UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

Estudo comparativo de métodos geoestatísticos de
estimativas e simulações estocásticas condicionais

RAFAEL DE AGUIAR FURUIE

Orientador: Prof. Dr. Jorge Kazuo Yamamoto

Dissertação de Mestrado

Programa de Pós-Graduação em Recursos Minerais e Hidrogeologia

São Paulo

2009

À minha família

Aproveito este espaço para agradecer a todos aqueles que participaram da construção deste trabalho.

Em primeiro lugar agradeço ao meu orientador Jorge Kazuo Yamamoto que desde o início sempre me apoiou no desenvolvimento do projeto com toda a confiança e paciência necessárias para tal. Também agradeço aos professores, colegas e funcionários do IGc-USP que me ajudaram nessa jornada, seja por seus comentários construtivos, pela ajuda em resolver problemas ou simplesmente, e não menos importante, pela sua amizade. Em especial cito os colegas do LIG e ds sala 105 pela ajuda prestada nesta etapa final de trabalho, a citar: Antônio, Sidney e Fabricio.

Sou grato também à ajuda prestada pelo colega Gustavo Saka, pelo auxílio na utilização da Simulação por Bandas Rotativas e por me fornecer a sua versão do algoritmo.

Também gostaria de demonstrar meu agradecimento a todos os colegas e chefes que tive ao longo de minha vida profissional, desde os tempos da Yamana Gold, passando pela CPRM – Serviço Geológico do Brasil, até atualmente na Petrobras, que foram compreensivos com o esforço e tempo por mim despendidos nesta dissertação de mestrado e sempre tentaram me ajudar ao máximo na conquista dos meus objetivos. Cabe citar aqui especialmente meus superiores imediatos, pela sua compreensão: Mauricio Assis e Marcelo Juliano de Carvalho (Yamana Gold) e Luiz Antônio Chieregati (CPRM). Também a todos os colegas que conheci no “trecho”, por todo o companheirismo e amizade, além do apoio e incentivo na construção deste trabalho, lembrando aqui apenas alguns: André e Luiz Fernando (Yamana); Fabrizio, Márcio, Luiz Gustavo, Frederico e Sérgio (CPRM).

Agradeço também a todos os amigos que me incentivaram a sempre aprimorar meus conhecimentos, assim como pelos palpites e interesse demonstrados, cabendo citar aqui: João Eduardo, Tobias, Lucas e Iupanque.

Finalmente agradeço muito a meus pais Raquel e Sérgio, assim como meu irmão Vinicius pelo ambiente construtivo que sempre tive em casa, com todo o estímulo ao livre pensar e ao esforço pessoal, assim como ao amor que sempre me demonstraram, não esquecendo a sua grande ajuda prestada na edição final deste trabalho. Também agradeço à Ana.

Resumo

FURUIE, R. A. Estudo comparativo de métodos geoestatísticos de estimativas e simulações estocásticas condicionais. 2009. 166 f. Dissertação (Mestrado) – Instituto de Geociências, Universidade de São Paulo, São Paulo, 2009.

Diferentes métodos geoestatísticos são apresentados como a melhor solução para diferentes contextos de acordo com a natureza dos dados a serem analisados. Alguns dos métodos de estimativa mais populares incluem a krigagem ordinária e a krigagem ordinária lognormal, esta última requerendo a transformação dos dados originais para uma distribuição gaussiana. No entanto, esses métodos apresentam limitações, sendo uma das mais discutidas o efeito de suavização apresentado pelas estimativas obtidas. Alguns algoritmos recentes foram propostos como meios de se corrigir este efeito, e são avaliados neste trabalho para a sua eficiência, assim como alguns algoritmos para a transformação reversa dos valores convertidos na krigagem ordinária lognormal. Outra abordagem para o problema é por meio do grupo de métodos denominado de simulação estocástica, alguns dos mais populares sendo a simulação gaussiana sequencial e a simulação por bandas rotativas, que apesar de não apresentar o efeito de suavização da krigagem, não possuem a precisão local característica dos métodos de estimativa. Este trabalho busca avaliar a eficiência dos diferentes métodos de estimativa (krigagem ordinária, krigagem ordinária lognormal, assim como suas estimativas corrigidas) e simulação (simulação sequencial gaussiana e simulação por bandas rotativas) para diferentes cenários de dados. Vinte e sete conjuntos de dados exaustivos (em grid 50x50) foram amostrados em 90 pontos por meio da amostragem aleatória simples. Estes conjuntos de dados partiam de uma distribuição gaussiana (Log1) e tinham seus coeficientes de variação progressivamente aumentados até se chegar a uma distribuição altamente assimétrica (Log27). Semivariogramas amostrais foram computados e modelados para os processos geoestatísticos de estimativa e simulação. As estimativas ou realizações resultantes foram então comparadas com os dados exaustivos originais de maneira a se avaliar quão bem esses dados originais eram reproduzidos. Isto foi feito pela comparação de parâmetros estatísticos dos dados originais com os dos dados reconstruídos, assim como por meio de análise gráfica. Resultados demonstraram que o método que apresentou melhores resultados foi a krigagem ordinária lognormal, estes ainda melhores quando aplicada a transformação reversa de Yamamoto, com grande melhora principalmente nos resultados para os dados altamente assimétricos. A krigagem ordinária apresentou sérias limitações na reprodução da cauda inferior dos conjuntos de dados mais assimétricos, apresentando para estes resultados piores que as estimativas não corrigidas. Ambos os métodos de simulação utilizados apresentaram uma baixa correlação como os dados exaustivos, seus resultados também cada vez menos representativos de acordo com o aumento do coeficiente de variação, apesar de apresentar a vantagem de fornecer diferentes cenários para tomada de decisões.

Palavras-chave: krigagem ordinária, krigagem ordinária lognormal, simulação estocástica, efeito de suavização, simulação gaussiana sequencial, simulação por bandas rotativas.

Abstract

FURUIE, R. A. Estudo comparativo de métodos geoestatísticos de estimativas e simulações estocásticas condicionais. 2009. 166 p. Dissertação (Mestrado) – Instituto de Geociências, Universidade de São Paulo, São Paulo, 2009.

Different geostatistical methods present themselves as the optimal solution to different realities according to the characteristics displayed by the data in analysis. Some of the most popular estimation methods include ordinary kriging and lognormal ordinary kriging, this last one involving the transformation of data from their original space to a Gaussian distribution. However, these methods present some limitations, one of the most prominent ones being the smoothing effect observed in the resulting estimates. Some recent algorithms have been proposed as a way to correct this effect, and are tested in this work for their effectiveness, as well as some methods for the backtransformation of the lognormal converted values. Another approach to the problem is by means of the group of methods known as stochastic simulation, some of the most popular ones being the sequential Gaussian simulation and turning bands simulation, which although do not present the smoothing effect, lack the local accuracy characteristic of the estimation methods. This work seeks to assess the effectiveness of the different estimation (ordinary kriging, lognormal ordinary kriging, and their corrected estimates) and simulation (sequential Gaussian simulation and turning bands simulation) methods for different scenarios. Twenty seven exhaustive data sets (in a 50x50 grid) have been sampled at 90 points based on simple random sampling. These data sets started from a Gaussian distribution (Log1) and had their variation coefficients increased progressively, up to a highly asymmetrical distribution (Log27). Experimental semivariograms have been computed and modeled for geostatistical estimation and simulation processes. The resulting estimates or realizations were then compared to the original exhaustive data in order to assess how well these reproduced the original data. This was done by comparing statistical parameters of the original data and the ones of the reconstructed data, as well as graphically. Results showed that the method that presented the best correlation with the exhaustive data was lognormal ordinary kriging, even better when the backtransformation technique by Yamamoto is applied, which much improved the results for the more asymmetrical data sets. Ordinary kriging and its correction had some severe limitations in reproducing the lower tail of the more asymmetrical data sets, with worst results than those for the uncorrected estimates. Both simulation methods used presented a very small degree of correlation to the exhaustive data, their results also progressively less representative as the variation coefficient grew, even though it has the advantage of presenting several scenarios for decision making.

Keywords: ordinary kriging, lognormal ordinary kriging, stochastic simulation, smoothing effect, sequential gaussian simulation, turning bands simulation.

Lista de Figuras

Figura 1.1 Comparação entre histogramas para os dados amostrais (esquerda) e estimativas da krigagem ordinária (direita).....	3
Figura 1.2. Variogramas experimentais dos dados amostrais (círculos vazios) e dos valores estimados pela krigagem ordinária (círculos cheios) e modelo de variograma amostral (linha grossa contínua).....	3
Figura 2.1. Representação gráfica do enviesamento condicional existente na krigagem. O gráfico representa a distorção existente entre os valores reais e os estimados após a realização da Validação Cruzada, sendo possível observar que os valores reais maiores são subestimados, e os valores reais menores são superestimados. Tal distorção leva à obtenção de uma reta de regressão que se afasta da 1ª bissetriz (caso ideal). (Adaptado de Olea, 1999).....	8
Figura 2.2. Conjuntos de dados distintos, porém com mesma variância da krigagem (adaptado de Armstrong, 1994).....	10
Figura 2.3. Representação Gráfica da Transformação Gaussiana (adaptado de Olea, 1991).....	23
Figura 2.4. Semivariogramas experimentais (pontos) e semivariogramas induzidos por modelo Gaussiano (linhas contínuas) para os nove decis da distribuição acumulada (CDF).....	25
Figura 2.5. Bandas rotativas (adaptado de Journel & Huijbregts, 1978).	37
Figura 2.6 Bandas rotativas associadas à linha D ₁ . Os pontos da grade bidimensional são projetados nesta linha e assumem o valor da simulação unidimensional apropriada para a sua respectiva banda (adaptado de Brooker, 1985).	38
Figura 3.1. Mapa de localização dos 90 pontos de dados.....	42
Figura 4.1. Comparação da variação dos coeficientes de correlação para dados estimados pela KO e pela KO Corrigida para diferentes conjuntos de dados.....	58
Figura 4.2. Comparação da variação dos erros RMS para dados estimados pela KO e pela KO Corrigida para diferentes conjuntos de dados.....	58
Figura 4.3. Comparação da variação dos coeficientes de correlação para dados estimados pela KOL Corrigida pelos métodos de Journel e Yamamoto para os diferentes conjuntos de dados.....	66
Figura 4.4. Comparação da variação dos erros RMS para dados estimados pela KOL e pela KOL Corrigida pelos métodos de Journel e Yamamoto para os diferentes conjuntos de dados.....	66
Figura 4.5. Variação dos coeficientes de correlação entre as estimativas/média das simulações e os dados exaustivos para as diversas metodologias.....	72
Figura 4.6. Variação dos erros absolutos médios para as diversas metodologias utilizadas.....	72
Figura 4.7. Variação dos erros médios para as diversas metodologias utilizadas.	73
Figura 4.8. Variação dos erros RMS para as diversas metodologias utilizadas.....	73
Figura 4.9. Maiores coeficientes de correlação das simulações SGS e TBSim comparados com o método de estimativa de maiores coeficientes (KOL Corrigida por método de Yamamoto, 2007).....	81
Figura 4.10. Menores coeficientes de correlação das simulações SGS e TBSim comparados com os métodos de estimativa de menores coeficientes de correlação resultados (KO e KO Corrigida)	82
Figura 6.1. Imagens das estimativas da Krigagem Ordinária para os conjuntos Log1 a Log9.....	89
Figura 6.2. Imagens das estimativas da Krigagem Ordinária para os conjuntos Log10 a Log18.....	90

Figura 6.3. Imagens das estimativas da Krigagem Ordinária para os conjuntos Log19 a Log27.....	91
Figura 6.4. Imagens das estimativas da Krigagem Ordinária Corrigida do Efeito de Suavização para os conjuntos Log1 a Log9.	92
Figura 6.5. Imagens das estimativas da Krigagem Ordinária Corrigida do Efeito de Suavização para os conjuntos Log10 a Log18.	93
Figura 6.6. Imagens das estimativas da Krigagem Ordinária Corrigida do Efeito de Suavização para os conjuntos Log19 a Log27.	94
Figura 6.7. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis) e P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log1 a Log9.....	95
Figura 6.8. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis) e P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log10 a Log18.	96
Figura 6.9. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis) e P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log19 a Log27.	97
Figura 6.10. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis). P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log1 a Log9.....	98
Figura 6.11. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis). P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log10 a Log18.	99
Figura 6.12. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis). P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log19 a Log27.	100
Figura 6.13. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária. Log1 a Log9.	101
Figura 6.14. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária. Log10 a Log18.	102
Figura 6.15. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária. Log19 a Log27.	103
Figura 6.16. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária corrigida do efeito de suavização. Log1 a Log9.	104
Figura 6.17. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária corrigida do efeito de suavização. Log10 a Log18.	105
Figura 6.18. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária corrigida do efeito de suavização. Log19 a Log27.	106
Figura 6.19. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Journel (1980) para os conjuntos Log1 a Log9.	107

Figura 6.20. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Journal (1980) para os conjuntos Log10 a Log18.	108
Figura 6.21. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Journal (1980) para os conjuntos Log19 a Log27.	109
Figura 6.22. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007) para os conjuntos Log1 a Log9.	110
Figura 6.23. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007) para os conjuntos Log10 a Log18.	111
Figura 6.24. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007) para os conjuntos Log19 a Log27.	112
Figura 6.25. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log1 a Log9.	113
Figura 6.26. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log10 a Log18.	114
Figura 6.27. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log19 a Log27.	115
Figura 6.28. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log1 a Log9.	116
Figura 6.29. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log10 a Log18.	117
Figura 6.30. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log19 a Log27.	118
Figura 6.31. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Journal (1980). Log1 a Log9.	119
Figura 6.32. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Journal (1980). Log10 a Log18.	120

Figura 6.33. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Journel (1980). Log19 a Log27.....	121
Figura 6.34. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007). Log1 a Log9.	122
Figura 6.35. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007). Log10 a Log18.	123
Figura 6.36. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007). Log19 a Log27.	124
Figura 6.37. Comparação dos variogramas obtidos na Krigagem Ordinária. Estão representadas os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária (círculos vazios) e para a Krigagem Ordinária corrigida do efeito de suavização (círculos cheios). Log1 a Log9.....	125
Figura 6.38. Comparação dos variogramas obtidos na Krigagem Ordinária. Estão representadas os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária (círculos vazios) e para a Krigagem Ordinária corrigida do efeito de suavização (círculos cheios). Log10 a Log18.....	126
Figura 6.39. Comparação dos variogramas obtidos na Krigagem Ordinária. Estão representadas os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária (círculos vazios) e para a Krigagem Ordinária corrigida do efeito de suavização (círculos cheios). Log19 a Log27.....	127
Figura 6.40. Comparação dos variogramas obtidos na Krigagem Ordinária Lognormal. Estão representados os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária Lognormal corrigida pelo método de Journel (1980)(círculos vazios) e para a Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto, 2007 (círculos cheios). Log1 a Log9.	128
Figura 6.41. Comparação dos variogramas obtidos na Krigagem Ordinária Lognormal. Estão representados os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária Lognormal corrigida pelo método de Journel (1980) (círculos vazios) e para a Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto, 2007 (círculos cheios). Log10 a Log18.	129
Figura 6.42. Comparação dos variogramas obtidos na Krigagem Ordinária Lognormal. Estão representados os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária Lognormal corrigida pelo método de Journel (1980) (círculos vazios) e para a Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto, 2007 (círculos cheios). Log19 a Log27.	130
Figura 6.43. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journel (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log1 a Log3.	131
Figura 6.44. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journel (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log4 a Log6.	132

Figura 6.45. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log7 a Log9.	133
Figura 6.46. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log10 a Log12.	134
Figura 6.47. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log13 a Log15.	135
Figura 6.48. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log16 a Log18.	136
Figura 6.49. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto Log19 a Log21.	137
Figura 6.50. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log22 a Log24.	138
Figura 6.51. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log25 a Log27.	139
Figura 6.52. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log1 a Log3.	140
Figura 6.53. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log4 a Log6.	141
Figura 6.54. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log7 a Log9.	142
Figura 6.55. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log10 a Log12.	143
Figura 6.56. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log13 a Log15.	144

Figura 6.57. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log16 a Log18.	145
Figura 6.58. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log19 a Log21.	146
Figura 6.59. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log22 a Log24.	147
Figura 6.60. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log25 a Log27.	148
Figura 6.61. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log1 a Log3.	149
Figura 6.62. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log4 a Log6.	150
Figura 6.63. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto Log7 a Log9.	151
Figura 6.64. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log10 a Log12.	152
Figura 6.65. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log13 a Log15.	153
Figura 6.66. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto Log16 a Log18.	154
Figura 6.67. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log19 a Log21.	155
Figura 6.68. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e	

mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log22 a Log24.	156
Figura 6.69. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log25 a Log27.	157
Figura 6.70. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log1 a Log3.	158
Figura 6.71. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log4 a Log6.	159
Figura 6.72. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log7 a Log9.	160
Figura 6.73. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log10 a Log12.	161
Figura 6.74. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log13 a Log15.	162
Figura 6.75. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log16 a Log18.	163
Figura 6.76. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log19 a Log21.	164
Figura 6.77. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log22 a Log24.	165
Figura 6.78. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log25 a Log27.	166

Lista de Tabelas

Tabela 3.1. Estatísticas descritivas dos conjuntos completos com 2500 pontos de dados.	43
Tabela 3.2. Estatísticas descritivas dos conjuntos de amostras com 90 pontos aleatórios de dados. .	44
Tabela 4.1. Estatísticas Básicas das Estimativas da Krigagem Ordinária.....	50
Tabela 4.2. Estatísticas Básicas das Estimativas da Krigagem Ordinária Corrigida do Efeito de Suavização	51
Tabela 4.3. Distâncias Médias no PP-Plot das Estimativas da Krigagem Ordinária e da Krigagem Ordinária Corrigida com os Dados Exaustivos e Amostras.....	53
Tabela 4.4. Coeficientes de Correlação e Erros da Krigagem Ordinária em relação aos Dados Exaustivos	55
Tabela 4.5. Coeficientes de Correlação e Erros da Krigagem Ordinária Corrigida do Efeito de Suavização em relação aos Dados Exaustivos	57
Tabela 4.6. Distâncias das Estimativas da Krigagem Ordinária Lognormal e da Krigagem Ordinária Lognormal Corrigida pelos Métodos de Journel (1980) e Yamamoto (2007) com os Dados Exaustivos e Amostras	60
Tabela 4.7. Estatísticas Básicas das Estimativas da Krigagem Ordinária Lognormal Corrigida pelo método de Journel (1980)	61
Tabela 4.8. Estatísticas Básicas das Estimativas pela Krigagem Ordinária Lognormal Corrigida pelo método de Yamamoto (2007)	62
Tabela 4.9. Coeficientes de Correlação e Erros da Krigagem Ordinária Lognormal Corrigida pelo Método de Journel (1980) em relação aos Dados Exaustivos.....	63
Tabela 4.10. Coeficientes de Correlação e Erros da Krigagem Ordinária Lognormal Corrigida pelo Método de Yamamoto (2007) em relação aos Dados Exaustivos.....	65
Tabela 4.11. Médias, valores máximos e mínimos dos Coeficientes de Correlação das Realizações em relação aos Dados Exaustivos (100 para cada conjunto de dados) da Simulação Gaussiana Sequencial e da Simulação por Bandas Rotativas	68
Tabela 4.12. Médias, valores máximos e mínimos dos Erros Absolutos Médios das Realizações em relação aos Dados Exaustivos (100 para cada conjunto de dados) da Simulação Gaussiana Sequencial e da Simulação por Bandas Rotativas	69
Tabela 4.13. Médias, valores máximos e mínimos dos Erros Médios das Realizações em relação aos Dados Exaustivos (100 para cada conjunto de dados) da Simulação Gaussiana Sequencial e da Simulação por Bandas Rotativas	70
Tabela 4.14. Médias, valores máximos e mínimos dos Erros RMS das Realizações em relação aos Dados Exaustivos (100 para cada conjunto de dados) da Simulação Gaussiana Sequencial e da Simulação por Bandas Rotativas	71
Tabela 4.15. Realizações da Simulação Sequencial Gaussiana com Maiores e Menores Parâmetros Estatísticos.....	77
Tabela 4.16. Realizações da Simulação Sequencial Gaussiana com Maiores e Menores Parâmetros Estatísticos.....	78

Lista de Siglas

CDF	<i>Conditional Distribution Function</i> (Função de Distribuição Condicional)
CCDF	<i>Cumulative Conditional Distribution Function</i> (Função de Distribuição Condicional Acumulativa)
KO	Krigagem Ordinária
KOL	Krigagem Ordinária Lognormal
KOL (Journel)	Krigagem Ordinária Lognormal com a Transformada Reversa de Journel (1980)
KOL (Yamamoto)	Krigagem Ordinária Lognormal com a Transformada Reversa de Yamamoto (2007)
KS	Krigagem Simples
RMS	<i>Root Mean Square</i> (Média Quadrática)
SGS	Simulação Gaussiana Sequencial
TBSim	<i>Turning Bands Simulation</i> (Simulação por Bandas Rotativas)

Sumário

1.	Introdução	1
1.1.	Objetivos	5
1.2.	Justificativas	5
2.	Revisão da Metodologia Corrente.....	6
2.1	Krigagem Ordinária.....	6
2.1.1	Incertezas associadas às estimativas da Krigagem Ordinária	10
2.1.2	Correção do Efeito de Suavização da KO.....	11
2.2	Krigagem Lognormal.....	14
2.2.1	O Cálculo do Termo de Não Viés para a Transformação Reversa das Estimativas da Krigagem Lognormal	16
2.3	Avaliação das Incertezas.....	18
2.3.1	O Intervalo de Confiança Local.....	18
2.3.2	Distribuições Locais de Probabilidade.....	19
2.3.3	A Abordagem MultiGaussiana.....	21
2.3.4	Transformada Gaussiana	22
2.3.5	Verificando a suposição Multigaussiana	23
2.4	Simulação Estocástica	25
2.4.1	Simulação Gaussiana Sequencial.....	31
2.4.2	Simulação por Bandas Rotativas	36
3.	Materiais e Métodos	41
4.	Resultados e Discussão.....	47
4.1	Estimativas da Krigagem Ordinária e da Krigagem Ordinária Corrigida do Efeito de Suavização	47
4.2	Estimativas da Krigagem Ordinária Lognormal: Transformadas Reversas de Journel e Yamamoto	59
4.3	Realizações da Simulação Estocástica	67
4.4	Comparação dos Resultados Obtidos.....	79
5.	Conclusões.....	83
	Referências Bibliográficas	85
	Anexos	88
	Anexo I – Imagens das Estimativas da Krigagem Ordinária	89
	Anexo II – Imagens das Estimativas da Krigagem Ordinária Corrigida.....	92

Anexo III – Gráficos Acumulativos e PP Plots para Amostras.....	95
Anexo IV – Gráficos Acumulativos e P-P Plots para Exaustivos.....	98
Anexo V – Gráficos de Dispersão das Distribuições Calculadas pela KO x Referência.....	101
Anexo VI – Gráficos de Dispersão das Distribuições Calculadas pela KO Corrigida x Referência ...	104
Anexo VII – Imagens KO Lognormal Corrigidas (Journel, 1980)	107
Anexo VIII – Imagens KO Lognormal Corrigidas (Yamamoto, 2007)	110
Anexo IX – Comparação das Distribuições Calculadas pela KO Lognormal x Amostras.....	113
Anexo X – Comparação das Distribuições Calculadas pela KO Lognormal x Dados Exaustivos	116
Anexo XI – Gráficos de Dispersão KOLognormal Corrigida (Journel, 1980) x Exaustivo	119
Anexo XII – Gráficos de Dispersão KOLognormal Corrigida (Yamamoto, 2007) x Exaustivo	122
Anexo XIII – Variogramas Comparados para a Krigagem Ordinária	125
Anexo XIV – Variogramas Comparados KO Lognormal	128
Anexo XV – Variação dos Coeficientes de Correlação e Erros RMS das Simulações Gaussianas Sequenciais.....	131
Anexo XVI - Realizações da SGS de Maiores e Menores Coeficientes. de Correlação	140
Anexo XVII – Coeficientes de Correlação e Erros RMS das Simulações por Bandas Rotativas	149
Anexo XVIII - Realizações da TBSim de Maiores e Menores Coeficientes. de Correlação	158

CAPÍTULO 1.

1. Introdução

O estudo de variáveis regionalizadas, campo do conhecimento denominado Geoestatística, sempre se utilizou de diferentes métodos de predição para estimar valores das variáveis em pontos não amostrados, a partir de dados discretos pontuais (pontos amostrais). Desta forma, os métodos geoestatísticos de estimativa sempre constituíram uma importante ferramenta na interpretação de dados geológicos, notadamente conhecidos por serem escassos e pontuais. São, por definição, procedimentos matemáticos que ajustam uma função para determinados pontos de dados para assim se poder estimar o valor de uma variável em um ponto não amostrado (Yamamoto, 1998).

Entre estes métodos é possível diferenciar inicialmente métodos globais e locais: os primeiros encontrando uma função que seja capaz de interpolar um valor em qualquer local dentro do domínio dos dados, sendo que qualquer alteração em algum ponto se propaga por todo o restante do domínio; os métodos locais definindo funções válidas para pequenas porções do mapa, que são aplicadas sucessivamente até cobrir toda a área de interesse, com alterações pontuais afetando apenas os pontos próximos a estas alterações. Entre os métodos globais usados na interpolação de dados geológicos, é possível citar a regressão polinomial (usada no ajuste de Superfícies de Tendência) e as equações multiquádricas (que visa representar superfícies por meio da soma de superfícies variadas matematicamente definidas). Por outro lado, os métodos locais seriam representados pela interpolação linear em triângulos, o inverso da distância e as funções de base radial. Estas últimas, se valendo de algumas funções base conhecidas para a determinação do valor da variável de acordo com a distância desta ao ponto amostrado, sendo que no caso de a função usada ser baseada em um modelo de variograma esta função passa a ser denominada de Krigagem. Uma revisão destes métodos pode ser encontrada em Yamamoto (1998).

A krigagem é o nome geral dado às técnicas de regressão linear que minimizam a variância das estimativas definida por um modelo anterior para uma dada covariância (Olea, 1991 apud Yamamoto, 1998). Dentre as formas mais comuns de krigagem, pode-se citar a

krigagem simples, krigagem ordinária e krigagem universal, entre outras. Esta técnica, em função de algumas qualidades intrínsecas como sua precisão local e confiabilidade, tornou-se a mais popular entre os métodos de interpolação supracitados, sendo largamente empregada nas mais diversas aplicações onde algum tipo de modelagem espacial é requerida, a citar: cubagem de jazidas minerais/dimensionamento de reservatórios de petróleo, delimitação de áreas contaminadas, entre outras.

Apesar de ser um ótimo procedimento de interpolação para se determinar precisamente dados espacialmente localizados, a krigagem (em suas diversas variações) apresenta algumas limitações que restringem a sua aplicabilidade. A título de ilustração, as principais limitações da krigagem são: ocorrência de pesos negativos e, conseqüentemente, teores negativos; caráter homoscedástico da variância de krigagem; efeito de suavização das estimativas de krigagem. Talvez, entre estas limitações, a mais problemática e também mais discutida na literatura atual seja o seu chamado “efeito de suavização da krigagem ordinária”. Tal efeito justifica o fato da krigagem ser considerada uma técnica de interpolação de precisão local (em contraste com as técnicas consideradas de precisão global), uma vez que o método tende a “suavizar” os valores de atributos dos dados, acarretando em uma perda no modelo de variabilidade da distribuição como um todo, ou seja, seu modelo global. Este efeito de suavização pode ser ilustrado por meio de uma comparação de dois histogramas (Figura 1.1): o primeiro de dados amostrais; o segundo representando os valores estimados pela krigagem, com base nos pontos amostrais. O efeito de suavização ficará evidente neste último caso, já que o histograma apresentará uma “perda das caudas”, isto é, a krigagem irá fazer com que os valores extremos na distribuição sejam excluídos, tanto altos como baixos, comprometendo a representação correta da variabilidade do dado original. É também observada uma concentração na frequência dos valores ao redor da média. Conseqüentemente, o modelo de variograma das estimativas não reproduz o modelo introduzido (Figura 1.2). A Figura 1.1 mostra um histograma com forte assimetria positiva, porém devido ao coeficiente de variação ser igual a 1,082 a distribuição não pode ser considerada lognormal, cujo coeficiente de variação deve ser igual a 1,254 (elevando $e=2,71828$ aos valores de uma distribuição normal). Sabe-se que quanto maior a assimetria, ou seja, o coeficiente de variação, maior o efeito de suavização e conseqüentemente o erro. Entretanto, ainda não se tem uma idéia dos erros cometidos,

conforme o coeficiente de variação. Pretende-se, neste trabalho, verificar o erro para distribuições lognormais com coeficientes de variação crescentes. Além disso, confirmar a eficiência do algoritmo de pós-processamento do efeito de suavização da krigagem ordinária e da Krigagem Ordinária Lognormal, conforme proposta de Yamamoto (2005 e 2007).

Figura 1.1 Comparação entre histogramas para os dados amostrais (esquerda) e estimativas da krigagem ordinária (direita).

Figura 1.2. Variogramas experimentais dos dados amostrais (círculos vazios) e dos valores estimados pela krigagem ordinária (círculos cheios) e modelo de variograma amostral (linha grossa contínua).

Como contraponto a estas limitações do método da krigagem ordinária, comuns a todas as técnicas de interpolação que visam a minimização da variância dos erros, há outro método alternativo para se tentar chegar a um modelo a partir de alguns poucos dados pontuais, este método sendo chamado de Simulação Estocástica. Tal técnica consiste basicamente na criação de algoritmos para se simular o comportamento de determinada variável, no caso, o seu comportamento espacial. O termo “estocástico” especificamente se refere ao fato destes algoritmos não serem determinísticos, isto é, quando há mais de uma opção entre o valor a ser assumido em determinado ponto de execução do algoritmo tal escolha será aleatória, gerando inúmeras possíveis realizações a partir de um mesmo conjunto inicial de dados.

Ao contrário da krigagem, a simulação estocástica apresenta precisão global, isto é, reproduz a variabilidade espacial da amostragem em detrimento da manutenção da precisão local. Por reprodução da variabilidade espacial entende-se que as estatísticas globais das estimativas, tais como variogramas e histogramas, serão idênticas às da amostragem (Caers, 2000). Tal fato leva à reprodução da textura geológica sugerida para a área, o que também elimina deste método a problemática do efeito de suavização apresentada pela krigagem e em todos os demais métodos baseados na minimização da variância do erro. Um dos maiores reveses da simulação estocástica talvez seja o fato de que serão geradas n imagens a partir de um mesmo conjunto de dados, todas equiprováveis, o que dificulta a escolha de uma realização ótima para qualquer situação.

A formulação da simulação estocástica pode também ser considerada como um problema de otimização (*simulated annealing*). Tal feito pode ser concretizado objetivando-se a reprodução de um determinado modelo de histograma ou de variograma ao se analisar os valores simulados. Entretanto, tais técnicas são computacionalmente intensivas e têm sido deixadas em segundo plano. Existem diversos algoritmos usados na geração de simulações estocásticas, sendo neste trabalho usadas duas de uso bastante difundido, que seriam a Simulação Gaussiana Sequencial e a Simulação por Bandas Rotativas.

As realizações da simulação estocástica não estão livres de erros, conforme referido na literatura (e.g. Olea, 1999). Além disso, não se sabe como se comportam as realizações para distribuições lognormais com coeficientes de variação crescentes. Assim, pretende-se

mensurar erros de realizações individuais e compará-los com imagens corrigidas do efeito de suavização. Observe-se que as imagens simuladas e a imagem corrigida apresentam precisão global. O objetivo é a comparação da magnitude dos erros locais.

1.1. Objetivos

Este trabalho tem por objetivo analisar erros resultantes de estimativas e simulações estocásticas usando dados de distribuições com coeficientes de variação crescentes. Amostras aleatórias simples serão extraídas de 27 distribuições completas, com coeficientes de variação entre 0,100 a 8,108. Evidentemente, as amostras não apresentarão a mesma variabilidade, principalmente em distribuições fortemente assimétricas, pois a probabilidade de se amostrar os poucos valores altos que caracterizam estas distribuições é pequena, sendo os coeficientes de variação das amostras de 0,093 a 3,566.

Outra problemática a ser abordada a partir destes dados é a determinação de qual o grau de semelhança das imagens simuladas com a imagem de referência, assim como com a estimativa da krigagem ordinária corrigida de seu efeito de suavização e a krigagem ordinária lognormal também corrigida, isto é, com a variabilidade espacial introduzida.

1.2. Justificativas

Distribuições lognormais representam fenômenos espaciais que se caracterizam por uma grande quantidade de valores baixos e uma pequena quantidade de valores altos, que podem ser considerados anômalos, mas que têm uma importância muito grande ao fenômeno em estudo. Por exemplo, em jazidas de ouro os valores anômalos representam porções com grandes concentrações do metal e garantem lucros adicionais ao empreendimento, pois tais anomalias não foram consideradas no estudo da viabilidade econômica. Se por um lado, esses valores são de difícil previsão, por outro as estimativas ou realizações devem refletir com menor erro a possibilidade de ocorrência de tais situações. Por esses motivos, as distribuições lognormais requerem estudos que permitam fazer estimativas ou simulações com maior precisão.

CAPÍTULO 2.

2. Revisão da Metodologia Corrente

2.1 Krigagem Ordinária

Uma das maiores aplicações da estatística na análise de dados geológicos é o estudo do comportamento espacial de variáveis que assumem valores definidos no espaço. O problema reside no fato de que as variáveis regionalizadas apresentam comportamento complexo demais para serem avaliadas por métodos estatísticos “clássicos”, nos quais as posições relativas das amostras são ignoradas e se considera que valores amostrais têm a mesma probabilidade de serem escolhidos.

No entanto, métodos da estatística espacial consideram valores amostrais como função de sua posição espacial, assim como a posição relativa dos pontos amostrados, o que permite a obtenção de elementos ponderadores que quantificam a semelhança entre valores amostrais em função da distância entre estas amostras (Landim, 2003).

Neste contexto, a krigagem é essencialmente um dos mais importantes métodos locais de interpolação disponíveis, técnica esta que deu origem a todo o campo da Geoestatística (Matheron, 1962). Esta é definida como um conjunto de técnicas numéricas que lidam com a caracterização de atributos espaciais (ou variáveis regionalizadas) utilizando-se principalmente de modelos aleatórios baseados em valores amostrados e taxas de flutuação espacial para se tentar minimizar a incerteza da estimativa em pontos não amostrados. Assim, a krigagem como técnica é uma generalização de formas de regressão linear para se formular um estimador ótimo de tal forma que haja a minimização dos erros médios ao quadrado, ou seja, a minimização da variância do erro (Olea, 1999).

Dentre as variadas técnicas abrangidas pela krigagem, a krigagem simples (KS) é a sua forma mais básica já que o seu modelo é o mais simples em termos de sua formulação matemática. É basicamente uma combinação linear de variáveis aleatórias nas localidades x_i , sendo seu estimador dado por:

$$z_{KS}^*(x_0) = m + \sum_{i=1}^k \lambda_i (z(x_i) - m)$$

Onde $z_{KS}^*(x_0)$ é o estimador na localidade x_0 , $Z(x)$ é a função aleatória de média m , e λ_i são os ponderadores obtidos a partir da resolução de sistema linear de equações da krigagem. Neste caso, a média é considerada constante em todo o domínio do fenômeno espacial.

A krigagem ordinária (KO), por sua vez, se diferencia em relação à krigagem simples por não necessitar da média para se encontrar os pesos que minimizam a variância do erro da estimativa. No entanto, apresenta algumas exigências, tal como a estimativa da média local, que resulta na soma unitária dos ponderadores λ_i ; e que a variável regionalizada seja estacionária, ou seja, não apresente deriva ou resíduos resultantes dela. Sua fórmula geral é:

$$z_{KO}^*(x_0) = \sum_{i=1}^n \lambda_i z(x_i)$$

A krigagem ordinária é um método robusto que minimiza a variância do erro, não apresenta enviesamento, além de ser um interpolador exato, ou seja, honra os pontos dos dados amostrais. Outra característica da krigagem ordinária se refere à precisão local das estimativas resultantes em termos da correlação entre os valores estimados e os valores dos pontos de dados.

No entanto, em decorrência de ser um algoritmo de interpolação com base na minimização da variância do erro, a krigagem como um todo (em suas diversas variações) tende a “suavizar” a variação espacial da informação em questão em suas estimativas. Uma decorrência deste fato é que valores abaixo da média serão superestimados e valores maiores que a média serão subestimados (Goovaerts, 1998), gerando uma reta de regressão entre valores estimados e de referência (após o procedimento de validação cruzada) menos inclinada que a ideal, que seria a primeira bissetriz. Tal distorção nas estimativas é denominada de enviesamento condicional das mesmas (Figura 2.1). Outra propriedade do efeito de suavização seria que o grau de suavização é inversamente proporcional à densidade de dados disponíveis (Olea, 1999). Esta suavização se reflete em uma baixa precisão global das estimativas (i.e. a não reprodução das estatísticas) e estas distorções se

manifestam nas seguintes formas: a não reprodução do variograma, sendo que o variograma experimental para os valores estimados terá amplitude maior, assim como menor variância espacial em relação ao variograma experimental das amostras, o que significaria uma continuidade exagerada nos valores; e a não reprodução no histograma dos valores estimados da distribuição do histograma da amostragem, sendo que o dos valores estimados apresenta frequências muito maiores para as classes próximas à média. Assim, devido à não reprodução da variabilidade espacial das amostras, a krigagem ordinária não é apropriada para aplicações nas quais sejam fundamentais a representação de valores extremos e seus padrões de continuidade, tais como na avaliação de recursos em depósitos minerais ou a modelagem de fluxo em meios porosos (Goovaerts, 1997).

Figura 2.1. Representação gráfica do enviesamento condicional existente na krigagem. O gráfico representa a distorção existente entre os valores reais e os estimados após a realização da Validação Cruzada, sendo possível observar que os valores reais maiores são subestimados, e os valores reais menores são superestimados. Tal distorção leva à obtenção de uma reta de regressão que se afasta da 1ª bissetriz (caso ideal). (Adaptado de Olea, 1999).

Mesmo assim, métodos como a krigagem continuam sendo a escolha natural quando é necessário se estimar atributos para os quais a precisão local seja prioritária, tal qual a estimativa de teores de blocos, em detrimento de métodos com alta precisão global, porém

com uma baixa precisão local, tais como a simulação estocástica. Neste último caso, segundo Olea (1999), as realizações da simulação estocástica apresentam normalmente erros ainda maiores que para as estimativas da krigagem, cabendo a decisão sobre qual método utilizar à ponderação do que é mais relevante para cada aplicação: erros mínimos das estimativas (krigagem) ou a representação correta da continuidade espacial (simulação estocástica).

Tendo em vista os objetivos claramente conflitantes entre a krigagem e a simulação estocástica (a precisão local vs. a precisão global; a variação espacial suave vs. a variação espacial não suave), torna-se impossível conciliar ambos em uma única imagem resultante (Caers, 2000). Portanto, a única maneira de se chegar a uma realização que represente tanto a precisão global como a local é por meio de correções efetuadas sobre as imagens geradas por um dos dois métodos, adicionando-se as características desejáveis do outro método. Levando-se em conta que a precisão local em geral é um fator mais importante na estimativa de recursos naturais a partir de dados de amostragem, a maioria dos autores que desenvolveram trabalhos nesta área optou por propor possíveis soluções para se corrigir o efeito de suavização apresentado pela krigagem ordinária, todos baseados no pós-processamento da imagem krigada resultante (Yamamoto, 2005; Krige et al., 2004).

Alguns destes métodos incluem: o uso de regressão linear e validação cruzada entre os dados e as estimativas krigadas e suavizadas para se estabelecer a relação entre elas e então corrigir as estimativas, que Olea & Pawlowsky (1996) denominam de “krigagem compensada”; o pós-processamento de estimativas valendo-se de algoritmo que reproduza o histograma das amostras (Journel & Xu, 1994); a criação de uma função baseada na representação isofatorial da distribuição bivariada do valor real contra a estimada (Guertin, 1984); o uso do algoritmo de pós processamento espectral (Yao, 1998), que se propõe a reproduzir o modelo de variograma, apesar da perda de precisão local; e método derivado do de Olea & Pawlowsky (op. cit.) que também utiliza resultados da validação cruzada para se derivar os erros verdadeiros, que são então usados para se corrigir as estimativas da KO (Yamamoto, 2005).

2.1.1 Incertezas associadas às estimativas da Krigagem Ordinária

Uma das formas de se avaliar a qualidade do estimador é por meio da incerteza associada, ou seja, por meio da dispersão em torno da estimativa. Da krigagem ordinária, se pode derivar dois tipos de variância, usados para se avaliar o grau de confiabilidade das estimativas: a variância da krigagem e a variância de interpolação.

2.1.1.1 Variância da Krigagem

A variância da krigagem é resultante da resolução do sistema de equações de krigagem ordinária e, por definição, é igual a variância mínima, para um dado arranjo espacial:

$$\sigma_{KO}^2 = C(0) - \sum_i \lambda_i \text{Cov}\{Z(x_i), Z(x_0)\} - \mu$$

A variância de krigagem não depende dos valores dos pontos de dados utilizados na krigagem e, portanto, apresenta um caráter homoscedástico. A Figura 2.2, conforme Armstrong (1994) ilustra como 2 blocos com valores completamente diferentes podem resultar num único valor da variância de krigagem, haja vista não depender dos valores locais, mas tão somente do modelo de variograma utilizado.

Figura 2.2. Conjuntos de dados distintos, porém com mesma variância da krigagem (adaptado de Armstrong, 1994).

2.1.1.2 Variância da Interpolação

Tendo em vista a não representação da totalidade da incerteza pela variância da krigagem, Yamamoto (2000) propôs como alternativa a variância da interpolação. Ao contrário da variância da krigagem, a variância de interpolação depende dos valores dos dados, sendo definida como a média ponderada das diferenças ao quadrado entre os valores dos pontos dos dados e a estimativa $z_{KO}^*(x_0)$. Desta forma, é calculada como:

$$S_0^2 = \sum_{i=1}^n \lambda_i [z(x_i) - z_{KO}^*(x_0)]^2$$

Onde λ_i são os ponderadores da KO. Como a variância de interpolação é dependente dos dados, tem-se que:

- Corresponde à propriedade da exatidão da krigagem, ou seja, caso o ponto a ser estimado seja um ponto de dado, o peso deste dado será igual a 1, e os demais terão peso 0, o que resulta em uma variância de interpolação igual a 0.
- A variância da interpolação aumenta proporcionalmente à dispersão dos dados.
- Leva em consideração a distância estrutural do variograma por conta dos ponderadores da krigagem, ou seja, quanto mais influente o dado, maior seu peso.

Outra vantagem da variância de interpolação em comparação à variância da krigagem é o reconhecimento do efeito proporcional, condição que estabelece uma relação de proporção direta entre a variância do erro e alguma função da média local dos dados amostrais (Olea, 1991 apud Yamamoto, 2000). Tal fato é claramente observável em se tratando de fenômenos naturais tais como reservas minerais, medidas de porosidade, entre outras. Assim, enquanto o desvio padrão da krigagem (raiz quadrada da variância de krigagem) não apresenta nenhuma correlação com a estimativa, o desvio padrão de interpolação apresenta alta correlação positiva com a estimativa da krigagem ordinária, principalmente quando a distribuição apresentar assimetria positiva.

2.1.2 Correção do Efeito de Suavização da KO

A metodologia de Yamamoto (2005) foi escolhida para ser aplicada neste presente estudo, visando a avaliação de sua eficiência. O procedimento para a sua realização envolve os seguintes passos, conforme segue:

Proceder à validação cruzada para se obter em cada ponto amostrado o valor estimado, o desvio padrão da interpolação e o erro verdadeiro, respectivamente:

$$z_{KO}^*(x_0) = \sum_{i=1}^n \lambda_i z(x_i)$$

$$S_0 = \sqrt{\sum_{i=1}^n \lambda_i [z(x_i) - z_{KO}^*(x_0)]^2}$$

e

$$ErroVerdadeiro = [z_{KO}^*(x_0) - z(x_0)]$$

Com base nestes valores, torna-se possível calcular o valor de uma nova variável aleatória denominada “número de desvios padrão” (N_{S_0}), que é calculada como:

$$N_{S_0} = \frac{-ErroVerdadeiro}{S_0}$$

Nesta equação, o sinal negativo no numerador irá compensar o efeito de suavização, ou seja, um erro verdadeiro positivo mostra que o valor foi sobreestimado, e portanto este erro deve ser subtraído da estimativa. Já um erro verdadeiro negativo mostra uma subestimativa, e assim o erro deve ser somado à estimativa para se chegar ao valor corrigido. No denominador tem-se o desvio padrão de interpolação que é sempre diferente de zero, haja vista o valor conhecido no ponto ser removido durante a validação cruzada na sua localização.

Prossegue-se então a determinar o número de desvios padrão da interpolação (N_{S_0}) em todos os nós a serem estimados e corrigidos. Nesta etapa, deve-se usar um número mínimo de pontos vizinhos para se evitar a suavização do valor do número de desvios padrão de interpolação. Em seguida, realizar a krigagem ordinária para se estimar o valor desconhecido $z_{KO}^*(x_0)$, bem como do desvio padrão de interpolação, que é a medida de incerteza associada à estimativa, conforme Yamamoto (2000). Assim, após o duplo processamento da krigagem ordinária, tem-se para cada ponto estimado: o valor estimado $z_{KO}^*(x_0)$, o desvio padrão de interpolação S_0 e o número de desvios padrão N_{S_0} . Combinando-se estes valores, obtém-se a estimativa corrigida $Z_{KO}^{**}(x_0)$:

$$z_{KO}^{**}(x_0) = z_{KO}^*(x_0) + N_{S_0} \cdot S_0$$

Sendo $z_{KO}^{**}(x_0)$ o valor corrigido da estimativa $z_{KO}^*(x_0)$ para o ponto x_0 .

É introduzida uma limitação para se certificar que o valor corrigido se encontre dentro do intervalo compreendido pelo conjunto de dados vizinhos usados na KO, ou seja:

$$(z_{min} < z_{KO}^{**}(x_0) < z_{max}),$$

onde:

$$z_{min} = \min[z(x_i), i = 1, n] \text{ e}$$

$$z_{max} = \max[z(x_i), i = 1, n]$$

Caso os valores corrigidos caiam fora deste intervalo de valores, estes são recalculados, a quantidade de correção $N_{S_0} \cdot S_0$ devendo ser substituída por δ , onde $\delta = -(z_{KO}^*(x_0) - z_{min})$ se o sinal de N_{S_0} for negativo; e $\delta = z_{max} - z_{KO}^*(x_0)$ caso N_{S_0} for positivo. Assim:

$$z_{KO}^{**}(x_0) = z_{KO}^*(x_0) + \delta$$

Ambas as expressões de estimativas corrigidas não garantem que a variância das estimativas corrigidas seja igual à variância amostral, sendo necessário aplicar um fator f para tal. Este fator é dado como uma constante que torna a variância das estimativas corrigidas igual à variância amostral, sendo aplicada a todos os valores corrigidos (Yamamoto, 2007):

$$\begin{cases} z_{KO}^{**}(x_0) = z_{KO}^*(x_0) + N_{S_0} \cdot S_0 \cdot f \\ z_{KO}^{**}(x_0) = z_{KO}^*(x_0) + \delta \cdot f \end{cases}$$

Que podem ser escritas como:

$$z_{KO}^{**}(x_0) = z_{KO}^*(x_0) + z_{NS_0}^*(x_0) \cdot f$$

Onde $z_{NS_0}^*(x_0)$ é a variável que substitui os termos $N_{S_0} \cdot S_0$ e δ , e f é o fator ótimo que pode ser computado como:

$$f = \frac{-2Cov(z_{OK}^*(x_0), z_{NS_0}^*(x_0)) \pm \sqrt{\Delta}}{2Var[z_{NS_0}^*(x_0)]}$$

onde:

$$\Delta = 4Cov\left(Z_{KO}^*(x_0), Z_{NS_0}^*(x_0)\right)^2 - 4Var\left[Z_{NS_0}^*(x_0)\right](Var[Z_{KO}^*(x_0)] - Var[Z(x)])$$

A última correção procura tornar a média das estimativas corrigidas próxima à média amostral. No caso das médias não serem iguais, uma constante K é adicionada às estimativas corrigidas. Esta constante K é calculada como:

$$K = E[Z(x)] - E[Z_{KO}^{**}(x)]$$

Novamente deve ser verificado se os novos valores estão dentro do intervalo compreendido pela população amostral, $S_{min} < Z_{KO}^{**}(x_0) < S_{max}$, caso contrário os valores de S_{min} e S_{max} são atribuídos às estimativas caso estas estejam respectivamente abaixo ou acima do intervalo.

Tal procedimento torna a média das estimativas igual à amostral, porém sem alterar a sua variância, já que uma das propriedades fundamentais da variância é que a variância de uma variável somada à de uma constante é sempre igual à variância da variável.

2.2 Krigagem Lognormal

A krigagem lognormal é um método de estimativa por interpolação desenvolvido para tratar especificamente de distribuições de dados altamente enviesados, conforme ocorre com diversos fenômenos naturais, como na distribuição de ocorrência de metais raros (Souza, 2007). Envolve basicamente a transformação logarítmica (normalização) dos dados para uma distribuição lognormal, que por definição é uma distribuição cujos logaritmos das observações apresentam uma distribuição normal. Seu principal objetivo é a transformação dos dados de maneira a reduzir a influência dos poucos valores extremamente altos, ou seja, reduzindo a variância dos dados e facilitando o cálculo das estatísticas básicas e das médias ponderadas (como é o caso das estimativas da krigagem ordinária). A transformação é dada pela transformação logarítmica da variável aleatória original em uma nova variável aleatória:

$$Y(x) = \ln Z(x)$$

Segundo Yamamoto (2009), pode-se dividir a variável $z(x)$ pela mediana, sem alterar a forma da distribuição resultante:

$$y(x) = \ln\left(\frac{z(x)}{\text{mediana}}\right)$$

Todo o processo de krigagem é então realizado sobre os dados transformados, seguindo os passos da krigagem ordinária:

$$y_{KO}^*(x_0) = \sum_{i=1}^n \lambda_i y(x_i)$$

Após a transformação e a krigagem dos dados transformados, é necessária a volta dos dados para o seu domínio original para sua interpretação, processo este denominado de transformada reversa (*backtransformation*), consistindo no valor exponencial da estimativa da krigagem mais um termo de não viés (Journel, 1980), baseado na variância da krigagem (σ^2) e no multiplicador de Lagrange (μ). Cabe a ressalva que a variância da krigagem é função do modelo de variograma adotado e, portanto, sujeito a grandes flutuações devido à natureza relativamente arbitrária deste.

$$Z_{KSL}^*(x_0) = \exp(Y_{KS}^*(x_0) + \sigma_{KS}^2/2) \text{ (Krigagem Simples)}$$

$$Z_{KOL}^*(x_0) = \exp(Y_{KO}^*(x_0) + \sigma_{KO}^2/2 - \mu) \text{ (Krigagem Ordinária)} \quad (2.1)$$

Levando-se em conta a divisão pela mediana da variável $z(x)$ antes da transformação logarítmica, a expressão (2.1) acima pode ser escrita como:

$$Z_{KOL}^*(x_0) = \exp(Y_{KO}^*(x_0) + \sigma_{KO}^2/2 - \mu) \cdot \text{mediana} \quad (2.2)$$

No entanto estas expressões descritas por Journel (1980) apresentam dois problemas: a esperança (E) das estimativas submetidas à transformação reversa é enviesada, ou seja, é distinta (geralmente menor) da média amostral, fato explicado pela falta de robustez na hipótese lognormal multivariada; o segundo problema deriva do fato que o termo de não viés é totalmente dependente do modelo do variograma. A problemática consiste em encontrar um termo de não viés que garanta um valor esperado

das estimativas igual à média amostral e que esses valores não sofram perda da precisão local (Yamamoto, 2007).

2.2.1 O Cálculo do Termo de Não Viés para a Transformação Reversa das Estimativas da Krigagem Lognormal

Por se tratar de um método de médias ponderadas, as estimativas resultantes da krigagem lognormal irão apresentar suavização, consequentemente o histograma dos logaritmos não será fielmente reproduzido. O algoritmo já mencionado de Yamamoto (2005) é proposto para se reproduzir o histograma amostral assim como o variograma amostral, sem a perda da precisão local. Este algoritmo resulta em duas possíveis estimativas corrigidas do efeito de suavização:

$$y_{KO}^{**}(x_0) = y_{KO}^*(x_0) + N_{S_0} S_0 \cdot f \quad (2.3)$$

$$y_{KO}^{**}(x_0) = y_{KO}^*(x_0) + \delta \cdot f \quad (2.4)$$

Onde o fator (f) é uma constante que torna a variância das estimativas corrigidas igual à variância amostral, sendo aplicada a todos os valores corrigidos.

Yamamoto (2007) propôs nova abordagem para o termo de não viés, assim como para o erro da suavização, envolvendo o pós-processamento das estimativas da krigagem ordinária:

$$Y_{KO}^{**}(x_0) = Y_{KO}^*(x_0) + Y_{NS_0}^*(x_0) \cdot f$$

Onde $Y_{NS_0}^*(x_0) \cdot f$ é o erro da suavização que é adicionado a todas as estimativas da KO, garantindo a reprodução do histograma dos valores transformados, assim como do modelo de covariância. A partir da transformação reversa deste histograma são obtidas estimativas na escala original, reproduzindo de maneira fiel o histograma experimental.

Para se determinar o valor do fator a ser encontrado são feitos rearranjos, que determinam que o seu valor ótimo seja uma raiz de uma função quadrática, de discriminante sempre positivo (duas raízes, ponto mínimo – este sempre negativo), a semelhança do descrito para a correção do efeito de suavização (Capítulo 2.1.2):

$$f_{1,2} = \frac{-2cov(Y_{OK}^*(x_0), Y_{NS_0}^*(x_0)) \pm \sqrt{\Delta}}{2var[Y_{NS_0}^*(x_0)]}$$

Onde Δ é:

$$\Delta = 4cov(Y_{KO}^*(x_0), Y_{NS_0}^*(x_0))^2 - 4var[Y_{NS_0}^*(x_0)](var[Y_{KO}^*(x_0)] - var[Y(x)])$$

As estimativas corrigidas reproduzem o variograma logarítmico e o histograma dos logaritmos sem perder a precisão local, sendo então realizada a sua transformada reversa por meio da exponenciação:

$$Z_{KOL}^{**}(x_0) = \exp(Y_{KO}^*(x_0) + Y_{NS_0}^* \cdot f) \cdot \text{mediana} \quad (2.5)$$

Sendo $Y_{NS_0}^*$ o termo não enviesado. A esperança das estimativas corrigidas submetidas à transformada reversa será mais próxima à média amostral.

O uso da expressão (2.3) é preferível pois as estimativas corrigidas resultantes são melhores, e a equação (2.4) será usada quando $y_{KO}^{**}(x_0) < y_{min}$ ou $y_{KO}^{**}(x_0) > y_{max}$ (Yamamoto, 2007). Os valores de médias para os valores transformados pela expressão (2.5) serão mais próximos da média amostral, provando que os primeiros não são enviesados; ao contrário da expressão empregada por Journel (1980) para a transformada reversa. A expressão (2.5) também promove a reprodução do histograma amostral, e também apresenta maiores coeficientes de correlação com os valores reais em relação à expressão (2.2).

Yamamoto (2007, 2009) comparou os resultados obtidos com o uso da expressão (2.2) com as estatísticas da amostras e concluiu que o seu uso gera resultados enviesados pela mediana, devido ao termo de não viés derivar da variância da krigagem e do multiplicador de Lagrange. O mesmo autor também demonstrou que os valores transformados pela expressão (2.5) apresentam melhores resultados. As estimativas submetidas à transformação reversa após a correção do efeito de suavização apresentam melhores resultados em relação àquelas que não foram corrigidas, como maiores coeficientes de correlação com os dados exaustivos dos quais as amostras foram extraídas.

Os erros (erro absoluto, erro médio e erro RMS) também são sempre menores para as estimativas corrigidas do efeito de suavização.

2.3 Avaliação das Incertezas

Goovaerts (1997) discute a modelagem da incerteza sobre o desconhecido, com o objetivo de se avaliar o risco envolvido em qualquer processo de tomada de decisão. A partir do modelo de incerteza é então possível derivar diferentes estimativas, cada uma otimizada de acordo com diferentes critérios.

Supondo o problema de se modelar dois valores desconhecidos $z(x_1)$ e $z(x_2)$, como realizações de duas variáveis aleatórias espacialmente dependentes $Z(x_1)$ e $Z(x_2)$, é possível se avaliar a incerteza nestas duas localidades de duas maneiras:

- Computar estimativas que minimizem a variância do erro (por krigagem) e derivar intervalos de confiança arbitrariamente definidos como gaussianos.
- Modelar a distribuição local de probabilidades das duas variáveis aleatórias $Z(x_1)$ e $Z(x_2)$.

2.3.1 O Intervalo de Confiança Local

A abordagem tradicional para a avaliação da incerteza em um local não amostrado x_0 envolve o cálculo de uma estimativa que minimize a variância do erro (krigagem), sendo $z^*(x_0)$, com erro de variância associado $\sigma_E^2(x_0) = \text{Var}\{Z^*(x_0) - Z(x_0)\}$. A estimativa e a variância do erro são então combinadas para se derivar um intervalo de confiança de tipo gaussiano centrado no valor estimado. Por exemplo, para um intervalo de confiança de 95%:

$$\text{Prob}\{Z(x_0) \in [z^*(x_0) - 2\sigma_E(x_0), z^*(x_0) + 2\sigma_E(x_0)]\} = 0,95$$

Onde $\sigma_E^2(x_0)$ é a variância do erro da krigagem.

É um método simples, que exige apenas uma medida da correlação entre os valores de z (variograma ou função covariância), no entanto possui duas grandes deficiências:

- O erro de estimativa $z^*(x_0) - z(x_0)$ é modelado como uma realização de um erro de variável aleatória gaussiana. Isto leva a se assumir a simetria da distribuição local dos erros (o que pode até ser verdade se considerarmos o conjunto global de erros de estimativa, mas que localmente não é verdade), o que gera enviesamento positivo ou negativo dos erros.
- A variância do erro $\sigma_E^2(x_0)$ independe dos valores dos dados. Isto implica que a variância dos erros independe dos valores dos dados, e dependem apenas da configuração dos dados (característica denominada de homoscedasticidade, raramente existente no mundo real, que pode ser avaliada pela variância da krigagem).

Assim, valores semelhantes localizados perto uns dos outros e valores bastante variáveis próximos não terão qualquer influência sobre o erro local. Desta forma intervalos de confiança baseados apenas na variância da estimativa não são soluções satisfatórias para se avaliar corretamente a incerteza local.

2.3.2 Distribuições Locais de Probabilidade

Goovaerts (1997) discute uma abordagem mais rigorosa para se avaliar as incertezas das estimativas, que seria primeiramente avaliar a incerteza sobre o desconhecido e então se deduzir uma estimativa “ótima” em algum sentido apropriado. Assim, são criados intervalos de probabilidade independentes de qualquer estimativa particular do tipo $z^*(x_0)$ do valor desconhecido $z(x_0)$, já que a incerteza depende da informação disponível (n) , e não do critério “ótimo” usado no processo de estimativa. A função distribuição $F(x_0; z|(n)) = \text{Prob}\{Z(x_0) \leq z|(n)\}$ condicionada à informação (n) gera intervalos do tipo:

$$\text{Prob}\{Z(x_0) \in (a, b]|(n)\} = F(x_0; b|(n)) - F(x_0; a|(n))$$

ou

$$\text{Prob}\{Z(x_0) > b|(n)\} = 1 - F(x_0; b|(n))$$

Cada função de distribuição de probabilidade condicional $F(x_0; z|(n))$ fornece uma medida da incerteza local, já que depende diretamente da localização específica x_0 .

A informação mínima para a determinação deste intervalo é o intervalo de limites físicos para os valores a serem assumidos pela variável, dado por $[z_{min}, z_{max}]$, como valores máximo e mínimo para os valores da variável na área, constituindo a função de distribuição acumulativa (CDF).

$$F(x_1; z) = F(x_2; z) = \begin{cases} 0 & \text{se } z \leq 0 \\ z/n & \text{se } z \in (0, n] \\ 1 & \text{se } z > n \end{cases}$$

A distribuição marginal de amostras $F^*(z)$ – que é a probabilidade de o desconhecido assumir um valor até o designado - pode ser vista como informação global disponível em qualquer localidade não amostrada:

$$F(x_0; z) = \text{Prob}\{Z \leq z\} = F^*(z) \quad \forall x_0 \in A$$

A incerteza na localidade não amostrada (x_0) é modelada a partir do histograma acumulativo dos valores de dados disponíveis. A probabilidade então do valor desconhecido exceder o valor arbitrário escolhido seria então igual à real proporção dos dados que excedem este valor.

Estes procedimentos diminuem as incertezas, porém ainda as tornam independentes da localização dos dados desconhecidos (dados vizinhos, etc.). Torna-se necessário capitalizar na dependência espacial entre 2 pontos para se atualizar o modelo global de incertezas, ou seja, torná-lo condicional às informações locais disponíveis. O uso desta informação (n) leva a uma distribuição posterior à CDF, denominada de função de distribuição acumulativa condicional (CCDF):

$$F(x_0; z|(n)) = \text{Prob}\{Z(x_0) \leq z|(n)\}$$

Onde o “|(n)” condiona a função à informação local, como dados vizinhos $z(x_\alpha)$. A rigor, deveria ser escrita também como $F(x_0; z|\{n(x)\})$, porém é referida como acima.

2.3.3 A Abordagem MultiGaussiana

A maneira mais fácil de se derivar CCDFs consiste em assumir um modelo para toda a lei espacial (distribuição multivariada) da função aleatória $Z(x)$. A lei deve ser abrangente o suficiente para que todas CCDFs $F(x_0; z|(n))$, $\forall x_0 \in A$ tenham a mesma expressão analítica e sejam completamente descritas com base em alguns poucos parâmetros (como média e variância), reduzindo o problema a se determinar estes, daí o nome abordagem paramétrica. O modelo de função aleatória multivariada gaussiana é de longe o modelo paramétrico mais usado, pois identifica de forma simples e abrangente os parâmetros da CCDF.

O modelo multigaussiano possui algumas propriedades descritas em Goovaerts (1997). Se $\{Y(x_0), x \in A\}$ for uma função aleatória multivariada Gaussiana padrão, com função covariância $CY(h)$, temos:

- Todos os subconjuntos de pontos da função aleatória $Y(x)$ também são multivariados normais, assim como a CDF de um ponto de qualquer combinação linear de componentes de variáveis aleatórias é também normal.
- A distribuição de dois pontos de quaisquer pares de variáveis aleatórias $Y(x)$ e $Y(x+h)$ é normal e determinada pela função covariância.
- Se duas variáveis aleatórias $Y(x)$ e $Y(x')$ não forem correlacionadas (ou seja, covariância = 0), elas são também independentes.

Todas as distribuições condicionais de qualquer subconjunto da função aleatória, dadas realizações de quaisquer outros subconjuntos deste, são normais (múltiplos pontos). A distribuição da variável $Y(x)$ é dada pelos parâmetros média e variância, que são a média condicional e a variância condicional da variável aleatória $Y(x)$ dada a informação (n) . A CDF padrão normal é dada por:

$$G(x_0; y|(n)) = G\left(\frac{y - E\{Y(x_0)|(n)\}}{\sqrt{\text{Var}\{Y(x_0)|(n)\}}}\right)$$

No modelo, a média e a variância da CCDF na localização x_0 são idênticas respectivamente à estimativa $(y_{SK}^*(x_0))$ e à variância $(\sigma_{SK}^2(x_0))$ da krigagem simples, determinados a partir dos dados $n(x_0)$ e $y(x_\alpha)$. A CCDF é então modelada:

$$[G(x_0; y|(n))]_{KS}^* = G\left(\frac{y - y_{KS}^*(x_0)}{\sigma_{KS}(x_0)}\right)$$

2.3.4 Transformada Gaussiana

A abordagem multigaussiana reduz o problema a resolver um sistema de krigagem simples (KS) no local x_0 , porém ao custo de que os dados devem assumir uma distribuição multigaussiana, o que implica primeiro que a distribuição dos dados de 1 ponto (histograma amostral) seja normal. Envolve os seguintes passos:

- Os dados originais z são primeiramente transformados em y com um histograma normal, processo chamado de normal score transform. $y(x_\alpha) = \phi(z(x_\alpha))$, onde y é chamado de normal score.
- Se a hipótese de bigaussianidade for válida, o modelo multigaussiano é aplicado aos normal scores permitindo a derivação da CCDF gaussiana em qualquer localidade não amostrada.

$$G(x_0; y|(n)) = \text{Prob}\{Y(x_0) \leq y|(n)\}$$

- A CCDF da variável original é então (sob a condição que ϕ é monotônica crescente):

$$F(x_0; z|(n)) = \text{Prob}\{Z(x_0) \leq z|(n)\}$$

$$= \text{Prob}\{Y(x_0) \leq y|(n)\}$$

$$= G(x_0; \phi(z)|(n))$$

A função da transformada Gaussiana pode ser derivada por meio de correspondência gráfica entre as distribuições acumuladas de um ponto das variáveis original e normal padrão. A transformada que permite se sair de uma função aleatória $Z(x_0)$ com qualquer CDF $F(z)$ para uma função aleatória $Y(x_0)$ com uma CDF gaussiana $G(y)$ é gráfica (Figura 2.3) e representada por:

$$Y(x_0) = \phi(Z(x_0)) = G^{-1}[F(Z(x_0))]$$

A transformada Gaussiana pode ser vista como uma correspondência entre quantis iguais z_p e y_p de duas distribuições, ou seja, estes dois correspondem à mesma probabilidade cumulativa p .

Em resumo, a transformação ocorre em três passos:

- ranquear os n dados originais $z(x_\alpha)$ em ordem ascendente, sendo que valores iguais devem ser desempatados pela introdução de valores aleatórios;
- a frequência de distribuição acumulada do datum $z(x_\alpha)$ com rank k é então computado;
- a transformada normal score do datum z com rank k corresponde ao quantil p_k^* da CDF normal:

$$y(x_\alpha) = G^{-1}[F^*(z(x_\alpha))] = G^{-1}(p_k^*)$$

A probabilidade de ser menor que o datum mínimo z ou ser maior que o datum máximo z é não-zero, portanto a transformada nunca gera valores infinitos.

Figura 2.3. Representação Gráfica da Transformação Gaussiana (adaptado de Olea, 1991).

2.3.5 Verificando a suposição Multigaussiana

A distribuição normal da CDF de um ponto é necessária, porém, para se determinar o modelo da função aleatória como gaussiano multivariado é necessário também se confirmar a normalidade das CDFs de 2 pontos, 3 pontos, ..., n pontos dos dados gaussianos.

2.3.5.1 Verificando para distribuição Gaussiana de dois pontos

Para verificação da hipótese de multigaussianidade dos dados, testa-se se a distribuição gaussiana de dois pontos é também normal; em caso positivo o modelo gaussiano multivariado pode ser adotado, caso contrário outro modelo deve ser buscado.

Existem vários métodos para se checar se a distribuição de dois pontos dos dados $\{y(x_\alpha), \alpha = 1, \dots, n\}$ é normal. Um deles consiste em se verificar que valores experimentais da CDF de dois pontos para quaisquer pares de dados separados pelo mesmo vetor h , tenham padrão semelhante ao modelo teórico (Goovaerts, 1997). Isto é feito pela comparação do semivariograma das indicadoras amostral $\gamma_1(h; p)$ e das indicadoras gaussianas induzidos pelo modelo Gaussiano são comparados graficamente para vários quantis p (Figura 2.4). Caso o ajuste seja bom, então se assume a hipótese de normalidade de dois pontos. Um critério de bom ajuste seria comparar a magnitude dos desvios entre variogramas de indicadoras experimentais e do modelo com aqueles observados para valores simulados conhecidamente bigaussianos.

A CDF gaussiana de dois pontos é dada por (Deutsch & Journel, op cit):

$$\begin{aligned} G(h; y_p) &= \text{Prob}\{Y(x) \leq y_p, Y(x+h) \leq y_p\} \\ &= E\{I(x; p) \cdot I(x+h; p)\} = p - \gamma_1(h; p) \end{aligned}$$

Onde $I(x; p) = 1$ caso $Y(x) \leq y_p$; zero caso contrário; e $\gamma_1(h; p)$ é o semivariograma das indicadoras para o quantil p do limiar y_p .

O modelo gaussiano não permite correlação espacial significativa para valores extremamente altos ou baixos, propriedade conhecida como efeito de desestruturação, associado à propriedade de máxima entropia do modelo de função aleatória gaussiana. Isto é ilustrado quando a frequência acumulada p tende a 0 ou 1, a CDF de dois pontos tende ao produto p^2 das duas probabilidades marginais, sendo que o variograma da indicadora tende ao seu sill (variância máxima, campo aleatório). O padrão de correlação espacial das indicadoras é simétrico com relação à mediana, o valor limiar $y_{0,5}$, ou seja, por exemplo: os 10% valores menores apresentam o mesmo padrão de variabilidade espacial que os 10% valores maiores.

Figura 2.4. Semivariogramas experimentais (pontos) e semivariogramas induzidos por modelo Gaussiano (linhas contínuas) para os nove decis da distribuição acumulada (CDF).

Nas geociências, padrões de baixa entropia, como filões conectados ou agrupamentos de valores extremos correspondem a feições que exigem atenção especial (veios, contaminação). Assim, a abordagem por funções aleatórias Multigaussianas pode não ser apropriada onde análises adicionais indicam que os valores extremos são espacialmente correlacionados. O modelo gaussiano é não conservador, de maneira que sua característica de entropia máxima leva a se subestimar a associação de valores extremos. A abordagem das indicadoras permite se levar em consideração a correlação espacial dos extremos.

2.4 Simulação Estocástica

Em resposta à problemática do efeito de suavização no uso das técnicas de interpolação derivadas de métodos de minimização da variância dos erros, diversos autores propuseram como alternativa o uso da simulação estocástica. Tal técnica consiste basicamente na criação de algoritmos para se simular o comportamento de determinada variável, no caso, o seu comportamento espacial. O termo “estocástico” especificamente se refere ao fato destes algoritmos serem não determinísticos, isto é, quando há mais de uma

opção entre o valor a ser assumido em determinado ponto de execução do algoritmo tal escolha será aleatória, gerando inúmeras possíveis realizações a partir de um mesmo conjunto inicial de dados.

É definido como o processo de se extrair realizações conjuntas, alternativas e equiprováveis de uma variável aleatória a partir de um modelo de uma função aleatória. Sendo possíveis L imagens possíveis da distribuição espacial dos valores de atributos $z(x)$ no campo A , cada imagem é denominada de “imagem estocástica”, refletindo as propriedades que foram impostas no modelo de função aleatória $Z(x)$. Assim sendo, quanto mais propriedades forem inferidas do conjunto amostral e incorporados no modelo $Z(x)$, melhor o modelo resultante (Deutsch & Journel, 1992).

Ao contrário da krigagem, tal método de estimativa possui precisão global, isto é, reproduz a variabilidade espacial da amostragem em detrimento da manutenção da precisão local. Por reprodução da variabilidade espacial entende-se que as estatísticas multiponto globais das estimativas, tais como variogramas e histogramas, serão idênticas às da amostragem (Caers, 2000). Tal fato leva à reprodução da textura geológica sugerida para a área, o que também elimina deste método a problemática do efeito de suavização presente na krigagem e em todos os demais métodos baseados na minimização da variância do erro. Apesar de modelos estocásticos serem raramente utilizados quantitativamente, a sua variabilidade auxilia na compreensão da complexidade que é encontrada nos fenômenos geológicos naturais, desafiando a crença na simplificação exacerbada dos outros modelos (como as estimativas). As diversas realizações apresentadas normalmente atraem ceticismo dos diversos profissionais envolvidos na caracterização geológica (como seções geológicas, caracterização de reservatórios) devido à sua complexidade, porém acabam gerando discussões construtivas que levam ao aprimoramento dos modelos existentes, chegando-se a um quadro mais próximo da “realidade”. A modelagem estocástica também pode ser usada para gerar apenas uma realização como base para se prever a realidade geológica. A vantagem de tal uso seria a sua habilidade de honrar um nível realístico de heterogeneidade, evitando modelos muito simplistas.

Se diferencia dos algoritmos de interpolação por dois aspectos basicamente (Deutsch & Journel, 1992):

1. Na maior parte dos métodos de interpolação o objetivo é fornecer uma “melhor” estimativa local $z^*(x_0)$ de cada valor não amostrado, um por vez, sem uma preocupação específica para as estatísticas espaciais resultantes das estimativas. Já na simulação as feições globais resultantes (textura) e as estatísticas dos valores simulados são preponderantes sobre a precisão local.
2. Para um dado conjunto de dados locais e estatísticas condicionantes, a krigagem usada como um algoritmo de interpolação fornece um único modelo numérico, que é considerado “ótimo” em algum sentido de precisão local. No entanto, a simulação fornece vários modelos numéricos alternativos, sendo cada um uma representação “ótima” no sentido da representação da realidade global. As diferenças entre os modelos e realizações alternativas fornece uma medida da incerteza espacial conjunta.

Um dos maiores reveses da Simulação Estocástica talvez seja o fato de que serão geradas n imagens a partir de um mesmo conjunto de dados, todas equiprováveis, o que dificulta a escolha de uma realização ótima para qualquer situação. Modelos estocásticos, oferecendo uma gama de possibilidades, permitem gerar múltiplos cenários de possibilidades, com modelos otimistas e pessimistas, permitindo se lidar com as incertezas (Srivastava, 1994). No entanto, a ressalva de que podem haver várias realizações se aplica igualmente aos mapas interpolados: pode haver imagens estimadas alternativas a depender do método de interpolação/estimativa usado.

A partir da década de 1980 e meados da década de 1990 que se iniciou a consolidação dos métodos estocásticos como de grande utilidade nas geociências, notadamente no campo de caracterização de reservatórios de hidrocarbonetos. Diversas metodologias específicas foram apresentadas nesta época, abrangendo inúmeros estudos de caso, permitindo a abertura de um amplo espectro de abordagens possíveis tanto do ponto de vista teórico como testadas na prática. Tendo em vista tal diversidade de métodos cabe perguntar se antes de iniciar os trabalhos qual seria o objetivo a ser atingido, tendo em vista se tal metodologia será apropriada para os fins desejados. Além disto, os diversos métodos apresentam grande variabilidade em termos de demanda de tempo, dinheiro, recursos computacionais e humanos para seu uso, devendo ser escolhido aquele que apresente um

resultado ótimo considerando as condições existentes e os objetivos em questão (Srivastava, 1994).

A formulação da Simulação Estocástica pode também ser considerada como um problema de otimização. Tal feito pode ser concretizado objetivando-se a reprodução de um determinado modelo de histograma ou de variograma ao se analisar os valores estimados. Tipicamente o modelo da função aleatória $Z(x)$ é limitado apenas pelo modelo de covariância- z inferido da covariância amostral correspondente, e o desenho é tal que todas as realizações honram os valores dos dados z nas suas localidades, estas realizações sendo denominadas de “condicionais” (aos valores dos dados) (Deutsch & Journel, 1992). Mesmo a simulação estocástica tendo sido desenvolvida para fornecer medidas de incerteza espacial, algoritmos de simulação são cada vez mais usados para fornecer um único mapa estimado. Realmente, algoritmos de simulação tem se provado muito mais versáteis na reprodução de todo o espectro de variabilidade dos dados espaciais e em considerar dados de diferentes tipos e fontes. Assim, essas realizações são usadas em aplicações onde a reprodução de feições espaciais seja mais importante que a precisão local.

Outra característica da simulação é a ergodicidade (“flutuações ergódicas”), que se refere às discrepâncias entre as estatísticas da realização e os parâmetros do modelo correspondente. A função aleatória estacionária $Z(x)$ é dita ser ergódica se as estatísticas das realizações tenderem ao parâmetro μ conforme se aumenta o tamanho do campo A , sendo este parâmetro geralmente considerado como uma média estacionária da função aleatória, ou seja, a média da sua CDF $F(z)$. Assim, provado que $Z(x)$ seja estacionária e ergódica, e o campo de simulação seja grande o suficiente, deve-se esperar que as estatísticas de qualquer realização reproduzam exatamente os parâmetros do modelo. O tamanho do campo A para determinar a ergodicidade depende do tipo de função aleatória escolhida e na amplitude das suas heterogeneidades, conforme caracterizado pelo modelo de covariância (quanto mais contínua a covariância em $h = 0$ e quanto maior a amplitude da covariância, o campo A deve ser maior para que a ergodicidade seja atingida).

Questões de ergodicidade surgem já que normalmente os usuários apenas retêm uma única realização, ou seja, usando a simulação como um algoritmo melhorado de interpolação. Muitas vezes, esta realização retida não reproduz as estatísticas do modelo

(sendo este o pior cenário possível). A única maneira de garantir a reprodução das estatísticas seria uma média levando-se em conta um número grande de realizações, sendo que quanto menos ergódico o modelo de função aleatória escolhido, maior o número de realizações necessárias para se aproximar dos valores esperados. Assim, para se ter uma noção da incerteza associada com as simulações é sempre bom reter mais de uma realização (e a depender da variabilidade entre elas, mais realizações podem dar uma dimensão melhor desta incerteza) (Deutsch & Journal, 1992).

A modelagem estocástica também permite realizar a análise de risco de Monte Carlo já que as várias realizações resultantes não são apenas plausíveis no sentido de que honram a informação disponível (condicionamento), mas que todas as realizações são equiprováveis. Tais modelos alternativos são gerados e processados para se gerar uma distribuição de possíveis valores para os parâmetros de interesse, estas distribuições sendo então usadas para se otimizar a tomada de decisão pela minimização de uma função objetiva. A análise de risco de Monte Carlo envolve a noção da distribuição de probabilidade, ao contrário dos Modelos Estocásticos para Avaliação do Impacto da Incerteza – que se baseiam nos casos extremos: otimista e pessimista. No caso a análise de risco de Monte Carlo, depende da noção de uma distribuição completa de probabilidade de possíveis valores de saída, e na esperança de que a abordagem estocástica produz valores de saída que representem razoavelmente a distribuição por completa (Srivastava, 1994).

Já para se avaliar risco pela análise de Monte Carlo dois critérios são necessários serem avaliados: a velocidade computacional e se o método estocástico está amostrando razoavelmente o espaço de incerteza. Alguns autores defendem que o espaço de incerteza deve ser teoricamente definido fora do contexto do algoritmo, para que o usuário tenha alguma esperança de checar que a família de realizações represente uma amostragem razoável do espaço de incerteza. Outros advogam que o espaço de incerteza pode apenas ser definido com base no algoritmo e isto é por definição a família de realizações que é produzida com todos os números aleatórios alimentados ao algoritmo.

Novamente, devido às precisões local e global serem objetivos conflitantes, a melhor alternativa para se obter uma estimativa que alie estas duas propriedades é a combinação de ambas, no caso da simulação estocástica através da adição de precisão local aos valores

estimados. Tal procedimento foi abordado por Caers (2000) e Goovaerts (1998), que sugeriram duas alternativas para se concretizar tal objetivo: o primeiro autor por meio da realização de simulação direta sequencial com posterior adição de spikes à distribuição local condicional; o segundo por uma modificação progressiva de uma imagem inicial por meio da técnica chamada de “simulated annealing” de modo a conferir à realização melhor condicionamento local e global, em especial em relação às informações amostrais.

São citados ainda outros métodos de simulação estocástica:

- Booleano, marked-point process e simulação baseada em objetos
- Estimativa mais erro simulado: simulação fractal e por bandas rotativas
- Simulated annealing
- Simulação de campo de probabilidade
- Métodos de decomposição de matrizes
- Métodos iterativos
- Métodos híbridos

No caso do uso dos modelos estocásticos como um catalisador visual para análise crítica, então o apelo visual deve ser o critério principal para se avaliar o resultado final. Normalmente, a abordagem booleana é a que apresenta melhores resultados neste sentido (este método envolve na modelagem o estabelecimento das formas básicas das geometrias a serem descritas, assim como parâmetros diversos – tamanho, anisotropia, que descrevem estas formas e sua distribuição por meio de algoritmos). As outras técnicas tendem a produzir imagens muito descontínuas e heterogêneas, o que para interpretação pode se tornar um fator complicador. No entanto, para modelagem petrofísica e de fluidos, onde a heterogeneidade é interessante, as imagens se tornam mais plausíveis quando apresentam uma variabilidade e algum padrão de continuidade espacial, sendo esta modelada por variogramas. Normalmente os variogramas experimentais apresentam um efeito pepita, no entanto os modelos estocásticos possuem maior apelo visual se o modelo usado não apresentar um efeito pepita. Para a modelagem de propriedades contínuas (como espessuras) as técnicas sequenciais não funcionam de maneira muito satisfatória, gerando resultados não satisfatórios visualmente devido a esquemas caóticos de atribuição de pesos.

Nestes casos, os mais recomendados seriam de estimativa mais erro simulado, de campo de probabilidade ou simulated annealing.

São descritas a seguir os métodos de simulação utilizados no presente trabalho: a simulação sequencial gaussiana e a simulação por bandas rotativas.

2.4.1 Simulação Gaussiana Sequencial

A Simulação Gaussiana Sequencial (SGS) consiste em um método para se gerar realizações parciais utilizando funções aleatórias normais multivariadas (Olea, 1999). Seu objetivo é a reprodução de propriedades da distribuição multivariada gaussiana da população amostral nas estimativas por meio do uso sequencial de distribuições condicionais, que seriam funções que descrevem completamente a distribuição de probabilidade para uma determinada variável aleatória (Caers, 2000). O fato de ser gaussiana se deve ao fato de ser necessário que os dados amostrais sejam primeiramente transformados para um espaço gaussiano normal. O modelo de variável aleatória gaussiana é único na Estatística por sua extrema simplicidade analítica e por ser a distribuição limite de muitos teoremas analíticos coletivamente conhecidos como “teorema do limite central”.

É uma alternativa para se eliminar o efeito de suavização, no entanto apresenta diversas limitações, pois não é livre de erros, e em média os erros são maiores que aqueles da krigagem (Olea, 1999). Os valores simulados reproduzem o modelo de covariância, ao custo de menor correlação entre valores simulados e valores reais. Resultam em imagens equiprováveis reproduzindo todo o espectro da variabilidade espacial dos dados (Yamamoto, 2008).

Considerando-se determinado conjunto $Z(x_j)$ de N variáveis aleatórias definidas em N localidades x_j , sendo o objetivo gerar diversas realizações conjuntas $(z(x_j), j = 1, \dots, N)$ condicionais aos dados disponíveis e a algum modelo estrutural tal como o variograma. Segundo Goovaerts (1998; Olea, 1999), pode-se demonstrar que uma distribuição multivariada de N pontos pode ser decomposta em um conjunto de N funções de distribuição acumulada (CDFs). Isto pode ser escrito de forma:

$$F(x_1, x_2, \dots, x_N; t_1, t_2, \dots, t_N | n) = F(x_1; t_1 | n) F(x_2; t_2 | n + 1) \dots F(x_n; t_n | n + N - 1)$$

Assim, temos a geração de N funções do tipo CDF levando em consideração o conjunto de valores originais e também as realizações prévias obtidas. Esta decomposição da distribuição que torna possível a geração de uma imagem visitando cada nó sequencialmente, daí o nome da técnica. Tal relação também estabelece que qualquer produto de uma série de distribuições condicionais resulta em uma distribuição multivariada específica, não necessariamente Gaussiana multivariada, fato este que se reflete nas realizações das simulações. Isto se traduz no fato de que se a sequência de CCDFs for alterada, cada realização será baseada em uma distribuição multivariada diferente. No entanto, por trás de todas as realizações existe uma classe de funções aleatórias pouco compreendidas (todas baseadas no mesmo modelo de variograma) cada uma com distribuição multivariada distinta e igualmente prováveis de serem escolhidas.

Desta forma, as sequências de Funções de Distribuição Acumulada Condicionais (CCDFs) não são usadas para a identificação de algum modelo específico de distribuição multivariada, mas apenas como meio para se atingir determinado objetivo, tal como obtenção de precisão local e reprodução do modelo de variograma. No caso específico deste último, Journel (1994, apud Caers, 2000) estabeleceu teorema que diz: “Para o algoritmo de Simulação Sequencial reproduzir determinado modelo de covariância, todas as CCDFs usadas na Simulação Sequencial devem identificar a média da krigagem simples assim como a variância derivada deste modelo de covariância”.

Resumidamente, se o fenômeno espacial contínuo $\{z(x), x \in A\}$ for gerado pela soma de um número de fontes independentes $\{y_k(x), x \in A\}$, $k = 1, \dots, K$, com distribuições espaciais semelhantes, então a sua distribuição espacial pode ser modelada pelo modelo de função aleatória multivariada gaussiana:

$$Z(x) = \sum_{k=1}^K Y_k(x) \approx \text{Gaussiano}$$

O fator limitante não é o número K, nem o fato de que as componentes $Y_K(x)$ estão igualmente distribuídas, mas sim a hipótese de independência entre as $Y_K(x)$.

Se erros humanos e de medidas podem ser consideradas como eventos ou processos independentes, nas geociências raramente os processos geológicos observados são independentes entre si ou aditivos. Modelos gaussianos multivariados são bem entendidos, e tem um histórico de aplicações de sucesso, sendo a melhor escolha para a modelagem de variáveis contínuas a menos que se provem inapropriadas (Deutsch & Journel, 1992)

A função aleatória gaussiana multivariada pode ser definida por suas propriedades características. Assim, a função aleatória $Y(x) = \{Y(x), x \in A\}$ é multivariada normal se e somente se todos os subconjuntos desta função aleatória, por exemplo, $Y(x) = \{Y(x), x \in B \subset A\}$ são também normais multivariados.

Todas as combinações lineares das componentes das variáveis aleatórias de $Y(x)$ apresentam distribuição normal (univariada), por exemplo, se

$$X = \sum_{\alpha=1}^n \omega_{\alpha} Y(x_{\alpha})$$

$\forall \omega_{\alpha}$, desde que $x_{\alpha} \in A$

Covariância (ou correlação) igual a zero determina independência completa.

$$\text{Se } \text{Cov}\{Y(x), Y(x')\} = 0,$$

Então as duas variáveis aleatórias $Y(x)$ e $Y(x')$ não são apenas não correlatas, como são também independentes.

- Todas as distribuições condicionais de qualquer subconjunto da função aleatória $Y(x)$, dadas realizações de qualquer outro subconjunto, são normais multivariadas.

O caso de $K = 1$, $x'_1 = x_0$, onde a variável aleatória $Y(x_0)$, onde a variável aleatória $Y(x_0)$ modela a incerteza sobre um valor específico não amostrado é de interesse especial: a CCDF de $Y(x_0)$, dados os n dados y_{α} , é normal e totalmente caracterizado por:

- Sua média, identificada à estimativa pela krigagem simples (regressão linear) de $y(u_0)$

$$E\{Y(x_0)|y(x_{\alpha}) = y_{\alpha}, \alpha = 1, \dots, n\} \equiv [y(x_0)]_{KS}^*$$

$$= m(x_0) + \sum_{\alpha=1}^n \lambda_{\alpha} [y_{\alpha} - m(x_{\alpha})]$$

Onde $m(x) = E\{Y(x)\}$ é o valor esperado da variável aleatória não necessariamente estacionária $Y(x)$. Os n pesos λ_{α} são dados pelo sistema da krigagem simples:

$$\sum_{\beta=1}^n \lambda_{\beta} C(x_{\beta}, x_{\alpha}) = C(x_0, x_{\alpha}), \alpha = 1, \dots, n$$

Onde $C(x, x') = \text{Cov}\{Y(x), Y(x')\}$ é a covariância, não necessariamente estacionária, da função aleatória $Y(x)$.

- Sua variância, ou variância condicional é a variância da krigagem simples:

$$\text{Var}\{Y(x_0) | y(x_{\alpha}) = y_{\alpha}, \alpha = 1, \dots, n\}$$

$$C(x_0, x_0) - \sum_{\alpha=1}^n \lambda_{\alpha} C(x_0, x_{\alpha})$$

Notar a homocedasticidade da variância condicional, que independe dos valores dos dados y_{α} , mas apenas na configuração dos dados e no modelo de covariância (Deutsch & Journel, 1992).

O algoritmo mais direto para se gerar realizações de um campo Gaussiano multivariado é com base no princípio sequencial, sendo cada variável simulada sequencialmente de acordo com sua CCDF normal caracterizada por um sistema da krigagem simples. Os dados condicionantes são aqueles originais e todos os previamente simulados encontrados na vizinhança da localidade sendo simulada (Deutsch & Journel, 1992). Para a simulação estocástica sequencial a normalidade de todas as CCDFs é uma grande vantagem já que a determinação da sequência de CCDFs consecutivas é reduzida a se resolver a sequência correspondente de sistemas da KS.

Os seguintes passos constituem os algoritmos mais comuns para a simulação sequencial (Deutsch & Journel, 1992):

1. Definição de um caminho aleatório entre os nós

2. Para cada nó encontrar os dados vizinhos mais próximos, incluindo os dados originais e valores de nós previamente simulados
3. Usar a Krigagem Simples (KS) para se estimar $Z_{KS}^*(x)$ e o desvio padrão da KS σ_{KS} para se computar o termo de erro
4. Adicionar o valor simulado $Z_{SGS}^l(x)$ ao conjunto de dados
5. Ir ao próximo nó e repetir até que todos os nós tenham sido visitados

A simulação condicional de uma variável contínua $z(x)$ modelada por uma função aleatória estacionária gaussiana $Z(x)$ procede da seguinte forma (Deutsch & Journel, 1992):

1. Determinar a CDF univariada $F_z(z)$ representativa de toda a área de estudo e não apenas dos dados amostrais z disponíveis. Declustering pode ser necessário se os dados z estiverem locados preferencialmente.
2. Usando a cdf $F_z(z)$, fazer a transformação normal dos dados z em dados y com uma CDF normal padrão.
3. Checar para a bigaussianidade dos dados normais transformados y . Caso o modelo gaussiano multivariado não puder ser obtido, considerar modelos alternativos como a mistura de populações gaussianas ou um algoritmo baseado em indicadoras para a simulação estocástica.
4. Se um modelo gaussiano multivariado puder ser adotado para a função aleatória para a variável y , procede-se com a simulação sequencial, conforme mencionado anteriormente (Deutsch & Journel, 1992).
5. Retrotransformar os valores normais simulados $\{y^{(l)}(x), x \in A\}$ em valores simulados para a variável original $\{z^{(l)}(x) = \phi^{-1}(y^{(l)}(x), x \in A)\}$. Interpolações dentro das classes e extrapolações de caudas normalmente são necessárias.

Caso sejam necessárias múltiplas realizações $\{y^{(l)}(x), x \in A\}, l = 1, \dots, L$ o algoritmo da simulação sequencial é repetido L vezes.

De fato, o termo “sequencial” deriva do fato que o conjunto de dados condicionantes é progressivamente atualizado conforme os valores são simulados. Este modelo se baseia na

premissa da “gaussianidade” multivariada, implicando na normalidade não só da CDF de um ponto, mas também das CDFs de dois, três pontos dos dados normalizados. O modelo multivariado gaussiano pode ser adotado na prática se os dados normalizados forem gaussianos bivariados (Goovaerts, 1997), usando para se checar isto o procedimento proposto pelo mesmo autor conforme visto na seção 2.3.5.1.

2.4.2 Simulação por Bandas Rotativas

O método de Simulação Estocástica por Bandas Rotativas, ou *Turning Bands* (Matheron, 1973), assim como outros métodos de simulação, envolve a transformação não linear dos dados contínuos para o domínio Gaussiano, constituindo a distribuição espacial Gaussiana Multivariada (ou MultiGaussiana). O método das bandas rotativas é, ao lado da já citada simulação gaussiana sequencial, considerado um algoritmo aproximado, ou seja, que reproduz realizações de campos aleatórios cujas distribuições espaciais são próximas da Multigaussiana, já que a distribuição do campo aleatório simulado não é ergódico nem Multigaussiana.

Este método se fundamenta na simulação de realizações de dados 2D ou 3D por meio da sua simplificação em uma série de simulações unidimensionais independentes em linhas que cruzam o espaço R^2 ou R^3 . O campo aleatório gerado será o resultado da soma de diversos campos aleatórios gerados ao longo de linhas cujas orientações espaciais permeiam a esfera unitária R^3 . Uma das suas propriedades permite que se substitua a simulação de um campo aleatório multidimensional de covariância C_y pela simulação de um campo unidimensional de covariância C_x , determinando a covariância de C_x associada a uma dada covariância isotrópica C_y (Emery & Lantuéjoul, 2006):

$$C_x(r) = \frac{d}{dr} [rC_y(r)]$$

Journel & Huijbregts (1978) detalham este método ilustrando o seu funcionamento (Figura 2.5). Considerando uma linha D_1 em um espaço tridimensional R^3 , assim como uma função aleatória unidimensional $Y(x_{D_1})$ definida nesta mesma linha D_1 . Tal função aleatória será estacionária de segunda ordem, com esperança zero $E\{Y(x_{D_1})\} = 0$ e terá uma

covariância unidimensional, $C^{(1)}(h_{D_1})$. Seja x_{D_1} a projeção de um ponto x qualquer na linha D_1 , e considerando a função aleatória tridimensional definida por $Z_1(x) = Y(x_{D_1})$, $\forall x \in R^3$. Esta função aleatória $Z_1(x)$ será estacionária de segunda ordem, com esperança zero e uma covariância tridimensional igual a:

$$E\{Z_1(x)Z_1(x+h)\} = E\{Y(x_{D_1}) \cdot Y(x_{D_1} + h_{D_1})\} = C^{(1)}(h_{D_1})$$

Onde h_{D_1} é a projeção do vetor h na linha D_1 .

Figura 2.5. Bandas rotativas (adaptado de Journel & Huijbregts, 1978).

Para se produzir a realização $z_1(x)$, o valor $z(x_{D_1})$ simulado no ponto x_{D_1} , na linha D_1 , é atribuído a todos os pontos dentro da Banda centrada no plano perpendicular a D_1 no ponto x_{D_1} , sendo que a largura da Banda é definida pela equidistância entre os valores simulados na linha D_1 (Journel & Huijbregts, 1978).

Figura 2.6 Bandas rotativas associadas à linha D_1 . Os pontos da grade bidimensional são projetados nesta linha e assumem o valor da simulação unidimensional apropriada para a sua respectiva banda (adaptado de Brooker, 1985).

Consideram-se N linhas $D_1, D_2, \dots, D_N$ correspondendo às direções dos vetores unitários distribuídos uniformemente sobre a esfera unitária. Em cada linha D_i é gerada uma realização $y(x_{D_i})$ da função aleatória $Y(x_{D_i})$, as N funções aleatórias $\{Y(x_{D_i}), i = 1 \text{ a } N\}$ sendo independentes. Desta maneira, uma realização tridimensional $z_i(x) = y(x_{D_i}), \forall x \in R^3$ corresponderá a cada realização unidimensional $y(x_{D_i})$. O valor atribuído a cada ponto x será a soma das N contribuições das N linhas:

$$z_s(x) = \frac{1}{\sqrt{N}} \sum_{i=1}^N z_i(x)$$

A covariância ou variograma são reproduzidos apenas no sentido médio. Existem diversos métodos para se simular um campo aleatório unidimensional de dada covariância, no entanto algoritmos de simulação gaussiana exata (como de médias móveis) se aplicam apenas à simulação de localidades regularmente espaçadas. Este não é o caso da simulação por bandas rotativas, onde o que se quer são as projeções de nós a serem simulados em uma linha de orientação aleatória, e isso é obtido por algoritmos de simulação contínua ao longo da linha, gerando a covariância desejada sem a necessidade de localidades regularmente espaçadas (Emery & Lantuéjoul, 2006).

Diversos modelos de covariância existem para a geração das simulações unidimensionais. Caso a escolha deste algoritmo seja precisa, o método das bandas rotativas reproduz o modelo teórico de covariância sem viés, independentemente do número de linhas N definidas. É de grande importância a decisão de quantas linhas serão geradas, sendo necessário considerar: a distribuição das linhas na esfera; a importância relativa da estrutura básica no modelo de covariância, quanto maior a importância da estrutura, maior o número de linhas; o critério usado para decidir se o modelo multigaussiano é bem reproduzido ou não; e o tipo de covariância e a técnica de simulação associada ao longo das linhas. O uso de linhas insuficientes pode gerar artefatos nas realizações simuladas.

Emery & Lantuéjoul (2006) introduzem método para o condicionamento das realizações aos dados Gaussianos. Isto é feito pela conversão das realizações não condicionais por meio da krigagem, obtida da seguinte forma: primeiramente escolhe-se uma realização não condicional na localidade x e nas localidades dos dados disponíveis pelo método de bandas rotativas; calculam-se os desvios (resíduos) entre os valores dos dados e os valores simulados nas localidades dos dados; executa-se a krigagem simples dos resíduos nas localidades com dados; e adiciona-se o resultado à realização não condicional de média zero.

Em relação à simulação sequencial gaussiana, a simulação por bandas rotativas apresenta três vantagens (Emery & Lantuéjoul, 2006): apenas uma krigagem é necessária para que as realizações sejam condicionadas aos dados, pois os pesos da krigagem serão os mesmos; a matriz da krigagem envolve apenas os dados condicionantes originais, sendo mais rápida que o algoritmo sequencial, que leva em conta também os valores simulados; e caso a simulação seja feita em um grid fino, instabilidades numéricas nos sistemas da krigagem são evitadas.

Uma limitação é que a superfície krigada será gerada usando os dados originais e o variograma/covariância modelado a partir destes dados, então os valores simulados retrotransformados são adicionados à superfície krigada. Portanto é necessário modelar a covariância/variograma tanto para os dados originais como transformados.

Emery & Lantuéjoul (2006) propuseram o algoritmo TBSIM.M para a Simulação por Turning Bands, obtendo com este resultados satisfatórios com base na análise dos valores

simulados, por meio da análise visual, análise das estatísticas das distribuições simuladas uni e bivariadas, e pela inspeção das flutuações ao longo das realizações.

CAPÍTULO 3

3. Materiais e Métodos

O trabalho iniciou-se com uma revisão bibliográfica centrada em livros e artigos científicos recentes relevantes na área de desenvolvimento da dissertação. Esta revisão, somada às disciplinas regulares de pós-graduação frequentadas, serviram para estabelecer uma base teórica com relação à teoria e fundamentos da Geoestatística, assim como uma introdução a algumas das ferramentas usadas neste trabalho.

Como o foco principal desse trabalho se refere ao estudo de métodos de estimativa e de simulação estocástica em dados com distribuições assimétricas positivas, decidiu-se trabalhar com dados sintéticos gerados em computador. O ponto de partida foi a variável secundária do arquivo true.dat (Deutsch & Journel, 1992). Esta variável foi inicialmente transformada em uma variável normal, por meio da transformação Gaussiana. Assim, utilizou-se a expressão seguinte, para obter diferentes distribuições com assimetria positiva:

$$Z \log_i = \exp(i \times 0.1 \times Z_{gauss})$$

Dessa forma, foram obtidas 27 distribuições (Log1 a Log27) com assimetria positiva, das quais as variáveis Log10 até Log27 representam distribuições lognormais. A variável Log10 representa exatamente uma distribuição lognormal, pois o seu logaritmo resulta em uma distribuição normal. Esta variável apresenta um coeficiente de variação igual a 1,254. Log27 é uma distribuição com forte assimetria positiva e coeficiente de variação igual 8,108. Embora, distribuições como estas sejam raras na natureza, considerou-se esta variável (Log27) para testar a eficiência dos métodos de estimativa e simulação estocástica. As distribuições assim obtidas representam os conjuntos completos (2500 pontos em uma malha regular de 50 por 50 pontos), dos quais foram extraídas amostras aleatórias simples com 90 pontos, ou seja, 3,6% da população total, utilizando o programa REAMOSTRA. Inicialmente foi obtida uma amostra aleatória simples com 90 pontos do conjunto completo para a variável Log1. A localização destes 90 pontos foi utilizada para extrair as demais amostras para as variáveis Log2 até Log27. A Figura 3.1 apresenta o mapa de localização dos 90 pontos de dados comum a todas as amostras.

O processamento dos conjuntos de dados foi totalmente feito no ambiente do software Geovisual – Licença Acadêmica (Yamamoto, 2002).

As estatísticas descritivas dos conjuntos completos (Tabela 3.1) são comparadas às mesmas das amostras (Tabela 3.2).

Figura 3.1. Mapa de localização dos 90 pontos de dados.

Tabela 3.1. Estatísticas descritivas dos conjuntos completos com 2500 pontos de dados.

Exaustivo	Média	D Padrão	Coef. Var.	Máx	Q Superior	Mediana	Q Inferior	Mín
Log1	1,005	0,100	0,100	1,398	1,070	1,000	0,935	0,715
Log2	1,020	0,205	0,201	1,955	1,144	1,000	0,874	0,511
Log3	1,046	0,319	0,306	2,734	1,224	1,000	0,817	0,366
Log4	1,083	0,448	0,414	3,823	1,309	1,000	0,763	0,262
Log5	1,132	0,598	0,529	5,347	1,400	1,000	0,714	0,187
Log6	1,196	0,778	0,651	7,476	1,498	1,000	0,667	0,134
Log7	1,275	0,997	0,782	10,455	1,602	1,000	0,624	0,096
Log8	1,373	1,271	0,925	14,619	1,714	1,000	0,583	0,068
Log9	1,494	1,616	1,082	20,442	1,833	1,000	0,545	0,049
Log10	1,640	2,057	1,254	28,586	1,961	0,999	0,509	0,035
Log11	1,818	2,626	1,444	39,973	2,098	0,999	0,476	0,025
Log12	2,035	3,367	1,655	55,896	2,244	0,999	0,445	0,018
Log13	2,299	4,339	1,887	78,161	2,400	0,999	0,416	0,013
Log14	2,621	5,621	2,145	109,297	2,568	0,999	0,389	0,009
Log15	3,016	7,323	2,429	152,835	2,746	0,999	0,363	0,007
Log16	3,501	9,593	2,740	213,716	2,938	0,999	0,340	0,005
Log17	4,099	12,631	3,081	298,849	3,143	0,999	0,318	0,003
Log18	4,841	16,714	3,453	417,895	3,361	0,999	0,297	0,002
Log19	5,764	22,221	3,855	584,362	3,596	0,999	0,277	0,002
Log20	6,919	29,668	4,288	817,140	3,846	0,999	0,259	0,001
Log21	8,370	39,769	4,752	1142,646	4,114	0,999	0,242	0,001
Log22	10,200	53,501	5,245	1597,814	4,401	0,999	0,227	0,001
Log23	12,521	72,212	5,767	2234,298	4,707	0,999	0,212	0,000
Log24	15,476	97,757	6,317	3124,324	5,035	0,999	0,198	0,000
Log25	19,256	132,696	6,891	4368,887	5,386	0,999	0,185	0,000
Log26	24,110	180,562	7,489	6109,219	5,762	0,999	0,173	0,000
Log27	30,368	246,236	8,108	8542,805	6,163	0,999	0,162	0,000

Tabela 3.2. Estatísticas descritivas dos conjuntos de amostras com 90 pontos aleatórios de dados.

Amostra	Média	D Padrão	Coef. Var.	Máx	Q Superior	Mediana	Q Inferior	Mín
Log1	1,001	0,093	0,093	1,246	1,052	0,998	0,933	0,772
Log2	1,010	0,189	0,187	1,552	1,108	0,996	0,871	0,596
Log3	1,028	0,291	0,283	1,933	1,166	0,995	0,813	0,460
Log4	1,056	0,403	0,382	2,408	1,227	0,993	0,759	0,355
Log5	1,093	0,530	0,485	3,000	1,291	0,991	0,708	0,274
Log6	1,142	0,676	0,592	3,738	1,359	0,990	0,661	0,211
Log7	1,202	0,848	0,705	4,656	1,431	0,988	0,617	0,163
Log8	1,277	1,053	0,825	5,801	1,506	0,986	0,576	0,126
Log9	1,367	1,299	0,950	7,226	1,586	0,984	0,537	0,097
Log10	1,476	1,597	1,082	9,002	1,669	0,983	0,502	0,075
Log11	1,606	1,959	1,220	11,214	1,757	0,981	0,468	0,058
Log12	1,761	2,401	1,364	13,970	1,850	0,979	0,437	0,045
Log13	1,945	2,941	1,512	17,404	1,948	0,977	0,408	0,034
Log14	2,165	3,604	1,664	21,681	2,051	0,976	0,381	0,027
Log15	2,427	4,416	1,820	27,009	2,160	0,974	0,355	0,021
Log16	2,739	5,415	1,977	33,647	2,274	0,972	0,332	0,016
Log17	3,112	6,643	2,135	41,915	2,395	0,971	0,309	0,012
Log18	3,557	8,153	2,292	52,216	2,522	0,969	0,289	0,009
Log19	4,090	10,013	2,448	65,049	2,656	0,967	0,270	0,007
Log20	4,729	12,305	2,602	81,035	2,797	0,966	0,252	0,006
Log21	5,495	15,128	2,753	100,951	2,946	0,964	0,235	0,004
Log22	6,417	18,610	2,900	125,760	3,103	0,962	0,219	0,003
Log23	7,526	22,904	3,043	156,667	3,268	0,960	0,205	0,003
Log24	8,865	28,202	3,181	195,169	3,443	0,959	0,191	0,002
Log25	10,481	34,741	3,315	243,133	3,626	0,957	0,178	0,002
Log26	12,436	42,814	3,443	302,885	3,820	0,955	0,166	0,001
Log27	14,804	52,784	3,566	377,321	4,024	0,954	0,155	0,001

Com as amostras disponíveis foram feitos variogramas experimentais para cada uma das 27 amostras com o programa VARCOM, que calcula variogramas experimentais para dados de distribuição irregular, gerando arquivos tipo logXX.var (para variogramas originais), logXXn01.var (para variogramas gaussianos) e logXXlog.var (para variogramas lognormais). Foram usados os seguintes parâmetros: 10 passos com abertura igual a 3 e omnidirecional, ou seja, com tolerância angular igual a 90°. Em seguida foi feita a modelagem dos mesmos utilizando o programa MODVARG, com a definição dos parâmetros básicos (número de estruturas, modelo, variância espacial, amplitude mínima e máxima), gerando arquivos do

tipo logXX.mod (para variogramas originais), logXXn01.mod (para variogramas gaussianos) e logXXlog.mod (para variogramas lognormais).

A seguir, com base nos variogramas modelados, foi feita a interpolação dos dados usando o programa CROSSORDKRIG2E sendo feita a krigagem normal e lognormal com os respectivos variogramas modelados, de acordo com os parâmetros descritos na etapa anterior. Como parâmetros da Krigagem foram usados um grid de 50x50 nós, com abertura dx e dy de 1, totalizando 2500 nós (como o conjunto original). Também foi definido um número de 3 pontos por setor.

Este procedimento também envolveu a aplicação do método de correção do efeito de suavização proposto por Yamamoto (2005), optando-se pelo uso da correção tipo delta e a opção para NS0 como nppq=3 (número de pontos por quadrante). Tais procedimentos resultaram em estimativas corrigidas, assim como a estimativa dos erros dados pela mesma com base nos parâmetros Desvio de Interpolação (Yamamoto, 2000) e Desvio da Krigagem (Journel & Huijbregts, 1978). Os parâmetros de krigagem mostraram-se bastante variados conforme o conjunto de dados do qual a amostra fora retirada.

No caso da obtenção das estimativas pela krigagem ordinária lognormal foi adotado o mesmo procedimento usado para a KO, porém com a geração de arquivos no CROSSORDKRIG2E que apresentavam os resultados conforme as diferentes técnicas de transformação reversa mencionadas. No caso, foram geradas estimativas transformadas pela simples exponenciação de Y_{KO}^* , pelo algoritmo de Journel (1980) e pelo algoritmo de Yamamoto (2007).

Os procedimentos então adotados se basearam na rotina do estudo de Yamamoto (2008), em que visa comparar os resultados obtidos por meio de um algoritmo de pós-processamento para correção do efeito de suavização da krigagem ordinária (KO) e entre aqueles resultantes de realizações da simulação gaussiana sequencial (SGS). A obtenção das simulações gaussianas sequenciais se deu com o uso do programa SGSIM2D, baseado no programa de domínio público SGSIM (Deutsch & Journel, 1992), em que os valores originais são transformados em valores normalizados e então retro-transformados em valores na escala original, exigindo assim o variograma para os dados normais.

Com base nos variogramas modelados no procedimento da krigagem, foram feitas simulações para os mesmos conjuntos de dados, no ambiente do SGSIM2D. No caso, optou-se pela simulação gaussiana sequencial (SGS) usando como variograma para os dados gaussianos aquele anteriormente modelado. Como fronteira convexa foi escolhida a área total, determinada anteriormente no arquivo logXX.cvx. Foram feitas 100 realizações para cada conjunto de dados, usando como tipo de krigagem a krigagem simples, e gravada a variável original. Estas realizações foram salvas em arquivos em que eram juntados grupos de 10 realizações em um único arquivo.

Outra abordagem para a Simulação Estocástica foi o uso da simulação por bandas rotativas, por meio da rotina TBSim proposta por Emery & Lantuéjoul (2006). Tal rotina foi compilada em Delphi por Saka & Yamamoto (2008), sendo usado o programa Turning_Bands.exe para a geração de 100 realizações para cada um dos 27 conjuntos amostrais. Foram usados os mesmos parâmetros usados na simulação gaussiana sequencial, incluindo o variograma para os dados transformados para o espaço gaussiano. A transformada reversa para o espaço original foi feita por meio de tabela de conversão .trn. As realizações também foram juntadas em arquivos contendo 10 realizações cada para a análise estatística da qualidade das mesmas com relação aos dados exaustivos.

De maneira a se analisar o conjunto de resultados obtidos tanto para as simulações gaussianas sequenciais como as simulações por bandas rotativas foram feitas as médias das 100 realizações de cada conjunto de dados para alguns parâmetros estatísticos, tais como: coeficiente de correlação com os dados exaustivos, erro absoluto médio, erro médio e erro RMS. Tais médias que foram tomadas como base para comparar a qualidade dos resultados destes métodos com os demais.

Com base nos resultados tanto para a krigagem ordinária corrigida (uma única estimativa), para a krigagem ordinária lognormal e nas realizações das simulações para os mesmos conjuntos de dados, foram analisadas as diferentes relações existentes entre os resultados, com base em gráficos de correlação entre os mesmos, tomando como base o procedimento adotado por Yamamoto (2008).

CAPÍTULO 4

4. Resultados e Discussão

Neste capítulo serão abordados os resultados obtidos por meio do uso dos procedimentos descritos. Os resultados são apresentados na forma de figuras, tabelas e gráficos de maneira a ilustrar da melhor maneira as diferenças observadas no uso de diferentes métodos geoestatísticos para os diferentes conjuntos de dados.

Embora o pós-processamento proposto por Yamamoto (2005) procure reproduzir as características das amostras, optou-se também por comparar os resultados com os dados completos, ou exaustivos.

4.1 Estimativas da Krigagem Ordinária e da Krigagem Ordinária Corrigida do Efeito de Suavização

Os resultados da interpolação da malha regular bidimensional de 50x50 nós por meio da krigagem ordinária, assim como os resultados decorrentes da aplicação do algoritmo de correção do efeito de suavização da mesma (Yamamoto, 2005) são apresentados respectivamente nos Anexos I e II. Já as estatísticas descritivas dos resultados obtidos nas imagens interpoladas por estes métodos encontram-se nas Tabelas 4.1 e 4.2.

É observável que as imagens representam um progressivo aumento na variância, em acordo com o que ocorre com as amostras, notado pelos valores existentes nas escalas apresentadas, que tendem a aumentar a diferença entre os valores mínimo e máximo conforme a amostra em estudo, sendo que a menor variância corresponde ao conjunto de dados Log1 e a maior ao conjunto Log27. Comparando as diferenças entre as imagens resultantes da krigagem ordinária e da krigagem ordinária corrigida para cada amostra, temos que para todos os 27 conjuntos de dados na imagem corrigida há uma maior representatividade dos valores extremos com relação à imagem não corrigida, sendo inclusive os valores máximos e mínimos da primeira respectivamente maiores e menores que da segunda, conforme constam na comparação das Tabelas 4.1 e 4.2. Outro parâmetro estatístico que demonstra a melhora a representatividade dos dados após a correção do

efeito de suavização é o aumento do coeficiente de variação comparando-se os mesmos conjuntos de dados.

Comparando-se os parâmetros estatísticos das imagens interpoladas pela KO com aqueles das amostras é possível determinar uma coerência maior para os valores da KO corrigida do efeito de suavização com os dados amostrais, especialmente no que se refere ao desvio padrão, ao coeficiente de variação (devido ao efeito de suavização) e aos valores atribuídos aos quartis, máximo e mínimo. Em linhas gerais, o que se observa com a correção do efeito de suavização é uma aproximação da variância calculada em relação à variância amostral. Estes fatores levam a Krigagem Ordinária corrigida do efeito de suavização a representar melhor o histograma amostral e, conseqüentemente, o variograma, em relação à krigagem suavizada. A correspondência com os parâmetros estatísticos para os dados exaustivos é pequena, especialmente para os dados de maior coeficiente de variação, porém isso se deve à dificuldade de modelar dados de grande variabilidade com base em um pequeno espaço amostral.

São também apresentadas as curvas acumulativas para a Krigagem Ordinária, para a Krigagem Ordinária Corrigida e para as amostras (Anexo III) e dados exaustivos (Anexo IV). É possível observar pelas curvas acumulativas que a KO Corrigida apresenta boa fidelidade com as amostras e dados exaustivos para os conjuntos de dados que apresentam baixo coeficiente de variação (até Log7 ou Log8), ao contrário da KO sem a correção, cujo efeito de suavização se encontra bem ilustrado nestes gráficos. Os resultados ausentes da Tabela 4.2, correspondentes aos quartis inferiores e das medianas para os conjuntos de dados a partir de Log14 podem ser explicados nas curvas acumulativas (Anexos III e IV), que para esses dados se inicia em mais que 25% ou 50%, o que impossibilita de se saber o quartil inferior ou a mediana.

É possível também perceber que a KO Corrigida, na parte que corresponde à parte inferior da cauda, apresenta maior desvio com relação aos valores exaustivos e amostrais. Esse problema ocorre devido à natureza da distribuição de frequências, qual seja, de assimetria positiva, não necessariamente uma distribuição lognormal, pois o problema ocorre para coeficientes de variação inferiores a 1,254. Na realidade, o tipo de dado sintético gerado pela exponenciação de dados gaussianos resulta em 50% de valores

menores que 1 e os outros 50% em valores maiores que 1. Assim, a mediana teórica das distribuições completas é igual a um (Tabela 3.1). Da mesma forma, a mediana das distribuições amostrais deverá ser próxima de um (Tabela 3.2). Quando aplicado o algoritmo de correção do efeito de suavização para valores pertencentes à cauda inferior da distribuição, verifica-se que o valor estimado pela krigagem é superestimado e que uma quantidade de correção deve ser subtraída. Porém, como os valores da cauda inferior estão compreendidos entre zero e 1, a correção não é efetiva, resultando na maioria dos casos que $z_{KO}^{**}(x_0) < z_{min}$ e assim $z_{KO}^{**}(x_0) = z_{min}$. Quando isso ocorre a frequência de valores iguais a z_{min} aumenta e a distribuição de frequências acumuladas começa em porcentagens muito superiores à porcentagem amostral. Por exemplo, a variável Log9 (Anexo III – Figura 6.7) estimada com a KO corrigida é de 5%, enquanto a porcentagem acumulada amostral é de apenas 1%.

Aparentemente este problema pode ser devido ao algoritmo de correção da suavização, mas é de natureza dos dados sintéticos (de mediana igual a 1). Na verdade, a principal causa desse problema é a forte assimetria dos dados, o que deve ser corrigido por meio de uma transformação apropriada dos dados, no caso, a transformação logarítmica. A transformação logarítmica amplia quando os valores originais estão situados entre zero e 1, e reduz quando os valores originais são maiores que 1. As mesmas considerações feitas para o Anexo III são válidas para o Anexo IV, pois a comparação neste último caso é feita com a distribuição do conjunto exaustivo.

Nos gráficos P-P, apresentados em quadro reduzido nas curvas acumulativas nos mesmos Anexos III e IV, é possível comparar respectivamente as probabilidades dos dados amostrais e exaustivos e das estimativas, exibindo o claro efeito de suavização nos círculos verdes (estimativas da KO), que se distanciam da primeira bissetriz (que representam uma probabilidade igual à dos dados de referência). As estimativas corrigidas, representadas em quadrados azuis, apresentam resultados melhores neste aspecto, sendo que as probabilidades entre os dados exaustivos e estimativas tornam-se semelhantes, ao menos para as distribuições mais simétricas (até Log10). Neste caso há uma atenuação do efeito de suavização da krigagem ordinária. À medida que aumenta a assimetria da distribuição de frequências, o desvio observado na cauda inferior da distribuição de frequências dos valores

corrigidos fica mais evidente, com o distanciamento cada vez maior dos pontos em relação à primeira bissetriz do gráfico P-P. Isto indica a não reprodução adequada do histograma de distribuição amostral ou exaustivo pelas estimativas corrigidas a partir da variável Log10.

Tabela 4.1. Estatísticas Básicas das Estimativas da Krigagem Ordinária

Conjunto	Nº Dados	Média	D. Padrão	Coef. Var.	Máx	Q. Superior	Mediana	Q. Inferior	Mín
Log1	2500	1,002	0,058	0,058	1,226	1,039	1,003	0,961	0,790
Log2	2500	1,013	0,122	0,120	1,512	1,087	1,008	0,927	0,632
Log3	2500	1,033	0,188	0,182	1,866	1,138	1,016	0,901	0,508
Log4	2500	1,063	0,260	0,245	2,307	1,200	1,030	0,881	0,413
Log5	2500	1,102	0,342	0,310	2,856	1,268	1,048	0,869	0,340
Log6	2500	1,153	0,436	0,378	3,540	1,348	1,071	0,860	0,285
Log7	2500	1,216	0,546	0,449	4,392	1,434	1,101	0,856	0,243
Log8	2500	1,293	0,676	0,523	5,452	1,536	1,134	0,858	0,211
Log9	2500	1,386	0,832	0,601	6,771	1,654	1,172	0,863	0,188
Log10	2500	1,496	0,980	0,656	8,373	1,743	1,245	0,903	0,158
Log11	2500	1,630	1,249	0,766	10,457	1,938	1,272	0,884	0,160
Log12	2500	1,789	1,526	0,853	13,001	2,103	1,333	0,906	0,152
Log13	2500	1,978	1,865	0,943	16,167	2,287	1,399	0,931	0,139
Log14	2500	2,202	2,280	1,035	20,108	2,506	1,477	0,945	0,129
Log15	2500	2,469	2,787	1,129	25,012	2,754	1,569	0,974	0,121
Log16	2500	2,786	3,410	1,224	31,117	3,041	1,658	0,999	0,115
Log17	2500	3,164	4,176	1,320	38,715	3,356	1,760	1,031	0,111
Log18	2500	3,615	5,116	1,415	48,173	3,762	1,876	1,065	0,104
Log19	2500	4,154	6,274	1,510	59,944	4,189	2,004	1,101	0,097
Log20	2500	4,800	7,699	1,604	74,596	4,719	2,147	1,138	0,091
Log21	2500	5,573	9,455	1,697	92,833	5,259	2,287	1,182	0,085
Log22	2500	6,501	11,620	1,787	115,533	5,906	2,445	1,225	0,081
Log23	2500	7,618	14,290	1,876	143,789	6,756	2,619	1,270	0,077
Log24	2500	8,898	16,950	1,905	178,194	8,066	2,863	1,433	0,054
Log25	2500	10,584	21,651	2,046	222,742	8,801	3,017	1,372	0,070
Log26	2500	12,542	26,673	2,127	277,240	10,047	3,247	1,434	0,065
Log27	2500	14,912	32,877	2,205	345,079	11,555	3,495	1,518	0,061

Tabela 4.2. Estatísticas Básicas das Estimativas da Krigagem Ordinária Corrigida do Efeito de Suavização

Conjunto	Nº Dados	Média	D. Padrão	Coef. Var.	Máx	Q. Superior	Mediana	Q. Inferior	Mín
Log1	2500	1,001	0,093	0,093	1,240	1,060	0,998	0,937	0,772
Log2	2500	1,010	0,188	0,186	1,537	1,128	0,995	0,881	0,596
Log3	2500	1,028	0,290	0,282	1,907	1,201	0,993	0,826	0,460
Log4	2500	1,056	0,401	0,380	2,370	1,281	0,992	0,772	0,355
Log5	2500	1,093	0,527	0,482	2,947	1,371	0,989	0,720	0,274
Log6	2500	1,142	0,672	0,589	3,666	1,464	0,984	0,672	0,211
Log7	2500	1,202	0,843	0,701	4,562	1,570	0,981	0,619	0,163
Log8	2500	1,277	1,046	0,819	5,679	1,691	0,972	0,557	0,126
Log9	2500	1,367	1,289	0,943	7,070	1,806	0,959	0,492	0,097
Log10	2500	1,476	1,581	1,071	8,823	1,904	0,952	0,421	0,075
Log11	2500	1,608	1,931	1,201	10,958	2,096	0,902	0,330	0,058
Log12	2500	1,765	2,358	1,336	13,643	2,243	0,858	0,226	0,045
Log13	2500	1,955	2,878	1,472	16,986	2,423	0,804	0,107	0,034
Log14	2500	2,184	3,513	1,608	21,148	2,592	0,716		0,027
Log15	2500	2,461	4,288	1,742	26,330	2,753	0,619		0,021
Log16	2500	2,798	5,237	1,872	32,782	3,001	0,475		0,016
Log17	2500	3,207	6,401	1,996	40,817	3,166	0,268		0,012
Log18	2500	3,706	7,830	2,113	50,826	3,289	0,036		0,009
Log19	2500	4,312	9,587	2,224	63,293	3,393			0,007
Log20	2500	5,046	11,751	2,329	78,826	3,676			0,006
Log21	2500	5,942	14,419	2,427	98,182	3,800			0,004
Log22	2500	7,031	17,711	2,519	122,303	3,904			0,003
Log23	2500	8,357	21,780	2,606	152,356	3,995			0,003
Log24	2500	9,957	26,914	2,703	189,864	3,799			0,002
Log25	2500	11,950	33,026	2,764	236,463	3,999			0,002
Log26	2500	14,356	40,721	2,857	294,610	3,738			0,001
Log27	2500	17,296	50,246	2,905	367,060	3,665			0,001

As distâncias médias no P-P Plot, representadas na Tabela 4.3 são definidas por:

$$Distância\ Média = \frac{100 \cdot \cos(45)}{n} \sum_{i=1}^n abs(P_{ref_i} - P_{obs_i})^2$$

Onde P_{ref_i} e P_{obs_i} são percentis das distribuições de referência (dado amostral ou exaustivo) e estimada, respectivamente. Por serem multiplicadas por 100 estas distâncias podem ser interpretadas como um erro percentual, sendo que valores em torno de 1-2% são erros perfeitamente aceitáveis para duas distribuições semelhantes (Yamamoto, 2005). Uma distância de zero significa que a distribuição observada e a referência são idênticas.

Inicialmente é preciso considerar que a distância entre a amostra e o dado completo é sempre igual a 1,75 (Tabela 4.3), pois todas as amostras foram extraídas usando a mesma localização de pontos, conforme a Figura 3.1. Analisando a Tabela 4.3 é possível observar que as distâncias entre o valor de referência e as estimativas corrigidas são bastante inferiores com relação às distâncias para as estimativas não corrigidas, tomando como referência tanto o conjunto exaustivo como o conjunto amostral. A partir da amostra Log11 começa a aumentar de forma bastante significativa a distância, sendo que a partir de Log17 as distâncias entre os valores estimados corrigidos e os valores de referência são maiores que as distâncias entre os valores estimados não corrigidos e a referência. Além disso, pode-se utilizar o valor de 1,75 como limite de aceitação para a semelhança entre distribuições dos valores calculados e distribuições dos dados exaustivos ou amostrais. Segundo esse critério, nenhuma distribuição calculada pela KO é semelhante à respectiva distribuição amostral ou exaustiva. Por outro lado, em relação às distribuições de valores corrigidos, verifica-se que até a variável Log9 as distribuições são semelhantes às distribuições exaustivas, e que até a variável Log7 as distribuições são semelhantes às distribuições amostrais. No caso da variável Log27 a distância é superior a 19, que fica muito próximo do valor encontrado na comparação com uma distribuição aleatória (22, 122) por Yamamoto (2005), ou seja, isso significa que a distribuição corrigida nada tem a ver com a distribuição exaustiva ou amostral.

Tabela 4.3. Distâncias Médias no PP-Plot das Estimativas da Krigagem Ordinária e da Krigagem Ordinária Corrigida com os Dados Exaustivos e Amostras

	Exhaust x Amostra	Conjunto Exaustivo		Conjunto Amostral	
		Est KO	Est Corrigidas	Est KO	Est Corrigidas
Log1	1,75	6,030	0,990	4,660	1,170
Log2	1,75	5,800	0,980	4,500	1,150
Log3	1,75	5,940	0,940	4,680	1,220
Log4	1,75	6,100	0,880	4,920	1,270
Log5	1,75	6,270	0,830	5,280	1,370
Log6	1,75	6,460	0,850	5,690	1,490
Log7	1,75	6,660	0,950	6,080	1,670
Log8	1,75	6,890	1,210	6,520	1,970
Log9	1,75	7,120	1,620	6,940	2,460
Log10	1,75	7,880	2,040	7,820	2,850
Log11	1,75	7,640	2,930	7,730	3,830
Log12	1,75	7,890	3,820	8,110	4,730
Log13	1,75	8,150	4,890	8,580	5,750
Log14	1,75	8,400	6,100	9,110	6,910
Log15	1,75	8,700	7,420	9,610	8,150
Log16	1,75	8,960	8,750	10,100	9,420
Log17	1,75	9,210	10,090	10,480	10,710
Log18	1,75	9,440	11,320	10,820	11,850
Log19	1,75	9,760	12,460	11,260	12,890
Log20	1,74	10,000	13,660	11,670	14,060
Log21	1,74	10,230	14,660	11,940	15,020
Log22	1,74	10,560	15,600	12,230	15,920
Log23	1,74	10,810	16,490	12,530	16,790
Log24	1,74	11,840	17,320	13,830	17,590
Log25	1,74	11,350	17,910	12,830	18,240
Log26	1,74	11,660	18,540	13,050	18,910
Log27	1,74	11,960	19,080	13,380	19,410

Como parâmetros estatísticos para a avaliação da qualidade dos dados estimados ou simulados foram usados o coeficiente de correlação, o coeficiente de correlação rankeado, o erro absoluto médio, o erro médio e o erro RMS (*root mean square*).

O coeficiente de correlação de Pearson é definido por:

$$\rho_{x,y} = \frac{1}{n} \sum_{i=1}^n \left(\frac{x_{1,i} - \bar{x}_1}{\sigma_{x_1}} \right) \left(\frac{x_{2,i} - \bar{x}_2}{\sigma_{x_2}} \right)$$

Sendo $\bar{x}_1$ o valor esperado para os valores exaustivos, $\bar{x}_2$ o valor esperado para os valores estimados ou simulados, σ_{x_1} e σ_{x_2} o desvio padrão dos valores exaustivos e calculados respectivamente.

O coeficiente de correlação rankeado é usa a mesma fórmula de Pearson, porém os dados das variáveis x_1 e x_2 são classificados em ordem crescente, independentemente. Ao invés dos valores propriamente ditos, usa-se o *rank* (variando de 1 a n) para calcular a correlação rankeada.

O erro absoluto médio por sua vez é:

$$EAM = \frac{1}{n} \sum_{i=1}^n |x_{1,i} - x_{2,i}|$$

Onde $x_{1,i}$ corresponde ao valor esperado para o ponto em questão (valor exaustivo) e $x_{2,i}$ é o valor estimado ou simulado.

O erro médio é simplesmente:

$$EM = \frac{1}{n} \sum_{i=1}^n (x_{1,i} - x_{2,i})$$

Por último, o erro RMS (*root mean square*) ou erro quadrático médio é por sua vez calculado por:

$$ERMS = \sqrt{\frac{\sum_{i=1}^n (x_{1,i} - x_{2,i})^2}{n}}$$

Tabela 4.4. Coeficientes de Correlação e Erros da Krigagem Ordinária em relação aos Dados Exaustivos

Conjunto	Coeficiente Correlação	Correlação Rankeada	Erro Absoluto Médio	Erro Médio	Erro RMS
Log1	0,851	0,852	0,044	-0,003	0,060
Log2	0,865	0,869	0,085	-0,007	0,117
Log3	0,863	0,869	0,130	-0,012	0,184
Log4	0,858	0,868	0,180	-0,020	0,263
Log5	0,849	0,867	0,235	-0,030	0,358
Log6	0,838	0,863	0,299	-0,043	0,478
Log7	0,824	0,859	0,373	-0,059	0,632
Log8	0,808	0,853	0,460	-0,081	0,831
Log9	0,790	0,846	0,565	-0,108	1,091
Log10	0,771	0,815	0,716	-0,144	1,451
Log11	0,749	0,829	0,845	-0,188	1,892
Log12	0,726	0,818	1,032	-0,246	2,502
Log13	0,703	0,807	1,261	-0,321	3,221
Log14	0,679	0,796	1,545	-0,419	4,423
Log15	0,655	0,784	1,895	-0,547	5,912
Log16	0,631	0,771	2,333	-0,715	7,928
Log17	0,608	0,759	2,880	-0,935	10,665
Log18	0,585	0,747	3,566	-1,225	14,389
Log19	0,562	0,736	4,433	-1,610	19,467
Log20	0,540	0,725	5,529	-2,119	26,402
Log21	0,520	0,714	6,922	-2,797	35,889
Log22	0,500	0,704	8,698	-3,699	48,886
Log23	0,481	0,694	10,971	-4,904	66,715
Log24	0,480	0,656	14,027	-6,578	91,082
Log25	0,446	0,675	17,655	-8,672	124,868
Log26	0,429	0,667	22,520	-11,567	171,204
Log27	0,414	0,659	28,825	-15,457	235,040

A comparação entre valores estimados e valores reais só é possível pois há os conjuntos completos ou exaustivos. Assim, os resultados desta comparação encontram-se nas Tabelas 4.4 e 4.5. Além disso, os diagramas de dispersão entre os valores reais versus valores estimados e corrigidos encontram-se nos Anexos V e VI respectivamente. Nas Tabelas 4.4 e 4.5 pode-se observar que até o conjunto Log15 as estimativas produzidas pela KO Corrigida apresentam maiores coeficientes de correlação com os dados exaustivos em relação aos valores sem correção, denunciando uma maior precisão local por parte destas estimativas em relação às não corrigidas. Também há um grande aumento nos erros absoluto médio, médio e RMS (*Root Mean Square*) para as amostras de maior coeficiente de

variação. Novamente aqui percebe-se que os valores estimados pela KO corrigida do efeito de suavização apresentam melhora significativa com relação aos valores estimados sem a correção apenas para conjuntos de dados de valores baixos para os coeficientes de variação, sendo a diferença desprezível para conjuntos com maiores coeficientes de variação, como a partir de Log14. A variação dos coeficientes de correlação é representada na Figura 4.1, sendo perceptível uma progressiva diminuição dos coeficientes de correlação de acordo com o aumento do coeficiente de variação das amostras (Log1 possui o menor coeficiente de variação, que aumenta gradualmente até Log27, que possui o maior coeficiente). Na Figura 4.2 é possível analisar os erros RMS (*Root Mean Square*) das estimativas, e o seu aumento progressivo conforme a amostra usada, provavelmente decorrente da alta variância dos últimos conjuntos de dados.

Os diagramas de dispersão da KO (Anexo V) mostram claramente o enviesamento condicional aumentando com o coeficiente de variação. Observa-se que a comparação é feita com os dados reais, e não com as amostras. Por exemplo, a média populacional no caso da variável Log27 é de 30,368 (Tabela 3.1), a média amostral é 14,804 (Tabela 3.2) e a média calculada pela KO é 14,912 (Tabela 4.1). Assim a comparação pode ser feita, tendo essa ressalva em mente. Quando se compara os diagramas de dispersão do Anexo VI, ou seja, os dados exaustivos em relação às estimativas corrigidas, verifica-se que para as variáveis iniciais (Log1 a Log9) o enviesamento condicional verificado anteriormente é corrigido. No entanto, para as variáveis superiores a Log10, além do enviesamento observa-se o efeito mencionado anteriormente em relação à cauda inferior, ou seja, a grande quantidade de valores iguais a z_{min} .

Tabela 4.5. Coeficientes de Correlação e Erros da Krigagem Ordinária Corrigida do Efeito de Suavização em relação aos Dados Exaustivos

Conjunto	Coeficiente Correlação	Correlação Rankeada	Erro Absoluto Médio	Erro Médio	Erro RMS
Log1	0,881	0,904	0,035	-0,004	0,048
Log2	0,881	0,898	0,071	-0,010	0,098
Log3	0,884	0,899	0,108	-0,018	0,151
Log4	0,882	0,900	0,148	-0,027	0,213
Log5	0,875	0,899	0,194	-0,039	0,292
Log6	0,865	0,897	0,248	-0,054	0,394
Log7	0,851	0,893	0,313	-0,073	0,529
Log8	0,834	0,889	0,392	-0,096	0,708
Log9	0,814	0,884	0,489	-0,126	0,947
Log10	0,794	0,884	0,600	-0,164	1,261
Log11	0,767	0,873	0,754	-0,210	1,699
Log12	0,741	0,866	0,936	-0,269	2,281
Log13	0,713	0,858	1,161	-0,343	3,068
Log14	0,685	0,848	1,444	-0,437	4,135
Log15	0,655	0,834	1,798	-0,554	5,583
Log16	0,626	0,819	2,239	-0,703	7,553
Log17	0,597	0,798	2,791	-0,892	10,237
Log18	0,568	0,771	3,484	-1,135	13,900
Log19	0,541	0,743	4,358	-1,453	18,905
Log20	0,514	0,703	5,469	-1,873	25,754
Log21	0,489	0,666	6,884	-2,428	35,141
Log22	0,464	0,635	8,689	-3,169	48,021
Log23	0,441	0,614	11,001	-4,164	65,712
Log24	0,427	0,582	13,911	-5,519	89,788
Log25	0,399	0,558	17,811	-7,306	123,506
Log26	0,380	0,542	22,776	-9,754	169,617
Log27	0,362	0,522	29,229	-13,072	233,187

Figura 4.1. Comparação da variação dos coeficientes de correlação para dados estimados pela KO e pela KO Corrigida para diferentes conjuntos de dados.

Figura 4.2. Comparação da variação dos erros RMS para dados estimados pela KO e pela KO Corrigida para diferentes conjuntos de dados.

São também comparados os variogramas dos diversos dados: o experimental a partir das amostras, o modelo adotado no MODVARG, o variograma das estimativas da KO, e o variograma das estimativas da KO corrigidas. Os resultados são visualizados no Anexo XIII – Variogramas Comparados para a Krigagem Ordinária, sendo claramente observado que a correção do efeito de suavização faz com que os valores estimados tenham uma variância muito mais próxima da variância amostral, assim como da variância da população, que os valores estimados pela KO sem a mesma correção. O resultado obtido é bastante constante

e satisfatório para as amostras de menor coeficiente de variação (caso das primeiras amostras), principalmente considerando até Log18. Para aquelas com maior coeficiente estes resultados já são mais destoantes do modelo adotado, talvez com exceção das amostras Log23, Log24 e Log26. Tais observações refletem uma boa qualidade da estimativa corrigida no sentido da precisão global, com a reprodução de uma das características mais importantes, que seria o modelo de variância espacial da população em estudo.

4.2 Estimativas da Krigagem Ordinária Lognormal: Transformadas Reversas de Journel e Yamamoto

Análise semelhante à feita para as estimativas da Krigagem Ordinária foi feita para aquelas obtidas a partir da krigagem ordinária lognormal (KOL). Neste caso levamos em consideração para cada um dos 27 conjuntos de dados, duas estimativas distintas: krigagem ordinária lognormal com a transformada reversa pelo método de Journel (1980) e krigagem ordinária lognormal com a transformada reversa pelo método de Yamamoto (2007). As imagens correspondentes às estimativas obtidas pelas duas metodologias se encontram respectivamente nos Anexos VII e VIII. Nos Anexos IX e X são apresentados gráficos acumulativos e gráficos P-P para as estimativas obtidas pela krigagem ordinária lognormal, sendo que nos gráficos acumulativos estão representadas as distribuições de frequência das estimativas obtidas por meio da KOL (Journel) e pela KOL (Yamamoto) comparados com as distribuições amostrais e exaustivas respectivamente. Aqui a krigagem ordinária lognormal transformada pelo método de Journel apresenta claro efeito de suavização, este bastante atenuado quando se estima pela krigagem ordinária transformada pelo método de Yamamoto (2007). Além disso, as distâncias no diagrama P-P (Tabela 4.6) confirmam a melhor reprodução dos histogramas exaustivo e amostral pelo método de transformação reversa de Yamamoto (2007). Todas as distâncias obtidas pelo método de Journel (1980) foram muito superiores a 1,75 e, portanto, resultando em distribuições de frequência muito diferentes em relação às respectivas distribuições exaustivas e amostrais.

Tabela 4.6. Distâncias das Estimativas da Krigagem Ordinária Lognormal e da Krigagem Ordinária Lognormal Corrigida pelos Métodos de Journal (1980) e Yamamoto (2007) com os Dados Exaustivos e Amostras

		Conjunto Exaustivo			Conjunto Amostral		
	Exaustivo x Amostra	Estimativa KOL	Estimativa Corrigida Journal	Estimativa Corrigida Yamamoto	Estimativa KOL	Estimativa Corrigida Journal	Estimativa Corrigida Yamamoto
Log1	1,75	5,94	5,97	1,02	4,57	4,61	1,15
Log2	1,75	5,62	5,68	1,03	4,28	4,36	1,07
Log3	1,75	5,62	5,75	1,03	4,28	4,46	1,07
Log4	1,75	5,62	5,82	1,03	4,28	4,62	1,07
Log5	1,75	5,62	5,92	1,03	4,28	4,84	1,07
Log6	1,75	5,62	6,02	1,03	4,28	5,16	1,07
Log7	1,75	5,62	6,15	1,03	4,28	5,50	1,07
Log8	1,75	5,62	6,30	1,03	4,28	5,84	1,07
Log9	1,75	5,62	6,46	1,03	4,28	6,16	1,07
Log10	1,75	5,62	6,64	1,03	4,27	6,48	1,07
Log11	1,75	5,62	6,82	1,03	4,27	6,78	1,07
Log12	1,75	5,62	7,05	1,03	4,27	7,13	1,07
Log13	1,75	5,62	7,29	1,03	4,27	7,46	1,07
Log14	1,75	5,62	7,51	1,03	4,27	7,79	1,07
Log15	1,75	5,62	7,79	1,03	4,27	8,22	1,07
Log16	1,75	5,62	8,06	1,03	4,27	8,62	1,07
Log17	1,75	5,62	8,31	1,03	4,27	9,03	1,07
Log18	1,75	5,62	8,65	1,03	4,27	9,46	1,06
Log19	1,75	5,62	8,98	1,03	4,27	9,90	1,06
Log20	1,74	5,62	9,24	1,03	4,27	10,23	1,06
Log21	1,74	5,62	9,62	1,03	4,27	10,70	1,06
Log22	1,74	5,62	9,90	1,03	4,27	11,07	1,06
Log23	1,74	5,62	10,24	1,03	4,27	11,52	1,06
Log24	1,74	5,62	10,57	1,03	4,27	11,94	1,06
Log25	1,74	5,62	10,92	1,03	4,27	12,37	1,06
Log26	1,74	5,62	11,26	1,03	4,27	12,78	1,06
Log27	1,74	5,62	11,62	1,03	4,27	13,20	1,06

As estatísticas descritivas apresentadas nas Tabelas 4.7 e 4.8, respectivamente para valores após transformação reversa de Journal (1980) e Yamamoto (2007), confirmam a superioridade deste último método em relação ao primeiro. A análise parte da verificação da coluna da mediana, que no caso de Journal (1980) aumenta com o coeficiente de variação (Log1 a Log27), enquanto a mediana para os valores obtidos pelo método de Yamamoto (2007) é sempre próximo de 1, condizente com os conjuntos amostrais e exaustivos.

Tabela 4.7. Estatísticas Básicas das Estimativas da Krigagem Ordinária Lognormal Corrigida pelo método de Journal (1980)

Conjunto	Nº Dados	Média	Desvio Padrão	Coef. Var.	Máx	Quartil Superior	Mediana	Quartil Inferior	Mín
Log1	2500	1,002	0,058	0,058	1,224	1,039	1,003	0,960	0,788
Log2	2500	1,011	0,122	0,120	1,505	1,086	1,007	0,923	0,624
Log3	2500	1,027	0,186	0,181	1,848	1,139	1,017	0,893	0,494
Log4	2500	1,051	0,255	0,243	2,271	1,198	1,031	0,865	0,391
Log5	2500	1,084	0,331	0,305	2,794	1,266	1,050	0,844	0,310
Log6	2500	1,125	0,415	0,369	3,439	1,343	1,072	0,825	0,245
Log7	2500	1,176	0,511	0,434	4,236	1,431	1,099	0,810	0,195
Log8	2500	1,239	0,621	0,501	5,223	1,529	1,131	0,800	0,155
Log9	2500	1,313	0,748	0,570	6,444	1,638	1,169	0,791	0,123
Log10	2500	1,402	0,898	0,640	7,957	1,767	1,213	0,787	0,098
Log11	2500	1,507	1,073	0,712	9,832	1,909	1,262	0,783	0,078
Log12	2500	1,633	1,284	0,786	12,162	2,076	1,321	0,786	0,062
Log13	2500	1,781	1,536	0,862	15,055	2,267	1,380	0,789	0,049
Log14	2500	1,946	1,832	0,941	18,632	2,471	1,450	0,791	0,039
Log15	2500	2,150	2,196	1,021	23,101	2,720	1,535	0,804	0,031
Log16	2500	2,386	2,633	1,103	28,650	2,999	1,634	0,816	0,025
Log17	2500	2,663	3,162	1,187	35,554	3,306	1,731	0,825	0,020
Log18	2500	3,004	3,819	1,271	44,199	3,669	1,856	0,849	0,016
Log19	2500	3,411	4,624	1,356	54,996	4,103	2,004	0,878	0,013
Log20	2500	3,865	5,583	1,444	68,360	4,569	2,151	0,903	0,010
Log21	2500	4,448	6,804	1,530	85,221	5,135	2,354	0,936	0,008
Log22	2500	5,101	8,261	1,619	106,103	5,754	2,536	0,968	0,007
Log23	2500	5,914	10,092	1,707	132,353	6,496	2,775	1,002	0,005
Log24	2500	6,878	12,344	1,795	165,144	7,351	3,028	1,052	0,004
Log25	2500	8,043	15,139	1,882	206,212	8,351	3,320	1,101	0,003
Log26	2500	9,478	18,648	1,968	257,828	9,615	3,683	1,162	0,003
Log27	2500	11,220	23,023	2,052	322,570	11,055	4,050	1,234	0,002

Tabela 4.8. Estatísticas Básicas das Estimativas pela Krigagem Ordinária Lognormal Corrigida pelo método de Yamamoto (2007)

Conjunto	Nº Dados	Média	D. Padrão	Coef. Var.	Máx	Quartil Superior	Mediana	Quartil Inferior	Mín
Log1	2500	1,002	0,093	0,093	1,240	1,059	0,999	0,937	0,772
Log2	2500	1,010	0,187	0,185	1,534	1,123	0,996	0,877	0,596
Log3	2500	1,027	0,287	0,279	1,900	1,190	0,994	0,821	0,460
Log4	2500	1,054	0,395	0,375	2,353	1,262	0,992	0,769	0,355
Log5	2500	1,091	0,516	0,473	2,915	1,337	0,991	0,720	0,274
Log6	2500	1,138	0,653	0,574	3,610	1,417	0,989	0,675	0,211
Log7	2500	1,196	0,812	0,679	4,471	1,502	0,987	0,632	0,163
Log8	2500	1,268	0,999	0,788	5,538	1,592	0,985	0,592	0,126
Log9	2500	1,354	1,219	0,901	6,859	1,687	0,983	0,554	0,097
Log10	2500	1,456	1,482	1,018	8,496	1,788	0,981	0,519	0,075
Log11	2500	1,578	1,798	1,139	10,523	1,895	0,979	0,486	0,058
Log12	2500	1,722	2,178	1,265	13,033	2,008	0,977	0,455	0,045
Log13	2500	1,891	2,637	1,394	16,142	2,128	0,976	0,426	0,034
Log14	2500	2,091	3,193	1,527	19,994	2,256	0,974	0,399	0,027
Log15	2500	2,327	3,868	1,662	24,763	2,391	0,972	0,374	0,021
Log16	2500	2,605	4,690	1,801	30,671	2,534	0,970	0,350	0,016
Log17	2500	2,932	5,691	1,941	37,989	2,685	0,968	0,328	0,012
Log18	2500	3,319	6,912	2,082	47,051	2,846	0,966	0,307	0,009
Log19	2500	3,776	8,402	2,225	58,277	3,016	0,965	0,288	0,007
Log20	2500	4,318	10,225	2,368	72,180	3,197	0,963	0,269	0,006
Log21	2500	4,960	12,454	2,511	89,400	3,388	0,961	0,252	0,004
Log22	2500	5,721	15,183	2,654	110,728	3,590	0,959	0,236	0,003
Log23	2500	6,627	18,526	2,795	137,144	3,805	0,957	0,221	0,003
Log24	2500	7,706	22,624	2,936	169,863	4,033	0,955	0,207	0,002
Log25	2500	8,992	27,650	3,075	210,387	4,274	0,954	0,194	0,002
Log26	2500	10,529	33,818	3,212	260,579	4,530	0,952	0,182	0,001
Log27	2500	12,367	41,391	3,347	322,745	4,801	0,950	0,170	0,001

A correlação entre as estimativas pela KOL e suas correções com os valores exaustivos, assim como os erros apresentados também foi feita. Esta se encontra na forma de gráficos de dispersão nos Anexos XI e XII, com seus valores apresentados nas Tabelas 4.9 e 4.10. Ilustrativamente a variação dos coeficientes de correlação e dos erros RMS de acordo com os conjuntos de dados são apresentadas nas Figuras 4.3 e 4.4.

Comparando-se os diagramas de dispersão dos Anexos XI e XII verifica-se que os resultados dos dados transformados pelo método de Yamamoto (2007) não estão

enviesados. As estatísticas comparativas com os dados exaustivos (Tabelas 4.9 e 4.10) também confirmam o melhor desempenho do método de Yamamoto (2007) em relação ao de Journal (1980).

A Figura 4.3 mostra correlações maiores com os dados exaustivos para o método de Yamamoto (2007) e menores para aqueles obtidos segundo Journal (1980). Com relação aos erros RMS, há uma leve tendência do método de Yamamoto (2007) apresentar erros menores em relação a Journal (1980).

Tabela 4.9. Coeficientes de Correlação e Erros da Krigagem Ordinária Lognormal Corrigida pelo Método de Journal (1980) em relação aos Dados Exaustivos

Conjunto	Coeficiente Correlação	Correlação Rankeada	Erro Absoluto Médio	Erro Médio	Erro RMS
Log1	0,848	0,850	0,044	-0,003	0,060
Log2	0,862	0,865	0,085	-0,009	0,118
Log3	0,861	0,864	0,131	-0,018	0,186
Log4	0,857	0,864	0,180	-0,032	0,267
Log5	0,851	0,864	0,234	-0,048	0,365
Log6	0,843	0,864	0,295	-0,071	0,488
Log7	0,832	0,863	0,364	-0,099	0,646
Log8	0,819	0,863	0,445	-0,134	0,852
Log9	0,804	0,862	0,540	-0,180	1,122
Log10	0,787	0,861	0,652	-0,238	1,479
Log11	0,767	0,861	0,787	-0,311	1,954
Log12	0,746	0,860	0,951	-0,402	2,588
Log13	0,723	0,859	1,150	-0,518	3,348
Log14	0,699	0,858	1,394	-0,675	4,585
Log15	0,673	0,857	1,697	-0,865	6,127
Log16	0,648	0,856	2,073	-1,114	8,215
Log17	0,621	0,855	2,544	-1,436	11,045
Log18	0,594	0,854	3,138	-1,837	14,881
Log19	0,568	0,852	3,890	-2,354	20,099
Log20	0,542	0,851	4,844	-3,054	27,225
Log21	0,516	0,849	6,063	-3,922	36,931
Log22	0,491	0,848	7,623	-5,099	50,224
Log23	0,467	0,847	9,634	-6,608	68,407
Log24	0,444	0,845	12,233	-8,598	93,339
Log25	0,421	0,844	15,605	-11,212	127,559
Log26	0,399	0,842	20,001	-14,632	174,570
Log27	0,379	0,840	25,735	-19,148	239,238

No Anexo XIV são apresentados os variogramas obtidos com a Krigagem Ordinária Lognormal, comparando os resultados da Krigagem Ordinária Lognormal com a transformada reversa pelo método de Journel (1980), a Krigagem Ordinária Lognormal com a transformada reversa pelo método de Yamamoto (2007), o variograma experimental a partir das amostras e o modelo adotado. No caso das curvas correspondentes à KOL (Yamamoto) percebe-se a forte uniformidade dos dados, principalmente da coerência das estimativas obtidas pelo método de correção com relação ao modelo adotado, em relação à KO, à KO Corrigida e à KOL (Journel). A reprodução da variância amostral é bem corrigida para todas as amostras, sem perda de consistência nos dados com o aumento do coeficiente de variação das amostras. Este fato contrasta com o comportamento errático apresentado pelo variograma das estimativas da krigagem ordinária corrigida do efeito de suavização para as amostras de maior coeficiente de variação (evidente de Log18 em diante) conforme observado no Anexo XIII.

Tabela 4.10. Coeficientes de Correlação e Erros da Krigagem Ordinária Lognormal Corrigida pelo Método de Yamamoto (2007) em relação aos Dados Exaustivos

Conjunto	Coeficiente Correlação	Correlação Rankeada	Erro Aboluto Médio	Erro Médio	Erro RMS
Log1	0,878	0,901	0,035	-0,004	0,049
Log2	0,875	0,891	0,072	-0,010	0,100
Log3	0,878	0,891	0,109	-0,018	0,154
Log4	0,877	0,891	0,149	-0,028	0,217
Log5	0,873	0,891	0,192	-0,041	0,295
Log6	0,867	0,891	0,241	-0,058	0,393
Log7	0,857	0,891	0,297	-0,079	0,521
Log8	0,846	0,891	0,363	-0,106	0,690
Log9	0,832	0,891	0,442	-0,140	0,916
Log10	0,816	0,891	0,537	-0,184	1,218
Log11	0,799	0,891	0,652	-0,240	1,625
Log12	0,781	0,891	0,793	-0,313	2,174
Log13	0,761	0,891	0,967	-0,408	2,920
Log14	0,741	0,891	1,184	-0,530	3,934
Log15	0,720	0,891	1,455	-0,689	5,316
Log16	0,699	0,891	1,796	-0,896	7,203
Log17	0,679	0,891	2,226	-1,167	9,784
Log18	0,658	0,891	2,771	-1,522	13,319
Log19	0,638	0,891	3,465	-1,988	18,168
Log20	0,618	0,891	4,352	-2,601	24,825
Log21	0,600	0,891	5,490	-3,410	33,977
Log22	0,581	0,891	6,955	-4,479	46,567
Log23	0,564	0,891	8,847	-5,894	63,902
Log24	0,547	0,891	11,300	-7,771	87,791
Log25	0,532	0,891	14,489	-10,263	120,733
Log26	0,517	0,891	18,648	-13,581	166,192
Log27	0,503	0,891	24,088	-18,002	228,964

Figura 4.3. Comparação da variação dos coeficientes de correlação para dados estimados pela KOL Corrigida pelos métodos de Journal e Yamamoto para os diferentes conjuntos de dados.

Figura 4.4. Comparação da variação dos erros RMS para dados estimados pela KOL e pela KOL Corrigida pelos métodos de Journal e Yamamoto para os diferentes conjuntos de dados.

4.3 Realizações da Simulação Estocástica

Por meio dos métodos de simulação estocástica: gaussiana sequencial e por bandas rotativas, foram geradas realizações para cada um dos 27 conjuntos de dados, totalizando 5400 realizações. Como as amostras foram extraídas de conjuntos completos e as realizações foram feitas sobre as mesmas malhas regulares dos conjuntos completos, é possível comparar os resultados das simulações estocásticas com os valores reais. Essa comparação foi feita através dos seguintes parâmetros estatísticos: coeficiente de correlação, erro absoluto médio, erro médio e erro RMS. Tendo em vista as 100 realizações para cada método de simulação, as estatísticas mencionadas são apresentadas em termos de seus valores médios, bem como os valores máximos e mínimos encontrados.

Os resultados comparativos encontram-se nas Tabelas 4.11 a 4.14, respectivamente coeficiente de correlação, erro absoluto médio, erro médio e erro RMS. Esses resultados são também representados graficamente nas Figuras 4.5 a 4.8 na mesma sequência.

Tabela 4.11. Médias, valores máximos e mínimos dos Coeficientes de Correlação das Realizações em relação aos Dados Exaustivos (100 para cada conjunto de dados) da Simulação Gaussiana Sequencial e da Simulação por Bandas Rotativas

Conjunto	Simulação Gaussiana Sequencial				Simulação por Bandas Rotativas			
	Realizações	Média	Máx	Mín	Realizações	Média	Máx	Mín
Log1	100	0,655	0,741	0,563	100	0,629	0,706	0,491
Log2	100	0,654	0,714	0,570	100	0,628	0,750	0,496
Log3	100	0,649	0,728	0,497	100	0,623	0,720	0,491
Log4	100	0,635	0,717	0,518	100	0,609	0,700	0,495
Log5	100	0,623	0,724	0,483	100	0,594	0,709	0,352
Log6	100	0,614	0,708	0,470	100	0,584	0,710	0,463
Log7	100	0,594	0,696	0,485	100	0,564	0,696	0,404
Log8	100	0,575	0,683	0,469	100	0,543	0,693	0,381
Log9	100	0,555	0,701	0,460	100	0,525	0,695	0,392
Log10	100	0,535	0,682	0,360	100	0,490	0,640	0,359
Log11	100	0,504	0,628	0,346	100	0,476	0,675	0,299
Log12	100	0,476	0,643	0,290	100	0,447	0,657	0,309
Log13	100	0,465	0,620	0,321	100	0,429	0,643	0,268
Log14	100	0,423	0,577	0,288	100	0,398	0,560	0,221
Log15	100	0,414	0,626	0,248	100	0,383	0,614	0,211
Log16	100	0,376	0,573	0,180	100	0,356	0,560	0,202
Log17	100	0,365	0,544	0,215	100	0,335	0,595	0,153
Log18	100	0,343	0,559	0,165	100	0,326	0,520	0,155
Log19	100	0,338	0,555	0,109	100	0,292	0,475	0,109
Log20	100	0,321	0,514	0,118	100	0,270	0,491	0,072
Log21	100	0,293	0,448	0,089	100	0,271	0,540	0,060
Log22	100	0,283	0,563	0,099	100	0,246	0,493	0,043
Log23	100	0,260	0,518	0,085	100	0,246	0,500	0,047
Log24	100	0,240	0,485	0,098	100	0,211	0,498	0,036
Log25	100	0,237	0,442	0,045	100	0,190	0,485	0,036
Log26	100	0,212	0,547	0,050	100	0,194	0,407	0,025
Log27	100	0,227	0,448	0,042	100	0,194	0,465	0,024

Tabela 4.12. Médias, valores máximos e mínimos dos Erros Absolutos Médios das Realizações em relação aos Dados Exaustivos (100 para cada conjunto de dados) da Simulação Gaussiana Sequencial e da Simulação por Bandas Rotativas

Conjunto	Simulação Gaussiana Sequencial				Simulação por Bandas Rotativas			
	Realizações	Média	Máx	Mín	Realizações	Média	Máx	Mín
Log1	100	0,061	0,067	0,053	100	0,062	0,070	0,056
Log2	100	0,122	0,134	0,114	100	0,126	0,145	0,111
Log3	100	0,187	0,219	0,166	100	0,193	0,223	0,171
Log4	100	0,257	0,288	0,231	100	0,266	0,294	0,237
Log5	100	0,333	0,371	0,302	100	0,344	0,405	0,293
Log6	100	0,417	0,473	0,380	100	0,431	0,503	0,370
Log7	100	0,516	0,570	0,466	100	0,535	0,604	0,480
Log8	100	0,632	0,691	0,561	100	0,654	0,780	0,582
Log9	100	0,764	0,848	0,666	100	0,785	0,877	0,669
Log10	100	0,915	1,056	0,809	100	0,963	1,098	0,803
Log11	100	1,112	1,271	0,988	100	1,142	1,329	0,986
Log12	100	1,333	1,533	1,131	100	1,382	1,629	1,115
Log13	100	1,603	1,782	1,404	100	1,651	1,982	1,425
Log14	100	1,944	2,148	1,752	100	2,005	2,363	1,735
Log15	100	2,330	2,667	1,985	100	2,418	2,784	2,099
Log16	100	2,878	3,446	2,399	100	2,957	3,421	2,599
Log17	100	3,496	4,104	3,108	100	3,592	4,065	3,107
Log18	100	4,278	5,192	3,894	100	4,386	5,082	3,866
Log19	100	5,170	6,293	4,638	100	5,409	6,303	4,735
Log20	100	6,394	7,174	5,670	100	6,661	7,891	5,777
Log21	100	7,922	9,135	7,136	100	8,238	9,811	6,807
Log22	100	9,890	11,234	8,417	100	10,229	12,001	8,559
Log23	100	12,365	14,840	11,028	100	12,566	15,120	10,460
Log24	100	15,546	17,602	14,006	100	16,115	18,506	13,800
Log25	100	19,449	22,763	17,454	100	20,099	24,005	17,403
Log26	100	24,824	29,203	20,584	100	25,647	29,781	22,127
Log27	100	30,874	36,172	27,213	100	32,057	38,558	27,556

Tabela 4.13. Médias, valores máximos e mínimos dos Erros Médios das Realizações em relação aos Dados Exaustivos (100 para cada conjunto de dados) da Simulação Gaussiana Sequencial e da Simulação por Bandas Rotativas

Conjunto	Simulação Gaussiana Sequencial				Simulação por Bandas Rotativas			
	Realizações	Média	Máx	Mín	Realizações	Média	Máx	Mín
Log1	100	-0,003	0,010	-0,019	100	-0,003	0,011	-0,016
Log2	100	-0,007	0,020	-0,026	100	-0,008	0,012	-0,036
Log3	100	-0,017	0,037	-0,058	100	-0,017	0,018	-0,054
Log4	100	-0,030	0,020	-0,070	100	-0,023	0,022	-0,067
Log5	100	-0,046	0,040	-0,109	100	-0,044	0,003	-0,114
Log6	100	-0,068	0,010	-0,164	100	-0,059	0,005	-0,150
Log7	100	-0,087	0,007	-0,180	100	-0,077	0,038	-0,176
Log8	100	-0,109	0,015	-0,261	100	-0,097	0,014	-0,211
Log9	100	-0,146	-0,023	-0,284	100	-0,138	-0,024	-0,310
Log10	100	-0,203	0,041	-0,458	100	-0,165	0,038	-0,331
Log11	100	-0,269	-0,045	-0,469	100	-0,246	-0,006	-0,482
Log12	100	-0,352	0,061	-0,602	100	-0,344	-0,047	-0,666
Log13	100	-0,452	-0,178	-0,722	100	-0,435	0,065	-0,718
Log14	100	-0,591	-0,123	-0,995	100	-0,536	0,009	-0,969
Log15	100	-0,757	-0,274	-1,229	100	-0,697	-0,163	-1,339
Log16	100	-0,971	-0,084	-1,521	100	-0,875	-0,289	-1,380
Log17	100	-1,197	-0,108	-1,756	100	-1,161	-0,087	-1,899
Log18	100	-1,606	-0,536	-2,407	100	-1,423	-0,360	-2,311
Log19	100	-1,969	-1,064	-2,980	100	-1,895	-0,816	-3,133
Log20	100	-2,666	-1,411	-4,056	100	-2,490	-0,622	-3,947
Log21	100	-3,533	-1,970	-5,034	100	-3,201	-1,158	-4,866
Log22	100	-4,415	-2,156	-5,977	100	-4,312	-2,019	-6,122
Log23	100	-5,974	-3,228	-8,043	100	-5,769	-2,802	-8,150
Log24	100	-7,656	-4,142	-11,685	100	-7,264	-3,308	-10,554
Log25	100	-10,309	-6,492	-13,762	100	-9,791	-5,236	-13,062
Log26	100	-13,370	-7,770	-17,832	100	-12,436	-6,316	-16,954
Log27	100	-17,907	-12,244	-22,757	100	-16,888	-10,954	-23,225

Tabela 4.14. Médias, valores máximos e mínimos dos Erros RMS das Realizações em relação aos Dados Exaustivos (100 para cada conjunto de dados) da Simulação Gaussiana Sequencial e da Simulação por Bandas Rotativas

Conjunto	Simulação Gaussiana Sequencial				Simulação por Bandas Rotativas			
	Realizações	Média	Máx	Mín	Realizações	Média	Máx	Mín
Log1	100	0,079	0,089	0,069	100	0,082	0,095	0,074
Log2	100	0,162	0,180	0,150	100	0,169	0,194	0,146
Log3	100	0,253	0,296	0,225	100	0,263	0,304	0,234
Log4	100	0,360	0,406	0,318	100	0,377	0,420	0,332
Log5	100	0,487	0,555	0,432	100	0,506	0,617	0,435
Log6	100	0,637	0,742	0,574	100	0,662	0,745	0,568
Log7	100	0,831	0,922	0,730	100	0,862	0,976	0,740
Log8	100	1,077	1,194	0,945	100	1,119	1,321	0,932
Log9	100	1,386	1,540	1,181	100	1,436	1,615	1,180
Log10	100	1,784	2,101	1,519	100	1,877	2,118	1,614
Log11	100	2,331	2,650	2,066	100	2,396	2,759	1,961
Log12	100	3,034	3,446	2,605	100	3,117	3,483	2,555
Log13	100	3,920	4,307	3,438	100	4,033	4,572	3,348
Log14	100	5,191	5,621	4,622	100	5,300	5,871	4,696
Log15	100	6,763	7,547	5,729	100	6,913	7,611	5,880
Log16	100	9,026	9,903	8,000	100	9,139	10,034	7,985
Log17	100	11,896	13,136	10,737	100	12,090	13,155	10,229
Log18	100	15,829	17,208	13,988	100	15,984	17,288	14,324
Log19	100	21,054	22,752	18,911	100	21,492	23,398	19,613
Log20	100	28,219	30,381	25,745	100	28,839	31,200	26,271
Log21	100	38,180	40,727	35,926	100	38,539	41,434	34,153
Log22	100	51,458	54,623	46,398	100	52,145	56,055	47,941
Log23	100	69,939	73,330	65,540	100	70,247	74,728	64,733
Log24	100	95,180	98,537	89,011	100	95,929	100,508	87,349
Log25	100	129,321	135,571	122,306	100	130,750	135,837	119,956
Log26	100	176,911	183,460	164,231	100	177,556	185,246	167,992
Log27	100	240,567	250,398	229,453	100	242,180	250,363	225,409

Figura 4.5. Variação dos coeficientes de correlação entres as estimativas/média das simulações e os dados exaustivos para as diversas metodologias.

Figura 4.6. Variação dos erros absolutos médios para as diversas metodologias utilizadas.

Figura 4.7. Variação dos erros médios para as diversas metodologias utilizadas.

Figura 4.8. Variação dos erros RMS para as diversas metodologias utilizadas.

De forma geral, as simulações quando comparadas às estimativas apresentam os menores coeficientes de correlação, maiores erros absolutos médios e erros RMS. Isso reflete a falta de precisão local dos métodos de simulação em relação a qualquer um dos métodos de estimativa.

Comparando-se os dois métodos de simulação, a SGS apresenta resultados levemente superiores em relação à TBSim, esta última apresentando normalmente erros maiores e menores coeficientes de correlação em relação às simulações da SGS. Os coeficientes de correlação com os dados exaustivos para a TBSim também eram sistematicamente menores que os da SGS para todos os conjuntos de dados, mantendo uma diferença entre eles semelhante para todos os conjuntos de diferentes coeficientes de variação. Ao mesmo tempo os erros absolutos médios e RMS são maiores para a simulação por bandas rotativas com relação à SGS, sendo que a diferença dos erros tende a aumentar conforme se aumenta a assimetria do conjunto amostral. Apenas o erro médio que é pouco menor em módulo para as realizações da simulação por bandas rotativas em relação às da simulação sequencial gaussiana.

Em termos dos valores máximos e mínimos encontrados no conjunto de 100 realizações verifica-se que a correlação máxima é 0,741 para a SGS e variável Log1, enquanto a correlação mínima é de 0,024 para a TBSim na variável Log27. Observa-se que mesmo entre 100 realizações, a máxima correlação é 0,741, o que não representa uma alta correlação (bastante inferior aos coeficientes obtidos pelas estimativas – valores a partir de 0,848). Então, é muito difícil que esse valor aumente, mesmo que o número de realizações seja bastante elevado, haja vista tratar-se de processo de Monte Carlo.

Os erros obtidos, (absoluto médio, erro médio e erro RMS) de uma maneira geral, possuem uma tendência a erros maiores que aqueles encontrados nas estimativas. No caso específico dos erros RMS (Figura 4.8) estes eram pouco maiores para as simulações gaussianas sequenciais, com erros semelhantes encontrados para as simulações TBSIM. No entanto isso é válido apenas para os conjuntos de dados de pequeno coeficiente de variação, sendo que a partir de determinado valor deste coeficiente estas diferenças de erros tornam-se relativamente pouco significativas. Em compensação, ao se analisar na Figura 4.6 as variações nos erros absolutos médios, observa-se uma tendência de serem

bastante semelhantes para as amostras de baixa variância, com esta diferença aumentando à medida que os erros aumentam exponencialmente de acordo com a variância da amostra. Para os erros médios (Figura 4.7) houve fenômeno semelhante, com as diferenças entre erros aumentando conforme o coeficiente de variação da amostra, porém aqui a simulação gaussiana sequencial e a simulação por bandas rotativas apresentavam erros menores que os da KOL, da KOL (Journel), e erros semelhantes àqueles observados para a KOL (Yamamoto). Os erros médios são negativos para todas as metodologias utilizadas, sendo que a KOL (mera exponenciação de Y_{KO}^*) apresentou os maiores erros em módulo, enquanto a KO corrigida do efeito de suavização apresentou os menores erros em módulo.

Na Tabela 4.15 estão discriminadas as realizações da simulação gaussiana sequencial que apresentaram os valores extremos (máximo e mínimo) para as médias de coeficiente de correlação, erro RMS, erro médio absoluto e erro médio. Observa-se uma boa correspondência para as simulações que apresentam os maiores coeficientes de correlação e os menores erros RMS, caracterizando assim a melhor realização para determinado conjunto de dados. Da mesma forma há uma correspondência entre as simulações com os menores coeficientes de correlação e maiores erros RMS, indicando as piores simulações do ponto de vista de fidelidade ao dado exaustivo. Já as simulações com erros médios e erros médios absolutos extremos não apresentam nenhum tipo de correlação aparente com os demais dados.

Os resultados das realizações individuais, conforme coeficiente de correlação com os dados exaustivos para as 100 realizações de cada conjunto de dados pode ser observada no Anexo XV (Figuras 6.43 a 6.51). As flutuações destes valores indicam a natureza altamente variável das diferentes realizações, verificando-se a aleatoriedade do processo de Monte Carlo. As flutuações são mais intensas conforme o aumento do coeficiente de variação da variável em questão. Também é possível por meio destes gráficos afirmar que os coeficientes de correlação das simulações da SGS são em média inferiores àqueles obtidos por meio das estimativas da KOL com ambas as variações da transformada reversa para todos os conjuntos de dados (Figura 4.5). Porém, também é possível afirmar que conforme o aumento do coeficiente de variação do conjunto de dados (a partir de Log19), há casos em que algumas realizações apresentam coeficientes de correlação com os dados exaustivos

superiores às da KOL, e inclusive às da KOL Corrigida (caso da variável Log22), apesar de serem casos raros e isolados.

Uma avaliação semelhante pode ser feita no caso dos erros RMS (Figuras 6.43 a 6.51), sendo que as realizações da SGS possuem geralmente erros maiores que os da KOL, porém são maiores que os da KOL corrigida. Uma variação dos erros também é observável, sendo possível afirmar uma aproximação dos erros RMS em relação aos erros da KOL com o aumento do coeficiente de variação da variável em questão (sendo evidente a partir de Log20), muitas vezes os erros da simulação sendo menores que os da KOL, porém quase na totalidade das vezes maiores que os da KOL com a transformada reversa de Yamamoto (2007).

No Anexo XVI (figuras 6.52 a 6.60) podem ser observadas as realizações da simulação gaussiana sequencial que apresentaram os maiores e menores coeficientes de correlação em relação aos conjuntos exaustivos para os 27 conjuntos de dados. Pode-se especular que as imagens que apresentam os maiores coeficientes de correlação são aquelas cujos valores apresentam maior continuidade espacial, com alguns agrupamentos representativos de valores correlatos aos dados exaustivos, porém não há nada conclusivo a respeito. Mesmo assim, há tanto para a “melhor” como para a “pior” imagem a preservação de feições comuns, como a localização dos *clusters* de valores extremos assim como a proporção entre os diferentes valores no mapa, o que corresponderia à reprodução do histograma dos dados exaustivos.

Na Tabela 4.16 estão também discriminadas as realizações da simulação por bandas rotativas que apresentaram os extremos (máximo e mínimo) para as médias de coeficiente de correlação, erro RMS, erro médio absoluto e erro médio. Assim como para a simulação gaussiana sequencial é possível afirmar uma boa correspondência para as simulações que apresentam os maiores coeficientes de correlação e os menores erros RMS, caracterizando assim as melhores realizações. Da mesma forma há uma correspondência razoável entre as simulações com os menores coeficientes de correlação e maiores erros RMS, que seriam as piores simulações. Ocasionalmente as simulações com erros médios e erros médios absolutos extremos apresentam algum tipo de correlação aparente com os demais parâmetros (coeficientes de correlação e erros RMS).

Tabela 4.15. Realizações da Simulação Sequencial Gaussiana com Maiores e Menores Parâmetros Estatísticos

Conjunto	Coef. Correlação		RMS		Erro Médio Absoluto		Erro Médio	
	Mínimo	Máximo	Mínimo	Máximo	Mínimo	Máximo	Mínimo	Máximo
Log1	48	19	19	92	19	48	51	36
Log2	91	56	56	91	17	21	19	56
Log3	11	95	95	11	95	11	71	70
Log4	29	82	82	29	82	29	54	95
Log5	56	71	71	56	99	56	49	73
Log6	23	5	5	23	5	23	58	64
Log7	92	99	99	92	99	84	51	31
Log8	4	59	59	85	59	4	18	28
Log9	64	92	92	94	92	82	46	49
Log10	7	28	28	7	28	7	26	29
Log11	52	49	28	52	91	39	86	67
Log12	19	81	81	19	44	68	58	68
Log13	88	15	15	88	12	88	53	42
Log14	28	65	71	49	5	14	56	11
Log15	61	82	82	53	82	53	18	53
Log16	92	74	15	58	51	58	9	58
Log17	92	54	54	81	80	81	44	56
Log18	67	100	100	92	66	92	26	81
Log19	66	83	95	89	35	89	19	16
Log20	68	15	15	68	50	68	56	89
Log21	57	16	16	57	88	47	94	47
Log22	87	76	76	82	76	82	67	61
Log23	64	6	13	92	6	71	69	46
Log24	46	91	91	43	57	6	46	6
Log25	83	2	2	83	75	83	44	52
Log26	12	47	47	42	47	42	22	19
Log27	33	95	95	40	95	40	4	68

Tabela 4.16. Realizações da Simulação Sequencial Gaussiana com Maiores e Menores Parâmetros Estatísticos

Conjunto	Coef. Correlação		RMS		Erro Médio Absoluto		Erro Médio	
	Mínimo	Máximo	Mínimo	Máximo	Mínimo	Máximo	Mínimo	Máximo
Log1	44	43	43	44	92	44	45	20
Log2	10	32	32	96	52	96	36	19
Log3	67	20	20	12	20	67	90	92
Log4	59	99	99	14	70	22	75	44
Log5	48	71	71	48	71	48	75	33
Log6	68	93	93	24	71	97	75	52
Log7	47	48	48	47	92	63	88	69
Log8	37	76	76	60	95	60	67	87
Log9	59	72	72	98	71	98	65	20
Log10	88	37	37	1	37	1	53	17
Log11	87	29	29	87	12	3	58	80
Log12	78	49	49	79	49	13	32	13
Log13	15	32	32	15	79	15	66	48
Log14	99	92	24	89	94	91	87	21
Log15	32	76	76	32	76	14	34	15
Log16	55	73	73	68	21	34	6	46
Log17	24	32	32	59	32	59	24	2
Log18	68	70	70	4	89	48	26	57
Log19	48	43	43	48	76	47	69	13
Log20	63	39	87	38	66	26	8	83
Log21	54	2	62	54	68	54	76	28
Log22	81	61	31	81	15	81	4	23
Log23	58	92	68	56	21	56	3	15
Log24	91	2	2	91	2	12	57	21
Log25	33	2	2	12	2	89	38	89
Log26	79	36	36	79	49	79	11	92
Log27	90	50	50	53	91	75	1	85

A variação dos coeficientes de correlação com os dados exaustivos e os erros RMS para as 100 realizações de cada conjunto de dados são observados no Anexo XVII (Figuras 6.61 a 6.69). Como para a SGS, também há altas flutuações destes valores, que indicam a natureza variável das diferentes realizações, assim como os menores coeficientes de correlação com os dados exaustivos quando comparados contra as estimativas da KOL e da KOL com a transformada reversa por Yamamoto (2007). A partir do conjunto Log20 passam a ser observados algumas realizações com coeficientes de correlação maiores que os da KOL transformada por Journel (1980), porém são quase sempre inferiores aos da KOL com a transformada reversa de Yamamoto (2007) . Conforme observado para a SGS, também há

um aumento na variabilidade dos valores dos coeficientes de correlação conforme o aumento do coeficiente de variação da variável em questão.

Também à semelhança da SGS, os erros RMS das simulações TBSim são menores que os da KOL e da KOL com a transformada reversa de Yamamoto (2007) para as variáveis de pequeno coeficiente de variação (até Log17), e comparáveis aos erros das estimativas pela KOL para as variáveis de maior coeficiente de variação. No entanto as realizações possuem erros RMS superiores àqueles da KOL com a transformada Yamamoto.

Assim como para a simulação gaussiana sequencial, no Anexo XVIII (figuras 6.70 a 6.78) podem ser observadas as realizações da simulação por bandas rotativas que apresentaram os maiores e menores coeficientes de correlação com os 27 conjuntos exaustivos. Aqui também é possível especular que as imagens com os maiores coeficientes de correlação são aquelas cujos valores apresentam maior continuidade espacial, com a presença também de agrupamentos semelhantes aos dados exaustivos. Mesmo assim, há para ambos os extremos de imagens a preservação de feições comuns, como a localização dos *clusters* de valores máximos e mínimos assim como a proporção entre os diferentes valores no mapa, o que também corresponderia à reprodução do histograma dos dados exaustivo, confirmando a precisão global da simulação por bandas rotativas.

4.4 Comparação dos Resultados Obtidos

É possível comparar a qualidade dos dados estimados/simulados por meio dos Gráficos e Tabelas comparativos que demonstram a variação dos parâmetros estatísticos usados (coeficiente de correlação com os dados exaustivos, erro médio absoluto, erro médio e erro RMS). A observação das Figuras 4.5, 4.6, 4.7 e 4.8, comparando os parâmetros estatísticos, permite afirmar que de um modo geral os resultados apresentados pelas estimativas são muito superiores àqueles obtidos pelas diversas realizações. Geralmente os coeficientes de correlação máximos observados para as simulações (seja pela simulação das bandas rotativas, seja pela simulação gaussiana sequencial) eram próximos aos coeficientes obtidos pelos diversos métodos de krigagem, sendo que os coeficientes de correlação médios para as simulações eram muito menores.

Como observado por Souza (2007), os resultados da krigagem ordinária lognormal com a transformada reversa de Yamamoto são melhores para distribuições lognormais, com os melhores resultados (em relação à krigagem ordinária) obtidos com a maior assimetria entre os dados. No presente estudo este método de estimativa apresentou coeficientes melhores para todas as distribuições, desde a normal (Log1) até a mais assimétrica (Log27), provando ser a metodologia mais eficaz para a estimativa de uma realidade qualquer que seja o seu coeficiente de variação.

As estimativas também são superiores às simulações em relação aos erros apresentados, que no caso do erro médio absoluto e do erro RMS eram sempre maiores para as realizações por ambas os métodos de simulação, com erros pouco maiores para a simulação por bandas rotativas. Os erros apresentados pelos métodos são de certa forma bastante semelhantes, com as médias dos erros absolutos médios e dos erros RMS sendo geralmente pouco maiores para as simulações por SGS e por TBSim. A flutuação dos valores destes erros é relativamente pequena, sendo que bastante frequentemente os erros das estimativas se encontram no intervalo entre os extremos dos erros das realizações. De uma maneira geral os erros RMS são minimizados aplicando-se os algoritmos de correção das estimativas (tanto da krigagem ordinária como ambas as metodologias para a krigagem ordinária lognormal), sendo que os melhores resultados para o erro absoluto médio e para o erro RMS (minimização do erro) são obtidos pela KOL com a transformada reversa pelo método de Yamamoto (2007), conforme pode ser observado nas Figuras 4.6 e 4.8.

Figura 4.9. Maiores coeficientes de correlação das simulações SGS e TBSim comparados com o método de estimativa de maiores coeficientes (KOL Corrigida por método de Yamamoto, 2007)

Apenas comparando-se as melhores realizações das simulações SGS e TBSim com as estimativas da KOL (método que apresentou maiores coeficientes de correlação) é possível observar uma gradual melhoria dos resultados da simulação de acordo com o aumento do coeficiente de variação da variável, chegando a valores muito semelhantes a partir de Log22, conforme observado na Figura 4.9. Cabe lembrar porém que estes são casos extremos das simulações, sendo portanto apenas um dentre diversos possíveis cenários apresentados por meio das diferentes realizações. Comparando os métodos de simulação entre si, não há uma clara diferenciação entre as duas metodologias, com resultados variáveis para ambos.

Análise semelhante é feita para as realizações da SGS e TBSim que apresentaram os menores coeficientes de correlação com os dados exaustivos (Figura 4.10), comparando-as com os coeficientes de correlação das estimativas que apresentaram menores coeficientes (KO – para variáveis de pequeno coeficiente de variação; e KO Corrigida – para variáveis de alto coeficiente de variação). É possível observar que há um decréscimo generalizado dos

coeficientes de correlação conforme o aumento do coeficiente de variação das amostras, chegando-se a valores muito próximos de zero no caso das simulações, o que representariam basicamente cenários aleatórios, nada condizentes com os dados exaustivos correspondentes à realidade a ser representada.

Figura 4.10. Menores coeficientes de correlação das simulações SGS e TBSim comparados com os métodos de estimativa de menores coeficientes de correlação resultados (KO e KO Corrigida)

Os erros médios apresentam um comportamento distinto daquele observado para os demais parâmetros. Pela Figura 4.7 é possível discernir erros maiores (em módulo) para as estimativas da KOL. Com uma sensível diferença para os erros médios da KOL vem alguns outros resultados relativamente semelhantes: KOL (Journel) e KOL (Yamamoto), vindo a seguir a média das realizações da SGS, então as da TBSim e então as da krigagem ordinária. Os menores erros médios foram aqueles observados na krigagem ordinária corrigida do efeito de suavização, apresentando os menores erros em termos de módulo para os conjuntos amostrais a partir da amostra Log16 e daí para os demais conjuntos de elevada variância, enquanto a KO apresentava melhores resultados para os dados de menor variância (Log1 até Log15).

CAPÍTULO 5

5. Conclusões

O objetivo deste estudo foi comparar a eficiência de diferentes métodos de se reconstruir uma realidade previamente conhecida utilizando modelos construídos a partir de uma amostragem dos dados exaustivos. Foram duas abordagens para o problema: a geração de estimativas ótimas e a geração de múltiplas realizações de simulações estocásticas. Para o primeiro caso foram usadas duas metodologias, a krigagem ordinária e a krigagem ordinária lognormal, cada uma submetida a correções e transformações visando a minimização do efeito de suavização intrínseco a ambos os métodos. Já as simulações foram geradas por meio de dois algoritmos, a simulação gaussiana sequencial e a simulação por bandas rotativas.

De maneira geral as estimativas obtidas por meio da krigagem ordinária e da krigagem ordinária lognormal, por sua própria natureza de minimizarem a variância das estimativas, foram os métodos que apresentaram maior fidelidade aos dados exaustivos em termos de apresentarem maiores coeficientes de correlação e minimização dos erros. Para estes dois métodos as correções apresentadas para o efeito de suavização foram todas benéficas para as estimativas, melhorando os seus coeficientes de correlação e minimizando seus erros. No caso específico da krigagem ordinária lognormal a correção proposta por Journel (1980) apresentou pouca melhora em relação às estimativas não corrigidas, ao contrário daquelas corrigidas pelo método de Yamamoto (2007) que apresentaram sensível melhora nos resultados.

Pelo lado das realizações das simulações estocásticas foram geradas 100 realizações para cada conjunto de dados, conforme os métodos da simulação gaussiana sequencial e por simulação por bandas rotativas. Apesar de apresentarem resultados muito semelhantes entre si, a simulação gaussiana sequencial apresentou resultados médios um pouco superiores em relação aos resultados médios da simulação por bandas rotativas em termos de coeficientes de correlação e minimização de erros. No entanto, ambas apresentaram resultados médios dos parâmetros estatísticos avaliados bastante inferiores a quaisquer

daqueles obtidos utilizando-se dos métodos de estimativas. A confiabilidade dos resultados obtidos nas simulações, a maneira dos resultados das estimativas, também decresce de acordo com a variância das distribuições de dados, com a diminuição dos coeficientes de correlação e aumento exponencial dos erros.

Os resultados obtidos corroboram a maior eficácia dos métodos de estimativa como métodos de maior precisão local e a dos métodos de simulação como de maior precisão global. Em casos onde seja necessário se obter um valor relativamente confiável para determinado ponto no espaço (tal como no planejamento de lavra) as estimativas por meio da KO ou da KOL são recomendadas por minimizarem os erros associados a esta estimativa. Os respectivos algoritmos de correção de ambos os métodos apenas contribuem para um incremento na confiabilidade destas estimativas com relação à realidade por meio da melhor representação dos valores extremos.

Por outro lado, as realizações obtidas por meio das simulações possuem qualidades intrínsecas que as tornam preferíveis em alguns casos com relação às estimativas. O fato de gerarem múltiplas realizações possibilita que sejam considerados múltiplos cenários possíveis para determinado contexto, o que é bastante importante na hora de se avaliar incertezas. Outro aspecto positivo das simulações, por mais que percam muito a sua precisão local, é o fato de se trabalhar com realizações com uma variância correspondente à realidade, permitindo muitas vezes uma modelagem melhor das heterogeneidades existentes na natureza e permitindo a construção de modelos geológicos mais robustos. Tais propriedades são desejados por exemplo na modelagem de reservatórios de hidrocarbonetos, permitindo a construção de cenários variáveis e heterogêneos que permitam uma modelagem adequada de fluxo.

A possibilidade de se usar um método ou outro dependerá da problemática em questão e do julgamento do usuário. Mesmo com as correções aqui propostas as estimativas ainda constroem modelos bastante simplistas para representar a realidade geológica, assim como as simulações são pouco confiáveis quando se deseja estabelecer um resultado ótimo previsto para determinada localidade.

Referências Bibliográficas

- ARMSTRONG, M., 1994. Is Research in Mining Geostats as Dead as a Dodo? In Dimitrakopoulos, R. (Ed.), *Geostatistics for the Next Century*. Kluwer Academic. p. 303-312.
- BROOKER, I. B., 1985. Two Dimensional Simulation by Turning Bands. *Mathematical Geology*, vol. 17, nº 1. p. 81-90.
- CAERS, J., 2000. Adding Local Accuracy to Direct Sequential Simulation. *Mathematical Geology*, vol. 32, nº 7. p. 815-850.
- DEUTSCH, C.V., JOURNEL, A.G. *GSLIB Geostatistical Software Library and User's Guide*. Oxford University Press. 1992. 340 p.
- EMERY, X., LANTUÉJOL, C., 2006. TBSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Computers & Geosciences* 32, p. 1615–1628
- GOOVAERTS, P., 1998. Accounting for Estimation Optimality Criteria in Simulated Annealing. *Mathematical Geology*, vol. 30, nº 5. p. 511-534.
- GOOVAERTS, P., 1997. Chapter 7 – Assessment of Local Uncertainty in *Geostatistics for Natural Resources Evaluation*. Oxford University Press. 496 p.
- GUERTIN, K., 1984. Correcting conditional bias in Verly, G., David, M., Journel, A.G., Marechal, A. (Eds.), *Geostatistics for Natural Resources Characterization*, Part. 1. Reidel Publishing Co. p. 245-260.
- ISAAKS, E. H. & SRIVASTAVA, R. M., 1989. *Applied Geostatistics*. Oxford University Press. 561 p.
- JOURNEL, A. G.; KYRIAKIDIS, P. C. e MAO, S., 2000. Correcting the Smoothing Effect of Estimators: A Spectral Postprocessor. *Mathematical Geology*, vol. 32, nº 7. p. 787-813.
- JOURNEL, A. G.; 1980. The Lognormal Approach to Predicting Local Distributions of Selective Mining Unit Grades. *Mathematical Geology*, vol. 12. p. 285-303.
- JOURNEL, A. G.; XU, W., 1994. Posterior Identification Of Histograms Conditional To Local Data. *Mathematical Geology*, vol. 22, nº3, p. 323-359.
- JOURNEL, A. G.; HUIJBREGTS, Ch. J., 1978. *Mining Geostatistics*. Academic Press. 600 p.

- KRIGE, D. G.; ASSIBEY-BONSU, W.; TOLMAY, L., 2004. Post Processing of SK Estimations and Simulations for Assessment of Recoverable Resources and Reserves for South African Gold Mines in Leuangthong, O & Deutsch, C. V. (Eds.), *Geostatistics Banff 2004*, Springer. p. 375-386.
- LANDIM, P. M. B., 2003. *Análise Estatística de Dados Geológicos*. São Paulo: Editora Unesp. 253 p.
- MATHERON, G., 1973. The Intrinsic Random Functions and Their Applications. *Advances in Applied Probability* 5, p. 439-468.
- MATHERON, G., 1962. *Traité de Géostatistique appliquée*. Paris: Editions Technip.
- OLEA, R. A., 1999. *Geostatistics for Engineers and Earth Scientists*. Kluwer Academic Publishers. 303 p.
- OLEA, R. A. (Ed.), 1991. Geostatistical Glossary and Multilingual Dictionary. International Association for Mathematical Geology Studies in *Mathematical Geology* No. 3. Oxford University Press.
- OLEA, R. & PAWLOWSKY, V. Compensating for Estimation Smoothing in Kriging. *Mathematical Geology*, vol. 28, Nº 4, 1996. p. 407-417.
- SAITO, H. & GOOVARTS P., 2000. Geostatistical Interpolation of Positively Skewed and Censored Data in a Dioxin-Contaminated Site. *Environmental Science & Technology*, vol. 34, nº 19. p. 4228-4235.
- SAKA, G.H. & YAMAMOTO, J.K., 2009. *Avaliação de recursos minerais utilizando métodos estocásticos multigaussianos*. Relatório Final. São Paulo, IGc-USP. Fomento a Bolsas de Graduação DIPM – Vale.
- SOUZA, A. L., 2007. *Estudo do Efeito de Suavização da Krigagem Ordinária em Diferentes Distribuições Estatísticas*. São Paulo, IGc-USP, Dissertação de Mestrado. 101 f.
- SRIVASTAVA, R.M., 1994. An Overview of Stochastic Methods for Reservoir Characterization in Yarus, J.M. & Chambers, R.L. (Eds.), *Stochastic Modeling and Geostatistics: Principles, Methods and Case Studies*. *American Association of Petroleum Geologists – Computer Applications in Geology*, No. 3, 1994.
- VANN, J.; BERTOLI, O.; JACKSON, S. *An Overview of Geostatistical Simulation for Quantifying Risk* In: GEOSTATISTICAL ASSOCIATION OF AUSTRALIA SYMPOSIUM “QUANTIFYING RISK AND ERROR” Março de 2002.
- YAMAMOTO, J. K., 2009. *Data Transforms and Unbiased Back Transforms in Geostatistics*. In: INTERNATIONAL ASSOCIATION FOR MATHEMATICAL GEOLOGY ANNUAL CONFERENCE, Stanford University, August 23-29th 2009.

- YAMAMOTO, J. K., 2008. Estimation or Simulation? That is the Question. *Comput. Geosci.*, vol. 12, p. 573-591.
- YAMAMOTO, J. K., 2007. On Unbiased Backtransform of Lognormal Kriging Estimates. *Computers & Geosciences*. vol. 11., p. 219-234.
- YAMAMOTO, J. K., 2005. Correcting the Smoothing Effect of Ordinary Kriging Estimates. *Mathematical Geology*, vol. 37, nº 1. p. 69-94.
- YAMAMOTO, J. K., 2002. *Guia do usuário Geovisual*. Instituto de Geociências – USP, São Paulo. 82 p.
- YAMAMOTO, J. K., 2000. An Alternative Measure of the Reliability of Ordinary Kriging Estimates. *Mathematical Geology*, vol. 32, nº 4. p. 489-509.
- YAMAMOTO, J. K., 1998. A Review of Numerical Methods for the Interpolation of Geological Data. *Anais da Academia Brasileira de Ciências*, vol. 70, nº 1.
- YAO, T., 1998. Conditional Spectral Simulation with Phase Identification. *Mathematical Geology*, vol. 30, nº 3. p. 285-308.

Anexos

Anexo I – Imagens das Estimativas da Krigagem Ordinária

Figura 6.1. Imagens das estimativas da Krigagem Ordinária para os conjuntos Log1 a Log9.

Figura 6.2. Imagens das estimativas da Krigagem Ordinária para os conjuntos Log10 a Log18.

Figura 6.3. Imagens das estimativas da Krigagem Ordinária para os conjuntos Log19 a Log27.

Anexo II – Imagens das Estimativas da Krigagem Ordinária Corrigida

Figura 6.4. Imagens das estimativas da Krigagem Ordinária Corrigida do Efeito de Suavização para os conjuntos Log1 a Log9.

Figura 6.5. Imagens das estimativas da Krigagem Ordinária Corrigida do Efeito de Suavização para os conjuntos Log10 a Log18.

Figura 6.6. Imagens das estimativas da Krigagem Ordinária Corrigida do Efeito de Suavização para os conjuntos Log19 a Log27.

Anexo III – Gráficos Acumulativos e PP Plots para Amostras

Figura 6.7. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis) e P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log1 a Log9.

Figura 6.8. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis) e P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log10 a Log18.

Figura 6.9. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis) e P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log19 a Log27.

Anexo IV – Gráficos Acumulativos e P-P Plots para Exaustivos

Figura 6.10. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis). P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log1 a Log9.

Figura 6.11. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis). P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log10 a Log18.

Figura 6.12. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KO (círculos verdes) e da KO corrigida do efeito de suavização (quadrados azuis). P-P Plots das estimativas da KO (círculos verdes) e da KO corrigida (quadrados azuis). Estimativas para Log19 a Log27.

Anexo V – Gráficos de Dispersão das Distribuições Calculadas pela KO x Referência

Figura 6.13. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária. Log1 a Log9.

Figura 6.14. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária. Log10 a Log18.

Figura 6.15. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária. Log19 a Log27.

Anexo VI – Gráficos de Dispersão das Distribuições Calculadas pela KO Corrigida x Referência

Figura 6.16. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária corrigida do efeito de suavização. Log1 a Log9.

Figura 6.17. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária corrigida do efeito de suavização. Log10 a Log18.

Figura 6.18. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária corrigida do efeito de suavização. Log19 a Log27.

Anexo VII – Imagens KO Lognormal Corrigidas (Journal, 1980)

Figura 6.19. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Journal (1980) para os conjuntos Log1 a Log9.

Figura 6.20. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Journel (1980) para os conjuntos Log10 a Log18.

Figura 6.21. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Journel (1980) para os conjuntos Log19 a Log27.

Anexo VIII – Imagens KO Lognormal Corrigidas (Yamamoto, 2007)

Figura 6.22. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007) para os conjuntos Log1 a Log9.

Figura 6.23. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007) para os conjuntos Log10 a Log18.

Figura 6.24. Imagens das estimativas pela Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007) para os conjuntos Log19 a Log27.

Anexo IX – Comparação das Distribuições Calculadas pela KO Lognormal x Amostras

Figura 6.25. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log1 a Log9.

Figura 6.26. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log10 a Log18.

Figura 6.27. Gráficos acumulativos para as amostras (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log19 a Log27.

Anexo X – Comparação das Distribuições Calculadas pela KO

Lognormal x Dados Exaustivos

Figura 6.28. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log1 a Log9.

Figura 6.29. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log10 a Log18.

Figura 6.30. Gráficos acumulativos para os dados exaustivos (cruzes vermelhas), estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis) e P-P Plots das estimativas da KOL corrigida pelo método de Journal (círculos verdes) e da KOL corrigida pelo método de Yamamoto (quadrados azuis). Estimativas para Log19 a Log27.

Anexo XI – Gráficos de Dispersão KOLognormal Corrigida (Journal, 1980) x Exaustivo

Figura 6.31. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Journal (1980). Log1 a Log9.

Figura 6.32. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Journal (1980). Log10 a Log18.

Figura 6.33. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Journal (1980). Log19 a Log27.

Anexo XII – Gráficos de Dispersão KOLognormal Corrigida (Yamamoto, 2007) x Exaustivo

Figura 6.34. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007). Log1 a Log9.

Figura 6.35. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007). Log10 a Log18.

Figura 6.36. Gráficos de dispersão entre os valores exaustivos e as estimativas da Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto (2007). Log19 a Log27.

Anexo XIII – Variogramas Comparados para a Krigagem Ordinária

Figura 6.37. Comparação dos variogramas obtidos na Krigagem Ordinária. Estão representadas os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária (círculos vazios) e para a Krigagem Ordinária corrigida do efeito de suavização (círculos cheios). Log1 a Log9.

Figura 6.38. Comparação dos variogramas obtidos na Krigagem Ordinária. Estão representadas os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária (círculos vazios) e para a Krigagem Ordinária corrigida do efeito de suavização (círculos cheios). Log10 a Log18.

Figura 6.39. Comparação dos variogramas obtidos na Krigagem Ordinária. Estão representadas os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária (círculos vazios) e para a Krigagem Ordinária corrigida do efeito de suavização (círculos cheios). Log19 a Log27.

Anexo XIV – Variogramas Comparados KO Lognormal

Figura 6.40. Comparação dos variogramas obtidos na Krigagem Ordinária Lognormal. Estão representados os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária Lognormal corrigida pelo método de Journel (1980)(círculos vazios) e para a Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto, 2007 (círculos cheios). Log1 a Log9.

Figura 6.41. Comparação dos variogramas obtidos na Krigagem Ordinária Lognormal. Estão representados os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária Lognormal corrigida pelo método de Journel (1980) (círculos vazios) e para a Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto, 2007 (círculos cheios). Log10 a Log18.

Figura 6.42. Comparação dos variogramas obtidos na Krigagem Ordinária Lognormal. Estão representados os variogramas para as amostras (asteriscos), o modelo adotado (linha contínua), o variograma para a Krigagem Ordinária Lognormal corrigida pelo método de Journel (1980) (círculos vazios) e para a Krigagem Ordinária Lognormal corrigida pelo método de Yamamoto, 2007 (círculos cheios). Log19 a Log27.

Anexo XV – Variação dos Coeficientes de Correlação e Erros RMS das Simulações Gaussianas Sequenciais

Figura 6.43. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log1 a Log3.

Figura 6.44. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log4 a Log6.

Figura 6.45. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log7 a Log9.

Figura 6.46. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log10 a Log12.

Figura 6.47. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log13 a Log15.

Figura 6.48. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log16 a Log18.

Figura 6.49. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto Log19 a Log21.

Figura 6.50. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log22 a Log24.

Figura 6.51. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação gaussiana sequencial. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log25 a Log27.

Anexo XVI - Realizações da SGS de Maiores e Menores Coeficientes de Correlação

Figura 6.52. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log1 a Log3.

Figura 6.53. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log4 a Log6.

Figura 6.54. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log7 a Log9.

Figura 6.55. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log10 a Log12.

Figura 6.56. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log13 a Log15.

Figura 6.57. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log16 a Log18.

Figura 6.58. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log19 a Log21.

Figura 6.59. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log22 a Log24.

Figura 6.60. Realizações da Simulação Gaussiana Sequencial de Maiores e Menores Coeficientes de Correlação. Log25 a Log27.

Anexo XVII – Coeficientes de Correlação e Erros RMS das Simulações por Bandas Rotativas

Figura 6.61. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log1 a Log3.

Figura 6.62. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log4 a Log6.

Figura 6.63. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto Log7 a Log9.

Figura 6.64. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log10 a Log12.

Figura 6.65. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log13 a Log15.

Figura 6.66. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto Log16 a Log18.

Figura 6.67. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log19 a Log21.

Figura 6.68. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log22 a Log24.

Figura 6.69. Variação dos coeficientes de correlação com os dados exaustivos e erros RMS das 100 realizações da simulação por bandas rotativas. Valores máximos (em círculos cheios) e mínimos (círculos vazios) representados, assim como os coeficientes de correlação para a KOL corrigida por Journal (linha fraca com *) e a KOL corrigida por Yamamoto (linha grossa com **). Log25 a Log27.

Anexo XVIII - Realizações da TBSim de Maiores e Menores Coeficientes. de Correlação

Figura 6.70. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log1 a Log3.

Figura 6.71. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log4 a Log6.

Figura 6.72. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log7 a Log9.

Figura 6.73. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log10 a Log12.

Figura 6.74. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log13 a Log15.

Figura 6.75. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log16 a Log18.

Figura 6.76. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log19 a Log21.

Figura 6.77. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log22 a Log24.

Figura 6.78. Realizações da Simulação por Bandas Rotativas de Maiores e Menores Coeficientes de Correlação. Log25 a Log27.