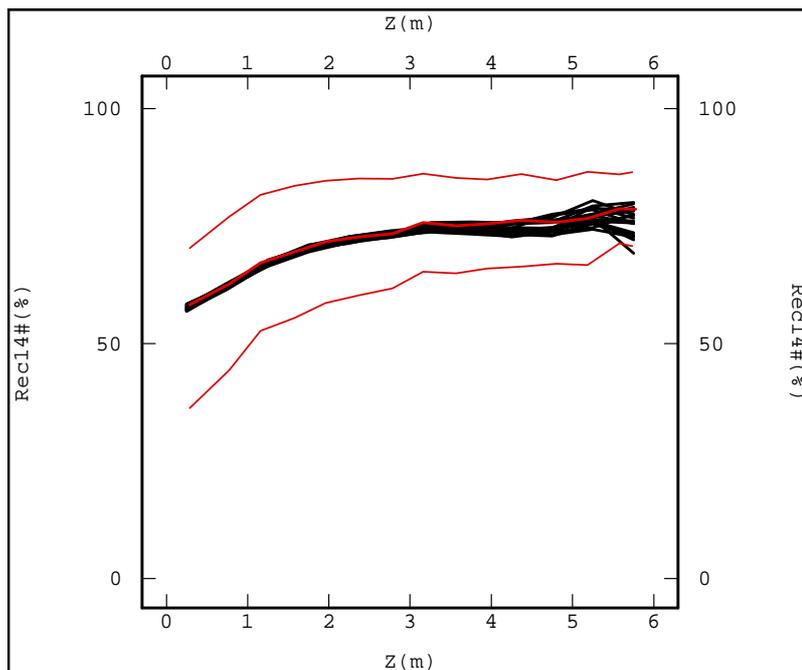


FIGURA 4.6: Análise de deriva ao longo do eixo "z". As médias simuladas da variável Rec14# (linhas pretas) seguem a mesma tendência das médias originais da variável (linha contínua vermelha) com uma leve discrepância para os valores maiores do que 3m de cota em Z por falta de dados nessa cota. As linhas vermelhas pontilhadas representam a envoltória do desvio padrão das amostras.

Considerando que a reprodução da estatística básica, dos correlogramas e médias locais foi adequada, é possível afirmar que os modelos simulados pelo método de bandas rotativas, em pontos posteriormente reblocados, representam adequadamente o comportamento da variável Rec14#.

#### 4.1.3. VALIDAÇÃO DA SIMULAÇÃO DIRETA EM BLOCOS

A validação dos resultados obtidos na simulação direta em blocos da variável Rec14# foi realizada com base na reprodução dos correlogramas, da média global dos teores, e, das médias locais obtidas para os cenários gerados quando comparadas com os dados amostrados.

##### 4.1.3.1. REPRODUÇÃO DA ESTATÍSTICA BÁSICA

A tabela 4.3 mostra a estatística básica da variável Rec14# amostrada (no topo) e dos vinte cenários simulados para a mesma. Note que as médias dos valores simulados são próximas à média dos dados originais (66,63%).

TABELA 4.3: Estatística básica dos dados originais (topo da tabela) e dos vinte cenários simulados para a variável Rec14#.

VARIÁVEL	DADOS	MÍNIMO(%)	MÁXIMO(%)	MÉDIA(%)	DESVIO PADRÃO(%)
<b>DADOS ORIGINAIS</b>	<b>3735</b>	<b>3,21</b>	<b>97,52</b>	<b>67,62</b>	<b>15,84</b>
REC14DB[00001]	64886	8,1	92,56	66,38	15,26
REC14DB[00002]	64886	9,21	92,56	64,99	15,34
REC14DB[00003]	64886	11,98	92,56	65,9	15,63
REC14DB[00004]	64886	12,48	92,56	68,45	15,24
REC14DB[00005]	64886	10,15	92,56	67,17	15,06
REC14DB[00006]	64886	12,78	92,56	68,24	15,98
REC14DB[00007]	64886	15,05	92,56	68,22	15,87
REC14DB[00008]	64886	13,84	92,56	67,73	15,84
REC14DB[00009]	64886	13,73	92,56	67,9	15,56
REC14DB[00010]	64886	10,3	92,56	67,3	15,23
REC14DB[00011]	64886	9,7	92,56	65,53	15,54
REC14DB[00012]	64886	8,39	92,56	64,9	15,26
REC14DB[00013]	64886	12,42	92,56	66,78	15,29
REC14DB[00014]	64886	11,62	92,56	65,73	15,87
REC14DB[00015]	64886	10,72	92,56	66,15	15,98
REC14DB[00016]	64886	5,66	92,56	65,01	15,32
REC14DB[00017]	64886	7,58	92,56	66,35	15,23
REC14DB[00018]	64886	9,83	92,56	66,72	15,48
REC14DB[00019]	64886	10,55	92,56	64,15	15,63
REC14DB[00020]	64886	13,01	92,56	66,56	15,78

#### 4.1.3.2. REPRODUÇÃO DOS CORRELOGRAMAS

Os correlogramas dos cenários simulados (linhas pretas) aderem adequadamente tanto aos correlogramas experimentais (linha fina vermelha) quanto aos modelados (linha grossa vermelha) para a variável Rec14#, na análise onmidirecional (figura 4.7) e vertical (figura 4.8).

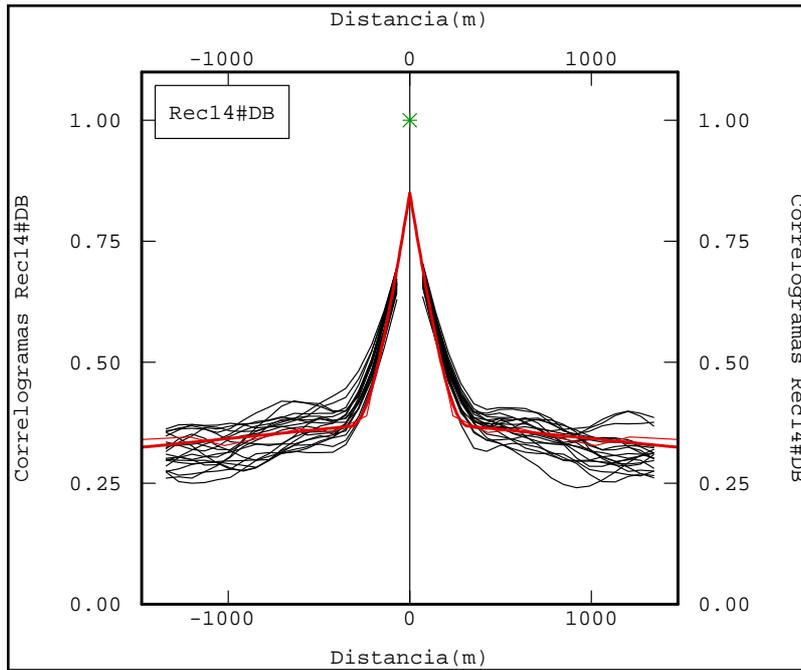


FIGURA 4.7: Correlogramas simulados (linhas pretas), experimental (linha fina vermelha) e modelado (linha grossa vermelha) para a variável Rec14#.

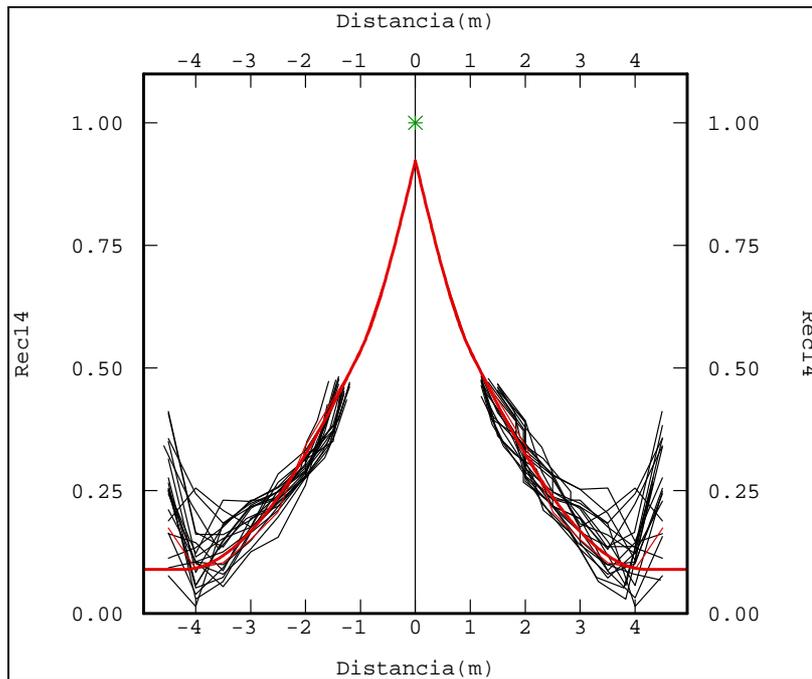


FIGURA 4.8: Correlogramas ao longo do furo simulados (linhas pretas), experimental (linha fina vermelha) e modelado (linha grossa vermelha) para a variável Rec14#.

#### 4.1.3.3. REPRODUÇÃO DAS MÉDIAS LOCAIS

As médias locais da variável Rec14# foram bem reproduzidas na simulação direta em blocos (figuras 4.7, 4.8 e 4.9). Os blocos simulados (linhas pretas) apresentam médias locais bastante semelhantes às dos dados originais (linha vermelha), acompanhando inclusive as oscilações mais bruscas das mesmas.

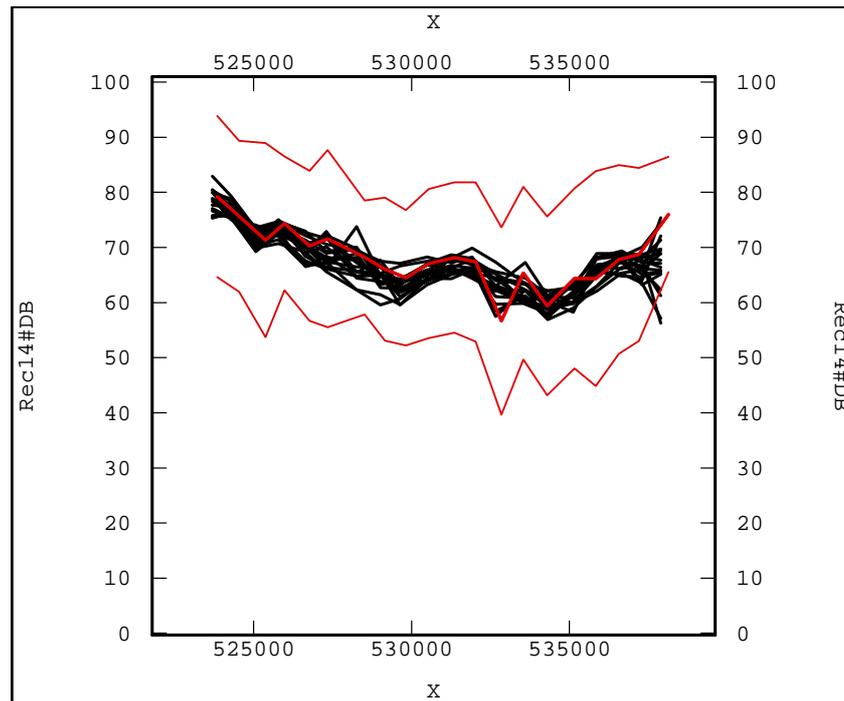


FIGURA 4.9: Análise de deriva em "x". As linhas pretas representam os blocos simulados e a linha vermelha os dados originais. Os blocos simulados (linhas pretas) apresentam médias locais bastante semelhantes às dos dados originais (linha vermelha), acompanhando inclusive as oscilações mais bruscas das mesmas. As linhas vermelhas pontilhadas representam a envoltória do desvio padrão das amostras.

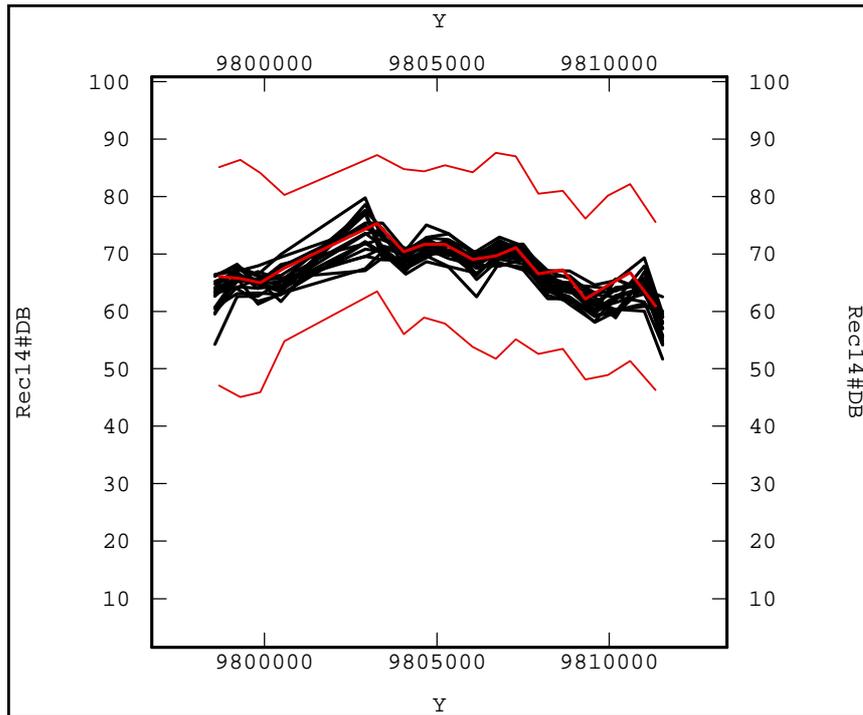


FIGURA 4.10: Análise de deriva em "y". As linhas pretas representam os blocos simulados e a linha vermelha os dados originais. Os blocos simulados (linhas pretas) apresentam médias locais bastante semelhantes às dos dados originais (linha vermelha), acompanhando inclusive as oscilações mais bruscas das mesmas. As linhas vermelhas pontilhadas representam a envoltória do desvio padrão das amostras.

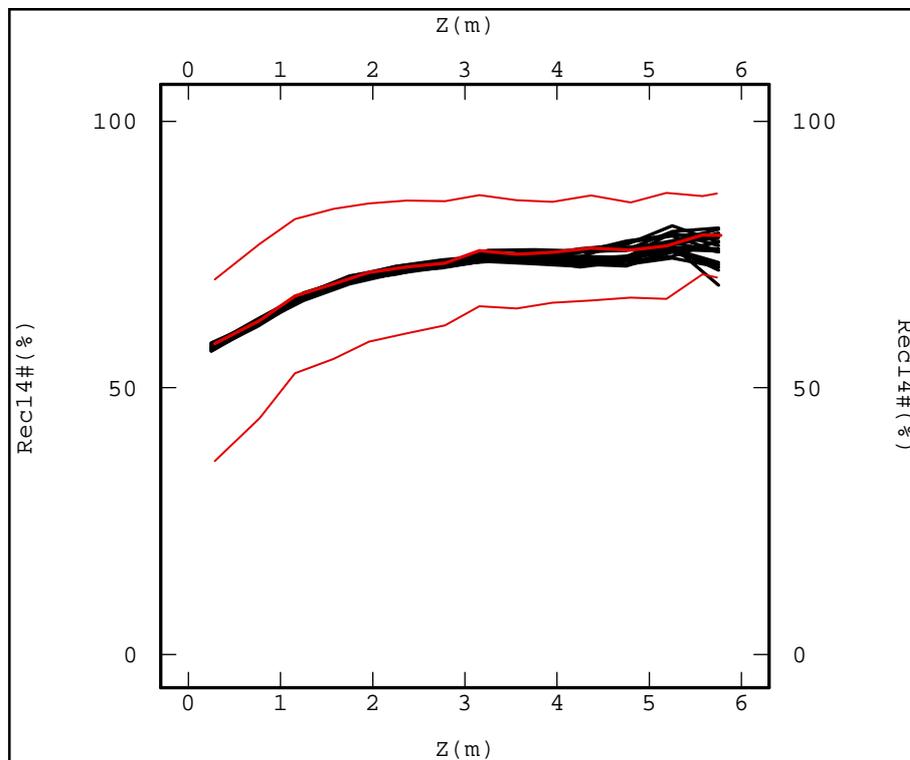


FIGURA 4.11: Análise de deriva em "z". As linhas pretas representam os blocos simulados e a linha vermelha os dados originais. Os blocos simulados (linhas pretas) apresentam médias locais bastante semelhantes às dos dados originais (linha vermelha), acompanhando inclusive as oscilações mais bruscas das mesmas.

Considerando a adequada reprodução dos modelos gerados às principais características do depósito estudado, é possível dizer que os cenários simulados diretamente em suporte de blocos são adequados para a posterior utilização na desacumulação das variáveis acumuladas simuladas.

#### 4.2. SIMULAÇÃO DOS DADOS ORIGINAIS

É o primeiro método testado (figura 1.2). Em um primeiro momento, pensou-se em utilizar a cossimulação dos dados originais, mas o ajuste do modelo linear de correionalização (MLC) para as cinco variáveis envolvidas se mostrou impraticável. Fato este que comprova que dificilmente o método é utilizado rotineiramente na indústria mineira, quando tratamos de depósitos com três ou mais variáveis de interesse que sejam correlacionadas. Como a simulação individual dos atributos parece ser mais próxima e facilmente implementável no dia a dia do planejamento de longo prazo, esta foi escolhida para utilização nesta combinação que envolveu os seguintes passos:

- i. Acumulação das variáveis analisadas (teor da amostraRec14# da amostra);
- ii. Normalização dos dados;
- iii. Cálculo e ajuste dos correlogramas individuais;
- iv. Simulação individual em pontos, com uso do método de bandas rotativas:
  - a. Malha de 10x10x0,5m;
  - b. Mesmo modelo e parâmetros ajustados para a variável Rec14# normalizada.
- v. Retrotransformação dos valores simulados normalizados;
- vi. Desacumulação (teor acumulado simulado/recuperação simulada);
- vii. Validação dos cenários gerados frente aos dados originais da variável;
- viii. Reblocação para um *malha de estimativa (grid)* de 50x50x0,5m
- ix. Verificação da reprodução das correlações entre as variáveis.
- x. Verificação da reprodução da soma dos teores nos blocos simulados.

Os pontos (i), (ii) e (iii) estão no terceiro capítulo, a validação das simulações frente aos dados originais está no anexo A e os pontos (ix) e (x) fazem parte do presente capítulo.

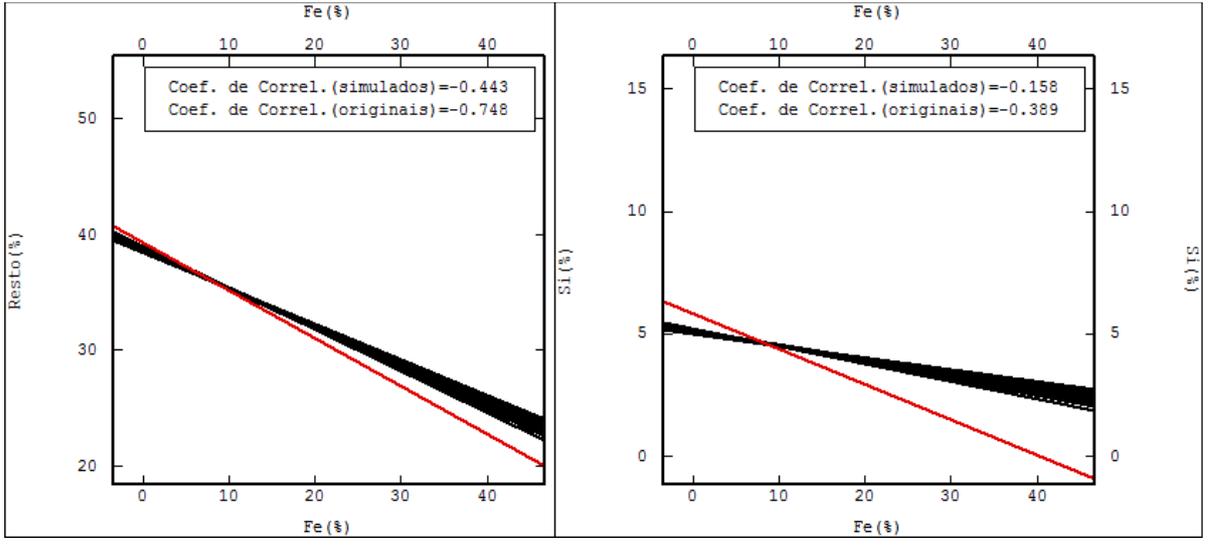
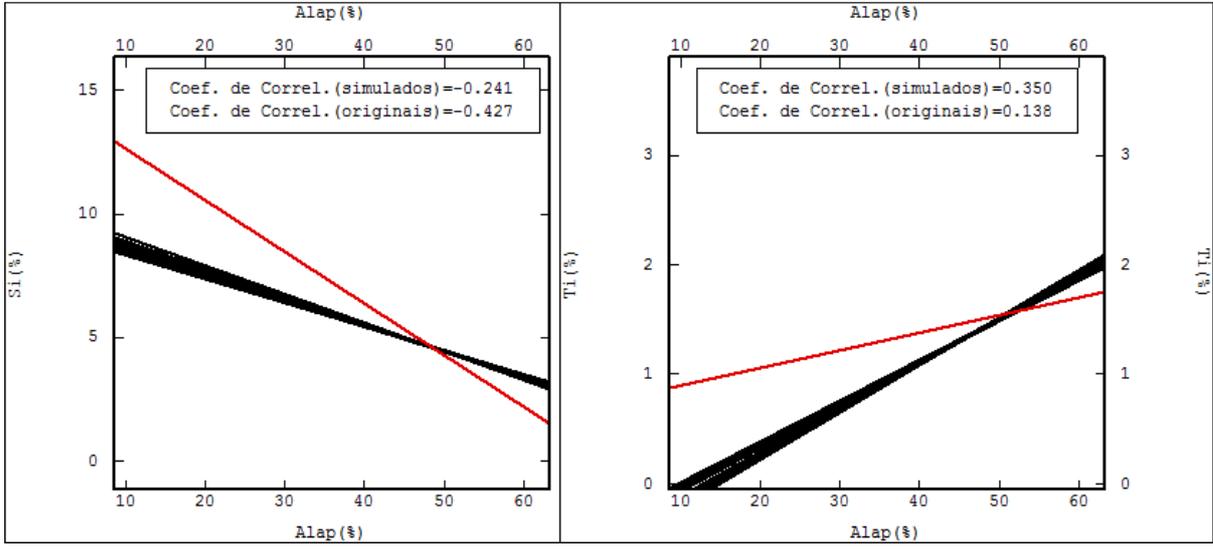
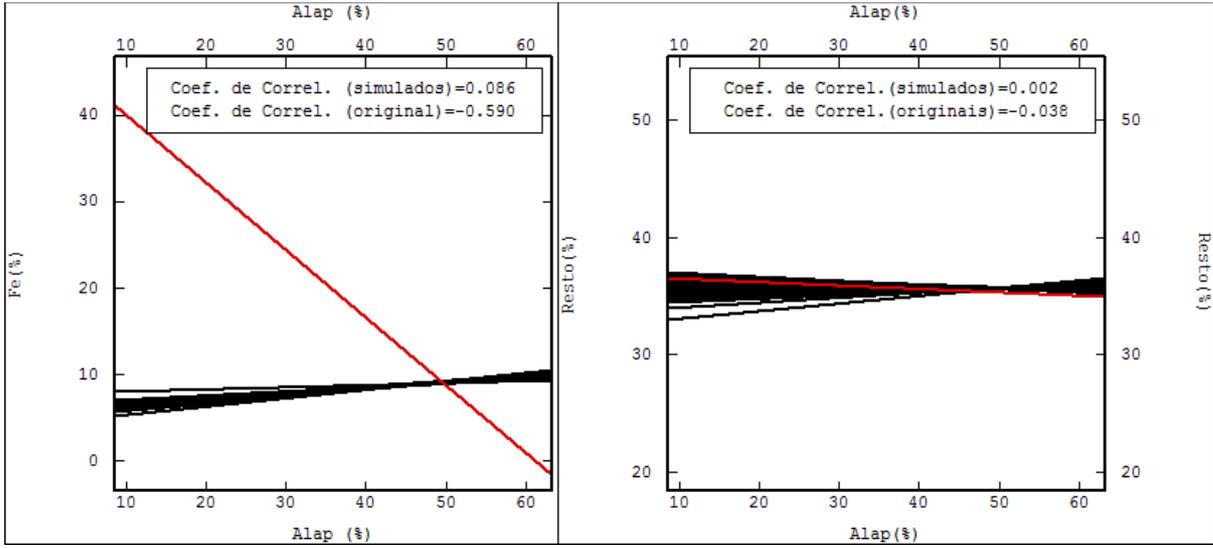
#### 4.2.1. CORRELAÇÕES

Quando variáveis pertencentes a um depósito apresentam comportamentos espaciais parecidos, com zonas de altos e baixos teores coincidentes, é possível dizer que estas são correlacionadas positivamente. Quando o comportamento de uma é justamente o inverso do da outra, zonas de alto teor para uma equivalendo a zonas de baixo teor para a outra, estas também se correlacionam, mas esta correlação será dita negativa.

Se, dois ou mais atributos analisados possuírem correlações significativas verificadas nos dados originais, este comportamento também deve ser reproduzido pelos cenários simulados. Também é verdade que variáveis que não apresentem correlação verificada pelo comportamento das amostras, não devem exibir correlações nos modelos simulados.

Na figura 4.10, é possível observar as correlações lineares entre as variáveis originais (linhas vermelhas) e entre as variáveis simuladas (linhas pretas). O coeficiente de correlação exibido para os valores simulados representa a média aritmética simples dos vinte cenários simulados. As correlações  $AlapxResto$ ,  $SixResto$  e  $TixResto$  foram perfeitamente reproduzidas nos cenários simulados. No entanto, as correlações  $AlapxSi$ ,  $AlapxTi$ ,  $FexResto$ ,  $FexSi$  e  $SixTi$ ,  $AlapxFE$  e  $FexTi$  possuem correlações distintas quando comparamos os resultados dos dados originais com os obtidos nos cenários gerados.

A falha na reprodução das correlações pode ser justificada pelo fato de cada uma das variáveis de interesse ter sido simulada de modo independente das demais. Este problema poderia ter sido evitado se em vez da simulação individual, se tivesse aplicado o método de cossimulação com utilização do Modelo de Corregionalização Intrínseco (MCI).



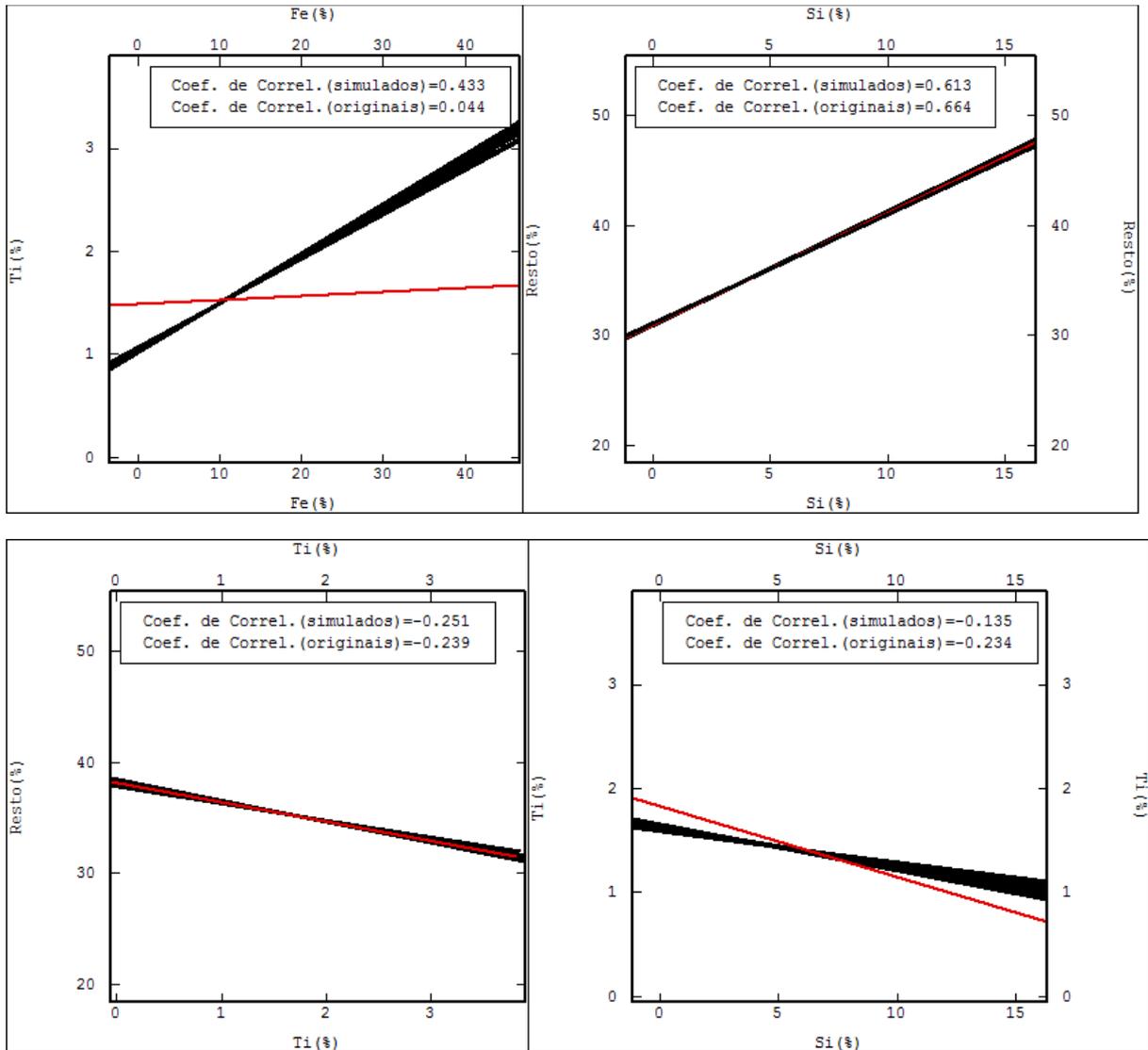


FIGURA 4.12: Correlações lineares entre as variáveis originais (linhas vermelhas) e cenários simulados (linhas pretas). As correlações  $Alap \times Resto$ ,  $Six \times Resto$  e  $Tix \times Resto$  foram perfeitamente reproduzidas nos cenários simulados. As correlações  $Alap \times Si$ ,  $Alap \times Ti$ ,  $Fex \times Resto$ ,  $Fex \times Si$  e  $Six \times Ti$ , apesar de não apresentarem uma reprodução perfeita apresentam um comportamento próximo ao das correlações entre os dados originais. Já  $Alap \times Fe$  e  $Fex \times Ti$  possuem correlações bastante distintas sob o olhar dos dados originais e dos cenários obtidos.

#### 4.2.2. FECHAMENTO

Outro item importante de ser analisado é o fechamento das somas dos teores simulados para cada bloco frente à soma dos teores originais em cada amostra ou faixa granulométrica, que também deve ser reproduzido.

A tabela 4.4 contém informações sobre o número de blocos estimados dentro do intervalo permitido (entre 95% e 105%) e a média da soma dos teores simulados para a faixa 14#. O intervalo definido como permitido foi pensado

para este estudo, considerando como aceitável um erro de até 5% para mais ou para menos na soma dos teores. Em nenhum dos cenários simulados, neste método, foram encontrados blocos com valores de fechamento superiores ou inferiores a este intervalo.

Analisando os dados exibidos, é possível perceber que apesar de a soma dos teores simulados não ser exatamente igual à soma dos teores dos dados originais (100%), os valores obtidos são bastante próximos e suficientes para uma análise preliminar.

Tabela 4.4: Informações sobre o número de blocos estimados dentro do intervalo permitido (entre 95% e 100%) e a média da soma dos teores simulados para a faixa 14#. É possível perceber que apesar de a soma dos teores simulados não ser exatamente igual à soma dos teores dos dados originais (100%), os valores obtidos são bastante próximos e suficientes para uma análise preliminar.

<i>Variável</i>	<i>Número de blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>
<i>soma[00001]</i>	64886	95	105	99,63
<i>soma[00002]</i>	64886	95	105	99,69
<i>soma[00003]</i>	64886	95	105	99,63
<i>soma[00004]</i>	64886	95	105	99,77
<i>soma[00005]</i>	64886	95	105	99,78
<i>soma[00006]</i>	64886	95	105	99,7
<i>soma[00007]</i>	64886	95	105	99,78
<i>soma[00008]</i>	64886	95	105	99,59
<i>soma[00009]</i>	64886	95	105	99,6
<i>soma[00010]</i>	64886	95	105	99,74
<i>soma[00011]</i>	64886	95	105	99,77
<i>soma[00012]</i>	64886	95	105	99,66
<i>soma[00013]</i>	64886	95	105	99,78
<i>soma[00014]</i>	64886	95	105	99,69
<i>soma[00015]</i>	64886	95	105	99,72
<i>soma[00016]</i>	64886	95	105	99,82
<i>soma[00017]</i>	64886	95	105	99,7
<i>soma[00018]</i>	64886	95	105	99,78
<i>soma[00019]</i>	64886	95	105	99,69
<i>soma[00020]</i>	64886	95	105	99,65

#### 4.2.3. ANÁLISE DO MÉTODO

A análise do método foi feita considerando três fatores principais: tempo dispendido, facilidade de aplicação e qualidade dos resultados obtidos.

Em relação ao fator tempo pode-se dizer que o método de simulação em pontos, quando realizado de modo individual e envolvendo mais do que três variáveis a serem estudadas é bastante moroso. Ainda mais quando a simulação for realizada primeiramente em suporte de pontos que, em um segundo momento, precisam ser reblocados. O ajuste dos modelos variográficos individuais, o processo de simulação em si repetido seis vezes (uma para cada variável e a Rec14#, utilizada na desacumulação) e o processo de reblocagem tomam um tempo considerável, o que prejudica a aplicação do método no cotidiano das empresas de mineração.

Quanto ao grau de facilidade de implementação, é justo dizer que o método aqui aplicado é absolutamente viável, principalmente por ser razoavelmente conhecido e envolver poucos processos de transformação de variáveis.

Olhando mais atentamente para a qualidade dos resultados obtidos, pode-se concluir que o fechamento da soma dos teores nos blocos simulados é satisfatório. Se considerarmos que as simulações, por serem independentes, não consideraram em nenhum momento as relações entre as variáveis e ainda, que não houve a utilização de nenhum método de restrição para os valores obtidos nos modelos, o método produz resultados que poderiam de fato ser utilizados nas situações em que haja um certo grau de tolerância sobre os resultados obtidos em termos de correlações e fechamento de teores.

#### 4.3. SIMULAÇÃO DAS ILRS

A proposta da segunda combinação testada (figura 1.3) é considerar os dados originais como composições pertencentes ao Simplex de D dimensões, trazê-los para um subespaço equivalente do Real de D-1 dimensões por meio da transformação logarítmica isométrica (ilr), simular as ilrs e retornar para o espaço restrito. Com isso, espera-se principalmente a obtenção de resultados

que obedecem à condição de soma constante dos teores da faixa estudada (100%), adicionando o benefício da diminuição do número de variáveis a serem simuladas em uma unidade.

O método foi aplicado seguindo os seguintes passos:

- i. Transformação ilr;
- ii. Acumulação das ilrs (ilr X Rec14#);
- iii. Normalização dos dados;
- iv. Cálculo e ajuste dos correlogramas individuais;
- v. Simulação individual em pontos pelo método de bandas rotativas:
  - a. Malha de 10x10x0,5;
  - b. Mesmo modelo e parâmetros ajustados para a variável Rec14# normalizada;
- vi. Retrotransformação dos ilrs acumulados simulados normalizados para não normalizados acumulados;
- vii. Desacumulação (ilr acumulada simulada/recuperação simulada);
- viii. Transformação igl (retorno do Real para o Simplex);
- ix. Validação dos cenários gerados frente aos dados originais da variável;
- x. Reblocagem para um *malha de estimativa (grid)* de 50x50x0,5m.
- xi. Verificação da reprodução das correlações entre as variáveis;
- xii. Verificação da reprodução da soma dos teores nos blocos simulados.

O item (ix.) está no Anexo A e os itens (xi.) e (xii.) são analisados a seguir.

#### 4.3.1. CORRELAÇÕES

As correlações lineares entre as variáveis originais (linhas vermelhas) são confrontadas com as correlações entre as variáveis simuladas (linhas pretas) na figura 4.11. Assim como na primeira combinação, aqui, o coeficiente de correlação referente aos valores simulados também representa a média aritmética simples dos vinte cenários simulados.

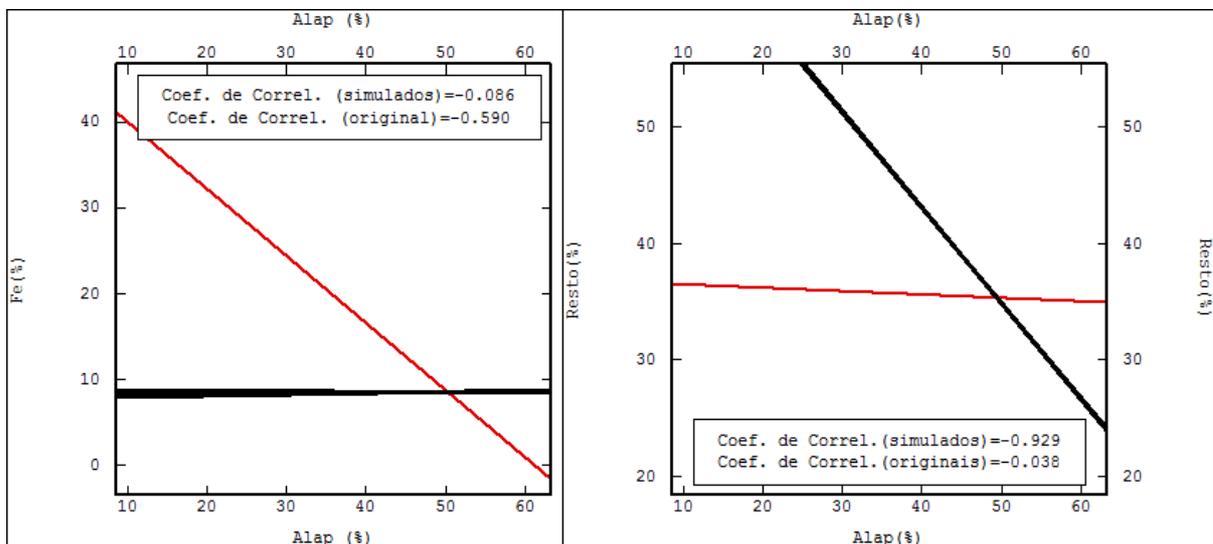
Observando os gráficos, é possível perceber que as relações RestoxFe e FexSi são mantidas quase que perfeitamente após as simulações. Já AlapxSi, RestoxSi, SixTi, AlapxFe, AlapxResto, AlapxTi, FexTi e RestoxTi apresentam

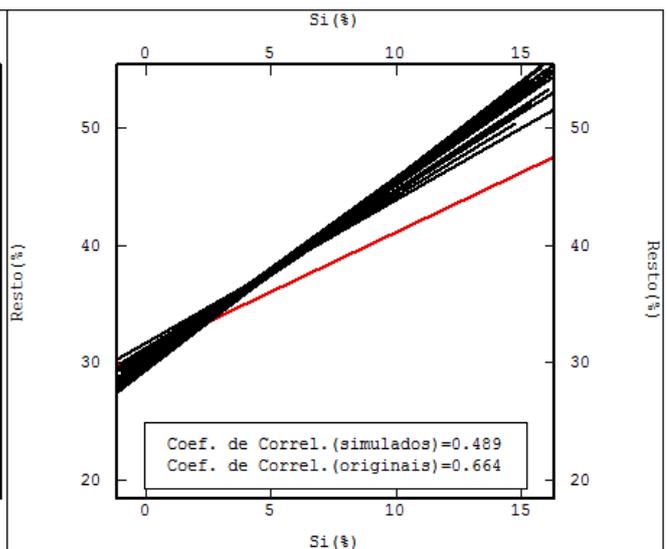
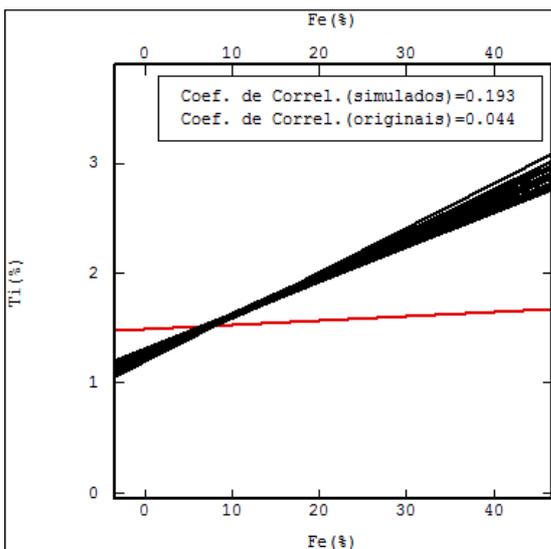
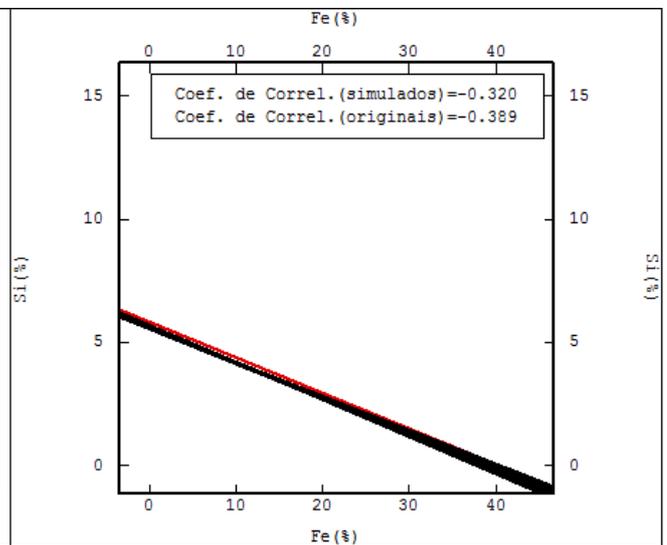
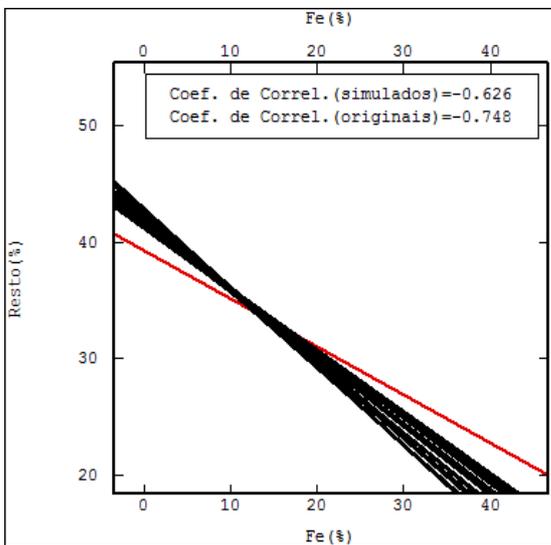
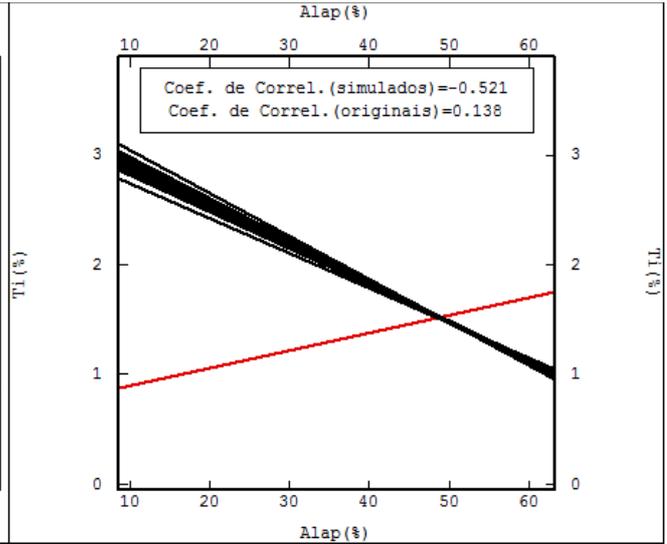
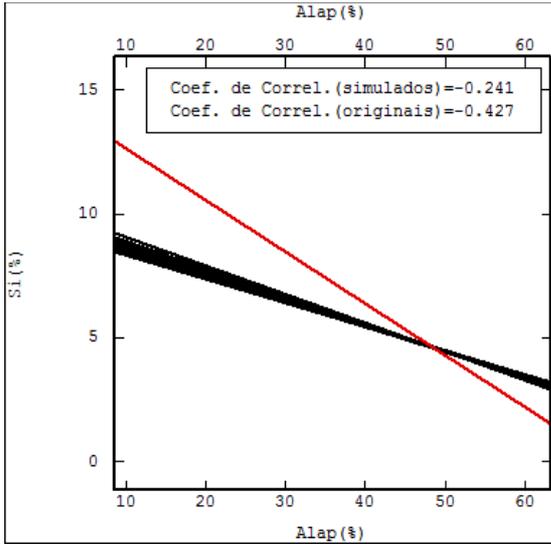
comportamento significativamente distinto na comparação entre as correlações originais e as simuladas.

Considerando o acima exposto, podemos dizer que a utilização da transformação ilr seguida da simulação individual dos irls calculados implica na não reprodução das correlações entre as variáveis.

Alguns fatores contribuíram para que isto acontecesse:

- i. As ilrs foram simuladas de modo independente, sem a consideração conjunta das correlações cruzadas entre as variáveis;
- ii. Para garantir que o processo de desacumulação fosse eficiente, na alimentação do sistema de simulação, foram utilizados o modelo e os fatores inicialmente ajustados para a variável Rec14 normalizada;
- iii. Nesta combinação de métodos, esteve presente a variável artificial Resto, que se correlaciona com qualquer uma das demais de forma irreal, uma vez que esta não existe de fato (correlação espúria).





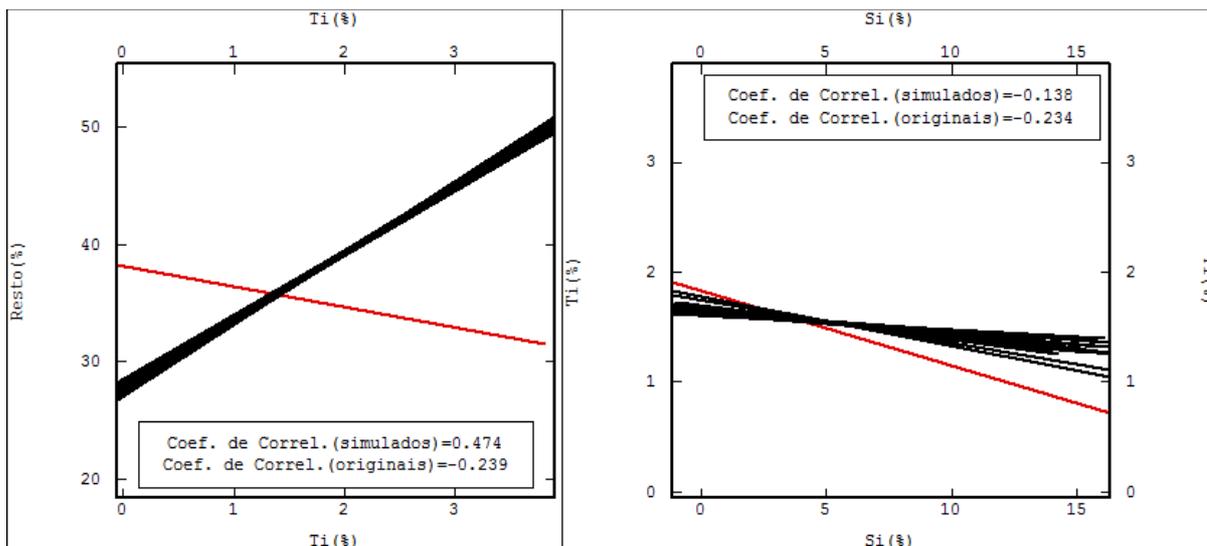


FIGURA 4.13: Correlações lineares entre as variáveis originais (linhas vermelhas) e cenários simulados (linhas pretas). As correlações AlapxResto, SixResto e TixResto foram perfeitamente reproduzidas nos cenários simulados. As correlações AlapxSi, AlapxTi, FexResto, FexSi e SixTi, apesar de não apresentarem uma reprodução perfeita apresentam um comportamento próximo ao das correlações entre os dados originais. Já AlapxFe e FexTi possuem correlações bastante distintas sob o olhar dos dados originais e dos cenários obtidos.

#### 4.3.2. FECHAMENTO

O segundo item a ser analisado é o fechamento das somas dos teores nos blocos simulados com a utilização da transformação ilr, quando comparados com a soma dos teores originais.

Observe na tabela 4.5 que todos os blocos simulados, para todas as rodadas, possuem soma única e constante (100%). Este fechamento já era esperado uma vez que as amostras foram consideradas como partes de um conjunto de composições pertencentes a um espaço restrito de soma constante, o Simplex.

TABELA 4.5: Informações sobre o número de blocos estimados dentro do intervalo permitido (entre 95% e 100%) e a média da soma dos teores simulados para a faixa 14#. É possível perceber que a soma dos teores simulados é exatamente igual à soma dos teores dos dados originais (100%).

<i>Variável</i>	<i>Número de blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>
<i>soma[00001]</i>	69484	100	100	100
<i>soma[00002]</i>	69484	100	100	100
<i>soma[00003]</i>	69484	100	100	100
<i>soma[00004]</i>	69484	100	100	100
<i>soma[00005]</i>	69484	100	100	100
<i>soma[00006]</i>	69484	100	100	100
<i>soma[00007]</i>	69484	100	100	100
<i>soma[00008]</i>	69484	100	100	100
<i>soma[00009]</i>	69484	100	100	100
<i>soma[00010]</i>	69484	100	100	100
<i>soma[00011]</i>	69484	100	100	100
<i>soma[00012]</i>	69484	100	100	100
<i>soma[00013]</i>	69484	100	100	100
<i>soma[00014]</i>	69484	100	100	100
<i>soma[00015]</i>	69484	100	100	100
<i>soma[00016]</i>	69484	100	100	100
<i>soma[00017]</i>	69484	100	100	100
<i>soma[00018]</i>	69484	100	100	100
<i>soma[00019]</i>	69484	100	100	100
<i>soma[00020]</i>	69484	100	100	100

#### 4.3.3. ANÁLISE DO MÉTODO

Na análise dos resultados obtidos com a aplicação da simulação individual das ilrs, é claramente perceptível que a grande vantagem da aplicação do método está na reprodução e garantia da soma correta dos teores dos blocos ou faixas granulométricas simuladas. É válido ressaltar que, para que o método possa ser aplicado, os dados de entrada devem possuir soma constante, o que dificilmente ocorre em dados amostrados em depósitos minerais. No entanto, este problema é facilmente resolvido com a inclusão de uma variável artificial ao banco de dados ou com a utilização da operação de fechamento.

Ao mesmo tempo em que a soma dos teores é perfeitamente mantida nos cenários simulados, não podemos fazer a mesma afirmação quando nos referirmos à qualidade das correlações entre as variáveis simuladas frente às

originais. Apesar de algumas das relações terem se mantido próximas às reais, o método acaba por não ser o mais indicado para os casos em que a reprodução das correlações precise ser perfeitamente garantida no final do processo.

Quanto ao tempo dispendido e o grau de facilidade de aplicação da combinação de métodos é possível dizer que o ponto positivo é a utilização da transformação ilr, que implica na diminuição de uma variável a ser simulada, fazendo com que o total de simulações individuais realizadas passe a ser cinco (quatro variáveis mais a recuperação da faixa de interesse). O que desfavorece a aplicação do método é, assim como para a primeira combinação estudada, a simulação em pontos que, ao final do processo, precisam ser transformados para o suporte de blocos. A utilização do método também é perfeitamente viável quando o usuário está atento à ordem em que os procedimentos de acumulação, transformação, normalização e retro transformação precisam ser executados.

#### 4.4. SIMULAÇÃO DIRETA EM BLOCOS DOS FATORES MAF

A terceira combinação testada é a simulação direta em blocos dos fatores MAF (figura 1.4). A transformação dos dados originais em fatores de autocorrelação mínimos e máximos foi incluída por permitir que cada variável possa ser estimada de modo independente das demais, mantendo as correlações existentes entre os dados originais, também nos modelos gerados. Já, a simulação direta em blocos foi utilizada com o objetivo de aumentar a eficiência computacional envolvida nos processos de simulação, sem perda de qualidade nos resultados.

O método foi aplicado seguindo os seguintes passos:

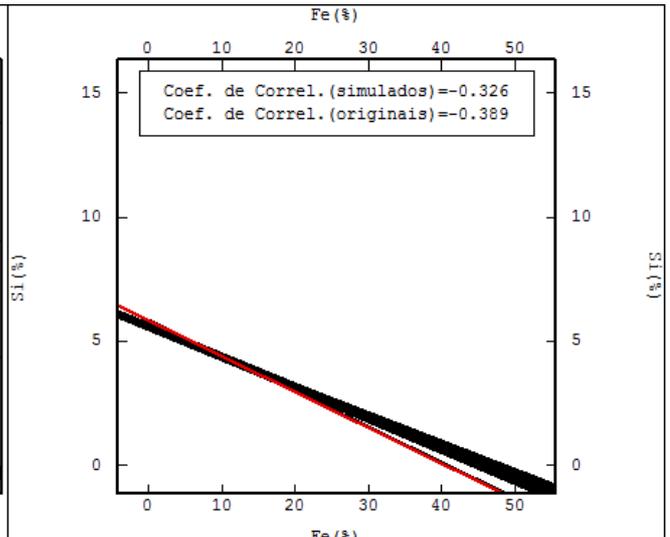
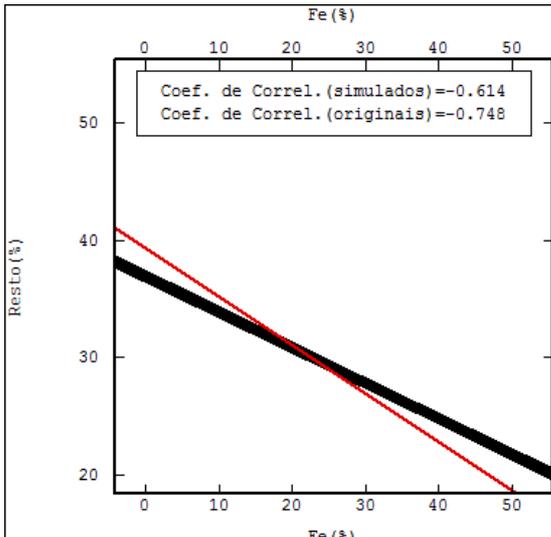
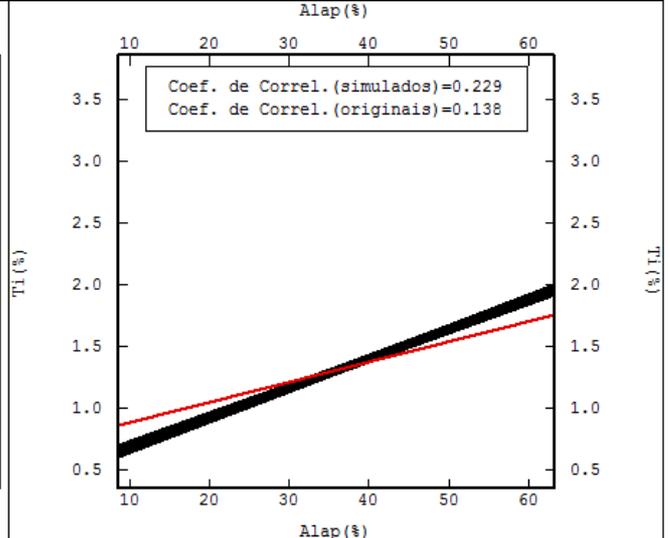
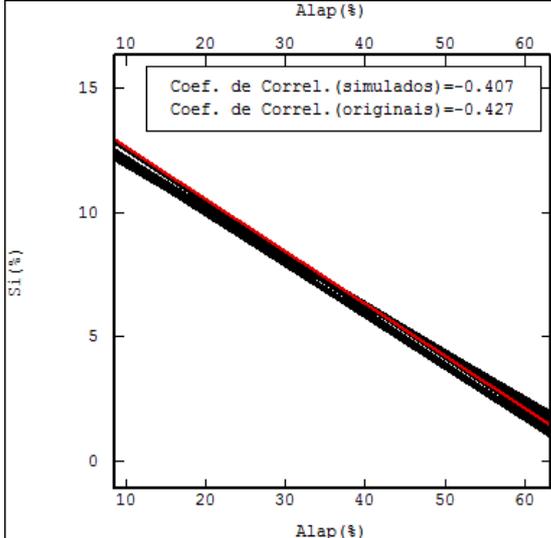
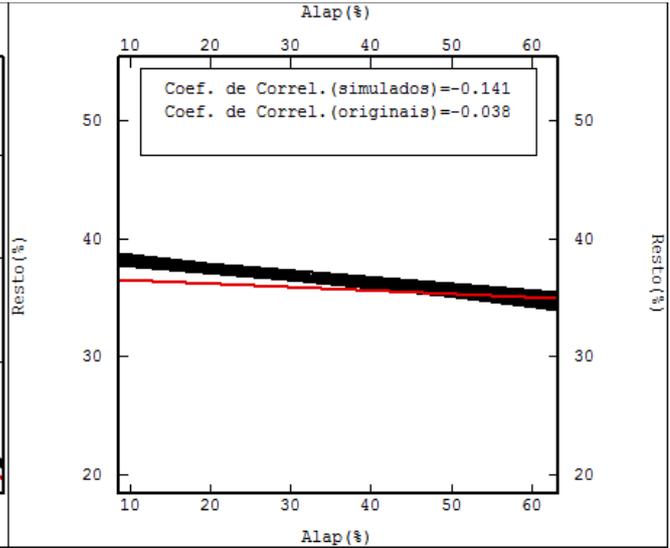
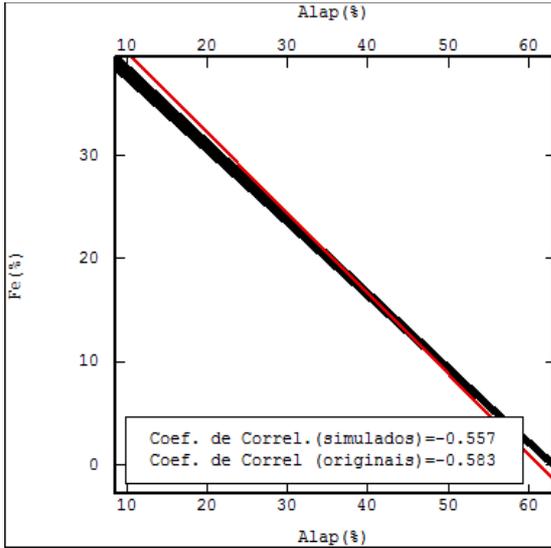
- i. Acumulação das variáveis originais;
- ii. Ortogonalização dos dados originais via transformação MAF;
- iii. Normalização dos MAFs;
- iv. Simulação direta em blocos dos fatores MAF:
  - a. Malha de 50x50x0,5m;
  - b. Mesmo modelo e parâmetros ajustados para a variável Rec14# normalizada;
- v. Retrotransformação dos fatores MAF normalizados simulados;

- vi. Operação MAF inversa, trazendo os MAFs simulados para o espaço original, ainda acumulados;
- vii. Desacumulação (variável simulada acumulada/recuperação simulada);
- viii. Validação dos cenários simulados frente às variáveis originais (anexo A);
- ix. Verificação da reprodução das correlações entre as variáveis (item 4.4.1.);
- x. Verificação da reprodução da soma dos teores nos blocos simulados (item 4.4.2.).

#### 4.4.1. CORRELAÇÕES

Na figura 4.12, é possível observar as correlações lineares entre as variáveis originais (vermelho) e os cenários simulados (preto). Assim como para as situações anteriores, o coeficiente de correlação referente aos valores simulados representa a média aritmética simples dos vinte cenários simulados. Note que, a grande maioria das correlações foi quase que perfeitamente reproduzida. Somente quando observamos RestoxSi percebemos resultados simulados um pouco destoantes do original.

Estes resultados já eram esperados, uma vez que, por definição, a decomposição MAF garante a reprodução das correlações entre variáveis após a retrotransformação da estimativa/simulação. Partindo de dados correlacionados, os ortogonalizando, e, posteriormente retornando para a base original, o método possibilita a estimativa/simulação individual dos dados independentes, seguida do retorno para o espaço original correlacionando duas a duas.



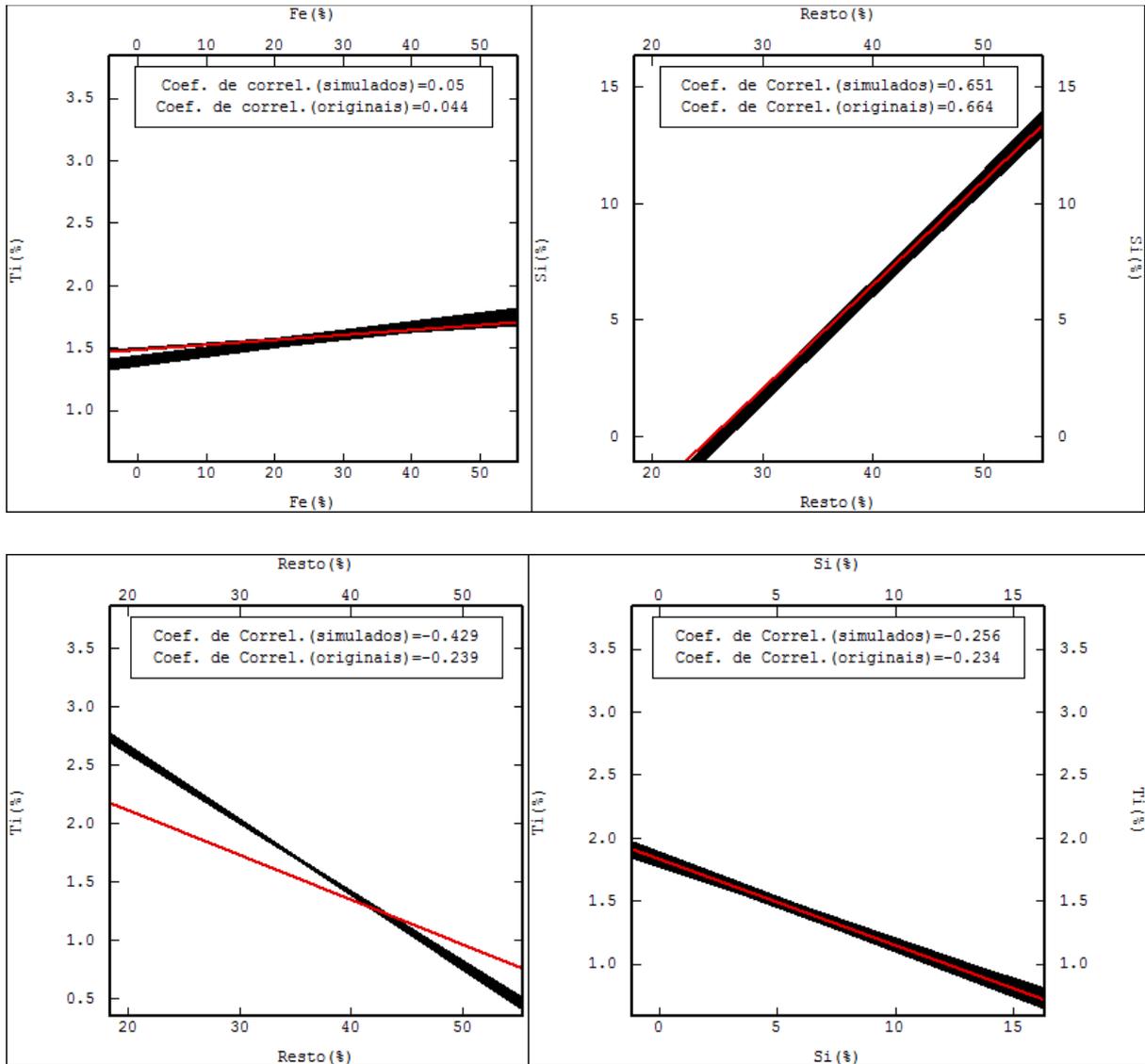


FIGURA 4.14: Correlações lineares entre as variáveis originais (linhas vermelhas) e cenários simulados (linhas pretas). Note que, a grande maioria das correlações foi quase que perfeitamente reproduzida. Somente quando olhamos Resto vs Ti percebemos resultados simulados um pouco destoantes do original.

#### 4.4.2. FECHAMENTO

Analisando a reprodução do fechamento da soma dos teores nos blocos simulados, na tabela 4.6, podemos perceber que os valores obtidos nos cenários gerados são satisfatórios, no entanto, é importante salientar que, em todas as simulações realizadas, um percentual de aproximadamente 25% dos blocos estimados apresentou valores de soma, ou inferiores a 95%, ou superiores a 105%.

Como solução, foi realizado um ajuste nos valores simulados para a variável artificial Resto, ou seja, resultados cujas somas dos teores eram inferiores a 95% foram trazidos para este patamar com adição do resíduo no valor simulado para o "Resto" e valores de soma superiores também foram ajustados para, no máximo, 105%. A escolha da variável Resto se deu pelo fato de ela não existir fisicamente e somente ter sido criada como forma de garantir o fechamento em 100%. Na comparação das médias globais antes e após o ajuste, pode-se perceber somente uma pequena variação nos valores calculados, da ordem de casas decimais, portanto, o ajuste neste contexto, não alterou significativamente os resultados das simulações para o depósito como um todo, mas atingiu um percentual relevante de blocos.

TABELA 4.6: Informações sobre o número de blocos estimados dentro do intervalo permitido (entre 95% e 100%), após o ajuste da variável Resto, e, a média da soma dos teores simulados para a faixa 14#. É possível perceber que apesar de a soma dos teores simulados não ser exatamente igual à soma dos teores dos dados originais (100%), os valores obtidos são bastante próximos e suficientes para uma análise preliminar.

<i>Variável</i>	<i>Número de blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>
<i>soma[00001]</i>	64749	95	105	99,73
<i>soma[00002]</i>	64843	95	105	100,39
<i>soma[00003]</i>	64841	95	105	99,94
<i>soma[00004]</i>	64761	95	105	98,65
<i>soma[00005]</i>	64800	95	105	99,33
<i>soma[00006]</i>	64773	95	105	98,75
<i>soma[00007]</i>	64845	95	105	98,76
<i>soma[00008]</i>	64769	95	105	99,03
<i>soma[00009]</i>	64744	95	105	98,94
<i>soma[00010]</i>	64753	95	105	99,24
<i>soma[00011]</i>	64789	95	105	100,17
<i>soma[00012]</i>	64871	95	105	100,42
<i>soma[00013]</i>	64816	95	105	99,51
<i>soma[00014]</i>	64842	95	105	100,07
<i>soma[00015]</i>	64835	95	105	99,82
<i>soma[00016]</i>	64848	95	105	100,43
<i>soma[00017]</i>	64787	95	105	99,76
<i>soma[00018]</i>	64405	95	105	99,4
<i>soma[00019]</i>	64854	95	105	100,84
<i>soma[00020]</i>	64771	95	105	99,69

#### 4.4.3. ANÁLISE DO MÉTODO

De acordo com os resultados obtidos, é possível afirmar que a utilização do método seria indicada para situações em que o maior objetivo é a reprodução das correlações entre as variáveis. A possibilidade de simular cada variável de modo independente e mesmo assim garantir a correlação entre elas nos resultados pôde ser claramente percebida por meio da reprodução quase que perfeita das correlações existentes entre as variáveis amostradas nos cenários gerados por simulação.

A proposta também se mostrou eficiente na rapidez de implementação, uma vez que a simulação diretamente em suporte de blocos não precisa do posterior processo de ajuste de suporte de pontos para blocos. Contribui também para a velocidade de implementação o fato de a simulação em si operar com um número significativamente menor de dados condicionantes, uma vez que a simulação dos valores de um novo bloco será realizada de acordo com a interação ponto/bloco e bloco/bloco.

Embora o método seja válido pelos motivos citados acima e apesar de resultar em somas de teores satisfatoriamente próximas a 100%, a condição de soma constante dos teores simulados na faixa estudada não foi atendida. Os resultados apresentados permitem dizer que um número razoável de blocos apresentou soma de teores fora do intervalo permitido, necessitando, portanto, de ajuste posterior. Então, o método sozinho não é indicado para situações em que o fechamento a soma dos teores precise estar perfeitamente garantido, apesar de retornar resultados muito próximos aos desejados.

#### 4.5. SIMULAÇÃO DIRETA EM BLOCOS DOS FATORES MAFS DAS ILRS

A quarta combinação testada é a simulação direta em blocos dos fatores MAF das ilrs (figura 1.5). O objetivo da inclusão da decomposição MAF, assim como da simulação direta blocos, também utilizados na combinação anterior, é garantir as correlações entre as variáveis originais nos cenários simulados e garantir a rapidez computacional do processo de simulação em si. Já a transformação ilr foi adicionada para garantir somas fechadas de teores, umas

das desvantagens apresentadas pela simulação direta em blocos dos fatores MAFs.

Com esta combinação espera-se obter uma solução acabada para os principais problemas da simulação de teores em depósitos multivariados complexos: a reprodução das correlações e a dificuldade de implementação dos métodos clássicos de cossimulação; a garantia de somas constantes e fechadas dos teores dos blocos e/ou faixas granulométricas e o tempo e o trabalho envolvidos no processo de obtenção dos cenários.

O método foi aplicado seguindo os seguintes passos:

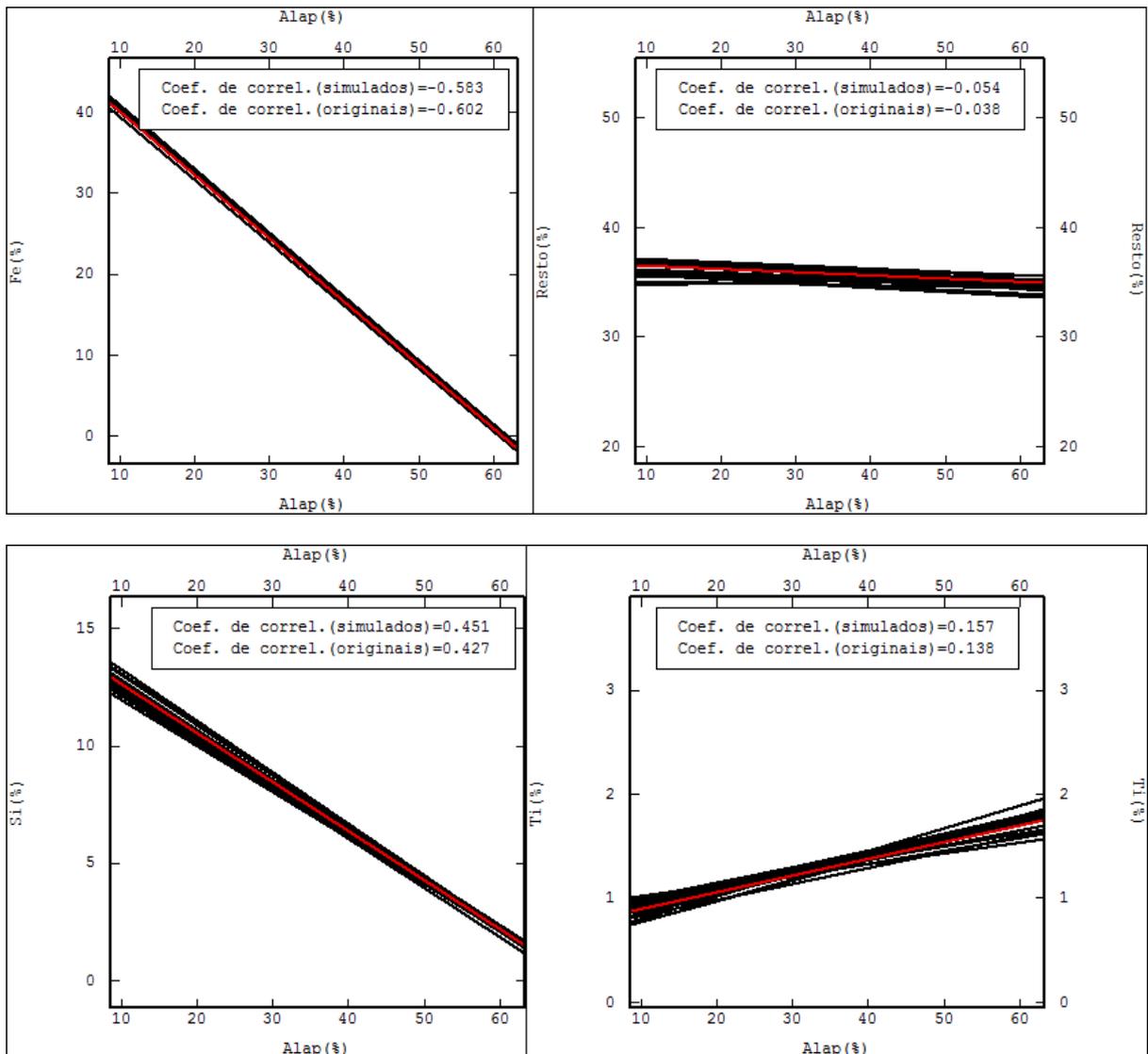
- i. Transformação ilr sobre as variáveis originais não acumuladas;
  - ii. Acumulação das ilrs;
  - iii. Ortogonalização das ilrs normalizadas via transformação MAF;
  - iv. Normalização dos MAFs das ilrs acumuladas;
  - v. Simulação direta em blocos:
    - v.i. Malha de 50x50x0,5m;
    - v.ii. Mesmo modelo e parâmetros ajustados para a variável Rec14# normalizada;
  - vi. Retrotransformação dos dados MAFs das ilrs normalizados para MAFs das ilrs acumuladas;
  - vii. Retrotransformação, aplicação do MAF inverso para obtenção das ilrs acumuladas;
  - viii. Desacumulação (ilr acumulada simulada/recuperação simulada);
  - ix. Transformação igl (retorno do Real para o Simplex);
  - x. Validação dos cenários simulados frente às variáveis originais (anexo A);
  - xi. Verificação da reprodução das correlações entre as variáveis;
- Verificação da reprodução da soma dos teores nos blocos simulados.

#### 4.5.1. CORRELAÇÕES

A utilização da decomposição MAF, assim como na combinação 4, garantiu a reprodução, praticamente fiel, das correlações entre as variáveis

originais nos cenários simulados. Quando observamos, na figura 4.13, para: AlapxFe, AlapxResto, AlapxSi, AlapxTi, FexSi, FexTi, RestoxTi e SixTi, percebemos uma quase perfeita correlação dos valores simulados aos valores originais.

E, apesar de FexResto e SixResto não exibirem este comportamento praticamente perfeito, as correlações entre as variáveis nos cenários simulados são suficientemente reproduzidas. Novamente, a presença da variável artificial Resto prejudica, mesmo que minoritariamente, a reprodução exata das correlações nas simulações.



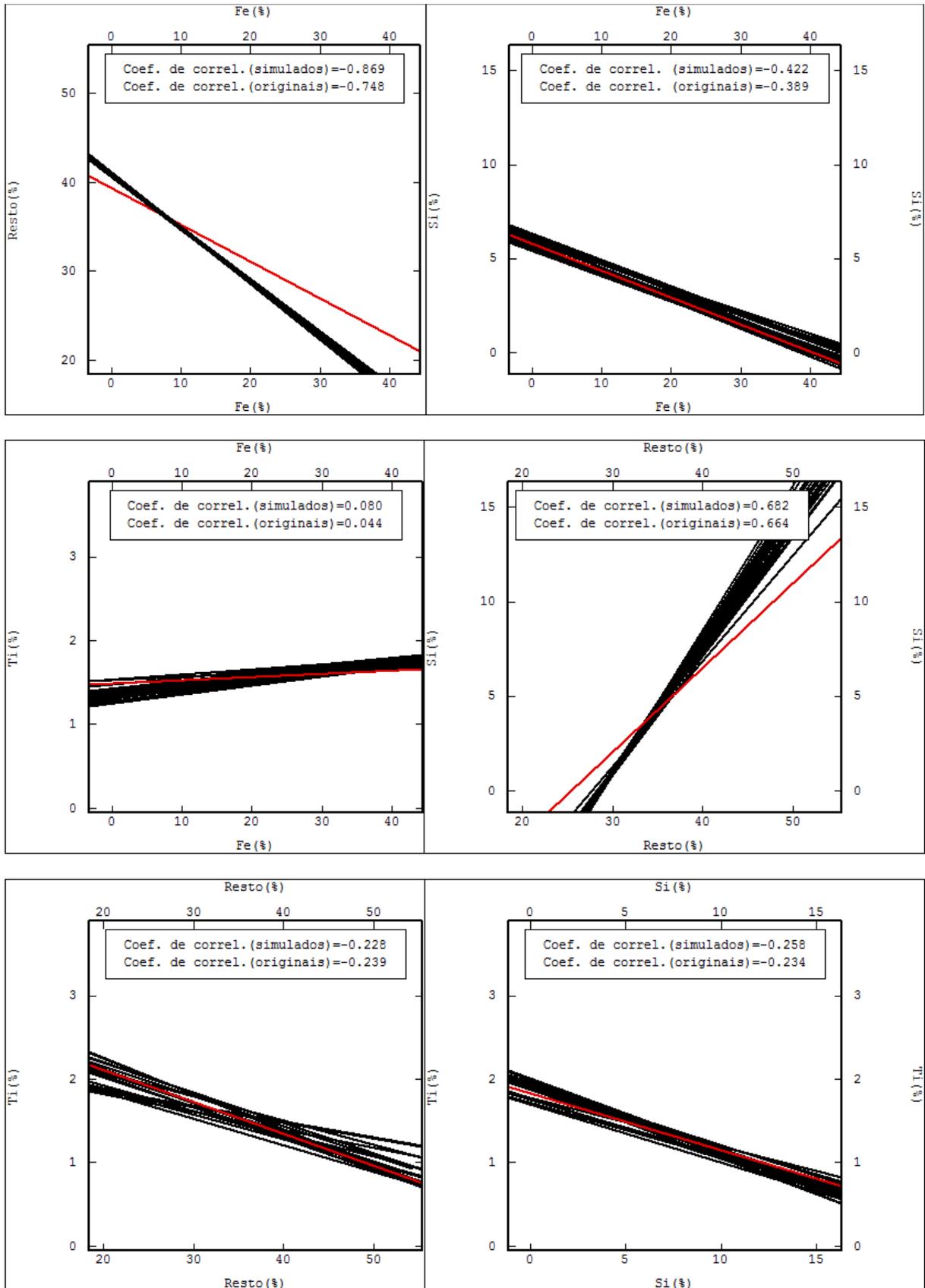


FIGURA 4.15: Correlações lineares entre as variáveis originais (linhas vermelhas) e cenários simulados (linhas pretas). AlapxFe, AlapxResto, AlapxSi, AlapxTi, FexSi, FexTi, RestoxTi e SixTi, exibem uma aderência quase que completa dos valores simulados aos valores originais. E, FexResto e SixResto, apesar de não exibirem este comportamento praticamente perfeito, mostram correlações suficientes.

#### 4.5.2. FECHAMENTO

Observe na tabela 4.7, que todos os blocos simulados, para todas as rodadas, possuem soma única e constante (100%). Este fechamento já era esperado uma vez que as amostras foram consideradas como composições do Simplex.

Tabela 4.7: Análise do fechamento da soma dos teores simulados para a faixa 14#.

<i>Variável</i>	<i>Número de blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>
<i>soma[00001]</i>	69484	100	100	100
<i>soma[00002]</i>	69484	100	100	100
<i>soma[00003]</i>	69484	100	100	100
<i>soma[00004]</i>	69484	100	100	100
<i>soma[00005]</i>	69484	100	100	100
<i>soma[00006]</i>	69484	100	100	100
<i>soma[00007]</i>	69484	100	100	100
<i>soma[00008]</i>	69484	100	100	100
<i>soma[00009]</i>	69484	100	100	100
<i>soma[00010]</i>	69484	100	100	100
<i>soma[00011]</i>	69484	100	100	100
<i>soma[00012]</i>	69484	100	100	100
<i>soma[00013]</i>	69484	100	100	100
<i>soma[00014]</i>	69484	100	100	100
<i>soma[00015]</i>	69484	100	100	100
<i>soma[00016]</i>	69484	100	100	100
<i>soma[00017]</i>	69484	100	100	100
<i>soma[00018]</i>	69484	100	100	100
<i>soma[00019]</i>	69484	100	100	100
<i>soma[00020]</i>	69484	100	100	100

#### 4.5.3. ANÁLISE DO MÉTODO

Os itens considerados na análise da qualidade do método empregado foram classificados seguindo os mesmos critérios estabelecidos para as combinações anteriores.

A análise dos resultados obtidos permite afirmar que a combinação da simulação dos MAFs dos ilrs diretamente em suporte de blocos é uma excelente alternativa para viabilizar a simulação de teores de diversas variáveis em um depósito multivariado complexo. Além disto, também se apresenta como possível

alternativa para a estimativa dos teores destes depósitos (retendo o E-type), uma vez que soluciona o problema do viés existente nas estimativas de composições baseadas em métodos lineares.

A combinação gerou cenários válidos e que de forma geral reproduziram as correlações existentes entre as variáveis e o fechamento da soma dos teores da faixa estudada. Crê-se que, para os casos em que a acumulação não seja necessária, a decomposição MAF traga não somente o benefício da reprodução das correlações, mas também a possibilidade de simular cada variável utilizando o seu respectivo modelo.

É importante observar que a aplicação, em sequência, da combinação proposta não é trivial. Assim sendo, é importante que o usuário esteja familiarizado com todos os métodos envolvidos de forma que as operações sejam aplicadas na sequência correta.

#### 4.6. SIMULAÇÃO DIRETA EM BLOCOS DOS MAFs DAS ILRS COM A OPERAÇÃO DE FECHAMENTO

Apesar de a combinação 4 prover resultados bastante satisfatórios, pensou-se ainda em mais uma combinação de métodos: a simulação direta em blocos, dos MAFs das ilrs, após a operação de fechamento (figura 1.6).

Como já salientado no capítulo 3, quando utilizamos a operação de fechamento, passamos a trabalhar com as proporções das variáveis de interesse dentro de cada amostra, solucionando, portanto, a restrição de soma única e constante em todo o banco de dados.

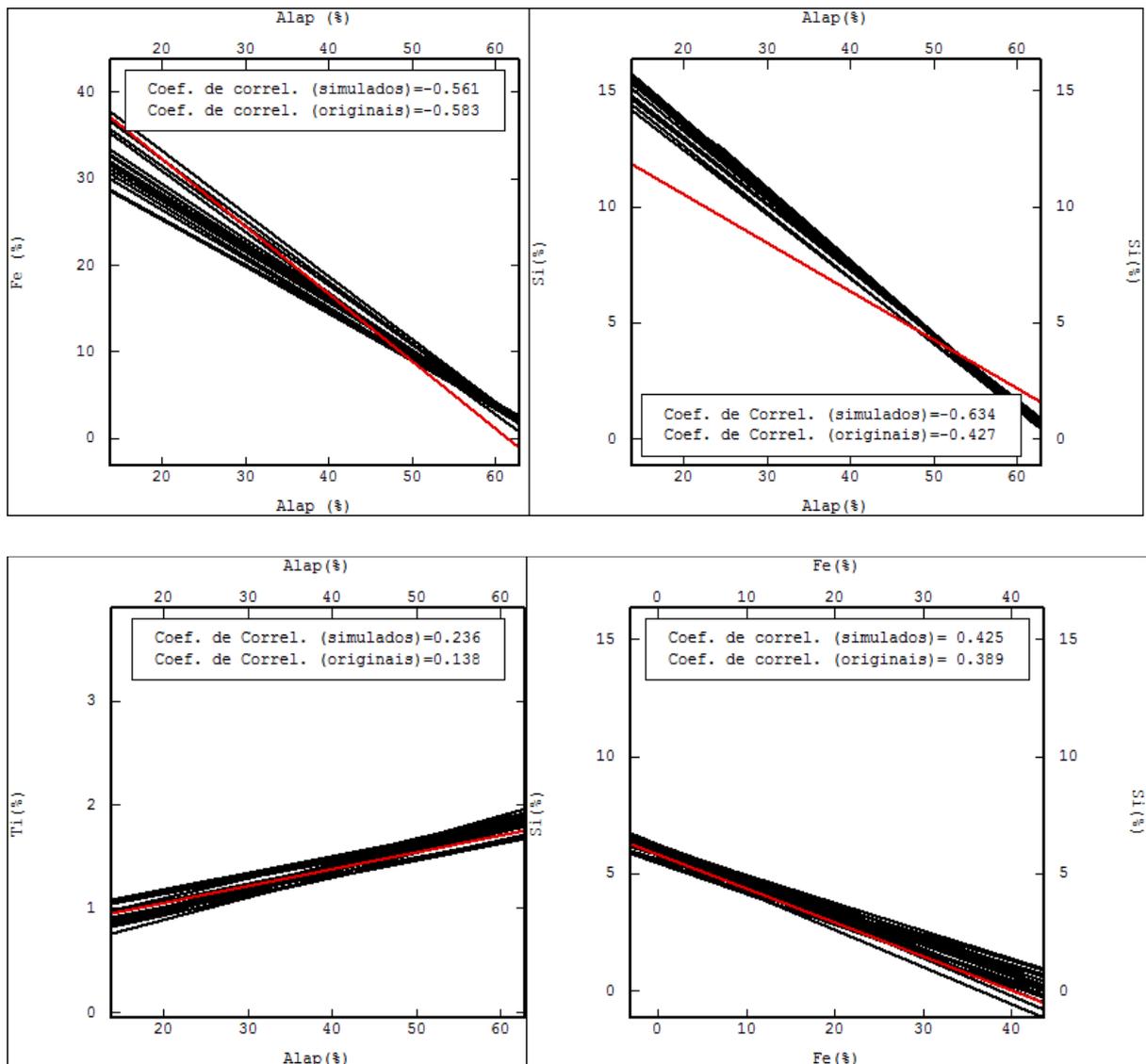
Por estar baseado em proporções em vez de valores absolutos, o método também traz consigo a vantagem de não precisar da utilização de variáveis artificiais utilizadas para induzir o fechamento dos teores. E, estando o fechamento aliado à utilização da transformação ilr, o número de variáveis que precisam ser simuladas fica restrito às originais menos uma. Ou seja, foram efetivamente simuladas apenas três razões logarítmicas isométricas, o que reduziu de maneira significativa o trabalho necessário e o tempo de implementação da combinação.

Os passos necessários para a implementação do método são:

- i. Operação de fechamento;
- iii. Transformação ilr;
- iv. Acumulação das ilrs;
- iv. Transformação MAF;
- v. Normalização dos MAFs;
- vi. Cálculo e ajuste dos correlogramas normalizados;
- vii. Simulação direta em blocos:
  - v.i. Malha de 50x50x0,5m;
  - v.ii. Mesmo modelo e parâmetros ajustados para a variável Rec14# normalizada;
- viii. Validação das simulações (Anexo A)
- viii. Retro transformação dos resultados normalizados para MAFs;
- ix. Operação MAF inversa;
- x. Desacumulação dos ilrs;
- xi. Operação igl;
- xii. Operação inversa do fechamento.

#### 4.6.1. CORRELAÇÕES

Na figura 4.14, é possível observar as correlações entre Alap, Fe, Si e Ti, nos cenários simulados (linhas pretas) e nos dados originais (linhas vermelhas). Analisando os gráficos, percebe-se que todas as correlações ou ausências de, são reproduzidas adequadamente pelas simulações realizadas. Somente o comportamento de AlapxSi simulados é um pouco destoante do que ocorre nos dados originais, mas de forma não significativa.



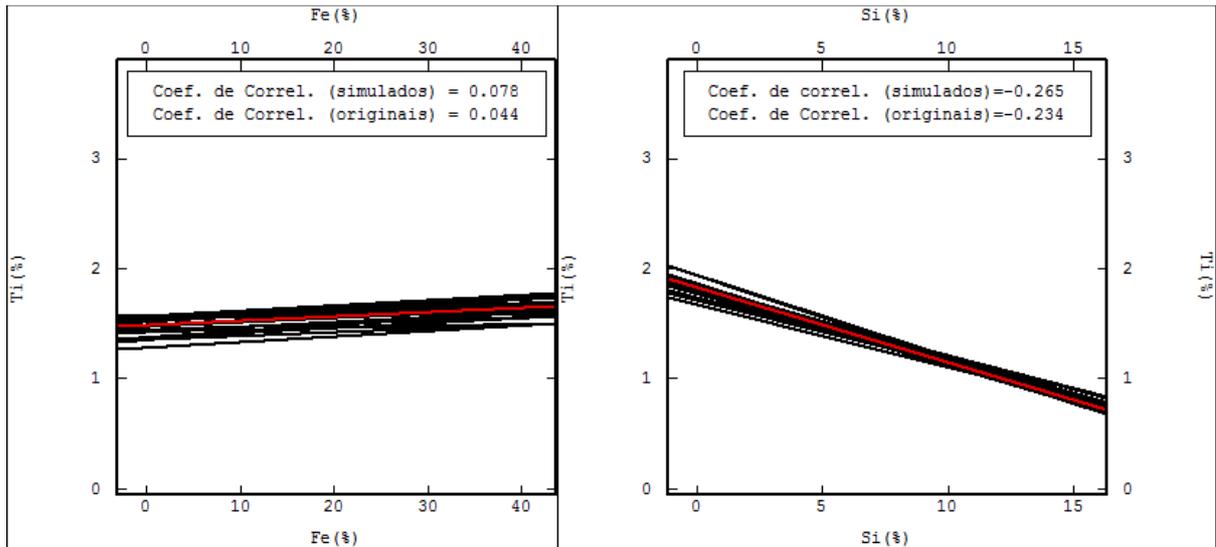


FIGURA 4.16: Correlações lineares entre as variáveis originais (linhas vermelhas) e cenários simulados (linhas pretas). AlapxFe, AlapxTi, FexSi, FexTi, e SixTi, exibem uma aderência quase que completa dos valores simulados aos valores originais. E, AlapxSi, apesar de não exibirem este comportamento praticamente perfeito, mostram correlações suficientes.

#### 4.6.2. FECHAMENTO

Sempre que um conjunto qualquer de amostras que possua dados pertencentes ao espaço Simplex, ou seja, com valores maiores do que zero e com soma constante, for considerado como tal, trazido para o espaço Real (irrestrito, positivo e negativo), depois novamente retornado ao Simplex, as condições de restrição estarão garantidas.

Assim como ocorreu nas combinações 2 e 4, nesta última combinação de métodos que envolveu a transformação *ilr*, também as somas das proporções simuladas se mantiveram constantes nos resultados obtidos por meio das simulações (tabela 4.8).

Tabela 4.8: Análise do fechamento da soma das proporções simuladas para a faixa 14#.

<i>Variável</i>	<i>Número de blocos</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Média</i>
<i>soma[00001]</i>	69484	1	1	1
<i>soma[00002]</i>	69484	1	1	1
<i>soma[00003]</i>	69484	1	1	1
<i>soma[00004]</i>	69484	1	1	1
<i>soma[00005]</i>	69484	1	1	1
<i>soma[00006]</i>	69484	1	1	1
<i>soma[00007]</i>	69484	1	1	1
<i>soma[00008]</i>	69484	1	1	1
<i>soma[00009]</i>	69484	1	1	1
<i>soma[00010]</i>	69484	1	1	1
<i>soma[00011]</i>	69484	1	1	1
<i>soma[00012]</i>	69484	1	1	1
<i>soma[00013]</i>	69484	1	1	1
<i>soma[00014]</i>	69484	1	1	1
<i>soma[00015]</i>	69484	1	1	1
<i>soma[00016]</i>	69484	1	1	1
<i>soma[00017]</i>	69484	1	1	1
<i>soma[00018]</i>	69484	1	1	1
<i>soma[00019]</i>	69484	1	1	1
<i>soma[00020]</i>	69484	1	1	1

Em vez de uma soma constante em termos de porcentagem, quando utilizamos a operação de fechamento, o que será constante será a soma das proporções das variáveis que compõe a amostra ou bloco. Por este motivo, além da verificação da manutenção da soma das proporções, nesta quinta combinação, foi especialmente verificado o comportamento das médias globais dos teores dos blocos simulados em comparação com os das variáveis originais amostradas.

Analisando as tabelas 4.9, 4.10, 4.11 e 4.12, que contêm a estatística básica dos dados originais (no topo) e dos resultados obtidos na simulação (abaixo), é possível perceber que os valores resultantes da simulação respeitam o comportamento exibido pelas amostras, com valores simulados flutuando em torno dos originais para todas as quatro variáveis estudadas.

TABELA 4.9: Médias globais dos dados originais (no topo) e dos 20 cenários simulados para Alap (abaixo).

<b>VARIÁVEL ALAP</b>	<b>NÚMERO DE DADOS/BLOCOS</b>	<b>MÍNIMO(%)</b>	<b>MÁXIMO(%)</b>	<b>MÉDIA(%)</b>	<b>DESVIO PADRÃO(%)</b>
<b>DADOS ORIGINAIS</b>	<b>3735</b>	<b>13,07</b>	<b>58,78</b>	<b>49,17</b>	<b>3,83</b>
ALAP[00001]	69484	13,07	58,78	49,42	3,65
ALAP [00002]	69484	13,07	58,78	49,98	3,32
ALAP [00003]	69484	13,07	58,78	49,69	3,65
ALAP[00004]	69484	13,07	58,78	49,02	3,54
ALAP[00005]	69484	13,07	58,78	49,26	3,25
ALAP[00006]	69484	13,07	58,78	49,1	3,21
ALAP[00007]	69484	13,07	58,78	48,84	3,98
ALAP[00008]	69484	13,07	58,78	49,1	3,65
ALAP[00009]	69484	13,07	58,78	49,05	3,87
ALAP[00010]	69484	13,07	58,78	49,32	3,49
ALAP[00011]	69484	13,07	58,78	50,15	3,69
ALAP[00012]	69484	13,07	58,78	49,89	3,35
ALAP[00013]	69484	13,07	58,78	49,28	3,24
ALAP[00014]	69484	13,07	58,78	49,5	3,21
ALAP[00015]	69484	13,07	58,78	49,59	3,69
ALAP[00016]	69484	13,07	58,78	49,97	3,98
ALAP[00017]	69484	13,07	58,78	49,44	3,67
ALAP[00018]	69484	13,07	58,78	49,37	3,95
ALAP[00019]	69484	13,07	58,78	50,34	3,84
ALAP[00020]	69484	13,07	58,78	49,34	3,56

TABELA 4.10: Médias globais dos dados originais (no topo) e dos 20 cenários simulados para Fe (abaixo).

<b>Variável Fe</b>	<b>Número de dados/blocos</b>	<b>Mínimo(%)</b>	<b>Máximo(%)</b>	<b>Média(%)</b>	<b>Desvio padrão(%)</b>
<b>Dados originais</b>	<b>3735</b>	<b>0,69</b>	<b>54,65</b>	<b>9,49</b>	<b>4,98</b>
Fe[00001]	69484	0,69	54,65	9,24	4,58
Fe[00002]	69484	0,69	54,65	9,08	4,96
Fe[00003]	69484	0,69	54,65	9,23	5
Fe[00004]	69484	0,69	54,65	9,44	5,01
Fe[00005]	69484	0,69	54,65	9,47	5,03
Fe[00006]	69484	0,69	54,65	9,4	5,06
Fe[00007]	69484	0,69	54,65	9,53	4,95
Fe[00008]	69484	0,69	54,65	9,49	4,86
Fe[00009]	69484	0,69	54,65	9,42	4,76
Fe[00010]	69484	0,69	54,65	9,44	4,98
Fe[00011]	69484	0,69	54,65	9,03	4,99
Fe[00012]	69484	0,69	54,65	9,07	5,00
Fe[00013]	69484	0,69	54,65	9,56	5,01
Fe[00014]	69484	0,69	54,65	9,31	5,06

<i>Fe[00015]</i>	69484	0,69	54,65	9,31	4,99
<i>Fe[00016]</i>	69484	0,69	54,65	9,11	4,78
<i>Fe[00017]</i>	69484	0,69	54,65	9,43	4,85
<i>Fe[00018]</i>	69484	0,69	54,65	9,34	4,91
<i>Fe[00019]</i>	69484	0,69	54,65	9,92	4,93
<i>Fe[00020]</i>	69484	0,69	54,65	9,35	5,04

TABELA 4.11: Médias globais dos dados originais (no topo) e dos 20 cenários simulados para Si (abaixo).

<b>Variável Si</b>	<b>Número de dados/blocos</b>	<b>Mínimo(%)</b>	<b>Máximo(%)</b>	<b>Média(%)</b>	<b>Desvio padrão(%)</b>
<b>Dados originais</b>	<b>3735</b>	<b>0,27</b>	<b>14,90</b>	<b>4,44</b>	<b>1,97</b>
<i>Si[00001]</i>	69484	0,27	14,90	4,71	2,01
<i>Si[00002]</i>	69484	0,27	14,90	4,41	2,02
<i>Si[00003]</i>	69484	0,27	14,90	4,46	1,96
<i>Si[00004]</i>	69484	0,27	14,90	4,89	1,95
<i>Si[00005]</i>	69484	0,27	14,90	4,69	1,96
<i>Si[00006]</i>	69484	0,27	14,90	4,85	2,06
<i>Si[00007]</i>	69484	0,27	14,90	4,90	1,97
<i>Si[00008]</i>	69484	0,27	14,90	4,72	1,96
<i>Si[00009]</i>	69484	0,27	14,90	4,83	1,95
<i>Si[00010]</i>	69484	0,27	14,90	4,59	1,99
<i>Si[00011]</i>	69484	0,27	14,90	4,33	2,01
<i>Si[00012]</i>	69484	0,27	14,90	4,44	2,09
<i>Si[00013]</i>	69484	0,27	14,90	4,60	2,04
<i>Si[00014]</i>	69484	0,27	14,90	4,53	2,00
<i>Si[00015]</i>	69484	0,27	14,90	4,52	1,94
<i>Si[00016]</i>	69484	0,27	14,90	4,37	1,93
<i>Si[00017]</i>	69484	0,27	14,90	4,52	1,98
<i>Si[00018]</i>	69484	0,27	14,90	4,66	1,99
<i>Si[00019]</i>	69484	0,27	14,90	4,26	2,01
<i>Si[00020]</i>	69484	0,27	14,90	4,65	2,02

TABELA 4.12: Médias globais dos dados originais (no topo) e dos 20 cenários simulados para Ti (abaixo).

<b>Variável Ti</b>	<b>Número de dados/blocos</b>	<b>Mínimo(%)</b>	<b>Máximo(%)</b>	<b>Média(%)</b>	<b>Desvio Padrão(%)</b>
<b>Dados originais</b>	<b>3735</b>	<b>0,27</b>	<b>3,57</b>	<b>1,53</b>	<b>0,45</b>
<i>Ti[00001]</i>	69484	0,27	3,57	1,51	0,44
<i>Ti[00002]</i>	69484	0,27	3,57	1,42	0,46
<i>Ti[00003]</i>	69484	0,27	3,57	1,48	0,41
<i>Ti[00004]</i>	69484	0,27	3,57	1,64	0,43
<i>Ti[00005]</i>	69484	0,27	3,57	1,55	0,44
<i>Ti[00006]</i>	69484	0,27	3,57	1,62	0,46

<i>Ti[00007]</i>	69484	0,27	3,57	1,62	0,48
<i>Ti[00008]</i>	69484	0,27	3,57	1,58	0,41
<i>Ti[00009]</i>	69484	0,27	3,57	1,6	0,47
<i>Ti[00010]</i>	69484	0,27	3,57	1,55	0,41
<i>Ti[00011]</i>	69484	0,27	3,57	1,45	0,44
<i>Ti[00012]</i>	69484	0,27	3,57	1,42	0,45
<i>Ti[00013]</i>	69484	0,27	3,57	1,52	0,46
<i>Ti[00014]</i>	69484	0,27	3,57	1,46	0,47
<i>Ti[00015]</i>	69484	0,27	3,57	1,49	0,43
<i>Ti[00016]</i>	69484	0,27	3,57	1,43	0,48
<i>Ti[00017]</i>	69484	0,27	3,57	1,48	0,44
<i>Ti[00018]</i>	69484	0,27	3,57	1,53	0,45
<i>Ti[00019]</i>	69484	0,27	3,57	1,37	0,49
<i>Ti[00020]</i>	69484	0,27	3,57	1,51	0,42

#### 4.6.3. ANÁLISE DO MÉTODO

A análise da qualidade dos resultados obtidos em termos de reprodução de correlações e fechamento de soma de teores, tempo e facilidade de implementação será apresentada a seguir.

Com relação às correlações observadas nas amostras e observadas nos cenários simulados, podemos dizer que o conjunto de operações de fechamento, operação *ilr* e operação *MAF*, seguida da simulação dos teores da faixa 14# diretamente em suporte de blocos, retornou resultados adequados. Ou seja, os comportamentos cruzados exibidos por todas as variáveis na análise dos dados originais são mantidos nos resultados das simulações, principalmente devido ao fato de a decomposição *MAF* fazer parte dos métodos aplicados nesta combinação.

A análise da soma dos teores quando a operação de fechamento é utilizada acaba sendo um pouco diferente da realidade a que estamos acostumados no meio mineiro. Em vez de termos somas de teores em blocos ou faixas granulométricas fechadas em 100% (ou outra constante qualquer) temos que a soma das proporções das variáveis presentes em cada amostra, bloco ou faixa simulados é unitária.

Verificando os resultados obtidos após a aplicação de todos os métodos que fazem parte da quinta combinação, é possível perceber que a condição de soma unitária das proporções é mantida, como esperado. Para trazer estes

resultados para uma realidade mais próxima, foram especialmente analisadas as reproduções das médias globais para cada variável de interesse. E, durante esta verificação pôde-se perceber que as médias simuladas apresentam valores muito próximos, variando um pouco para mais e um pouco para menos, em relação aos valores das médias originais.

Por fim, considerando que a operação de fechamento não requer a criação de uma variável artificial para indução de fechamento e que, a utilização da transformação ilr implica na redução de uma variável a ser simulada, o método se destaca pela necessidade real de somente quatro rodadas serem necessárias, três referentes aos atributos de interesse e uma referente à variável de acumulação/desacumulação, Rec14#.

Este fato contribui para que o tempo necessário para grande parte da aplicação da combinação fechamento+ilr+MAF+simulação diretamente em blocos, seja reduzido, o que pode ser percebido principalmente nas etapas de modelagem de correlogramas, processo de simulação em si e validações.

Assim como ocorre com a combinação quatro, aqui também o usuário precisa estar familiarizado e atento à sequência em que as operações devem ser aplicadas para que os resultados sejam adequados e não ocorra nenhum tipo de manipulação imprópria dos resultados. Atendidas estas condições, é possível afirmar que a combinação cinco apresenta resultados coerentes para simulação, correlações e fechamento, sendo, portanto, perfeitamente indicada para aplicação na modelagem de teores e no planejamento de mina.

#### 4.7. COMPARAÇÃO ENTRE OS MÉTODOS

A análise comparativa entre os métodos empregados inclui: a verificação da reprodução das estatísticas de primeira e segunda ordem dos dados originais do depósito pelos cenários simulados em cada combinação de métodos e; a verificação das principais características de interesse dos usuários, fechamento da soma dos teores, tempo e facilidade de implementação.

Relembrando, os métodos ou combinações de métodos estudados nesta tese foram:

i. *Combinação 1*: simulação individual das variáveis de interesse em suporte de pontos posteriormente reblocados, com a utilização do método de bandas rotativas;

ii. *Combinação 2*: transformação ilr + simulação individual das razões em suporte de pontos posteriormente reblocados, com a utilização do método de bandas rotativas;

iii. *Combinação 3*: decomposição MAF + simulação direta em blocos;

iv. *Combinação 4*: transformação ilr + decomposição MAF + simulação direta em blocos;

v. *Combinação 5*: operação de fechamento + transformação ilr + decomposição MAF + simulação direta em blocos.

#### 4.7.1. VALIDAÇÃO DOS CENÁRIOS

As médias dos cenários simulados apresentam valores que flutuam em torno dos originais em todas as combinações propostas. Percebe-se que, a média global de todas as variáveis foi adequadamente reproduzida pelos métodos propostos, assim como as médias locais, e, o comportamento espacial original foi também mantido nos cenários gerados em todos os métodos ou combinação de métodos, para todas as variáveis (Anexo A).

#### 4.7.2. CORRELAÇÕES

Com relação à reprodução das correlações nos cenários simulados quando comparados com os originais é possível afirmar que para todas as combinações de métodos aplicados em que a decomposição MAF esteve presente (*combinações iii., iv. e v.*) estas foram mantidas de maneira adequada nos resultados obtidos.

Quando verificamos os dados de saída das combinações i. e ii, percebemos que os resultados obtidos são destoantes dos originais. Como já dito anteriormente, esta diferença se dá por uma soma de motivos e o principal deles é a simulação individual do teores para i. e a utilização de uma transformação do tipo ilr seguida da simulação individual dos teores para ii..

Fazendo a comparação entre os métodos aplicados é válido dizer que, para garantir a reprodução de correlações quando não for utilizado um método clássico de cossimulação, a decomposição MAF deve estar presente entre os métodos aplicados. Além disto, é importante salientar que a decomposição MAF também carrega consigo a vantagem de permitir a manipulação de cada variável de forma independente, sem perda na qualidade dos resultados finais, o que é muito importante quando tratamos de depósitos multivariados complexos com mais de três variáveis envolvidas.

#### 4.7.3. FECHAMENTO

A soma dos teores simulados para a faixa 14# deve ser de 100% para todas as combinações, com exceção da última em que este valor deve ser unitário.

Na combinação 1, todos os blocos simulados apresentam soma pertencente ao intervalo estabelecido como aceitável (entre 95% e 105%) e as médias das somas dos teores nos blocos, apesar de não serem exatamente 100%, possuem valores muito próximos deste. O resultado pode ser considerado razoável uma vez que todo o trabalho foi dentro do espaço Real, que não restringe as somas.

Já na terceira combinação, cerca de 25% dos blocos simulados em todos os cenários simulados apresentaram valores de soma ou superiores a 105% ou inferiores a 95%. Como solução, os valores da variável artificial Resto foram ajustados, o que não implica em maiores prejuízos, já que este somente é utilizada para indução do fechamento.

O procedimento de “carregar” o erro das estimativas ou simulações sobre a variável menos importante do depósito é rotineiro e não implica em prejuízos do ponto de vista prático ou operacional às mineradoras. No entanto, é inegável que este é um ajuste sem embasamento teórico algum.

Como, por construção, as transformações que levam as composições do Simplex de D dimensões para um subespaço equivalente do Real de D-1 dimensões, implicam em somas constantes, a simulação em pontos das ilrs, a simulação direta em blocos dos MAFs das ilrs e a simulação direta em blocos dos MAFs das ilr, após a operação de fechamento, geraram resultados com soma de

teores em 100%, para ii., iv. e v., e, 1, para v.. Assim, a inclusão deste método é fortemente indicada para atendimento da condição de soma fechada.

Também é importante salientar que as operações de mudança de espaço abarcam ainda a redução no número de variáveis a serem estimadas/simuladas em uma unidade, o que reduz significativamente o tempo dispendido para obtenção dos resultados finais.

#### 4.7.4. INDICAÇÕES DE UTILIZAÇÃO

Com base na análise dos resultados obtidos nas cinco combinações estudadas, pode-se afirmar que o método que atende de forma mais eficiente e completa às questões de reprodução de correlações, fechamento de soma de teores, tempo e facilidade de implementação é a simulação direta em blocos dos fatores MAF das ilrs precedida da operação de fechamento.

A inclusão da decomposição MAF permitiu que as correlações fossem reproduzidas de forma razoável, permitindo a simulação de cada uma de modo independente das demais. A transformação ilr trouxe consigo a vantagem da garantia de somas constantes de teores para todos os blocos, em todos os cenários simulados, adicionado do número de variáveis simuladas, que foi reduzido em duas unidades. Por fim, a simulação direta em suporte de blocos faz com que o método seja rapidamente aplicado, o que corrobora para sua indicação de utilização na rotina do planejamento de longo prazo.

As principais desvantagens são:

i. As transformações ilrs ainda não estão disponíveis no software utilizado, mas as operações de ida e volta não se mostraram penosas, quando consideramos a qualidade dos resultados obtidos;

ii. O fato de trabalharmos com proporções em vez de teores em porcentagem pode causar estranheza ao usuário que não esteja plenamente familiarizado com o método;

iii. A sequência em que as operações precisam ser aplicadas também deve ser familiar ao usuário, mas tendo este o domínio do porquê da realização de cada passo, a sequência também acabará por ser intuitiva.

Apesar de a combinação cinco ser a que melhor se apresenta como alternativa aos métodos clássicos de cossimulação, passível de fato à

implementação na indústria, as demais combinações estudadas também podem ser utilizadas com a garantia de resultados satisfatórios, o que já ocorre em algumas aplicações.

#### 4.8. COMENTÁRIOS

O presente capítulo apresentou os resultados obtidos em cada combinação de métodos estudada. Os resultados foram discutidos e analisados de forma individual e em grupo, do ponto de vista de qualidade e facilidade de implementação. O próximo capítulo consta das conclusões obtidas a partir do estudo e ainda, das recomendações para trabalhos futuros.

# Capítulo 5

## 1. CONCLUSÕES

O estudo das variáveis presentes em um depósito mineral deve contemplar não somente o atributo de maior interesse econômico, mas também os demais elementos presentes, seja por sua possível influência no processo de beneficiamento, pela possibilidade de trazer contaminações ao meio ambiente ou por quaisquer outros motivos relevantes para o caso em questão.

Quando tratamos da estimativa de teores em depósitos multivariados complexos, com mais de três variáveis envolvidas, os métodos clássicos acabam por se tornar demasiadamente trabalhosos, o que prejudica sua aplicação na rotina da indústria mineira. Esta, por sua vez, termina buscando soluções mais rápidas, não tão precisas, como a distribuição do erro das estimativas entre todas as variáveis presentes, ou ainda, o carregamento deste erro sobre a variável de menor interesse. Estas soluções, apesar de não serem tão robustas matematicamente, de modo geral, geram resultados suficientes para a estimativa dos teores destes depósitos.

No entanto, quando em vez da busca pelo melhor valor estimado para os teores, em um dado local, queremos informações sobre a variabilidade que estes podem sofrer. Nesse caso, nos referimos, portanto, à geração de diversos cenários simulados equiprováveis e a dificuldade na implementação dos métodos clássicos de cossimulação se torna ainda maior, uma vez que é como se estivéssemos repetindo o processo de estimativa inúmeras vezes, para cada elemento de interesse.

É importante percebermos que quando o custo de extração for baixo, ou for alto o preço de venda do minério e há grande a disponibilidade de reservas minerais, a utilização de aproximações, ou seja, de correções práticas aos valores estimados/simulados, não causará, perceptivamente, grande impacto nos resultados econômicos da empresa.

No entanto, na conjuntura atual, onde os preços dos minérios têm sofrido queda significativa e as reservas facilmente lavráveis têm se exaurido, é

importante garantirmos que os resultados obtidos na estimativa ou simulação de teores sejam os mais corretos possíveis, de modo a garantir uma exploração sustentável. Tudo isto motiva a busca por alternativas aos métodos clássicos, com resultados igualmente precisos, mas que sejam mais atrativos e facilmente implementáveis.

Nesta tese, foram propostas cinco possibilidades de combinações de métodos para a simulação da faixa #14 de um depósito multivariado complexo de bauxita, com quatro variáveis de interesse.

Os resultados obtidos por meio da aplicação de cada combinação proposta foram analisados em termos da qualidade da reprodução das características estatísticas dos dados originais, da manutenção ou não das correlações ou falta de correlações entre as variáveis originais nos cenários simulados, e, da garantia da soma correta dos teores dos dados originais também nos blocos simulados.

A tabela 5.1 consta de um quadro resumo com as informações sobre a qualidade da reprodução dos resultados obtidos em cada combinação testada, onde "Reproduziu" significa que os resultados obtidos na referida combinação reproduzem fielmente o comportamento exibido pelos dados originais e "Não reproduziu" significa que a combinação em questão não gerou resultados suficientemente semelhantes aos dos dados originais.

Com relação ao grau de facilidade de aplicação da metodologia os índices escolhidos foram: "Razoável" quando os métodos aplicados são de conhecimento relativamente comum para os profissionais do meio, ou ainda, exigem poucos passos de transformação de dados; e, "Exige atenção" para quando a combinação proposta é a soma da aplicação de duas transformações que por sua vez, exigem atenção e execução de mais passos de transformação.

Tabela 5.1: quadro resumo com dados sobre a qualidade de reprodução dos resultados simulados quando comparados com os dados originais e sobre o grau de facilidade de implementação da metodologia.

<i>Combinação</i>	<i>Características estatísticas</i>	<i>Correlações</i>	<i>Fechamento</i>	<i>Facilidade de aplicação</i>
1	Reproduziu	Não reproduziu	Reproduziu	Razoável
2	Reproduziu	Não reproduziu	Reproduziu	Razoável
3	Reproduziu	Reproduziu	Não reproduziu	Razoável
4	Reproduziu	Reproduziu	Reproduziu	Exige atenção
5	Reproduziu	Reproduziu	Reproduziu	Exige atenção

Observando os resultados exibidos no quadro (Tabela 5.1) se pode perceber que as combinações 4 e 5 são únicas que reproduziram plenamente as características estatísticas, as correlações e o fechamento. Este resultado era esperado uma vez que ambas as combinações envolvem a aplicação de metodologias que carregam consigo os benefícios buscados.

A consideração dos dados originais como composições pertencentes ao Simplex de D dimensões, a aplicação da transformação logarítmica isométrica e a posterior realização das simulações sobre estas razões seguida do retorno ao espaço original, traz consigo o benefício da manutenção das somas fechadas constantes e pertencentes ao espaço restrito, além da redução do número de variáveis simuladas em uma unidade.

Já a decomposição em fatores de autocorrelação mínimos e máximos viabiliza a simulação individual dos atributos, uma vez que faz com que estes fiquem temporariamente descorrelacionados e, posteriormente, quando da aplicação da operação MAF inversa, voltem a ser correlacionados.

Por fim, a simulação quando executada diretamente em suporte de blocos, além de fornecer os resultados diretamente no suporte de interesse, requer menos memória computacional, por armazenar somente a média centrada de cada bloco e utilizar este valor como condicionante para os demais.

Apesar de ambas as combinações 4 e 5 fornecerem resultados com igual qualidade, a última pode ser considerada como menos trabalhosa. Isto ocorre porque ao contrário do que acontece quando utilizamos a variável Resto para induzir o fechamento, a operação *closure* permite que trabalhemos com as proporções relativas em base unitária, para cada elemento dentro da amostra ou faixa granulométrica de interesse. Esta consideração implica na redução do número de variáveis a serem efetivamente simuladas, passando de quatro (seriam cinco, mas a transformação *ilr* já incorpora a redução de uma variável) para três.

Com base no exposto, é possível afirmar que a simulação diretamente em suporte de blocos, dos fatores de autocorrelação mínimos e máximos obtidos após a transformação logarítmica isométrica, sobre os dados resultantes da operação de fechamento, gera resultados que atendem plenamente os requisitos de reprodução das características estatísticas, das correlações e dos fechamentos das somas dos teores para o depósito estudado, provendo portanto, uma solução acabada para a simulação de teores em depósitos multivariados complexos com mais de três variáveis de interesse.

É inegável que o usuário deve estar atento à sequência de aplicação de cada passo e que a metodologia proposta abarca uma série de transformações e considerações dos dados de forma bastante diferente daquela a que estamos acostumados, onde nos afastamos das porcentagens durante a efetiva aplicação do método, mas, após a realização de todos os passos envolvidos, retornamos para os teores da forma como os conhecemos.

## 2. RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Para trabalhos futuros se recomenda:

- (i) aplicação da metodologia em um banco de dados que não precise da operação de acumulação, onde então, cada variável poderá ser simulada com o seu respectivo modelo variográfico;
- (ii) aplicação e comparação com outras combinações de métodos de fatorização de variáveis e demais transformações relativas aos dados composicionais;

(iii) estudo da viabilidade de utilização da simulação de dados composicionais como alternativa para a estimativa de teores em depósitos multivariados complexos, com mais de três variáveis envolvidas.

A consideração de dados amostrais como composições pertencentes ao espaço restrito Simplex, ou seja, como partes de um todo que possui soma constante e somente valores positivos, seguida da aplicação da krigagem ou suas variações, pode gerar resultados enviesados. Neste caso, os métodos de simulação viriam solucionar este problema de viés, surgindo como alternativa aos métodos de krigagem.

Sugere-se para trabalho futuro o estudo aprofundado da metodologia proposta na combinação 5 (ilr+MAF+simulação direta em suporte de blocos+operação de fechamento) como alternativa para a estimativa de teores de composições. O estudo deveria englobar:

(iii.i) estimativa das variáveis de interesse via cokrigagem;

(iii.ii) estimativa das variáveis de interesse via krigagem individual;

(iii.iii) estimativa das variáveis após a transformação em razões logarítmicas (alr ou ilr);

(iii.iv) simulação das variáveis após a transformação em razões logarítmicas (alr ou ilr) retendo o E-Type de cada variável como sendo o cenário indicado para representar o valor único estimado para o atributo;

(iii.v) aplicação da combinação 5 (ilr+MAF+simulação direta em suporte de blocos+operação de fechamento), retendo o E-Type de cada variável como sendo o cenário indicado para representar o valor único estimado para o atributo.

(iii.vi) validação e comparação dos resultados obtidos em (i, ii., iii., iv., e v.) com relação aos valores individuais estimados para cada variável, ao fechamento das somas dos teores nos blocos estimados e a presença de valores negativos e/ou enviesados.

(iv.) substituir a proposta da combinação i. pela cossimulação dos teores com a utilização do Modelo de Corregionalização Intrínseco (MCI);

(v.) verificar a reprodução entre os ilrs para as combinações (ii.), (iv) e (v);

(vi.) verificar a reprodução das características estatística com a geometria do Simplex: centro da composição, matriz de variação e variância da composição;

(vii.) verificar as reproduções dos variogramas cruzados e biplots para todas as combinações.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Abichequer, L., Costa, J., Pasti, H., & Koppe, J. (2011). Design of blending piles by geostatistical simulated models. A real case reconciliation. *International Journal of Mineral Processing*, 99, 21-26.
- Aitchison, J. (1981). A new approach to null correlations of proportions. *Mathematical Geology*, 13(2), 175-189.
- Aitchison, J. (1982). The statistical analysis of compositional data (with discussion). *Journal of Royal Statistical Society*, 44(2), 139-177.
- Aitchison, J. (1983). Principal Component Analysis of Compositional Data. *Biometrika*, 70(1), 57-65.
- Aitchison, J. (1984). The statistical analysis of geochemical compositions. *Mathematical Geology*, 16(2), 531-564.
- Aitchison, J. (1986). *The statistical analysis of compositional data, monographs on statistics and applied probability*. Chapman & Hall Ltd.
- Aitchison, J. (1997). The one-hour course in compositional data analysis or compositional data analysis is simple. *Proceedings of IAMG'97 – The third annual conference of the International* (pp. 3-35). Barcelona: International Center for Numerical Methods in Engineering (CIMNE).
- Aitchison, J., & Egozcue, J. (2005). Compositional Data Analysis: Where are we and where should we be heading? *Mathematical Geology*, 37(7), 829-850.
- Aitchison, J., & Kay, J. (2003). Possible solution of some essential zero problems in compositional data analysis. *Compositional Data Analysis Workshop - Coda Work'03, Proceedings*. Universidad de Girona.
- Alabert, F. (1987). The practice of fast conditional simulations through the LU decomposition of the covariance-matrix. *Mathematical Geology*, 19(2), 369-386.
- Almeida, A. (1993). *Joint simulation of multiple variables with a Markov-type*. University of Standford Press.

- Almeida, A., & Journel, A. (1994). Joint simulation of multiple-variables with a Markov-type coregionalization model. *Mathematical geology*, 26(5), 565-588.
- Bacon-Shone, J. (2003). Modelling structural zeros in compositional data . *Compositional Data Analysis Workshop - Codawork'03*. Universidad de Girona.
- Bandarian, E., Bloom, L., & Muller, U. (2008). Direct minimum/maximum autocorrelation factors within the framework of a two structure linear model of coregionalisation. *Computers and Geoscience*, 34, 190-200.
- Barnett, R., & Deutsch, C. (2012). Practical Implementation of Non-linear transforms for Modeling Geometallurgical Variables. *Geostatistics Oslo 2012, Quantitative Geology and Geostatistics*, 17, 409-422.
- Beretta, F., Costa, J., & Koppe, J. (2010). Reducing coal attributes variability using properly designed blending piles helped by geostatistical simulation. *International Journal of Coal Geology*, 84, 83-93.
- Berman, M. (1985). *The statistical properties of three noise removal procedures for multichannel remotely sense data*. Csiro.
- Billheimer, D., Guttorp, P., & Fagan, W. (1997). *Statistical Analysis and interpretation of discrete compositional data, Technical Report*. Seattle: University of Washington.
- Billheimer, D., Guttorp, P., & Fagan, W. (2001). Statistical interpretation of species composition. *J.Am.Stat. Assoc.*, 1205-1214.
- Blackwell, G., & A.J., S. (2002). *Applied Mineral Inventory Estimation*. Cambridge: Cambridge University Press.
- Blackwell, G., Anderson, H., & Ronson, K. (1999). Simulated Grades and Open Pit Mine Planning - Resolving Opposed Positions. *28 International Symposium on Computer Applications in the Mineral Industries.*, (pp. 205-215). Colorado.
- Boezio, M., Abichequer, L., & Costa, J. (2012). MAF decomposition of compositional data to estimate grades in iron ore deposit. *Ninth International Congress on Geostatistics*. Oslo.

- Boezio, M., Costa, J., & Koppe, J. (2012). Cokrigagem de razões logarítmicas aditivas para estimativa de teores em depósito de ferro. *REM, Revista Escola de Minas*, 401-411.
- Bonato, F., Costa, J., & Koppe, J. (2000). Planejamento de Lavra Baseado na Variabilidade de Atributos Geológicos. *REM - Revista Escola de Minas*, 275-284.
- Borgman, L., Taheri, M., & Hagan, R. (1984). Three-dimensional, frequency domain simulations of geological variables. In G. D. Verly, *Geostatistics for natural resources characterization* (pp. 517-541). Reidel Publishing Company.
- Boucher, A. (2003). *Conditional Joint Simulation of Random Fields on Block Support*. Brisbane: Queensland University.
- Boucher, A., & Dimitrakopoulos, R. (2007). A new efficient joint simulation framework and application in a multivariable deposit. *Orebody modelling and strategic mine planing* (pp. 345-354). Melbourne: The Australasian Institute of Mining and Metallurgy.
- Boucher, A., & Dimitrakopoulos, R. (2009). Block simulation of multiple correlated variables. *Mathematical Geoscience*, 41, 215-237.
- Boucher, A., & Dimitrakopoulos, R. (2012). Multivariate Block-Support Simulation of the Yandi Iron Ore Deposit. *Mathematical Geoscience*, 44, 449-468.
- Buccianti, A., Mateu-Figueras, G., & Pawlowsky-Glahn, V. (. (2006). *Compositional Data Analysis in the Geosciences: From Theory to Practice*. London: Geological Society, Special Publications.
- Butler, J. (1979). The effects of closure on the moments of a distribution. *Mathematical Geology*, 75-84.
- Carr, J., & Meyers, D. (1985). COSIM, A FORTRAN IV program for coconditional simulation. *Computer and Geosciences*, 11, 675-705.
- Carvalho, A., Boulangé, B., Melfi, A., & Lucas, Y. (1997). *Brazilian Bauxites*. São Paulo: ORSTOM.
- Chaves, A. P. (1999). *Teoria e prática do tratamento de minérios - britagem e moagem*. (Vol. 3). São Paulo: Signus.

- Chayes, F. (1960). On correlation between variables of constant sum. *Journal of Geophysical Research*, 65, 4185-4193.
- Chayes, F. (1971). *Ratio correlation*. Chicago: Chicago Press.
- Chilès, J. (1977). *Géostatistique des phénomènes non stationnaires (dans le plan)*. Tese (Doutorado), Université de Nancy I, Nancy.
- Chilès, J. (1984). Simulation of a Nickel Deposit: Problems Encountered and Practical Solutions. *Geostatistics for Natural Resources Characterization*, 1015-1030.
- Chilès, J. (2012). Validity Range of the Discret Gaussian Change-of-Support Model and its variant. *Geostats 2012*. Oslo: Ninth International Geostatistics Congress.
- Chilès, J., & Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty*. Nova Iorque: John Wiley & Sons.
- Chilès, B., & Delfiner, P. (1999). *Geostatistics: modeling spatial uncertainty*. New York: Wiley-Interscience.
- Clark, I., & White, B. (1976). Geostatistical Modelling of an Orebody as an Aid to Mine Planning. *14 International Symposium on Computer Applications in the Mineral Industries*, (pp. 1004-1012). Colorado.
- Dagbert, M. (1981). The Simulation of Space Dependent Data in Geology. In R. e. Craig, *Future Trends in Geomathematics* (pp. 29-47). London: Pion Limited.
- David, M. (1973). Tools for Planning: Variances and Conditional Simulations. *11 International Symposium on Computer Applications in Mineral Industries*, (pp. D10-D23). Arizona.
- David, M. (1977). *Geostatistical ore reserve estimation*. Elsevier Scientific Publishing Company.
- David, M. (1988). *Handbook of applied advanced geostatistical ore reserve estimation*. Amsterdam: Elsevier.
- David, M., Dagbert, M., Sergerie, G., & Cupic, F. (1984). Complete estimation of the tonnage, shape and grade of a Saskatchewan Uranium deposit. *27th International Geology Congress* (pp. 154-186). VNU Science Press.

- David, M., Dowd, P., & Korobov, S. (1974). Forecasting Departure from Planning in Open Pit Design and Grade Control. *12 International Symposium on Computer Applications in the Mineral Industries*, (pp. F131-F153). Colorado.
- Davis, J. (1986). *Statistics and Data Analysis en Geology*. New York: Wiley.
- Davis, M. (1987). Production of conditional simulations via the LU decomposition of the covariance matrix. *Mathematical Geology*, *19*(2), 91-98.
- Definer, P. (1976). Linear estimation of non stationary phenomena. In M. D. Guarascio, *Advanced geostatistics in the mining industry* (pp. 49-68). Roma: D.Reidel Publishing Company.
- Deraisme, J., De Fouquet, C., & Fraise, H. (1984). Geostatistical Orebody Model for Computer Optimization of Profits from Different Underground Mining Methods. *18 International Symposium on Computer Applications on the Mineral Industries*, (pp. 583-590). London.
- Deraisme, J., Rivoirard, J., & Castelli, P. (2008). Multivariate Uniform Conditioning and Block Simulations with Discrete Gaussian Model: Application to Chuquicamata Deposit. *Geostats*. Santiago.
- Desbarats, A., & Dimitrakopoulos, R. (2000). Geostatistical simulation of regionalized pore-size distributions using Min/Max autocorrelation factors. *Mathematical Geology*, *32*(8), 919-942.
- Deutsch, C. &. (1998). *GSLIB: Geostatistical Software Library and User's Guide*. . New York: Oxford University Press.
- Dimitrakopoulos, R. &. (2004). Uncertainty based production scheduling in open pit mining. In *SME Transactions* (p. 316).
- Dimitrakopoulos, R. &. (2003). Assessing risk in grade-tonnage curves in a complex copper deposit, Northern Brazil, based on an efficient joint simulation of multiple correlated variables. *31st International Symposium on the Application of Computers and Operations Research in the Minerals Industries*. Cape Town.
- Dimitrakopoulos, R. (1990). Conditional simulation of intrinsic random functions of order k. *Mathematical geology*, *22*(3), 361-380.

- Dowd, P. (1983). Conditional simulation of inter related beds in an oil deposit. In G. D. Verly, *Geostatistics for natural resources characterization* (pp. 1031-1043). Reidel Publishing Co.
- Dowd, P. (1988). Generalised cross-covariances. In M. Armstrong, *Geostatistics* (pp. 151-162). Avignon: Kluwer Academic Publisher.
- Ecole des Mines de Paris & Geovariances. (2012). *Isatis Software - Users guide*. Fontainebleau: Geovariances.
- Egozcue, J., & Pawlowsky-Glahn, V. (2005). Groups of parts and their balances in compositional data analysis. *Mathematical Geology*, 37(7), 795-828.
- Egozcue, J., Pawlowsky-Glahn, V., Mateu-Figueras, G., & Barceló-Vidal, C. (2003). Isometric logratios transformations for compositional data analysis. *Mathematical geology*, 35(3), 279-300.
- Emery, X., & Ortiz, J. (2005). Internal Consistency and Inference of Change-of-Support Isofactorial Models. In O. a. Leuangthong (Ed.), *Geostatistics Banff. 2*, pp. 1057-1056. Banff: Springer.
- Fry, J., Fry, T., & McLaren, K. (1996). Compositional data analysis and zeros in micro data. *Centre of Polici Studies (COPS)*.
- Galton, F. (1879). The geometric mean, in vital and social statistics. *Proc.R.Soc.Lond.*, (pp. 365-366).
- Gambin, F., Costa, J., & Koppe, J. (2005). Estratégia de Controle de Qualidade de Minério na Lavra Utilizando Simulação Geoestatística. *REM - Revista Escola de Minas*, 193-200.
- Gneiting, T. (1999, Feb.). The correlation bias for two-dimensional simulations by turning bands. *Mathematical Geology*, 195-211.
- Godoy, M. (2002). *The effective management of geological risk in long-term scheduling for open pit mines (PhD tesis)*. Brisbane: University of Queensland.
- Gomez-Hernandez, J. (1992). Regularization of hydraulic conductivities: a numerical approach . In A. Soares, *Geostatistics Troia'92* (pp. 767-778). Troia: Kluwer Academic Publishers.

- Gomez-Hernandez, J., & Journel, A. (1993). Joint sequential simulation of multi-gaussian fields. In S. & Amilcar, *Geostatistics* (pp. 85-94). Dordrecht: Kluwer Academic Publishers.
- Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*. New York: Oxford University Press.
- Goovarts, P. (1993). Spatial orthogonality of the principal components computed from coregionalized variables. *Mathematical Geology*, 25(13), 281-302.
- Green, A., Berman, M., Switzer, P., & Craig, M. (1988). A transformation for ordering multispectral data in terms of image quality with implication for noise removal. *IEEE transactions on geoscience and remote sensing*, 65-74.
- Guibal, D., Humphryes, M., Sanguinetti, H., & Srivastava, P. (1996). Geostatistical conditional simulation of a large iron orebody of the Pilbara region in Western Australia. In E. & Baafi, *Geostatistics Wollongong* (pp. 695-706). Wollongong: Kluwer Academic Press.
- Gutjahr, A., Bullard, B., & Hatch, S. (1997). General joint conditional simulations using a fast Fourier transform method. *Mathematical Geology*, 29(3), 361-389.
- Gutjahr, A., Hatch, S., & Bullard, B. (1994). Joint conditional simulations and flow modelling. In R. Dimitrakopoulos, *Geostatistics for the next century* (pp. 185-196). Montreal: Kluwer academic press.
- Horta, A., & Soares, A. (2010). Direct Sequential Co-simulation with joint probability distributions. *Mathematical Geoscience*, 42, 269-292.
- Isaaks, E., & Srivastava, R. (1989). *An Introduction to Applied Geostatistics*. New York: Oxford University Press.
- Johnson, M. (1987). *Multivariate statistical simulation*. New York: John Wiley & Sons.
- Johnson, R., & Wichern, D. (1998). *Applied Multivariate Statistical Analysis*. New Jersey: Prentice Hall.
- Journel, A. (1974). Geostatistics for conditional simulation of ore bodies, n.5, p. 673-687. 1974. *Economic Geology*, v. 69, 673-687.

- Journel, A. (1994). Geostatistics for Conditional Simulation of Ore Bodies. *Economic Geology*, 673-687.
- Journel, A. (1999). Markov Models for Cross-Covariances. *Mathematical Geology*, 31(8), 955-964.
- Journel, A., & Huijbregts, C. (1978). *Mining Geostatistics*. London: Academic Press.
- Kenedy, W., & Gentle, J. (1980). *Statistical Computing*. New York: Marcel Dekker.
- Koppe, V., Costa, J., Peroni, R., & Koppe, J. (2011). Choosing Between Two Kinds of Sampling Patterns Using Geostatistical Simulation. *Natural Resources Research*, 131-142.
- Krige, D. (1951). *A statistical approach to some mine valuation and allied problems on the Witwatersrand [dissertação de mestrado]*. South Africa: University of South Africa.
- Krumbein, W. (1962). Open and closed number systems in stratigraphic mapping. *American Association of Petroleum Geologists Bulletin*, 2229-2245.
- Lajaunie, C. (1992). Indicator principal component kriging. *Mathematical Geology*, 24(5), 555-565.
- Lantuejoul, C. (2002). *Geostatistical Simulation, models and algorithms*. Berlin: Springer-Verlag.
- Leuangthong, O., Lyall, G., & Deutsch, C. (2002). Multivariate simulation of a nickel laterite deposit. *APCOM*, (pp. 261-173). Phoenix.
- Lloyd, C., Pawlowsky-Glahn, V., & Egozcue, J. (2012). Compositional Data Analysis in Population Studies. *Annals of the Association of American Geographers* (pp. 1251-1266). Taylor & Francis.
- Marcotte, D. (1993). Direct simulation of block grades. In R. Dimitrakopoulos, *Geostatistics for the next century* (pp. 245-258). Montreal: Kluwer Academic Publishers.
- Marcotte, D., Naraghy, K., Bellehumeur, C., & Gloaguen, E. (2005). An application of multivariate simulation in the cement industry. *Mathematical Geology*, 37(5), 493-512.

- Marechal, A. (1970). *Cokrigage et Regression em Correlation Intrinsique*. Fontainebleau: Centre de Geostatistique de Fontainebleau.
- Marques, D. C., & Koppe, J. (2010). A Comprovação da Relação VolumexVariância na Homogeneização da Sílica no Minério de Ferro. *REM - Revista Escola de Minas*, 355-361.
- Martín-Fernandez, J. (2001). *Medidas de diferencia y clasificación no paramétrica de dados composicionales [tese de doutorado]*. Barcelona: Universidade Politècnica da Catalunya.
- Martín-Fernandez, J., Barceló-Vidal, C., & Pawlowsky-Glahn, V. (2000). Zero Replacement in Compositional Data Sets. *Studies in Classification, Data Analysis, and Knowledge Organization (Proceedings of the 7th Conference of the International Federation of Classification Societies (IFCS'2000))* (pp. 155-160). Namur: Springer-Verlag.
- Martín-Fernandez, J., Barceló-Vidal, C., & Pawlowsky-Glahn, V. (2003). Dealing With Zeros and Missing Values in Compositional Data Sets Using Non-parametric Imputation. *Mathematical Geology*, 35(3), 253-278.
- Mateu-Figueras, G. (2003). *Models de distribuició sobre el simplex: PhD thesis*. Barcelona: Universitat Politècnica da Catalunya.
- Matheron, G. (1965). *Les variables regionalisées et leur estimation*. Paris: Ed. Masson.
- Matheron, G. (1973). The intrinsic random functions and their application. *Advances in Applied Probability*, 439-468.
- Matheron, G. (1975). *Les fonctions de transfert de petits panneaux*. França: CGMM.
- Matheron, G. (1976). Forecasting block grade distributions: the transfer functions. In M. D. Guarascio, *Advanced geostatistics in the mining industry* (pp. 237-251). Dordrecht: Reidel Publishing company.
- Mosimann, J. (1962). On the compound multinomial distribution, the multivariate beta-distribution and correlations among proportions. *Biometrika*, 65-82.
- Mueller, U., & Ferreira, J. (2012). The U-Wedge Transformation Method for Multivariate Geostatistical Simulation. *Mathematical Geoscience*, 44, 427-448.

- Mueller, U., & Ward, C. (2012). Multivariate Estimation Using Log Ratios: A Worked Alternative. *Geostatistics Oslo 2012, Quantitative Geology and Geostatistics*, 333-343.
- Mueller, U., Tolosana Delgado, R., & van den Boogaart, K. (2014). Approaches to the Simulation of Compositional Data – A Nickel-Laterite Comparative Case Study. *Proceedings Orebody Modelling and Strategic Mine Planning Symposium* (pp. 61-72). Melbourne: The Australasian Institute of Mining and Metallurgy.
- Myers, D. (1988). Vector Conditional Simulation. In M. Armstrong, *Geostatistics* (pp. 283-292). Avignon: Kluwer Academic Publishers.
- Pardo-Iguzquiza, E., & Chica-Olmo, M. (1993). The Fourier integral method - an efficient spectral method for simulation of random fields. *Mathematical Geology*, 25, 177-217.
- Pardo-Iguzquiza, E., & Chica-Olmo, M. (1994). Spectral simulation of multivariable stationary random functions using covariance Fourier transforms. *Mathematical Geology*, 26(3), 277-299.
- Pawlowisky-Glahn, V., & Egozcue, J. (2001). Geometric approach to statistical analysis on the simplex: Stochastic Environ. Res. Risk Assess. *SERRA*, 384-398.
- Pawlowisky-Glahn, V., & Buccianti, A. (2011). *Compositional Data Analysis: Theory and applications*. John Wiley & Sons, Ltd.
- Pawlowisky-Glahn, V. (2003). Stistical modelling on coordinates.
- Pawlowisky-Glahn, V., & Egozcue, J. (2002). Blu Estimators and Compositional Data. *Mathematical Geology*, 34(3), 259-274.
- Pawlowisky-Glahn, V., & Egozcue, J. (2011). Exploring Compositional Data with the Coda-Dendogram . *Austrian Journal of Statistics*, 103-113.
- Pawlowisky-Glahn, V., & Olea, R. (2004). *Geostatistical Analysis of Compositional Data, Studies in Mathematical Geology*. Oxford University Press.
- Pawlowisky-Glahn, V., Egozcue, J., & Tolosana-Delgado, R. (2013). *Lectures Notes on Compositional Data Analysis with an Introduction to Compositional Geostatistics*.

- Pearson, K. (1897). Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London* (pp. 489-502). London: LX.
- Pilger, G. (2000). *Critérios para Locação Amostral Baseados em Simulação Estocástica. Dissertação de mestrado*. Porto Alegre: Programa de Pós-Graduação em Engenharia de Minas, Metalúrgica e de Materiais (PPGEM/UFRGS).
- Ripley, B. (1987). *Stochastic simulation*. New York: John Wiley & Sons.
- Rivoirard, J. (1994). *Introduction to disjunctive kriging and non-linear geostatistics*. Clarendon Press.
- Rivoirard, J. (2002). Which Models for Collocated Cokriging? *Mathematical Geology*, 117-131.
- Rivoirard, J. (2004). On some simplifications of cokriging neighborhood. *Mathematical Geology*, 899-915.
- Robin, M., Gutjahr, A., Sudicky, E., & Wilson, J. (1993). Cross-correlated random-field generation with the direct fourier-transform method. *Water Resources Research*, 2385-2397.
- Rock, N. (1988). *Numerical Geology*. Springer-Verlag.
- Rollinson, H. (1995). *Using geochemical data: Evaluation, presentation, interpretation*. Longman Geochemistry Series, Longman Group Ltd.
- Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Annals of mathematical statistics*, (pp. 470-472).
- Rubio, R., Costa, J., & Bassani, M. (2016). A Geostatistical Framework for Estimating Compositional Data Avoiding Bias in Back Transformation. *REM - Revista Escola de Minas*, 219-226.
- Sarmanov, O., & Vistelius, A. (1959). On the correlation of percentage values. *Doklady of the Academy of Sciences of the USSR; Earth Science Section*, 22-25.
- Shive, P., Lowry, T., Easley, D., & Borgman, L. (1990). Geostatistical simulation for geophysical applications. *Geophysical modeling*, 1441-1446.

- Soares, A. (2001). Direct sequential simulation and cosimulation. *Mathematical Geology*, 33(8), 911-926.
- Souza, L., Costa, J., & Koppe, J. (2004). Uncertainty Estimate in Resources Assessment: A Geostatistical Contribution. *Natural Resources Research*, 1-5.
- Suro-Perez, V., & Journel, A. (1990). Stochastic simulation of lithofacies: an improved sequential indicator approach. *2nd European conference on the mathematics of oil recovery*, (pp. 3-10).
- Suro-Perez, V., & Journel, A. (1991). Indicator principal component kriging. *Mathematical Geology*, 23(5), 759-788.
- Switzer, P., & Green, A. A. (1984). *Min/Max autocorrelation factors for multivariate spatial imagery*. Stanford: Stanford University Press.
- Tercan, A. (1999). Importance of orthogonalization algorithm in modeling conditional distributions by orthogonal indicator methods. *Mathematical Geology*, 31(2), 155-173.
- Tolosana Delgado, R. M., van den Boogaart, K., C., W., & Gutzmer, J. (2015). Improving processing by adaption to conditional geostatistical simulation of block compositions. *Journal of the Southern African Institute of Mining and Metallurgy*, 13-26.
- Tolosana Delgado, R., Mueller, U., & van den Boogaart, K. (2016). Compositionally Compliant Contact Analysis. *Geostatistical and Geospatial Approaches for the Characterization of Natural Resources in the Environment - Challenges, Processes and Strategies* (pp. 11-14). Springer International Publishing.
- Tolosana Delgado, R., Mueller, U., van den Boogaart, K., & Ward, C. (2014). Compositional Block Cokriging. *Mathematics of Planet Earth - Proceedings of the 15th Annual Conference of the International Association for Mathematical Geosciences* (pp. 713-716). Springer Berlin Heidelberg.
- Tolosana-Delgado, R. (2006). *Geostatistics for constrained variables: positive data, compositions and probabilities. Application to Environmental Hazard Monitoring [tese de doutorado]*. University of Girona.

- Tolosana-Delgado, R., Mueller, U., Boogaart, v. d., & Ward, C. (2013). Block Cokriging of a Whole Composition. *Proceedings of APCOM 2013*, (pp. 267-277). Porto Alegre.
- Tozzo-Martins, A., Ribeiro Júnior, P., & Bonat, W. (2009). Um modelo geoestatístico bivariado para dados composicionais. *Revista Brasileira de Biometria*, 456-477.
- Verly, G. (1993). Sequential gaussian cosimulation; a simulation method integrating several types of information. In S. & Amilcar, *Geostatistics* (pp. 543-554). Dordrecht: Kluwer Academic Publishers.
- Wackernagel, H. (1994). *Multivariate Geostatistics. An Introduction with Applications*. Berlin: Springer-Verlag.
- Wackernagel, H., Petitgas, Y., & Touffait, Y. (1989). Overview of methods for coregionalization analyses. In M. Armstrong, *Geostatistics* (pp. 409-420). Kluwer Academic.
- Xu, W., Tran, R., Srivastava, M., & Journel, A. (1992). The collocated cokriging alternative. *Anais do 67th Annual Technical Conference of The Society of Petroleum Engineers*, (pp. 833-842).

# ANEXO A – VERIFICAÇÃO DA REPRODUÇÃO DAS CARACTERÍSTICAS ESTATÍSTICAS DE CONTINUIDADE ESPACIAL

## 1. MÉDIAS GLOBAIS

As tabelas a seguir (A.1 a A.24) apresentam as médias globais obtidas para os vinte cenários simulados em cada combinação testada e sua semelhança com as médias das variáveis nos dados originais.

### 1.1. MÉDIAS GLOBAIS – COMBINAÇÃO 1

TABELA A.2: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL ALAP (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão (%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>13,07</b>	<b>58,78</b>	<b>49,17</b>	<b>3,83</b>
<i>Alap[00001]</i>	69485	13,07	58,78	48,69	3,84
<i>Alap[00002]</i>	69485	13,07	58,78	48,72	3,89
<i>Alap[00003]</i>	69485	13,07	58,78	49,03	3,86
<i>Alap[00004]</i>	69485	13,07	58,78	48,84	3,87
<i>Alap[00005]</i>	69485	13,07	58,78	49,13	3,84
<i>Alap[00006]</i>	69485	13,07	58,78	48,73	3,81
<i>Alap[00007]</i>	69485	13,07	58,78	48,77	3,86
<i>Alap[00008]</i>	69485	13,07	58,78	49,12	3,85
<i>Alap[00009]</i>	69485	13,07	58,78	48,71	3,87
<i>Alap[00010]</i>	69485	13,07	58,78	48,77	3,89
<i>Alap[00011]</i>	69485	13,07	58,78	48,82	3,84
<i>Alap[00012]</i>	69485	13,07	58,78	49,15	3,86
<i>Alap[00013]</i>	69485	13,07	58,78	48,75	3,81
<i>Alap[00014]</i>	69485	13,07	58,78	49,02	3,85
<i>Alap[00015]</i>	69485	13,07	58,78	48,72	3,87
<i>Alap[00016]</i>	69485	13,07	58,78	48,78	3,88
<i>Alap[00017]</i>	69485	13,07	58,78	49,01	3,89
<i>Alap[00018]</i>	69485	13,07	58,78	48,72	3,83
<i>Alap[00019]</i>	69485	13,07	58,78	49,9	3,81
<i>Alap[00020]</i>	69485	13,07	58,78	48,72	3,83

TABELA A.3: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL FE (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão (%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,69</b>	<b>54,65</b>	<b>9,49</b>	<b>4,98</b>
<i>Fe[00001]</i>	69485	0,69	54,65	8,97	4,99
<i>Fe[00002]</i>	69485	0,69	54,65	9,98	4,97
<i>Fe[00003]</i>	69485	0,69	54,65	8,91	4,93
<i>Fe[00004]</i>	69485	0,69	54,65	9,08	4,95
<i>Fe[00005]</i>	69485	0,69	54,65	9,09	4,96
<i>Fe[00006]</i>	69485	0,69	54,65	9,03	9,91
<i>Fe[00007]</i>	69485	0,69	54,65	9,05	4,99
<i>Fe[00008]</i>	69485	0,69	54,65	8,9	4,97
<i>Fe[00009]</i>	69485	0,69	54,65	9,85	4,96
<i>Fe[00010]</i>	69485	0,69	54,65	9,01	5
<i>Fe[00011]</i>	69485	0,69	54,65	9,06	5,06
<i>Fe[00012]</i>	69485	0,69	54,65	9,92	5,04
<i>Fe[00013]</i>	69485	0,69	54,65	9,12	4,97
<i>Fe[00014]</i>	69485	0,69	54,65	9,96	4,96
<i>Fe[00015]</i>	69485	0,69	54,65	9,06	5,01
<i>Fe[00016]</i>	69485	0,69	54,65	9,13	5,06
<i>Fe[00017]</i>	69485	0,69	54,65	8,91	4,89
<i>Fe[00018]</i>	69485	0,69	54,65	9,17	4,96
<i>Fe[00019]</i>	69485	0,69	54,65	9,99	4,99
<i>Fe[00020]</i>	69485	0,69	54,65	9,98	5,02

TABELA A.4: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL RESTO (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>21,47</b>	<b>52,32</b>	<b>35,76</b>	<b>2,88</b>
<i>Resto[00001]</i>	69485	21,47	52,32	35,63	2,84
<i>Resto[00002]</i>	69485	21,47	52,32	35,63	2,86
<i>Resto[00003]</i>	69485	21,47	52,32	35,66	2,87
<i>Resto[00004]</i>	69485	21,47	52,32	35,66	2,89
<i>Resto[00005]</i>	69485	21,47	52,32	35,64	2,54
<i>Resto[00006]</i>	69485	21,47	52,32	35,60	2,86
<i>Resto[00007]</i>	69485	21,47	52,32	35,70	2,84

<i>Resto[00008]</i>	69485	21,47	52,32	35,62	2,89
<i>Resto[00009]</i>	69485	21,47	52,32	35,66	2,85
<i>Resto[00010]</i>	69485	21,47	52,32	35,67	2,83
<i>Resto[00011]</i>	69485	21,47	52,32	35,63	2,81
<i>Resto[00012]</i>	69485	21,47	52,32	35,68	2,87
<i>Resto[00013]</i>	69485	21,47	52,32	35,59	2,86
<i>Resto[00014]</i>	69485	21,47	52,32	35,69	2,89
<i>Resto[00015]</i>	69485	21,47	52,32	35,63	2,85
<i>Resto[00016]</i>	69485	21,47	52,32	35,63	2,84
<i>Resto[00017]</i>	69485	21,47	52,32	35,76	2,85
<i>Resto[00018]</i>	69485	21,47	52,32	35,59	2,86
<i>Resto[00019]</i>	69485	21,47	52,32	35,60	2,81
<i>Resto[00020]</i>	69485	21,47	52,32	35,63	2,96

TABELA A.5: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL SI (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b><i>Dados originais</i></b>	<b><i>3735</i></b>	<b><i>0,27</i></b>	<b><i>14,90</i></b>	<b><i>4,44</i></b>	<b><i>1,97</i></b>
<i>Si[00001]</i>	69485	0,27	14,9	4,55	1,99
<i>Si[00002]</i>	69485	0,27	14,9	4,37	1,96
<i>Si[00003]</i>	69485	0,27	14,9	4,56	1,98
<i>Si[00004]</i>	69485	0,27	14,9	4,58	1,97
<i>Si[00005]</i>	69485	0,27	14,9	4,37	1,96
<i>Si[00006]</i>	69485	0,27	14,9	4,56	1,99
<i>Si[00007]</i>	69485	0,27	14,9	4,58	2
<i>Si[00008]</i>	69485	0,27	14,9	4,55	2,01
<i>Si[00009]</i>	69485	0,27	14,9	4,47	1,97
<i>Si[00010]</i>	69485	0,27	14,9	4,57	1,99
<i>Si[00011]</i>	69485	0,27	14,9	4,38	1,98
<i>Si[00012]</i>	69485	0,27	14,9	4,55	1,94
<i>Si[00013]</i>	69485	0,27	14,9	4,58	2,05
<i>Si[00014]</i>	69485	0,27	14,9	4,57	2,03
<i>Si[00015]</i>	69485	0,27	14,9	4,27	2,00
<i>Si[00016]</i>	69485	0,27	14,9	4,58	2,01
<i>Si[00017]</i>	69485	0,27	14,9	4,37	1,99
<i>Si[00018]</i>	69485	0,27	14,9	4,56	1,98
<i>Si[00019]</i>	69485	0,27	14,9	4,46	1,96
<i>Si[00020]</i>	69485	0,27	14,9	4,55	1,94

TABELA A.6: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL TI (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,27</b>	<b>3,57</b>	<b>1,53</b>	<b>0,45</b>
<i>Ti[00001]</i>	69485	0,27	3,57	1,45	0,47
<i>Ti[00002]</i>	69485	0,27	3,57	1,56	0,48
<i>Ti[00003]</i>	69485	0,27	3,57	1,45	0,49
<i>Ti[00004]</i>	69485	0,27	3,57	1,56	0,45
<i>Ti[00005]</i>	69485	0,27	3,57	1,46	0,41
<i>Ti[00006]</i>	69485	0,27	3,57	1,55	0,46
<i>Ti[00007]</i>	69485	0,27	3,57	1,45	0,42
<i>Ti[00008]</i>	69485	0,27	3,57	1,44	0,47
<i>Ti[00009]</i>	69485	0,27	3,57	1,44	0,46
<i>Ti[00010]</i>	69485	0,27	3,57	1,55	0,42
<i>Ti[00011]</i>	69485	0,27	3,57	1,46	0,41
<i>Ti[00012]</i>	69485	0,27	3,57	1,54	0,48
<i>Ti[00013]</i>	69485	0,27	3,57	1,46	0,41
<i>Ti[00014]</i>	69485	0,27	3,57	1,45	0,49
<i>Ti[00015]</i>	69485	0,27	3,57	1,55	0,45
<i>Ti[00016]</i>	69485	0,27	3,57	1,46	0,46
<i>Ti[00017]</i>	69485	0,27	3,57	1,44	0,47
<i>Ti[00018]</i>	69485	0,27	3,57	1,56	0,44
<i>Ti[00019]</i>	69485	0,27	3,57	1,45	0,42
<i>Ti[00020]</i>	69485	0,27	3,57	1,55	0,46

## 1.2. MÉDIAS GLOBAIS – COMBINAÇÃO 2

TABELA A.7: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL ALAP (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>13,07</b>	<b>58,78</b>	<b>49,17</b>	<b>3,83</b>
<i>Alap[00001]</i>	69485	13,07	58,78	49,62	3,84
<i>Alap[00002]</i>	69485	13,07	58,78	49,76	3,86
<i>Alap[00003]</i>	69485	13,07	58,78	48,58	3,85
<i>Alap[00004]</i>	69485	13,07	58,78	49,88	3,88
<i>Alap[00005]</i>	69485	13,07	58,78	48,96	3,82
<i>Alap[00006]</i>	69485	13,07	58,78	48,79	3,81
<i>Alap[00007]</i>	69485	13,07	58,78	49,68	3,80

<i>Alap[00008]</i>	69485	13,07	58,78	49,47	3,83
<i>Alap[00009]</i>	69485	13,07	58,78	48,32	3,87
<i>Alap[00010]</i>	69485	13,07	58,78	49,62	3,89
<i>Alap[00011]</i>	69485	13,07	58,78	48,95	3,84
<i>Alap[00012]</i>	69485	13,07	58,78	49,39	3,86
<i>Alap[00013]</i>	69485	13,07	58,78	48,95	3,85
<i>Alap[00014]</i>	69485	13,07	58,78	49,39	3,80
<i>Alap[00015]</i>	69485	13,07	58,78	48,61	3,84
<i>Alap[00016]</i>	69485	13,07	58,78	48,92	3,86
<i>Alap[00017]</i>	69485	13,07	58,78	49,44	3,84
<i>Alap[00018]</i>	69485	13,07	58,78	48,89	3,89
<i>Alap[00019]</i>	69485	13,07	58,78	49,64	3,82
<i>Alap[00020]</i>	69485	13,07	58,78	49,43	3,81

TABELA A.8: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL FE (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b><i>Dados originais</i></b>	<b><i>3735</i></b>	<b><i>0,69</i></b>	<b><i>54,65</i></b>	<b><i>9,49</i></b>	<b><i>4,98</i></b>
<i>Fe[00001]</i>	69485	0,69	54,65	9,56	4,99
<i>Fe[00002]</i>	69485	0,69	54,65	9,55	5,01
<i>Fe[00003]</i>	69485	0,69	54,65	9,54	4,98
<i>Fe[00004]</i>	69485	0,69	54,65	9,53	4,97
<i>Fe[00005]</i>	69485	0,69	54,65	9,55	4,94
<i>Fe[00006]</i>	69485	0,69	54,65	9,53	4,96
<i>Fe[00007]</i>	69485	0,69	54,65	9,5	4,98
<i>Fe[00008]</i>	69485	0,69	54,65	9,55	4,99
<i>Fe[00009]</i>	69485	0,69	54,65	9,52	5,01
<i>Fe[00010]</i>	69485	0,69	54,65	9,54	5,03
<i>Fe[00011]</i>	69485	0,69	54,65	9,54	5,01
<i>Fe[00012]</i>	69485	0,69	54,65	9,53	4,99
<i>Fe[00013]</i>	69485	0,69	54,65	9,52	4,97
<i>Fe[00014]</i>	69485	0,69	54,65	9,51	4,96
<i>Fe[00015]</i>	69485	0,69	54,65	9,53	4,95
<i>Fe[00016]</i>	69485	0,69	54,65	9,54	4,91
<i>Fe[00017]</i>	69485	0,69	54,65	9,49	4,93
<i>Fe[00018]</i>	69485	0,69	54,65	9,58	4,92
<i>Fe[00019]</i>	69485	0,69	54,65	9,56	4,91
<i>Fe[00020]</i>	69485	0,69	54,65	9,53	4,93

TABELA A.9: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL RESTO (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão (%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>21,47</b>	<b>52,32</b>	<b>35,76</b>	<b>2,88</b>
<i>Resto[00001]</i>	69485	21,47	52,32	36,84	2,89
<i>Resto[00002]</i>	69485	21,47	52,32	36,71	2,87
<i>Resto[00003]</i>	69485	21,47	52,32	35,84	2,90
<i>Resto[00004]</i>	69485	21,47	52,32	36,61	2,84
<i>Resto[00005]</i>	69485	21,47	52,32	36,57	2,89
<i>Resto[00006]</i>	69485	21,47	52,32	35,67	2,84
<i>Resto[00007]</i>	69485	21,47	52,32	36,79	2,85
<i>Resto[00008]</i>	69485	21,47	52,32	36,05	2,83
<i>Resto[00009]</i>	69485	21,47	52,32	36,14	2,86
<i>Resto[00010]</i>	69485	21,47	52,32	36,86	2,81
<i>Resto[00011]</i>	69485	21,47	52,32	36,53	2,86
<i>Resto[00012]</i>	69485	21,47	52,32	36,08	2,87
<i>Resto[00013]</i>	69485	21,47	52,32	35,54	2,86
<i>Resto[00014]</i>	69485	21,47	52,32	36,04	2,80
<i>Resto[00015]</i>	69485	21,47	52,32	36,80	2,84
<i>Resto[00016]</i>	69485	21,47	52,32	36,51	2,86
<i>Resto[00017]</i>	69485	21,47	52,32	36,02	2,84
<i>Resto[00018]</i>	69485	21,47	52,32	35,62	2,86
<i>Resto[00019]</i>	69485	21,47	52,32	35,77	2,85
<i>Resto[00020]</i>	69485	21,47	52,32	35,94	2,84

TABELA A.10: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL SI (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão (%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,27</b>	<b>14,90</b>	<b>4,44</b>	<b>1,97</b>
<i>Si[00001]</i>	69485	0,27	14,9	4,38	1,99
<i>Si[00002]</i>	69485	0,27	14,9	4,39	1,98
<i>Si[00003]</i>	69485	0,27	14,9	4,36	1,99
<i>Si[00004]</i>	69485	0,27	14,9	4,34	1,96
<i>Si[00005]</i>	69485	0,27	14,9	4,38	1,97
<i>Si[00006]</i>	69485	0,27	14,9	4,38	2,1
<i>Si[00007]</i>	69485	0,27	14,9	4,47	1,97
<i>Si[00008]</i>	69485	0,27	14,9	4,37	1,99
<i>Si[00009]</i>	69485	0,27	14,9	4,38	2,02
<i>Si[00010]</i>	69485	0,27	14,9	4,46	2,03

<i>Si[00011]</i>	69485	0,27	14,9	4,37	1,99
<i>Si[00012]</i>	69485	0,27	14,9	4,47	1,98
<i>Si[00013]</i>	69485	0,27	14,9	4,46	1,97
<i>Si[00014]</i>	69485	0,27	14,9	4,37	2,00
<i>Si[00015]</i>	69485	0,27	14,9	4,36	2,01
<i>Si[00016]</i>	69485	0,27	14,9	4,37	2,03
<i>Si[00017]</i>	69485	0,27	14,9	4,47	2,05
<i>Si[00018]</i>	69485	0,27	14,9	4,36	1,98
<i>Si[00019]</i>	69485	0,27	14,9	4,48	2,97
<i>Si[00020]</i>	69485	0,27	14,9	4,37	2,96

TABELA A.11: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL TI (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,27</b>	<b>3,57</b>	<b>1,53</b>	<b>0,45</b>
<i>Ti[00001]</i>	69485	0,27	3,57	1,57	0,44
<i>Ti[00002]</i>	69485	0,27	3,57	1,57	0,45
<i>Ti[00003]</i>	69485	0,27	3,57	1,57	0,46
<i>Ti[00004]</i>	69485	0,27	3,57	1,56	0,41
<i>Ti[00005]</i>	69485	0,27	3,57	1,56	0,49
<i>Ti[00006]</i>	69485	0,27	3,57	1,56	0,47
<i>Ti[00007]</i>	69485	0,27	3,57	1,56	0,45
<i>Ti[00008]</i>	69485	0,27	3,57	1,57	0,46
<i>Ti[00009]</i>	69485	0,27	3,57	1,57	0,47
<i>Ti[00010]</i>	69485	0,27	3,57	1,55	0,44
<i>Ti[00011]</i>	69485	0,27	3,57	1,54	0,42
<i>Ti[00012]</i>	69485	0,27	3,57	1,58	0,43
<i>Ti[00013]</i>	69485	0,27	3,57	1,54	0,45
<i>Ti[00014]</i>	69485	0,27	3,57	1,58	0,44
<i>Ti[00015]</i>	69485	0,27	3,57	1,56	0,48
<i>Ti[00016]</i>	69485	0,27	3,57	1,55	0,46
<i>Ti[00017]</i>	69485	0,27	3,57	1,57	0,47
<i>Ti[00018]</i>	69485	0,27	3,57	1,55	0,44
<i>Ti[00019]</i>	69485	0,27	3,57	1,57	0,43
<i>Ti[00020]</i>	69485	0,27	3,57	1,57	0,44

### 1.3. MÉDIAS GLOBAIS – COMBINAÇÃO 3

TABELA A.12: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL ALAP (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>13,07</b>	<b>58,78</b>	<b>49,17</b>	<b>3,83</b>
<i>Alap[00001]</i>	69485	13,07	58,78	48,33	3,86
<i>Alap[00002]</i>	69485	13,07	58,78	50,02	3,84
<i>Alap[00003]</i>	69485	13,07	58,78	49,58	3,86
<i>Alap[00004]</i>	69485	13,07	58,78	47,63	3,85
<i>Alap[00005]</i>	69485	13,07	58,78	48,88	3,84
<i>Alap[00006]</i>	69485	13,07	58,78	47,93	3,86
<i>Alap[00007]</i>	69485	13,07	58,78	47,93	3,85
<i>Alap[00008]</i>	69485	13,07	58,78	47,30	3,86
<i>Alap[00009]</i>	69485	13,07	58,78	47,58	3,85
<i>Alap[00010]</i>	69485	13,07	58,78	48,23	3,81
<i>Alap[00011]</i>	69485	13,07	58,78	49,01	3,89
<i>Alap[00012]</i>	69485	13,07	58,78	50,25	3,82
<i>Alap[00013]</i>	69485	13,07	58,78	47,38	3,85
<i>Alap[00014]</i>	69485	13,07	58,78	48,98	3,86
<i>Alap[00015]</i>	69485	13,07	58,78	48,13	3,85
<i>Alap[00016]</i>	69485	13,07	58,78	49,10	3,82
<i>Alap[00017]</i>	69485	13,07	58,78	46,35	3,84
<i>Alap[00018]</i>	69485	13,07	58,78	48,31	3,82
<i>Alap[00019]</i>	69485	13,07	58,78	49,60	3,86
<i>Alap[00020]</i>	69485	13,07	58,78	48,51	3,82

TABELA A.13: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL FE (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,69</b>	<b>54,65</b>	<b>9,49</b>	<b>4,98</b>
<i>Fe[00001]</i>	69485	0,69	54,65	9,12	4,99
<i>Fe[00002]</i>	69485	0,69	54,65	9,61	4,97
<i>Fe[00003]</i>	69485	0,69	54,65	9,54	4,96
<i>Fe[00004]</i>	69485	0,69	54,65	8,87	4,99
<i>Fe[00005]</i>	69485	0,69	54,65	9,33	4,94
<i>Fe[00006]</i>	69485	0,69	54,65	9,01	4,96

<i>Fe[00007]</i>	69485	0,69	54,65	9,01	5,02
<i>Fe[00008]</i>	69485	0,69	54,65	8,77	5,00
<i>Fe[00009]</i>	69485	0,69	54,65	8,91	5,04
<i>Fe[00010]</i>	69485	0,69	54,65	9,12	4,99
<i>Fe[00011]</i>	69485	0,69	54,65	9,35	4,98
<i>Fe[00012]</i>	69485	0,69	54,65	9,74	4,97
<i>Fe[00013]</i>	69485	0,69	54,65	8,77	4,96
<i>Fe[00014]</i>	69485	0,69	54,65	9,32	4,99
<i>Fe[00015]</i>	69485	0,69	54,65	9,09	4,98
<i>Fe[00016]</i>	69485	0,69	54,65	9,41	5,02
<i>Fe[00017]</i>	69485	0,69	54,65	8,50	5,00
<i>Fe[00018]</i>	69485	0,69	54,65	9,10	4,98
<i>Fe[00019]</i>	69485	0,69	54,65	9,45	4,99
<i>Fe[00020]</i>	69485	0,69	54,65	9,18	4,97

TABELA A.14: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL RESTO (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b><i>Dados originais</i></b>	<b><i>3735</i></b>	<b><i>21,47</i></b>	<b><i>52,32</i></b>	<b><i>35,76</i></b>	<b><i>2,88</i></b>
<i>Resto[00001]</i>	69485	21,47	52,32	36,44	2,86
<i>Resto[00002]</i>	69485	21,47	52,32	37,57	2,87
<i>Resto[00003]</i>	69485	21,47	52,32	37,12	2,88
<i>Resto[00004]</i>	69485	21,47	52,32	35,41	2,89
<i>Resto[00005]</i>	69485	21,47	52,32	36,38	2,90
<i>Resto[00006]</i>	69485	21,47	52,32	35,64	2,85
<i>Resto[00007]</i>	69485	21,47	52,32	35,64	2,86
<i>Resto[00008]</i>	69485	21,47	52,32	35,59	2,87
<i>Resto[00009]</i>	69485	21,47	52,32	35,67	2,88
<i>Resto[00010]</i>	69485	21,47	52,32	36,12	2,89
<i>Resto[00011]</i>	69485	21,47	52,32	37,06	2,90
<i>Resto[00012]</i>	69485	21,47	52,32	37,68	2,85
<i>Resto[00013]</i>	69485	21,47	52,32	35,97	2,86
<i>Resto[00014]</i>	69485	21,47	52,32	36,88	2,89
<i>Resto[00015]</i>	69485	21,47	52,32	36,45	2,85
<i>Resto[00016]</i>	69485	21,47	52,32	37,29	2,87
<i>Resto[00017]</i>	69485	21,47	52,32	35,73	2,86
<i>Resto[00018]</i>	69485	21,47	52,32	36,30	2,89
<i>Resto[00019]</i>	69485	21,47	52,32	37,75	2,84
<i>Resto[00020]</i>	69485	21,47	52,32	36,45	2,86

TABELA A.15: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL SI (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,27</b>	<b>14,90</b>	<b>4,44</b>	<b>1,97</b>
<i>Si[00001]</i>	69485	0,27	14,9	4,76	1,98
<i>Si[00002]</i>	69485	0,27	14,9	4,81	1,99
<i>Si[00003]</i>	69485	0,27	14,9	4,72	1,96
<i>Si[00004]</i>	69485	0,27	14,9	4,59	2,00
<i>Si[00005]</i>	69485	0,27	14,9	4,61	2,01
<i>Si[00006]</i>	69485	0,27	14,9	4,56	1,99
<i>Si[00007]</i>	69485	0,27	14,9	4,57	1,98
<i>Si[00008]</i>	69485	0,27	14,9	4,69	1,97
<i>Si[00009]</i>	69485	0,27	14,9	4,65	1,95
<i>Si[00010]</i>	69485	0,27	14,9	4,65	2,01
<i>Si[00011]</i>	69485	0,27	14,9	4,82	2,02
<i>Si[00012]</i>	69485	0,27	14,9	4,79	2,03
<i>Si[00013]</i>	69485	0,27	14,9	4,81	1,99
<i>Si[00014]</i>	69485	0,27	14,9	4,80	1,98
<i>Si[00015]</i>	69485	0,27	14,9	4,82	1,97
<i>Si[00016]</i>	69485	0,27	14,9	4,89	1,99
<i>Si[00017]</i>	69485	0,27	14,9	4,94	1,99
<i>Si[00018]</i>	69485	0,27	14,9	4,75	1,98
<i>Si[00019]</i>	69485	0,27	14,9	4,98	1,94
<i>Si[00020]</i>	69485	0,27	14,9	4,71	1,96

TABELA A.16: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL TI (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,27</b>	<b>3,57</b>	<b>1,53</b>	<b>0,45</b>
<i>Ti[00001]</i>	69485	0,27	3,57	1,57	0,44
<i>Ti[00002]</i>	69485	0,27	3,57	1,63	0,43
<i>Ti[00003]</i>	69485	0,27	3,57	1,61	0,46
<i>Ti[00004]</i>	69485	0,27	3,57	1,53	0,45
<i>Ti[00005]</i>	69485	0,27	3,57	1,57	0,44
<i>Ti[00006]</i>	69485	0,27	3,57	1,54	0,46
<i>Ti[00007]</i>	69485	0,27	3,57	1,54	0,47
<i>Ti[00008]</i>	69485	0,27	3,57	1,53	0,49
<i>Ti[00009]</i>	69485	0,27	3,57	1,54	0,41
<i>Ti[00010]</i>	69485	0,27	3,57	1,56	0,42

<i>Ti[00011]</i>	69485	0,27	3,57	1,61	0,44
<i>Ti[00012]</i>	69485	0,27	3,57	1,64	0,45
<i>Ti[00013]</i>	69485	0,27	3,57	1,55	0,43
<i>Ti[00014]</i>	69485	0,27	3,57	1,60	0,43
<i>Ti[00015]</i>	69485	0,27	3,57	1,58	0,49
<i>Ti[00016]</i>	69485	0,27	3,57	1,62	0,41
<i>Ti[00017]</i>	69485	0,27	3,57	1,54	0,46
<i>Ti[00018]</i>	69485	0,27	3,57	1,57	0,41
<i>Ti[00019]</i>	69485	0,27	3,57	1,64	0,49
<i>Ti[00020]</i>	69485	0,27	3,57	1,57	0,48

#### 1.4. MÉDIAS GLOBAIS – COMBINAÇÃO 4

TABELA A.17: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL ALAP (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b><i>Dados originais</i></b>	<b><i>3735</i></b>	<b><i>13,07</i></b>	<b><i>58,78</i></b>	<b><i>49,17</i></b>	<b><i>3,83</i></b>
<i>Alap[00001]</i>	69485	13,07	58,78	49,62	3,84
<i>Alap[00002]</i>	69485	13,07	58,78	49,43	3,83
<i>Alap[00003]</i>	69485	13,07	58,78	49,18	3,85
<i>Alap[00004]</i>	69485	13,07	58,78	48,75	3,81
<i>Alap[00005]</i>	69485	13,07	58,78	48,74	3,82
<i>Alap[00006]</i>	69485	13,07	58,78	48,66	3,85
<i>Alap[00007]</i>	69485	13,07	58,78	48,64	3,86
<i>Alap[00008]</i>	69485	13,07	58,78	49,13	3,82
<i>Alap[00009]</i>	69485	13,07	58,78	48,99	3,82
<i>Alap[00010]</i>	69485	13,07	58,78	49,03	3,83
<i>Alap[00011]</i>	69485	13,07	58,78	49,65	3,85
<i>Alap[00012]</i>	69485	13,07	58,78	49,28	3,84
<i>Alap[00013]</i>	69485	13,07	58,78	49,61	3,86
<i>Alap[00014]</i>	69485	13,07	58,78	49,49	3,85
<i>Alap[00015]</i>	69485	13,07	58,78	49,66	3,82
<i>Alap[00016]</i>	69485	13,07	58,78	49,74	3,81
<i>Alap[00017]</i>	69485	13,07	58,78	50,14	3,85
<i>Alap[00018]</i>	69485	13,07	58,78	49,34	3,86
<i>Alap[00019]</i>	69485	13,07	58,78	50,13	3,87
<i>Alap[00020]</i>	69485	13,07	58,78	49,30	3,85

TABELA A.18: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL FE (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,69</b>	<b>54,65</b>	<b>9,49</b>	<b>4,98</b>
<i>Fe[00001]</i>	69485	0,69	54,65	9,84	4,99
<i>Fe[00002]</i>	69485	0,69	54,65	9,90	5,02
<i>Fe[00003]</i>	69485	0,69	54,65	8,63	4,97
<i>Fe[00004]</i>	69485	0,69	54,65	8,13	5,03
<i>Fe[00005]</i>	69485	0,69	54,65	8,49	4,99
<i>Fe[00006]</i>	69485	0,69	54,65	8,50	5,00
<i>Fe[00007]</i>	69485	0,69	54,65	8,83	4,99
<i>Fe[00008]</i>	69485	0,69	54,65	8,63	5,03
<i>Fe[00009]</i>	69485	0,69	54,65	8,36	4,97
<i>Fe[00010]</i>	69485	0,69	54,65	8,15	4,99
<i>Fe[00011]</i>	69485	0,69	54,65	9,76	5,00
<i>Fe[00012]</i>	69485	0,69	54,65	8,85	5,04
<i>Fe[00013]</i>	69485	0,69	54,65	8,22	4,97
<i>Fe[00014]</i>	69485	0,69	54,65	8,47	4,99
<i>Fe[00015]</i>	69485	0,69	54,65	8,17	4,98
<i>Fe[00016]</i>	69485	0,69	54,65	9,32	4,91
<i>Fe[00017]</i>	69485	0,69	54,65	8,42	5,00
<i>Fe[00018]</i>	69485	0,69	54,65	8,05	5,04
<i>Fe[00019]</i>	69485	0,69	54,65	8,33	5,06
<i>Fe[00020]</i>	69485	0,69	54,65	9,84	4,99

TABELA A.19: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL RESTO (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>21,47</b>	<b>52,32</b>	<b>35,76</b>	<b>2,88</b>
<i>Resto[00001]</i>	69485	21,47	52,32	36,15	2,86
<i>Resto[00002]</i>	69485	21,47	52,32	35,77	2,84
<i>Resto[00003]</i>	69485	21,47	52,32	35,85	2,87
<i>Resto[00004]</i>	69485	21,47	52,32	35,55	2,88
<i>Resto[00005]</i>	69485	21,47	52,32	35,88	2,89
<i>Resto[00006]</i>	69485	21,47	52,32	36,26	2,90
<i>Resto[00007]</i>	69485	21,47	52,32	35,46	2,91
<i>Resto[00008]</i>	69485	21,47	52,32	35,89	2,87
<i>Resto[00009]</i>	69485	21,47	52,32	35,70	2,95

<i>Resto[00010]</i>	69485	21,47	52,32	35,86	2,96
<i>Resto[00011]</i>	69485	21,47	52,32	35,03	2,91
<i>Resto[00012]</i>	69485	21,47	52,32	35,80	2,85
<i>Resto[00013]</i>	69485	21,47	52,32	35,88	2,84
<i>Resto[00014]</i>	69485	21,47	52,32	35,87	2,92
<i>Resto[00015]</i>	69485	21,47	52,32	36,15	2,93
<i>Resto[00016]</i>	69485	21,47	52,32	35,77	2,84
<i>Resto[00017]</i>	69485	21,47	52,32	35,85	2,87
<i>Resto[00018]</i>	69485	21,47	52,32	35,55	2,88
<i>Resto[00019]</i>	69485	21,47	52,32	35,88	2,99
<i>Resto[00020]</i>	69485	21,47	52,32	36,26	2,85

TABELA A.20: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL SI (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b><i>Dados originais</i></b>	<b><i>3735</i></b>	<b><i>0,27</i></b>	<b><i>14,90</i></b>	<b><i>4,44</i></b>	<b><i>1,97</i></b>
<i>Si[00001]</i>	69485	0,27	14,9	5,04	1,98
<i>Si[00002]</i>	69485	0,27	14,9	5,18	1,95
<i>Si[00003]</i>	69485	0,27	14,9	5,17	1,98
<i>Si[00004]</i>	69485	0,27	14,9	5,37	2,01
<i>Si[00005]</i>	69485	0,27	14,9	5,31	2,03
<i>Si[00006]</i>	69485	0,27	14,9	5,29	1,99
<i>Si[00007]</i>	69485	0,27	14,9	4,87	1,98
<i>Si[00008]</i>	69485	0,27	14,9	5,02	1,97
<i>Si[00009]</i>	69485	0,27	14,9	5,10	1,96
<i>Si[00010]</i>	69485	0,27	14,9	4,78	2,00
<i>Si[00011]</i>	69485	0,27	14,9	5,12	2,03
<i>Si[00012]</i>	69485	0,27	14,9	4,54	2,04
<i>Si[00013]</i>	69485	0,27	14,9	4,86	2,01
<i>Si[00014]</i>	69485	0,27	14,9	4,62	2,00
<i>Si[00015]</i>	69485	0,27	14,9	4,72	1,98
<i>Si[00016]</i>	69485	0,27	14,9	4,05	1,95
<i>Si[00017]</i>	69485	0,27	14,9	4,87	1,97
<i>Si[00018]</i>	69485	0,27	14,9	4,50	1,96
<i>Si[00019]</i>	69485	0,27	14,9	4,93	1,95
<i>Si[00020]</i>	69485	0,27	14,9	5,04	1,96

TABELA A.21: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL TI (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,27</b>	<b>3,57</b>	<b>1,53</b>	<b>0,45</b>
<i>Ti[00001]</i>	69485	0,27	3,57	1,43	0,44
<i>Ti[00002]</i>	69485	0,27	3,57	1,63	0,43
<i>Ti[00003]</i>	69485	0,27	3,57	1,62	0,46
<i>Ti[00004]</i>	69485	0,27	3,57	1,56	0,47
<i>Ti[00005]</i>	69485	0,27	3,57	1,56	0,41
<i>Ti[00006]</i>	69485	0,27	3,57	1,43	0,45
<i>Ti[00007]</i>	69485	0,27	3,57	1,48	0,46
<i>Ti[00008]</i>	69485	0,27	3,57	1,45	0,47
<i>Ti[00009]</i>	69485	0,27	3,57	1,46	0,48
<i>Ti[00010]</i>	69485	0,27	3,57	1,40	0,49
<i>Ti[00011]</i>	69485	0,27	3,57	1,42	0,45
<i>Ti[00012]</i>	69485	0,27	3,57	1,51	0,44
<i>Ti[00013]</i>	69485	0,27	3,57	1,34	0,43
<i>Ti[00014]</i>	69485	0,27	3,57	1,50	0,44
<i>Ti[00015]</i>	69485	0,27	3,57	1,43	0,45
<i>Ti[00016]</i>	69485	0,27	3,57	1,63	0,46
<i>Ti[00017]</i>	69485	0,27	3,57	1,62	0,47
<i>Ti[00018]</i>	69485	0,27	3,57	1,56	0,48
<i>Ti[00019]</i>	69485	0,27	3,57	1,56	0,44
<i>Ti[00020]</i>	69485	0,27	3,57	1,43	0,49

### 1.5. MÉDIAS GLOBAIS – COMBINAÇÃO 5

TABELA A.22: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL ALAP (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>13,07</b>	<b>58,78</b>	<b>49,17</b>	<b>3,83</b>
<i>Alap[00001]</i>	69485	13,07	58,78	49,98	3,84
<i>Alap[00002]</i>	69485	13,07	58,78	49,69	3,83
<i>Alap[00003]</i>	69485	13,07	58,78	49,02	3,81
<i>Alap[00004]</i>	69485	13,07	58,78	49,26	3,82
<i>Alap[00005]</i>	69485	13,07	58,78	49,10	3,85
<i>Alap[00006]</i>	69485	13,07	58,78	48,84	3,82
<i>Alap[00007]</i>	69485	13,07	58,78	49,10	3,81

<i>Alap[00008]</i>	69485	13,07	58,78	49,05	3,82
<i>Alap[00009]</i>	69485	13,07	58,78	49,32	3,83
<i>Alap[00010]</i>	69485	13,07	58,78	50,15	3,84
<i>Alap[00011]</i>	69485	13,07	58,78	49,89	3,86
<i>Alap[00012]</i>	69485	13,07	58,78	49,28	3,82
<i>Alap[00013]</i>	69485	13,07	58,78	49,50	3,82
<i>Alap[00014]</i>	69485	13,07	58,78	49,59	3,87
<i>Alap[00015]</i>	69485	13,07	58,78	49,97	3,86
<i>Alap[00016]</i>	69485	13,07	58,78	49,44	3,87
<i>Alap[00017]</i>	69485	13,07	58,78	49,37	3,85
<i>Alap[00018]</i>	69485	13,07	58,78	50,34	3,81
<i>Alap[00019]</i>	69485	13,07	58,78	49,34	3,85
<i>Alap[00020]</i>	69485	13,07	58,78	49,98	3,81

TABELA A.23: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL FE (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b><i>Dados originais</i></b>	<b><i>3735</i></b>	<b><i>0,69</i></b>	<b><i>54,65</i></b>	<b><i>9,49</i></b>	<b><i>4,98</i></b>
<i>Fe[00001]</i>	69485	0,69	54,65	9,08	5,00
<i>Fe[00002]</i>	69485	0,69	54,65	9,44	4,99
<i>Fe[00003]</i>	69485	0,69	54,65	9,40	4,98
<i>Fe[00004]</i>	69485	0,69	54,65	9,49	4,96
<i>Fe[00005]</i>	69485	0,69	54,65	9,44	5,02
<i>Fe[00006]</i>	69485	0,69	54,65	9,07	4,97
<i>Fe[00007]</i>	69485	0,69	54,65	9,31	5,03
<i>Fe[00008]</i>	69485	0,69	54,65	9,31	4,99
<i>Fe[00009]</i>	69485	0,69	54,65	9,11	4,98
<i>Fe[00010]</i>	69485	0,69	54,65	9,43	5,02
<i>Fe[00011]</i>	69485	0,69	54,65	9,34	5,01
<i>Fe[00012]</i>	69485	0,69	54,65	8,92	4,96
<i>Fe[00013]</i>	69485	0,69	54,65	9,35	4,97
<i>Fe[00014]</i>	69485	0,69	54,65	9,08	4,99
<i>Fe[00015]</i>	69485	0,69	54,65	9,44	5,00
<i>Fe[00016]</i>	69485	0,69	54,65	9,40	5,03
<i>Fe[00017]</i>	69485	0,69	54,65	9,49	5,04
<i>Fe[00018]</i>	69485	0,69	54,65	9,44	5,01
<i>Fe[00019]</i>	69485	0,69	54,65	9,07	5,02
<i>Fe[00020]</i>	69485	0,69	54,65	9,31	5,00

TABELA A.24: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL SI (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,27</b>	<b>14,90</b>	<b>4,44</b>	<b>1,97</b>
<i>Si[00001]</i>	69485	0,27	14,9	4,41	1,98
<i>Si[00002]</i>	69485	0,27	14,9	4,46	1,97
<i>Si[00003]</i>	69485	0,27	14,9	4,89	1,99
<i>Si[00004]</i>	69485	0,27	14,9	4,69	1,96
<i>Si[00005]</i>	69485	0,27	14,9	4,85	2,01
<i>Si[00006]</i>	69485	0,27	14,9	4,90	1,99
<i>Si[00007]</i>	69485	0,27	14,9	4,72	2,03
<i>Si[00008]</i>	69485	0,27	14,9	4,83	1,98
<i>Si[00009]</i>	69485	0,27	14,9	4,59	2,00
<i>Si[00010]</i>	69485	0,27	14,9	4,33	2,01
<i>Si[00011]</i>	69485	0,27	14,9	4,44	1,98
<i>Si[00012]</i>	69485	0,27	14,9	4,60	1,97
<i>Si[00013]</i>	69485	0,27	14,9	4,53	2,03
<i>Si[00014]</i>	69485	0,27	14,9	4,52	2,04
<i>Si[00015]</i>	69485	0,27	14,9	4,37	1,99
<i>Si[00016]</i>	69485	0,27	14,9	4,52	1,98
<i>Si[00017]</i>	69485	0,27	14,9	4,66	1,96
<i>Si[00018]</i>	69485	0,27	14,9	4,26	1,97
<i>Si[00019]</i>	69485	0,27	14,9	4,65	1,99
<i>Si[00020]</i>	69485	0,27	14,9	4,41	1,97

TABELA A.25: MÉDIA DOS DADOS ORIGINAIS PARA A VARIÁVEL TI (%) E MÉDIAS SIMULADAS PARA OS VINTE CENÁRIOS GERADOS PARA A MESMA VARIÁVEL.

<i>Variável</i>	<i>Número de dados/blocos</i>	<i>Mínimo(%)</i>	<i>Máximo(%)</i>	<i>Média(%)</i>	<i>Desvio Padrão(%)</i>
<b>Dados originais</b>	<b>3735</b>	<b>0,27</b>	<b>3,57</b>	<b>1,53</b>	<b>0,45</b>
<i>Ti[00001]</i>	69485	0,27	3,57	1,42	0,44
<i>Ti[00002]</i>	69485	0,27	3,57	1,48	0,42
<i>Ti[00003]</i>	69485	0,27	3,57	1,64	0,47
<i>Ti[00004]</i>	69485	0,27	3,57	1,55	0,45
<i>Ti[00005]</i>	69485	0,27	3,57	1,62	0,44
<i>Ti[00006]</i>	69485	0,27	3,57	1,62	0,46
<i>Ti[00007]</i>	69485	0,27	3,57	1,58	0,43
<i>Ti[00008]</i>	69485	0,27	3,57	1,60	0,44
<i>Ti[00009]</i>	69485	0,27	3,57	1,55	0,45

<i>Ti[00010]</i>	69485	0,27	3,57	1,45	0,46
<i>Ti[00011]</i>	69485	0,27	3,57	1,42	0,47
<i>Ti[00012]</i>	69485	0,27	3,57	1,52	0,44
<i>Ti[00013]</i>	69485	0,27	3,57	1,46	0,45
<i>Ti[00014]</i>	69485	0,27	3,57	1,49	0,46
<i>Ti[00015]</i>	69485	0,27	3,57	1,43	0,45
<i>Ti[00016]</i>	69485	0,27	3,57	1,48	0,44
<i>Ti[00017]</i>	69485	0,27	3,57	1,53	0,48
<i>Ti[00018]</i>	69485	0,27	3,57	1,37	0,47
<i>Ti[00019]</i>	69485	0,27	3,57	1,51	0,46
<i>Ti[00020]</i>	69485	0,27	3,57	1,42	0,44

## 2. MÉDIAS LOCAIS

As figuras a seguir (A.1 a A.24) apresentam as médias locais ao longo dos eixos *x*, *y* e *z*, obtidas para os vinte cenários simulados em cada combinação testada. É possível perceber que as médias locais simuladas exibem comportamento semelhante ao das médias locais dos dados originais.

### 2.1. MÉDIAS LOCAIS – COMBINAÇÃO 1

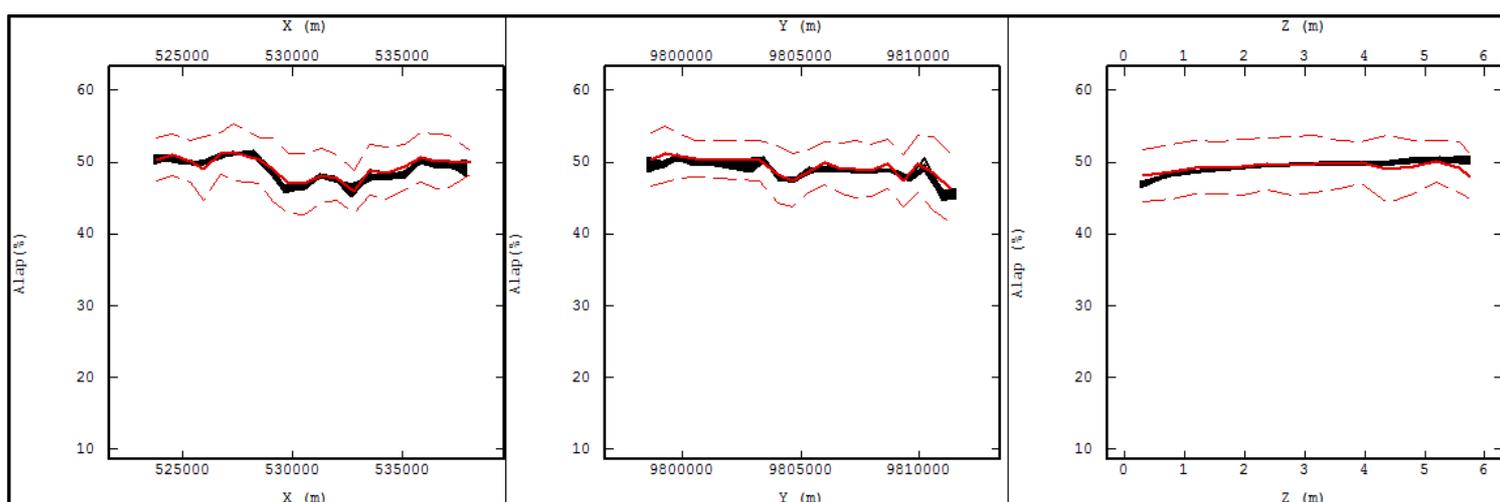


FIGURA A.0.1: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL ALAP (%) NA COMBINAÇÃO 1. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

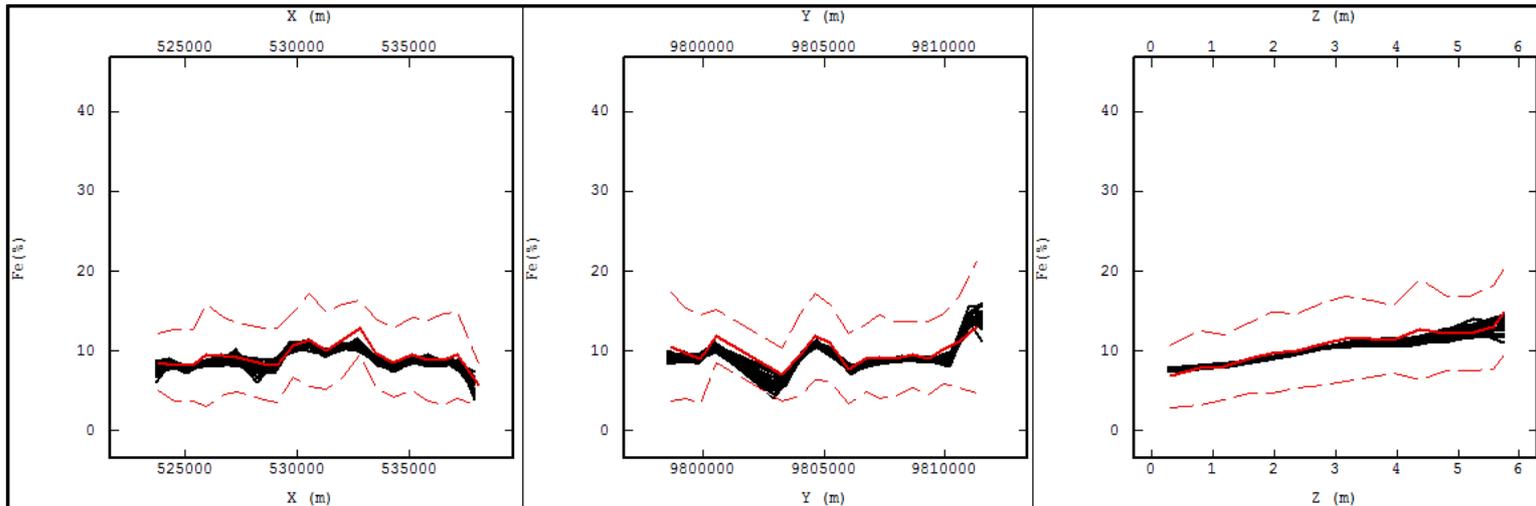


FIGURA A.0.2: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL FE (%) NA COMBINAÇÃO 1. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

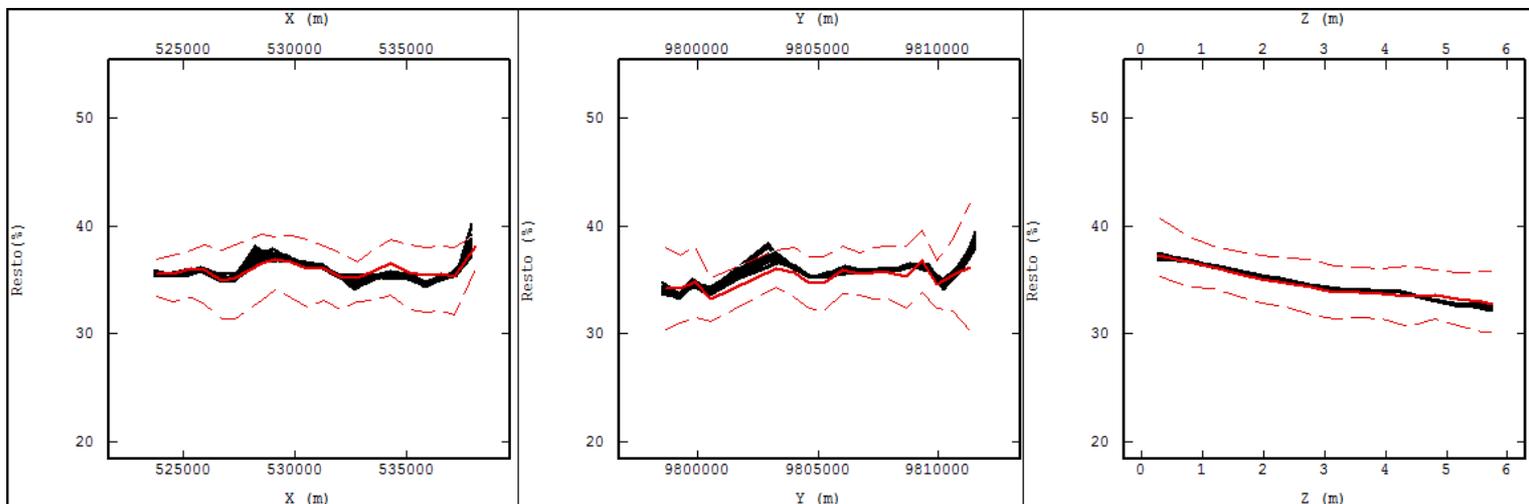


FIGURA A.0.3: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL RESTO (%) NA COMBINAÇÃO 1. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

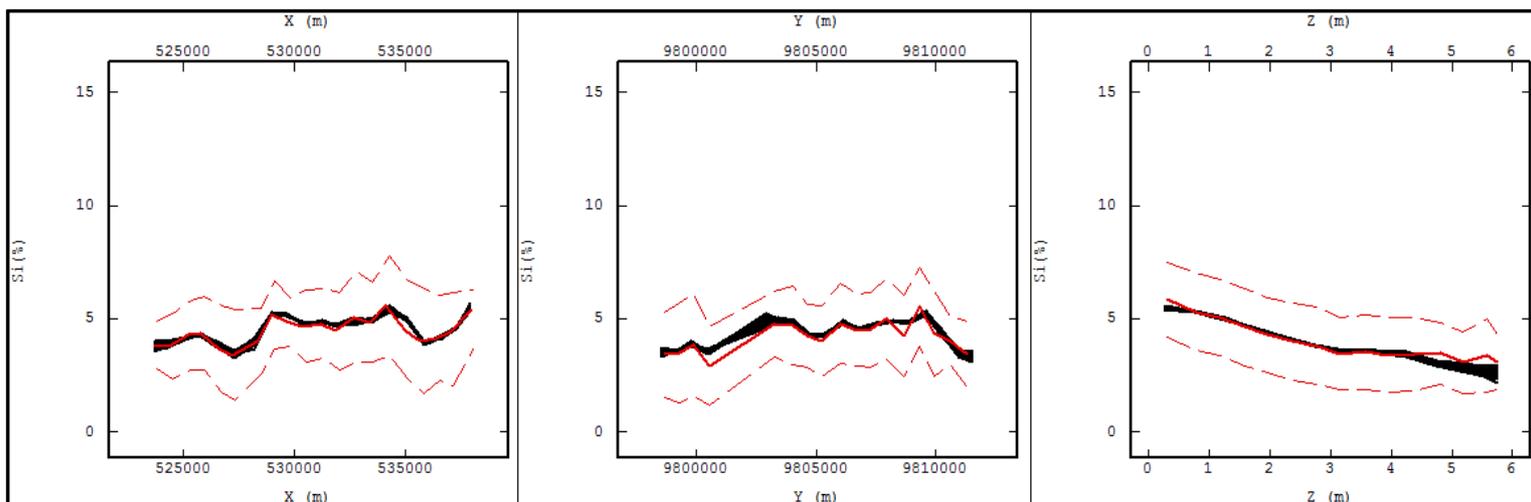


FIGURA A.0.4: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL SI (%) NA COMBINAÇÃO 1. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

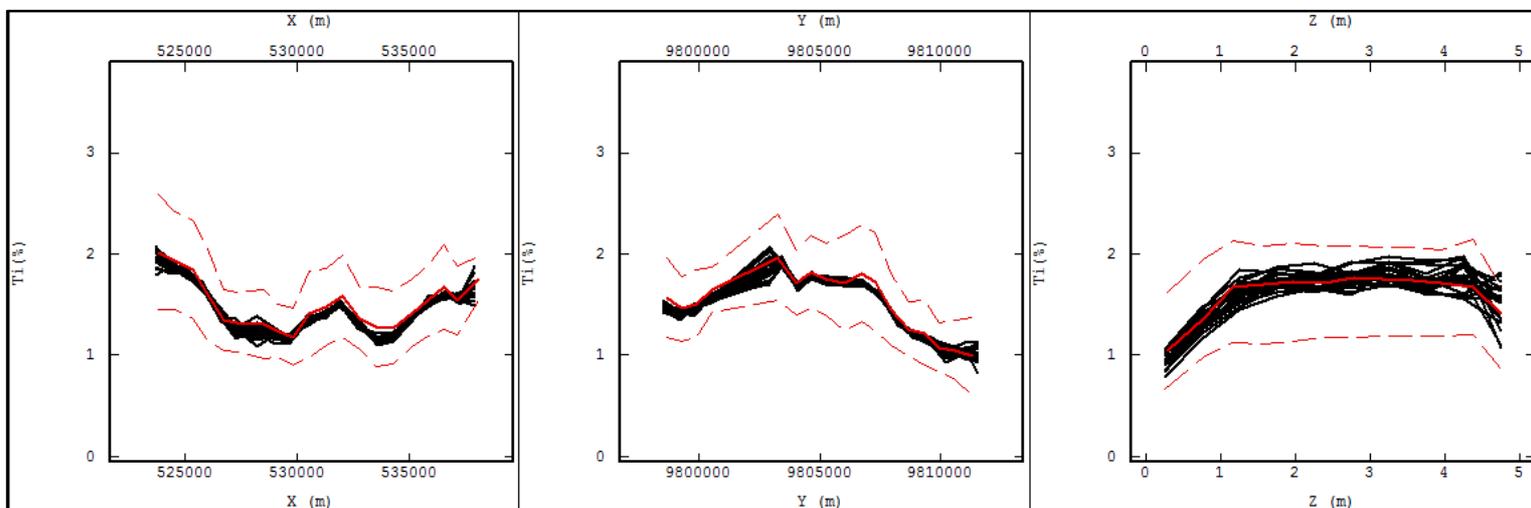


FIGURA A.0.5: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL TI (%) NA COMBINAÇÃO 1. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

## 2.2. MÉDIAS LOCAIS – COMBINAÇÃO 2

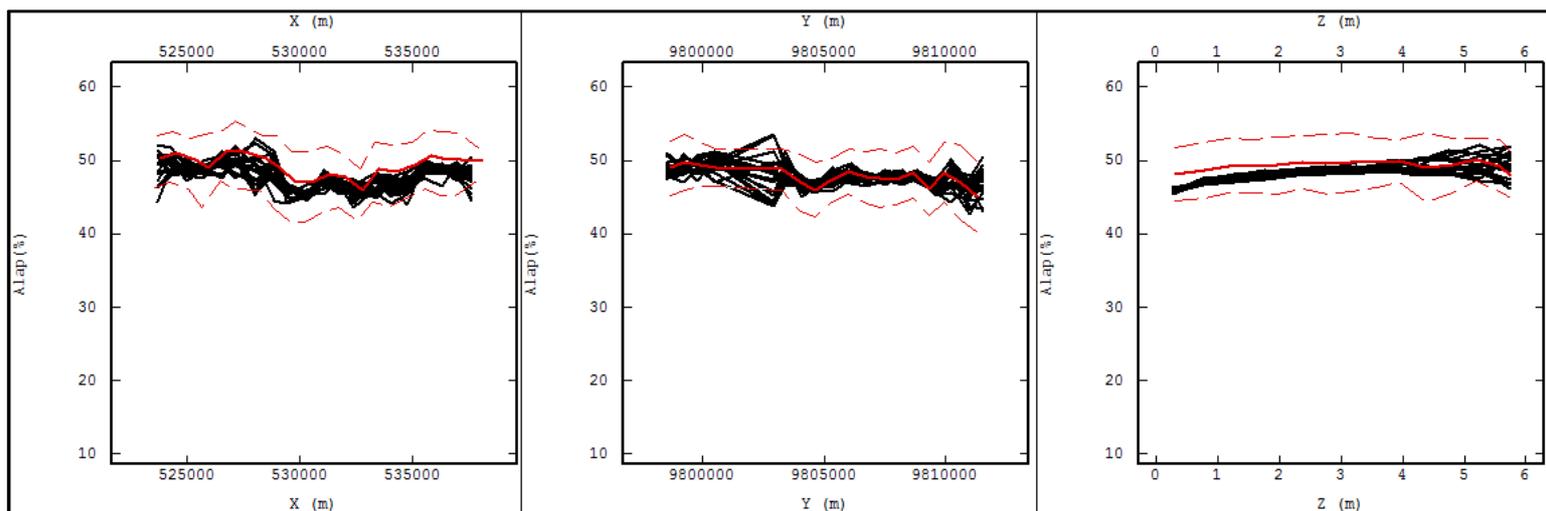


FIGURA A.0.6: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL ALAP (%) NA COMBINAÇÃO 2. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

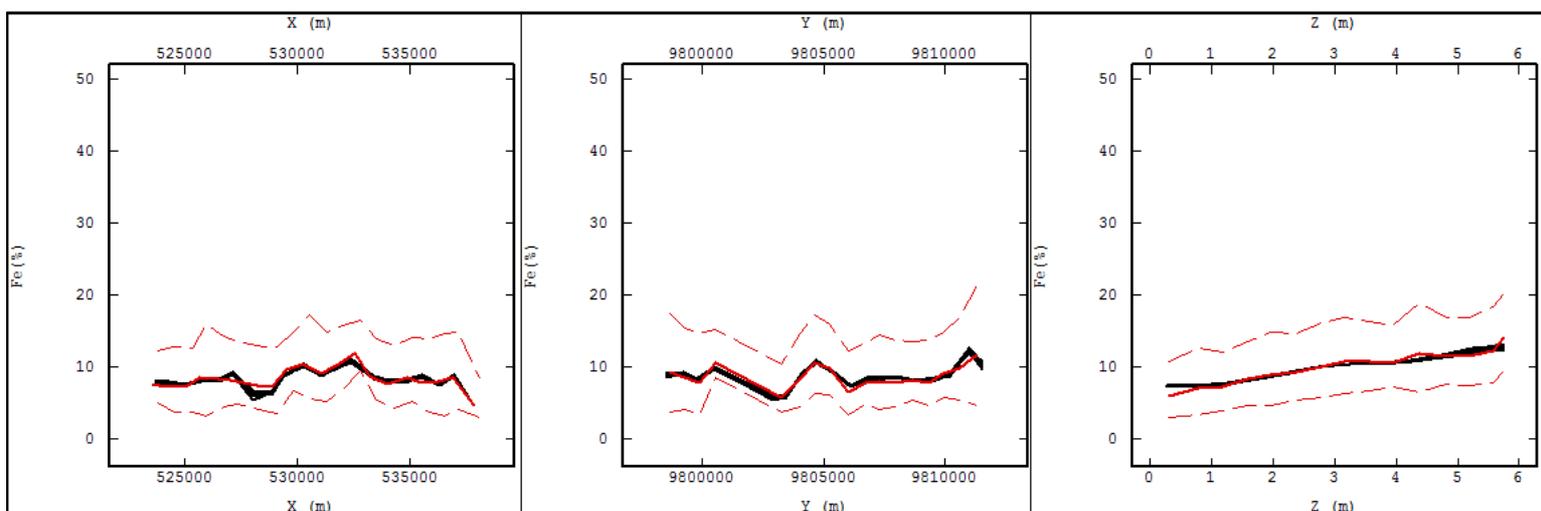


FIGURA A.0.7: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL FE (%) NA COMBINAÇÃO 2. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E

AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

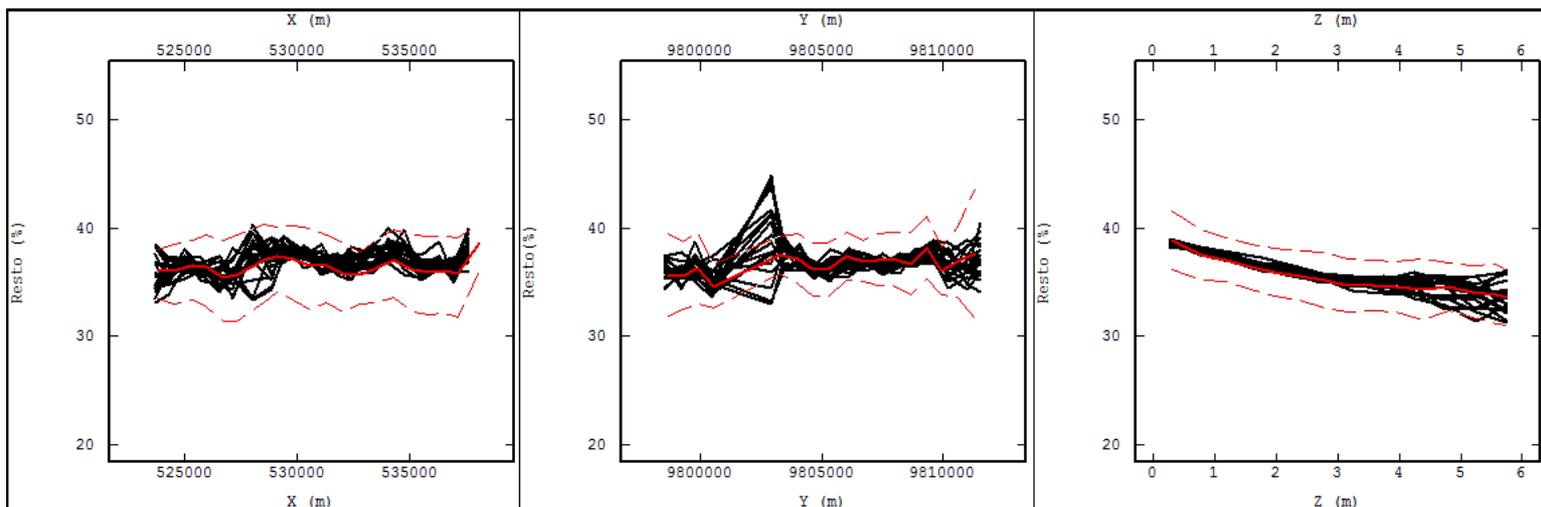


FIGURA A.0.8: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL RESTO (%) NA COMBINAÇÃO 2. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

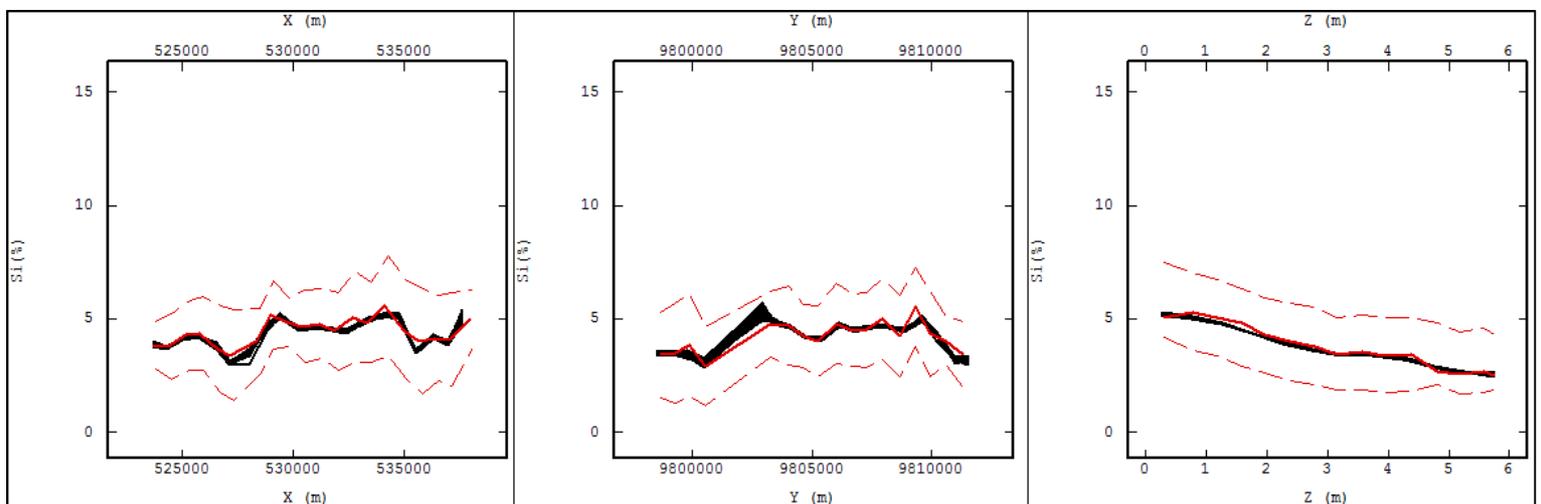


FIGURA A.0.9: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL SI (%) NA COMBINAÇÃO 2. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

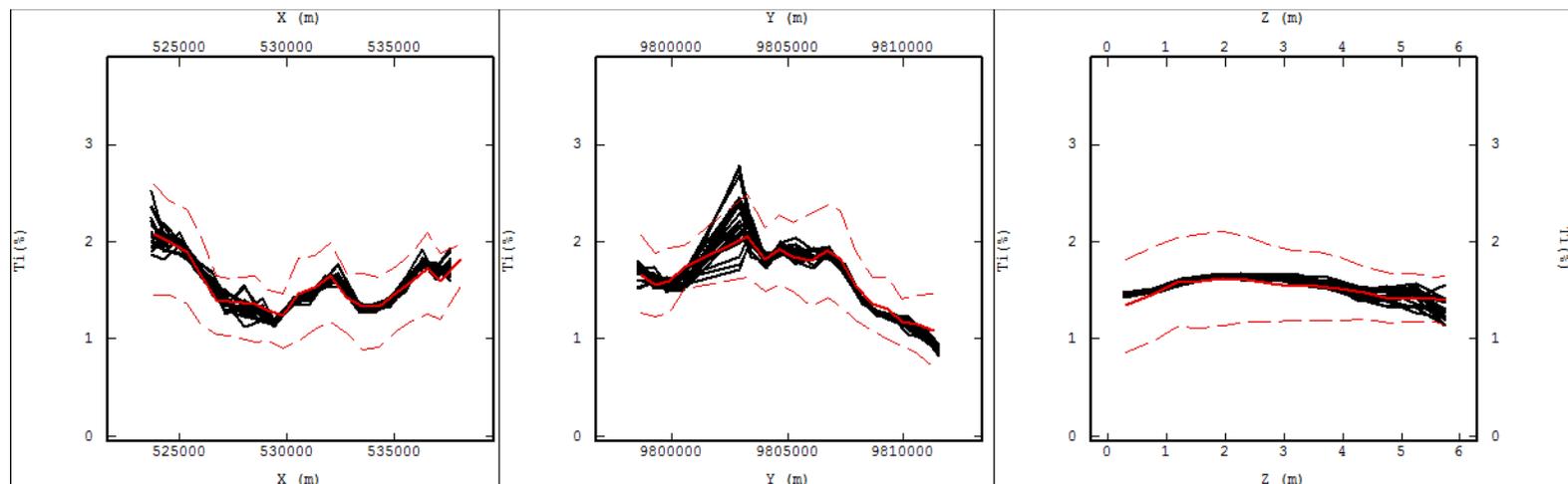


FIGURA A.0.10: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL TI (%) NA COMBINAÇÃO 2. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

### 2.3. MÉDIAS LOCAIS – COMBINAÇÃO 3

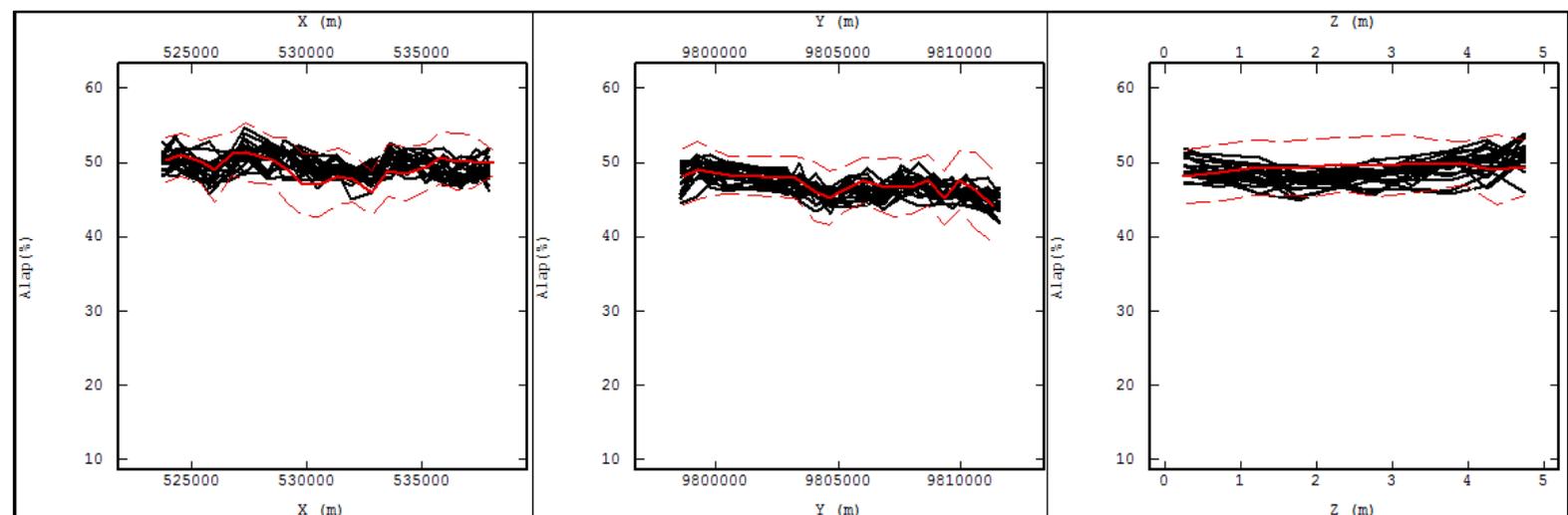


FIGURA A.0.11: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL ALAP (%) NA COMBINAÇÃO 3. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

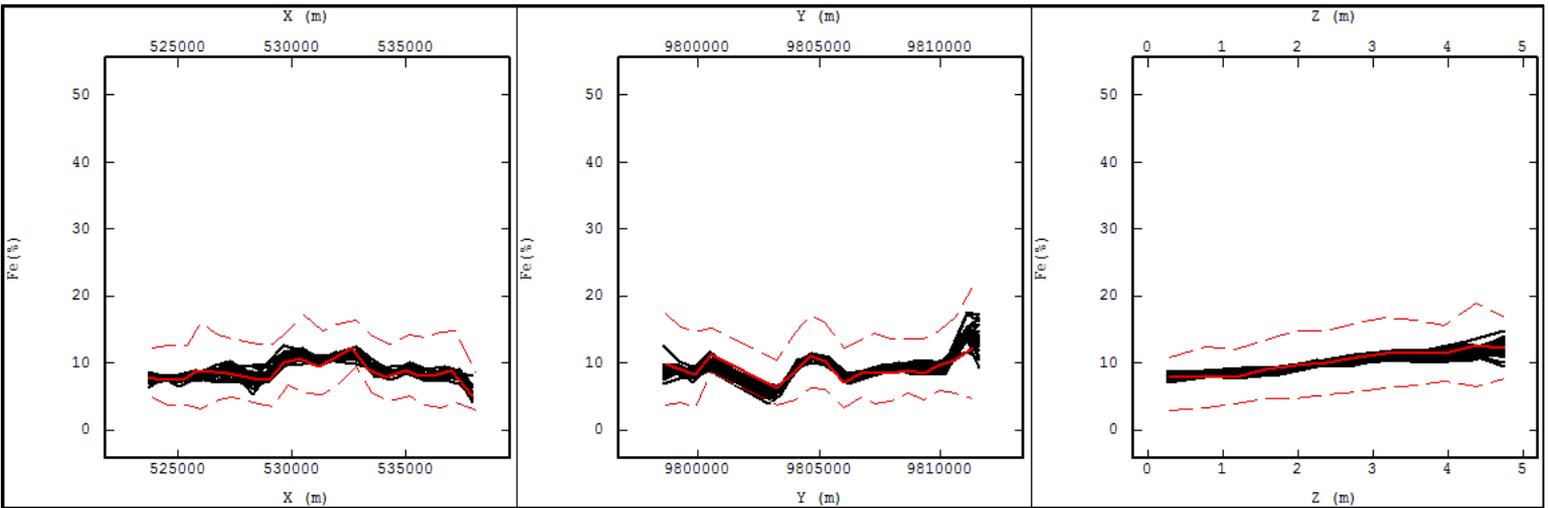


FIGURA A.0.12: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL FE (%) NA COMBINAÇÃO 3. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

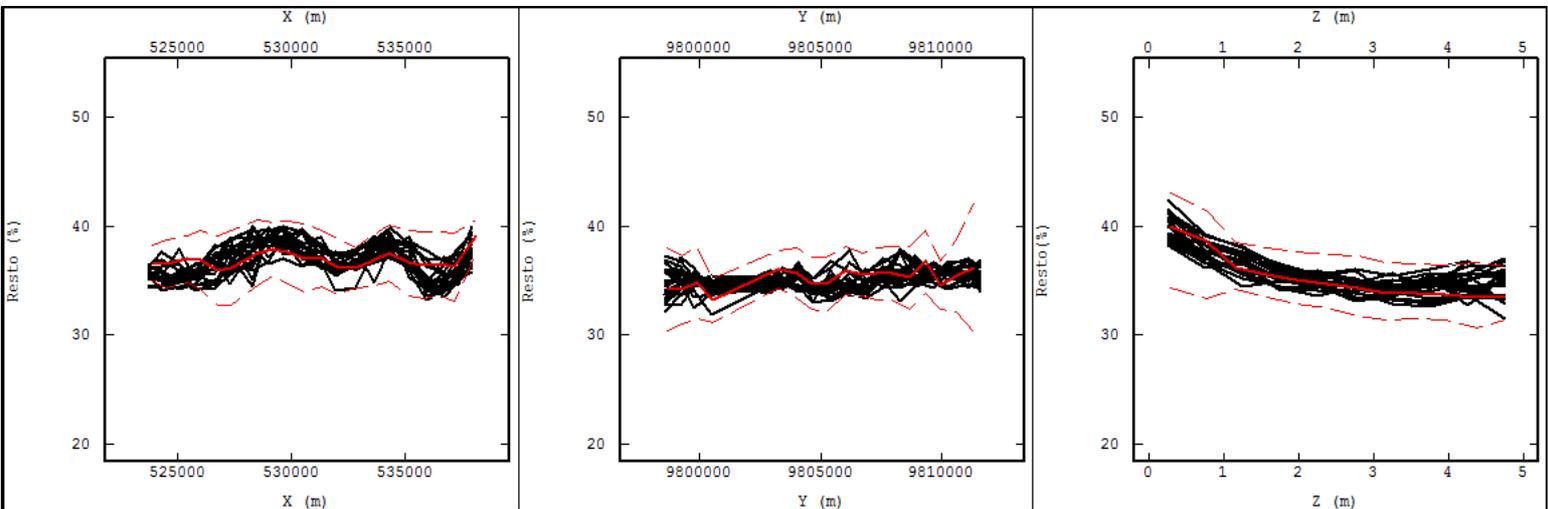


FIGURA A.0.13: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL RESTO (%) NA COMBINAÇÃO 3. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

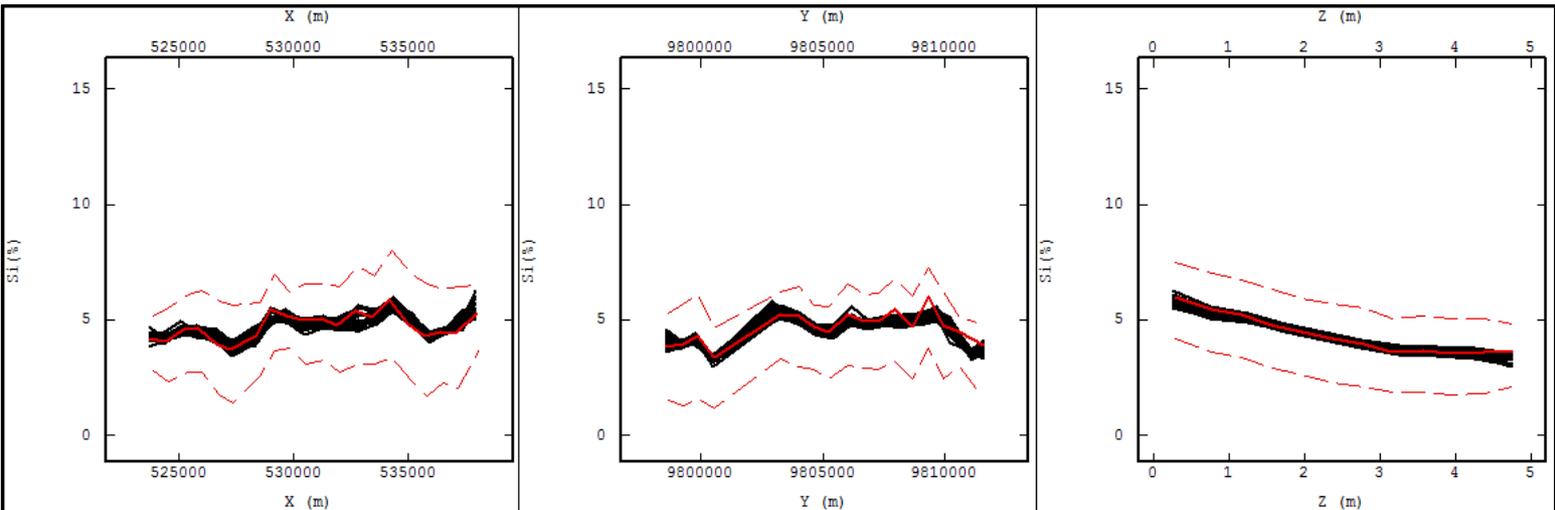


FIGURA A.0.14: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL SI (%) NA COMBINAÇÃO 3. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

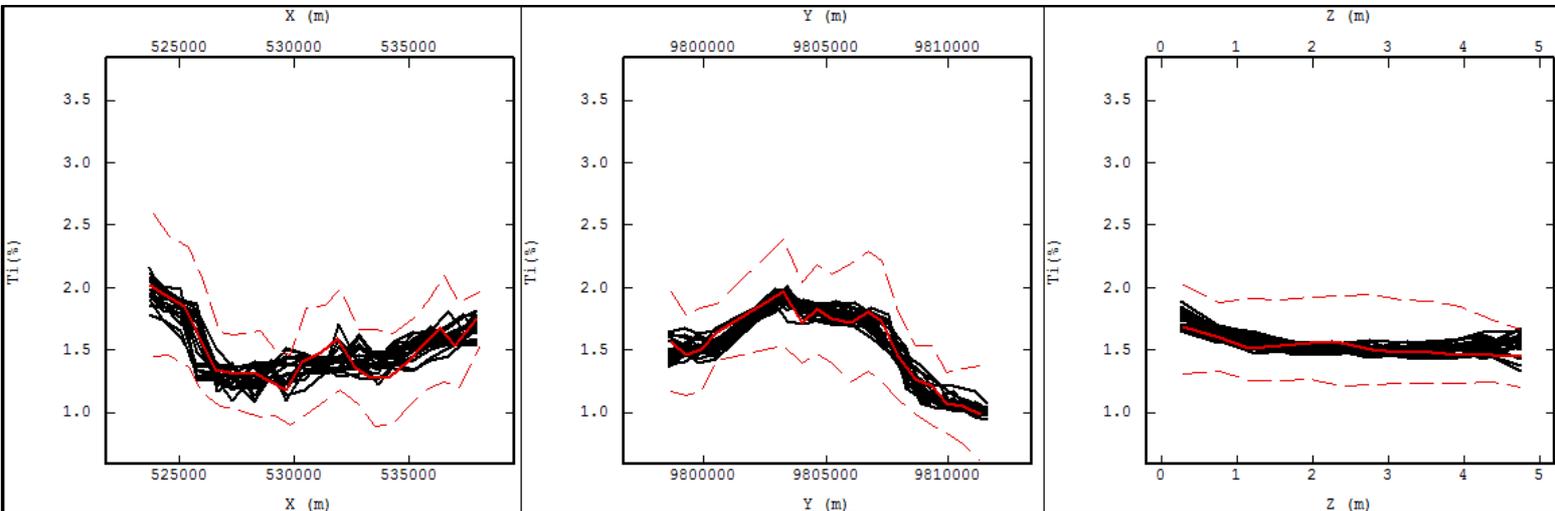


FIGURA A.0.15: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL TI (%) NA COMBINAÇÃO 3. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

## 2.4. MÉDIAS LOCAIS – COMBINAÇÃO 4

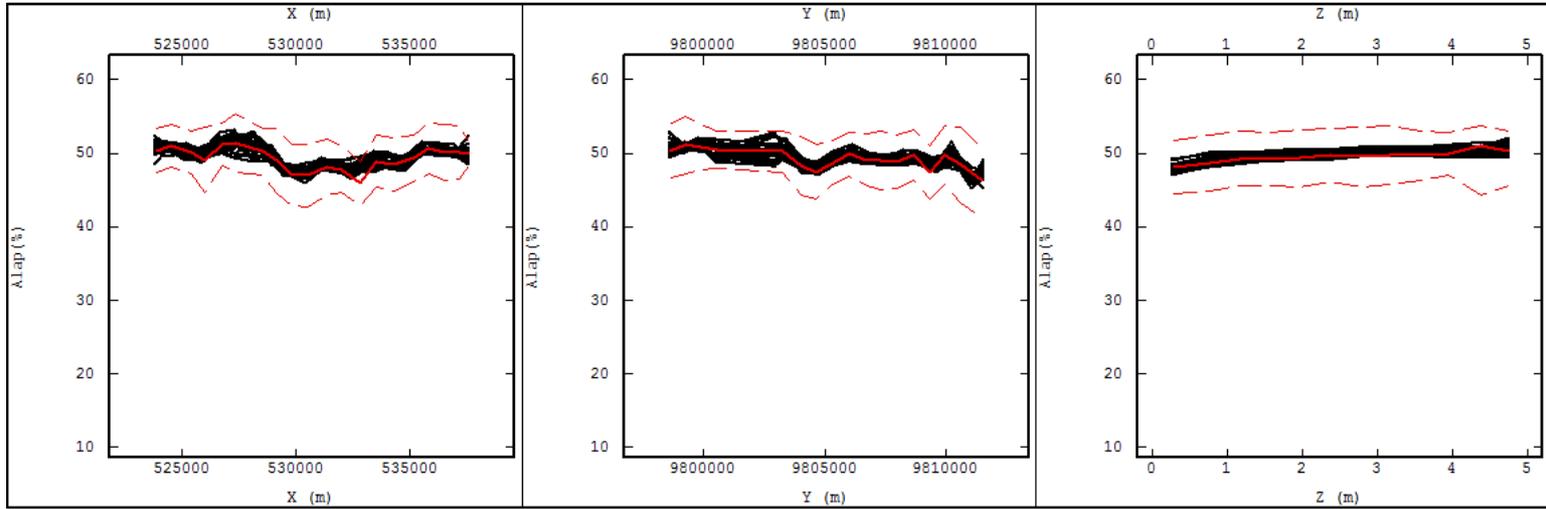


FIGURA A.0.16: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL ALAP (%) NA COMBINAÇÃO 4. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

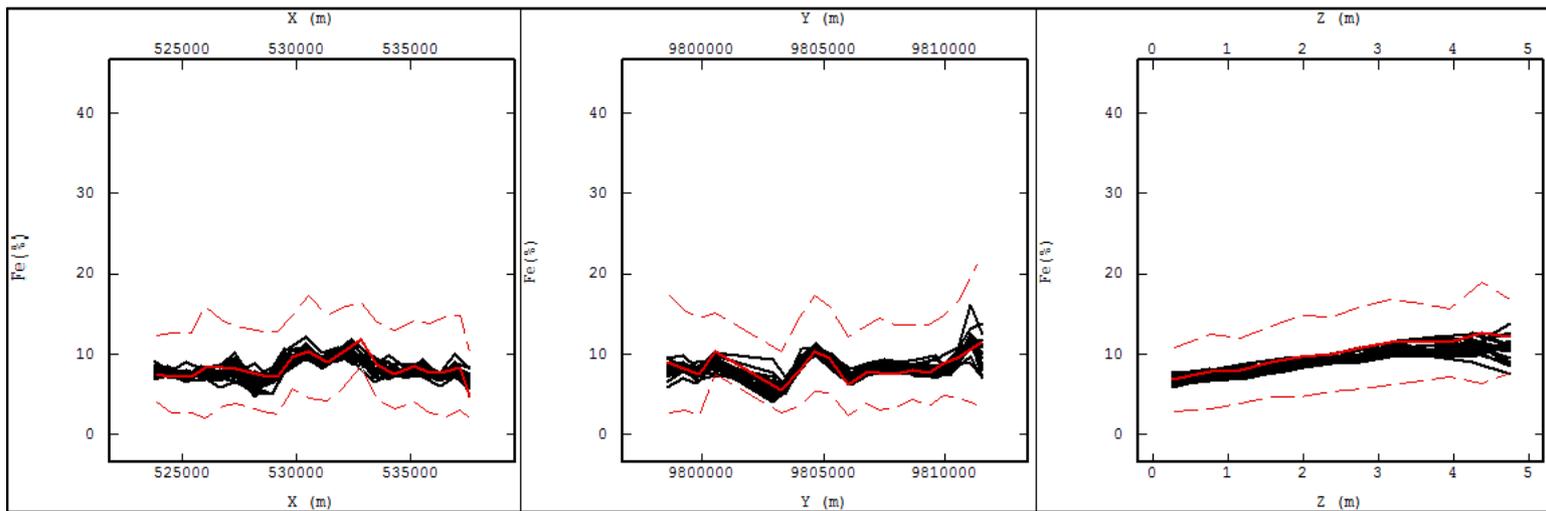


FIGURA A.0.17: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL FE (%) NA COMBINAÇÃO 4. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

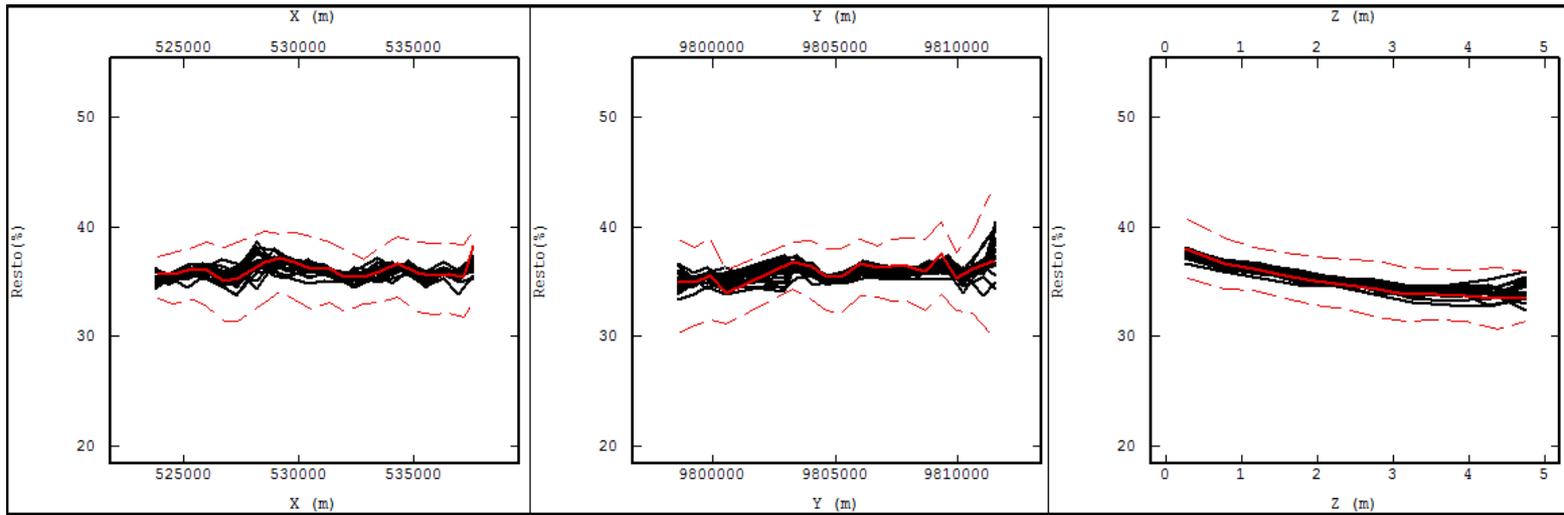


FIGURA A.0.18: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL RESTO (%) NA COMBINAÇÃO 4. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

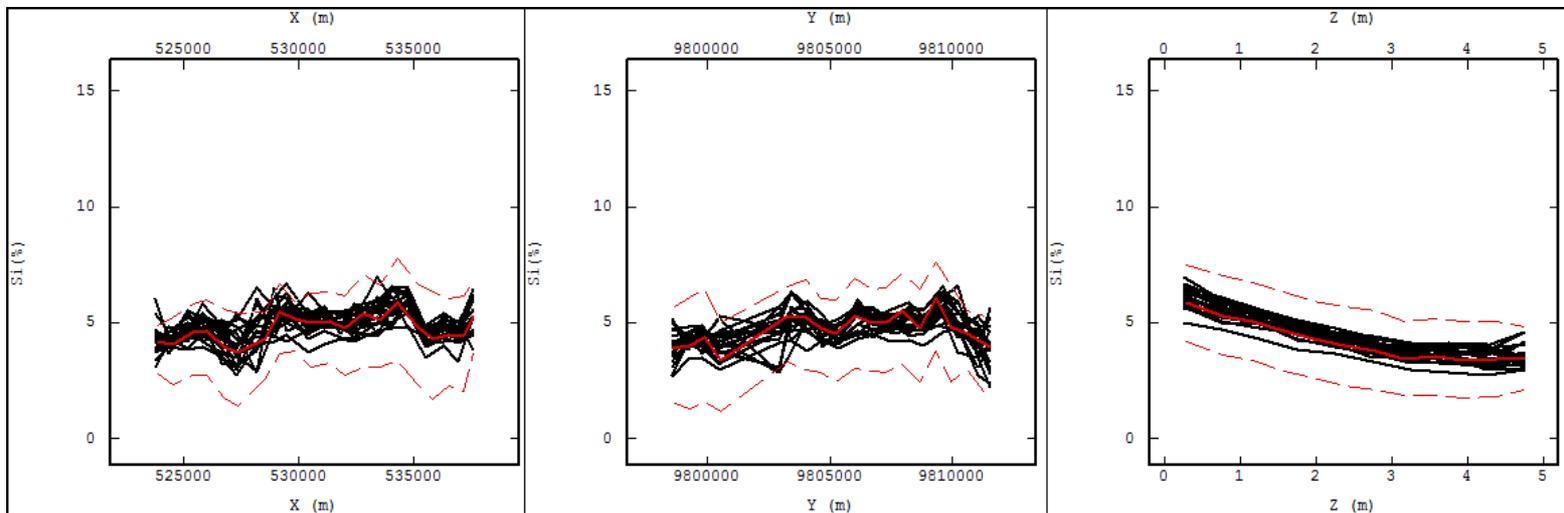


FIGURA A.0.19: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL SI (%) NA COMBINAÇÃO 4. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

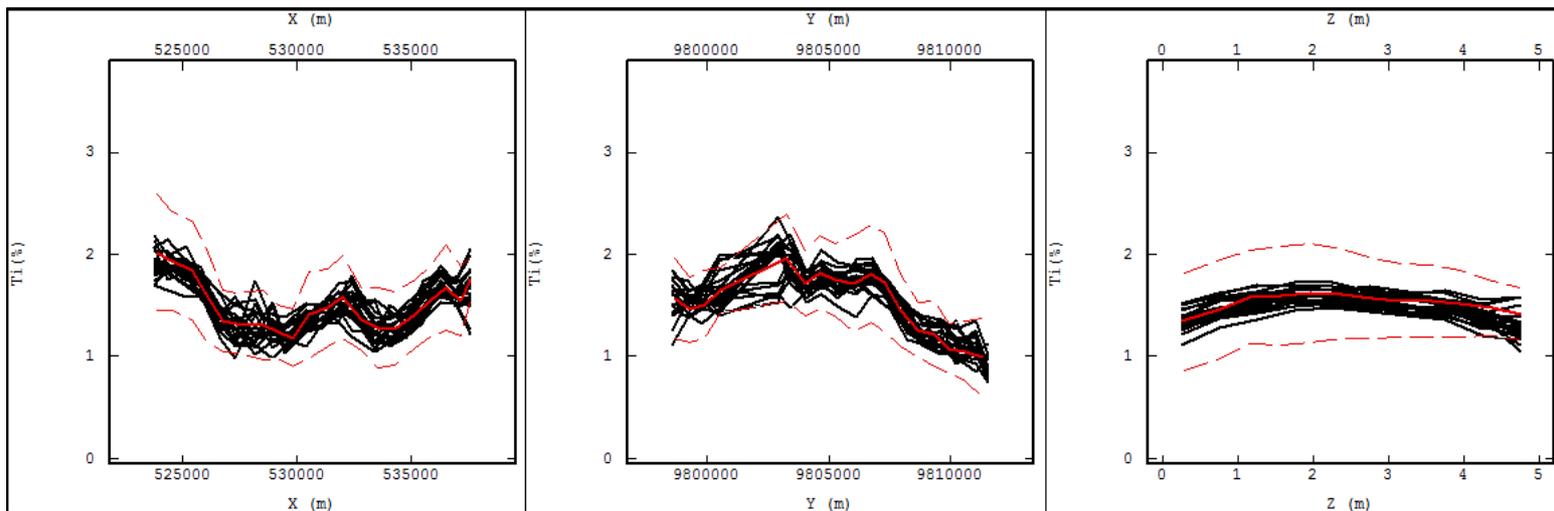


FIGURA A.0.20: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL TI (%) NA COMBINAÇÃO 4. A LINHA VERMELHA ESPESSA REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

## 2.5. MÉDIAS LOCAIS – COMBINAÇÃO 5

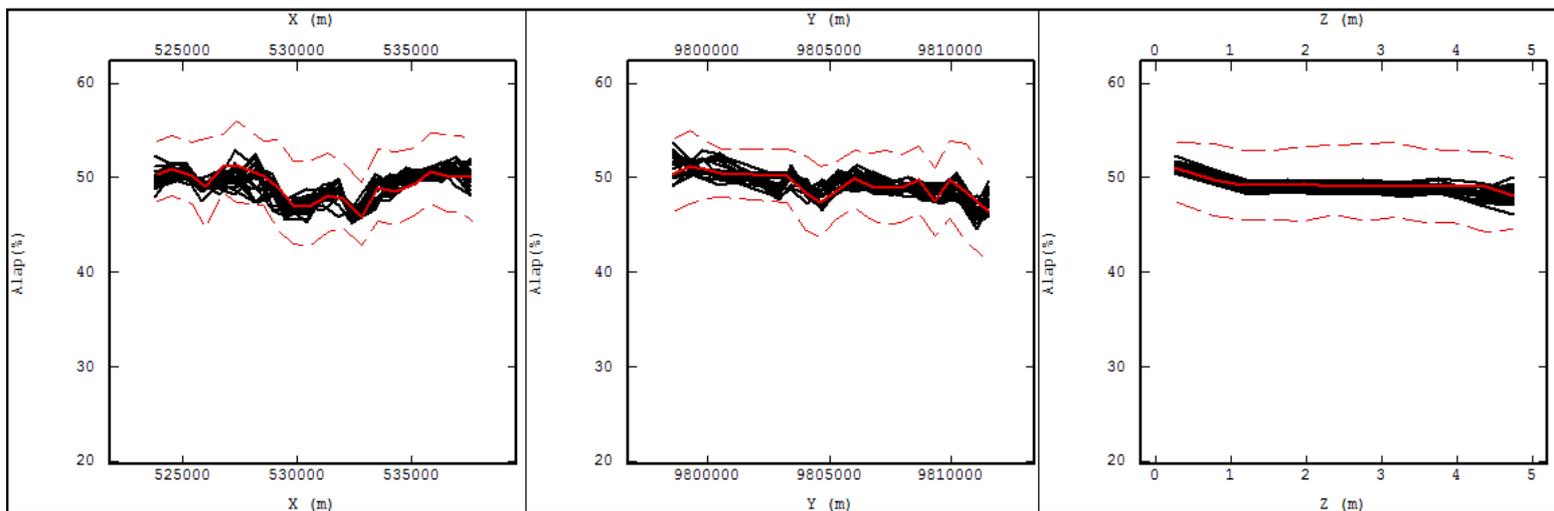


FIGURA A.0.21: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL ALAP (%) NA COMBINAÇÃO 5. A LINHA VERMELHA ESPESSA REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

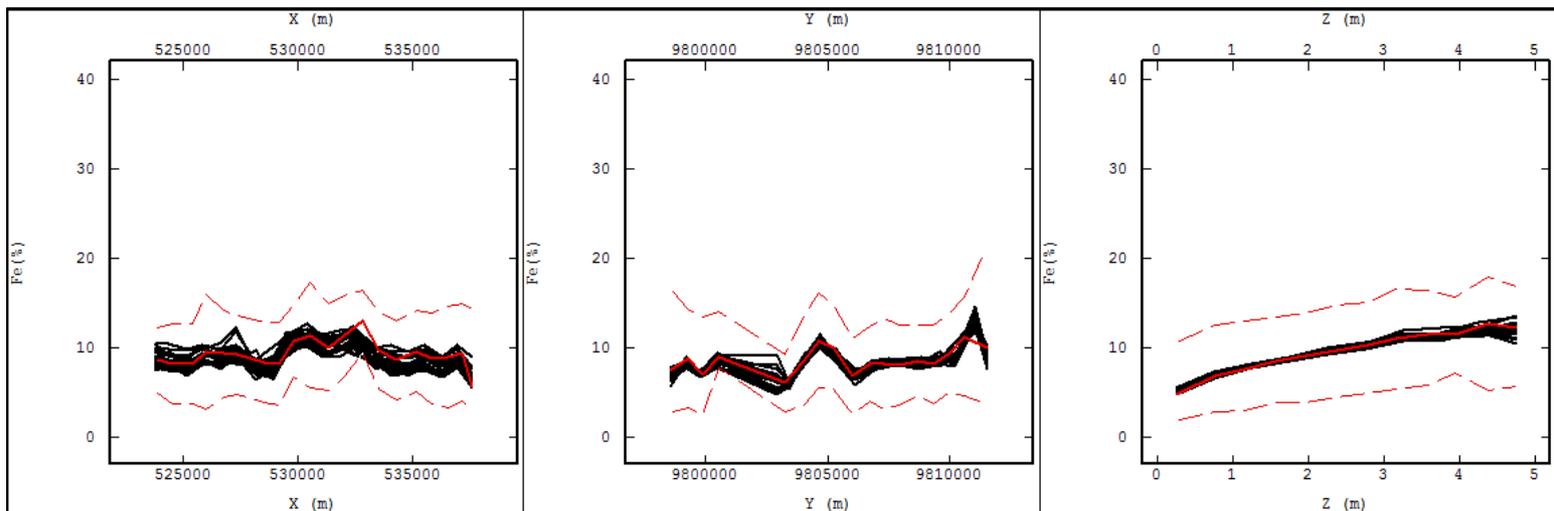


FIGURA A.0.22: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL FE (%) NA COMBINAÇÃO 5. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

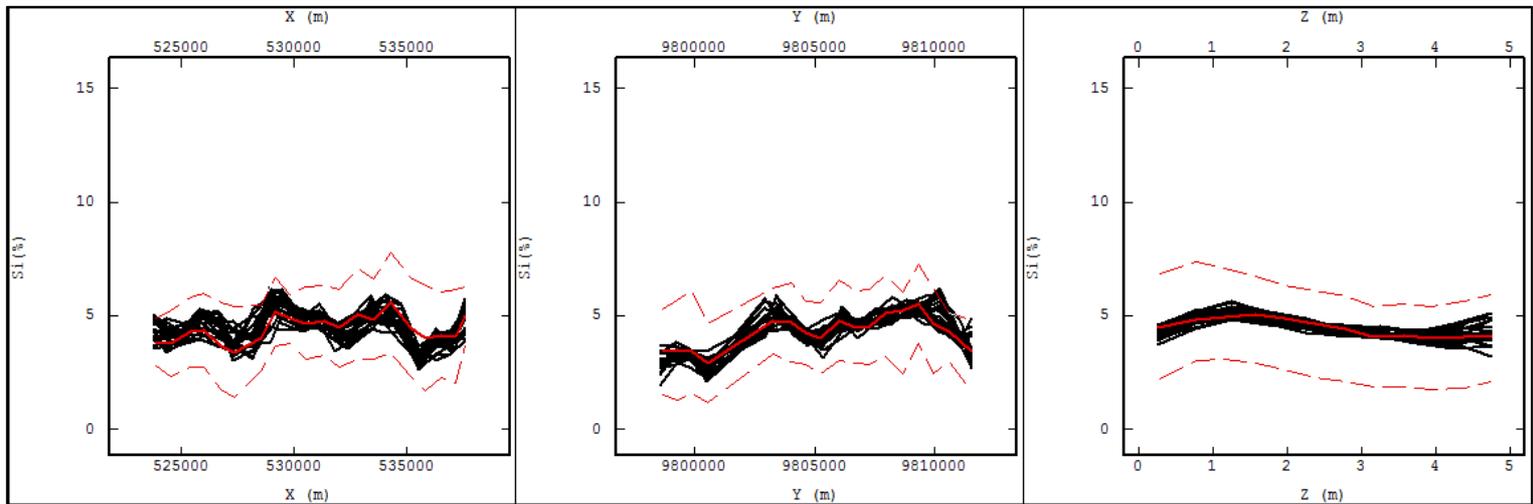


FIGURA A.0.23: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL SI (%) NA COMBINAÇÃO 5. A LINHA VERMELHA ESPESSE REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

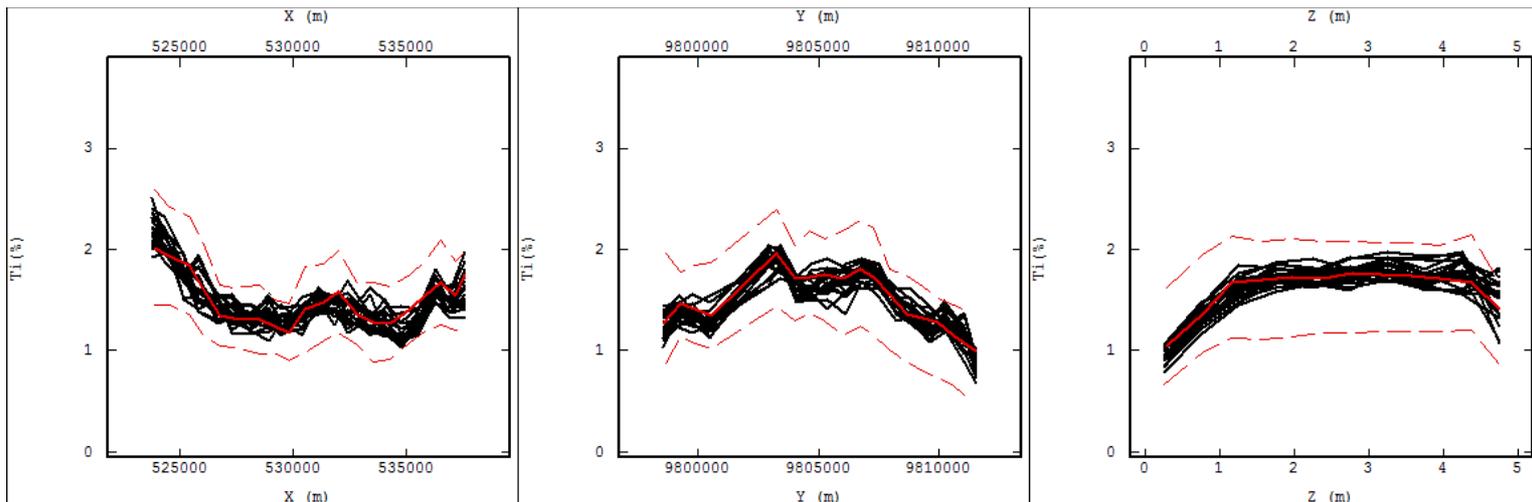


FIGURA A.0.24: MÉDIAS LOCAIS AO LONGO DOS EIXOS X, Y E Z, PARA A VARIÁVEL TI (%) NA COMBINAÇÃO 5. A LINHA VERMELHA ESPESSA REPRESENTA AS MÉDIAS LOCAIS DOS DADOS ORIGINAIS, AS LINHAS VERMELHAS TRACEJADAS REPRESENTAM O DESVIO PADRÃO DAS MÉDIAS LOCAIS DOS DADOS ORIGINAIS E AS LINHAS PRETAS REPRESENTAM AS MÉDIAS LOCAIS DOS VINTE CENÁRIOS SIMULADOS NA REFERIDA COMBINAÇÃO TESTADA.

### 3. CONTINUIDADE ESPACIAL

As figuras A.25 a A.28 exibem a adequada reprodução dos correlogramas obtidos nos vinte cenários simulados quando comparados com os correlogramas dos dados originais, para as variáveis de interesse em cada combinação analisada.

#### 3.1. CORRELOGRAMAS PARA A COMBINAÇÃO 1

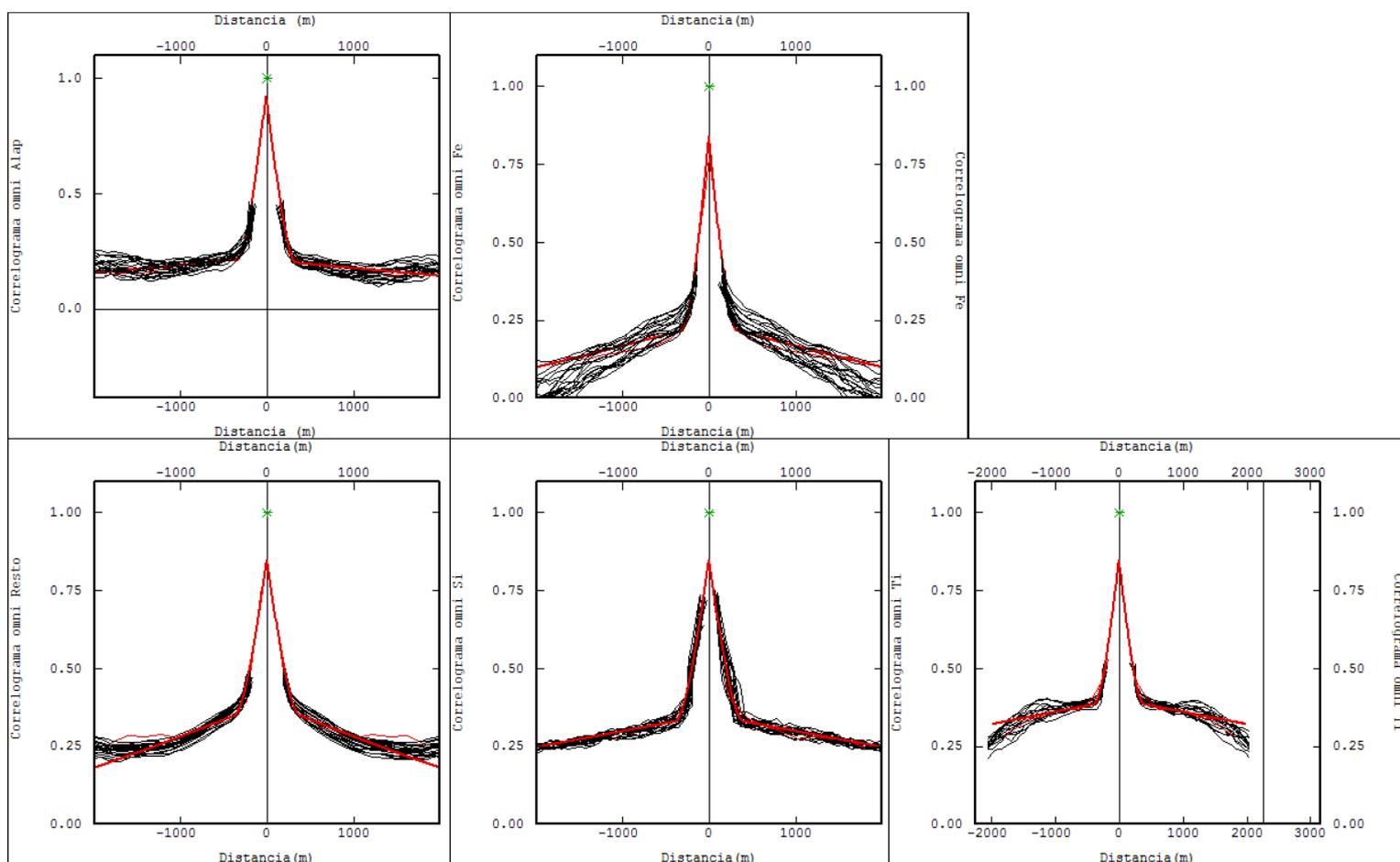


FIGURA A.0.25: CORRELOGRAMAS EXPERIMENTAIS (LINHA VERMELHA FINA), CORRELOGRAMAS MODELADOS (LINHA VERMELHA ESPESSA) E CORRELOGRAMAS OBTIDOS NAS SIMULAÇÕES (LINHAS PRETAS) PARA AS VARIÁVEIS ALAP, FE, RESTO, SI E TI.

### 3.2. CORRELOGRAMAS PARA A COMBINAÇÃO 2

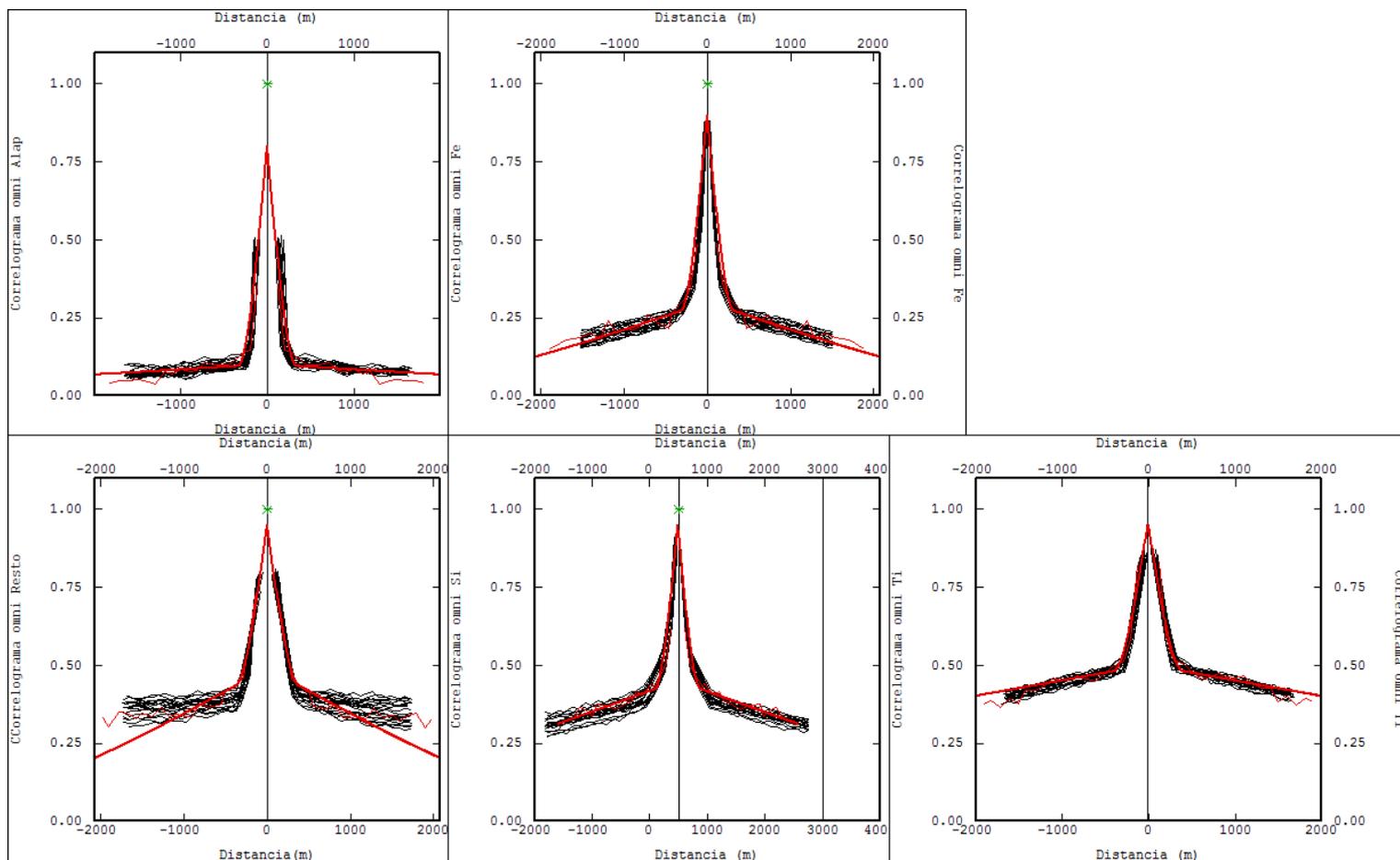


FIGURA A.0.26: CORRELOGRAMAS EXPERIMENTAIS (LINHA VERMELHA FINA), CORRELOGRAMAS MODELADOS (LINHA VERMELHA ESPESSE) E CORRELOGRAMAS OBTIDOS NAS SIMULAÇÕES (LINHAS PRETAS) PARA AS VARIÁVEIS ALAP, FE, RESTO, SI E TI.

### 3.3. CORRELOGRAMAS PARA A COMBINAÇÃO 3

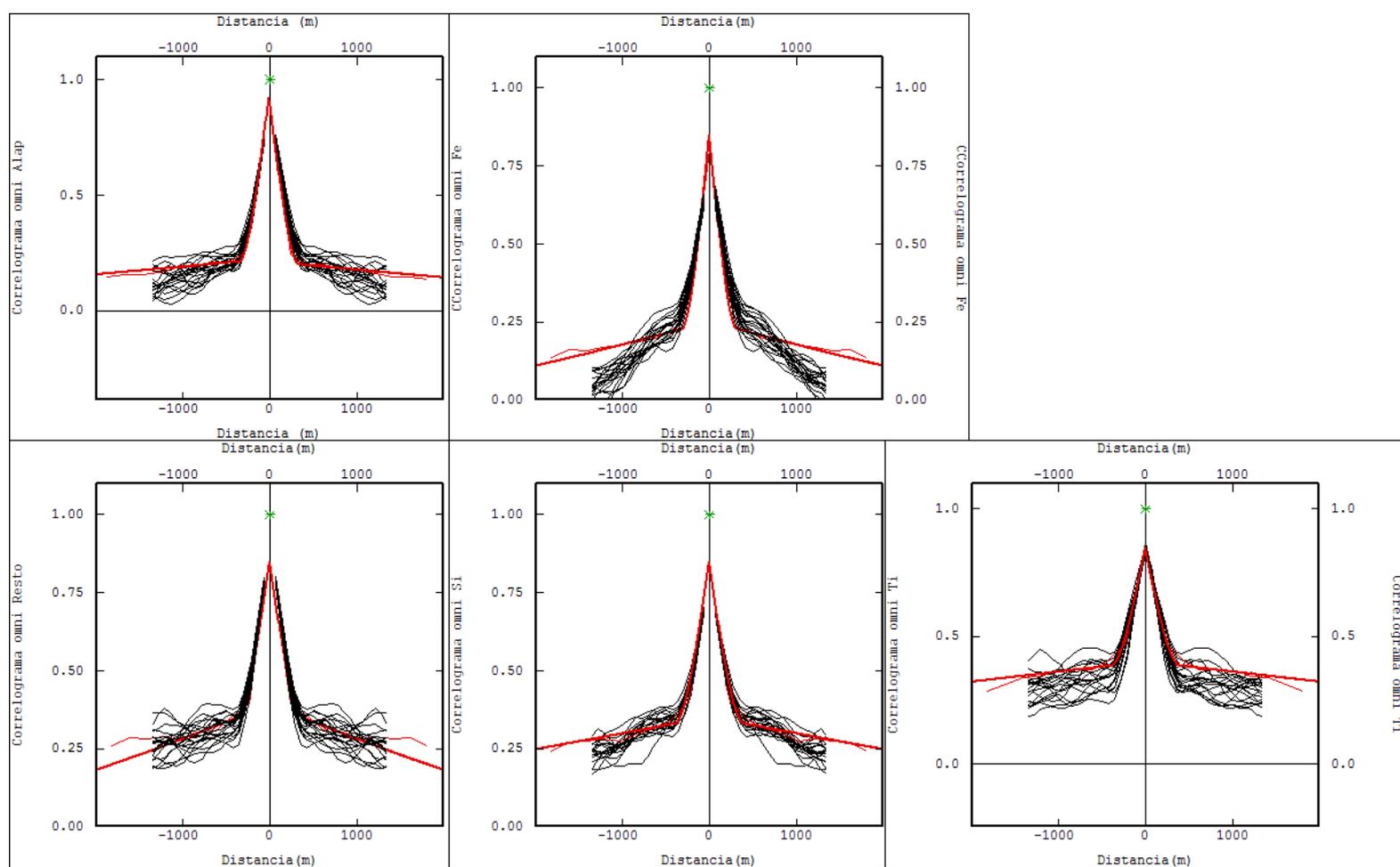


FIGURA A.0.27: CORRELOGRAMAS EXPERIMENTAIS (LINHA VERMELHA FINA), CORRELOGRAMAS MODELADOS (LINHA VERMELHA ESPESSE) E CORRELOGRAMAS OBTIDOS NAS SIMULAÇÕES (LINHAS PRETAS) PARA AS VARIÁVEIS ALAP, FE, RESTO, SI E TI.

### 3.4. CORRELOGRAMAS PARA A COMBINAÇÃO 4

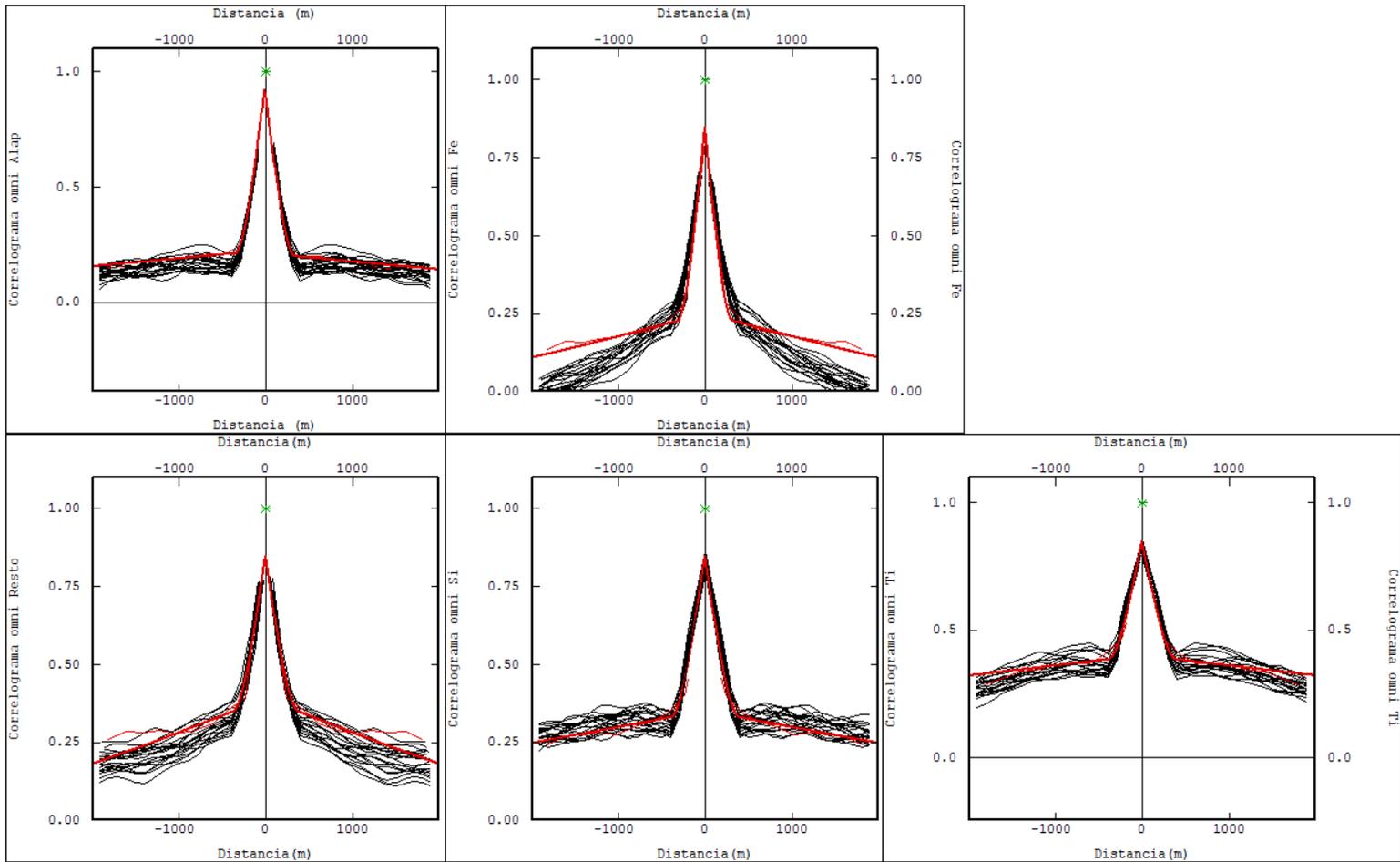


FIGURA A.0.28: CORRELOGRAMAS EXPERIMENTAIS (LINHA VERMELHA FINA), CORRELOGRAMAS MODELADOS (LINHA VERMELHA ESPESSE) E CORRELOGRAMAS OBTIDOS NAS SIMULAÇÕES (LINHAS PRETAS) PARA AS VARIÁVEIS ALAP, FE, RESTO, SI E TI.

### 3.5. CORRELOGRAMAS PARA A COMBINAÇÃO 5

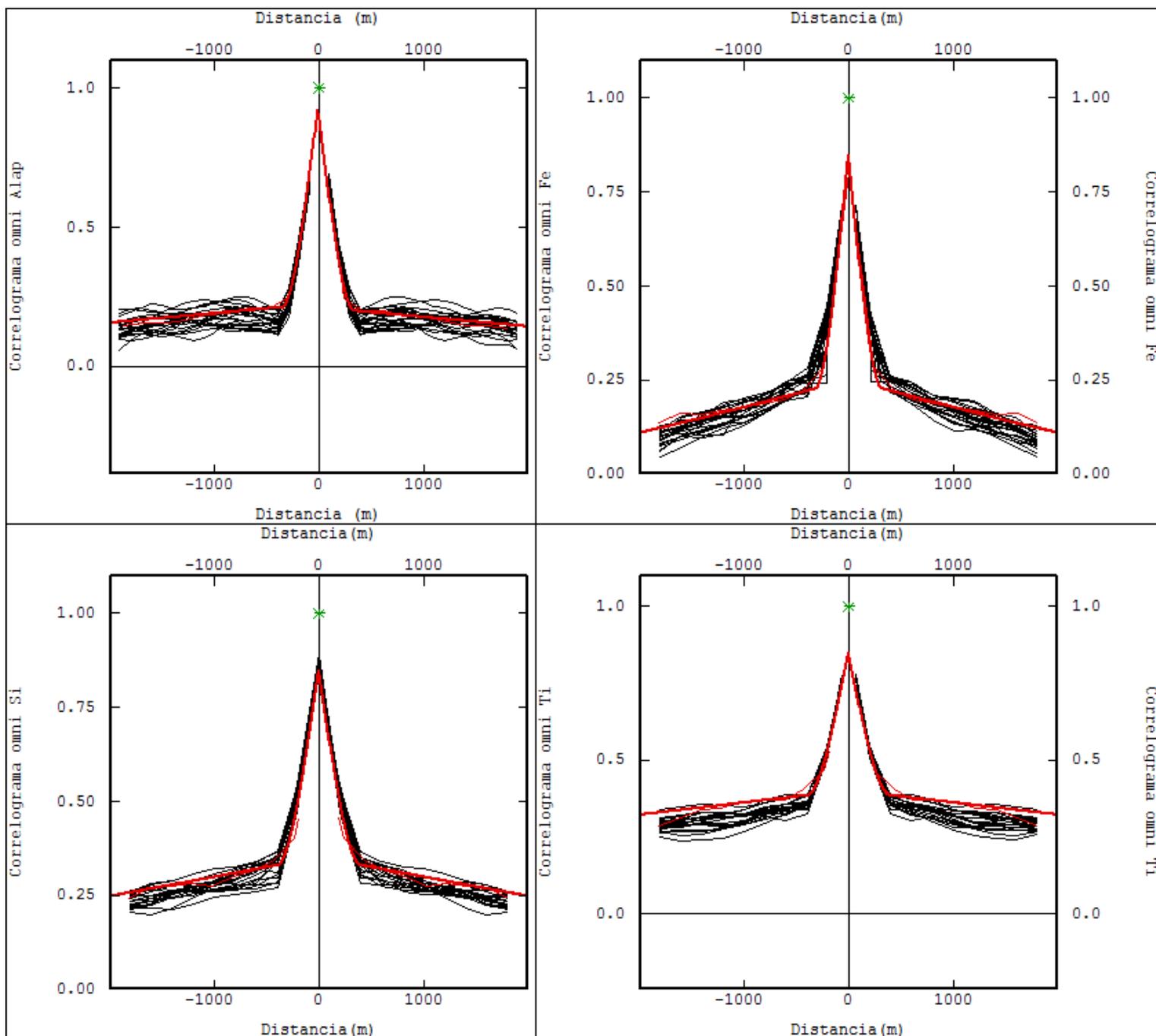


FIGURA A.0.29: CORRELOGRAMAS EXPERIMENTAIS (LINHA VERMELHA FINA), CORRELOGRAMAS MODELADOS (LINHA VERMELHA ESPESSE) E CORRELOGRAMAS OBTIDOS NAS SIMULAÇÕES (LINHAS PRETAS) PARA AS VARIÁVEIS ALAP, FE, SI E TI.

