

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

**ANÁLISE DO IMPACTO DOS MÉTODOS DE
AMOSTRAGEM NA REPRODUÇÃO DA TEXTURA
GEOLÓGICA COM BASE EM DIFERENTES
DISTRIBUIÇÕES ESTATÍSTICAS**

Marcelo Monteiro da Rocha

Orientador: Prof. Dr. Jorge Kazuo Yamamoto

TESE DE DOUTORAMENTO

Programa de Pós-Graduação em Recursos Minerais e Hidrogeologia

SÃO PAULO
2004

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

**ANÁLISE DO IMPACTO DOS MÉTODOS DE AMOSTRAGEM NA
REPRODUÇÃO DA TEXTURA GEOLÓGICA COM BASE EM
DIFERENTES DISTRIBUIÇÕES ESTATÍSTICAS**

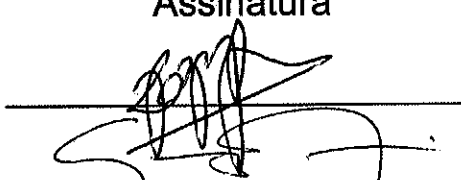
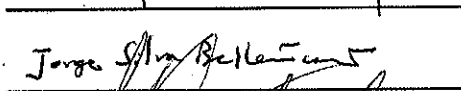

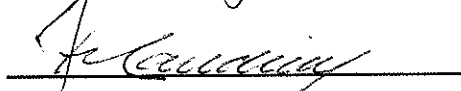
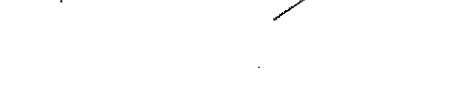
MARCELO MONTEIRO DA ROCHA

Orientador: Prof. Dr. Jorge Kazuo Yamamoto

TESE DE DOUTORAMENTO



COMISSÃO JULGADORA

	Nome	Assinatura
Presidente	Prof. Dr. Jorge Kazuo Yamamoto	
Examinadores:	Prof. Dr. Giorgio Francesco Cesare de Tomi	
	Prof. Dr. Jorge Silva Bettencourt	
	Prof. Dr. José Alberto Quintanilha	
	Prof. Dr. Paulo Milton Barbosa Landim	

SÃO PAULO
2004

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS**

**ANÁLISE DO IMPACTO DOS MÉTODOS DE AMOSTRAGEM NA
REPRODUÇÃO DA TEXTURA GEOLÓGICA COM BASE EM
DIFERENTES DISTRIBUIÇÕES ESTATÍSTICAS**

Marcelo Monteiro da Rocha

Orientador: Prof. Dr. Jorge Kazuo Yamamoto



TESE DE DOUTORAMENTO

Programa de Pós-Graduação em Recursos Minerais e Hidrogeologia

São Paulo

2004

DEDALUS - Acervo - IGC



30900015046

À Ariadne, minha mãe.

Resumo

Este trabalho discute o impacto de métodos de amostragem na reprodução da textura geológica (pelo histograma e variograma) com base em três diferentes distribuições estatísticas, quais sejam: distribuições assimétricas, positiva e negativa, e distribuição simétrica.

Os três métodos de amostragem testados foram: aleatória simples, aleatória estratificada e sistemática. Utilizou-se a simulação estocástica geoestatística para gerar os três modelos de distribuição. Uma única realização, de cada modelo de distribuição, foi escolhida, ao acaso, para ser a base de dados de referência. Os modelos de distribuição simulam distribuições de teores de ouro (lognormal), P_2O_5 (simétrica) e Fe (distribuição assimétrica negativa). Estas bases de dados foram amostradas, gerando amostras de diferentes tamanhos. Foram realizadas a análise estatística e geoestatística e a análise de superfícies de tendência, para todas as amostras. Os resultados, da análise geoestatística, permitiram o cálculo dos recursos geológicos, por meio da krigagem ordinária.

Os resultados foram cruzados com os dados de referência e, por meio desta comparação, concluiu-se que a amostragem aleatória estratificada gerou resultados melhores que os outros dois métodos. A amostragem aleatória simples e a amostragem aleatória estratificada foram métodos que apresentaram resultados muito próximos, mas a amostragem aleatória estratificada foi ligeiramente superior na reprodução da textura geológica, que pode ser considerada como a pedra fundamental em uma avaliação segura de recursos minerais. A amostragem sistemática poderia ser o melhor método se não apresentasse um revés. Um número muito grande de unidades amostrais precisa ser coletado para que a abertura da malha regular seja inferior à amplitude do modelo de variograma subjacente.

Abstract

This work discusses the impact of sampling methods for reproduction of the geological texture (by histogram and variogram). This study was based on three different statistical distributions: positively and negatively skewed distributions and symmetric distribution.

The three sampling methods that were tested were: simple random sampling, stratified random sampling and systematic sampling. Stochastic simulation was used to generate these three distribution models. One single simulation, of each distribution model, was randomly chosen to be the reference data set. The distribution models simulate grade distributions of gold (lognormal distribution), P_2O_5 (symmetrical) and Fe (negatively skewed distribution). These data sets were sampled drawing samples of different sizes. Statistical and geostatistical analysis and trend surfaces analysis were done for all samples. Results of geostatistical analysis allowed a geological resource computation through ordinary kriging.

The results were then checked with the reference data sets, and from this comparison it was concluded that stratified random sampling leads to better results than the other two sampling methods. Simple random sampling and stratified random sampling are considered methods where the results are close together, but stratified random sampling is slightly better for reproducing the geological texture, which can be considered the cornerstone of a reliable mineral resource evaluation. Systematic sampling would be the best method if it did not have a drawback. A very large number of sample units need to be drawn in order to produce a grid spacing less than the range of the underlying variogram model.

Agradecimentos

Gostaria de agradecer ao Prof. Dr. Jorge Kazuo Yamamoto pela orientação, a qual, sem dúvida, foi fundamental para o desenvolvimento deste trabalho.

À FAPESP, processos **99/01089-6** e **01/10948-4**, pelo fomento dado a esta pesquisa, na forma de uma bolsa de estudos e de um projeto de auxílio à pesquisa.

À Mineração Santa Elina, pela cessão da base de dados da Mina de São Vicente e pelo apoio logístico oferecido durante a visita à mina.

Ao Instituto de Geociências da Universidade de São Paulo, sobretudo ao Departamento de Geologia Sedimentar e Ambiental (GSA), pela acolhida calorosa e pela estrutura oferecida, sem as quais este trabalho não teria como ser desenvolvido.

Aos amigos e colegas funcionários, docentes e não docentes, e aos alunos do Instituto pela colaboração direta ou indireta durante esses anos de convívio.

À Ana Paula Cabanal Pentagna e à Magali Poli Fernandes Rizzo pela gentileza que lhes é peculiar e por tornarem a secretaria de pós-graduação tão agradável de ser visitada.

Ao amigo Prof. Dr. Teodoro Isnard Ribeiro de Almeida pelo apoio e incentivo dados nas horas difíceis. À amiga Profa. Dra. Lília Mascarenhas Sant'Agostino, pelas discussões técnicas a respeito do trabalho e, ainda, por estes anos de convívio edificante.

À amiga Neide Yoko Watanabe pela leitura criteriosa do volume final do trabalho, a amiga Veridiana Teixeira de Souza Martins pelos muitos almoços compartilhados e pela leitura e sugestões na elaboração do *abstract*. Aos colegas da Sala 106 pelos períodos de trabalho, ou não, que tivemos.

Aos amigos Jorge Watanabe e Sidney Schaberle Goveia, agradeço o inestimável auxílio dado, sem o qual, com certeza, esta tese não teria sido finalizada. Pelas infindáveis horas de trabalho nos finais de semana, pelas noites à frente do computador e pelas muitas horas de descontração que desfrutamos, tenham certeza que não existem palavras que possam expressar o meu sentimento de gratidão.

Aos meus irmãos, à minha mãe e aos meus pequenos, pela paciência, tolerância, apoio inconstitucional e amor que dedicaram a mim ao longo de nossas vidas. Ladi, você é peça fundamental da minha existência, te adoro.

Sumário

Resumo	i
Abstract.....	ii
Agradecimentos.....	iii
Sumário	iv
Índice de Figuras	vii
Índice de Tabelas	xii
Capítulo 1	1
Introdução.....	1
1.1 Objetivos.....	3
1.2 Justificativas.....	5
Capítulo 2	7
Trabalhos Anteriores.....	7
2.1 Revisão Bibliográfica	8
Capítulo 3	14
Materiais e Métodos.....	14
3.1 Materiais	14
3.2 Métodos	15
3.2.1 Geração dos Dados de Referência	15
3.2.1.1 Simulação da Distribuição Lognormal	18
3.2.1.2 Simulação da Distribuição Simétrica.....	20
3.2.1.3 Simulação da Distribuição Assimétrica Negativa	22
3.3 Re-amostragem dos Dados de Referência	24
3.3.1 Amostragem Aleatória Simples.....	28
3.3.2 Amostragem Aleatória Estratificada.....	29
3.3.3 Amostragem Sistemática.....	30
3.4 Análise Estatística	31
3.4.1 Diagramas Q-Q e P-P	31
3.5 Análise Geoestatística.....	32
3.6 Krigagem Ordinária.....	33
3.7 Avaliação de Recursos/reservas Minerais	34
3.7.1 Determinação da Fronteira Convexa dos Dados	35
3.8 Análise de Superfícies de Tendência.....	35
3.8.1 Atitude de Superfícies de Tendência de Grau 1	37
3.9 Análise Espectral.....	39
3.9.1 Matemática das Séries de Fourier.....	40
3.9.2 Análise Harmônica	42
Capítulo 4	45
Resultados.....	45
4.1 Amostragem.....	45
4.2 Análise Estatística dos Níveis Percentuais	50
4.2.1 Amostragem Aleatória Simples.....	50
4.2.1.1 Variável Assimétrica Positiva	51
4.2.1.2 Variável Simétrica	52
4.2.1.3 Variável Assimétrica Negativa.....	58
4.2.2 Amostragem Aleatória Estratificada.....	61

4.2.2.1 Variável Assimétrica Positiva	61
4.2.2.2 Variável Simétrica	64
4.2.2.3 Variável Assimétrica Negativa	67
4.2.3 Amostragem Sistemática.....	70
4.2.3.1 Variável Assimétrica Positiva	70
4.2.3.2 Variável Simétrica	73
4.2.3.3 Variável Assimétrica Negativa.....	75
4.3 Análise Geoestatística dos Níveis de Percentuais	77
4.3.1 Amostragem Aleatória Simples.....	77
4.3.1.1 Variável Assimétrica Positiva	78
4.3.1.2 Variável Simétrica	79
4.3.1.3 Variável Assimétrica Negativa.....	81
4.3.2 Amostragem Aleatória Estratificada.....	82
4.3.2.1 Variável Assimétrica Positiva	82
4.3.2.2 Variável Simétrica	84
4.3.2.3 Variável Assimétrica Negativa.....	85
4.3.3 Amostragem Sistemática.....	86
4.3.3.1 Variável Assimétrica Positiva	87
4.3.3.2 Variável Simétrica	88
4.3.3.3 Variável Assimétrica Negativa.....	89
4.4 Análise de Superfícies de Tendência.....	90
4.4.1 Amostragem Aleatória Simples.....	90
4.4.2 Amostragem Aleatória Estratificada.....	95
4.4.3 Amostragem Sistemática.....	100
4.5 Recursos Geológicos	104
4.5.1 Amostragem Aleatória Simples.....	105
4.5.1.1 Variável Assimétrica Positiva	105
4.5.1.2 Variável Simétrica	108
4.5.1.3 Variável Assimétrica Negativa.....	110
4.5.2 Amostragem Aleatória Estratificada.....	111
4.5.2.1 Variável Assimétrica Positiva	111
4.5.2.2 Variável Simétrica	113
4.5.2.3 Variável Assimétrica Negativa.....	116
4.5.3 Amostragem Sistemática.....	117
4.5.3.1 Variável Assimétrica Positiva	117
4.5.3.2 Variável Simétrica	119
4.5.3.3 Variável Assimétrica Negativa.....	121
4.6 Análise de Espectral	123
4.6.1 Amostragem Aleatória Simples.....	124
4.6.2 Amostragem Aleatória Estratificada.....	125
4.6.3 Amostragem Sistemática.....	126
Capítulo 5	127
Discussão dos Resultados.....	127
5.1 Recobrimento	127
5.2 Análise Estatística	127
5.2.1 Variável Assimétrica Positiva	128
5.2.2 Variável Simétrica	128
5.2.3 Variável Assimétrica Negativa.....	129
5.3 Análise Geoestatística.....	130

5.3.1 Amostragem Aleatória Estratificada x Amostragem Sistemática	133
5.3.2 Amostragem Aleatória Simples x Amostragem Sistemática	135
5.4 Superfícies de Tendência	137
5.5 Recursos Geológicos	138
5.6 Análise Espectral	140
Capítulo 6	141
Conclusões	141
Referências Bibliográficas	144

Índice de Figuras

<i>Figura 1: Variogramas experimentais apresentados por Caers (2001b), a linha cheia representa o variograma experimental da imagem de referência e a linha pontilhada representa o variograma experimental dos dados obtidos por re-amostragem.....</i>	<i>2</i>
<i>Figura 2: Fluxograma mostrando a seqüência de métodos aplicados.....</i>	<i>16</i>
<i>Figura 3: Histograma dos dados de referência (distribuição assimétrica positiva).....</i>	<i>19</i>
<i>Figura 4: Variograma experimental e respectivo modelo teórico ajustado aos dados de referência (distribuição assimétrica positiva), após a transformação reversa.</i>	<i>19</i>
<i>Figura 5: Variograma experimental e respectivo modelo de variograma ajustado aos dados de referência (variável simétrica), após a transformação reversa.</i>	<i>21</i>
<i>Figura 6: Histograma, para a variável simétrica, dos dados de referência.....</i>	<i>22</i>
<i>Figura 7: Histograma, da variável assimétrica negativa, dos dados de referência.</i>	<i>23</i>
<i>Figura 8: Variograma experimental e respectivo modelo teórico ajustado aos dados de referência (variável assimétrica negativa), após a transformação reversa.</i>	<i>23</i>
<i>Figura 9: Atitude da superfície de tendência de grau 1 (modificado de Marshak, 1998).....</i>	<i>37</i>
<i>Figura 10: Componentes de uma onda.</i>	<i>41</i>
<i>Figura 11: Superfície de espectro de potência bruto calculado a partir dos coeficientes de series duplas de Fourier.</i>	<i>44</i>
<i>Figura 12: Fronteira convexa de três níveis percentuais (A) 1%, (B) 5% e (C) 10% da amostragem aleatória simples.....</i>	<i>48</i>
<i>Figura 13: Fronteira convexa de três níveis percentuais (A) 1%, (B) 5% e (C) 10% da amostragem aleatória estratificada.....</i>	<i>49</i>
<i>Figura 14: Fronteira convexa de três níveis percentuais (A) 1%, (B) 5% e (C) 10% da amostragem sistemática.</i>	<i>50</i>
<i>Figura 15: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável assimétrica positiva, obtidos por amostragem aleatória simples.....</i>	<i>52</i>
<i>Figura 16: Diagramas Q-Q da variável assimétrica positiva referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória simples.</i>	<i>54</i>
<i>Figura 17: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável simétrica, obtidos por amostragem aleatória simples.</i>	<i>56</i>
<i>Figura 18: Diagramas Q-Q da variável simétrica referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória simples.</i>	<i>57</i>
<i>Figura 19: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável assimétrica negativa, obtidos por amostragem aleatória simples.....</i>	<i>58</i>
<i>Figura 20: Diagramas Q-Q da variável assimétrica negativa referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória simples.</i>	<i>60</i>
<i>Figura 21: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável assimétrica positiva, obtidos por amostragem aleatória estratificada. ...</i>	<i>61</i>
<i>Figura 22: Diagramas Q-Q da variável assimétrica positiva referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória estratificada.</i>	<i>63</i>
<i>Figura 23: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável simétrica, obtidos por amostragem aleatória estratificada.....</i>	<i>64</i>

- Figura 24: Diagramas Q-Q da variável simétrica referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória estratificada.66
- Figura 25: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável assimétrica negativa, obtidos por amostragem aleatória estratificada...67
- Figura 26: Diagramas Q-Q da variável assimétrica negativa referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória estratificada.68
- Figura 27: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável assimétrica positiva, obtidos por amostragem sistemática.71
- Figura 28: Diagramas Q-Q da variável assimétrica positiva referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem sistemática.....72
- Figura 29: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável simétrica, obtidos por amostragem sistemática.....73
- Figura 30: Diagramas Q-Q da variável simétrica referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem sistemática.74
- Figura 31: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável assimétrica negativa, obtidos por amostragem sistemática.75
- Figura 32: Diagramas Q-Q da variável assimétrica negativa referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem sistemática.....76
- Figura 33: Variogramas experimentais médios, da variável assimétrica positiva, amostrada por amostragem aleatória simples, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 7% e em (B) os níveis percentuais de 8% a 40%.....79
- Figura 34: Variogramas experimentais médios, da variável simétrica, amostrada por amostragem aleatória simples, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 7% e em (B) os níveis percentuais de 10% a 40%.....80
- Figura 35: Variogramas experimentais médios, da variável assimétrica negativa, amostrada por amostragem aleatória simples, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 7% e em (B) os níveis percentuais de 8% a 40%.....81
- Figura 36: Variogramas experimentais médios, da variável assimétrica positiva, amostrada por amostragem aleatória estratificada, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 7% e em (B) os níveis percentuais de 10% a 40%.....83
- Figura 37: Variogramas experimentais médios, da variável simétrica, amostrada por amostragem aleatória estratificada, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 8% e em (B) os níveis percentuais de 10% a 40%.....84
- Figura 38: Variogramas experimentais médios, da variável assimétrica negativa, amostrada por amostragem aleatória estratificada, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 8% e em (B) os níveis percentuais de 10% a 40%.....85
- Figura 39: Variogramas experimentais médios, da variável assimétrica positiva, coletada por amostragem sistemática, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 8% e em (B) os níveis percentuais de 10% a 50%.....87

- Figura 40: Variogramas experimentais médios, da variável simétrica, coletada por amostragem sistemática, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 8% e em (B) os níveis percentuais de 10% a 50%.88
- Figura 41: Variogramas experimentais médios, da variável assimétrica negativa, coletada por amostragem sistemática, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 8% e em (B) os níveis percentuais de 10% a 50%.89
- Figura 42: Polos dos planos das superfícies de tendência de grau 1 da variável assimétrica positiva (amostragem aleatória simples). O pólo da superfície dos dados de referência é o círculo (vermelho) de maior diâmetro.91
- Figura 43: Polos dos planos das superfícies de tendência de grau 1 da variável simétrica (amostragem aleatória simples). O pólo da superfície dos dados de referência é o círculo (vermelho) de maior diâmetro.92
- Figura 44: Polos dos planos das superfícies de tendência de grau 1 da variável assimétrica negativa (amostragem aleatória simples). O pólo da superfície dos dados de referência é o círculo (vermelho) de maior diâmetro.93
- Figura 45: Diagrama dos erros quadráticos médios em relação à superfície (grau 1) dos dados de referência para as três variáveis analisadas, coletadas por amostragem aleatória simples.94
- Figura 46: Diagrama dos erros quadráticos médios em relação à superfície (grau 2) dos dados de referência para as três variáveis analisadas, coletadas por amostragem aleatória simples.94
- Figura 47: Diagrama dos erros quadráticos médios em relação à superfície (grau 3) dos dados de referência para as três variáveis analisadas, coletadas por amostragem aleatória simples.95
- Figura 48: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável assimétrica positiva. O ponto de maior diâmetro é o pólo do plano dos dados de referência, para a amostragem aleatória estratificada.96
- Figura 49: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável simétrica. O ponto de maior diâmetro é o pólo do plano dos dados de referência, para a amostragem aleatória estratificada.97
- Figura 50: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável assimétrica negativa. O ponto de maior diâmetro é o pólo do plano dos dados de referência, para a amostragem aleatória estratificada.99
- Figura 51: RMS das superfícies de tendência de grau 1, calculadas para as amostras coletadas por amostragem aleatória estratificada.99
- Figura 52: RMS das superfícies de tendência de grau 2 calculadas para as amostras selecionadas por amostragem aleatória estratificada.100
- Figura 53: RMS calculado para as superfícies de tendência de grau 3 calculadas para as amostras selecionadas por amostragem aleatória estratificada.100
- Figura 54: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável assimétrica positiva (amostragem sistemática).101
- Figura 55: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável simétrica (amostragem sistemática).101
- Figura 56: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável assimétrica negativa (amostragem sistemática).103
- Figura 57: RMS calculado para as superfícies de tendência de grau 1 (amostragem sistemática).103
- Figura 58: RMS calculado para as superfícies de tendência de grau 2 (amostragem sistemática).104

<i>Figura 59: RMS calculado para as superfícies de tendência de grau 3 (amostragem sistemática).</i>	104
<i>Figura 60: Recursos geológicos calculados para as amostras da variável assimétrica positiva (amostragem aleatória simples).</i>	106
<i>Figura 61: Curva de parametrização dos recursos geológicos calculados para a 6ª realização do nível percentual de 9% (A) e da 6ª realização do nível percentual de 30% (B), para a variável assimétrica positiva, coletada por amostragem aleatória simples.</i>	107
<i>Figura 62: Recursos geológicos calculados para as amostras da variável simétrica (amostragem aleatória simples).</i>	108
<i>Figura 63: Curva de parametrização dos recursos geológicos calculados para a 5ª realização do nível percentual de 5% (A) e da 6ª realização do nível percentual de 30% (B), amostradas por amostragem aleatória simples.</i>	109
<i>Figura 64: Recursos geológicos calculados para as amostras da variável assimétrica negativa (amostragem aleatória simples).</i>	110
<i>Figura 65: Curva de parametrização dos recursos geológicos calculados da variável assimétrica negativa, coletada por amostragem aleatória simples para a 6ª realização do nível percentual de 2% (A) e da 5ª realização do nível percentual de 30% (B).</i>	111
<i>Figura 66: Recursos geológicos calculados para as amostras da variável assimétrica positiva (amostragem aleatória estratificada).</i>	112
<i>Figura 67: Curva de parametrização dos recursos geológicos calculados para a 6ª realização do nível percentual de 10% (A) e da 6ª realização do nível percentual de 30% (B), para a variável assimétrica positiva, coletada por amostragem aleatória estratificada.</i>	113
<i>Figura 68: Recursos geológicos calculados para as amostras da variável simétrica (amostragem aleatória estratificada).</i>	114
<i>Figura 69: Curva de parametrização dos recursos geológicos calculados para a 5ª realização do nível percentual de 5% (A) e da 6ª realização do nível percentual de 30% (B), para a variável simétrica, retirada por amostragem aleatória estratificada.</i>	115
<i>Figura 70: Recursos geológicos calculados para as amostras da variável assimétrica negativa (amostragem aleatória estratificada).</i>	116
<i>Figura 71: Curva de parametrização dos recursos geológicos calculados para a 6ª realização do nível percentual de 2% (A) e da 5ª realização do nível percentual de 30% (B), para a variável assimétrica negativa, coletada por amostragem aleatória estratificada.</i>	117
<i>Figura 72: Recursos geológicos calculados para as amostras da variável assimétrica positiva (amostragem sistemática).</i>	118
<i>Figura 73: Curva de parametrização dos recursos geológicos calculados para o nível percentual de 10% (A) e do nível percentual de 30% (B), para a variável assimétrica positiva, coletada por amostragem sistemática.</i>	119
<i>Figura 74: Recursos geológicos calculados para as amostras da variável simétrica (amostragem sistemática).</i>	120
<i>Figura 75: Curva de parametrização dos recursos geológicos calculados para o nível percentual de 5% (A) e do nível percentual de 30% (B) para a variável simétrica, coletada por amostragem sistemática.</i>	121
<i>Figura 76: Recursos geológicos calculados para as amostras da variável assimétrica negativa (amostragem sistemática).</i>	122

- Figura 77: Curva de parametrização dos recursos geológicos calculados para o nível percentual de 2% (A) e do nível percentual de 30% (B), para a variável assimétrica negativa, coletada por amostragem sistemática. 123
- Figura 78: Razão entre o somatório das potências dos dados estimados e a soma das potências dos dados de referência. Amostragem aleatória simples. 124
- Figura 79: Razão entre a soma das potências dos dados estimados e a soma das potências dos dados de referência. Amostragem aleatória estratificada. 125
- Figura 80: Razão entre o somatório dos espectros de potência de cada harmônico dos dados estimados e a soma das potências dos dados de referência para a amostragem sistemática. 126
- Figura 81: Diagramas em pizza, mostrando as tendências das médias e variâncias, da variável assimétrica positiva, para as três amostragens. (A) Amostragem aleatória simples; (B) Amostragem aleatória estratificada e; (C) Amostragem sistemática. 128
- Figura 82: Diagramas em pizza, mostrando as tendências das médias e variâncias, da variável simétrica, para as três amostragens. (A) Amostragem aleatória simples; (B) Amostragem aleatória estratificada e; (C) Amostragem sistemática. 129
- Figura 83: Diagramas em pizza, mostrando as tendências das médias e variâncias, da variável assimétrica negativa, para as três amostragens. (A) Amostragem aleatória simples; (B) Amostragem aleatória estratificada e; (C) Amostragem sistemática. 130
- Figura 84: Distância média de separação entre pares de pontos em função do nível percentual (A) e desvios padrão médios em função das distâncias (B). 132
- Figura 85: Variogramas experimentais e seus respectivos modelos calculados para uma malha regular de 20 x 29 nós, simulando uma amostragem sistemática. (A) é o variograma da variável assimétrica positiva. (B) é o variograma da variável assimétrica negativa e (C) é o variograma para a variável simétrica. 134
- Figura 86: Variogramas experimentais e modelos dos variogramas da Figura 85. (A) variável assimétrica positiva, (B) variável assimétrica negativa e (C) variável simétrica. ... 135
- Figura 87: Variogramas experimentais e os modelos de variogramas calculados a partir de pontos de dados gerados por amostragem aleatória simples sobre a malha regular de 20 x 29 nós. (A) variável assimétrica positiva, (B) variável assimétrica negativa e (C) variável simétrica. 136
- Figura 88: Mapa de pontos em cores proporcionais sobreposto à um mapa de krigagem. As elipses destacam a influência de valores anômalos em áreas sem pontos amostrais. 138

Índice de Tabelas

<i>Tabela 1: Erro médio absoluto da krigagem ordinária, para vários níveis percentuais (modificado de Lloyd & Atkinson, 1999).</i>	3
<i>Tabela 2: Parâmetros do modelo de blocos utilizado na Simulação Sequencial Gaussiana.</i>	16
<i>Tabela 3: Parâmetros do modelo de variograma ajustado aos dados de sondagem da mina de ouro.</i>	18
<i>Tabela 4: Estatísticas da base de dados de referência.</i>	19
<i>Tabela 5: Parâmetros do modelo de variograma ajustado aos dados de referência, após a transformação reversa.</i>	20
<i>Tabela 6: Parâmetros do modelo de variograma ajustado aos dados de P_2O_5 transformados.</i>	21
<i>Tabela 7: Estatísticas da base de dados de referência, variável simétrica.</i>	22
<i>Tabela 8: Parâmetros do modelo de variograma ajustado aos dados de Fe, transformados em uma distribuição Gaussiana normal.</i>	22
<i>Tabela 9: Estatísticas da base de dados de referência.</i>	23
<i>Tabela 10: Parâmetros do modelo de variograma ajustado aos dados de Fe.</i>	23
<i>Tabela 11: Níveis percentuais de retirada de amostras e respectivas quantidades de unidades amostrais coletadas.</i>	24
<i>Tabela 12: Porcentagens da área dos dados de referência contidas nas fronteiras convexas, para a amostragem aleatória simples.</i>	46
<i>Tabela 13: Porcentagens da área dos dados de referência contidas nas fronteiras convexas, para a amostragem aleatória estratificada.</i>	47
<i>Tabela 14: Porcentagens da área dos dados de referência contidas nas fronteiras convexas, para a amostragem sistemática.</i>	47
<i>Tabela 15: Estatísticas dos dados para os diferentes níveis percentuais obtidos por amostragem aleatória simples para a variável assimétrica positiva.</i>	53
<i>Tabela 16: Estatísticas dos dados para os diferentes níveis percentuais obtidos por amostragem aleatória simples da variável simétrica.</i>	55
<i>Tabela 17: Estatísticas dos dados para os diferentes níveis percentuais obtidos por amostragem aleatória simples da variável assimétrica negativa.</i>	59
<i>Tabela 18: Estatísticas descritivas dos dados coletados por amostragem aleatória estratificada da distribuição assimétrica positiva.</i>	62
<i>Tabela 19: Estatísticas descritivas dos dados coletados por amostragem aleatória estratificada. Distribuição simétrica.</i>	65
<i>Tabela 20: Estatísticas descritivas dos dados coletados por amostragem aleatória estratificada. Distribuição assimétrica negativa.</i>	69
<i>Tabela 21: Estatísticas descritivas das variáveis: assimétrica positiva, simétrica e assimétrica negativa.</i>	70
<i>Tabela 22: Parâmetros do cálculo dos variogramas experimentais.</i>	77
<i>Tabela 23: Modelos de variograma ajustados aos variogramas experimentais de cada variável.</i>	78
<i>Tabela 24: Parâmetros do cálculo dos variogramas experimentais.</i>	82
<i>Tabela 25: Modelos de variogramas ajustados aos dados selecionados por amostragem aleatória estratificada.</i>	83
<i>Tabela 26: Parâmetros para o cálculo dos variogramas experimentais.</i>	86
<i>Tabela 27: Modelos de variograma ajustados aos variogramas experimentais calculados para os dados coletados por amostragem sistemática.</i>	86
<i>Tabela 28: Índices médios R_1.</i>	139

<i>Tabela 29: Índices médios R2.....</i>	<i>140</i>
--	------------

Capítulo 1

Introdução

A reprodução da textura geológica, a partir de diferentes métodos de estimativa ou simulação estocástica, com ou sem pós-processamento, tem sido amplamente discutida, e.g. Journel & Rao (1997), Journel *et al.* (2000), Caers (2000), Caers (2001a) e Soares (2001). Contudo, nenhum destes trabalhos considera a representatividade da amostragem ou, simplesmente, se o número de unidades amostrais coletadas, em uma amostra, aproxima as estatísticas amostrais da populacional.

Autores como Journel & Rao (1997), Deutsch & Wen (1998), Caers (2000) e Caers (2001a) re-amostram os dados de referência, utilizados em seus estudos, em 0,5% do total de dados da referência. Journel *et al.* (2000) e Caers (2001b) utilizam uma re-amostragem de 1% do total dos dados da base de referência. Apenas Bourgault (1997) e Lloyd & Atkinson (1999) trabalharam com níveis percentuais superiores a 1%, sendo que o primeiro autor utilizou 9%, enquanto os últimos utilizaram seis níveis percentuais diferentes que variam entre 2% e 50%. Deve-se destacar que, em todos os artigos citados, as bases de dados de referência possuem 10000 ou mais pontos.

Deutsch & Journel (1996) utilizaram aproximadamente 5% dos dados de uma base de dados de referência. Observa-se que, para a amostragem utilizada por estes autores, a textura geológica não é reproduzida e, ainda, o variograma não é estruturado. Vindo o variograma apresentar estrutura, após dois procedimentos de manipulação matemática, respectivamente, o desagrupamento (*declustering*) e a transformação gaussiana (*normal score transformation*). Mesmo assim, o modelo original de variância espacial não é reproduzido.

Caers (2001b) mostra que o nível percentual por ele utilizado não reproduz a variância espacial dos dados de referência, conforme pode ser observado na Figura 1.

Caers (2001b) e Deutsch & Journel (1996) utilizam o modelo de variograma dos dados de referência para a estimativa por krigagem ordinária. Porém o modelo de variograma de um depósito mineral nunca é conhecido *a priori*, mas sim *a posteriori*. Deste modo, uma amostragem representativa da população estudada se faz necessária.

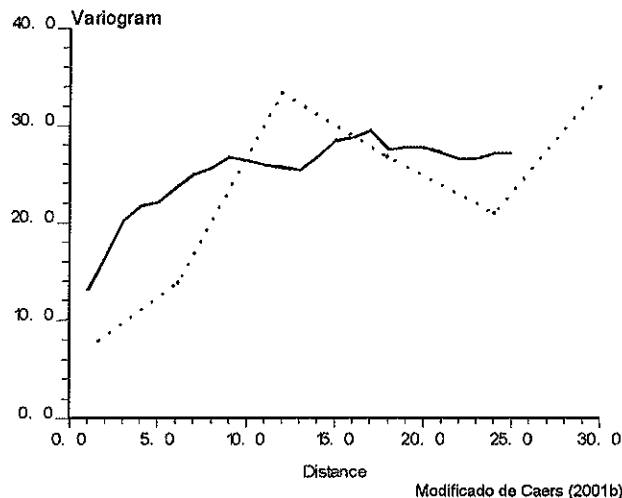


Figura 1: Variogramas experimentais apresentados por Caers (2001b), a linha cheia representa o variograma experimental da imagem de referência e a linha pontilhada representa o variograma experimental dos dados obtidos por re-amostragem.

A precisão dos resultados obtidos no cálculo de recursos/reservas minerais depende, dentre outros, do método de cálculo utilizado, da variabilidade natural que os dados apresentam e da incerteza associada à amostragem. A incerteza, relacionada à representatividade da amostragem, é desconhecida. Então, para uma dada amostragem, qual é o nível de incerteza a ela associada?

Parte da resposta a essa pergunta depende do tipo de variável analisada e, conseqüentemente, do tipo de distribuição estatística; do método de amostragem utilizado; do espaçamento médio e do número de unidades amostrais coletadas e sobretudo, da distribuição espacial dos pontos de dados.

A Tabela 1, modificada de Lloyd & Atkinson (1999), apresenta o erro absoluto da estimativa por krigagem ordinária. Observe-se que o erro absoluto aumenta à medida que os níveis percentuais utilizados na re-amostragem diminuem.

Tabela 1: Erro médio absoluto da krigagem ordinária, para vários níveis percentuais (modificado de Lloyd & Atkinson, 1999).

Nível percentual	Erro médio absoluto
50%	1,3
11%	1,52
6,25%	1,75
4%	1,83
2.78%	1,87
2.05%	2,01

A reprodução da textura geológica, ou seja, por meio do histograma e do modelo de variância espacial do fenômeno estudado, dependerá do número de unidades amostrais coletadas em uma amostra e, principalmente, de sua distribuição espacial. Variáveis com média a alta variabilidade jamais terão sua textura reproduzida através de número pequeno de unidades amostrais e largamente espaçadas. Por outro lado, variáveis com baixa variabilidade tenderão a reproduzir a textura geológica mais rapidamente, com menor número de unidades amostrais.

1.1 Objetivos

A avaliação de recursos/reservas minerais é função direta da qualidade e quantidade de unidades amostrais coletadas durante a fase de pesquisa mineral, dado que a estimativa dos teores, em pontos desconhecidos de um depósito, dependerá dos teores das unidades amostrais presentes em sua vizinhança.

O projeto de uma campanha de sondagem considera o bem mineral pesquisado, ou seja, sua variabilidade característica, além, é claro, de considerar os custos e o tempo envolvidos em sua execução. Porém, a variabilidade intrínseca do bem mineral assume papel secundário, em detrimento aos custos e tempo, na tomada de decisão sobre a adequação do projeto.

O objetivo deste estudo é mostrar a relação existente entre a representatividade de amostras de diferentes tamanhos, sobretudo, no que se refere à reprodução da textura geológica, retiradas de uma base de dados de referência, tomada como a população, e os resultados obtidos na avaliação dos recursos minerais destes dados.

A reprodução da textura geológica dependerá, considerando-se o histograma e as estatísticas associadas à população, apenas da quantidade de unidades amostrais em uma amostra, e dos valores que elas assumem. Já a reprodução do modelo de variância espacial, dependerá da distribuição espacial dos dados coletados. Deste modo, a amostragem utilizada exercerá um papel primordial. É óbvio que a variabilidade do bem mineral pesquisado é um fator preponderante no alcance desta reprodução. Outro fator que também deve ser levado em consideração é o posicionamento que as unidades amostrais têm no domínio. Assim sendo, deseja-se quantificar a proporção da população necessária para atingir a reprodução da textura geológica do domínio. Dessa forma, a base de dados de referência será amostrada segundo vários níveis percentuais, por exemplo, 1%, 2%, 3% etc.

Buscando esta quantificação, serão utilizadas três variáveis, cada uma representando um nível de variabilidade que descreva uma distribuição estatística diferente. A primeira variável testada será derivada do histograma e do modelo de variância espacial da variável ouro, de uma mina de ouro, cuja distribuição característica é lognormal e, conseqüentemente, de alta variabilidade. A segunda variável a ser estudada apresenta distribuição simétrica, ou seja, dados de média variabilidade derivada do histograma e do modelo de variância espacial da variável P_2O_5 , de uma mina de fosfato. A última variável, representando um depósito de baixa variabilidade, será resultante da reprodução do histograma e do modelo de variância espacial da variável Fe de uma mina de ferro. Esta última variável apresenta uma distribuição com assimetria negativa ou lognormal invertida.

Para alcançar os diferentes níveis percentuais, serão utilizados três diferentes métodos de amostragem: aleatória simples, aleatória estratificada e sistemática.

Assim, será possível determinar para cada tipo de distribuição estatística, o número de unidades amostrais, bem como o tipo de amostragem, necessários para reproduzir a textura geológica dos dados de referência.

Outro resultado a ser analisado será a influência do tamanho das amostras, bem como do método de amostragem, nos cálculos dos recursos, resultantes da estimativa por krigagem ordinária. Esta análise será realizada comparando os recursos totais e parciais dos dados de referência com os recursos obtidos em cada amostra. Entenda-se por recursos parciais os recursos calculados, utilizando-se a fronteira convexa gerada a partir dos pontos de coleta de dados em cada amostra.

1.2 Justificativas

A obtenção de precisão global, ou seja, a reprodução de estatísticas e variogramas, em uma base de dados geológica, será função da representatividade da amostra coletada. Para haver a reprodução da textura geológica (histograma e variograma), em uma base de dados, é necessário que haja a reprodução de características da população em estudo, devendo-se, desta forma, considerar uma amostragem capaz de gerar amostra(s) representativa(s) da população. Esta representatividade é resultante de fatores diversos, dentre os quais pode-se destacar: o tamanho da amostra a ser coletada; e os valores que a variável assume nos pontos de coleta para a reprodução do histograma da distribuição. Para a reprodução do variograma, deve-se garantir que a distribuição espacial (esquema de amostragem) dos locais de coleta seja tal que a variância espacial que a população apresenta seja capturada. Deve-se garantir, também, que o tamanho da amostra seja grande o suficiente gerando, assim, pouca variação nas médias dos pares de pontos tomados para o cálculo do variograma experimental, em um determinado passo.

A precisão local, ou seja, a reprodução de médias locais dentro de vizinhanças preestabelecidas será alcançada a depender do número de unidades amostrais contidas na vizinhança definida e, ainda, do nível de variância espacial que a população apresenta. Logo, a precisão local é função do variograma e do arranjo espacial que os pontos de amostragem apresentam na vizinhança definida.

Deste modo, fica claro que tanto a precisão global, por meio da reprodução da textura geológica, quanto a precisão local, com a reprodução de médias locais, é, de um modo ou de outro, função do esquema de amostragem escolhido para a coleta de uma amostra de tamanho n . Portanto, a escolha de um esquema de amostragem adequado é um passo fundamental para se obter sucesso na obtenção das referidas precisões.

Como na avaliação de recursos/reservas minerais a localização das amostras tem importância fundamental, a reprodução do histograma acaba sendo, também, vinculada à posição espacial que os pontos de amostragem possuem. Este vínculo existe, pois as variáveis geológicas não podem ser consideradas como variáveis puramente aleatórias. Estas variáveis são geradas por processos naturais que não são

independentes. Tome-se o exemplo: se em uma região determinada variável passou por um processo de concentração e acumulou-se em uma dada porção do domínio, não se pode amostrar apenas a região de acúmulo se o desejo é conhecer as características da distribuição da variável em todo o domínio. Caso a amostragem seja realizada apenas na porção de acúmulo, as estatísticas populacionais da variável jamais serão bem estimadas. Para estimar bem as estatísticas da população, deve-se amostrar todo o domínio do modo mais homogêneo o possível.

Assim sendo, uma amostragem bem planejada será condizente com o tipo de distribuição estatística que se está trabalhando, sendo, deste modo, o passo mais importante na rotina de avaliação de recursos/reservas minerais.

Capítulo 2

Trabalhos Anteriores

Durante os anos 50 e 60, do século passado, associado aos primórdios da utilização de computadores, houve o aumento de aplicações de estatística em geologia (Krumbein & Graybill, 1965). Ainda segundo estes autores, estas aplicações faziam uso crescente de métodos estatísticos analíticos e não mais, apenas, de estatísticas descritivas.

Neste período, uma das principais aplicações de métodos estatísticos analíticos buscava o reconhecimento e a representação numérica de feições globais, em dados geológicos, que permitissem identificar e separar as variações apresentadas por estes dados. Dos vários métodos existentes destacavam-se aqueles que separavam componentes de variação regional dos de variação local, por exemplo, superfícies de tendência por ajuste polinomial e análise harmônica, a partir de series de Fourier, aplicada a dados cíclicos.

Ainda durante a década de 1960, são publicados os primeiros trabalhos conceituando a teoria das variáveis regionalizadas e a geoestatística, empiricamente desenvolvidos por Krige e Sichel, no final da década de 50, visando estimar teores de ouro em uma mina com precisão local e mínima variância.

Nos anos 70, associado ao crescente aumento dos recursos computacionais, métodos que permitissem a geração automática de mapas de contorno proliferavam. Nesta mesma década, os primeiros trabalhos de geoestatística, por exemplo, Matheron (1971), são publicados em inglês e o método de estimativa por krigagem populariza-se.

A krigagem, considerada como o melhor estimador linear não enviesado (do inglês *B.L.U.E. – Best Linear Unbiased Estimator*), é um método de estimativa exato que possui precisão local, no sentido de mínima variância, porém devido ao efeito de suavização não apresenta precisão global. A utilização dos métodos de krigagem, simples ou ordinária, implica que os dados sejam estacionários de segunda ordem, ou seja, possuam variância espacial finita. Outra limitação da krigagem é a dificuldade de mensurar os erros associados às estimativas, pois estes derivam da variância de krigagem que é uma medida homocedástica, medindo apenas a configuração espacial dos dados na vizinhança em relação ao ponto/bloco estimado.

Ainda durante a década de 70, para aplicação em casos de dados não estacionários de segunda ordem, Matheron desenvolveu o método da Krigagem Universal. A Krigagem Universal consiste basicamente em filtrar a tendência dos dados, segundo Goovaerts (1997), por meio de um ajuste polinomial, com polinômios de baixo grau (não superiores a 2), e a estimativa dos resíduos resultantes por krigagem simples. Este método de estimativa, mesmo atualmente, não se difundiu dada a complexidade em aplicá-lo.

A busca por precisão global em predição de características de variáveis aleatórias regionalizadas teve seu apogeu na década de 80, sobretudo, na indústria do petróleo. A precisão global é alcançada por meio dos métodos geoestatísticos de simulação estocástica. Estes métodos permitem a geração de n mapas, diferentes e equiprováveis, que reproduzem histograma, variograma e textura geológica. Porém, o ônus é a perda da precisão local, alcançada pela krigagem.

As precisões local e global são propriedades conflitantes (Caers, 2000), porém vários autores, durante a década de 90 e início deste século, propuseram alternativas, ora para adicionar precisão global aos resultados da krigagem, ora precisão local aos resultados da simulação estocástica; porém todos sem conferir importância à amostragem ou ao tamanho da amostra utilizada na simulação ou estimativa.

2.1 Revisão Bibliográfica

Segundo Deutsch & Journel (1996), a simulação estocástica diferencia-se da krigagem em dois aspectos: o primeiro é o objetivo, que na krigagem é gerar a melhor estimativa local de cada ponto, dentro do domínio, com valor desconhecido; enquanto na simulação, as feições globais e as estatísticas precedem a precisão local; o segundo aspecto é a krigagem resultar em um único modelo numérico, o melhor no sentido da precisão local; enquanto a simulação, fornece vários modelos numéricos, cada um dos quais uma representação provável da realidade, em termos de precisão global.

Journel *et al.* (2000) citam que a krigagem prima pela precisão local, por intermédio da minimização da variância do erro baseado na covariância, enquanto a simulação visa à reprodução do modelo de covariância dos dados.

Embora a krigagem forneça um único modelo numérico este modelo será suavizado, devido ao efeito suavização da krigagem. O efeito de suavização caracteriza-se pela subestimativa dos valores maiores e superestimativa dos valores menores (Olea, 1991) . Já a simulação estocástica produz vários modelos numéricos, de um mesmo fenômeno, os quais não apresentam efeito de suavização. Porém, a precisão local destes modelos é comprometida, uma vez que apenas a reprodução do modelo de covariância não a garante.

Segundo Caers (2001b), uma única estimativa normalmente não é satisfatória, pois se procura sempre a determinação da incerteza associada ao valor desconhecido do ponto não amostrado. Ainda segundo este autor, em estatística, a incerteza é quantificada por meio de uma distribuição estatística que descreva as freqüências dos valores esperados em um dado local. Esta função de distribuição irá depender do número de dados amostrados, de sua configuração espacial, dos valores que as amostras assumem e do fenômeno espacial em estudo (Caers, 2001b).

Embora Caers (2001b), cite esta dependência, diversos autores, dentre os quais destacam-se Olea & Pawlowsky (1996), Journel *et al.* (2000), Caers (2000), Caers (2001a), e, inclusive, Caers (2001b) não questionam se as amostras utilizadas, ao menos, reproduzem satisfatoriamente a população estudada. Note-se que todos estes autores utilizam re-amostragens de dados de referência. Outro ponto, crítico e questionável, é que a maioria dos autores citados utiliza o modelo de variograma dos dados de referência, na krigagem ou simulação, ao invés do modelo ajustado ao variograma experimental obtido a partir das amostras. O problema deste tipo de abordagem é que, na prática, os dados de referência não são conhecidos e os dados existentes são frutos da amostragem e, conseqüentemente, o variograma calculado não reproduzirá o fenômeno espacial, se a amostra coletada não for representativa da população estudada.

Os objetivos dos trabalhos citados no parágrafo anterior não são testar se as amostras reproduzem o variograma dos dados de referência, mas sim adicionar precisão local aos resultados de diversos métodos de simulação estocástica ou adicionar precisão global ao resultado da krigagem.

Segundo Caers (2000), precisão local e global são propriedades conflitantes e propõe, então, a adição de precisão local às realizações da simulação, pois esta já possui precisão global. Para tal, a proposta é utilizar a abordagem da simulação

seqüencial direta, ou seja, um método de simulação que não requer a transformação prévia dos dados para uma distribuição gaussiana normal, com uma distribuição local conhecida. Resumidamente, o procedimento proposto por Caers (2000) é realizar uma krigagem seqüencial (krigagem realizada utilizando os pontos com valores amostrados e também os pontos previamente estimados), se o resultado for satisfatório, o valor do ponto estimado é retido; caso contrário, escolhe-se uma função de distribuição de frequência predefinida, que identifique a média e a variância da krigagem em cada ponto, e retira-se aleatoriamente um valor nela contido. Os resultados obtidos, ainda segundo Caers (2000), apresentam problemas, porém são satisfatórios.

Caers (2001b) testa o reconhecimento e reprodução de padrões geológicos por meio de vários métodos de estimativa e simulação, concluindo que a informação disponível, sobre um fenômeno com variação espacial, é sempre incompleta. Porém, segundo este autor, sempre existe uma variedade abundante de dados indiretos (obtidos por ferramentas de sensoriamento remoto). Caers (2001b) propõe, então, a integração de ambos os tipos de informação; seja pelo ajuste de um modelo de variograma ao variograma experimental dos dados; seja por uma imagem de treinamento modelada por um conceito ou por analogia, para quantificar a incerteza de um ponto desconhecido. Por fim, Caers (2000) propõe a realização de um pós-processamento utilizando uma função objetiva para determinar a relação custo/benefício e risco associados.

Na questão da obtenção de precisão local e global simultaneamente, Olea & Pawlowsky (1996), propõem a adição de precisão global à estimativa por meio da eliminação do efeito de suavização da krigagem. Esta eliminação é obtida após um pós-processamento, dos resultados da krigagem, que faz uso da validação cruzada. A conclusão destes autores é que o método proposto resulta em superfícies, cujos valores possuem propriedades intermediárias entre aquelas obtidas por krigagem e aquelas resultantes da simulação estocástica.

Na mesma linha, Journel *et al.* (2000) propõem a correção do efeito de suavização da estimativa por meio da utilização de um pós-processamento espectral, denominado pós-processador de Yao. Este pós-processamento, modifica, interativamente, dados estimados (krigagem) contendo propriedades de precisão local, introduzindo amplitudes espectrais correspondentes a um modelo de covariância.

Entretanto, o pós-processamento faz com que a precisão local, característica da krigagem, seja comprometida.

Journel *et al.* (2000), Caers (2001a), Caers (2000) Olea & Pawlowsky (1996) utilizam, conforme citado, estudos de caso que têm como base uma amostragem elaborada em dados de referência. Estes autores utilizam diversas técnicas de amostragem; por exemplo, Journel *et al.* (2000) baseiam-se na amostragem aleatória estratificada, enquanto Caers (2000) na amostragem aleatória simples. Destaca-se, ainda, que todos os artigos consideram imagens de referência com número de células superior a 10000, exceto Caers (2001b) que utiliza 2500 células e Olea & Pawlowsky (1996) com 1600.

O único artigo, dos revisados, que mais se assemelha, ao que será proposto aqui, é o apresentado por Lloyd & Atkinson (1999), onde os autores trabalham com a amostragem sistemática de uma base de dados de referência, utilizando vários incrementos regulares na coleta de amostras. Estes autores utilizam malhas com espaçamento, em metros, de 10 x 10, 15 x 15, 20 x 20, 25 x 25, 30 x 30 e 35 x 35 que perfazem, respectivamente, amostragens de 50%, 11%, 6,25%, 4%, 2,78% e 2,01% do total de dados da base de referência. O objetivo do trabalho de Lloyd & Atkinson (1999) é testar, entre os métodos de krigagem ordinária e krigagem de indicadoras, qual apresenta o menor erro real e o menor erro de estimativa.

Russo (1984), propõe um algoritmo para otimizar a malha de amostragem para o cálculo de variogramas experimentais. Este autor baseia esta otimização na obtenção de uma malha onde, utilizando um número constante de pares de pontos para cada classe de distâncias, haja a minimização de um desvio quadrático médio. Este desvio é baseado no desvio padrão das distâncias contidas em uma classe de passo para o variograma. Este autor realiza um teste de seu algoritmo, porém não leva em conta o tipo de distribuição estatística analisada.

Na mesma linha, Warrick & Myers (1987), também propõem um algoritmo para a otimização, ou geração, de malhas de amostragem para o cálculo de variogramas experimentais. Segundo estes autores, a configuração da malha de amostragem é crucial para a eficiência da estimativa do variograma experimental.

Webster & Oliver (1992), argumentam que a variância espacial estimada não estará enviesada, em relação à variância espacial que a variável regionalizada apresenta no domínio, se a amostra utilizada, para tal, não for enviesada. Estes

autores testam o número de unidades amostrais necessárias para a reprodução de um modelo de variância espacial ajustado a uma variável gaussiana normal e isotrópica, calculada por simulação estocástica em uma malha regular de 256 x 256 nós, perfazendo 65536 pontos. Para efeito de comparação, Webster & Oliver (1992), determinam limites de confiança, baseado nos desvios padrão das variâncias espaciais calculadas para cada passo. Com base nestes testes concluem que, para os exemplos utilizados, um variograma calculado a partir de 150 unidades amostrais reproduz satisfatoriamente o modelo de variograma utilizado, enquanto que um variograma calculado com 225 unidades amostrais é fidedigno ao modelo.

Conwell *et al.* (1997) testam a configuração de malhas de piezômetros no cálculo de variogramas experimentais. Estes autores, concluem que malhas regulares com configurações, de pontos amostrais, retangulares conduzem a dois problemas: um refere-se ao número de pares de pontos por classe de distâncias, que varia muito entre as classes e; o segundo é referente ao custo da aquisição de informações que utilizando esta configuração é muito cara, pois é necessário adquirir uma grande quantidade de medidas. Concluem ainda que malhas de amostragem proporcionais (*scaled sampling*) resultam em variogramas experimentais com um número uniforme de pares de pontos considerando-se uma grande amplitude de distâncias.

Silliman & Berkowitz (2000), estudando o impacto da amostragem enviesada na estimativa de variogramas experimentais de meios fraturados, procuram estabelecer como amostragens orientadas, em meios fraturados, afetam nas estatísticas espaciais de variáveis aleatórias que apresentam correlação com a orientação das fraturas. Ainda segundo estes autores, se em uma amostragem aleatória simples são amostradas zonas, ou porções, particulares do domínio e as unidades amostrais recebem pesos idênticos, as estatísticas amostrais não reproduzirão as populacionais. Ainda segundo Silliman & Berkowitz (2000), existe uma forte dependência entre o patamar de um variograma e a estratégia de amostragem utilizada e, ainda, que as amplitudes dos variogramas aparentemente não têm relação com o padrão espacial da amostragem.

Conforme o exposto, observa-se que nenhum trabalho, voltado à quantificação da incerteza associada à amostragem, foi realizado até o momento, sobretudo analisando diferentes tipos de distribuição estatística amostrados por diferentes métodos de amostragem. Assim, este trabalho tem o propósito de avaliar

sistematicamente o tamanho das amostras de uma base de dados de referência e, com esses dados, calcular as estatísticas, variogramas e recursos minerais associados, cujos resultados serão comparados àqueles da referência. A análise comparativa permitirá determinar os tamanhos das amostras que reproduzem as características estatísticas e geoestatísticas dos dados de referência.

Capítulo 3

Materiais e Métodos

3.1 Materiais

Durante a realização deste trabalho foram utilizados vários programas de computador de uso cotidiano e programas específicos de estatística e geoestatística, além daqueles desenvolvidos, pelo Professor Dr. Jorge K. Yamamoto e pelo candidato, especificamente para sanar questões peculiares a esta tese.

Os programas de estatística e geoestatística utilizados foram, o *Isatis r3.2* para a simulação da base de dados assimétrica negativa, o *GSLib* (Deutsch & Journel, 1996) na geração, por simulação estocástica, dos dados da variável assimétrica negativa e variável simétrica, além do *GeoVisual v.2.1* (Yamamoto, 2003) onde todas as análises estatística, geoestatística, de superfícies de tendência e estimativas foram realizadas. Destaca-se que a maior parte dos processamentos deste trabalho foram realizados no *GeoVisual*.

O estudo deverá ser conduzido em uma malha bidimensional representando dados sintéticos, gerados por meio de realizações de processos estocásticos. Optou-se por trabalhar com dados de referência bidimensionais, devido à facilidade de representar os resultados e, principalmente, porque o maior problema da coleta de unidades amostrais está no plano horizontal, ou seja, entre furos e não ao longo do furo (na vertical). Assim, os resultados do estudo podem ser estendidos para casos tridimensionais, considerando-se somente à distância entre furos de sonda, pois o intervalo de amostragem vertical será sempre suficiente para capturar a correlação espacial porventura existente nos dados.

Assim, o ponto de partida do trabalho será o que se denomina dados de referência. A base de dados de referência é composta por um único nível de uma malha regular tridimensional. Este nível possui espaçamento de 25x25 m contendo 1504 células, sendo 32 células ao longo do eixo X e 47 ao longo do eixo Y, sintetizada, por simulação estocástica condicional. Destaca-se que, este nível, bem como a realização selecionada da simulação estocástica, foram tomados ao acaso. Assim, o

nível escolhido apresenta uma textura geológica que será objeto de reprodução por meio da simulação de amostragem (considerando os métodos de amostragem aleatória simples, aleatória estratificada e sistemática, além do número de unidades amostrais coletadas).

As variáveis utilizadas para as simulações estocásticas condicionais, apresentam distribuições lognormal, normal e assimétrica negativa. Escolheu-se gerar os dados de referência por Simulação Seqüencial Gaussiana, pois este método garante a reprodução do histograma e do modelo de covariância espacial utilizado como parâmetro de entrada no processo estocástico.

3.2 Métodos

A Figura 2 ilustra esquematicamente os procedimentos adotados neste trabalho e envolve todos os métodos utilizados.

3.2.1 Geração dos Dados de Referência

Os dados de referência foram gerados por simulação estocástica não condicional em um modelo tridimensional de blocos com os parâmetros apresentados na Tabela 2. Como os dados foram simulados, a partir de bases de dados reais, optou-se por definir as dimensões horizontais dos blocos, baseando-as no espaçamento médio dos furos de sonda e a dimensão vertical na altura do bloco de lavra praticado na mina de ouro. Trata-se da Mina de Ouro de São Vicente, cujos dados foram cedidos gentilmente pela Mineração Santa Elina, proprietária da informação. Adicionalmente, como o objetivo do trabalho é verificar o comportamento de diferentes tipos de distribuição, submetidos a uma mesma amostragem, as duas outras distribuições foram geradas conforme o mesmo modelo de blocos adotado para a simulação da primeira variável. Conseqüentemente, foi necessário de realizar a simulação não condicional.

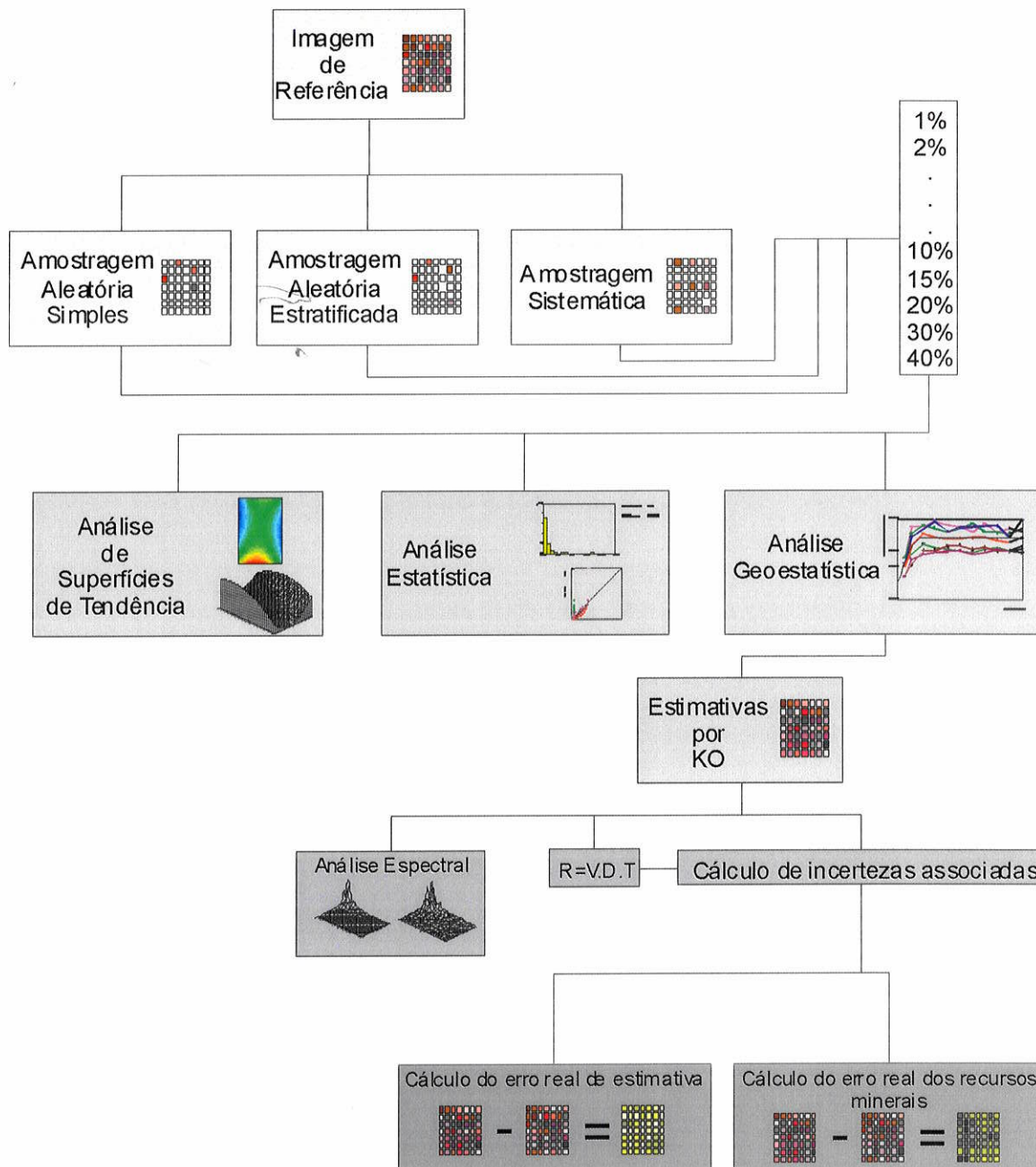


Figura 2: Fluxograma mostrando a seqüência de métodos aplicados.

Tabela 2: Parâmetros do modelo de blocos utilizado na Simulação Seqüencial Gaussiana.

	Eixo X	Eixo Y	Eixo Z
Dimensão (em metros)	25	25	5
Número de blocos	32	47	80

O método de simulação seqüencial gaussiana segue o conceito de que existe equivalência entre os resultados de uma distribuição multivariada e o resultado de uma seqüência de distribuições univariadas condicionadas a realizações univariadas (Olea, 1999).

Esquemáticamente os procedimentos para a realização da Simulação Seqüencial Gaussiana, segundo Deutsch & Journel (1996), Srivastava (1994) e Olea (1999), são:

1. realizar a transformação gaussiana dos dados; caso estes não sigam uma distribuição normal com média igual a zero e variância igual a um ($N[0,1]$);
2. proceder à análise estrutural dos dados transformados;
3. gerar a malha a ser simulada;
4. escolher aleatoriamente, da seqüência ordenada de nós, o ponto $Z^{(i)}(x_i)$ a ser simulado;
5. utilizando os resultados da análise estrutural realizar, por meio da krigagem simples, a estimativa de $Z^*(x_i)$ e calcular a respectiva variância de krigagem simples ($\sigma_{KS}^2(x_i)$) utilizando-se os pontos conhecidos (amostras) vizinhos $Z(x_n)$ mais os pontos $Z^{(i)}(x_i)$ anteriormente simulados;
6. utilizar $Z^*(x_i)$ e $\sigma_{KS}^2(x_i)$ para definir a distribuição normal com média $Z^*(x_i)$ e variância $\sigma_{KS}^2(x_i)$ ($N(Z^*(x_i), \sigma_{KS}^2(x_i))$) para retirar aleatoriamente o valor simulado $z^{(i)}(x_i)$;
7. adicionar $z^{(i)}(x_i)$ à vizinhança do próximo ponto a ser simulado;
8. refazer os passos de 5 a 6 até que o último nó da malha tenha sido simulado;
9. proceder a transformação reversa da anamorfose gaussiana.

Em termos práticos, o valor simulado $z^{(i)}(x_i)$ pela SGS pode ser escrito como sendo o valor da krigagem simples adicionado de um ruído aleatório que é calculado como a variância de krigagem simples vezes um número aleatório (ε), conforme:

$$z^{(i)}(x_0) = z_{KS}^*(x_0) + \sigma_{KS}^2 \varepsilon$$

onde ε é um número gerado aleatoriamente que em princípio segue uma distribuição uniforme entre 0 e 1 ($\varepsilon \sim U$) e posteriormente é transformado em uma distribuição gaussiana normal ($N[0,1]$).

3.2.1.1 Simulação da Distribuição Lognormal

Seguindo os passos da simulação seqüencial gaussiana, foi realizada a transformação dos dados de sondagem da mina de ouro para uma distribuição gaussiana normal ($N[0,1]$). A análise geoestatística dos dados resultou em um modelo teórico de variograma anisotrópico. Os parâmetros deste modelo de variograma são apresentados na Tabela 3. Observa-se, nesta tabela, que o modelo ajustado apresenta anisotropia geométrica na primeira estrutura e que a segunda estrutura é a componente zonal do modelo onde, nas direções horizontais, a variância espacial é maior que a da vertical. Deve-se destacar que, embora a simulação tenha sido não condicional, a transformação dos dados originais foi necessária uma vez que o método utilizado requer modelos de covariância no espaço gaussiano. Outra justificativa para a transformação é que a tabela de transformação também é utilizada na transformação reversa, pois os métodos gaussianos geram apenas distribuições gaussianas normais ($N[0,1]$). A transformação reversa faz com que a $N[0,1]$ seja transformada novamente em $Z[\mu, \sigma]$.

A simulação foi calculada com 100 realizações e como os dados foram transformados para uma distribuição $N[0,1]$, pôde-se utilizar a técnica da krigagem simples com média constante.

Tabela 3: Parâmetros do modelo de variograma ajustado aos dados de sondagem da mina de ouro.

Direção	Efeito Pepita	Estrutura 1		Estrutura 2	
		Amplitude	(C)	Amplitude	(C)
140°/0°/0°		83,00		170,00	
50°/0°/0°	0.30	83,00	0,655	146,00	0,045
140°/90°/0°		22,00		∞	

Realizadas as simulações, implementou-se a transformação reversa da variável para a distribuição original dos dados e tomou-se ao acaso um determinado nível do modelo de blocos, no caso o nível de 250m. Também, ao acaso, escolheu-se a realização de número treze para serem os dados de referência. Suas estatísticas são aquelas apresentadas na Tabela 4 e o histograma na Figura 3. O resultado da análise geoestatística, dos dados após a transformação reversa, encontra-se ilustrada na Figura 4 e os parâmetros do variograma na Tabela 5.

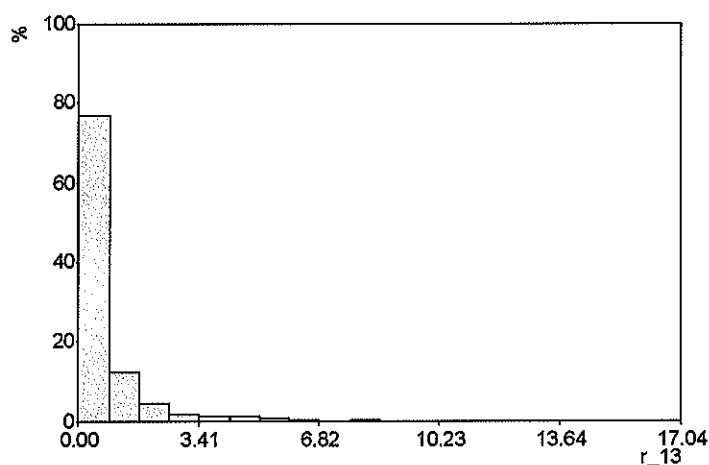


Tabela 4: Estatísticas da base de dados de referência.

Estatísticas	Resultados (ppm)
Média	0,828
Desvio Padrão	1,466
Coefficiente de Variação	1,770
Máximo	17,044
Quartil Superior	0,785
Mediana	0,376
Quartil Inferior	0,164
Mínimo	0,101

Figura 3: Histograma dos dados de referência (distribuição assimétrica positiva).

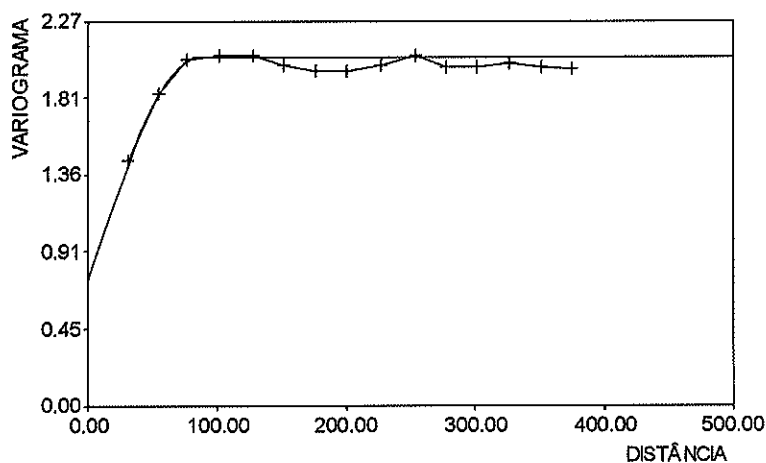


Figura 4: Variograma experimental e respectivo modelo teórico ajustado aos dados de referência (distribuição assimétrica positiva), após a transformação reversa.

Tabela 5: Parâmetros do modelo de variograma ajustado aos dados de referência, após a transformação reversa.

Direção	Efeito Pepita	Amplitude	C
Omnidirecional	0,757	85,00	1,300

Na Figura 4 observa-se que o variograma é omnidirecional. Justifica-se a utilização deste variograma, pois, pelos parâmetros definidos para a simulação condicional, todas as amostras utilizadas na simulação de um ponto possuíam distância máxima, do ponto a ser simulado, de 75 m. Logo, apenas a primeira estrutura do modelo apresentado na Tabela 3 foi utilizada para a simulação. Cabe lembrar que essa estrutura apresentou um componente de anisotropia geométrica apenas com relação à vertical, que está sendo desconsiderada na imagem de referência, e que nas direções horizontais esta estrutura é isotrópica. Assim, o modelo reproduzido pela realização da simulação é isotrópico, não sendo, portanto, necessário o cálculo de várias direções para o variograma dos dados de referência.

Comparando estes resultados àqueles do modelo de variograma fornecido para a simulação (Tabela 3), pode-se considerar que o variograma foi reproduzido.

Justifica-se a escolha de uma realização, como representativa dos dados, pelo próprio conceito de simulação estocástica, pois seu resultado pode ser considerado como uma imagem que reproduz a textura geológica. Além disso, a realização é equiprovável, ou seja, possui a mesma probabilidade de ocorrência que as demais realizações desta mesma simulação. Tem-se consciência de que esta não é a imagem mais representativa da realidade, mas sim uma imagem que possui a mesma probabilidade de ser tão realista quanto às demais.

3.2.1.2 Simulação da Distribuição Simétrica

A distribuição simétrica foi simulada conforme a malha descrita na Tabela 2. Neste caso, o modelo de variograma de entrada dos dados foi um modelo omnidirecional horizontal, pois, como explicado no item anterior, apenas o plano horizontal é o de interesse. A Tabela 6 apresenta os parâmetros do modelo teórico de variograma adotado para a realização da simulação não condicional. O modelo ajustado, aos dados transformados, foi o esférico.

Tabela 6: Parâmetros do modelo de variograma ajustado aos dados de P_2O_5 transformados.

Direção	Efeito Pepita	Amplitude	(C)
Ominidirecional	0,37	200,00	0,54

Realizada a simulação, uma realização foi tomada ao acaso no nível com cota em 250m. A realização escolhida foi a de número 52. O modelo de variograma ajustado aos dados, após a transformação reversa, pode ser observado na Figura 5. Este modelo é um modelo omnidirecional, esférico com efeito pepita de 3,59 e patamar de 5,6. A amplitude é igual a 137m.

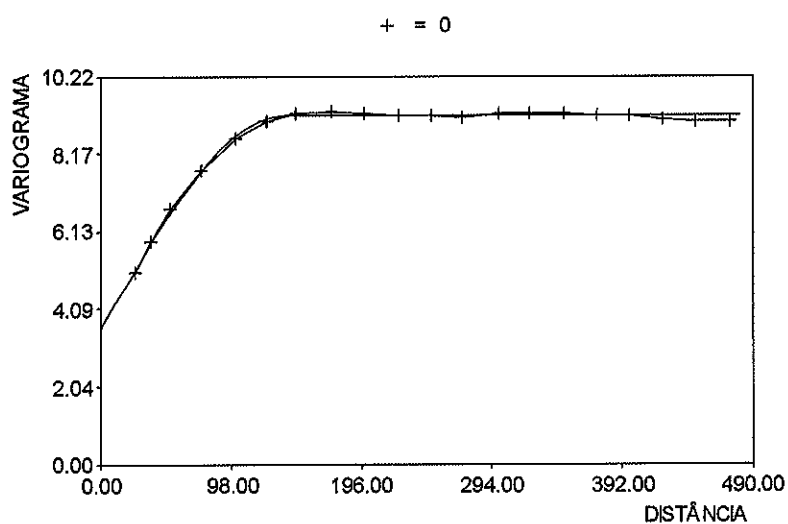


Figura 5: Variograma experimental e respectivo modelo de variograma ajustado aos dados de referência (variável simétrica), após a transformação reversa.

O histograma da Figura 6 representa a distribuição desta variável e a Tabela 7 suas estatísticas descritivas, observa-se que a distribuição apresenta uma pequena assimetria negativa, porém ela pode ser considerada perfeitamente como simétrica.

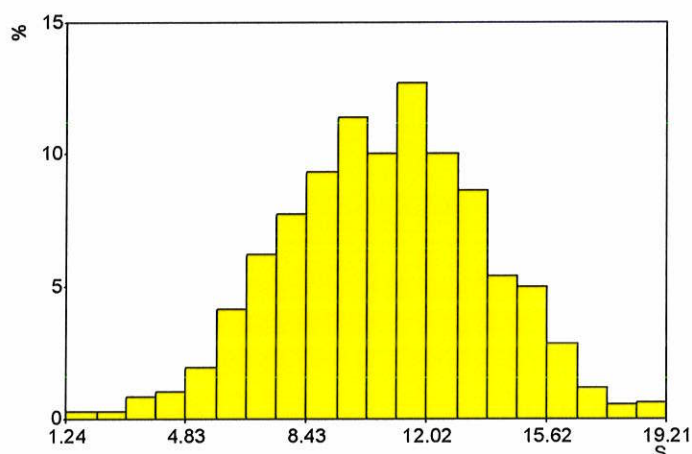


Figura 6: Histograma, para a variável simétrica, dos dados de referência.

Tabela 7: Estatísticas da base de dados de referência, variável simétrica.

Estatísticas	Resultados (%)
Média (%)	10,784
Desvio Padrão	3,012
Coeficiente de Variação	0,279
Máximo	19,212
Quartil Superior	12,837
Mediana	10,796
Quartil Inferior	8,743
Mínimo	1,236

3.2.1.3 Simulação da Distribuição Assimétrica Negativa

Do mesmo modo que as variáveis anteriores, esta distribuição foi gerada por simulação estocástica não condicional. O modelo de variograma fornecido para a simulação possui os parâmetros apresentados na Tabela 8. Assim como para a variável simétrica, o variograma experimental foi calculado como omnidirecional e a ele foi ajustado um modelo esférico com apenas uma estrutura. A simulação foi calculada com 100 realizações, das quais a 23ª realização foi selecionada no nível de 250 m para compor a base de dados de referência.

Tabela 8: Parâmetros do modelo de variograma ajustado aos dados de Fe, transformados em uma distribuição Gaussiana normal.

Direção	Efeito Pepita	Amplitude	(C)
Ominidirecional	0,449	160,00	0,5624

Esta realização apresenta uma distribuição com o histograma mostrado na Figura 7 e as estatísticas descritivas da Tabela 9. Note-se que há a predominância de valores relativamente altos e ainda que comparado à média, o desvio padrão é extremamente baixo.

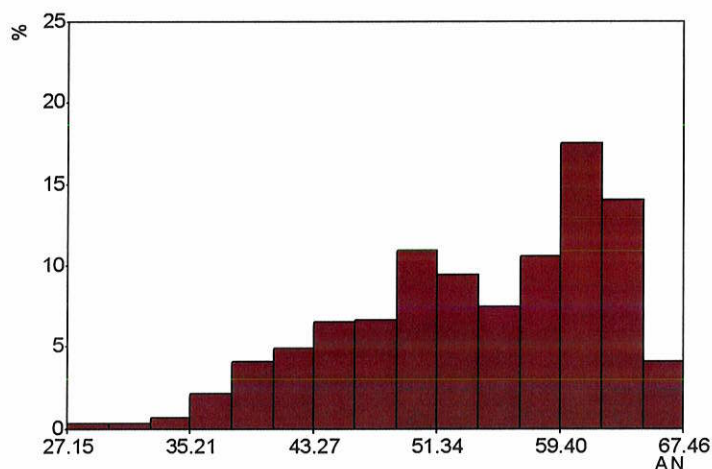


Figura 7: Histograma, da variável assimétrica negativa, dos dados de referência.

Tabela 9: Estatísticas da base de dados de referência.

Estatísticas	Resultados (%)
Média	54,124
Desvio Padrão	8,241
Coefficiente de Variação	0,152
Máximo	67,462
Quartil Superior	61,191
Mediana	56,296
Quartil Inferior	48,434
Mínimo	27,150

Nesta realização, após a transformação reversa, o modelo de variância espacial ajustado foi o esférico e seus parâmetros encontram-se especificados na Tabela 10. A Figura 8 apresenta este modelo.

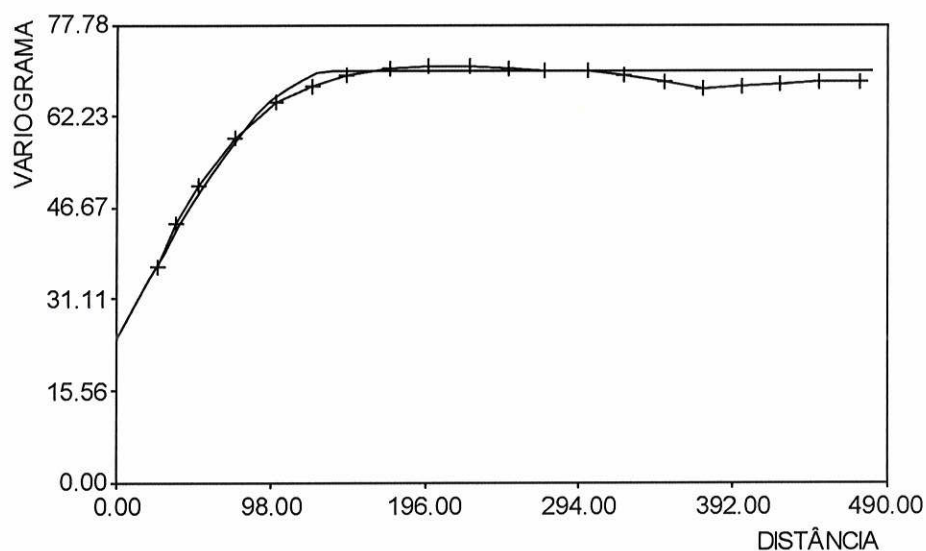


Figura 8: Variograma experimental e respectivo modelo teórico ajustado aos dados de referência (variável assimétrica negativa), após a transformação reversa.

Tabela 10: Parâmetros do modelo de variograma ajustado aos dados de Fe.

Direção	Efeito Pepita	Amplitude	(C)
Omnidirecional	24,5	137	45,5

Após a simulação de cada tipo de distribuição, as três variáveis foram integradas em um único arquivo para que se iniciasse o processo de re-amostragem.

3.3 Re-amostragem dos Dados de Referência

Gerada a base de dados de referência realizou-se a re-amostragem, segundo os níveis percentuais apresentados na Tabela 11.

Tabela 11: Níveis percentuais de retirada de amostras e respectivas quantidades de unidades amostrais coletadas.

Nível percentual	Número de amostras	Nível percentual	Número de amostras
1%	15	9%*	135
2%	30	10%	150
3%	45	15%	225
4%	60	20%	300
5%	75	30%	451
6%	90	40%	601
7%*	105	50%**	752
8%	120		

* Na amostragem aleatória estratificada e na sistemática estes níveis percentuais não foram calculados devido ao algoritmo implementado para o programa destas amostragens.

** Este nível percentual foi calculado apenas para a amostragem sistemática.

Segundo Yamane (1967), existem dois grupos de amostragem, o probabilístico e o não probabilístico. A amostragem não probabilística é aquela onde não se pode determinar objetivamente a probabilidade associada ao processo de retirada de unidades amostrais. Por consequência, não se pode determinar a segurança do resultado da amostra em termos de probabilidade. Ainda segundo Yamane (1967), a amostragem não probabilística pode ser ilustrada pela amostra de julgamento (*judgment sample*); amostra de quota (*quota sample*) e questionário por correspondência (*mail questionnaire*). Destaca-se que a amostragem não probabilística foi extensivamente utilizada nas décadas de 20 e 30 (Yamane, 1967).

Apenas a título de ilustração, estes três procedimentos de amostragem serão brevemente descritos. A amostragem por julgamento é feita com base na seleção de amostras, tidas como representativas, por um especialista, esta representatividade é assumida com base na experiência do especialista (Yamane, 1967). Barnett (1991)

destaca, ainda, que neste tipo de amostragem a escolha de amostras é subjetiva e deliberada pelo que o especialista considera uma amostra representativa. Tanto Yamane (1967) quanto Barnett (1991), destacam que esta amostragem pode produzir resultados muito bons e tal qualidade, como foi dito, dependerá exclusivamente da experiência do responsável por ela.

A amostragem por quotas, segundo Yamane (1967), é uma variação da amostragem por julgamento, onde as instruções explícitas são dadas em relação ao que deve ser selecionado. Barnett (1991), afirma que este procedimento é mais estruturado que o anterior.

A amostragem via questionário por correspondência, é um procedimento de amostragem considerado probabilístico ou não probabilístico, dependendo apenas do método de seleção de amostras; por exemplo, se a amostra é selecionada de maneira similar à amostragem por quotas, que é baseada na amostragem não probabilística, ela será não probabilística (Yamane, 1967).

Segundo Yamane (1967), a amostragem probabilística é um processo de retirada de amostras, onde cada indivíduo da amostra é retirado com probabilidade conhecida. Krumbein & Graybill (1965) afirmam que a amostragem probabilística é uma classe genérica de amostras que incluem um procedimento aleatório no plano de amostragem. Ainda segundo estes autores, o termo amostragem probabilística remete-se especificamente a um procedimento formal de seleção de uma ou mais amostras de uma população de modo que cada indivíduo ou unidade amostral na população tenha uma chance conhecida de ser selecionada na amostra.

Como visto, a amostragem probabilística implica no conhecimento da probabilidade associada à retirada de amostras. Deste modo, antes de abordar os métodos de amostragem cabe um breve resumo de alguns aspectos probabilísticos e conceitos estatísticos.

- Probabilidade

Define-se probabilidade como sendo a chance de ocorrência de um determinado evento em um experimento aleatório. Para a determinação da probabilidade de um evento aleatório é necessário o conhecimento preliminar de como e quantas amostras podem ser selecionadas em uma população. Segundo Yamane (1967), existem

basicamente dois modos de seleção de amostras, com e sem reposição. Ainda segundo este autor, o número de amostras possíveis irá variar de acordo com o modo de seleção utilizado.

Para ilustrar o número de amostras possíveis tome-se o exemplo que se segue (modificado de Yamane, 1967).

Sejam A, B e C três possíveis resultados de um dado evento, deve-se então, retirá-los dois a dois com reposição. Sabendo-se que o número de combinações possíveis é N^n então:

$$\begin{matrix} N = 3 \\ n = 2 \end{matrix} \Rightarrow N^n = 3^2 = 9$$

Ou seja, existem 9 modos diferentes de combinar os três resultados dois a dois.

Para o caso de retirada sem reposição, sabe-se que existem as combinações possíveis, conforme:

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

logo no exemplo ter-se-á:

$$\binom{N}{n} = \binom{3}{2} = \frac{3!}{2!(1)!} = \frac{3 * 2!}{2!} = 3$$

Os esquemas matriciais, a seguir, mostram os resultados das combinações com reposição e sem reposição. A retirada com reposição será:

	A	B	C
A	AA	AB	AC
B	BA	BB	BC
C	CA	CB	CC

caso não haja reposição:

	A	B	C
A	--	AB	AC
B	BA	--	BC
C	CA	CB	--

considerando-se a retirada onde a ordem dos resultados é importante ter-se-á:

	A	B	C
A	--	AB	AC
B	--	--	BC
C	--	--	--

Segundo Botter *et al.* (1996), a probabilidade de um evento A , associado a um experimento aleatório com espaço amostral S , será:

$$P(A) = \frac{\text{número de elementos de } A}{\text{número de elementos de } S}$$

onde o número de elementos de A é o número de resultados favoráveis ao evento A e o número de elementos de S é o número de resultados possíveis do experimento, considerando-se experimentos aleatórios e equiprováveis. Estes autores enumeram, ainda, algumas propriedades da probabilidade, como segue:

1. $0 \leq P(A) \leq 1$ qualquer que seja o evento A ,
2. $P(\emptyset) = 0$ e $P(S) = 1$,
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ quaisquer que sejam os eventos A e B ,
4. $P(A \cup B) = P(A) + P(B)$ se A e B forem eventos mutuamente exclusivos,
5. $P(\bar{A}) = 1 - P(A)$ para qualquer evento A ,

- Esperança matemática (média)

Segundo Yamane (1967), a esperança matemática de uma variável aleatória deve ser definida como: "Seja X_i uma variável aleatória com valores possíveis x_1, \dots, x_k e probabilidades p_1, \dots, p_k , respectivamente. Então o valor esperado de X_i é definido

como: $E[X_i] = \sum_{i=1}^n x_i p_i$ "

Sendo n o número total dos valores de X_i , então para ter a esperança amostral \bar{x} , pode-se escrever:

$$E[X_i] = \frac{1}{n} \sum^n x_i$$

onde n é o tamanho da amostra.

Como n é o tamanho da amostra, deve-se encontrar todos os valores possíveis de \bar{x} . Para que isto ocorra, é imperativo conhecer todas as possíveis amostras de tamanho n que podem ser selecionadas da população de tamanho N .

Yamane (1967) ressalta que, dada uma população de N valores de x_i e quando p_i não é dado, deve-se assumir que X_i tem a mesma possibilidade de ocorrência, ou seja:

$$p_i = \frac{1}{N}$$

e para outros possíveis valores de X_i que não ocorrem, deve-se estabelecer que $p_i=0$. O resultado destas inferências é que a soma de todos os valores possíveis de X_i é equivalente à soma dos N valores da população.

- Variância

A variância de dada população é uma medida de dispersão dos valores, que ela pode assumir, em torno da sua média. Esta medida é calculada como:

$$Var[X_i] = E[X_i^2] - E[X_i]^2$$

Revistos os conceitos estatísticos, segue-se para os métodos de amostragem probabilística.

Segundo Landim (1997), para que se possa estimar o comportamento de potenciais observações de uma população, é necessário que se retire amostras em subconjuntos que garantam que cada amostra tenha a mesma chance, ou probabilidade, de ser escolhida. Este autor cita, ainda, diversos métodos de planos de amostragem, dentre os quais destacar-se-ão:

3.3.1 Amostragem Aleatória Simples

Esta amostragem é a mais simples de ser realizada, pois requer apenas que se numere todos os elementos da população (N) e através de uma tabela de números aleatórios, ou pela geração computacional de números pseudo-aleatórios, escolhem-se

os n números que comporão a nova base amostral. Goes *et al.* (1991) destacam que, esta amostragem é normalmente utilizada quando se dispõe de pouca informação sobre o material a ser amostrado. Cochran (1963) cita que este tipo de amostragem garante que todas as amostras tenham a mesma probabilidade de serem retiradas da população.

Segundo Kim (1990), este tipo de amostragem tem como características prover estimadores não enviesados e, ainda, que os limites de confiança podem ser calculados para a média, daí a importância em se conhecer a média amostral e sua variância. Ainda segundo esse autor, se a variância populacional for conhecida, ou puder ser aproximadamente inferida, o tamanho da amostra (n) para um determinado limite de erro pode ser calculado antes da amostra ser efetivamente coletada.

Apresentar-se-á a seguir um possível esquema de como a amostragem aleatória simples pode ser realizada. No domínio, que se deseja amostrar, deve-se gerar uma malha regular, associa-se a cada centro de célula um valor numérico que é armazenado em um vetor. Sorteiam-se aleatoriamente posições deste vetor, cada posição corresponde a um número que indicará a posição das amostras individuais.

Outro possível esquema de amostragem aleatória simples, que foi o adotado neste trabalho, é semelhante ao descrito, porém a cada centro de célula são associados dois números que são armazenados em dois vetores distintos. Estes números quando combinados, dois a dois, darão a posição da célula em uma matriz bidimensional. Em seguida, sorteiam-se aleatoriamente posições nestes dois vetores, gerando pares ordenados que indicarão a posição da célula na matriz. Este procedimento garante maior representatividade espacial da amostragem, principalmente no caso em estudo em que a malha é bidimensional.

3.3.2 Amostragem Aleatória Estratificada

Quando há mistura de populações, a amostragem aleatória estratificada resulta em estimadores com menor variância do que os obtidos com a amostragem aleatória simples (Kim, 1990).

Segundo Landim (1997), esta amostragem é utilizada quando se supõe uma grande variabilidade nas observações. Neste tipo de amostragem, primeiro divide-se a população em subpopulações e cada uma destas é amostrada por amostragem

aleatória simples. As subpopulações devem possuir tamanhos constantes (Gy, 1992). Ainda segundo esse autor, esta amostragem é a mais segura quando flutuações cíclicas no atributo da variável são esperadas ou previsíveis.

Quando a variância dentro das populações for menor que a variância entre as populações a amostragem estratificada resulta em melhores estimadores do que os obtidos por amostragem aleatória simples (Kim, 1990).

Descreve-se, a seguir, um esquema de amostragem aleatória estratificada. Observe-se que a diferença deste sistema para a amostragem aleatória simples é que, naquele o domínio deve ser dividido em subdomínios e as amostras são coletadas aleatoriamente dentro de cada subdomínio. Neste tipo de amostragem, a seleção de amostras individuais pode ser proporcional ao tamanho do subdomínio.

Assim como na amostragem aleatória simples, o primeiro passo para realizar este tipo de amostragem é a geração de uma malha regular no domínio que se deseja estudar. O que diferencia esta amostragem da anterior é a divisão do domínio em subdomínios e para as células contidas em cada subdomínio anexa-se um vetor numérico. A escolha das amostras é feita pelo sorteio aleatório dos números do vetor em cada subdomínio.

No presente trabalho, a amostragem aleatória estratificada foi realizada utilizando estratos de tamanhos variáveis, ou seja, à medida que o n das amostras cresce o tamanho do estrato diminui. Por exemplo, para as amostras com n igual a 75 foram gerados 70 estratos com tamanho constante dos quais foi retirado um indivíduo por amostragem aleatória simples.

3.3.3 Amostragem Sistemática

Segundo Goes *et al.* (1991), a amostragem sistemática é aquela onde amostras geológicas são coletadas a intervalos regulares definidos *à priori*. Esta amostragem é implementada se definindo um padrão predeterminado, por exemplo, quadrados, e o ponto central destes quadrados são amostrados segundo um incremento preestabelecido.

A aleatoriedade da amostragem regular pode ser garantida escolhendo-se o ponto inicial da amostragem ao acaso. O arranjo espacial da população (no caso de variáveis regionalizadas) pode afetar a amostragem sistemática. Por exemplo, se os

elementos da população apresentam arranjo aleatório, o resultado da amostragem sistemática será igual àquele da amostragem aleatória simples. Porém, se os elementos da população estiverem ordenados, a amostragem sistemática será melhor que a aleatória simples e, ainda, se houver entre os elementos da população um arranjo cíclico ou periódico, os resultados da amostragem sistemática serão piores que aqueles para a amostragem aleatória simples (Kim, 1990).

3.4 Análise Estatística

Esta análise contou com o cálculo das estatísticas descritivas, quais sejam: média, variância e desvio padrão, coeficiente de variação, os valores de máximo e mínimo, quartis superior e inferior e mediana.

Como o intuito é comparar o resultado dos diversos níveis de amostragem com os dados de referência e observar o quanto as novas distribuições afastam-se da distribuição de referência calculou-se também diagramas do tipo Q-Q (quantil-quantil) e P-P (probabilidade-probabilidade).

3.4.1 Diagramas Q-Q e P-P

Estes diagramas são utilizados para comparar as similaridades entre duas ou mais distribuições.

a. Diagrama do tipo Q-Q

Neste diagrama, os quantis da variável de referência são representados no eixo das abscissas e os quantis da variável observada no eixo das ordenadas.

Neste gráfico é comum plotar a reta que marca a primeira bissetriz. Quanto maior a similaridade entre as distribuições analisadas, os pontos estarão mais próximos da primeira bissetriz.

b. Diagrama do tipo P-P

Semelhante ao gráfico Q-Q, este gráfico também compara duas ou mais variáveis, com a diferença que utiliza a probabilidade acumulada de ocorrência de cada valor dentro da distribuição. No eixo das abscissas são lançadas as probabilidades acumuladas da variável de referência e no eixo das ordenadas as probabilidades acumuladas da variável observada. A reta da primeira bissetriz também é plotada neste gráfico e quanto mais próximos estiverem os pontos desta bissetriz, maior a similaridade entre as distribuições estudadas.

3.5 Análise Geoestatística

O principal resultado da análise geoestatística é o modelo de variograma, que nada mais é do que um gráfico onde as variâncias espaciais dos dados são determinadas em função das distâncias que separam dois pontos. Os pontos do variograma experimental são calculados conforme:

$$2\gamma(h) = E\{[Z(x) - Z(x+h)]^2\}$$

onde: $2\gamma(h)$ é o valor do variograma para uma distância h (função variograma); $E[Z(x)]$ é a esperança matemática de ocorrência do atributo da variável no ponto x ; $E[Z(x+h)]$ é a esperança matemática de ocorrência do atributo da variável no ponto $(x+h)$.

Considerando-se a função variograma, em termos computacionais, pode-se escrevê-la como:

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

O variograma experimental pode ser derivado a partir da variância $C(0)$ e da covariância $C(h)$, como segue:

$$\gamma(h) = C(0) - C(h)$$

onde $\gamma(h)$ é a função semivariograma.

Para calcular um variograma experimental deve-se estabelecer alguns parâmetros, quais sejam: direção, tolerância angular da direção, largura máxima,

passo, tolerância do passo e número de passos. São estes os parâmetros de uma malha irregular.

Para uma malha de amostragem regular, escolhe-se apenas a direção, que já é predefinida segundo a topologia da malha. O passo é a abertura da malha na direção escolhida e o número de passos deve ser a metade do número de nós na direção escolhida.

O resultado final da análise geoestatística é o modelo teórico de variograma que é uma função matemática linear, cujo ajuste é obtido de modo iterativo.

Para poder utilizar o modelo de variograma, deve-se assumir que este seja válido para todo o domínio dos dados e ainda que a variância espacial dos dados seja função apenas da direção e da distância que separa duas amostras, isto é, deve-se assumir a Hipótese Intrínseca ou Hipótese de Quasi-estacionariedade. A Hipótese Intrínseca garante que a variância espacial obtida para dois pontos, separados por uma distância h , será válida para todo o domínio, desde que se observe àquela distância e também a direção utilizada para a tomada dos pontos. Segundo Huijbregts (1975), o variograma e seu respectivo modelo permitem descrever quantitativamente a variação no espaço de um fenômeno regionalizado.

A vantagem da geoestatística, sobre os demais métodos de interpolação, é poder se conhecer e definir diferenças de comportamento da variável conforme a direção analisada, ou seja, é poder definir se os dados apresentam anisotropia e utilizá-la na estimativa de pontos desconhecidos dentro do domínio estudado.

3.6 Krigagem Ordinária

Segundo Armstrong (1998), a krigagem é um método de estimativa que leva em consideração uma série de fatores, tais como:

- * número de amostras e a qualidade dos dados em cada ponto;
- * posição espacial das amostras no depósito;
- * distância entre as amostras e os pontos ou blocos a serem estimados;
- * continuidade espacial das variáveis.

O estimador da krigagem ordinária é calculado como:

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

onde: $z^*(x_0)$ é o valor estimado no ponto x_0 ; λ_i é o ponderador do i -ésimo ponto amostrado; $z(x_i)$ é a i -ésima amostra utilizada na estimativa de $z^*(x_0)$.

A condição de restrição imposta aos pesos da krigagem ordinária é:

$$\sum_{i=1}^n \lambda_i = 1$$

Minimizando-se a variância de estimativa, sujeita à condição de restrição dos pesos, obtém-se o sistema de equações de krigagem ordinária, que em termos matriciais pode ser escrito como:

$$\begin{bmatrix} \gamma(x_1, x_1) & \cdots & \gamma(x_1, x_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(x_n, x_1) & \cdots & \gamma(x_n, x_n) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma(x_1, x_0) \\ \vdots \\ \gamma(x_n, x_0) \\ 1 \end{bmatrix}$$

3.7 Avaliação de Recursos/reservas Minerais

Uma vez realizada a estimativa por krigagem ordinária, conhece-se o teor em todas as células da malha regular. Como as dimensões de cada célula são conhecidas, a aplicação da equação básica do cálculo de recursos/reservas ($R=VDT$) torna-se trivial.

Para fins de comparação de recursos geológicos dos dados de referência com os dados amostrais, procedeu-se à delimitação daqueles conforme a fronteira convexa estabelecida para as amostras. Este cuidado evita o erro geométrico causado por diferenças de áreas entre as imagens de referência e amostrada, pois esta última, especialmente para baixos níveis de percentuais, tenderá a apresentar uma área menor que a da referência.

3.7.1 Determinação da Fronteira Convexa dos Dados

Fronteira convexa é definida, segundo Sedgewick¹ (1983, *apud* Yamamoto, 1997), como sendo o menor polígono convexo que contenha a totalidade dos pontos amostrais em seu interior.

A determinação de uma fronteira convexa, a dados amostrais, na interpolação de valores, tem por objetivo evitar a interpretação de dados espúrios criados por extrapolação matemática (Yamamoto, 1997).

No presente trabalho a fronteira convexa dos dados é calculada pelo algoritmo desenvolvido por Yamamoto (1997). Este algoritmo é resumidamente descrito como segue (Yamamoto, 1997):

1. calcula-se o centróide do conjunto de dados;
2. determina-se os ângulos entre os pontos amostrais e o centróide calculado em 1;
3. verifica-se os extremos dos dados, em termos de coordenadas, e os marca com marcadores (*flags*), estes marcadores são na realidade uma variável booleana que recebe valor verdadeiro quando o ponto é extremo;
4. determina-se o menor ângulo, calculado em 2, dentre os pontos extremos;
5. ordena-se os pontos em ordem crescente de ângulos;
6. inicia-se a verificação nos dados ordenados, dois a dois, estabelecendo se este par de pontos pertence à fronteira. O par de pontos pertencerá à fronteira se, e somente se, houver pontos apenas à direita do vetor formado entre os dois pontos.

3.8 Análise de Superfícies de Tendência

Segundo Davis (1986), a análise de superfícies de tendência é realizada com o objetivo de separar dados, distribuídos em mapas, em dois componentes distintos, um de natureza regional e outro composto por flutuações locais. Segundo Harbaugh & Merriam (1968), a análise de superfícies de tendência invariavelmente envolve o ajuste de superfícies satisfazendo o critério dos mínimos quadrados. Desta feita, a superfície

¹ Sedgewick, R., 1983, *Algorithms*: Addison-Wesley, Reading, Massachusetts, 551 p.

ajustada aos dados deve ser tal que a soma dos quadrados das diferenças, entre a superfície e as observações, seja mínima.

Segundo Landim (1997), o componente regional é a própria superfície ajustada e o outro componente, representado pelas flutuações locais, é dado pelos valores residuais. Davis (1986), define tendência como uma função linear das coordenadas de um conjunto de observações, determinada de tal modo que o quadrado dos desvios seja minimizado. Ainda segundo este autor, desmembrando-se esta definição pode-se dizer que:

- uma observação é considerada como sendo, em parte, função de sua localização;
- a tendência sendo uma função linear nos coeficientes, ela assume a forma

$$F = \sum_{i=0}^m \sum_{j=0}^{m-i} c_{ij} X^i Y^j \quad (\text{segundo a notação utilizada em Yamamoto, 1998}) \text{ onde}$$

c_{ij} são os coeficientes do polinômio; X e Y são as coordenadas leste e norte, respectivamente; F é o valor da observação no ponto de coordenadas (X, Y) ; e

- a soma dos quadrados dos desvios remeteria à variância amostral se estes desvios fossem da observação com relação à média, assumindo a tendência como uma função média, pode-se considerá-la como uma função possuidora da menor variância possível (Davis, 1986).

O número de coeficientes (nc) de um polinômio bivariado de grau m é dado por (Yamamoto, 1998):

$$nc = \frac{(m+1)(m+2)}{2}$$

Yamamoto (1998), ilustra o cálculo dos coeficientes com um polinômio de grau 2, conforme a equação expandida:

$$F = c_{00} + c_{01}Y + c_{02}Y^2 + c_{10}X + c_{11}XY + c_{20}X^2$$

os coeficientes serão, então, obtidos a partir da resolução de um sistema de equações normais, conforme:

$$\begin{bmatrix} n & \sum Y & \sum Y^2 & \sum X & \sum XY & \sum X^2 \\ \sum Y & \sum Y^2 & \sum Y^3 & \sum XY & \sum XY^2 & \sum X^2Y \\ \sum Y^2 & \sum Y^3 & \sum Y^4 & \sum XY^2 & \sum XY^3 & \sum X^2Y^2 \\ \sum X & \sum XY & \sum XY^2 & \sum X^2 & \sum X^2Y & \sum X^3 \\ \sum XY & \sum XY^2 & \sum XY^3 & \sum X^2Y & \sum X^2Y^2 & \sum X^3Y \\ \sum X^2 & \sum X^2Y & \sum X^2Y^2 & \sum X^3 & \sum X^3Y & \sum X^4 \end{bmatrix} \cdot \begin{bmatrix} c_{00} \\ c_{01} \\ c_{02} \\ c_{10} \\ c_{11} \\ c_{20} \end{bmatrix} = \begin{bmatrix} \sum F \\ \sum FY \\ \sum FY^2 \\ \sum FX \\ \sum FXY \\ \sum FX^2 \end{bmatrix}$$

Ainda segundo Yamamoto (1998), o ajuste de superfícies a partir de polinômios de grau elevado apresenta, além de problemas de ordem computacional, problemas de precisão das superfícies interpoladas.

3.8.1 Atitude de Superfícies de Tendência de Grau 1

Este item baseia-se no algoritmo para o cálculo da atitude (azimute, rumo do mergulho e mergulho) de superfícies de tendência de grau 1, desenvolvido por Yamamoto (2003)². A Figura 9 ilustra os elementos que definem a atitude de um plano: azimute (φ), rumo do mergulho (ω) e mergulho (θ).

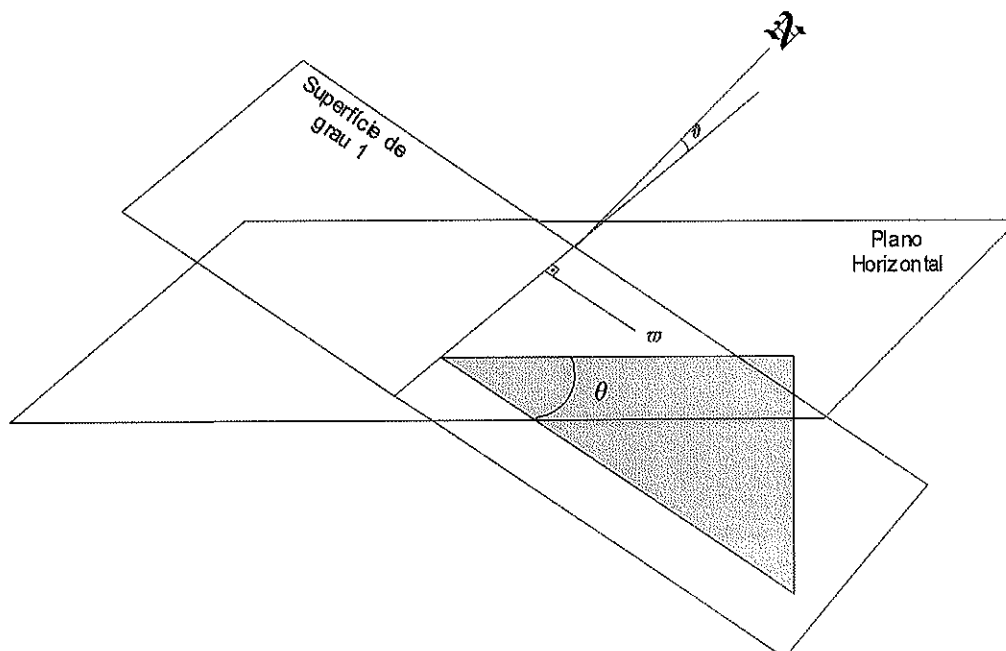


Figura 9: Atitude da superfície de tendência de grau 1 (modificado de Marshak, 1998).

² Comunicação pessoal.

Conhecidos os valores de mínimo e máximo das coordenadas de uma superfície, pode-se calcular o ponto médio de cada eixo (\bar{x} no eixo x e \bar{y} no eixo y) e, deste modo, obter o baricentro do plano. Deve-se calcular, ainda, o comprimento (d) do eixo X , como:

$$d = x_{\max} - x_{\min}$$

A cota (F^*) no baricentro do plano pode ser calculado:

$$F^* = c_{00} + c_{10}\bar{x} + c_{01}\bar{y}$$

onde \bar{x} e \bar{y} são as coordenadas do baricentro do plano e c_{ij} são os coeficientes do polinômio.

Com base nos coeficientes da equação do plano, pode-se calcular os coeficientes da reta de intersecção entre a superfície ajustada e um plano horizontal passando pelo baricentro. Sendo a equação da reta igual a $y = a + bx$, os coeficientes a e b serão:

$$a = \frac{(F^* - c_{00})}{c_{01}} \text{ e } b = -\frac{c_{10}}{c_{01}}$$

Obtida a equação da reta, o azimute (φ) da direção do plano é:

$$\varphi = \arctan(b)$$

como este ângulo é medido no sistema trigonométrico, ele deve ser convertido para o sistema azimutal, conforme:

$$\begin{cases} \varphi = |\varphi| + \pi/2 & \text{se } \varphi < 0 \\ \varphi = \pi/2 - \varphi & \text{se } \varphi > 0 \end{cases}$$

Como o rumo do mergulho não é conhecido, devem-se calcular dois ângulos, como segue:

$$\begin{cases} \omega_1 = \varphi - \pi/2 & \text{se } \omega_1 > 0 \\ \omega_1 = (\varphi - \pi/2) + 2\pi & \text{se } \omega_1 < 0 \end{cases}$$

e

$$\omega_2 = \varphi + \pi/2$$

Para a definição do caimento do rumo do mergulho as coordenadas de dois pontos devem ser calculadas. Estes pontos podem ser obtidos como a projeção do comprimento (d) a partir do baricentro, conforme ω_1 e ω_2 , ou seja:

$$x_1 = \bar{x} + d \cdot \text{sen}(\omega_1) \text{ e } y_1 = \bar{y} + d \cdot \text{cos}(\omega_1)$$

$$x_2 = \bar{x} + d \cdot \text{sen}(\omega_2) \text{ e } y_2 = \bar{y} + d \cdot \text{sen}(\omega_2)$$

As coordenadas destes pontos são utilizadas para a estimativa dos valores de F_1^* e F_2^* , pela resolução de:

$$F_1^* = c_{00} + c_{10}x_1 + c_{01}y_1 \text{ e } F_2^* = c_{00} + c_{10}x_2 + c_{01}y_2$$

Se F_1^* for menor que o valor de F^* , o rumo do mergulho será ω_1 ou se F_2^* for menor que F^* , o rumo do mergulho será ω_2 .

O ângulo do mergulho é calculado com base nas relações de um triângulo retângulo formado pela diferença entre F_1^* e F_2^* e a distância que separa os dois pontos. Deste modo, o ângulo do mergulho é:

$$\theta = \arcsen\left(\frac{F_1^* - F_2^*}{\sqrt{(F_1^* - F_2^*)^2 + (2d)^2}}\right)$$

Segundo Landim (1997), duas superfícies de tendência de grau um, provenientes de um mesmo domínio, podem ser comparadas por meio de duas medidas derivadas da direção e mergulho do plano. Estas medidas são feitas utilizando os co-senos entre os vetores que representam as superfícies, conforme (Landim, 1997):

$$DC_d = \cos|\varphi_1 - \varphi_2| \text{ e } DC_m = \cos|\theta_1 - \theta_2|$$

que irão variar de zero, no caso de perfeita correlação entre planos, a um, para a ortogonalidade.

3.9 Análise Espectral

Segundo Harbaugh & Merriam (1968), muitos fenômenos e processos geológicos são oscilatórios e a análise harmônica é uma das técnicas para analisar e sintetizar estes fenômenos. Ainda segundo estes autores, o objetivo geral da análise harmônica é a representação de um fenômeno por uma função ou funções que oscilam de modo periódico. Outro objetivo é distinguir quais componentes da observação são devidos a oscilações periódicas e quais são devidos a componentes aleatórios.

Existem diversas formas de realizar a análise harmônica e, segundo Harbaugh & Merriam (1968), a forma mais simples é a relacionada a métodos de suavização

aplicados a dados que podem ser representados como curvas. Dentre estes métodos, os autores citam a análise harmônica com séries de Fourier, a análise de autocovariância e a análise espectral.

Rayner (1971) define espectro como sendo a característica de qualquer fenômeno, quando estas características são ordenadas de acordo com comprimentos de onda crescentes ou decrescentes.

A análise espectral, segundo Davis (1986), é a partição da variação de uma série temporal em componentes de acordo com a duração ou comprimento dos intervalos nos quais a variação ocorre. Ainda segundo este autor, esta divisão é feita considerando-se a série temporal como sendo a soma de muitas séries temporais mais simples que têm a forma de senóides regulares com diferentes amplitudes, comprimentos de onda e pontos de início (fase).

Discussões de engenharia sobre a análise de Fourier sempre assumem que um sinal varia no tempo, porém o fato dos engenheiros examinarem o sinal como um fenômeno espacial no osciloscópio indica que o tempo e o espaço podem ser considerados equivalentes (Davis, 1986).

Séries de Fourier, segundo Harbaugh & Merriam (1968), são séries contendo termos de senos e co-senos que permitem representar formas complexas por meio da soma de termos senóides e co-senóides. Deste modo, ondas complexas ou qualquer outro tipo de dados que possam ser representados como uma onda, podem ser representadas como a soma de várias funções de senos e co-senos.

Basicamente, as séries de Fourier fornecem um modo de separar uma curva (por séries simples de Fourier) ou uma superfície (por séries duplas de Fourier) em um número de harmônicos simples (Harbaugh & Merriam, 1968). Um harmônico é um ciclo por intervalo básico e um intervalo básico é a frequência da onda (Davis, 1986).

3.9.1 Matemática das Séries de Fourier

Antes de definir matematicamente as séries de Fourier, cabem definições dos componentes de uma onda (Davis, 1986):

- comprimento de onda (λ): é a distância entre um ponto, em uma onda, até um ponto equivalente na onda seguinte;

- frequência (f): é o recíproco do comprimento de onda, ou seja $f = \frac{1}{\lambda}$. A frequência é o número de ondas em uma unidade de comprimento ou tempo;
- período: é o intervalo de tempo necessário para um sinal regular e cíclico se repetir, este termo é equivalente ao comprimento de onda com a diferença que o período é medido em unidade de tempo;
- amplitude (A): é a metade da altura entre uma crista e um vale da onda;
- fase ou ângulo de fase (ϕ): é a distância (em radianos) entre o início da crista ou vale da onda e sua origem, assumindo-se que uma única onda se completa em um intervalo básico (Rayner, 1971)

Estes componentes podem ser observados na Figura 10.

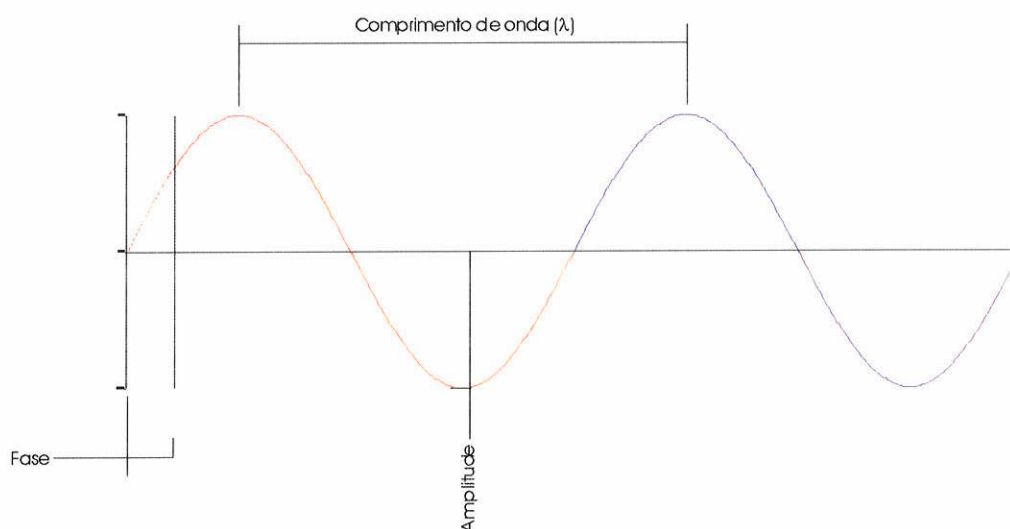


Figura 10: Componentes de uma onda.

Davis (1986) destaca que diferentes séries de ondas podem ser obtidas, alterando-se a amplitude, o comprimento de onda e o ângulo de fase. Pode-se ainda, a partir de uma senóide, deslocando-se sua origem para 90° , ou $\pi/2$, obter uma co-senóide. Por conveniência, algumas relações serão definidas para uma função co-seno. A equação de uma co-senóide é (Davis, 1986):

$$y = \cos(\theta)$$

Considerando-se alterar todos os componentes de uma onda, a equação da co-senóide pode ser escrita como (Davis, 1986):

$$Y_k = A_k \cos(k\theta - \phi_k)$$

aplicando-se a relação trigonométrica:

$$\cos(R - S) = \cos(S) \cdot \cos(R) + \sin(S) \cdot \sin(R)$$

a equação de Y_k pode ser reescrita:

$$Y_k = A_k \cos(\phi_k) \cos(k\theta) + A_k \sin(\phi_k) \sin(k\theta)$$

Observe-se que o índice k representa o número harmônico, ou o número de ciclos por intervalo básico (Davis, 1986).

Segundo Davis (1986), definindo-se:

$$\alpha_k = A_k \cos(\phi_k) \text{ e } \beta_k = A_k \sin(\phi_k)$$

pode-se escrever:

$$Y_k = \alpha_k \cos(k\theta) + \beta_k \sin(k\theta)$$

Ainda segundo Davis (1986), qualquer série temporal pode ser representada como a soma de co-senóides, a despeito do quão complexas elas sejam. Estas somas definem a Relação de Fourier:

$$Y = \sum_{k=0}^{\infty} \alpha_k \cos(k\theta) + \beta_k \sin(k\theta)$$

3.9.2 Análise Harmônica

Segundo Davis (1986), em análise harmônica, os coeficientes α_k e β_k podem ser obtidos a partir de:

$$\alpha_k = \frac{2}{k} \sum_{j=0}^{n-1} Y_j \cos\left(\frac{2\pi j k}{n}\right) \text{ e } \beta_k = \frac{2}{k} \sum_{j=0}^{n-1} Y_j \sin\left(\frac{2\pi j k}{n}\right)$$

Determinados os coeficientes α_k e β_k para o k -ésimo número harmônico, a amplitude da onda pode ser obtida por, conforme Rayner (1971):

$$A_k = \sqrt{\alpha_k^2 + \beta_k^2}$$

A variância, ou espectro de potência bruto, de cada co-senóide pode ser obtida como a metade do quadrado da amplitude (Rayner, 1971), logo:

$$s_k^2 = \frac{A^2}{2} = \frac{\alpha_k^2 + \beta_k^2}{2}$$

Segundo Davis (1986), a análise espectral bidimensional utiliza os mesmos conceitos que a unidimensional com a diferença que a superfície espectral calcula os

coeficientes por meio de séries duplas de Fourier, que resultam da expansão das séries simples de Fourier, como segue:

$$Y_{ij} = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \alpha_{ij} \cos\left(\frac{2k_1 \pi X_{1i}}{n_1}\right) \cdot \cos\left(\frac{2k_2 \pi X_{2j}}{n_2}\right) + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \beta_{ij} \cos\left(\frac{2k_1 \pi X_{1i}}{n_1}\right) \cdot \sin\left(\frac{2k_2 \pi X_{2j}}{n_2}\right) \\ + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \gamma_{ij} \sin\left(\frac{2k_1 \pi X_{1i}}{n_1}\right) \cdot \cos\left(\frac{2k_2 \pi X_{2j}}{n_2}\right) + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \delta_{ij} \sin\left(\frac{2k_1 \pi X_{1i}}{n_1}\right) \cdot \sin\left(\frac{2k_2 \pi X_{2j}}{n_2}\right)$$

Segundo Davis (1986), se os dados são coletados em uma malha regular, o comprimento de onda na direção X_1 pode ser definido arbitrariamente como sendo igual ao comprimento do conjunto de dados naquela direção, ou n_1 . Ainda segundo este autor, o comprimento de onda fundamental em X_2 pode ser definido, como n_2 , de modo similar a n_1 . Desta maneira, os coeficientes de qualquer número harmônico podem ser calculados por (Davis, 1986):

$$\alpha_{k_1 k_2} = \frac{\kappa}{n_1 n_2} \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} Y_{ij} \cos\left(\frac{2\pi k_1 X_{1i}}{n_1}\right) \cdot \cos\left(\frac{2\pi k_2 X_{2j}}{n_2}\right)$$

$$\beta_{k_1 k_2} = \frac{\kappa}{n_1 n_2} \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} Y_{ij} \cos\left(\frac{2\pi k_1 X_{1i}}{n_1}\right) \cdot \sin\left(\frac{2\pi k_2 X_{2j}}{n_2}\right)$$

$$\gamma_{k_1 k_2} = \frac{\kappa}{n_1 n_2} \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} Y_{ij} \sin\left(\frac{2\pi k_1 X_{1i}}{n_1}\right) \cdot \cos\left(\frac{2\pi k_2 X_{2j}}{n_2}\right)$$

$$\delta_{k_1 k_2} = \frac{\kappa}{n_1 n_2} \sum_{i=0}^{n_1-1} \sum_{j=0}^{n_2-1} Y_{ij} \sin\left(\frac{2\pi k_1 X_{1i}}{n_1}\right) \cdot \sin\left(\frac{2\pi k_2 X_{2j}}{n_2}\right)$$

com

$$\kappa = \begin{cases} 1 & \text{se } k_1 = 0 \text{ e } k_2 = 0 \\ 2 & \text{se } k_1 = 0 \text{ ou } k_2 = 0, \text{ mas não ambos} \\ 4 & \text{se } k_1 > 0 \text{ e } k_2 > 0 \end{cases}$$

onde k_1 e k_2 são os números harmônicos nas direções X_1 e X_2 , respectivamente; e n_1 e n_2 são os números de nós da malha regular nas direções X_1 e X_2 .

Os harmônicos devem ser calculados até $\frac{n_1}{2}$ e $\frac{n_2}{2}$, pois harmônicos maiores são estimados com poucos pontos por onda, logo são menos confiáveis que aqueles estimados para números harmônicos menores (Davis, 1986).

Tendo os coeficientes calculados, pode-se determinar o espectro de potência bruto (Figura 11) como (Rayner, 1971):

$$s_{k_1 k_2}^2 = \frac{\alpha_{k_1 k_2}^2 + \beta_{k_1 k_2}^2 + \gamma_{k_1 k_2}^2 + \delta_{k_1 k_2}^2}{4}$$

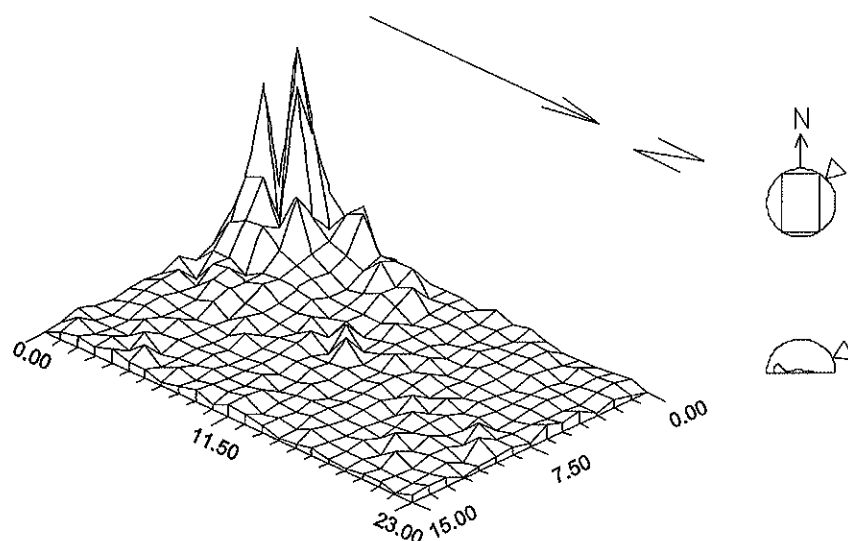


Figura 11: Superfície de espectro de potência bruto calculado a partir dos coeficientes de series duplas de Fourier.

Considerando-se a variância total de Y como s_Y^2 , a contribuição percentual de qualquer par harmônico será (Davis, 1986):

$$\frac{s_{k_1 k_2}^2}{s_Y^2} \cdot 100\%$$

Capítulo 4

Resultados

Os resultados serão apresentados conforme a seqüência lógica de operações envolvidas neste estudo. Após a geração dos dados de referência, realizou-se a análise estatística, a análise de superfícies de tendência, a análise geoestatística, a krigagem ordinária dos dados e a análise espectral dos dados das três distribuições.

4.1 Amostragem

A amostragem aleatória simples pode gerar inúmeras configurações espaciais de dados, que poderão levar a resultados distintos, notadamente em relação à medida da variabilidade espacial. Assim, para tentar superar este problema trabalhou-se com um número de realizações para cada nível percentual. Embora o número de combinações seja extremamente grande, optou-se por trabalhar com apenas seis realizações possíveis para cada nível percentual.

Cada realização foi gerada de forma que não houvesse repetição dos números pseudo-aleatórios utilizados, evitando, deste modo, duas realizações com a mesma configuração espacial de dados. Deve-se destacar que as realizações mencionadas não são realizações de simulação estocástica no sentido geoestatístico, mas sim realizações da simulação de amostragem, ou seja, a seleção de unidades amostrais através do sorteio aleatório de valores que descrevem uma distribuição uniforme.

Geradas as amostras, plotou-se os mapas de fronteira convexa de todos os níveis percentuais. Observa-se que quanto menor esse nível, menor a área da base de dados de referência amostrada e maior quantidade de formas e configurações espaciais definidas para esse subdomínio. A Figura 12 mostra as realizações para os níveis percentuais de 1%, 5% e 10% para a amostragem aleatória simples. Como seria de se esperar, a área do domínio é mais bem amostrada à medida que o n aumenta. A Figura 13 ilustra as realizações para os mesmos níveis percentuais, para a amostragem aleatória estratificada. Observa-se que, para este método de amostragem,

a superfície amostrada é superior àquela gerada pelo método de amostragem aleatória simples. Isto ocorre justamente devido à divisão do domínio em sub-domínios o que melhora a amostragem espacial. Nesse caso nos níveis percentuais menores, por exemplo 1%, a área amostrada é superior a 60%.

A amostragem sistemática, assim como nas anteriores, também recobre uma área pequena para os níveis percentuais baixos e à medida que o n aumenta a superfície dos dados de referência coberta também aumenta (Figura 14). A configuração deste tipo de amostragem faz com que a superfície recoberta pela amostragem não varie como nas amostragens anteriores. Mesmo em um nível percentual bastante elevado, apenas 89.76% da superfície são amostrados. As Tabelas 12, 13 e 14 apresentam a progressão percentual da área recoberta pela amostragem aleatória simples, aleatória estratificada e sistemática, respectivamente.

Tabela 12: Porcentagens da área dos dados de referência contidas nas fronteiras convexas, para a amostragem aleatória simples.

	Realizações dos níveis percentuais					
	1	2	3	4	5	6
1%	57.6%	55.3%	54.99%	49.00%	42.42%	56.45%
2%	81.84%	70.41%	64.82%	68.55%	76.26%	85.90%
3%	84.04%	67.35%	76.06%	89.16%	78.59%	78.32%
4%	82.17%	87.03%	85.57%	90.62%	79.98%	85.44%
5%	82.64%	90.49%	77.26%	83.51%	80.98%	87.36%
6%	87.50%	93.22%	87.17%	90.35%	90.42%	80.91%
7%	90.22%	93.68%	87.50%	93.41%	93.55%	90.69%
8%	88.63%	90.95%	96.14%	87.76%	89.42%	96.94%
9%	92.75%	91.55%	89.09%	93.81%	90.75%	95.81%
10%	94.55%	90.09%	92.22%	94.01%	96.47%	94.41%
15%	96.27%	95.14%	96.14%	96.34%	96.01%	91.62%
20%	96.67%	97.08%	99.20%	98.00%	99.26%	97.47%
30%	97.40%	98.60%	99.60%	98.60%	97.40%	96.27%
40%	99.26%	97.87%	97.34%	98.40%	99.80%	99.66%

Tabela 13: Porcentagens da área dos dados de referência contidas nas fronteiras convexas, para a amostragem aleatória estratificada.

	Realizações dos níveis percentuais						
	1	2	3	4	5	6	
Níveis Percentuais	1%	62.63%	61.57%	63.36%	65.56%	59.11%	61.10%
	2%	76.86%	66.82%	79.12%	72.01%	75.33%	76.40%
	3%	75.20%	84.18%	80.72%	67.62%	74.07%	77.59%
	4%	87.37%	84.44%	83.11%	87.30%	85.24%	82.78%
	5%	88.63%	90.76%	92.02%	87.30%	88.10%	85.51%
	6%	92.95%	96.01%	91.22%	92.62%	90.56%	92.89%
	8%	94.48%	92.62%	93.82%	94.55%	91.16%	92.35%
	10%	95.08%	96.41%	98.14%	97.87%	97.07%	96.34%
	15%	94.55%	94.95%	96.74%	98.01%	97.14%	97.01%
	20%	99.07%	96.54%	97.54%	98.47%	97.27%	95.81%
	30%	99.14%	99.40%	98.80%	99.00%	99.07%	98.54%
	40%	99.00%	98.47%	98.47%	98.47%	99.67%	98.54%

Tabela 14: Porcentagens da área dos dados de referência contidas nas fronteiras convexas, para a amostragem sistemática.

Nível Percentual	Porcentagem de Área
1%	46.54%
2%	62.23%
3%	62.23%
4%	70.88%
5%	70.88%
6%	80.05%
7%	--
8%	80.05%
9%	--
10%	80.05%
15%	80.05%
20%	80.05%
30%	89.76%
40%	89.76%
50%	89.76%

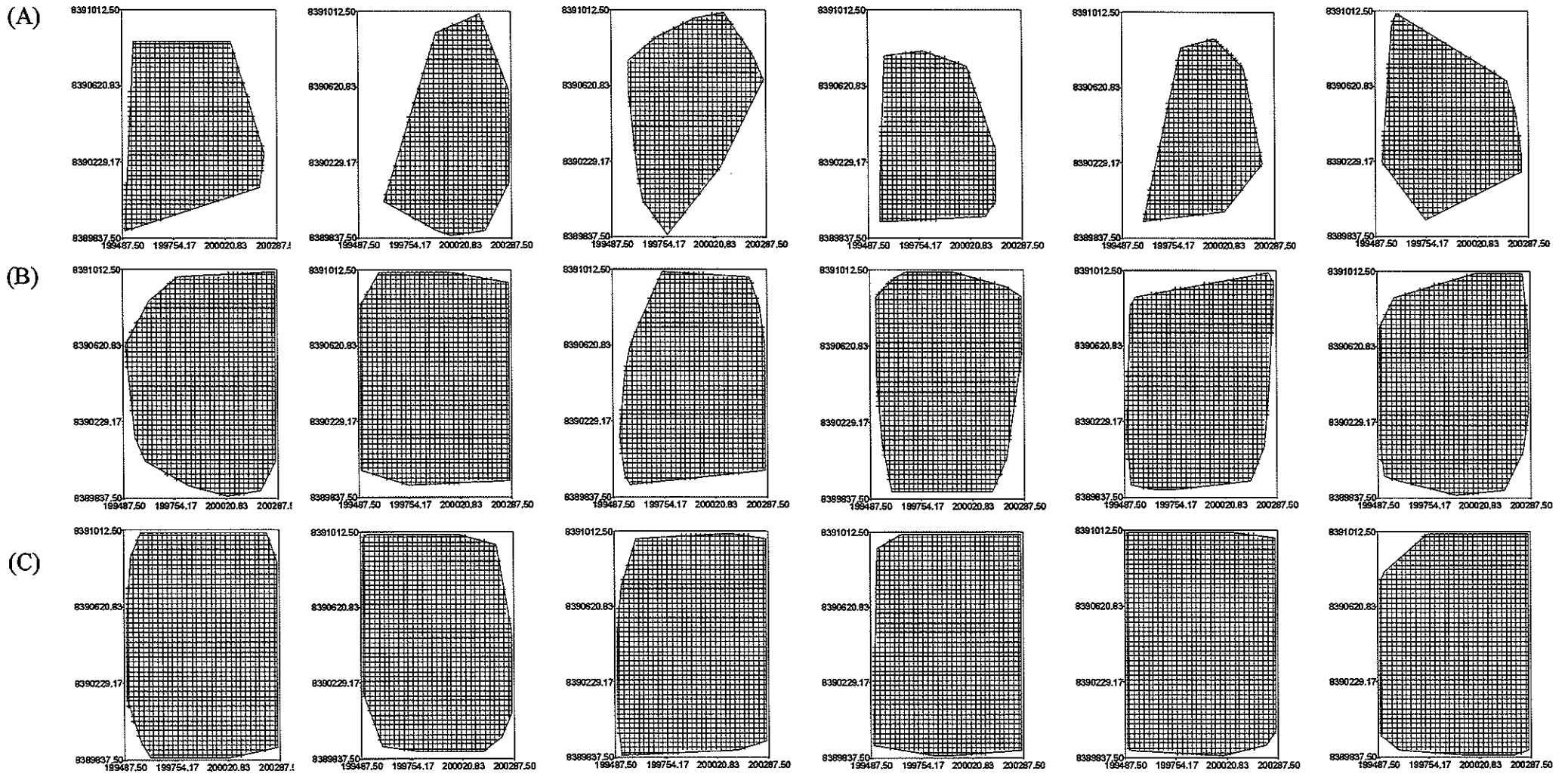


Figura 12: Fronteira convexa de três níveis percentuais (A) 1%, (B) 5% e (C) 10% da amostragem aleatória simples.

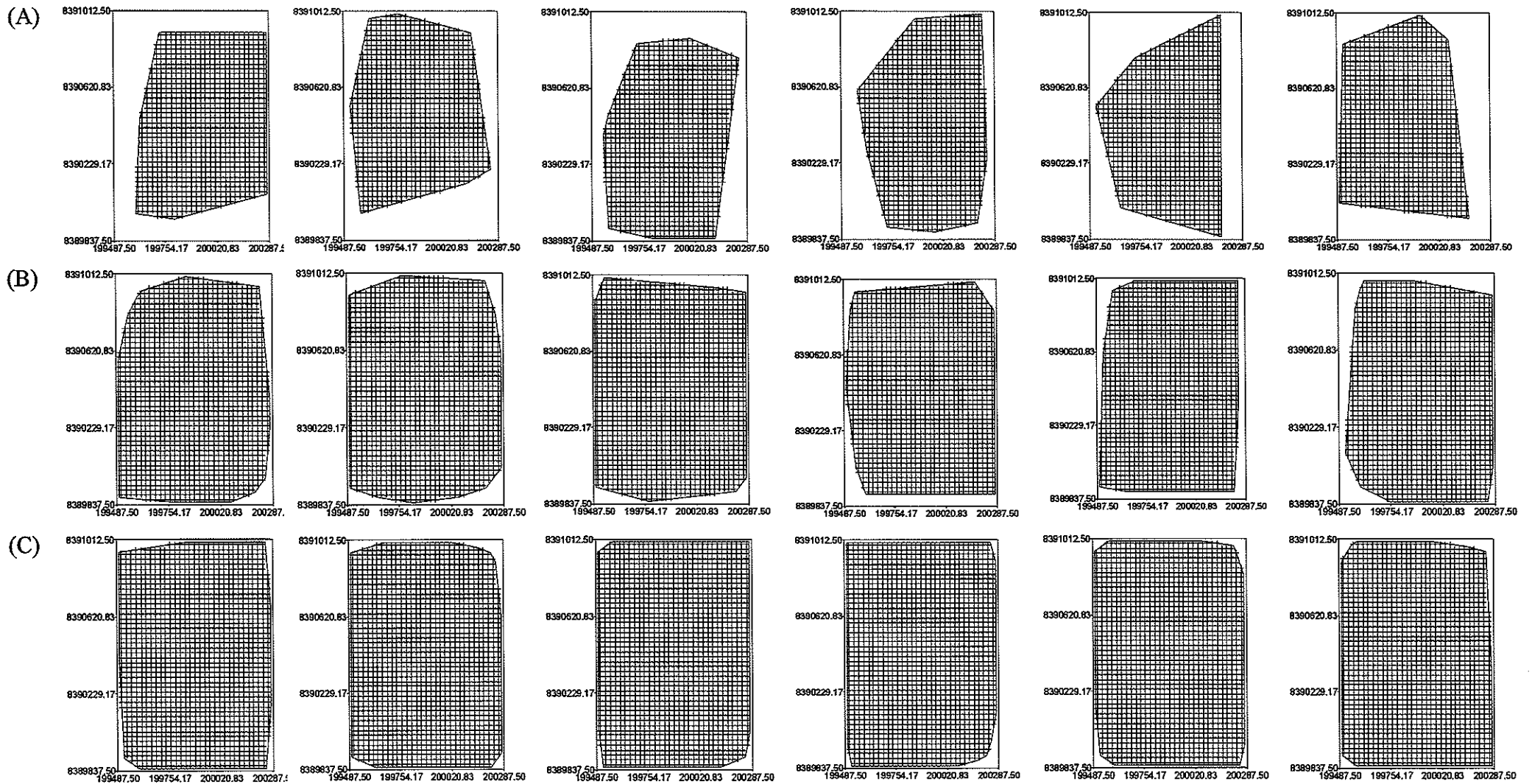


Figura 13: Fronteira convexa de três níveis percentuais (A) 1%, (B) 5% e (C) 10% da amostragem aleatória estratificada.

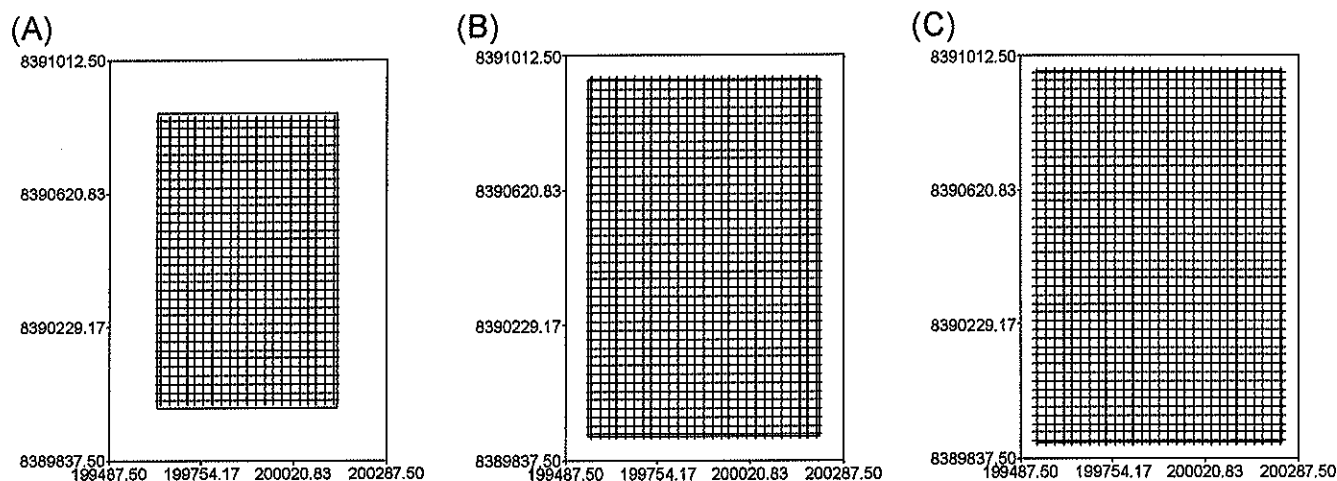


Figura 14: Fronteira convexa de três níveis percentuais (A) 1%, (B) 5% e (C) 10% da amostragem sistemática.

4.2 Análise Estatística dos Níveis Percentuais

Neste item, serão apresentados, a estatística descritiva e os diagramas de dispersão da média e da variância de cada nível percentual. Diagramas Q-Q e P-P serão mostrados, apenas para alguns destes níveis, dado o grande volume de dados manipulados. Será tratada, também, a convergência de média e variância, o nível percentual será considerado convergente se ao menos quatro valores das razões entre a média amostral e a média populacional, de cada realização, estiver dentro de uma faixa de tolerância de $\pm 10\%$ do valor da média populacional. A razão entre médias (RM) foi calculada como:

$$RM = \left(\frac{\bar{x}}{\mu} \right) \cdot 100\%$$

Para a porcentagem de variância total ($PVTO$), calculada como:

$$PVTO = \left(\frac{s^2}{\sigma^2} \right) \cdot 100\%,$$

foi aplicada uma faixa de tolerância de $\pm 20\%$.

4.2.1 Amostragem Aleatória Simples

4.2.1.1 Variável Assimétrica Positiva

A Tabela 15 sintetiza os resultados obtidos pela análise estatística das realizações para os diferentes níveis percentuais, gerados por amostragem aleatória simples, da variável assimétrica positiva. Observa-se, nesta tabela, que para níveis percentuais baixos, as estatísticas variam e em sua maioria, os valores da média somente iniciam a convergência às estatísticas populacionais a partir de níveis percentuais superiores a 10%. Este comportamento também pode ser observado na Figura 15, que é o diagrama de dispersão da razão entre a média amostral e a média populacional (A), expressa em porcentagem, e da porcentagem da variância total (B).

A média amostral converge para a média populacional, a partir do nível percentual de 10%. Por outro lado, a porcentagem da variância total converge para o valor populacional no nível percentual de 30%.

Os diagramas Q-Q e P-P confirmam que a reprodução da distribuição dos dados de referência ocorre apenas quando os níveis percentuais são superiores a 10%. Exemplifica-se esta tendência na Figura 16, que reproduz os diagramas Q-Q e P-P para os níveis percentuais de 1%, 10% e 20%. Na Figura 16, observa-se que, para a amostragem de 1% do total de dados, a dispersão das distribuições das 6 realizações, tanto no gráfico Q-Q quanto no P-P, é alta, mostrando que a distribuição dos dados de referência não foi reproduzida. Os diagramas para a amostragem de 10% mostram que, embora a dispersão em torno da reta 1:1 ainda seja alta, esta é menor que aquela de 1% dos dados. Considera-se, portanto, que a convergência ao histograma dos dados originais começa a ser alcançada. Para este nível percentual, no diagrama Q-Q, os quantis abaixo do quartil superior apresentam um espalhamento menor. Observando-se os diagramas do nível percentual de 20% (Figura 16), verifica-se que a dispersão em torno da 1ª bissetriz é menor e que o espalhamento dos valores aumenta apenas nos cinco últimos quantis, que correspondem aos valores encontrados na cauda direita da distribuição.

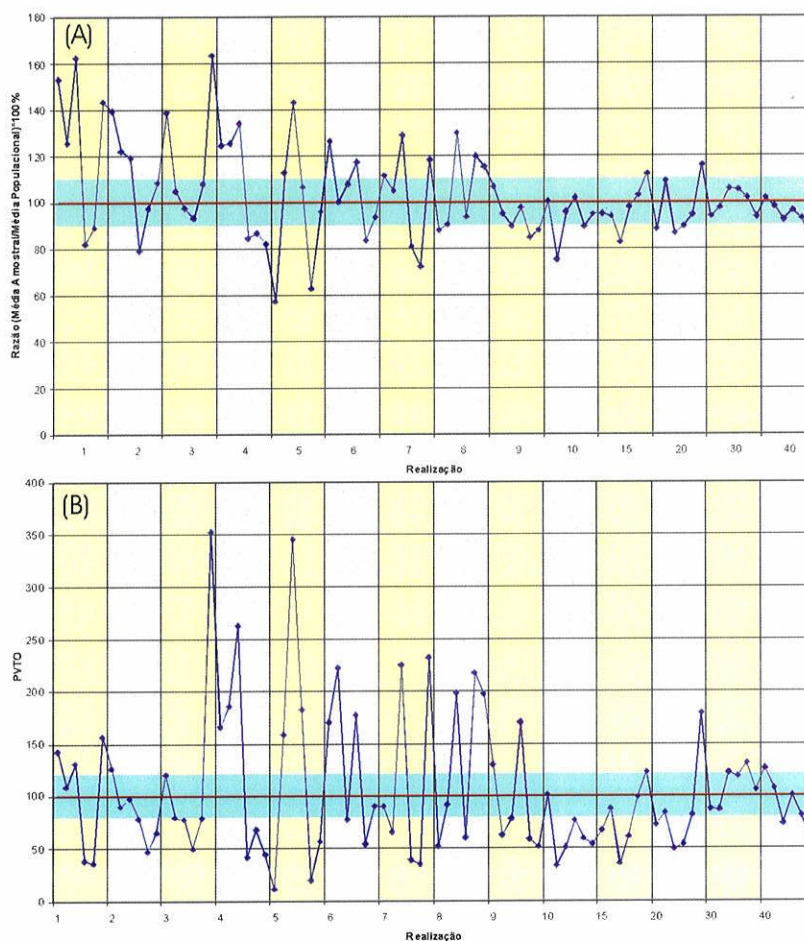


Figura 15: Dispersão das *RM* (A) e das *PVTO* (B) em torno da média e variância populacionais da variável assimétrica positiva, obtidos por amostragem aleatória simples.

4.2.1.2 Variável Simétrica

Os resultados obtidos para a variável simétrica indicam que, para esta variável, a média amostral converge mais rapidamente para a média populacional. O diagrama de dispersão das razões entre as médias (Figura 17-A), mostra que existe a convergência para todos os níveis percentuais. A *PVTO* converge no nível percentual de 4%, destaca-se que para níveis percentuais baixos a oscilação da *PVTO* é grande. A análise da Tabela 16 confirma estas interpretações, pois se vê que os resultados numéricos das estatísticas são muito próximos dos resultados da base de dados de referência.

Tabela 15: Estatísticas dos dados para os diferentes níveis percentuais obtidos por amostragem aleatória simples para a variável assimétrica positiva.

	Realizações																	
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
	1%						2%						3%					
n	15	15	15	15	15	15	30	30	30	30	30	30	45	45	45	45	45	45
\bar{x}	1.267	1.041	1.344	0.681	0.738	1.187	1.154	1.012	0.988	0.656	0.808	0.899	1.150	0.868	0.809	0.772	0.895	1.353
Med.	0.486	0.434	0.509	0.376	0.395	0.429	0.458	0.428	0.434	0.292	0.391	0.342	0.486	0.391	0.338	0.346	0.382	0.389
s	1.751	1.531	1.680	0.911	0.882	1.836	1.648	1.392	1.450	1.296	1.009	1.182	1.610	1.308	1.292	1.036	1.304	2.754
CV	1.383	1.471	1.250	1.337	1.197	1.547	1.429	1.375	1.468	1.975	1.248	1.314	1.399	1.507	1.597	1.342	1.457	2.035
	4%						5%						6%					
n	60	60	60	60	60	60	75	75	75	75	75	75	90	90	90	90	90	90
\bar{x}	1.032	1.039	1.110	0.700	0.719	0.680	0.475	0.936	1.185	0.884	0.521	0.795	1.047	0.830	0.893	0.972	0.692	0.775
Med.	0.404	0.421	0.386	0.370	0.359	0.388	0.334	0.344	0.366	0.312	0.307	0.404	0.432	0.354	0.427	0.379	0.359	0.373
s	1.888	2.000	2.377	0.947	1.209	0.980	0.498	1.846	2.724	1.982	0.656	1.104	1.913	2.185	1.292	1.952	1.082	1.395
CV	1.830	1.924	2.140	1.353	1.682	1.442	1.047	1.971	2.298	2.243	1.259	1.389	1.828	2.633	1.448	2.007	1.563	1.800
	7%						8%						9%					
n	105	105	105	105	105	105	120	120	120	120	120	120	135	135	135	135	135	135
\bar{x}	0.924	0.870	1.068	0.671	0.599	0.979	0.728	0.75	1.076	0.777	0.994	0.954	0.884	0.787	0.743	0.809	0.702	0.729
Med.	0.391	0.377	0.385	0.348	0.354	0.336	0.365	0.35	0.396	0.393	0.421	0.393	0.401	0.364	0.370	0.361	0.372	0.356
s	1.392	1.190	2.200	0.920	0.869	2.235	1.062	1.406	2.065	1.14	2.162	2.06	1.671	1.162	1.298	1.912	1.126	1.056
CV	1.506	1.368	2.061	1.370	1.453	2.284	1.459	1.874	1.919	1.467	2.176	2.159	1.890	1.476	1.747	2.364	1.602	1.449
	10%						15%						20%					
n	150	150	150	150	150	150	225	225	225	225	225	225	300	300	300	300	300	300
\bar{x}	0.832	0.623	1.070	0.843	0.741	0.786	0.787	0.778	0.686	0.812	0.853	0.93	0.733	0.904	0.718	0.743	0.784	0.962
Med.	0.357	0.358	0.381	0.36	0.375	0.407	0.377	0.374	0.379	0.39	0.377	0.383	0.37	0.391	0.369	0.384	0.359	0.378
s	1.479	0.85	2.111	1.286	1.129	1.083	1.205	1.376	0.877	1.151	1.461	1.624	1.248	1.346	1.029	1.076	1.327	1.961
CV	1.779	1.365	1.973	1.525	1.524	1.379	1.531	1.77	1.277	1.417	1.712	1.747	1.703	1.489	1.434	1.449	1.692	2.039
	30%						40%											
n	451	451	451	451	451	451	601	601	601	601	601	601						
\bar{x}	0.779	0.809	0.876	0.872	0.844	0.776	0.842	0.813	0.765	0.798	0.769	0.68						
Med.	0.359	0.382	0.384	0.376	0.351	0.36	0.368	0.379	0.361	0.36	0.373	0.355						
s	1.371	1.368	1.62	1.599	1.679	1.508	1.647	1.521	1.26	1.471	1.324	1.117						
CV	1.76	1.692	1.85	1.833	1.989	1.942	1.956	1.871	1.647	1.843	1.721	1.643						

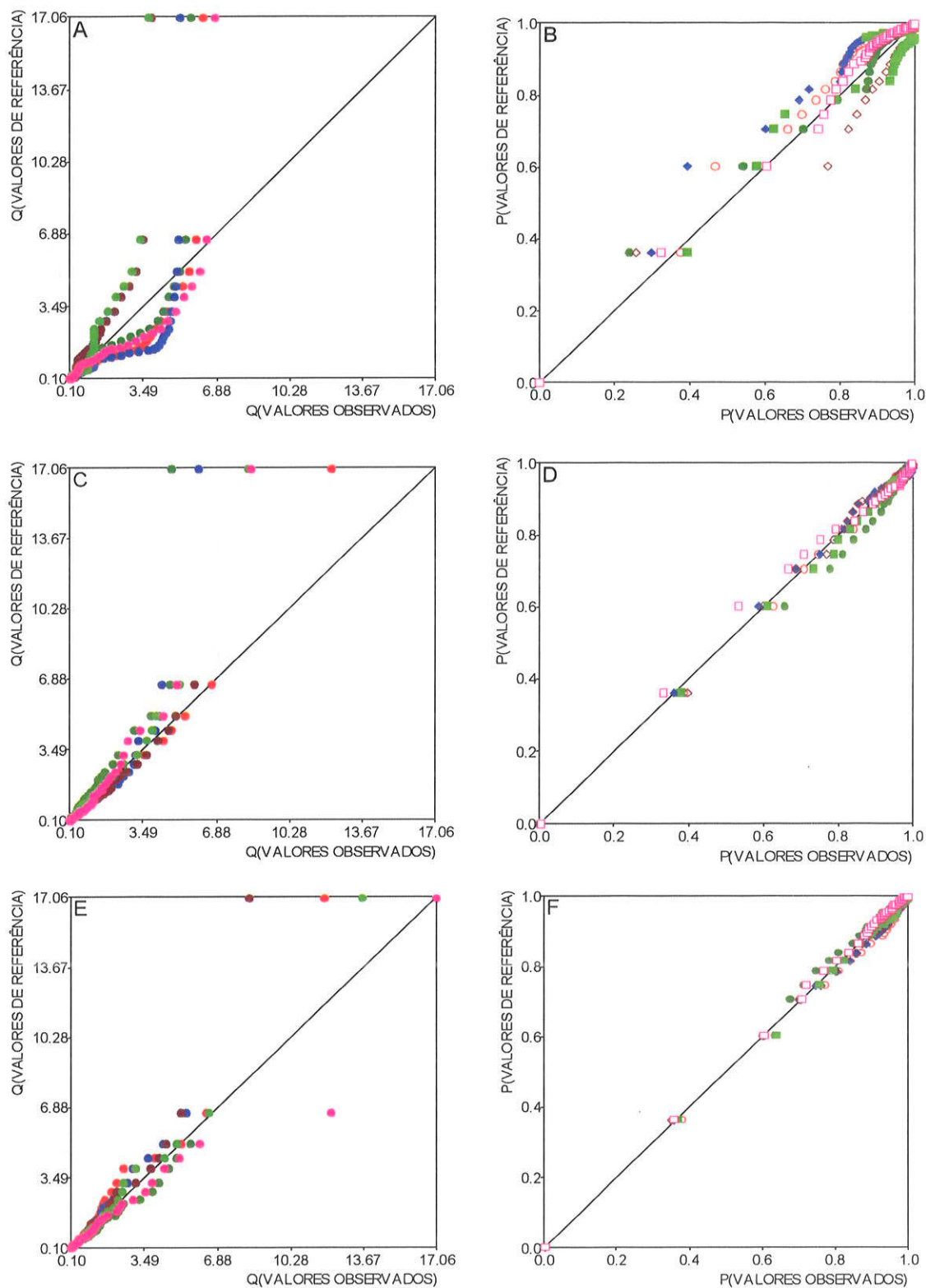


Figura 16: Diagramas Q-Q da variável assimétrica positiva referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória simples.

Tabela 16: Estatísticas dos dados para os diferentes níveis percentuais obtidos por amostragem aleatória simples da variável simétrica.

	Realizações																	
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
	1%						2%						3%					
n	15	15	15	15	15	15	30	30	30	30	30	30	45	45	45	45	45	45
\bar{x}	10.953	11.213	11.223	10.744	10.787	10.018	11.083	11.11	10.276	10.249	11.125	10.454	10.93	10.61	10.421	10.794	10.705	11.03
Med.	10.499	10.919	11.084	10.635	11.659	9.23	10.709	10.891	10.113	10.124	10.674	11.084	10.698	10.634	10.223	10.93	10.643	11.177
s	2.53	2.626	1.792	2.041	3.954	3.109	2.582	1.94	3.55	2.175	2.82	3.377	2.358	3.094	2.282	3.325	2.959	2.823
CV	0.231	0.234	0.16	0.19	0.367	0.31	0.233	0.175	0.345	0.212	0.253	0.323	0.216	0.292	0.219	0.308	0.276	0.256
	4%						5%						6%					
n	60	60	60	60	60	60	75	75	75	75	75	75	90	90	90	90	90	90
\bar{x}	11.685	11.087	10.903	10.943	11.153	10.659	10.314	10.191	10.567	10.844	10.798	10.633	11.025	10.968	10.953	11.085	10.699	11.392
Med.	12.285	11.446	11.117	11.22	10.837	11.33	9.989	9.979	10.388	11.062	10.136	10.699	10.897	11.055	11.093	11.133	10.335	11.664
s	2.635	2.708	2.897	2.973	2.939	2.991	2.643	3.199	2.964	2.977	2.716	2.816	2.866	3.048	3.095	3.048	3.231	2.725
CV	0.225	0.244	0.266	0.272	0.264	0.281	0.256	0.314	0.281	0.275	0.252	0.265	0.26	0.278	0.283	0.275	0.302	0.239
	7%						8%						9%					
n	105	105	105	105	105	105	120	120	120	120	120	120	135	135	135	135	135	135
\bar{x}	10.611	10.869	11.074	10.794	10.108	10.376	10.648	11.144	11.383	10.862	11.006	10.919	10.814	10.98	11.028	10.771	10.718	10.9
Med.	10.499	10.774	11.801	10.821	9.97	10.195	10.632	11.105	11.455	10.904	10.942	11.191	10.699	10.951	11.139	10.572	10.914	10.868
s	2.728	2.973	2.757	2.799	2.538	3.178	3.094	3.034	3.145	2.875	2.892	3.073	3.372	3.195	2.842	2.919	2.727	2.996
CV	0.257	0.274	0.249	0.259	0.251	0.306	0.291	0.272	0.276	0.265	0.263	0.281	0.312	0.291	0.258	0.271	0.254	0.275
	10%						15%						20%					
n	150	150	150	150	150	150	225	225	225	225	225	225	300	300	300	300	300	300
\bar{x}	10.942	10.934	11.124	10.653	10.783	10.845	10.872	11.046	10.669	11.016	10.972	10.663	10.903	10.919	10.697	11.117	10.768	10.856
Med.	10.847	11.036	11.062	10.599	10.856	10.716	11.114	11.245	11.015	11.035	10.955	10.464	11.128	10.992	10.76	11.171	10.771	11.094
s	3.032	3.007	3.067	2.94	2.844	2.887	2.943	2.915	3.279	2.842	3.091	2.931	3.106	2.804	3.052	3.013	3	3.059
CV	0.277	0.275	0.276	0.276	0.264	0.266	0.271	0.264	0.307	0.258	0.282	0.275	0.285	0.257	0.285	0.271	0.279	0.282
	30%						40%											
n	451	451	451	451	451	451	601	601	601	601	601	601						
\bar{x}	10.692	10.745	10.789	10.952	10.782	10.841	10.603	10.782	10.98	10.738	10.738	10.778						
Med.	10.572	10.701	10.736	10.963	10.659	10.766	10.539	10.966	11.094	10.758	10.758	10.824						
s	2.996	2.941	2.87	2.975	2.964	2.913	3.057	2.951	3.018	3.032	3.032	3.114						
CV	0.28	0.274	0.266	0.272	0.275	0.269	0.288	0.274	0.275	0.282	0.282	0.289						

A Figura 18 apresenta os diagramas Q-Q e P-P da distribuição simétrica. Estes gráficos confirmam as constatações feitas no parágrafo anterior. Destaca-se, ainda, que para o nível percentual de 1%, apenas os valores das caudas da curva estão dispersos em relação à distribuição populacional; os quantis intermediários já tendem a convergir para a reta 1:1. O diagrama Q-Q para o nível percentual de 10% mostra que, salvo por alguns poucos valores na extremidade inferior da cauda da distribuição, os demais convergem para a população. Os diagramas de Q-Q e P-P para o nível percentual de 20% apontam a reprodução quase que total da distribuição populacional, pois seus pontos estão muito pouco dispersos da reta 1:1.

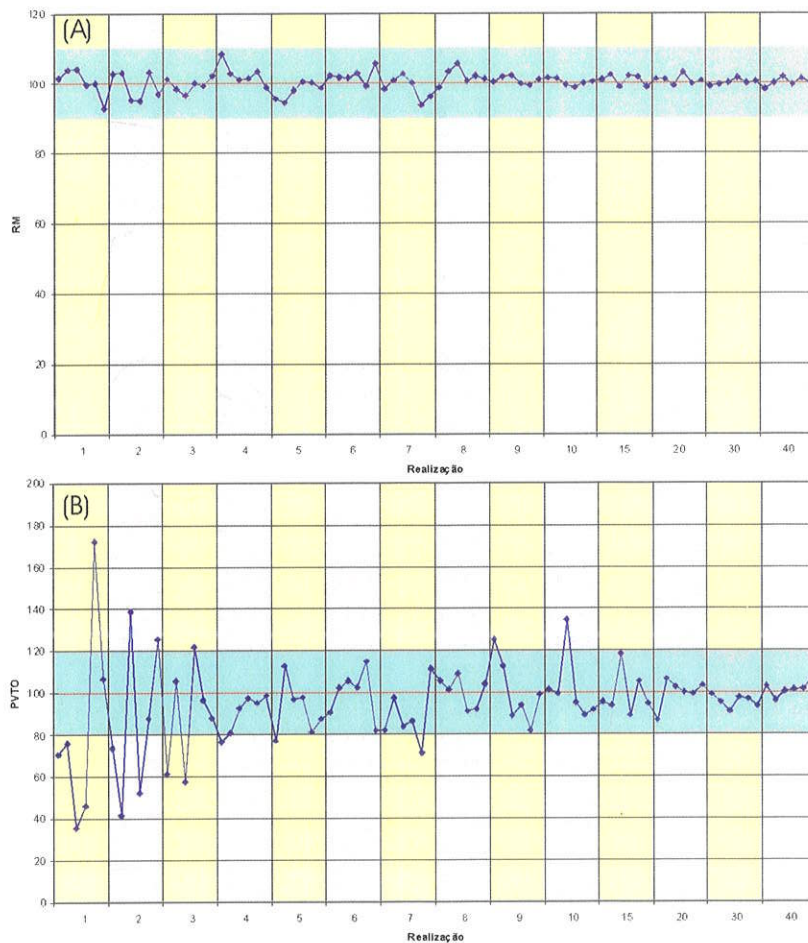


Figura 17: Dispersão das *RM* (A) e das *PVTO* (B) em torno da media e variância populacionais da variável simétrica, obtidos por amostragem aleatória simples.

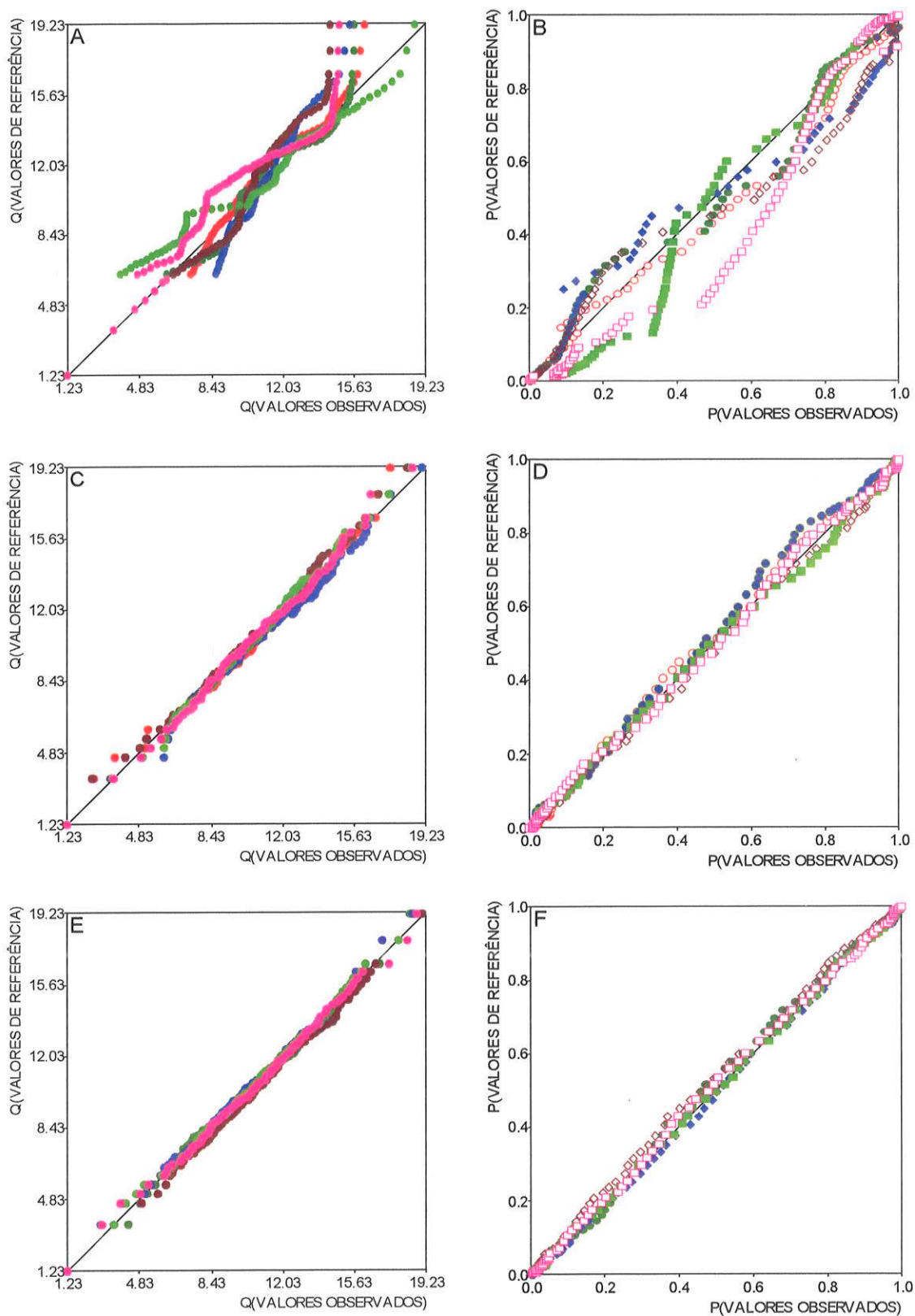


Figura 18: Diagramas Q-Q da variável simétrica referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória simples.

4.2.1.3 Variável Assimétrica Negativa

Na Tabela 17 encontram-se as estatísticas descritivas da variável assimétrica negativa. A média amostral converge para a média populacional, já no nível percentual de 1% (Figura 19-A), a variância converge no nível percentual de 4% (Figura 19-B). Os diagramas do tipo Q-Q e P-P (Figura 20) mostram que já em 10% os dados aproximam-se muito da população estudada, sendo que em 20% a distribuição dos dados de referência é quase totalmente reproduzida pelos dados amostrais.

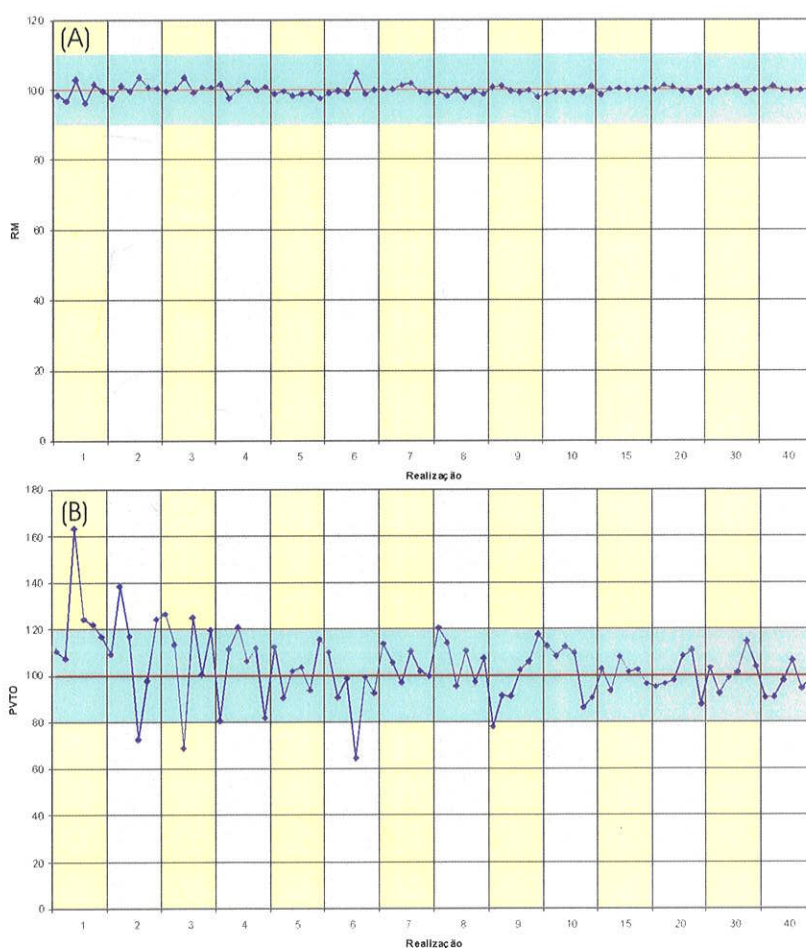


Figura 19: Dispersão das *RM* (A) e das *PVTO* (B) em torno da média e variância populacionais da variável assimétrica negativa, obtidos por amostragem aleatória simples.

Tabela 17: Estatísticas dos dados para os diferentes níveis percentuais obtidos por amostragem aleatória simples da variável assimétrica negativa.

	Realizações																	
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
	1%						2%						3%					
n	15	15	15	15	15	15	30	30	30	30	30	30	45	45	45	45	45	45
\bar{x}	53.273	52.376	55.657	52.059	54.948	53.902	52.824	54.753	53.899	56.051	54.515	54.363	53.863	54.280	55.985	53.688	54.499	54.449
Med.	55.435	50.885	60.571	52.246	56.319	56.429	53.114	58.441	55.166	57.261	56.314	56.937	56.008	56.319	56.958	56.310	54.601	56.459
s	8.670	8.548	10.533	9.195	9.108	8.911	8.621	9.701	8.918	7.028	8.159	9.192	9.274	8.780	6.844	9.224	8.269	9.022
CV	0.163	0.163	0.189	0.177	0.166	0.165	0.163	0.177	0.165	0.125	0.150	0.169	0.172	0.162	0.122	0.172	0.152	0.166
	4%						5%						6%					
n	60	60	60	60	60	60	75	75	75	75	75	75	90	90	90	90	90	90
\bar{x}	54.952	52.851	54.058	55.357	53.982	54.548	53.515	53.915	53.148	53.436	53.613	52.802	53.620	53.969	53.515	56.652	53.469	54.095
Med.	56.654	53.819	55.616	57.876	57.258	56.431	56.270	56.309	52.645	53.907	54.887	56.055	56.431	56.652	55.221	58.137	53.751	55.488
s	7.411	8.708	9.062	8.500	8.719	7.455	8.744	7.844	8.327	8.391	7.984	8.860	8.650	7.853	8.192	6.628	8.221	7.930
CV	0.135	0.165	0.168	0.154	0.162	0.137	0.163	0.145	0.157	0.157	0.149	0.168	0.161	0.146	0.153	0.117	0.154	0.147
	7%						8%						9%					
n	105	105	105	105	105	105	120	120	120	120	120	120	135	135	135	135	135	135
\bar{x}	54.166	54.149	54.837	55.140	53.790	53.575	53.789	53.131	53.997	52.921	53.828	53.454	54.512	54.722	53.933	53.714	54.024	52.973
Med.	56.319	55.801	56.655	57.778	56.270	56.631	54.592	53.818	55.893	54.998	56.572	55.772	56.589	56.319	55.209	56.429	56.429	53.246
s	8.794	8.470	8.117	8.667	8.326	8.239	9.054	8.804	8.051	8.673	8.135	8.544	7.286	7.885	7.870	8.346	8.492	8.943
CV	0.162	0.156	0.148	0.157	0.155	0.154	0.168	0.166	0.149	0.164	0.151	0.16	0.134	0.144	0.146	0.155	0.157	0.169
	10%						15%						20%					
n	150	150	150	150	150	150	225	225	225	225	225	225	300	300	300	300	300	300
\bar{x}	53.422	53.772	53.404	53.668	53.819	54.567	53.305	54.186	54.289	54.091	54.078	54.404	54.037	54.797	54.446	53.897	53.626	54.386
Med.	54.697	54.998	53.237	54.194	54.133	56.365	53.055	55.850	56.314	55.843	56.221	56.314	56.296	56.710	56.553	55.259	55.903	55.899
s	8.750	8.582	8.005	8.636	7.656	7.840	8.355	7.971	8.568	8.306	8.352	8.094	8.042	8.096	8.159	8.582	8.681	7.717
CV	0.164	0.160	0.150	0.161	0.142	0.144	0.157	0.147	0.158	0.154	0.154	0.149	0.149	0.148	0.15	0.159	0.162	0.142
	30%						40%											
n	451	451	451	451	451	451	601	601	601	601	601	601						
\bar{x}	53.645	54.099	54.314	54.539	53.500	54.029	54.076	54.643	54.035	53.869	53.869	54.199						
Med.	54.465	55.444	56.124	56.631	55.697	55.534	55.402	56.501	56.070	55.489	55.489	56.311						
s	8.382	7.920	8.202	8.301	8.827	8.401	7.844	7.853	8.158	8.511	8.511	8.191						
CV	0.156	0.146	0.151	0.152	0.165	0.155	0.145	0.144	0.151	0.158	0.158	0.151						

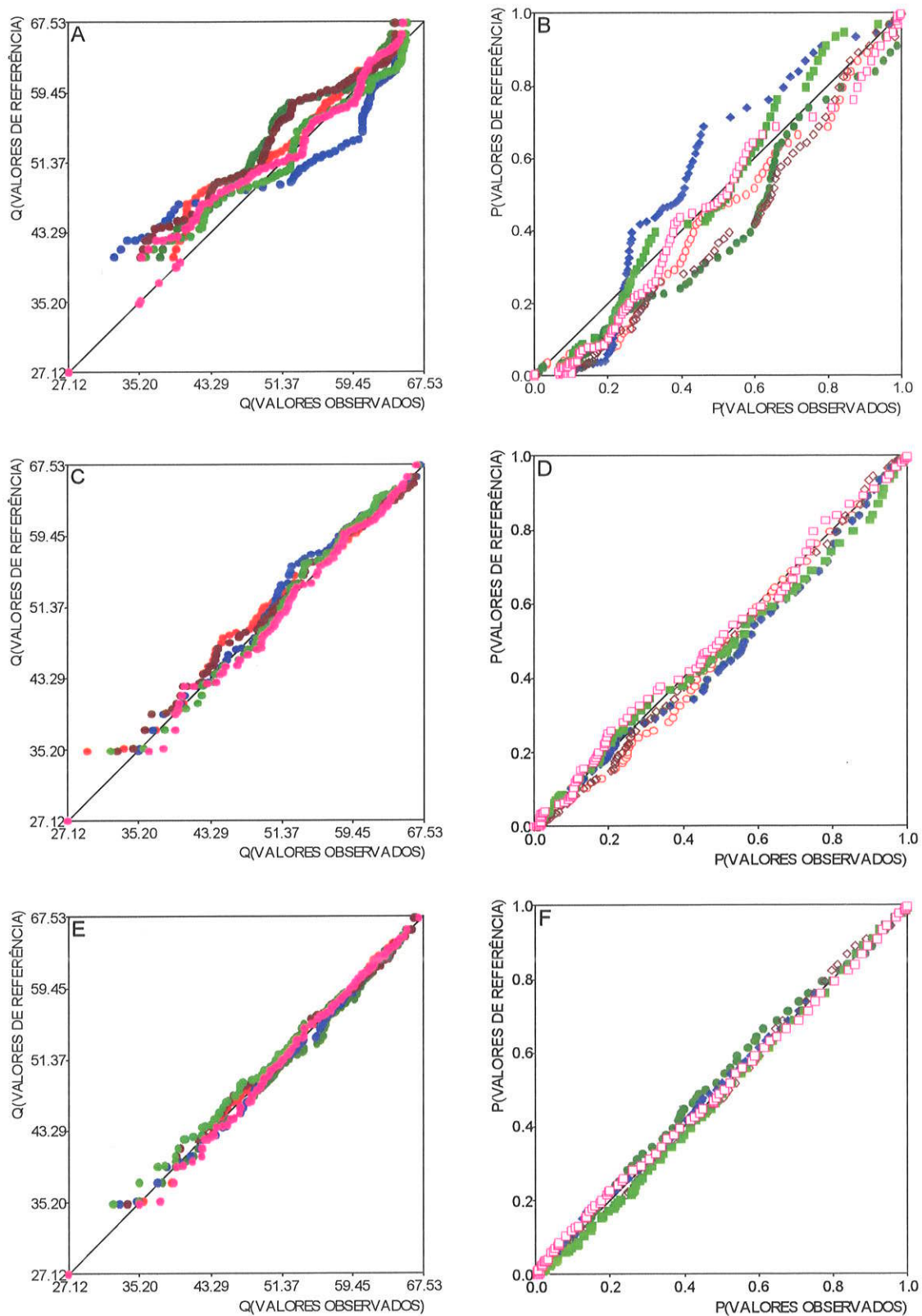


Figura 20: Diagramas Q-Q da variável assimétrica negativa referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória simples.

4.2.2 Amostragem Aleatória Estratificada

4.2.2.1 Variável Assimétrica Positiva

As estatísticas calculadas para a variável assimétrica positiva, coletadas por amostragem aleatória estratificada, encontram-se na Tabela 18. Para esta variável, as razões da média (Figura 21-A) indicam que a convergência da média amostral para a média populacional ocorre no nível percentual de 8% e a porcentagem total da variância indica convergência no nível percentual de 20%, conforme pode ser observado na Figura 21-B. Os diagramas Q-Q e P-P da variável assimétrica positiva (Figura 22), mostram que a dispersão dos valores é alta mesmo no nível percentual de 10%.

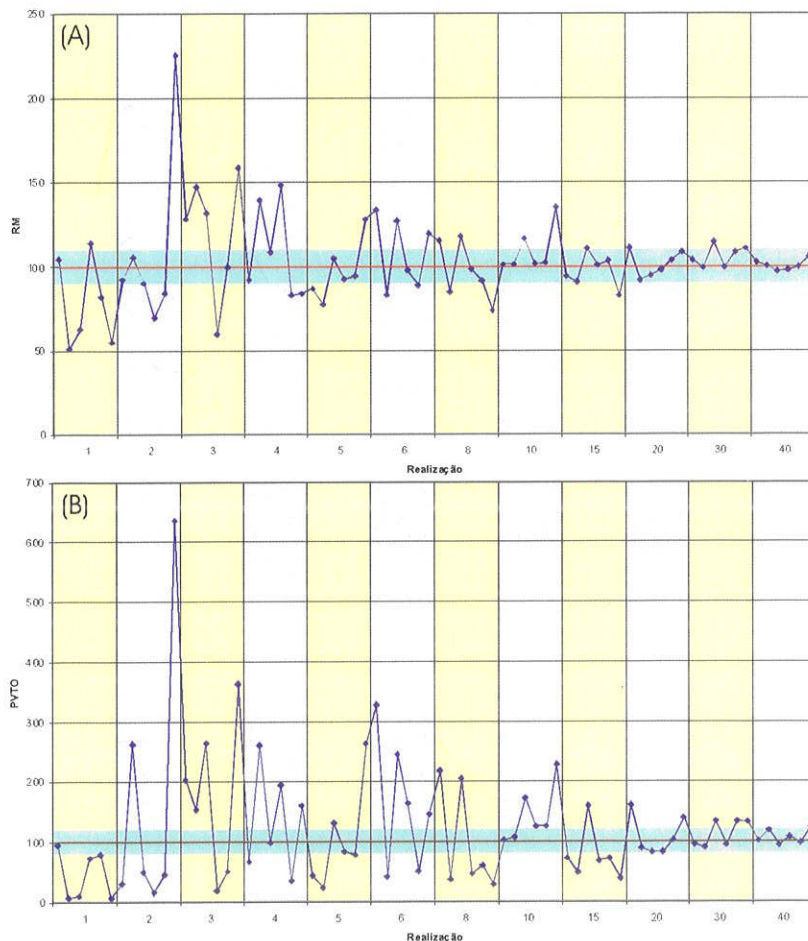


Figura 21: Dispersão das *RM* (A) e das *PVTO* (B) em torno da média e variância populacionais da variável assimétrica positiva, obtidos por amostragem aleatória estratificada.

Tabela 18: Estatísticas descritivas dos dados coletados por amostragem aleatória estratificada da distribuição assimétrica positiva

	Realizações																	
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
	1%						2%						3%					
n	15	15	15	15	15	15	30	30	30	30	30	30	45	45	45	45	45	45
\bar{x}	0.867	0.424	0.520	0.945	0.680	0.456	0.765	0.876	0.747	0.578	0.697	1.865	1.064	1.222	1.092	0.495	0.827	1.315
Med	0.410	0.259	0.349	0.347	0.234	0.376	0.407	0.324	0.333	0.342	0.330	0.373	0.434	0.387	0.365	0.225	0.428	0.385
s	1.429	0.387	0.454	1.256	1.301	0.365	0.811	2.376	1.028	0.581	0.984	3.696	2.089	1.819	2.383	0.634	1.050	2.792
CV	1.649	0.913	0.872	1.329	1.913	0.799	1.061	2.712	1.376	1.005	1.411	1.982	1.963	1.488	2.183	1.282	1.270	2.124
	4%						5%						6%					
n	60	60	60	60	60	60	75	75	75	75	75	75	90	90	90	90	90	90
\bar{x}	0.764	1.156	0.900	1.229	0.689	0.696	0.720	0.642	0.869	0.766	0.782	1.062	1.109	0.688	1.054	0.811	0.736	0.991
Med	0.378	0.344	0.420	0.428	0.361	0.346	0.386	0.389	0.393	0.390	0.420	0.335	0.360	0.388	0.400	0.372	0.395	0.353
s	1.193	2.368	1.454	2.042	0.865	1.855	0.965	0.710	1.677	1.339	1.294	2.378	2.654	0.944	2.295	1.872	1.042	1.765
CV	1.561	2.049	1.615	1.661	1.254	2.664	1.340	1.105	1.931	1.749	1.655	2.240	2.394	1.373	2.178	2.310	1.414	1.781
	8%						10%						15%					
n	120	120	120	120	120	120	150	150	150	150	150	150	225	225	225	225	225	225
\bar{x}	0.956	0.703	0.978	0.815	0.757	0.613	0.839	0.839	0.967	0.843	0.846	1.121	0.780	0.751	0.917	0.835	0.857	0.684
Med	0.356	0.376	0.397	0.423	0.382	0.374	0.359	0.379	0.385	0.380	0.394	0.390	0.378	0.360	0.375	0.391	0.395	0.347
s	2.161	0.894	2.099	1.002	1.132	0.790	1.484	1.517	1.925	1.637	1.639	2.213	1.248	1.023	1.847	1.209	1.238	0.907
CV	2.262	1.272	2.146	1.228	1.494	1.289	1.769	1.807	1.992	1.941	1.938	1.974	1.599	1.362	2.015	1.448	1.445	1.326
	20%						30%						40%					
n	300	300	300	300	300	300	451	451	451	451	451	451	601	601	601	601	601	601
\bar{x}	0.920	0.761	0.783	0.810	0.858	0.900	0.860	0.823	0.949	0.824	0.898	0.919	0.849	0.831	0.804	0.811	0.828	0.876
Med	0.376	0.360	0.378	0.373	0.404	0.374	0.381	0.377	0.389	0.374	0.384	0.379	0.387	0.377	0.377	0.392	0.379	0.383
s	1.858	1.385	1.327	1.332	1.481	1.731	1.433	1.395	1.692	1.422	1.693	1.689	1.475	1.597	1.414	1.516	1.445	1.636
CV	2.021	1.821	1.694	1.643	1.726	1.923	1.667	1.694	1.784	1.725	1.885	1.837	1.737	1.921	1.758	1.870	1.745	1.868

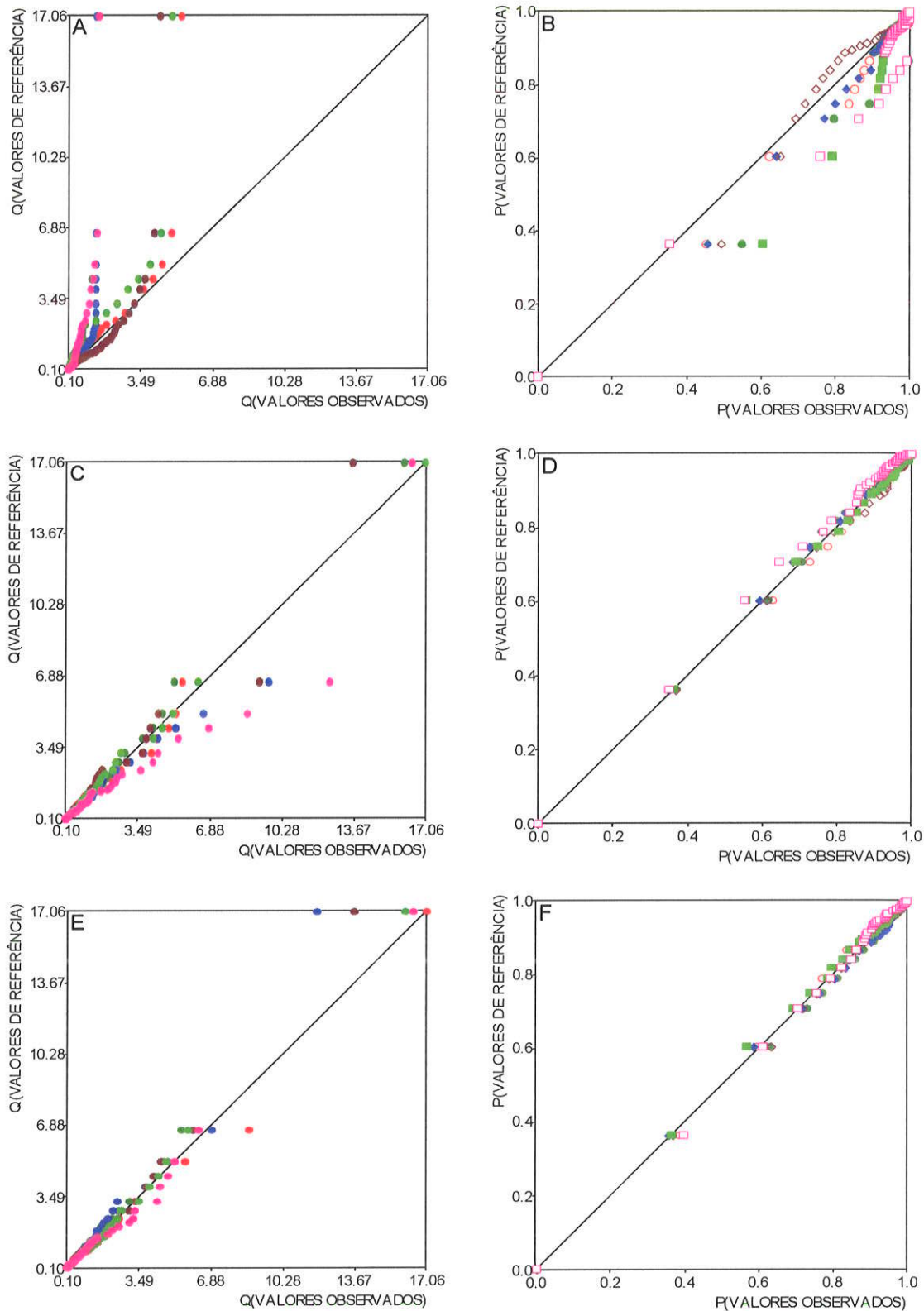


Figura 22: Diagramas Q-Q da variável assimétrica positiva referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória estratificada.

4.2.2.2 Variável Simétrica

As estatísticas descritivas da variável simétrica são as apresentadas na Tabela 19. As razões das médias, apresentadas na Figura 23-A, indicam as médias amostrais convergindo para as médias populacionais no nível percentual de 2%. A dispersão das porcentagens da variância total (Figura 23-B) indicam que no nível percentual de 4% ocorre a convergência.

Os diagramas Q-Q e P-P dos níveis percentuais de 1%, 10% e 20% (Figura 24) mostram que o nível percentual mais baixo embora pouco disperso na reta 1:1, tende a afastar-se da mesma quando próximo aos quantis das extremidades das caudas. Para o nível percentual de 10%, a maioria dos valores intermediários das distribuições concentra-se na 1ª bissetriz e apenas alguns quantis em ambos lados das caudas dispersam-se pouco. No nível percentual de 20%, o comportamento dos diagramas é semelhante àquele do nível de 10%, porém a dispersão é menor, mesmo nas caudas da distribuição.

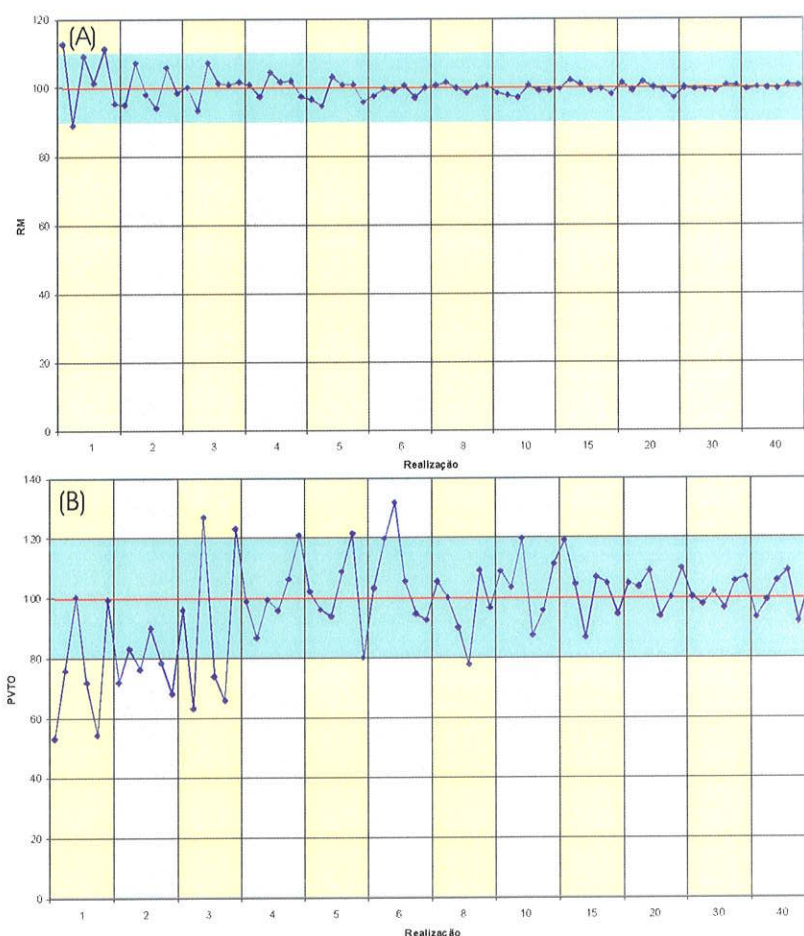


Figura 23: Dispersão das *RM* (A) e das *PVTO* (B) em torno da média e variância populacionais da variável simétrica, obtidos por amostragem aleatória estratificada.

Tabela 19: Estatísticas descritivas dos dados coletados por amostragem aleatória estratificada. Distribuição simétrica.

	Realizações																	
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
	1%						2%						3%					
n	15	15	15	15	15	15	30	30	30	30	30	30	45	45	45	45	45	45
\bar{x}	12.160	9.614	11.780	10.940	12.020	10.300	10.260	11.580	10.590	10.160	11.430	10.620	10.800	10.080	11.570	10.920	10.880	10.970
Med	12.240	10.150	11.490	11.450	12.710	9.152	10.390	11.480	10.550	10.570	11.500	10.820	11.140	9.957	11.470	11.510	11.250	10.820
s	2.197	2.625	3.019	2.555	2.220	3.004	2.555	2.747	2.632	2.859	2.669	2.487	2.954	2.394	3.395	2.589	2.443	3.344
CV	0.181	0.273	0.256	0.233	0.185	0.292	0.249	0.237	0.249	0.281	0.233	0.234	0.273	0.237	0.293	0.237	0.225	0.305
	4%						5%						6%					
n	60	60	60	60	60	60	75	75	75	75	75	75	90	90	90	90	90	90
\bar{x}	10.880	10.500	11.270	10.970	11.000	10.510	10.420	10.210	11.120	10.870	10.880	10.340	10.530	10.760	10.690	10.850	10.470	10.790
Med	11.030	10.130	11.580	11.130	10.770	10.370	10.780	10.100	11.560	10.900	10.660	10.190	10.670	11.070	10.590	10.830	10.570	10.800
s	2.996	2.806	3.004	2.949	3.106	3.315	3.045	2.953	2.918	3.143	3.321	2.695	3.060	3.298	3.458	3.094	2.931	2.900
CV	0.275	0.267	0.267	0.269	0.282	0.315	0.292	0.289	0.262	0.289	0.305	0.261	0.291	0.307	0.323	0.285	0.280	0.269
	8%						10%						15%					
n	120	120	120	120	120	120	150	150	150	150	150	150	225	225	225	225	225	225
\bar{x}	10.860	10.94	10.780	10.630	10.800	10.860	10.620	10.550	10.470	10.850	10.700	10.700	10.740	11.020	10.880	10.680	10.740	10.580
Med	11.190	10.750	10.800	10.540	11.280	10.880	10.650	10.450	10.490	10.840	10.760	10.870	10.890	11.200	11.140	10.600	10.930	10.620
s	3.093	3.014	2.860	2.657	3.147	2.962	3.144	3.063	3.299	2.818	2.950	3.177	3.290	3.080	2.806	3.114	3.084	2.930
CV	0.285	0.275	0.265	0.250	0.291	0.273	0.296	0.290	0.315	0.260	0.276	0.297	0.306	0.280	0.258	0.291	0.287	0.277
	20%						30%						40%					
n	300	300	300	300	300	300	451	451	451	451	451	451	601	601	601	601	601	601
\bar{x}	10.940	10.690	10.960	10.790	10.720	10.470	10.780	10.730	10.720	10.700	10.850	10.850	10.720	10.790	10.770	10.750	10.860	10.850
Med	11.020	10.930	11.080	10.790	11.030	10.530	10.730	10.770	10.780	10.650	10.890	10.600	10.710	10.890	10.860	10.880	10.860	10.840
s	3.084	3.067	3.145	2.919	3.013	3.157	3.016	2.980	3.044	2.960	3.094	3.113	2.912	3.002	3.097	3.146	2.890	3.073
CV	0.282	0.287	0.287	0.270	0.281	0.302	0.280	0.278	0.284	0.277	0.285	0.287	0.271	0.278	0.287	0.293	0.266	0.283

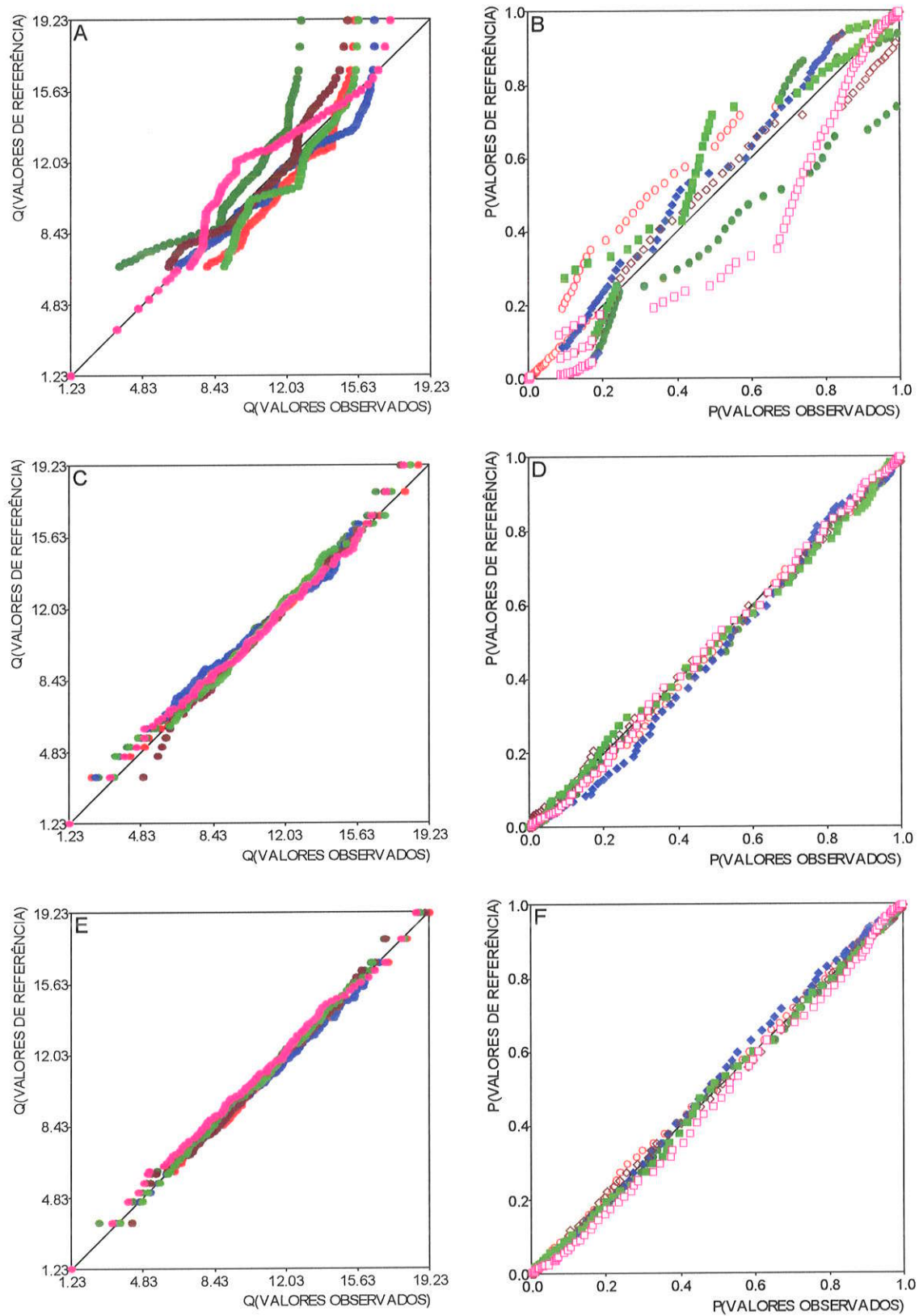


Figura 24: Diagramas Q-Q da variável simétrica referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória estratificada.

4.2.2.3 Variável Assimétrica Negativa

Para a variável assimétrica negativa, a dispersão dos valores das razões das médias indicam a convergência das médias já no nível percentual de 1% (Figura 25-A). As porcentagens totais da variância indicam que as variâncias amostrais convergem para a variância populacional no nível percentual de 5%, conforme pode ser observado na Figura 25-B.

Os diagramas Q-Q e P-P (Figura 26) indicam uma forte dispersão dos quantis, no nível percentual de 1%. O diagrama Q-Q, para o nível percentual de 20%, mostra a mesma tendência que o do nível percentual de 10%, porém dispersando menos. Já o diagrama P-P (20%) mostra que todos os percentis das distribuições das amostras agrupam-se junto à 1ª bissetriz que são os percentis dos dados de referência. Os resultados da análise estatística (Tabela 20) ratificam as informações obtidas nas análises anteriores.

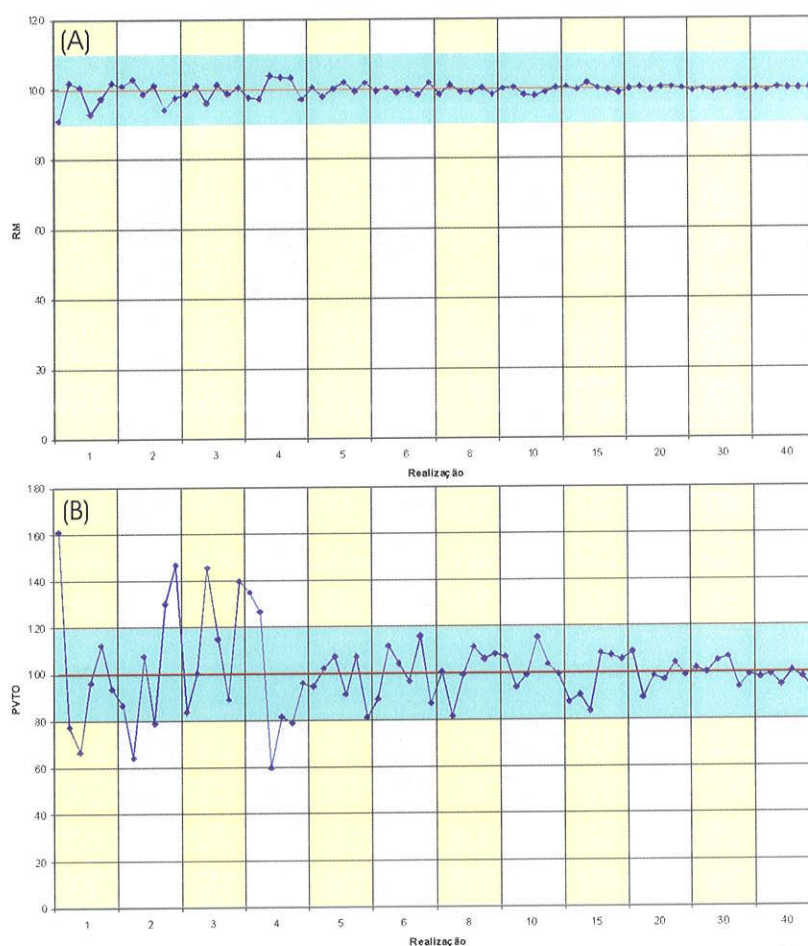


Figura 25: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável assimétrica negativa, obtidos por amostragem aleatória estratificada.

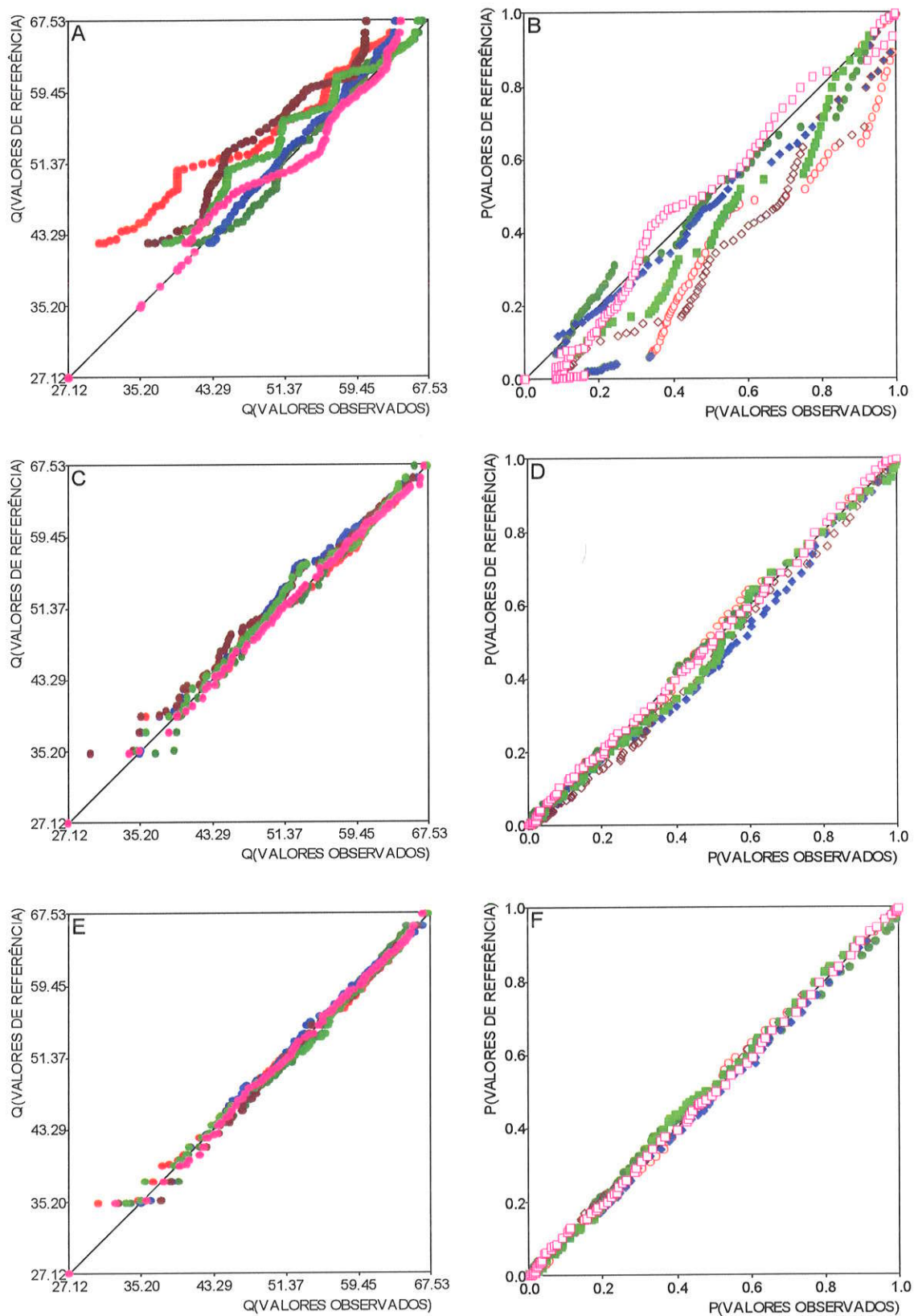


Figura 26: Diagramas Q-Q da variável assimétrica negativa referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem aleatória estratificada.

Tabela 20: Estatísticas descritivas dos dados coletados por amostragem aleatória estratificada. Distribuição assimétrica negativa.

	Realizações																	
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
	1%						2%						3%					
n	15	15	15	15	15	15	30	30	30	30	30	30	45	45	45	45	45	45
\bar{x}	49.445	55.262	54.548	50.439	52.872	55.216	54.801	55.791	53.499	54.870	51.091	52.939	53.501	54.780	52.112	54.928	53.550	54.474
Med	53.561	56.919	56.443	51.413	53.933	57.110	56.314	56.508	55.497	57.098	52.116	57.168	52.680	57.683	56.055	56.590	56.107	58.208
s.	10.457	7.255	6.732	8.090	8.741	7.984	7.674	6.606	8.557	7.323	9.405	9.985	7.553	8.254	9.945	8.842	7.779	9.745
CV	0.211	0.131	0.123	0.160	0.165	0.145	0.140	0.118	0.160	0.133	0.184	0.189	0.141	0.151	0.191	0.161	0.145	0.179
	4%						5%						6%					
n	60	60	60	60	60	60	75	75	75	75	75	75	90	90	90	90	90	90
\bar{x}	52.964	52.685	56.285	56.045	55.962	52.596	54.448	53.021	54.268	55.264	53.827	55.164	53.886	54.421	53.667	54.118	53.250	55.113
Med	55.180	55.779	58.139	58.185	57.566	52.573	56.847	52.481	56.375	58.044	56.906	56.559	53.613	56.305	53.545	55.804	55.326	56.638
s.	9.579	9.273	6.380	7.449	7.324	8.079	8.018	8.341	8.539	7.866	8.534	7.429	7.780	8.717	8.412	8.105	8.877	7.692
CV	0.181	0.176	0.113	0.133	0.131	0.154	0.147	0.157	0.157	0.142	0.159	0.135	0.144	0.160	0.157	0.150	0.167	0.140
	8%						10%						15%					
n	120	120	120	120	120	120	150	150	150	150	150	150	225	225	225	225	225	225
\bar{x}	53.334	54.760	53.713	53.621	54.301	53.275	54.211	54.377	53.216	53.080	53.606	54.310	54.475	54.043	55.067	54.214	54.015	53.515
Med	53.625	56.881	56.310	55.401	56.737	56.681	56.657	56.309	53.020	53.800	53.430	56.238	56.315	55.417	56.311	56.374	56.407	54.585
s.	8.272	7.445	8.218	8.698	8.491	8.575	8.529	7.994	8.212	8.851	8.391	8.224	7.715	7.849	7.545	8.583	8.546	8.478
CV	0.155	0.136	0.153	0.162	0.156	0.161	0.157	0.147	0.154	0.167	0.157	0.151	0.142	0.145	0.137	0.158	0.158	0.158
	20%						30%						40%					
n	300	300	300	300	300	300	451	451	451	451	451	451	601	601	601	601	601	601
\bar{x}	54.140	54.366	53.941	54.410	54.369	54.176	53.809	54.070	53.722	53.901	54.276	53.751	54.158	53.799	54.270	54.126	54.124	54.156
Med	56.337	56.316	55.871	56.465	56.482	56.032	55.402	55.945	55.470	56.319	56.315	55.501	55.990	55.221	55.898	56.310	56.120	55.453
s.	8.606	7.793	8.184	8.120	8.418	8.199	8.323	8.246	8.452	8.514	7.989	8.206	8.159	8.210	8.016	8.270	8.166	7.940
CV	0.159	0.143	0.152	0.149	0.155	0.151	0.155	0.152	0.157	0.158	0.147	0.153	0.151	0.153	0.148	0.153	0.151	0.147

4.2.3 Amostragem Sistemática

Dadas as características deste método de amostragem, para cada nível percentual, foi retirada apenas uma amostra, pois se optou por utilizar sempre a origem da malha dos dados de referência como coordenadas de origem para a tomada das amostras com espaçamento regular.

4.2.3.1 Variável Assimétrica Positiva

A Tabela 21 resume as estatísticas descritivas das três variáveis analisadas. A Figura 27-A apresenta a dispersão das razões das médias em relação a 100%, observe-se que existe convergência em praticamente todos os níveis percentuais, exceto nos níveis de 1 e 4%. A variância amostral converge para a variância populacional no nível percentual de 10%, embora nos níveis percentuais de 2 e 6% a convergência tenha sido alcançada (Figura 27-B).

A Figura 28 apresenta os diagramas Q-Q e P-P da variável assimétrica positiva para os níveis percentuais de 1%, 10% e 20%. Estes gráficos mostram que os valores desta variável, já no nível de 10%, aproximam-se muito dos valores dos dados de referência. A convergência é ainda mais elevada no nível percentual de 20%.

Tabela 21: Estatísticas descritivas das variáveis: assimétrica positiva, simétrica e assimétrica negativa.

	Variável Assimétrica Positiva					Variável Simétrica				Variável Assimétrica Negativa			
	<i>n</i>	\bar{x}	Med	<i>s</i>	CV	\bar{x}	Med	<i>s</i>	CV	\bar{x}	Med	<i>s</i>	CV
1%	15	0.605	0.300	0.981	1.622	10.109	10.116	2.952	0.292	53.642	54.793	7.572	0.141
2%	30	0.859	0.357	1.448	1.685	11.026	11.044	2.645	0.240	55.234	57.655	8.414	0.152
3%	45	0.763	0.409	1.131	1.483	10.499	10.337	3.010	0.287	51.497	51.975	9.372	0.182
4%	60	1.145	0.402	2.494	2.178	10.853	10.980	2.195	0.202	53.453	52.674	7.571	0.142
5%	75	0.834	0.429	1.220	1.463	10.488	10.711	3.319	0.317	54.874	56.723	7.921	0.144
6%	90	0.849	0.399	1.497	1.764	10.945	11.148	3.178	0.290	53.213	52.760	7.968	0.150
8%	105	0.812	0.362	1.261	1.552	10.672	10.736	3.120	0.292	54.727	56.157	7.794	0.142
10%	150	0.786	0.386	1.347	1.715	10.941	11.136	3.173	0.290	54.093	55.584	7.994	0.148
15%	225	0.820	0.372	1.325	1.615	10.689	11.089	3.072	0.287	54.134	56.638	8.617	0.159
20%	300	0.890	0.396	1.568	1.762	10.730	10.623	3.007	0.280	53.881	55.652	8.376	0.155
30%	451	0.821	0.353	1.565	1.906	10.898	11.096	3.065	0.281	53.785	55.33	8.278	0.154
40%	601	0.782	0.367	1.445	1.848	10.794	10.918	3.110	0.288	54.074	56.013	8.003	0.148
50%	752	0.864	0.378	1.628	1.884	10.880	11.016	3.033	0.279	54.129	56.302	7.991	0.148

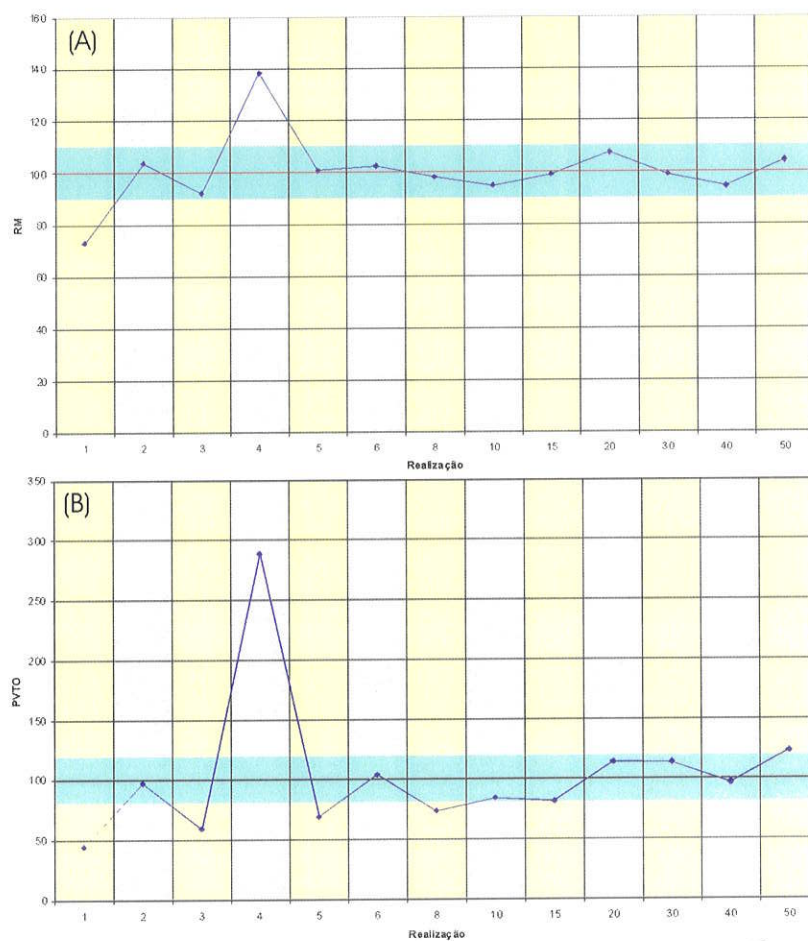


Figura 27: Dispersão das *RM* (A) e das *PVTO* (B) em torno da média e variância populacionais da variável assimétrica positiva, obtidos por amostragem sistemática.

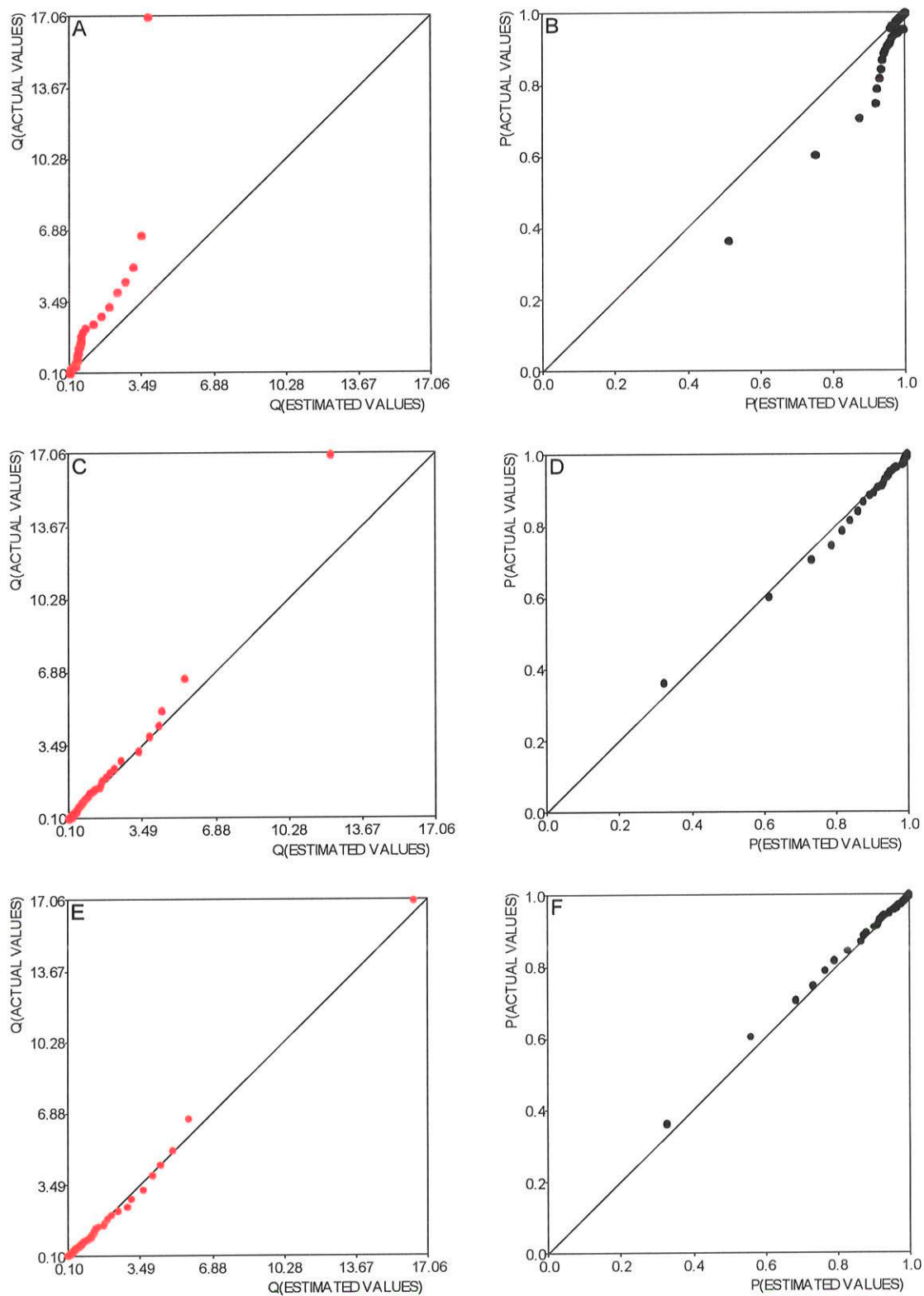


Figura 28: Diagramas Q-Q da variável assimétrica positiva referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem sistemática.

4.2.3.2 Variável Simétrica

As estatísticas descritivas da variável simétrica, coletada por amostragem sistemática, são apresentadas na Tabela 21. O diagrama de dispersão das razões das médias, Figura 29-A, indica que as médias amostrais convergem para a média populacional no nível percentual de 1%. A variância converge no nível percentual de 6%, conforme mostra a dispersão, da porcentagem da variância total, apresentada na Figura 29-B.

Os diagramas Q-Q e P-P (Figura 30) indicam que já a partir do nível percentual de 1% existe proximidade dos valores da variável com os dados de referência. A semelhança entre os dados da amostra com os valores dos dados de referência é maior no nível percentual de 10%, inclusive com a coincidência dos valores de mínimo e máximo da distribuição.

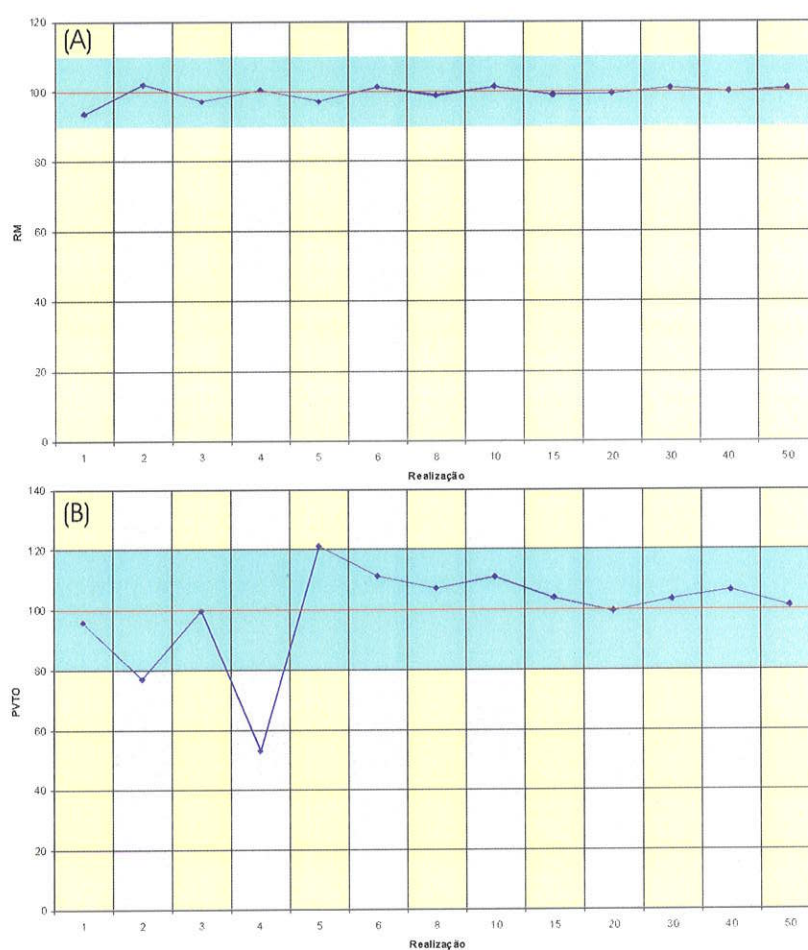


Figura 29: Dispersão das RM (A) e das PVTO (B) em torno da média e variância populacionais da variável simétrica, obtidos por amostragem sistemática.

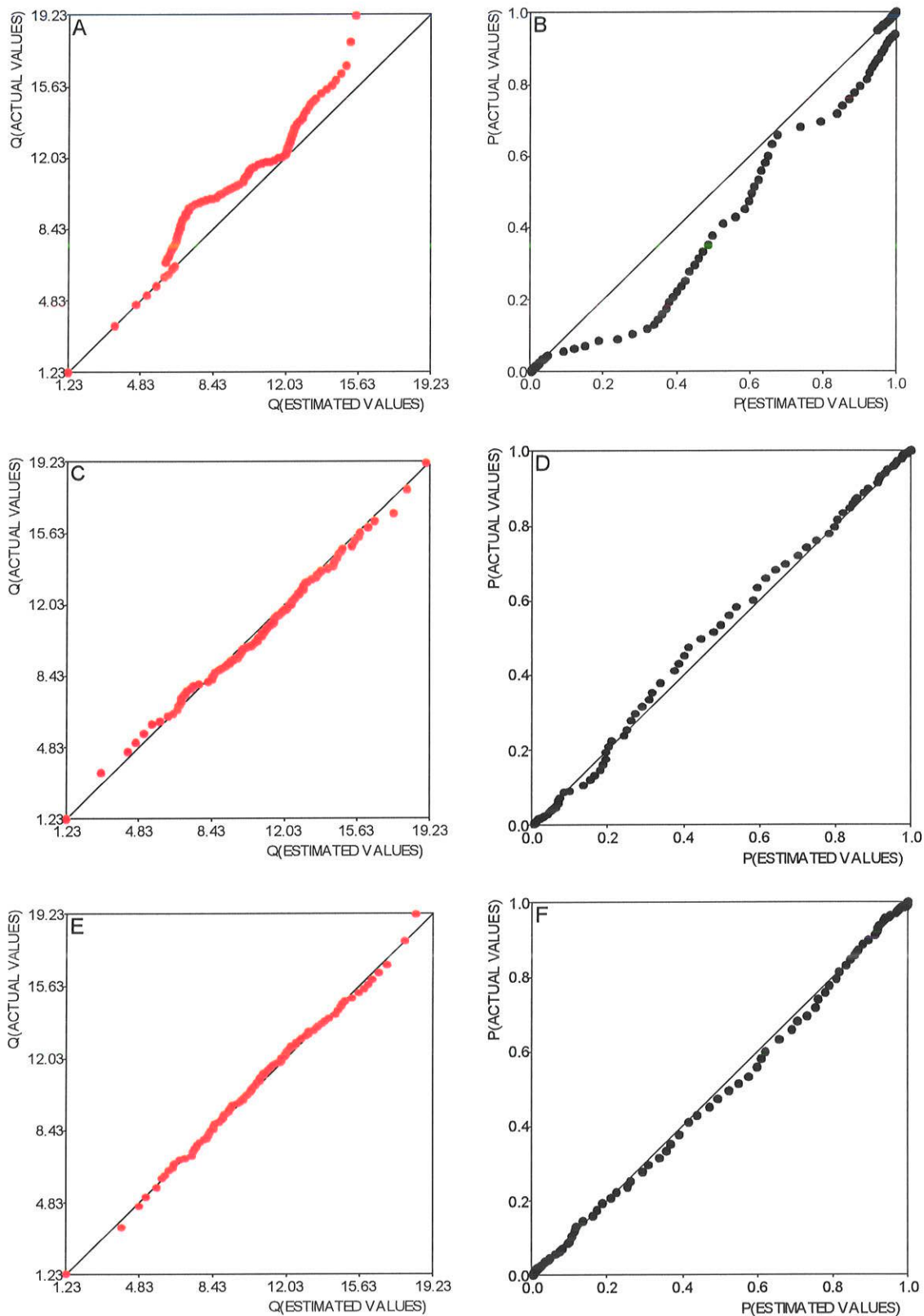


Figura 30: Diagramas Q-Q da variável simétrica referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem sistemática.

4.2.3.3 Variável Assimétrica Negativa

As estatísticas da variável assimétrica negativa são apresentadas na Tabela 21. A Figura 31-A ilustra a variação das razões das médias em torno de 100%, observa-se, nesta figura, que a média amostral converge para a populacional no nível percentual de 1%. Pelos resultados da dispersão da porcentagem da variância total, ilustrados na Figura 31-B, nota-se que apenas o nível percentual de 3% a variância amostral está fora da banda de tolerância. Os diagramas Q-Q e P-P (Figura 32) ratificam este comportamento, pois o nível percentual de 10% guarda grande similaridade entre a amostra e os dados populacionais.

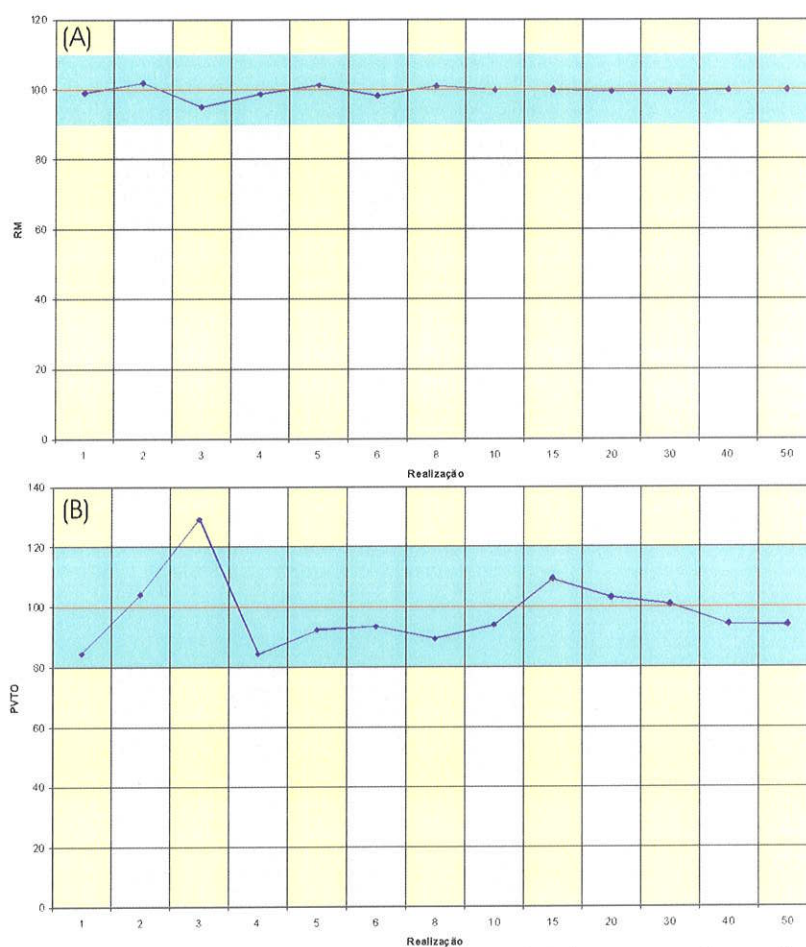


Figura 31: Dispersão das *RM* (A) e das *PVTO* (B) em torno da média e variância populacionais da variável assimétrica negativa, obtidos por amostragem sistemática.

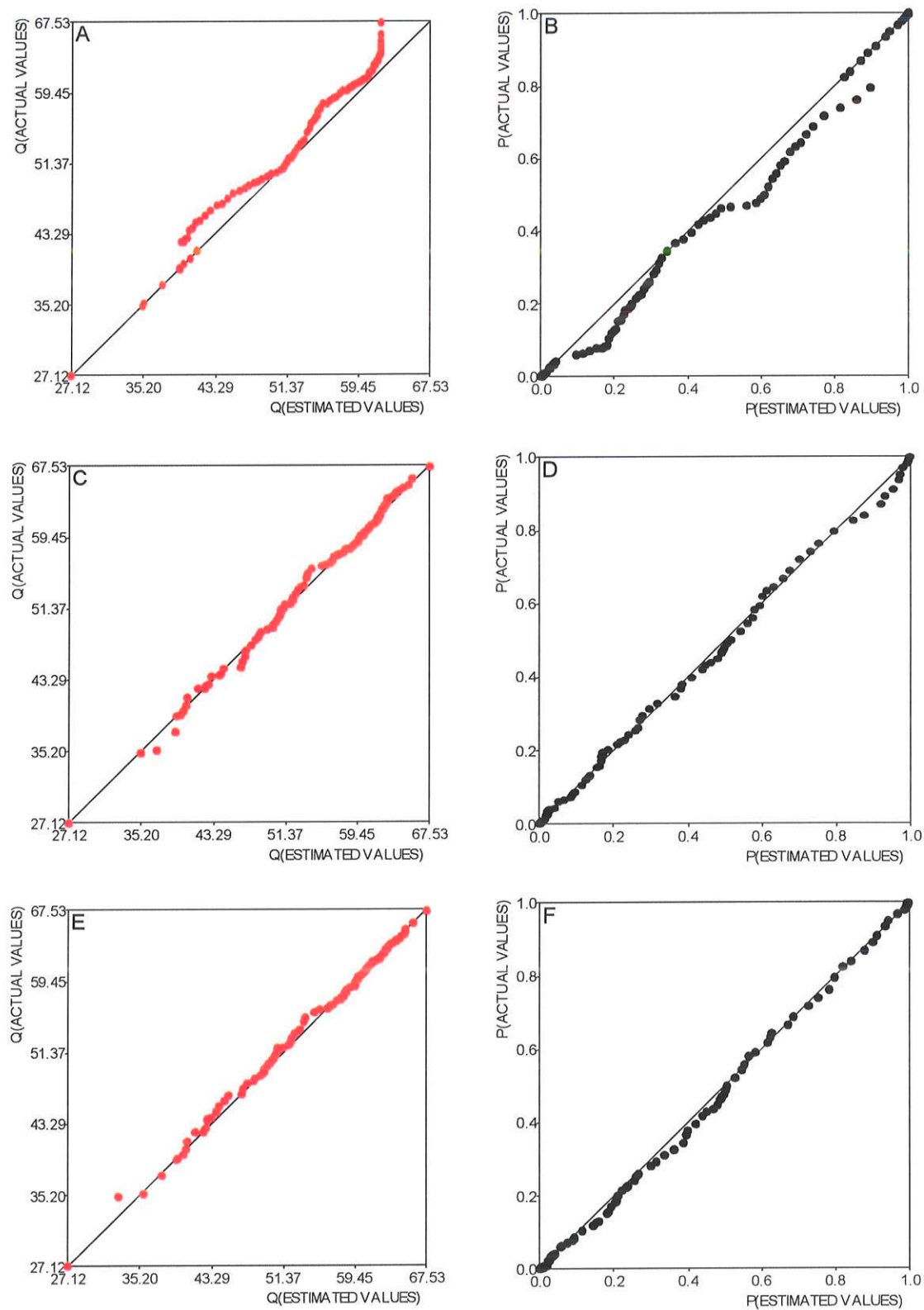


Figura 32: Diagramas Q-Q da variável assimétrica negativa referentes aos níveis percentuais 1% (A), 10% (C) e 20% (E) e diagramas P-P dos níveis percentuais de 1% (B), 10% (D) e 20% (F) obtidos por amostragem sistemática.

4.3 Análise Geoestatística dos Níveis de Percentuais

A análise geoestatística foi realizada visando atingir dois objetivos: checar em qual nível percentual o modelo de variância espacial seria reproduzido; fornecer o modelo de variograma teórico para o cálculo da krigagem ordinária. Adotar-se-á um critério visual para estabelecer se houve convergência dos variogramas experimentais das amostras em relação ao modelo de variograma dos dados de referência. Para facilitar a visualização os variogramas experimentais, dos níveis percentuais, aqueles foram divididos em grupos.

4.3.1 Amostragem Aleatória Simples

Os variogramas experimentais foram calculados segundo os parâmetros apresentados na Tabela 22. Optou-se por utilizar o maior número de parâmetros semelhantes para os variogramas experimentais como, por exemplo, direções, tolerâncias e passos. Os variogramas experimentais foram todos omnidirecionais. Para o ajuste do modelo teórico de variograma, calculou-se a média aritmética dos variogramas experimentais de cada nível percentual. Aos variogramas médios foram ajustados modelos contendo uma única estrutura, variando apenas o efeito pepita, o patamar e a amplitude. Dois modelos de variância espacial foram utilizados: o modelo esférico, para os variogramas experimentais com estrutura, e o efeito pepita puro, para os variogramas experimentais sem estrutura. As características destes modelos são as apresentadas na Tabela 23.

Tabela 22: Parâmetros do cálculo dos variogramas experimentais.

Nível percentual	Nº Passos	Passo
1%	5	100
2%	5	100
3%	5	100
4%	5	100
5%	5	100
6%	5	100
7%	7	75
8%	7	75
9%	7	75
10%	7	75
15%	7	75
20%	10	50
30%	10	50
40%	10	50

Tabela 23: Modelos de variograma ajustados aos variogramas experimentais de cada variável.

		Assimétrica Positiva			Simétrica			Assimétrica Negativa				
	Modelo	a	c ₀	c	Modelo	a	c ₀	c	Modelo	a	c ₀	c
1%	Pep. Puro	---	1.674	---	Esférico	127.5	1.750	4.900	Esférico	202.5	33.32	32.76
2%	Pep. Puro	---	1.500	---	Esférico	200.0	4.334	2.500	Esférico	160.0	18.36	58.00
3%	Esférico	115.2	1.050	1.070	Esférico	136.0	3.300	4.000	Esférico	160.0	13.62	58.50
4%	Esférico	140.0	0.490	1.950	Esférico	160.0	5.120	2.760	Esférico	168.8	35.10	31.00
5%	Esférico	142.4	0.900	1.060	Esférico	141.6	4.046	3.200	Esférico	140.0	30.96	43.00
6%	Esférico	160.0	1.548	1.070	Esférico	160.0	3.534	5.095	Esférico	160.0	44.50	18.00
7%	Esférico	140.0	0.600	1.460	Esférico	153.4	4.800	2.750	Esférico	123.2	24.03	44.00
8%	Pep. Puro	---	3.338	---	Esférico	191.1	3.387	4.840	Esférico	128.2	11.60	56.20
9%	Esférico	173.0	0.420	1.040	Esférico	144.0	4.200	4.500	Esférico	149.0	29.66	37.90
10%	Pep. Puro	20.0	1.588	0.000	Esférico	148.2	4.340	4.440	Esférico	150.4	26.55	39.50
15%	Esférico	121.0	0.540	0.660	Esférico	145.0	3.580	5.420	Esférico	150.4	35.50	34.30
20%	Esférico	124.0	0.636	0.932	Esférico	126.0	3.400	5.900	Esférico	137.0	26.63	42.00
30%	Esférico	109.8	0.649	1.350	Esférico	134.4	3.230	5.600	Esférico	139.0	20.60	50.80
40%	Esférico	123.3	0.792	0.800	Esférico	150.6	3.843	5.527	Esférico	156.6	28.60	39.90
Ref.	Esférico	85.0	0.757	1.300	Esférico	137.0	3.590	5.600	Esférico	137.0	24.50	45.50

Onde a= amplitude, c₀= efeito pepita e c= variância espacial.

4.3.1.1 Variável Assimétrica Positiva

Os variogramas experimentais médios da variável assimétrica positiva, comparados ao modelo de variograma dos dados de referência, podem ser observados na Figura 33. Como pode ser visto, nos variogramas dos níveis percentuais de 1 a 7% (Figura 33 A) não existe convergência, apenas o efeito pepita dos níveis percentuais de 5 e 7% aproximam-se daquele do modelo de referência.

Os variogramas experimentais dos níveis percentuais de 8 a 40% (Figura 33 B), embora mais próximos do modelo de variograma dos dados de referência, não convergem plenamente a esta. O efeito pepita, dos níveis percentuais superiores a 10%, converge ao valor do efeito pepita da referência. Destaca-se, ainda, que os patamares dos variogramas tendem a convergir a um valor menor que o dos dados de referência.

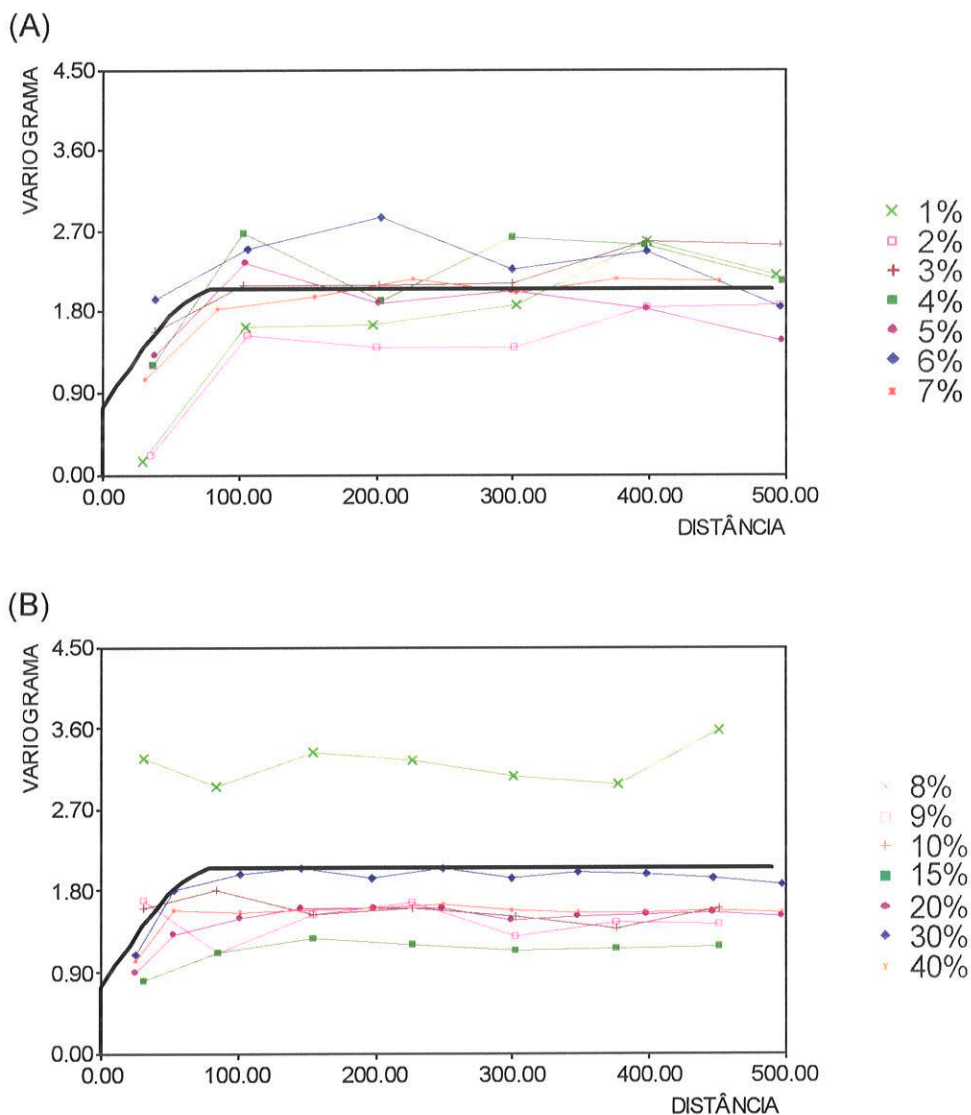


Figura 33: Variogramas experimentais médios, da variável assimétrica positiva, amostrada por amostragem aleatória simples, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 7% e em (B) os níveis percentuais de 8% a 40%.

4.3.1.2 Variável Simétrica

Os variogramas médios desta variável encontram-se representados na Figura 34, bem como o modelo de variograma dos dados de referência. Os variogramas médios, para os níveis percentuais de 1 a 7% (Figura 34-A), sempre apresentam patamar menor que a dos dados de referência, apenas o nível percentual de 6% aproxima-se daquele. As amplitudes convergem para a amplitude dos dados de

referência. De modo geral, os variogramas deste conjunto de níveis percentuais não apresentam uma convergência satisfatória.

Os variogramas médios convergem satisfatoriamente ao modelo de referência, para os níveis percentuais de 9 a 40% (Figura 34-B), tanto para a amplitude, quanto para o patamar e efeito pepita.

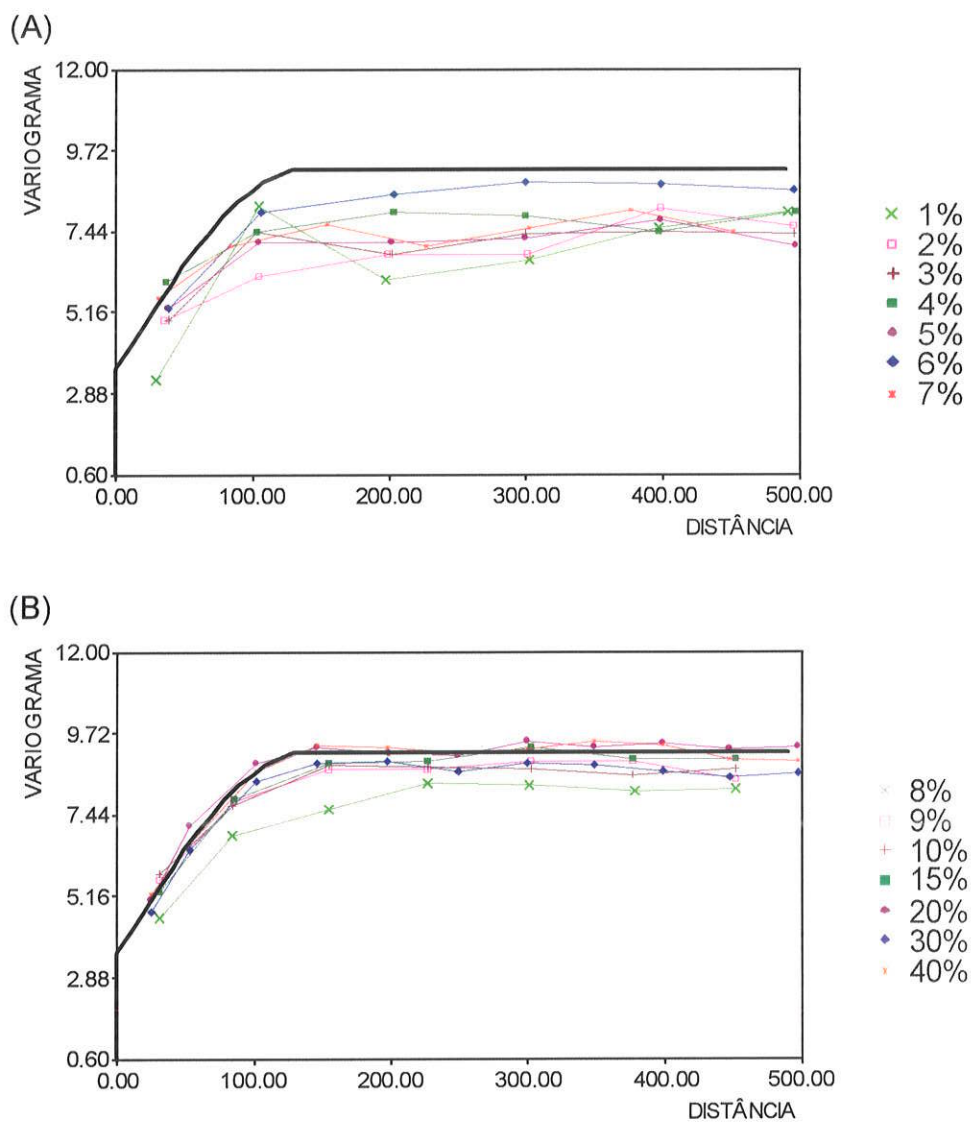


Figura 34: Variogramas experimentais médios, da variável simétrica, amostrada por amostragem aleatória simples, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 7% e em (B) os níveis percentuais de 10% a 40%.

4.3.1.3 Variável Assimétrica Negativa

A Figura 35 apresenta os variogramas experimentais e o modelo de variograma dos dados de referência, para a variável assimétrica negativa. As amplitudes, para os variogramas experimentais dos níveis percentuais entre 1 e 7%, tendem a convergir para a amplitude do modelo de referência. Os efeitos pepita não convergem, exceto para o nível percentual de 7%. Os patamares oscilam em torno do patamar do modelo de referência. Como os efeitos pepita e os patamares variam muito, em torno do modelo de referência, considera-se que os variogramas experimentais deste conjunto de níveis percentuais não convergem satisfatoriamente.

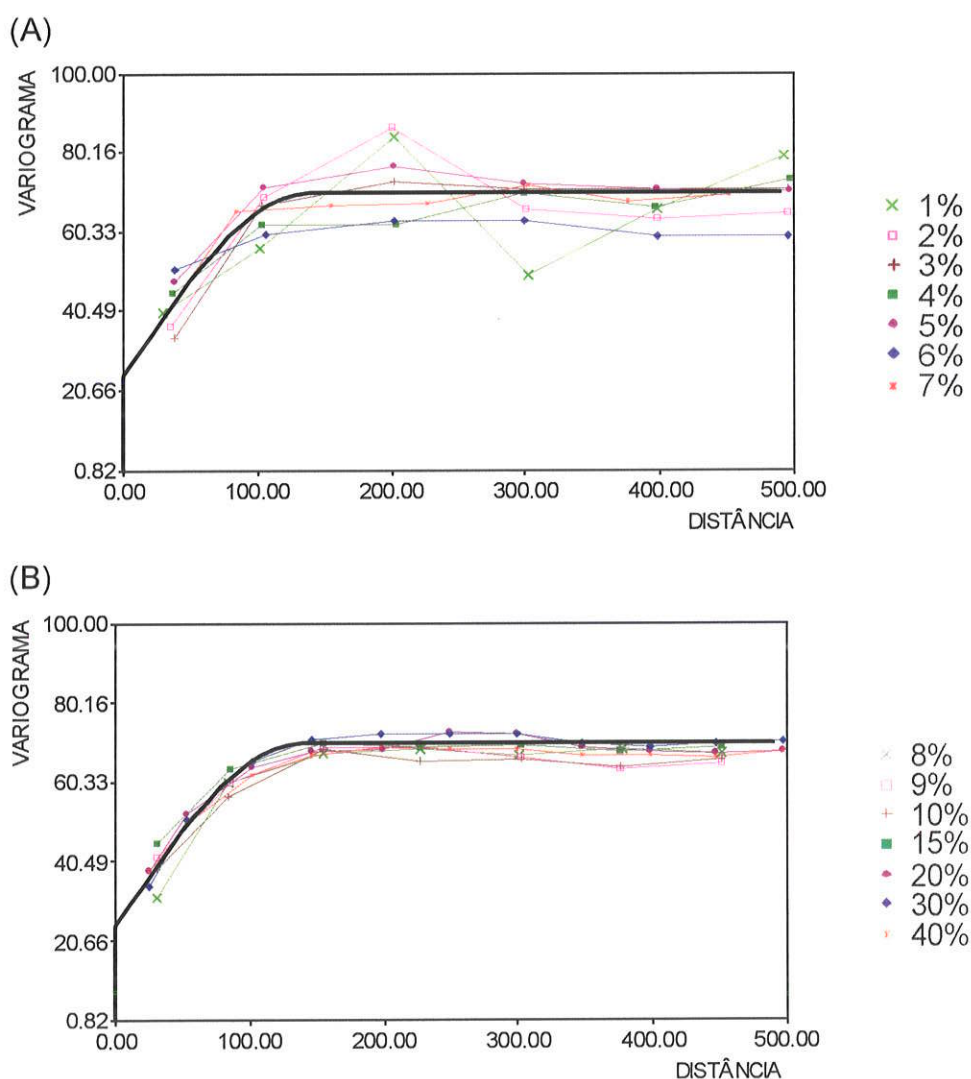


Figura 35: Variogramas experimentais médios, da variável assimétrica negativa, amostrada por amostragem aleatória simples, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 7% e em (B) os níveis percentuais de 8% a 40%.

A convergência, dos variogramas experimentais dos níveis percentuais de 8 a 40% (Figura 35-B), é melhor que a do conjunto anterior, de modo geral, pode-se considerar que os variogramas experimentais convergem ao modelo de variograma dos dados de referência. Observa-se que os efeitos pepita dos níveis percentuais de 8 e 15% são os únicos que não convergem.

4.3.2 Amostragem Aleatória Estratificada

Para esta amostragem, o procedimento adotado, para o cálculo dos variogramas experimentais, foi o mesmo que para a amostragem aleatória simples. Os parâmetros utilizados são os apresentados na Tabela 24. Na Tabela 25 podem ser observados os parâmetros dos modelos de variograma ajustados aos variogramas médios de cada nível percentual.

Tabela 24: Parâmetros do cálculo dos variogramas experimentais.

Nível Percentual	Nº Passos	Passo
1%	5	100
2%	5	100
3%	5	100
4%	5	100
5%	5	100
6%	5	100
8%	7	75
10%	7	75
15%	7	75
20%	10	50
30%	10	50
40%	10	50

4.3.2.1 Variável Assimétrica Positiva

A Figura 36-A apresenta os variogramas experimentais dos níveis percentuais de 1 a 8% e na Figura 36-B são apresentados os variogramas experimentais dos níveis de 10 a 40%. Como pode ser observado na Figura 36-A, o modelo de variograma dos dados de referência não é reproduzido em nenhum dos níveis percentuais.

Para os variogramas experimentais dos níveis percentuais de 10 a 40% (Figura 36-B), a convergência é melhor do que a dos níveis percentuais de 1 a 8%, exceto o variograma experimental do nível percentual de 30%, que apresenta efeito pepita e patamar muito mais elevados que o do modelo de referência.

Tabela 25: Modelos de variogramas ajustados aos dados selecionados por amostragem aleatória estratificada.

	Assimétrica Positiva			Simétrica			Assimétrica Negativa					
	Modelo	a	c ₀	c	Modelo	a	c ₀	c	Modelo	a	c ₀	c
1%	Esférico	288.0	0.540	0.550	P. Puro	---	5.660	---	Esférico	240.0	19.2	46.80
2%	Esférico	305.5	0.050	0.930	Esférico	132.0	2.130	3.710	Esférico	222.1	0.0	64.90
3%	P. Puro	---	2.651	---	P. Puro	---	7.640	---	P.Puro	---	75.0	---
4%	P. Puro	---	1.440	---	Esférico	260.4	5.133	3.100	Esférico	137.2	38.6	27.90
5%	Esférico	160.0	0.621	0.660	Esférico	200.0	1.665	7.000	Esférico	162.0	8.8	60.00
6%	P. Puro	---	2.490	---	Esférico	180.0	5.600	3.700	Esférico	200.0	40.9	30.40
8%	P. Puro	---	1.700	---	Esférico	140.0	4.800	3.475	Esférico	200.0	39.4	29.00
10%	Esférico	128.0	0.360	1.650	Esférico	160.0	4.180	5.000	Esférico	161.0	30.0	40.00
15%	Esférico	120.0	0.335	1.086	Esférico	151.2	4.060	5.200	Esférico	148.9	27.9	38.00
20%	Esférico	98.2	1.050	0.900	Esférico	136.8	3.640	5.800	Esférico	139.2	28.8	40.80
30%	Esférico	110.0	1.610	1.780	Esférico	137.0	3.410	5.800	Esférico	170.0	32.6	38.50
40%	Esférico	103.4	0.968	1.059	Esférico	143.0	3.380	5.900	Esférico	143.5	28.0	39.00
Ref.	Esférico	85.0	0.757	1.300	Esférico	137.0	3.590	5.600	Esférico	137.0	24.50	45.50

Onde a= amplitude, c₀= efeito pepita e c= variância espacial.

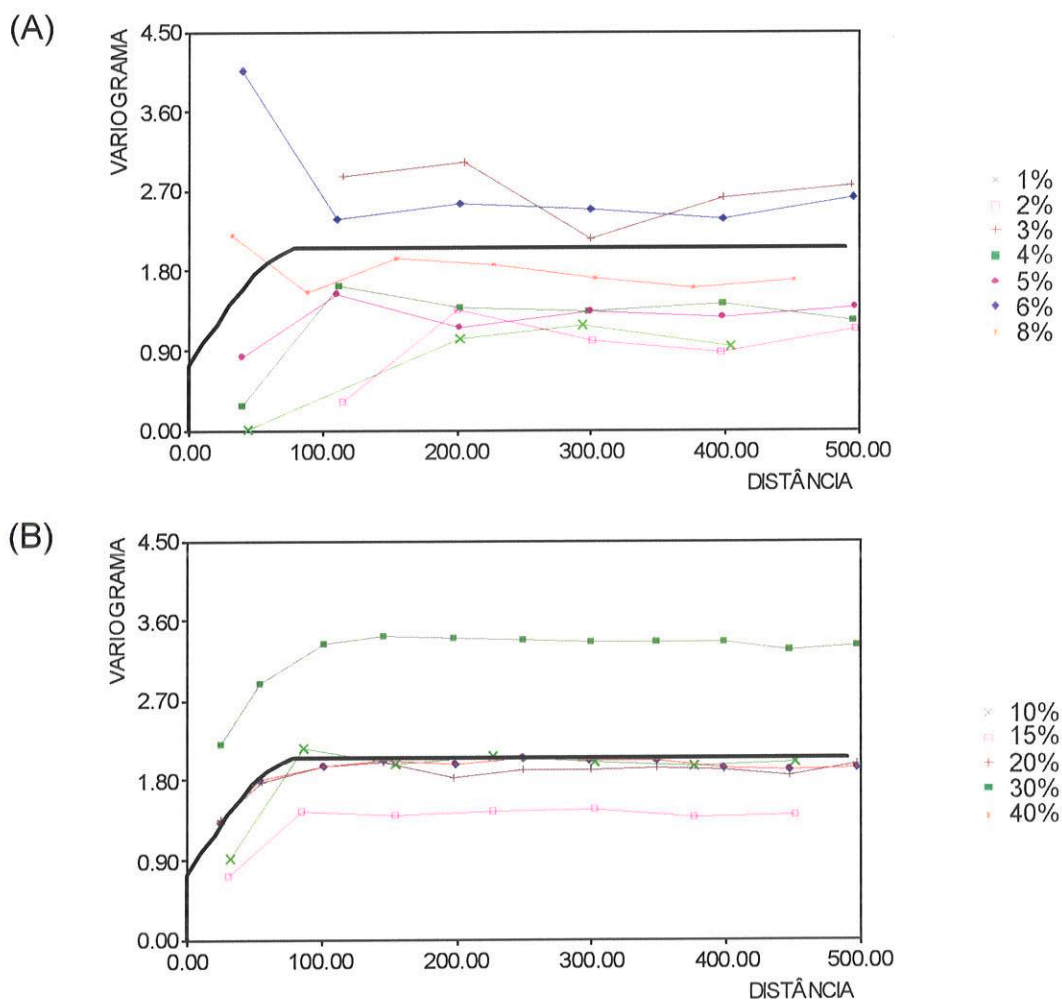


Figura 36: Variogramas experimentais médios, da variável assimétrica positiva, amostrada por amostragem aleatória estratificada, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 7% e em (B) os níveis percentuais de 10% a 40%.

4.3.2.2 Variável Simétrica

Os variogramas experimentais da variável simétrica são apresentados na Figura 37. Para os níveis percentuais entre 1 e 8% (Figura 37-A), não há convergência, dos variogramas experimentais calculados nos níveis percentuais, ao modelo de variograma dos dados de referência. Por outro lado, os variogramas experimentais dos níveis percentuais entre 10 e 40% (Figura 37-B), observa-se que existe convergência muito satisfatória. Esta convergência se dá para todos os parâmetros do variograma.

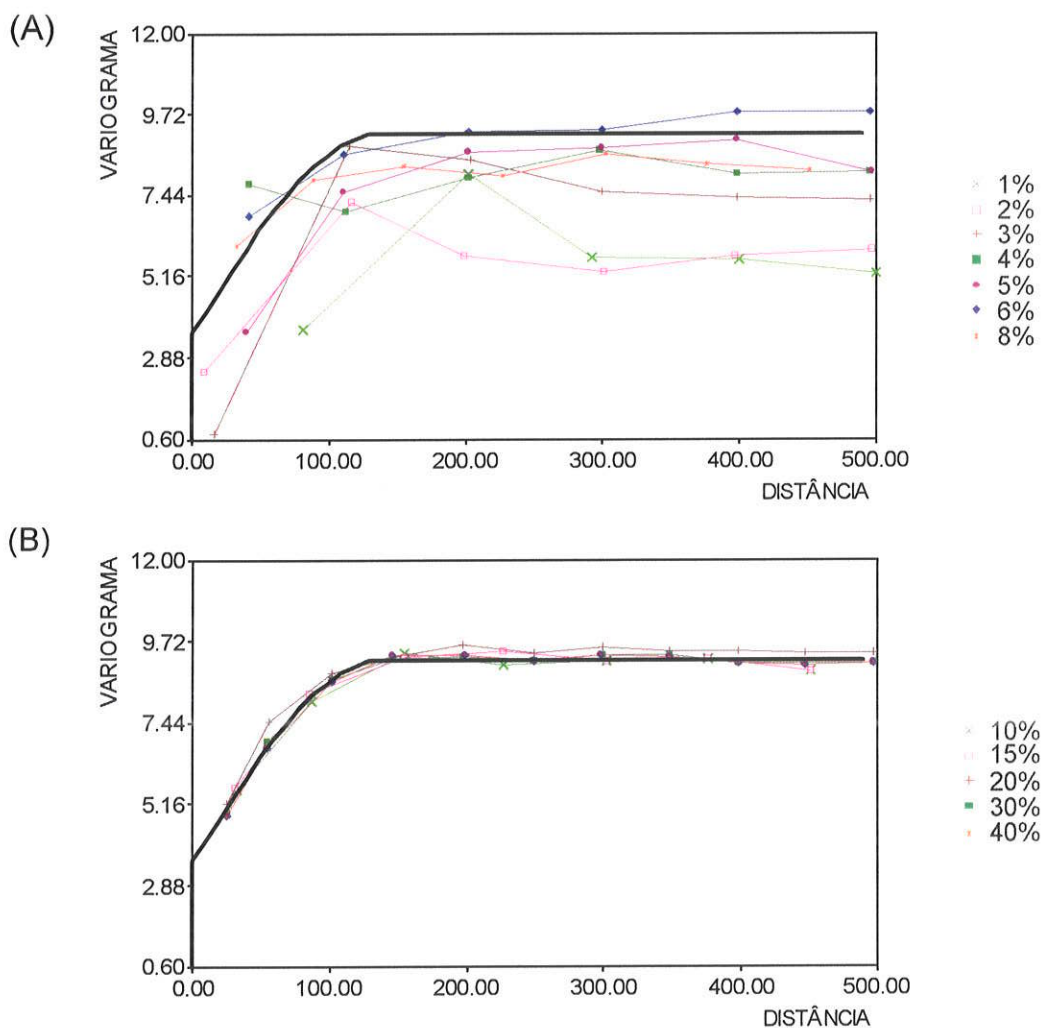


Figura 37: Variogramas experimentais médios, da variável simétrica, amostrada por amostragem aleatória estratificada, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 8% e em (B) os níveis percentuais de 10% a 40%.

4.3.2.3 Variável Assimétrica Negativa

Os variogramas experimentais da variável assimétrica negativa encontram-se representados na Figura 38, assim como o modelo de variograma ajustado ao variograma experimental desta variável nos dados de referência.

Nos níveis percentuais de 1 a 8% (Figura 38-A), apenas os patamares dos variogramas tendem a convergir ao patamar do modelo de variograma dos dados de referência, as amplitudes tendem a ser maiores e os efeitos pepita oscilam muito em torno do efeito pepita do modelo.

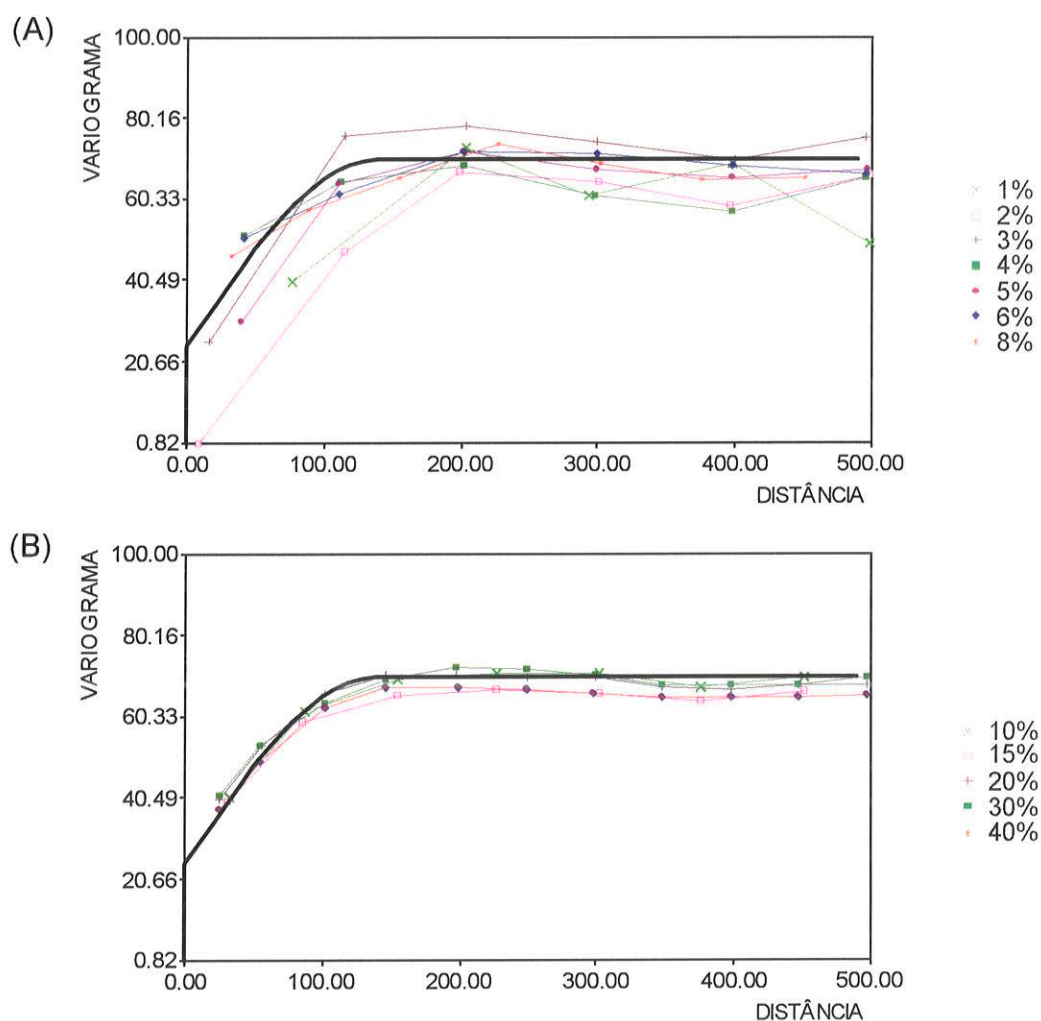


Figura 38: Variogramas experimentais médios, da variável assimétrica negativa, amostrada por amostragem aleatória estratificada, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 8% e em (B) os níveis percentuais de 10% a 40%.

Os níveis percentuais entre 10 e 40% (Figura 38-B) resultam em variogramas experimentais médios que convergem satisfatoriamente ao modelo de variograma dos dados de referência.

4.3.3 Amostragem Sistemática

As Tabelas 26 e 27 apresentam, respectivamente, os parâmetros utilizados no cálculo dos variogramas experimentais dos diversos níveis percentuais e os modelos teóricos de variogramas ajustados a cada variograma experimental.

Tabela 26: Parâmetros para o cálculo dos variogramas experimentais.

Nível percentual	Nº Passos	Passo
1%	4	200
2%	5	150
3%	5	100
4%	5	100
5%	7	75
6%	7	75
8%	7	75
10%	10	50
15%	10	50
20%	10	50
30%	10	40
40%	10	50
50%	15	40

Tabela 27: Modelos de variograma ajustados aos variogramas experimentais calculados para os dados coletados por amostragem sistemática.

Assimétrica Positiva				Simétrica				Assimétrica Negativa				
Modelo	a	c ₀	c	Modelo	a	c ₀	c	Modelo	a	c ₀	c	
1%	P. Puro	---	0.9620	---	P. Puro	---	8.717	---	P. Puro	---	53.64	---
2%	P. Puro	---	0.8188	---	P. Puro	---	7.648	---	Esférico	310.0	31.90	19.20
3%	P. Puro	---	0.9350	---	P. Puro	---	6.130	---	P. Puro	---	78.00	---
4%	Esférico	304.7	0.8300	0.410	P. Puro	---	4.520	---	P. Puro	---	57.60	---
5%	P. Puro	---	0.5410	---	Esférico	244.2	6.720	4.9	Esférico	213.0	47.50	17.80
6%	P. Puro	---	1.8900	---	Esférico	175.0	5.315	6.1	Esférico	203.0	39.60	30.00
8%	P. Puro	---	1.1000	---	P. Puro	---	9.060	---	Esférico	164.0	39.80	19.68
10%	Esférico	160.2	0.9200	0.230	Esférico	198.0	6.890	3.0	Esférico	140.0	32.00	31.50
15%	P. Puro	---	1.6400	---	Esférico	178.2	5.500	3.7	Esférico	194.7	51.30	28.00
20%	Esférico	112.8	1.9700	0.300	Esférico	163.0	4.910	4.3	Esférico	171.0	40.29	29.50
30%	Esférico	126.0	1.3100	0.600	Esférico	151.0	4.100	5.2	Esférico	157.0	34.06	39.00
40%	Esférico	131.2	1.3100	0.730	Esférico	172.9	4.350	5.1	Esférico	152.0	31.77	34.80
50%	Esférico	96.4	0.7570	1.350	Esférico	137.0	3.590	5.6	Esférico	137.0	24.50	45.50
Ref.	Esférico	85.0	0.757	1.300	Esférico	137.0	3.590	5.600	Esférico	137.0	24.50	45.50

Onde a= amplitude, c₀= efeito pepita e c= variância espacial.

4.3.3.1 Variável Assimétrica Positiva

Os variogramas experimentais da variável assimétrica positiva, para os níveis percentuais de 1 a 8% (figura 39-A), não convergem ao modelo de variância espacial dos dados de referência.

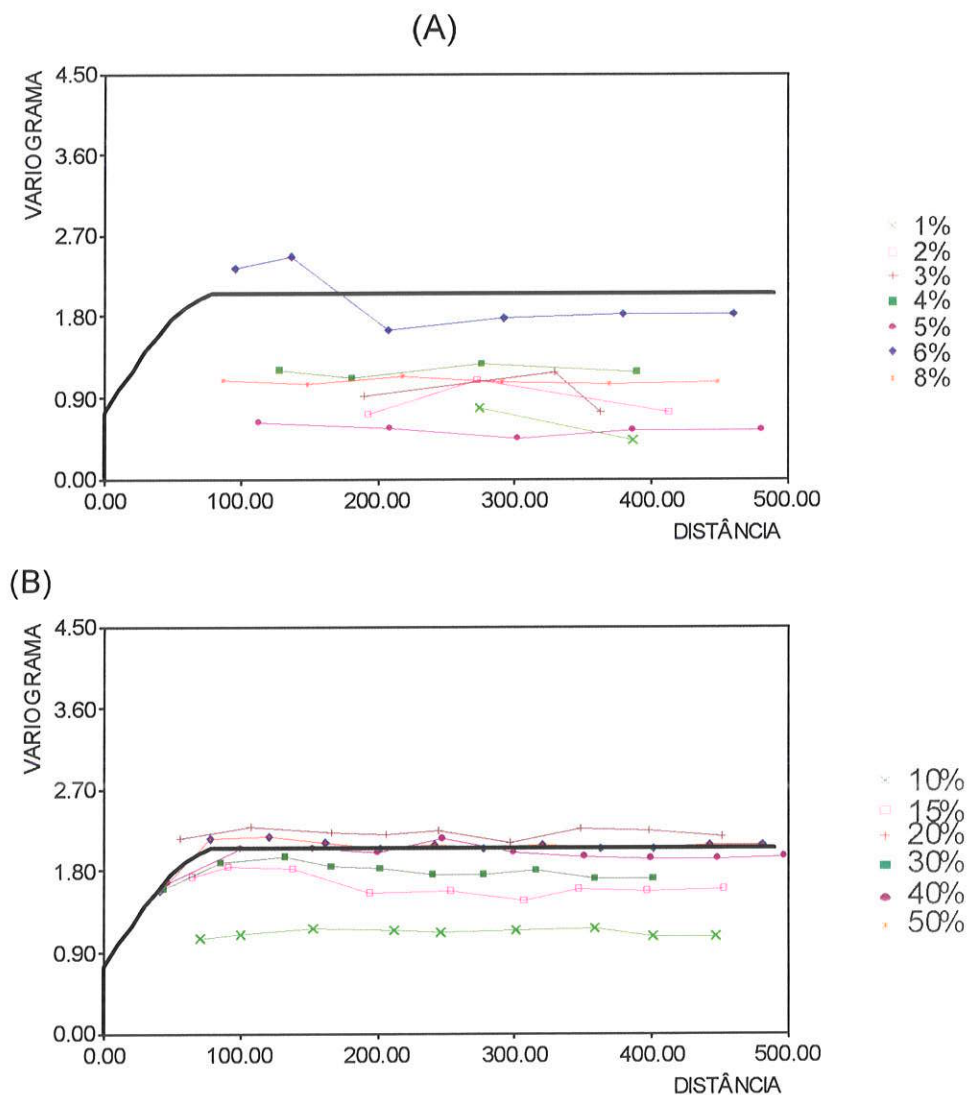


Figura 39: Variogramas experimentais médios, da variável assimétrica positiva, coletada por amostragem sistemática, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 8% e em (B) os níveis percentuais de 10% a 50%.

Os variogramas dos níveis percentuais de 10 a 30% (Figura 39-B) não convergem ao modelo de variograma dos dados de referência. Aqueles dos níveis de 40 e 50%, apresentam melhor aderência ao modelo de variograma de referência.

4.3.3.2 Variável Simétrica

Os variogramas experimentais da variável simétrica, calculados para os níveis percentuais de 1 a 8%, encontram-se representados na Figura 40-A. Nestes níveis percentuais, os variogramas experimentais não convergem ao modelo de variograma dos dados de referência.

Os níveis percentuais de 10 a 50% apresentam os variogramas experimentais mostrados na figura 40-B. Os variogramas destes níveis, aparentemente, convergem ao modelo de variograma dos dados de referência.

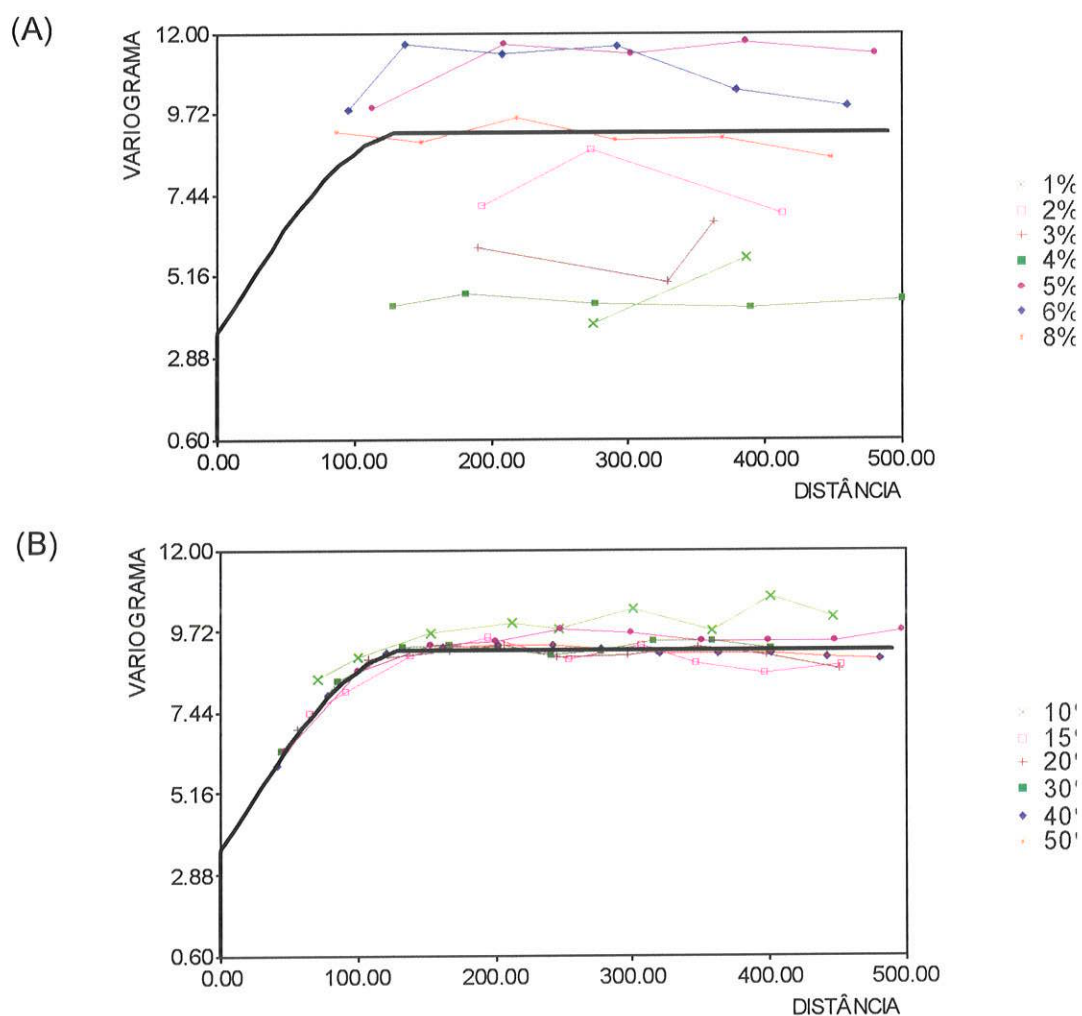


Figura 40: Variogramas experimentais médios, da variável simétrica, coletada por amostragem sistemática, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 8% e em (B) os níveis percentuais de 10% a 50%.

4.3.3.3 Variável Assimétrica Negativa

Para a variável assimétrica negativa, nos níveis percentuais de 1 a 8%, os variogramas experimentais (Figura 41-A) não convergem para o modelo de variograma de referência.

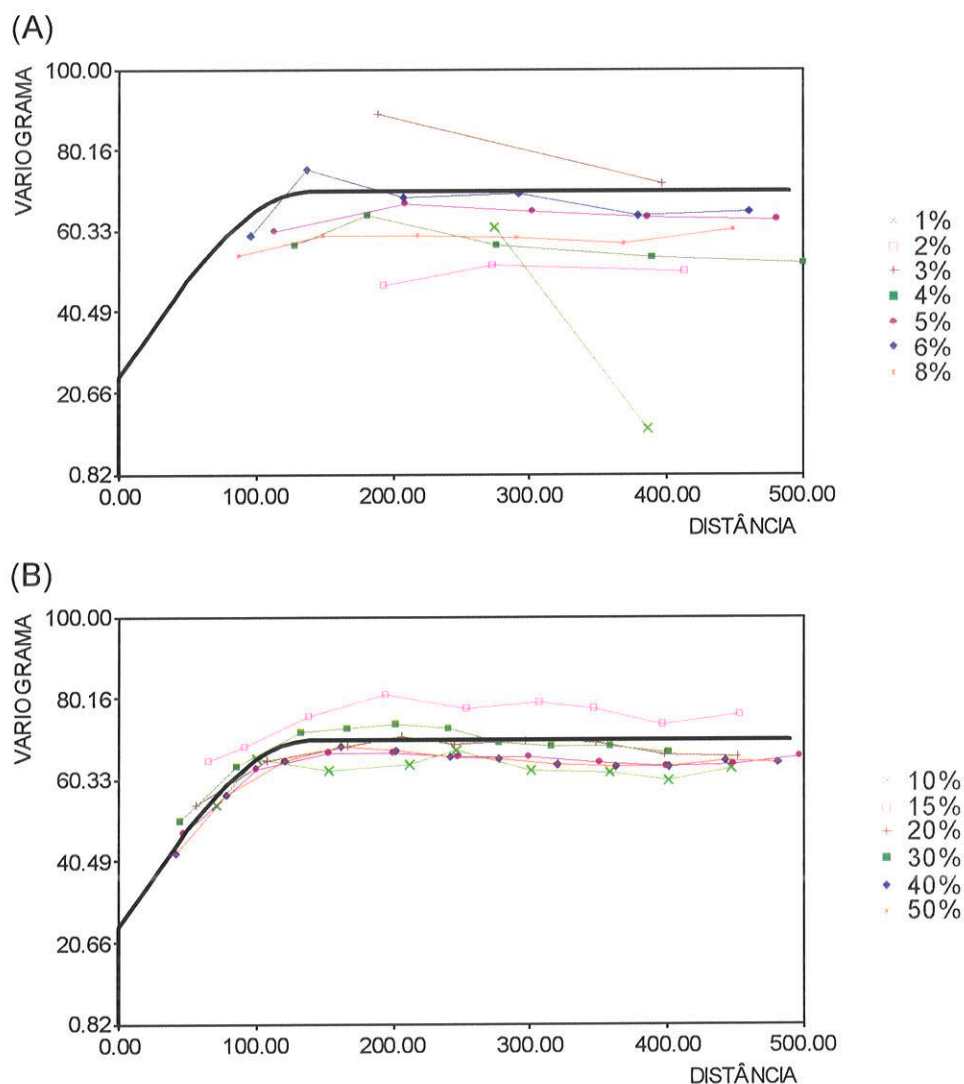


Figura 41: Variogramas experimentais médios, da variável assimétrica negativa, coletada por amostragem sistemática, e modelo teórico da imagem de referência. Em (A) os níveis percentuais de 1% a 8% e em (B) os níveis percentuais de 10% a 50%.

Os níveis percentuais mais elevados, entre 10 e 50%, apresentam variogramas experimentais estruturados (Figura 41-B) e convergentes ao modelo de variograma dos dados de referência. Destaca-se que o variograma do nível percentual de 15% é o único que não converge satisfatoriamente.

4.4 Análise de Superfícies de Tendência

A análise de superfícies de tendência foi realizada ajustando-se polinômios, de graus 1 a 3, aos dados dos níveis percentuais. O interesse por esta análise está em verificar se a componente de variação regional dos dados de referência é reproduzida pelos subconjuntos amostrados. A reprodução foi verificada por meio de estereogramas dos pólos dos planos resultantes das superfícies de grau 1 e por meio do erro quadrático médio entre as superfícies dos níveis percentuais e a superfície dos dados de referência. Destaca-se que o procedimento para o cálculo das distâncias entre os vetores, que representam a atitude das superfícies de tendência de grau um, sugerido em Landim (1997), foi realizado; porém, como esses resultados confirmaram aqueles resultados obtidos pelos estereogramas, optou-se por apresentar apenas os pólos dos planos nos estereogramas.

4.4.1 Amostragem Aleatória Simples

Os estereogramas dos pólos dos planos das superfícies de tendência de grau 1 das seis amostras selecionadas, em cada nível percentual, pelo método de amostragem aleatória simples encontram-se na Figura 42. Como se pode observar, a dispersão da variável assimétrica positiva é grande nos níveis mais baixos e diminui à medida que o tamanho n das amostras aumenta.

Similarmente, a Figura 43 representa os pólos dos planos da variável simétrica coletadas por amostragem aleatória simples. A dispersão dos pólos dos planos, em torno do pólo do plano dos dados de referência, comporta-se de maneira similar àquela da Figura 42, porém a aglutinação dos pólos é maior nesta variável. Cabe ressaltar que os planos, nesta figura, a partir do nível percentual de 10% mergulham todos para o mesmo quadrante do plano dos dados de referência. Para a variável assimétrica positiva isto ocorre apenas para os níveis percentuais de 8% , 30% e 40%.

Na Figura 44 podem ser observados os pólos dos planos das superfícies de tendência de grau 1, calculados para a variável assimétrica negativa. A dispersão dos pólos para esta variável é muito maior que a observada nos pólos das outras duas variáveis. Apenas no nível percentual de 9% todos os planos mergulham para o mesmo

quadrante e a dispersão dos pólos permanece alta até o nível percentual de 20%. Para os níveis percentuais mais baixos a dispersão é extremamente alta, sobretudo até o nível percentual de 15%. A maior dispersão para a variável assimétrica negativa deve-se ao formato da distribuição de frequência, que apresenta maior quantidade de valores na cauda esquerda. Isso faz com que aumente a probabilidade que os valores baixos da distribuição assimétrica negativa sejam amostrados e, conseqüentemente, maior dispersão, em termos de atitude das superfícies de grau 1.

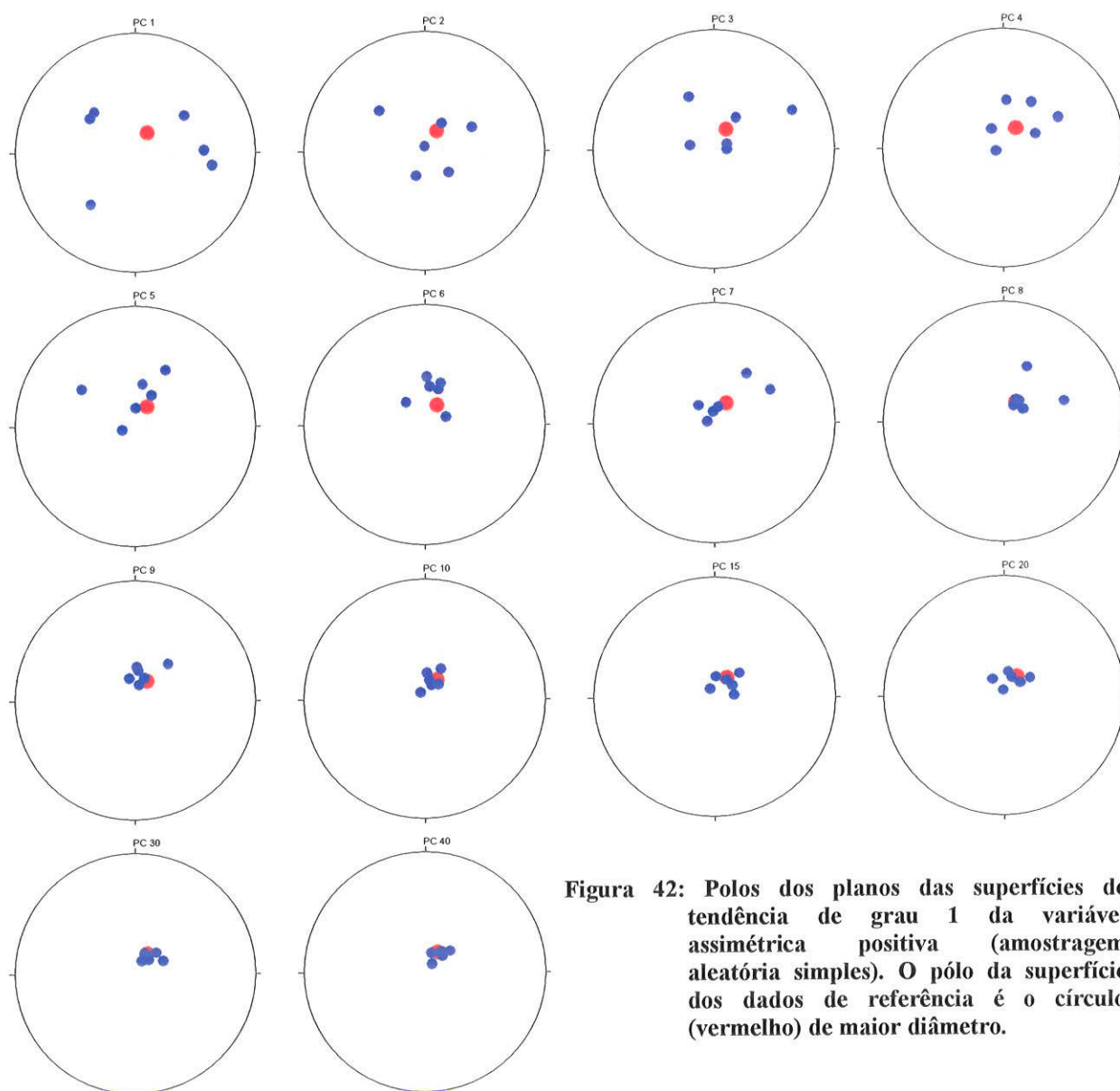


Figura 42: Polos dos planos das superfícies de tendência de grau 1 da variável assimétrica positiva (amostragem aleatória simples). O pólo da superfície dos dados de referência é o círculo (vermelho) de maior diâmetro.

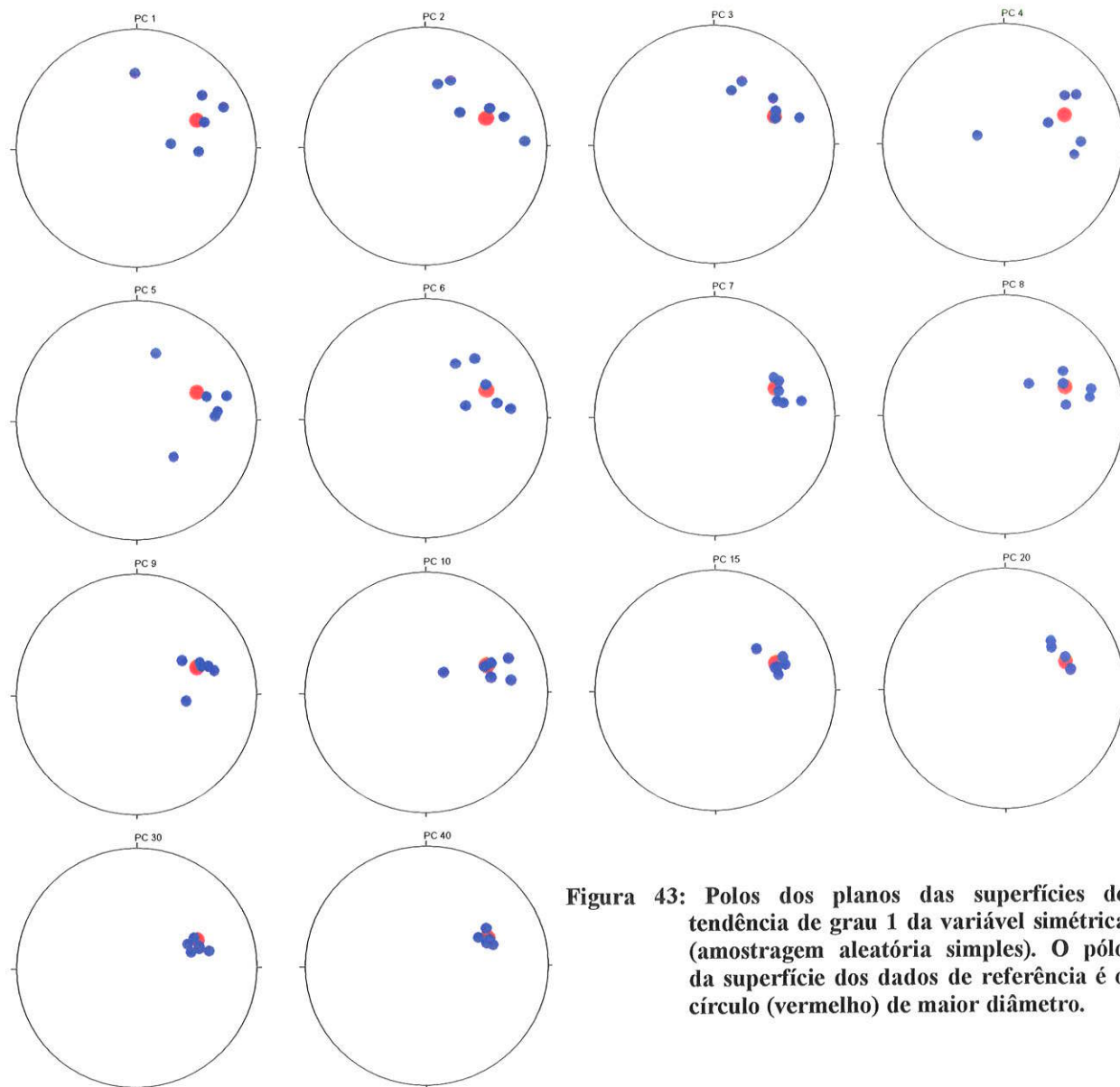


Figura 43: Polos dos planos das superfícies de tendência de grau 1 da variável simétrica (amostragem aleatória simples). O pólo da superfície dos dados de referência é o círculo (vermelho) de maior diâmetro.

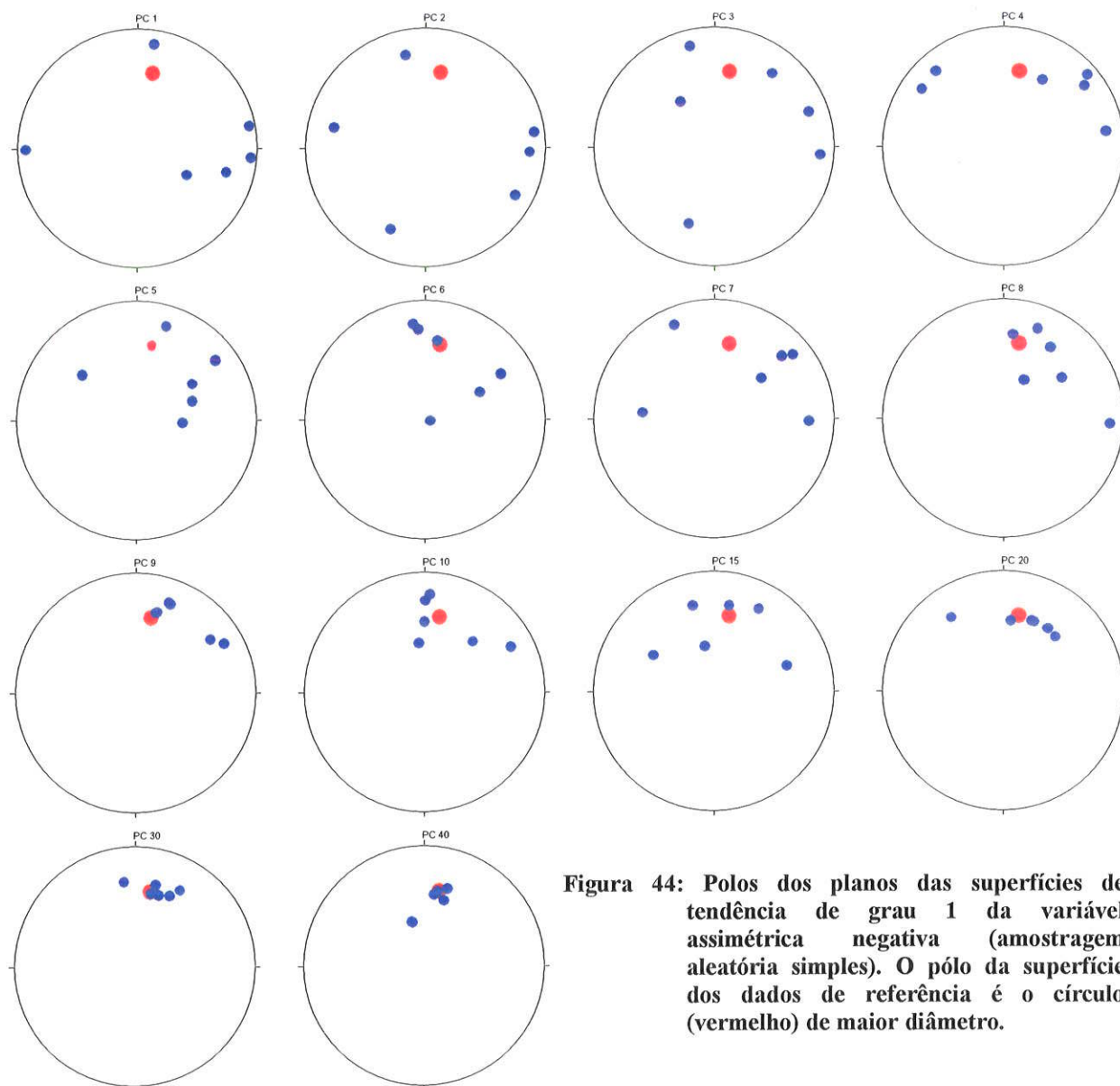


Figura 44: Polos dos planos das superfícies de tendência de grau 1 da variável assimétrica negativa (amostragem aleatória simples). O pólo da superfície dos dados de referência é o círculo (vermelho) de maior diâmetro.

O erro quadrático médio (RMS), em relação à superfície de referência, calculado para as superfícies de graus 1, 2 e 3, é mostrado, respectivamente, nas Figuras 45, 46 e 47. Na Figura 45, pode-se verificar que à medida que aumenta o n das amostras, o RMS diminui. Destaca-se, também, o comportamento deste erro para a variável assimétrica positiva que é menor que o das variáveis simétrica e assimétrica negativa. Os erros da superfície de grau 2 das três variáveis têm o mesmo comportamento da superfície de grau 1, como pode ser visto na Figura 46. Os erros para as superfícies de grau 3 (Figura 47) são extremos para o nível percentual de 1%. A partir desse valor, descontando-se o efeito de escala, os erros RMS das variáveis analisadas comportam-se da maneira semelhante às anteriores.

Nestes gráficos, verifica-se que a variável assimétrica negativa apresenta maiores erros que a simétrica e esta, por sua vez, maiores erros RMS que a assimétrica positiva. Aparentemente essa inversão ocorre devido à medida utilizada, no caso, o erro RMS, que é uma medida direta dos desvios e, portanto, correlacionável aos desvios padrão (Tabelas 4, 7 e 9).

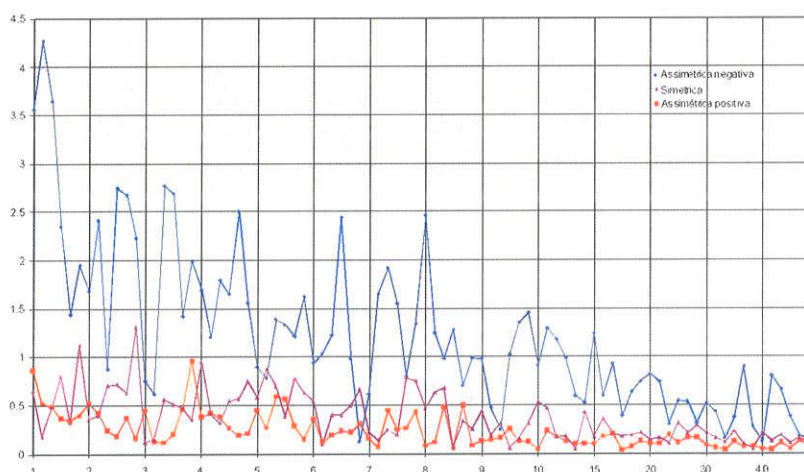


Figura 45: Diagrama dos erros quadráticos médios em relação à superfície (grau 1) dos dados de referência para as três variáveis analisadas, coletadas por amostragem aleatória simples.

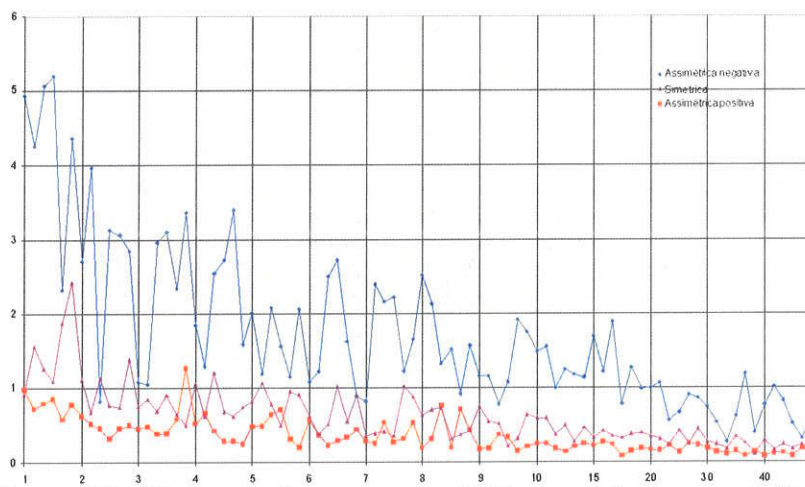


Figura 46: Diagrama dos erros quadráticos médios em relação à superfície (grau 2) dos dados de referência para as três variáveis analisadas, coletadas por amostragem aleatória simples.

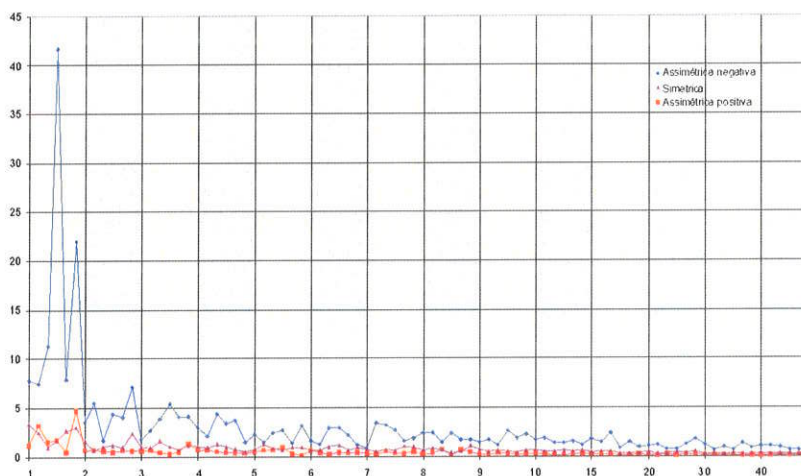


Figura 47: Diagrama dos erros quadráticos médios em relação à superfície (grau 3) dos dados de referência para as três variáveis analisadas, coletadas por amostragem aleatória simples.

4.4.2 Amostragem Aleatória Estratificada

Os estereogramas da Figura 48, representam os pólos dos planos das superfícies de tendência de grau 1, da variável assimétrica positiva. Nestes gráficos pode-se observar que a dispersão dos pólos, em torno do pólo de referência, diminui à medida que o n das amostras aumenta. A aglutinação dos pontos torna-se maior a partir do nível percentual de 8%, quando os pólos mergulham em sua maioria para o mesmo quadrante do pólo da superfície de tendência dos dados de referência.

Os estereogramas dos pólos dos planos da tendência apresentada pela variável simétrica (Figura 49), mostram o mesmo comportamento dos estereogramas da variável assimétrica positiva. Para a variável em questão, a dispersão diminui já a partir do nível percentual de 4%, onde os mergulhos dos pólos tendem para o mesmo quadrante do pólo de referência. Nos níveis percentuais superiores a 15%, inclusive, a aglutinação dos pólos é elevada.

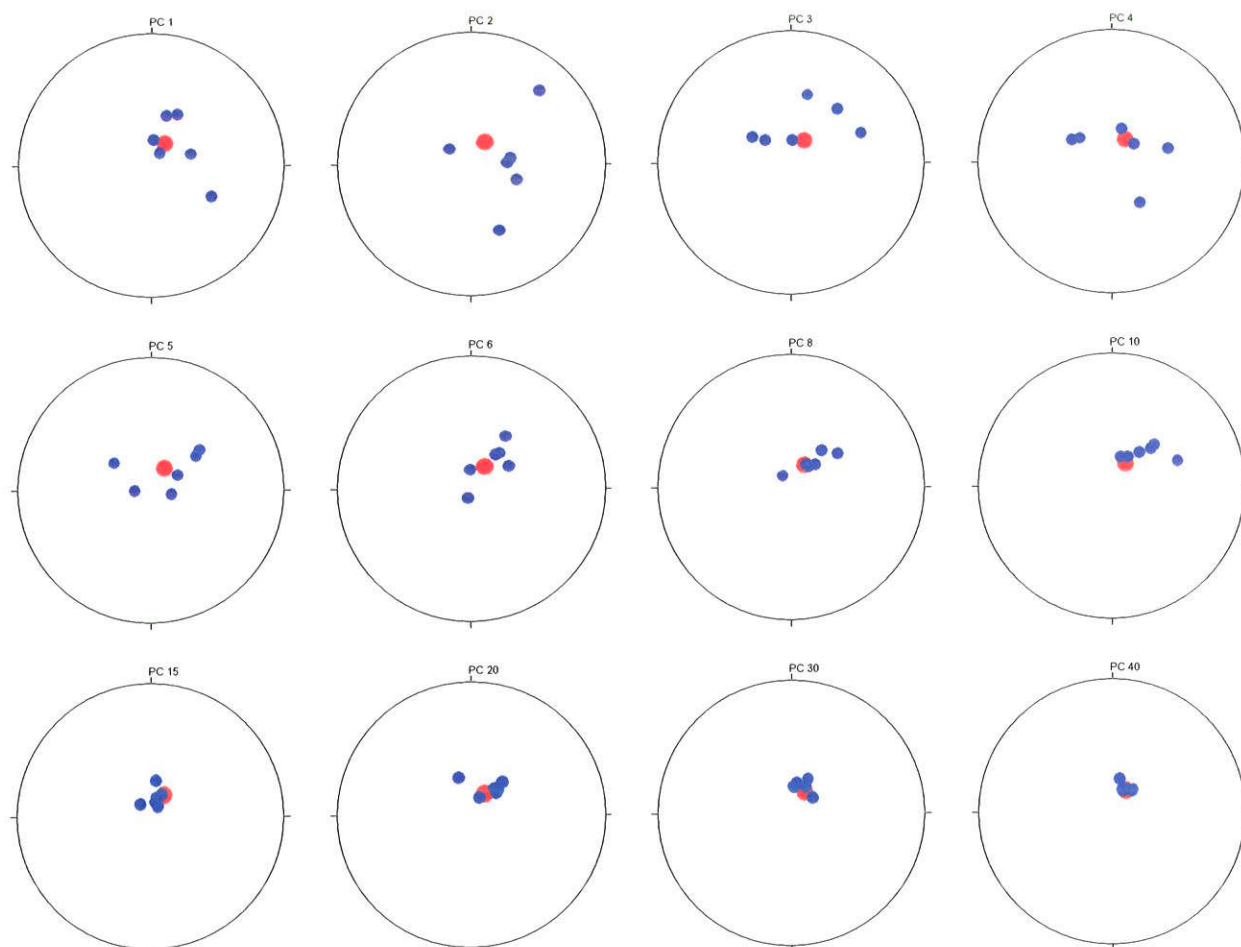


Figura 48: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável assimétrica positiva. O ponto de maior diâmetro é o pólo do plano dos dados de referência, para a amostragem aleatória estratificada.

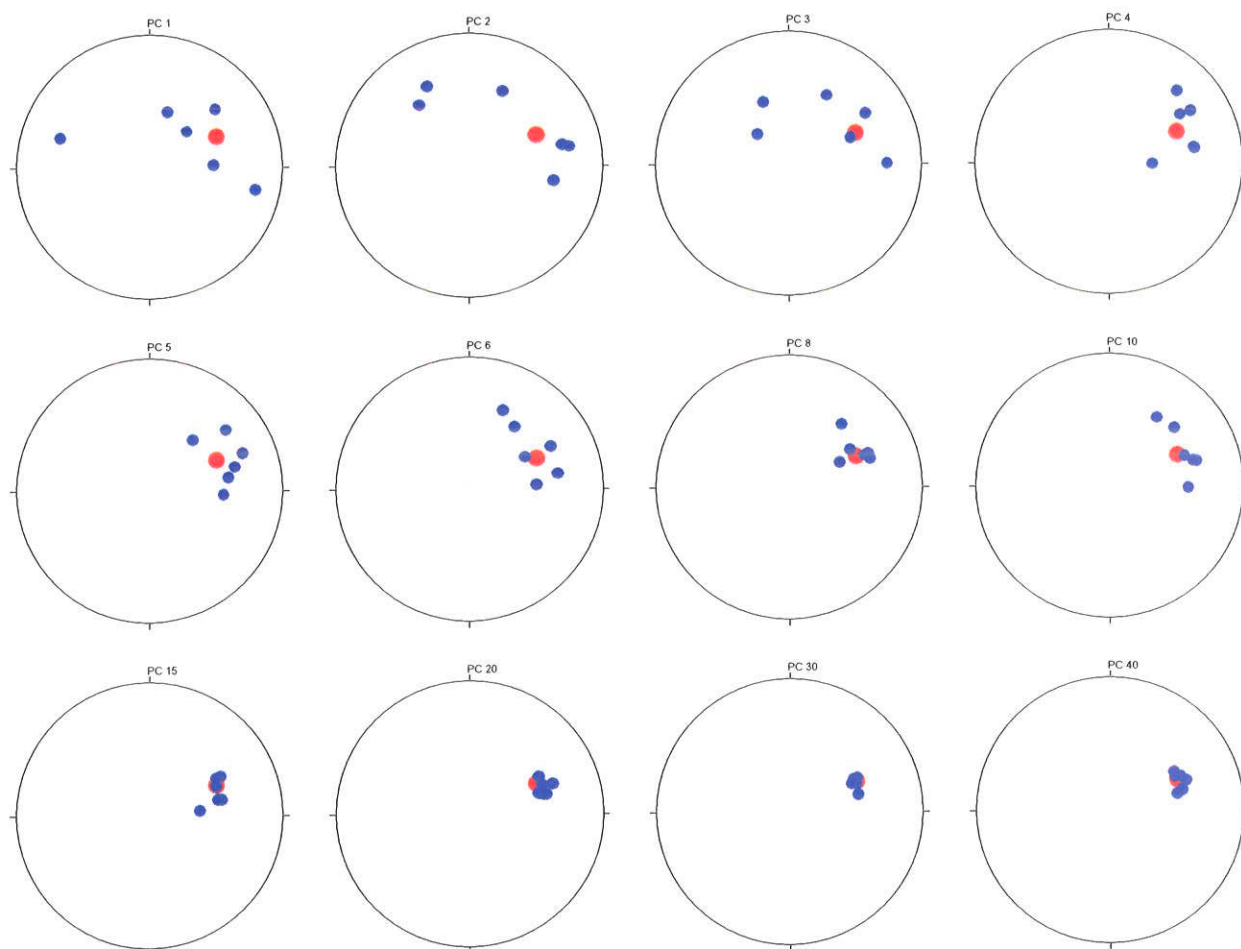


Figura 49: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável simétrica. O ponto de maior diâmetro é o pólo do plano dos dados de referência, para a amostragem aleatória estratificada.

A dispersão dos pólos das superfícies de tendência de grau 1, da variável assimétrica negativa (Figura 50), é elevada para os níveis percentuais entre 1% e 5%. Embora esta dispersão seja menor nos níveis percentuais maiores que os referidos, observa-se que esta ainda é alta se comparada à dispersão dos pólos dos planos da variável assimétrica positiva e da variável simétrica. No caso da variável assimétrica negativa, a dispersão somente diminui nos níveis percentuais de 30% e 40%.

Os valores de RMS das superfícies de tendência de grau 1 (Figura 51) confirmam a tendência apresentada nas Figuras 48, 49 e 50, pois, como pode ser observado, os erros diminuem à medida que o n das amostras aumenta e ainda que os erros da variável assimétrica negativa são mais elevados. Para a variável simétrica e para a variável assimétrica positiva, os erros são próximos em valor, diferenciando-se

apenas nos níveis percentuais mais baixos, mais especificamente 1% e 2%, onde a variável simétrica apresenta erros maiores e mais variáveis.

As superfícies de tendência de grau 2, para as três variáveis analisadas, apresentam erros quadráticos médios (Figura 52) com comportamento semelhante ao mesmo das superfícies de grau 1. Porém, para estas superfícies, os erros são levemente maiores e, portanto, com maiores variações.

A Figura 53 mostra o comportamento dos erros calculados para as superfícies de tendência de grau 3 de cada amostra. Como a dispersão dos erros, da variável assimétrica negativa, no primeiro nível percentual é extremamente elevada, os erros das demais variáveis aparentam ser pouco dispersos e pequenos. Os erros RMS das superfícies de grau 3, da variável assimétrica negativa, são extremamente elevados e variáveis nos níveis percentuais entre 1% e 5% e tendem a estabilizar nos níveis percentuais superiores a 8%. Nesta mesma figura, a variável assimétrica positiva é a que apresenta menores erros e também os erros com menor variação, sendo que para os níveis percentuais de 30% e 40% esta variação é mínima, menor do que 0,1. Os erros obtidos para a variável simétrica, no primeiro nível percentual, são elevados e apresentam alta variação, chegando ao máximo de 8,85. Esta variação diminui à medida que o n das amostras aumenta e a partir do nível percentual de 8% os valores dos erros permanecem abaixo de 1,0 e acima do nível de 30% este erro é menor que 0,4.

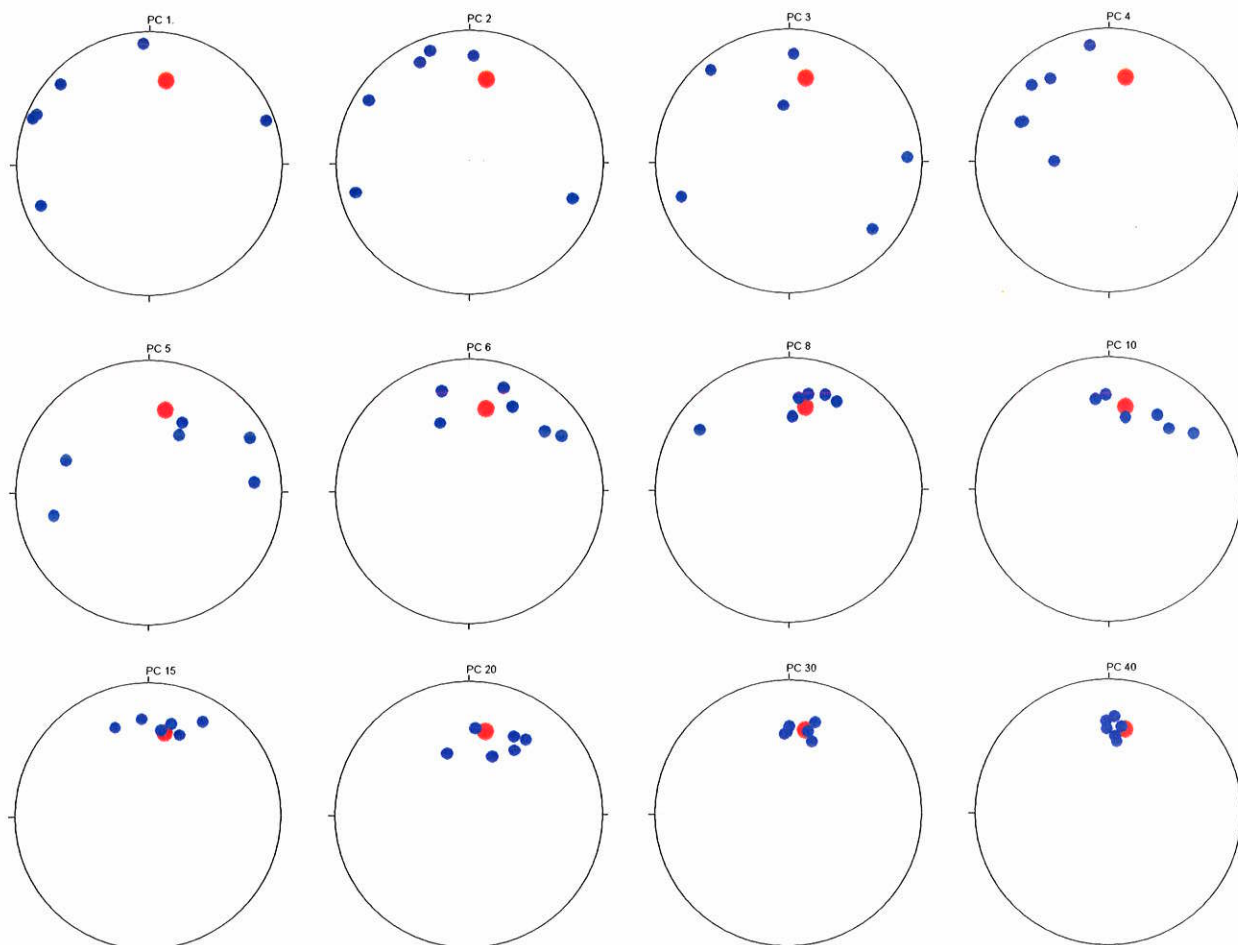


Figura 50: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável assimétrica negativa. O ponto de maior diâmetro é o pólo do plano dos dados de referência, para a amostragem aleatória estratificada.

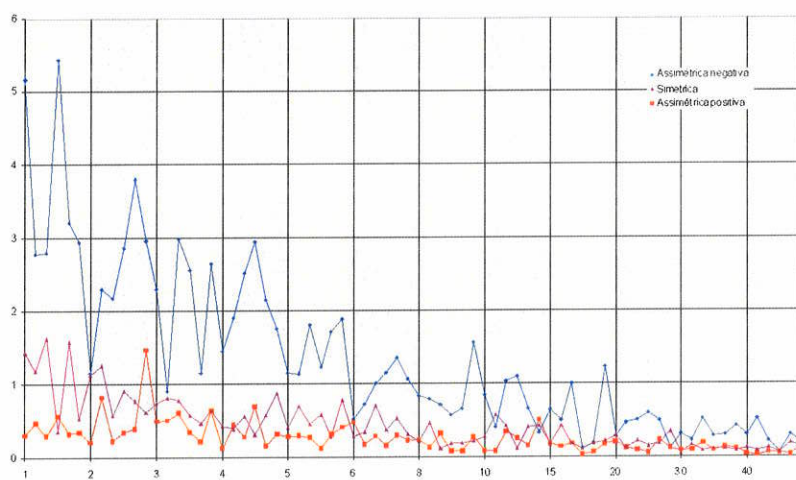


Figura 51: RMS das superfícies de tendência de grau 1, calculadas para as amostras coletadas por amostragem aleatória estratificada.

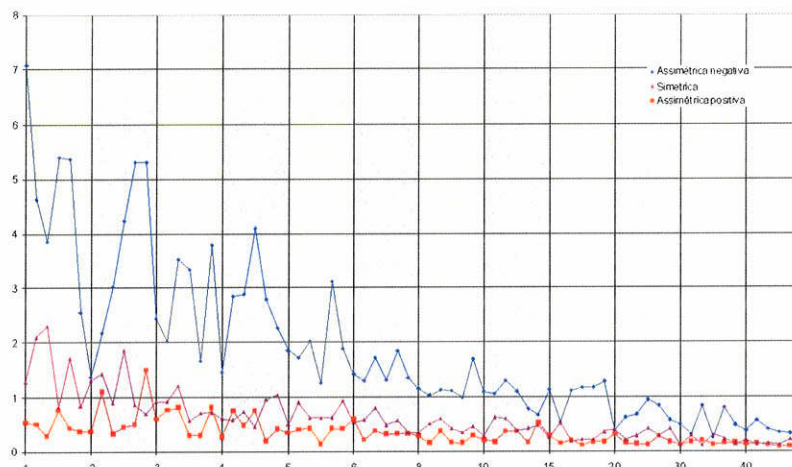


Figura 52: RMS das superfícies de tendência de grau 2 calculadas para as amostras selecionadas por amostragem aleatória estratificada.

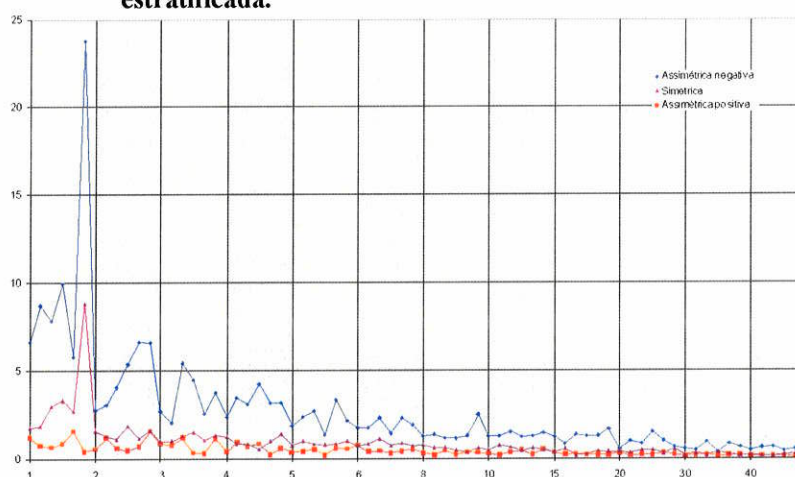


Figura 53: RMS calculado para as superfícies de tendência de grau 3 calculadas para as amostras selecionadas por amostragem aleatória estratificada.

4.4.3 Amostragem Sistemática

Uma vez que, para cada nível percentual, foi retirada apenas uma amostra, optou-se por fazer apenas três estereogramas, um para cada variável, contendo todos os níveis percentuais.

Os pólos das superfícies de grau 1, para variável assimétrica positiva, possuem dispersão média em torno do pólo do plano dos dados de referência (Figura 54). Os pólos dos planos dos níveis percentuais de 1%, 2% e 4% são os mais afastados do plano ajustado aos dados de referência. Embora exista a convergência dos planos das amostras ao plano da referência, ressalta-se que para os níveis percentuais mais altos,

acima de 20%, não há coincidência exata de nenhum destes níveis com os dados de referência.

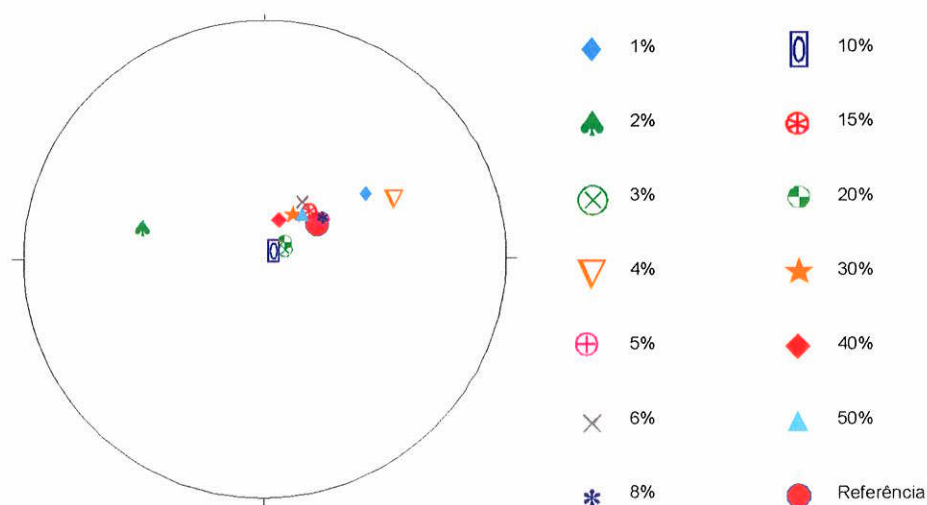


Figura 54: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável assimétrica positiva (amostragem sistemática).

Para a variável simétrica, a dispersão dos pólos dos planos (Figura 55) é menor que aquela observada na Figura 54. Observa-se, ainda, que a dispersão diminui à medida que o n das amostras aumenta e para os níveis percentuais de 30%, 40% e 50%. A convergência é elevada, sobretudo, para o último nível percentual. Outro ponto a ser destacado é que, exceto pelo nível percentual de 4%, todos os planos ajustados mergulham para o mesmo quadrante que a superfície de tendência dos dados de referência.

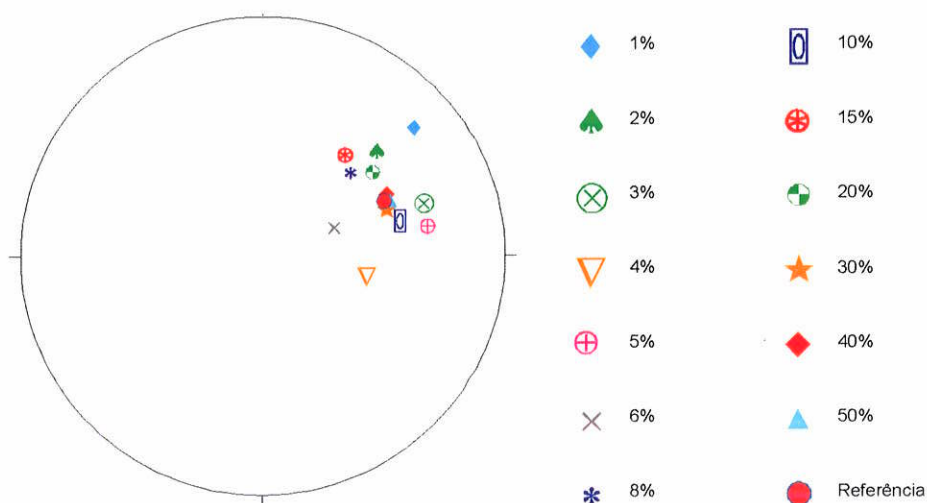


Figura 55: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável simétrica (amostragem sistemática).

O estereograma da Figura 56 representa os pólos dos planos de tendência de grau 1 da variável assimétrica negativa. Dos planos ajustados, apenas os dos níveis percentuais de 1%, 2% e 6% não se concentram próximo ao plano dos dados de referência. Os demais, embora não sejam coincidentes com este, espalham-se em seu entorno. É interessante notar que, o pólo do plano ajustado ao nível percentual de 50% está afastado do pólo do plano da referência.

A Figura 57, ilustra o comportamento dos erros RMS para a superfície de tendência de grau 1 em relação à superfície de referência. Pode-se observar, nesta figura, que o comportamento dos erros da variável assimétrica negativa é semelhante ao comportamento dos erros calculados, de maneira similar, com as amostras selecionadas pelos outros dois métodos de amostragem. Estes erros são elevados com grande variação para os primeiros níveis percentuais e decresce gradativamente à medida que o n das amostras aumenta. Os erros da variável simétrica variam pouco a partir do nível percentual de 2% e estão abaixo de 0,5 tendendo a aproximar-se de zero com o aumento do n . A variável assimétrica positiva possui erros variáveis nos primeiros níveis percentuais e estes erros diminuem já a partir do nível de 5% apresentando valores sempre menores que 0,15.

Os erros apresentados na Figura 58 foram calculados para superfícies de tendência de grau 2 em relação à mesma superfície dos dados de referência. Os erros da variável assimétrica negativa possuem comportamento similar àquele dos erros da superfície de grau 1, porém para esta superfície eles decrescem mais rapidamente que na superfície anterior. Embora a variável simétrica apresente erros menores que aqueles da variável assimétrica negativa, estes também são maiores nos níveis percentuais mais baixos e decrescem com o aumento do n das amostras. Observa-se, ainda, que para os níveis percentuais superiores a 6% os valores dos erros são inferiores a 0,5. Do mesmo modo, que os erros da variável assimétrica positiva para superfícies de grau 1, os de grau 2 apresentam os menores valores e a menor variação. Estes erros são menores que 0,5, já no nível percentual de 5%, a partir de onde mantêm-se praticamente constantes.

A Figura 59 mostra os erros calculados para as superfícies de tendência de grau 3 em relação à superfície, de mesmo grau, calculada para os dados de referência. Os erros da variável assimétrica negativa decrescem de modo semelhante aos erros

calculados para as superfícies de graus 1 e 2 à medida que o nível percentual aumenta. Os erros, tanto da variável simétrica quanto da variável assimétrica positiva, possuem comportamento semelhante ao das superfícies de graus 1 e 2, porém para ambas variáveis eles são ligeiramente mais altos que os anteriores.

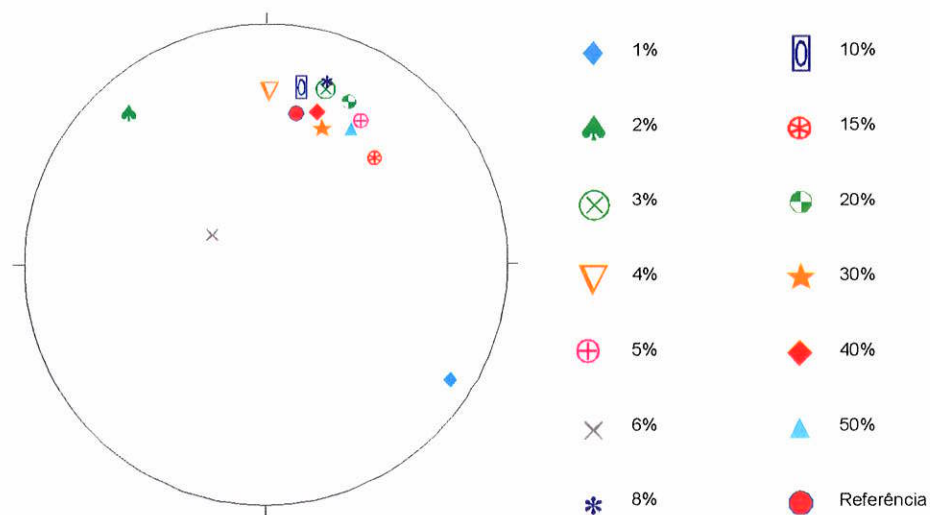


Figura 56: Pólos dos planos das superfícies de tendência de grau 1 calculados para a variável assimétrica negativa (amostragem sistemática).

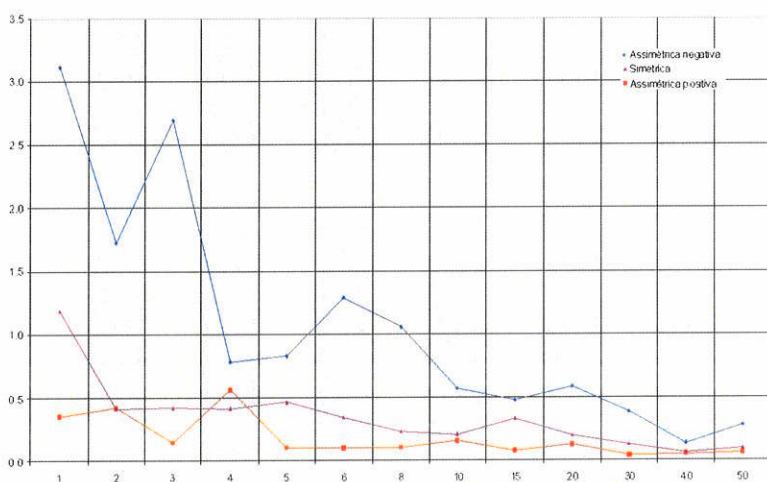


Figura 57: RMS calculado para as superfícies de tendência de grau 1 (amostragem sistemática).

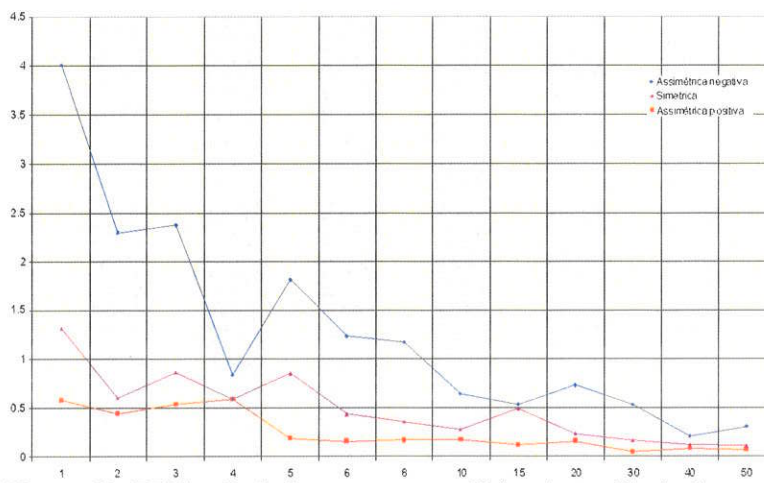


Figura 58: RMS calculado para as superfícies de tendência de grau 2 (amostragem sistemática).

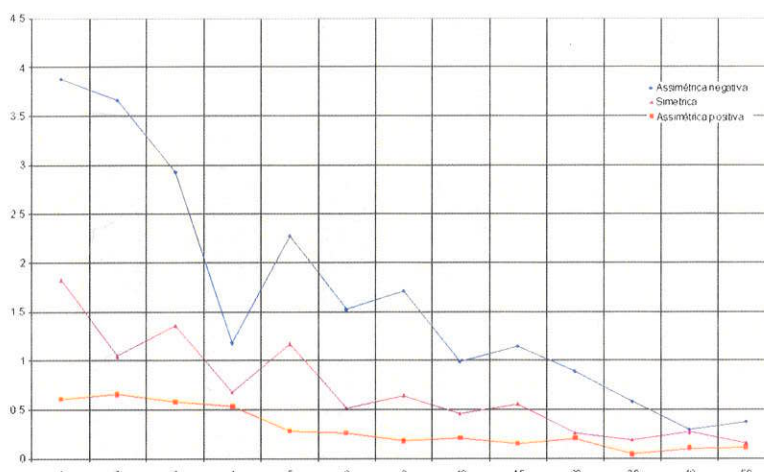


Figura 59: RMS calculado para as superfícies de tendência de grau 3 (amostragem sistemática).

4.5 Recursos Geológicos

Os recursos geológicos foram calculados como recursos totais das amostras, por meio do somatório dos recursos calculados para cada bloco, os recursos dos blocos foram obtidos pela aplicação direta da equação $R = Vdt$.

O cálculo dos recursos, da variável assimétrica positiva, utilizou como parâmetros: volume de cada bloco igual a 3.125m^3 , densidade aparente de $2,65\text{ ton/m}^3$ e o teor de corte aplicado é $0,356\text{ g/ton}$, os recursos desta variável estão expressos em quilogramas. Para a variável simétrica, representando a variável P_2O_5 apatítico, em uma depósito de fosfato, utilizou-se o mesmo volume que para a variável assimétrica positiva, a densidade aparente é $2,02\text{ ton/m}^3$ e o teor de corte adotado foi de 6%. Os recursos geológicos, neste caso, estão expressos em toneladas de minério.

Como a variável assimétrica negativa foi gerada com base em uma distribuição típica de um depósito de ferro, adotou-se um teor de corte de 53% e densidade aparente de 2,54 ton/m³, os recursos desta variável também estão expressos em tonelada de minério. O volume de cada bloco é igual a 3.125m³.

Neste ponto cabe abrir um parêntese, os recursos geológicos de alguns níveis percentuais, em todas as variáveis e em todos os métodos de amostragem, não foram calculados, pois seus variogramas experimentais não apresentaram estrutura, logo a krigagem não foi calculada para estes níveis percentuais. Justifica-se a não realização da krigagem para estes níveis utilizando-se o modelo de variograma dos dados de referência, pois, como já referido no Capítulo 1, o modelo de variograma de um depósito nunca é conhecido *a priori*, sempre *a posteriori*. Deste modo, os modelos de variogramas dos dados de referência, para todas as variáveis, são utilizados apenas como parâmetro de comparação e não como um parâmetro para o cálculo de estimativas.

4.5.1 Amostragem Aleatória Simples

4.5.1.1 Variável Assimétrica Positiva

Os recursos geológicos calculados, para os pontos selecionados por amostragem aleatória simples da variável assimétrica positiva, encontram-se ilustrados na Figura 60. Nesta figura, percebe-se que os recursos estimados são extremamente variáveis para todos os níveis percentuais e que apenas no nível de 30% os recursos estimados aproximam-se daqueles calculados para os dados de referência e contidos na fronteira convexa. Embora os recursos tenham apresentado grande variação, nota-se que esta variação decresce à medida que o *n* aumenta.

Os recursos, para alguns níveis percentuais, extrapolam os valores dos recursos totais e estes não acompanham os recursos da referência limitados pela fronteira convexa dos dados.

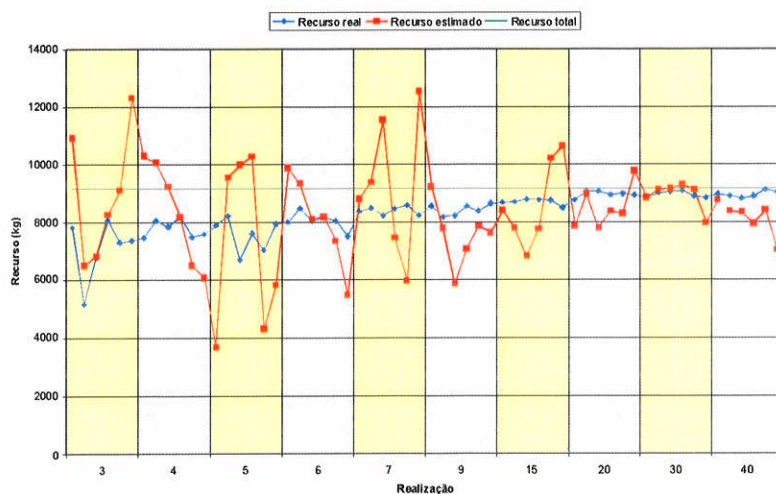


Figura 60: Recursos geológicos calculados para as amostras da variável assimétrica positiva (amostragem aleatória simples).

Foram calculadas curvas de parametrização dos recursos, para esta variável, estas curvas são apresentadas na Figura 61. Esta figura apresenta as curvas: *teor de corte x tonelagem* e *teor de corte x teor médio*, para o nível percentual de 9%, sexta realização (Figura 61-A). Observa-se, nesta figura, que os recursos estimados são sempre menores que os recursos reais dentro da fronteira convexa e que esta diferença é menor para os teores de corte baixos. De maneira semelhante, os teores médios dos dados estimados são sempre menores que aqueles dos dados reais e, ainda, que esta diferença é menor quando os teores de corte são baixos.

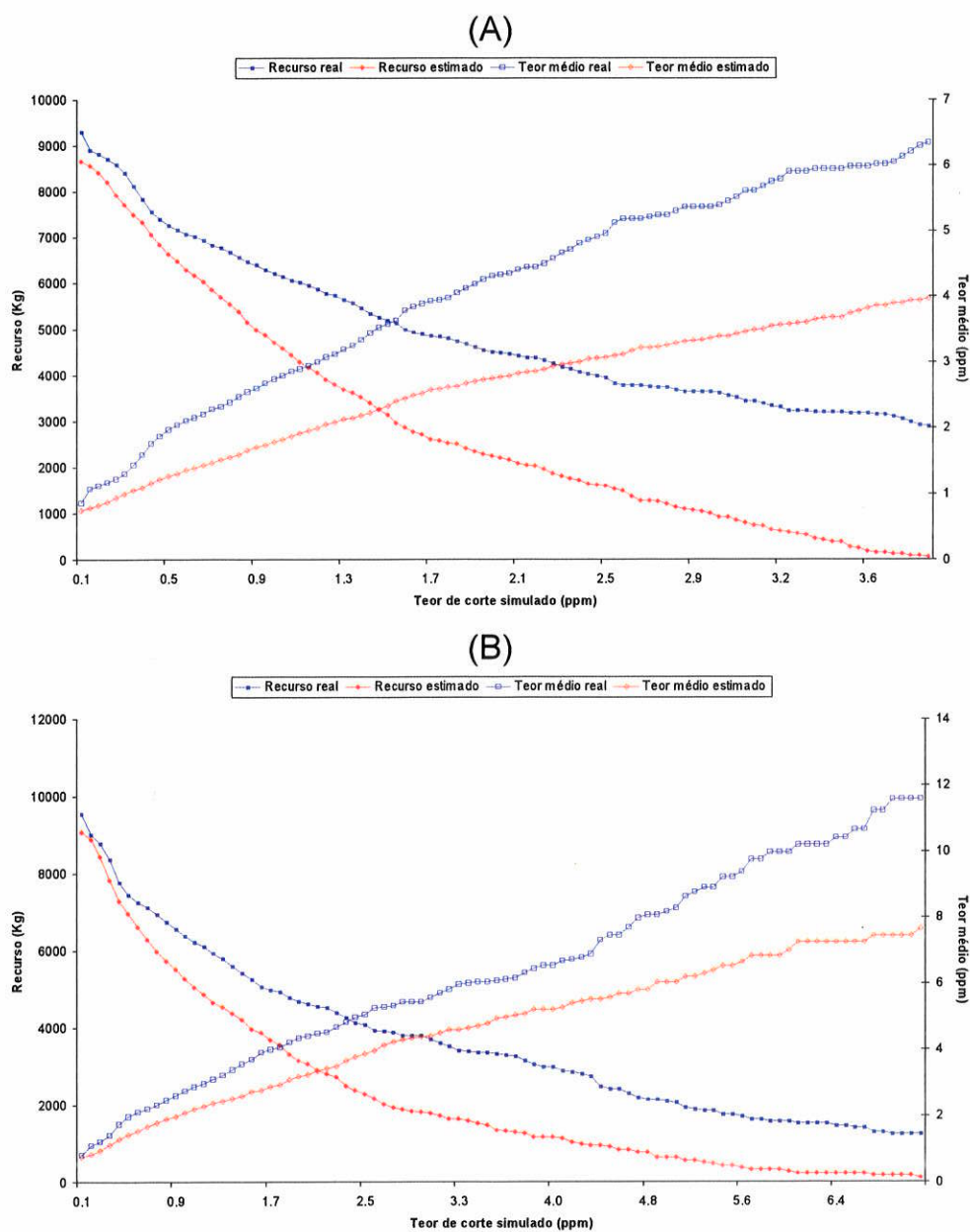


Figura 61: Curva de parametrização dos recursos geológicos calculados para a 6ª realização do nível percentual de 9% (A) e da 6ª realização do nível percentual de 30% (B), para a variável assimétrica positiva, coletada por amostragem aleatória simples.

O comportamento das curvas *teor de corte x tonelage* e *teor de corte x teor médio*, da sexta realização do nível percentual de 30% (Figura 61-B), são muito semelhantes àqueles das curvas da Figura 61-A, porém a diferença entre o estimado e real, na Figura 61-B, é menor que a da Figura 61-A. Destaca-se, ainda, que tanto os recursos para os primeiros teores de corte quanto o teor médio máximo são maiores no nível percentual de 30% que aqueles do nível percentual de 9%.

4.5.1.2 Variável Simétrica

Os recursos geológicos calculados para a variável simétrica, por amostragem aleatória simples, encontram-se na Figura 62. Estes recursos tendem a aproximar-se do recurso total da variável, à medida que o número de unidades amostrais aumenta. Cabe ressaltar, ainda, que os recursos dos blocos estimados são sempre próximos aos recursos dos dados de referência dentro das fronteiras convexas.

De modo geral, os recursos estimados para a variável simétrica (Figura 62) têm comportamento mais homogêneo que aquele apresentado pelos recursos estimados para a variável assimétrica positiva.

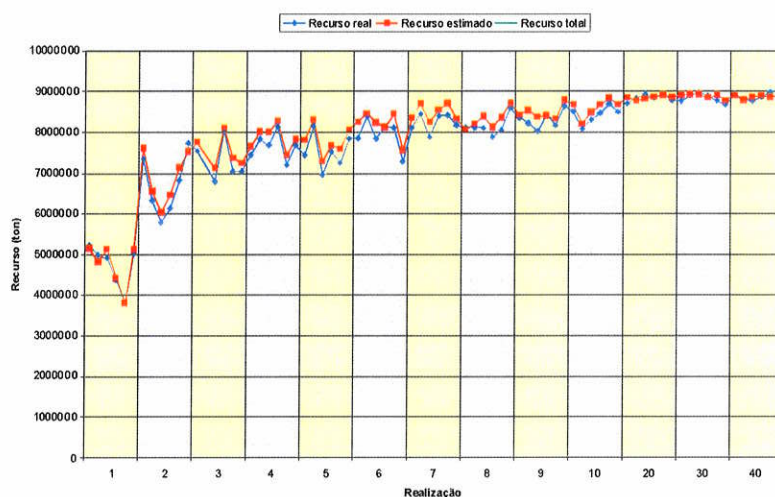
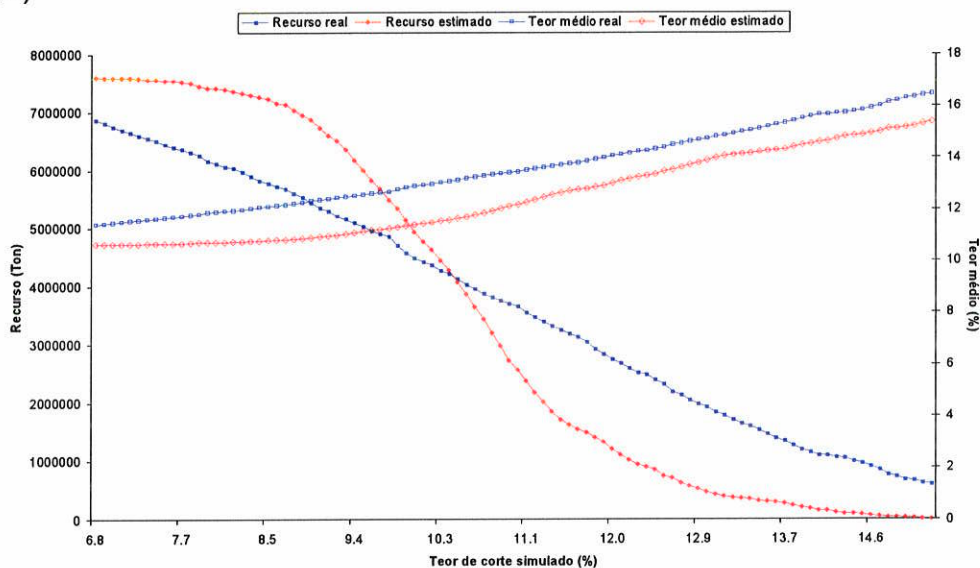


Figura 62: Recursos geológicos calculados para as amostras da variável simétrica (amostragem aleatória simples).

A Figura 63 (A e B) apresenta, respectivamente, as curvas de parametrização dos recursos geológicos da quinta realização do nível percentual de 5% e da sexta realização do nível percentual de 30%.

Nessa figura observa-se que as curvas de recursos estimado x real se interceptam em um valor de teor de corte (10,45% para o nível percentual de 5% e 10,53% para o nível percentual de 30%). Portanto, quando os teores são menores que o teor de intersecção, os recursos são superestimados e quando maiores, os recursos são subestimados, da mesma forma como acontece com a krigagem ordinária, devido ao efeito de suavização (Olea, 1991).

(A)



(B)

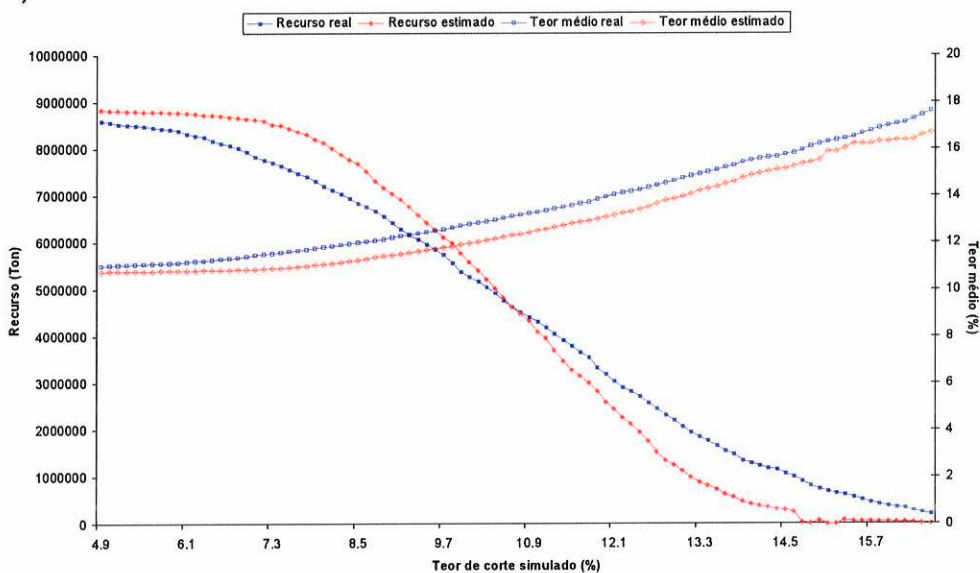


Figura 63: Curva de parametrização dos recursos geológicos calculados para a 5ª realização do nível percentual de 5% (A) e da 6ª realização do nível percentual de 30% (B), amostradas por amostragem aleatória simples.

Os teores médios, dos recursos estimados, são sempre menores que os teores médios dos recursos reais dentro da fronteira convexa, em ambos níveis percentuais. Os teores médios no nível percentual de 5% apresentam amplitude de variação menor que os do nível percentual de 30%. Observa-se que as diferenças entre os teores médios reais e estimados são sempre menores no nível percentual de 30%, quando comparada às do nível de 5%.

4.5.1.3 Variável Assimétrica Negativa

Analisando-se os recursos geológicos, desta variável, observa-se que aumentando o n das amostras, os recursos estimados aproximam-se dos recursos totais dos dados de referência. Observa-se, ainda, que algumas realizações resultaram em recursos maiores do que o recursos total dos dados de referência.

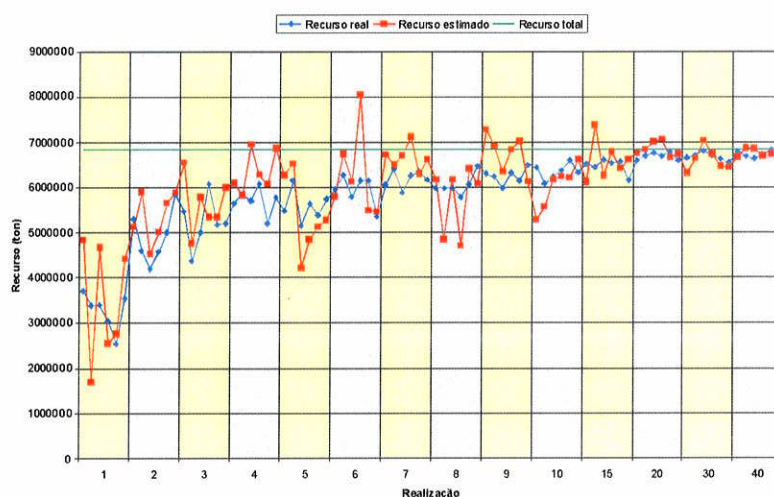


Figura 64: Recursos geológicos calculados para as amostras da variável assimétrica negativa (amostragem aleatória simples).

As curvas de parametrização são apresentadas na Figura 65, para a sexta realização do nível percentual de 2% (Figura 65-A) e para a quinta realização do nível percentual de 30% (Figura 65-B). As curvas *teor de corte x tonelagem*, em ambos níveis percentuais, mostram a mesma tendência geral, observada para a variável simétrica.

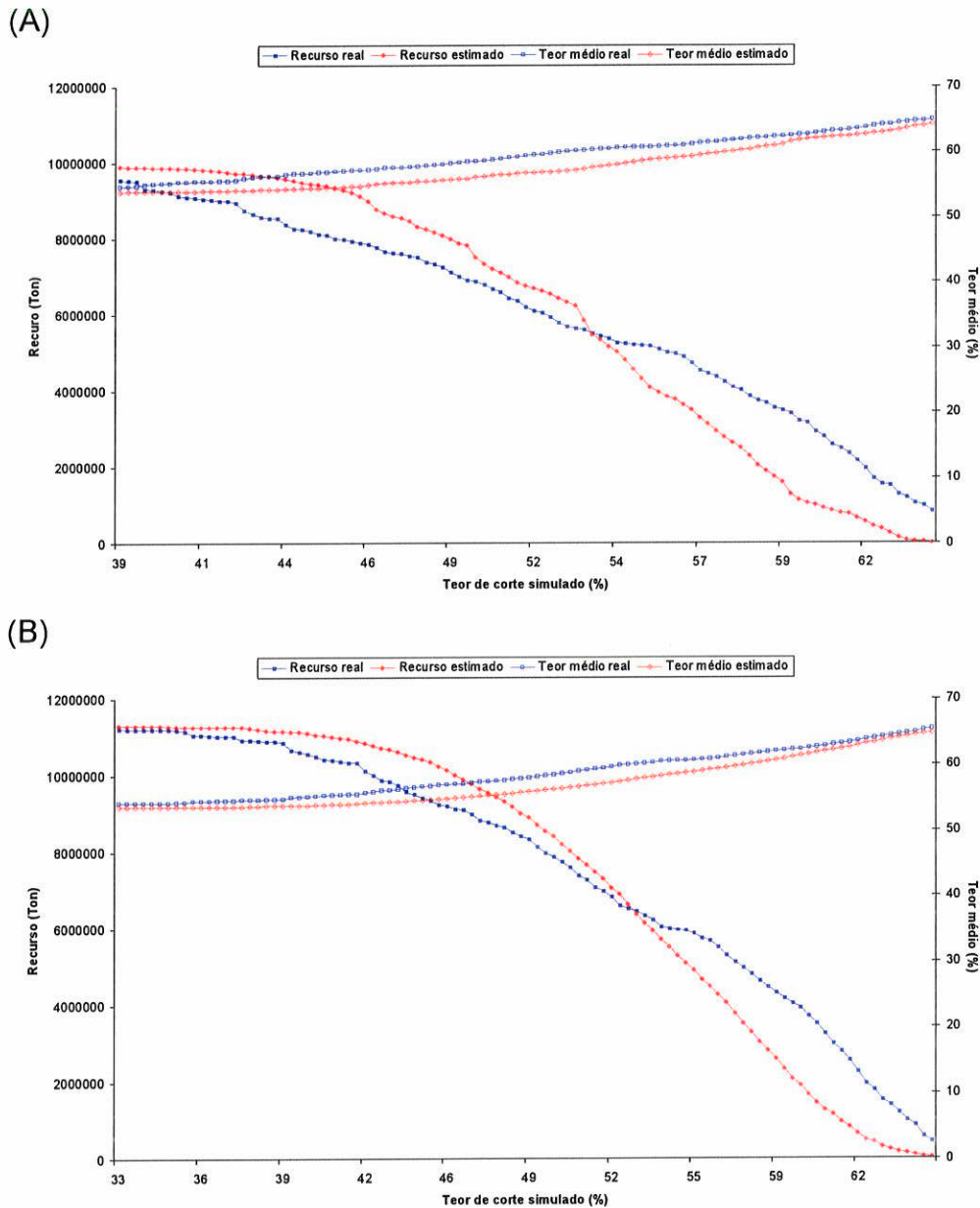


Figura 65: Curva de parametrização dos recursos geológicos calculados da variável assimétrica negativa, coletada por amostragem aleatória simples para a 6ª realização do nível percentual de 2% (A) e da 5ª realização do nível percentual de 30% (B).

4.5.2 Amostragem Aleatória Estratificada

4.5.2.1 Variável Assimétrica Positiva

A Figura 66 mostra a progressão dos valores dos recursos geológicos, da variável assimétrica positiva, à medida que os níveis percentuais são aumentados. Para esta variável, os recursos tendem a não estabilizar em torno de um valor dentro

de um mesmo nível percentual e oscilam até o nível percentual de 30%. Esta oscilação é maior nos níveis percentuais mais baixos.

Os recursos dos dados estimados, mesmo para níveis percentuais elevados, por vezes ultrapassam o recurso total da referência.

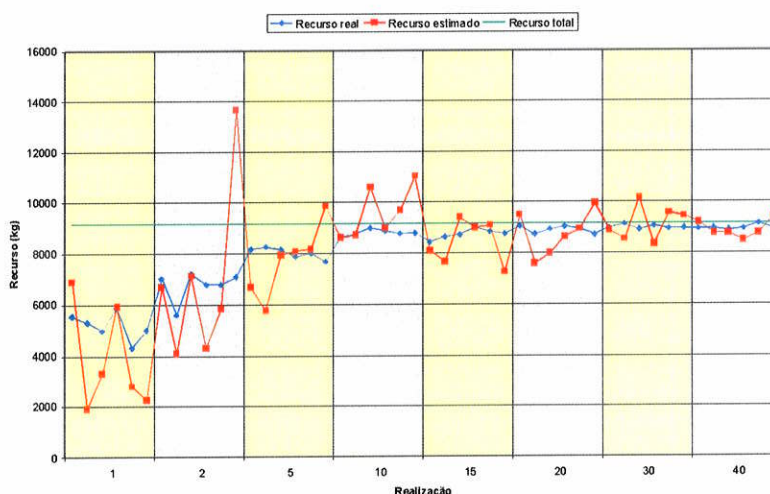


Figura 66: Recursos geológicos calculados para as amostras da variável assimétrica positiva (amostragem aleatória estratificada).

As curvas de parametrização, dos recursos geológicos, apresentadas na Figura 67 referem-se à sexta realização do nível percentual de 10% (A) e a sexta realização do nível percentual de 30% (B).

Observa-se que, para o nível percentual de 10%, na curva *teor de corte x tonelagem* os recursos estimados até o teor de corte de 7,9g/ton, são superestimados em relação aos recursos dos dados de referência dentro da fronteira convexa. Acima deste teor de corte, os recursos estão subestimados.

A curva *teor de corte x tonelagem* dos recursos estimados, para o nível percentual de 30%, comparada à dos recursos reais, mostra recursos superestimados até o teor de corte de aproximadamente 1,0 g/ton, a partir do qual estes passam a ser subestimados.

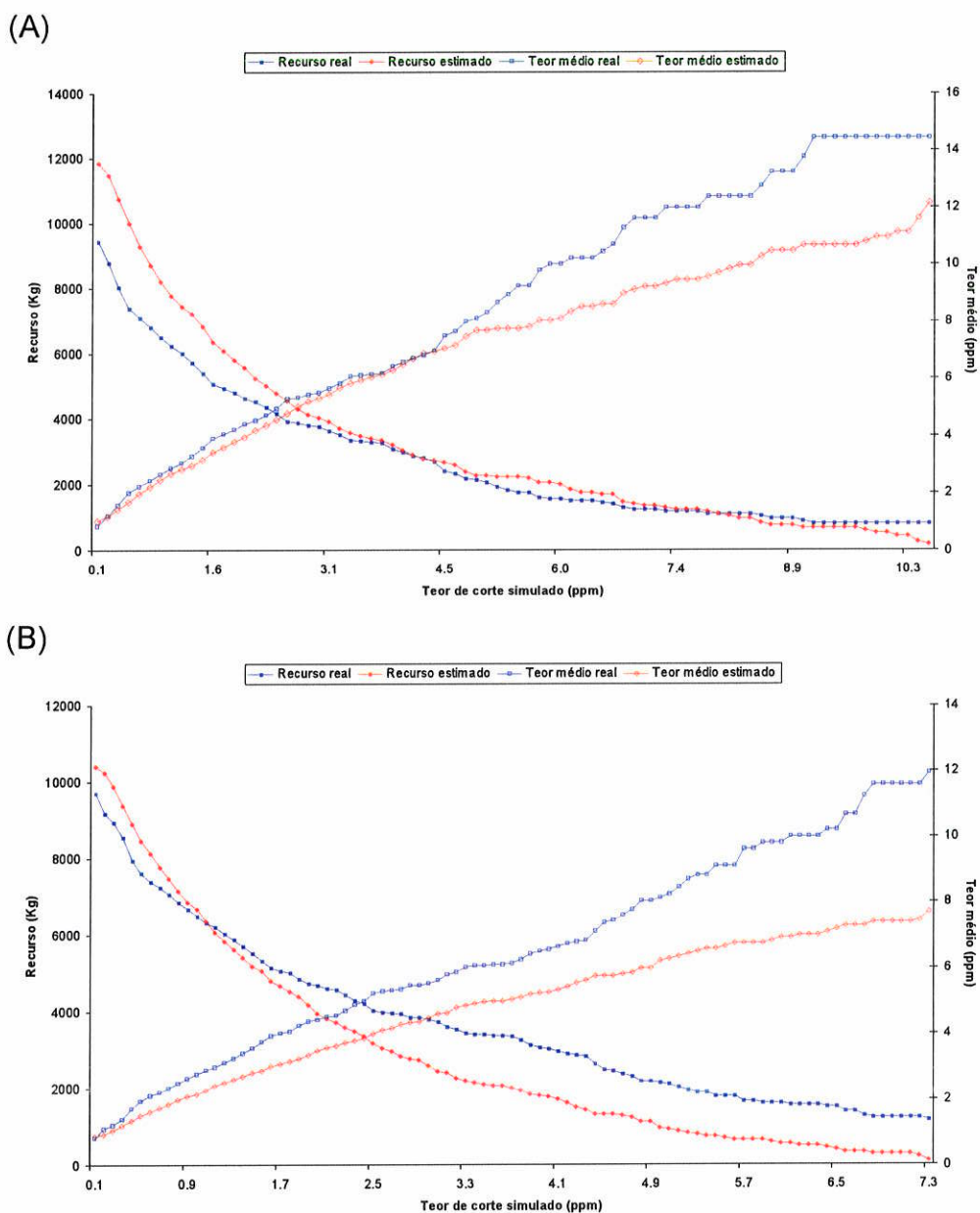


Figura 67: Curva de parametrização dos recursos geológicos calculados para a 6ª realização do nível percentual de 10% (A) e da 6ª realização do nível percentual de 30% (B), para a variável assimétrica positiva, coletada por amostragem aleatória estratificada.

4.5.2.2 Variável Simétrica

A Figura 68 apresenta a variação dos recursos geológicos para a variável simétrica em função dos níveis percentuais. Os recursos geológicos calculados apresentam pouca variação de toneladas. De modo geral, os recursos calculados a

partir das estimativas acompanham aqueles calculados para os dados de referência dentro da fronteira convexa permanecendo, contudo, na maioria dos níveis percentuais acima deste. Para esta variável, em nenhuma amostra, os recursos estimados são superiores ao recurso total dos dados de referência.

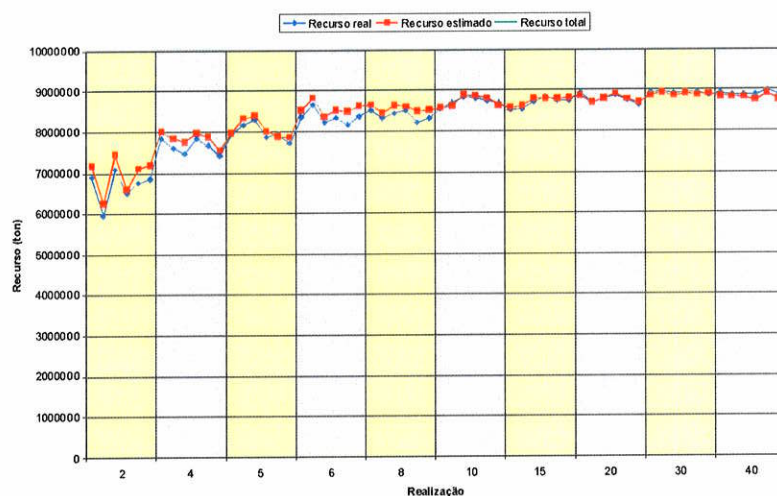


Figura 68: Recursos geológicos calculados para as amostras da variável simétrica (amostragem aleatória estratificada).

Os diagramas da Figura 69 (A e B) apresentam as curvas de parametrização dos recursos geológicos para, respectivamente, a quinta realização do nível percentual de 5% (Figura 69-A) e a sexta realização do nível percentual de 30% (Figura 69-B).

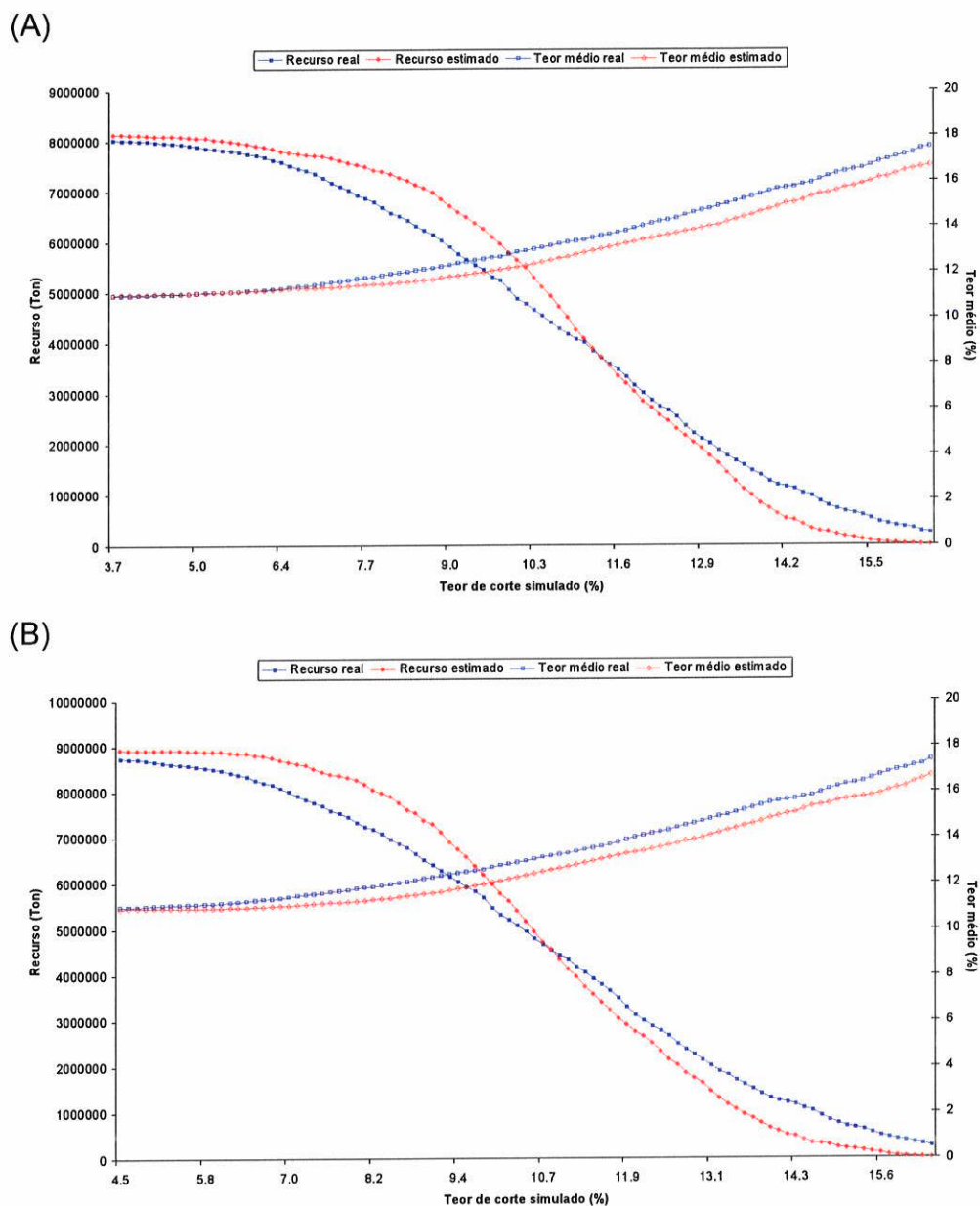


Figura 69: Curva de parametrização dos recursos geológicos calculados para a 5ª realização do nível percentual de 5% (A) e da 6ª realização do nível percentual de 30% (B), para a variável simétrica, retirada por amostragem aleatória estratificada.

Os recursos estimados, no nível percentual de 5%, são maiores que os recursos reais até o teor de corte de 11%, a partir do qual tornam-se menores. O mesmo comportamento é observado para a curva *teor de corte x tonelagem* referente à 6ª realização do nível percentual de 30%.

4.5.2.3 Variável Assimétrica Negativa

Para a variável assimétrica negativa, a dispersão dos recursos (Figura 70) é mais elevada nos níveis percentuais menores e diminui à medida que o tamanho das amostras aumenta. Os recursos geológicos estimados seguem a mesma tendência dos recursos reais dentro da fronteira convexa e, por vezes, extrapolam o valor do recurso total dos dados de referência.

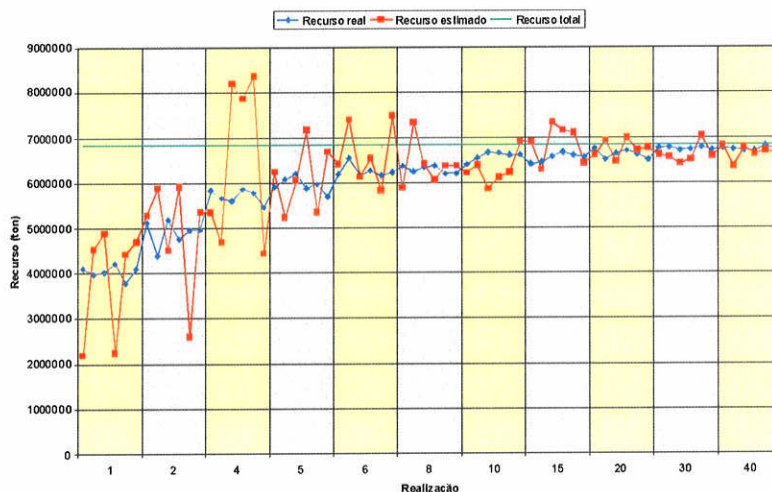


Figura 70: Recursos geológicos calculados para as amostras da variável assimétrica negativa (amostragem aleatória estratificada).

As curvas de parametrização calculadas para os recursos geológicos da sexta realização o nível percentual de 2% são apresentadas na Figura 71-A. Pode-se observar, nesta figura, que os recursos estimados são maiores que os recursos reais dentro da fronteira convexa até o teor de corte de 58%, a partir do qual os recursos estimados passam a ser menores. Na Figura 71-B, encontram-se representadas as curvas de parametrização dos recursos geológicos calculados para a quinta realização do nível percentual de 30%. A curva *teor de corte x tonelagem* dos recursos estimados apontam recursos maiores que aqueles da curva dos recursos reais. Esta superestimativa dos recursos ocorre até o teor de corte de 54%, a partir do qual os recursos são subestimados. Comparando-se as diferenças entre os recursos reais e estimados da Figura 71-A e Figura 71-B, observa-se que estas são maiores para os recursos do nível percentual de 30%.

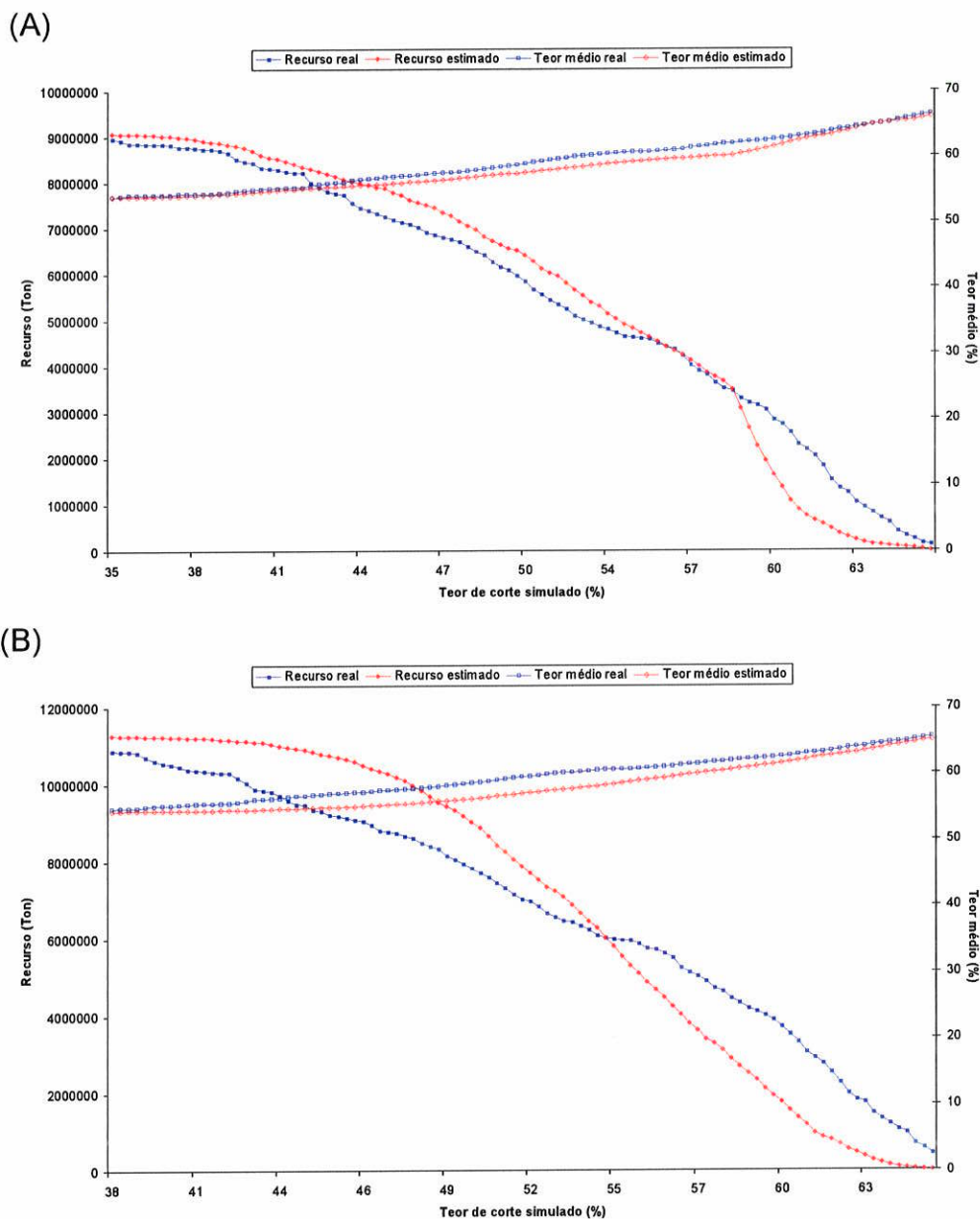


Figura 71: Curva de parametrização dos recursos geológicos calculados para a 6ª realização do nível percentual de 2% (A) e da 5ª realização do nível percentual de 30% (B), para a variável assimétrica negativa, coletada por amostragem aleatória estratificada.

4.5.3 Amostragem Sistemática

4.5.3.1 Variável Assimétrica Positiva

Os recursos calculados, para esta variável (Figura 72), podem ser considerados heterogêneos, pois não acompanham os recursos calculados utilizando os dados de referência dentro da fronteira convexa. A tendência dos recursos serem baixos para as

amostras com n pequeno e gradativamente se elevarem à medida que o n das amostras aumenta não se mantém para a amostragem desta variável.

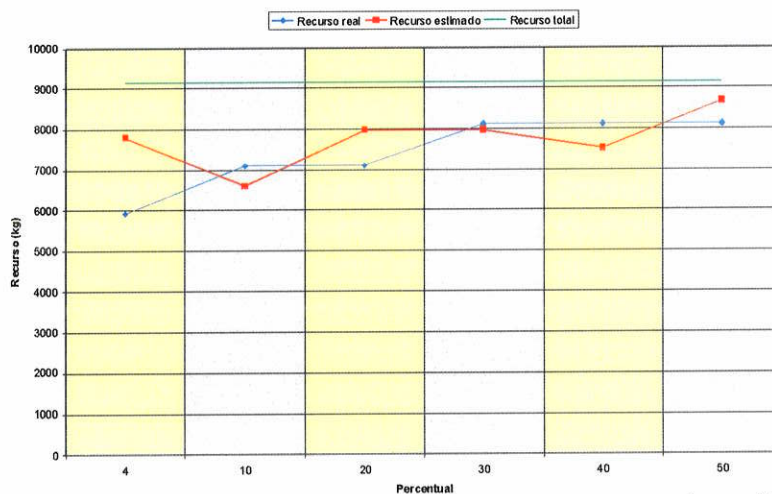


Figura 72: Recursos geológicos calculados para as amostras da variável assimétrica positiva (amostragem sistemática).

As curvas de parametrização dos recursos geológicos, desta variável, são as apresentadas na Figura 73, em (A) tem-se as curvas calculadas para o nível percentual de 10% e em (B) as do nível percentual de 30%.

Na Figura 73-A, os recursos geológicos estimados são sempre menores que os recursos reais dentro da fronteira convexa, a diferença entre recursos reais e estimados aumenta à medida que o teor de corte aumenta.

Comparando-se a curva *teor de corte x tonelagem* dos recursos estimados, do nível percentual de 30%, com a curva dos recursos reais mostra que, acima do sétimo teor de corte aplicado, os recursos estimados são sempre menores que os recursos reais dentro da fronteira convexa. As diferenças entre os recursos reais e estimados são maiores para o nível percentual de 10% do que aqueles observados no nível percentual de 30%.

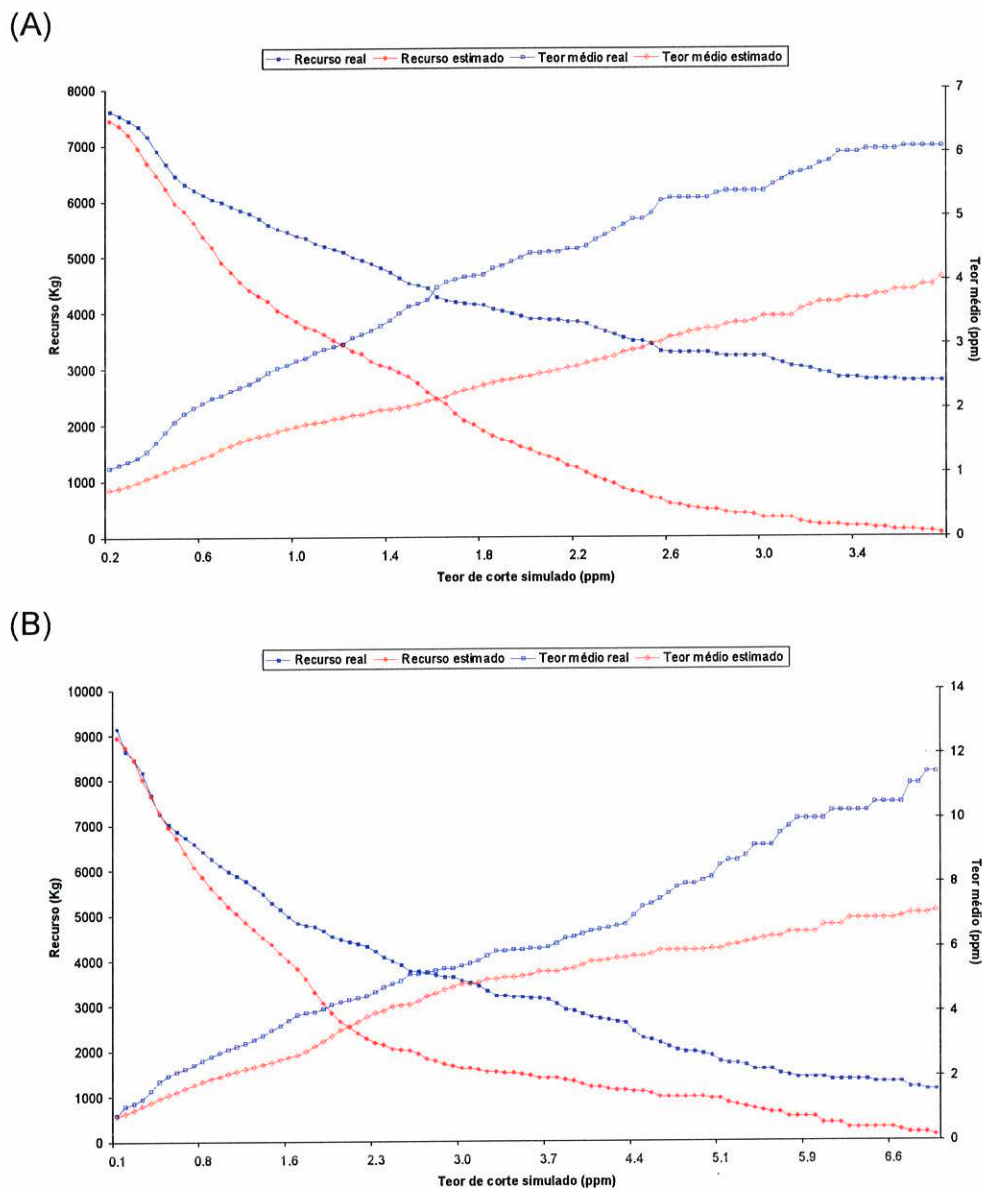


Figura 73: Curva de parametrização dos recursos geológicos calculados para o nível percentual de 10% (A) e do nível percentual de 30% (B), para a variável assimétrica positiva, coletada por amostragem sistemática.

4.5.3.2 Variável Simétrica

Os recursos da variável simétrica (Figura 74) são homogêneos se comparados aos recursos dos dados de referência, dentro da fronteira convexa, uma vez que estes apresentam valores muito semelhantes entre si. Destaca-se que, mesmo para o nível percentual de 50%, os recursos calculados são menores que o recurso total dos dados de referência, porém maiores que os recursos reais dentro da fronteira convexa.

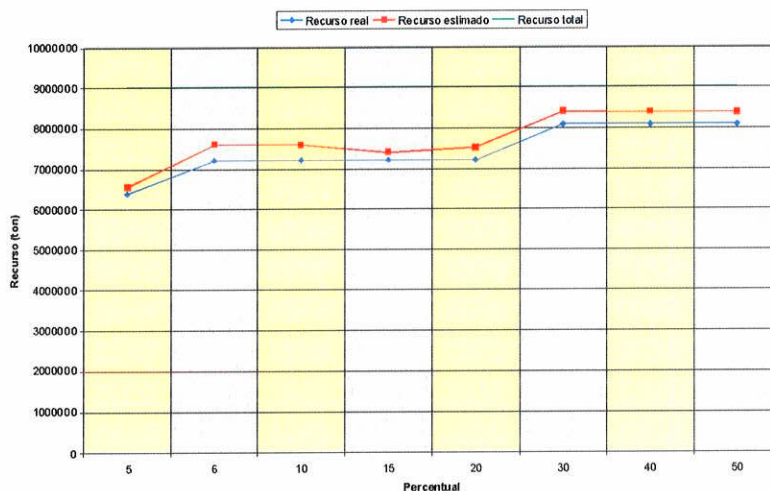


Figura 74: Recursos geológicos calculados para as amostras da variável simétrica (amostragem sistemática).

As curvas de parametrização apresentadas na Figura 75 (A e B) são as curvas calculadas, respectivamente, para os níveis percentuais de 5% e 30%.

Para o nível percentual de 5% (Figura 75-A), os recursos geológicos estimados possuem valores próximos aos valores dos recursos reais, dentro da fronteira convexa. A partir do teor de corte simulado de 10%, os recursos estimados são menores que os recursos reais e as diferenças entre eles aumentam à medida que o teor de corte aumenta. Os recursos geológicos, calculados para o nível percentual de 30%, têm comportamento diferente ao dos calculados para o nível percentual de 5%. Para os teores de corte inferiores a 10,8%, os recursos estimados são maiores que os recursos reais e, a partir deste teor, os recursos estimados são menores.

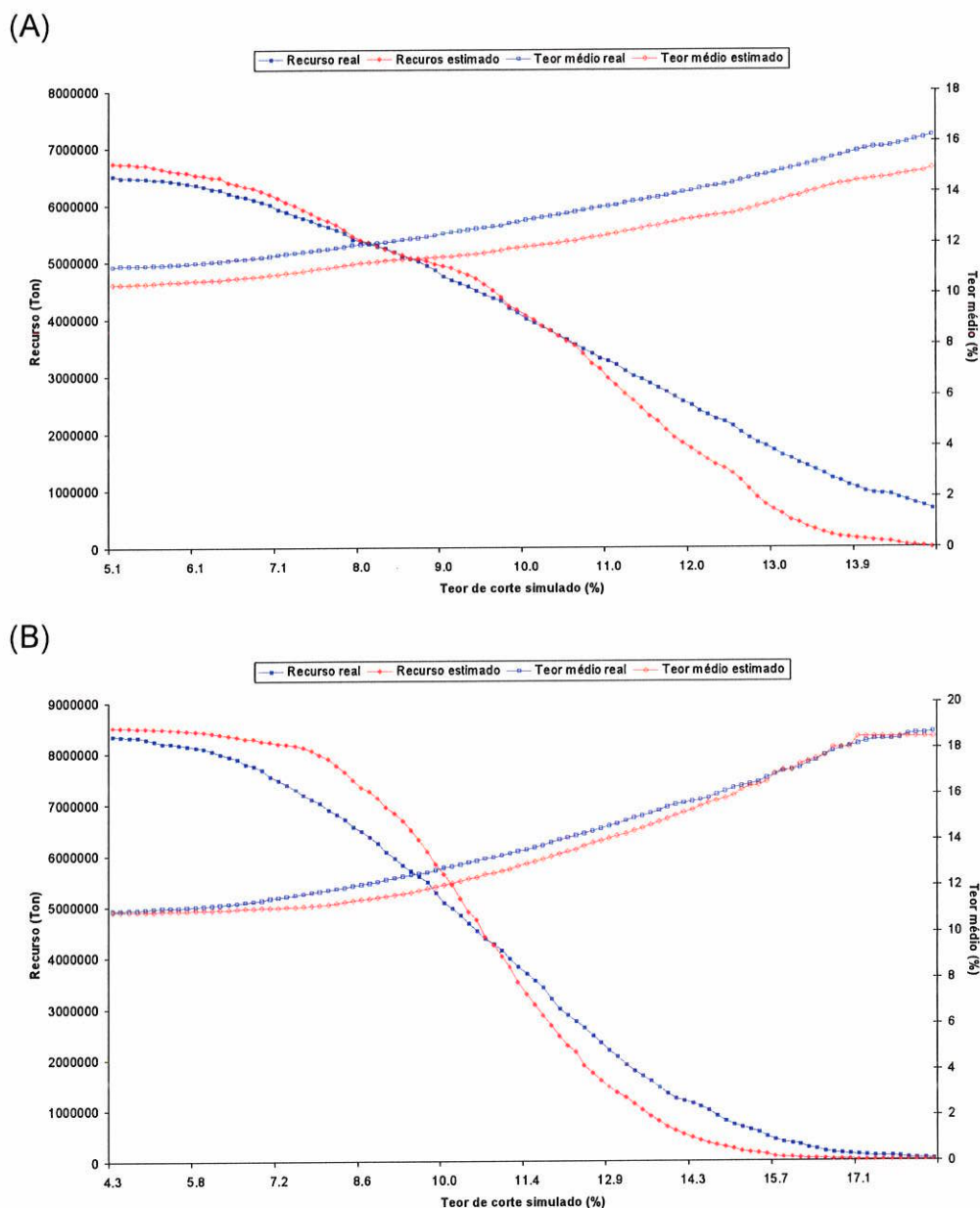


Figura 75: Curva de parametrização dos recursos geológicos calculados para o nível percentual de 5% (A) e do nível percentual de 30% (B) para a variável simétrica, coletada por amostragem sistemática.

4.5.3.3 Variável Assimétrica Negativa

Os recursos estimados, para esta variável (Figura 76), aproximam-se dos recursos reais, dentro da fronteira convexa, nos níveis percentuais superiores a 15%. Os recursos estimados para esta amostragem apresentam comportamento heterogêneo e a dispersão é grande nos níveis percentuais menores.

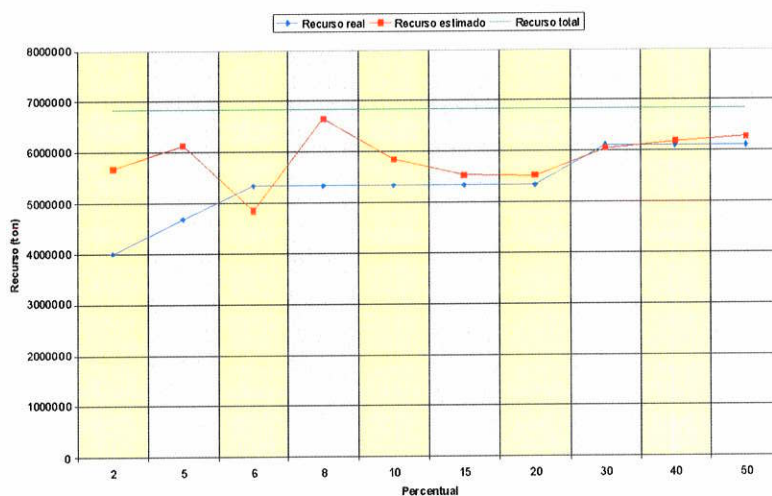


Figura 76: Recursos geológicos calculados para as amostras da variável assimétrica negativa (amostragem sistemática).

As curvas de parametrização dos recursos calculados, para os níveis percentuais de 2% e 30%, são apresentadas, respectivamente na Figura 77-A e Figura 77-B.

Para o nível percentual de 2%, os recursos estimados são maiores do que os recursos reais dentro da fronteira convexa até o teor de corte de 55%, a partir deste valor os recursos estimados são menores que os reais. No nível percentual de 30%, os recursos estimados são maiores que os reais até o teor de corte ser igual a 53%. A partir deste teor de corte, os recursos reais assumem valores maiores que os dos recursos estimados.

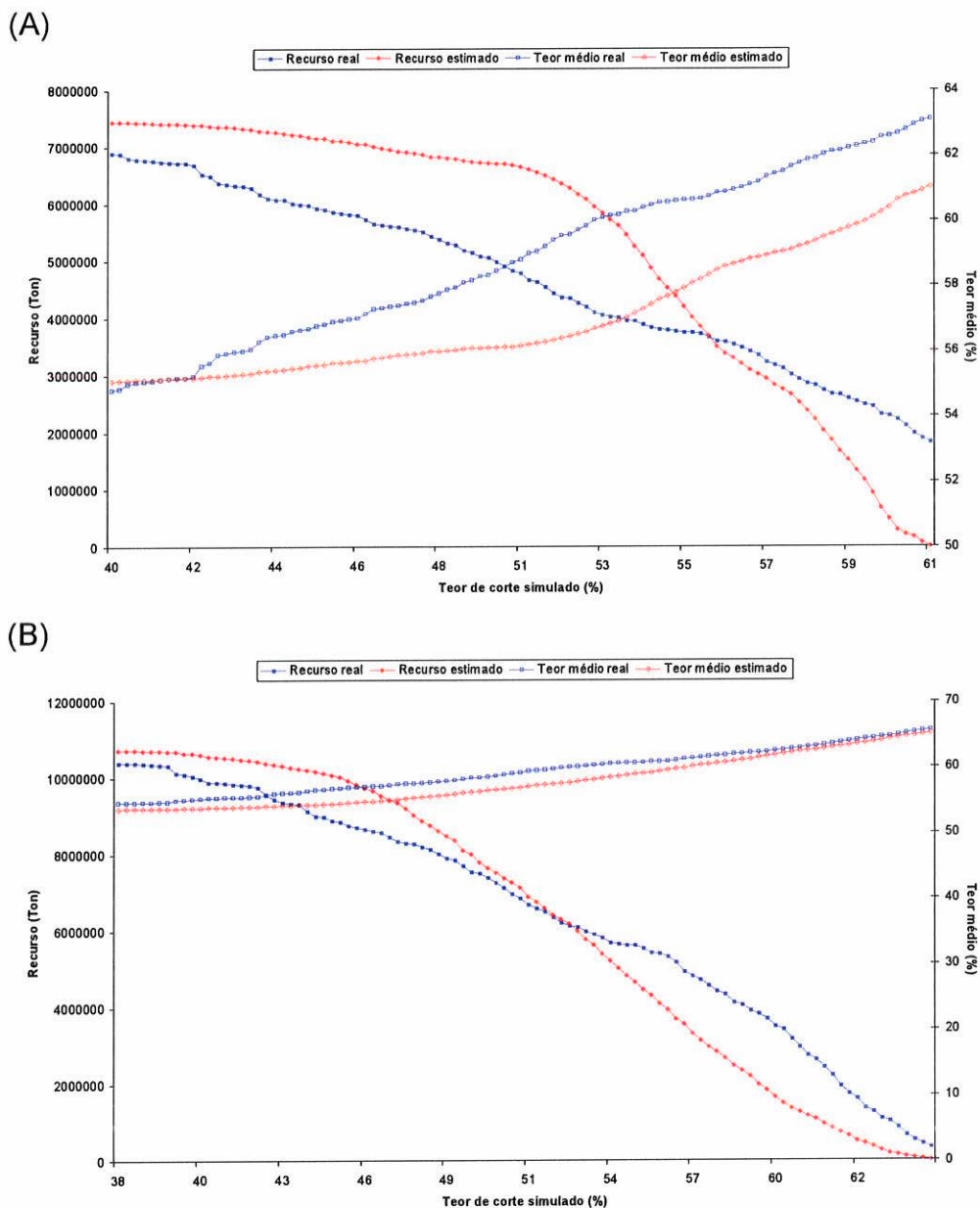


Figura 77: Curva de parametrização dos recursos geológicos calculados para o nível percentual de 2% (A) e do nível percentual de 30% (B), para a variável assimétrica negativa, coletada por amostragem sistemática.

4.6 Análise de Espectral

A análise espectral contou com o cálculo do espectro bruto, por séries duplas de Fourier, dos dados de referência e das estimativas por krigagem ordinária. Como o cálculo do espectro de potência deve ser realizado em malhas regulares, a krigagem realizada utilizou parâmetros de vizinhança (raio de 2000m) e fronteira relaxados em relação àqueles utilizados nas estimativas para o cálculo dos recursos.

Para comparar os espectros de potência das estimativas com o dos dados de referência, foi calculada a razão entre soma dos espectros brutos dos dados estimados pela soma dos espectros dos dados de referência. Esta razão multiplicada por 100, reflete, em termos percentuais, o quanto da energia dos dados de referência estão presentes nos espectros das estimativas.

4.6.1 Amostragem Aleatória Simples

A Figura 78 mostra a dispersão da razão entre espectros em relação ao nível percentual calculado.

Observa-se nesta figura que a variável assimétrica negativa apresenta as razões que oscilam menos e que todos os valores são superiores a 89%, os valores de mínimo e máximo destas razões são, respectivamente, 89,28% e 105,98%. A variável simétrica, com comportamento semelhante à variável assimétrica negativa, possui valores, de máximo e mínimo, iguais a 80,34% e 107,38%, respectivamente. As razões para esta variável têm comportamento mais errático, sobretudo, nos níveis percentuais mais baixos, tendendo à estabilidade a partir do nível percentual de 8% onde todos os valores da razão são superiores a 90%.

A variável assimétrica positiva é a que apresenta razões mais heterogêneas oscilando entre 9,71% e 94,10%. É interessante destacar que as porcentagens mais elevadas encontram-se nos níveis percentuais entre 3 e 8%. As razões mais baixas são encontradas nesta variável cuja razão é 42,18%.

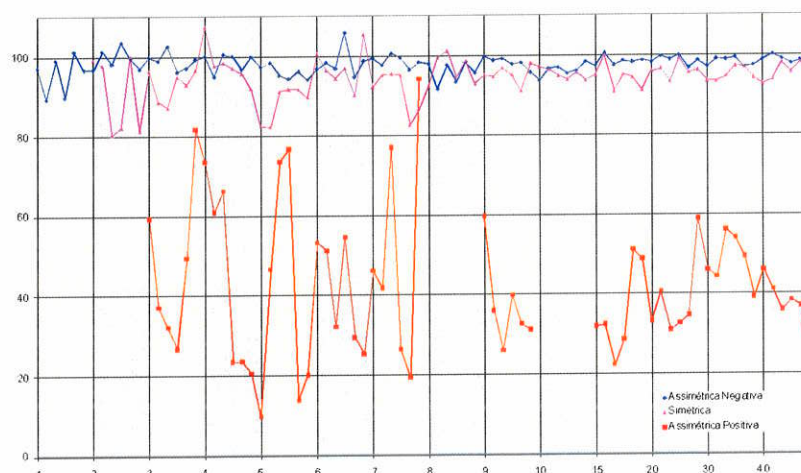


Figura 78: Razão entre o somatório das potências dos dados estimados e a soma das potências dos dados de referência. Amostragem aleatória simples.

4.6.2 Amostragem Aleatória Estratificada

O comportamento das razões para as amostras neste método de amostragem é mais errático que na amostragem aleatória simples, sobretudo nos níveis percentuais inferiores (Figura 79).

A variável assimétrica negativa apresenta máximo e mínimo de, respectivamente, 106,73% e 81,62% e a partir do nível percentual de 20% as razões estabilizam em valores próximos a 100%. A variável simétrica, de modo geral, possui comportamento semelhante àquele da variável assimétrica negativa, com a diferença que os valores destas razões são menores que na variável anterior. O valor de máximo para a variável simétrica é 107,51% e o valor de mínimo é 80,88%. A variável assimétrica positiva apresentou os menores valores de razão, sendo que em nenhum dos níveis percentuais este valor ultrapassou as razões obtidas para as outras duas variáveis. A dispersão destes valores é errática, mesmo em níveis percentuais elevados, como, por exemplo, 30 e 40%. Nota-se, ainda, uma certa tendência à estabilização nos níveis percentuais mais elevados e, ainda, pode-se notar que as razões apresentam alguma tendência à elevação de seus valores à medida que o tamanho das amostras cresce.

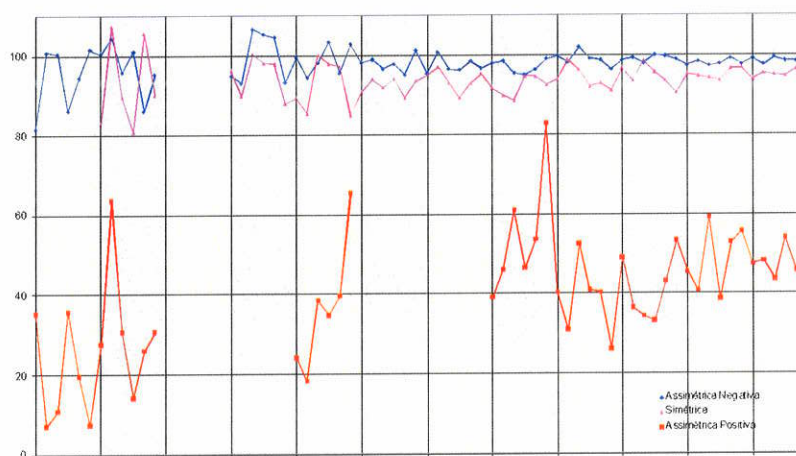


Figura 79: Razão entre a soma das potências dos dados estimados e a soma das potências dos dados de referência. Amostragem aleatória estratificada.

4.6.3 Amostragem Sistemática

O gráfico da Figura 80 mostra o comportamento das razões entre a somatória dos espectros de potência brutos dos números harmônicos das estimativas por krigagem ordinária e a soma dos espectros de potência dos dados de referência.

Pode-se observar, nesta figura, que a variável assimétrica positiva apresenta razões com valores homogêneos e sempre próximos a 100%, variando de um mínimo de 95,05% a um máximo de 102,13%. Para os níveis percentuais superiores a 15%, as razões praticamente não variam, sendo sempre superiores a 97% e apresentando uma amplitude de variação de, apenas, 1,3 pontos percentuais.

Pode-se observar que as razões calculadas para a variável simétrica, possuem comportamento mais errático que aquele obtido na variável assimétrica negativa, porém também não apresentam uma amplitude de variação elevada. O valor de mínimo das razões é 89,08% e o valor de máximo é 97,50%.

A variável assimétrica positiva, apresenta o comportamento mais errático e os valores mais baixos, se comparadas às outras duas variáveis, seu valor de mínimo é 29,33% e o de máximo é 66,63%. Mesmo para níveis percentuais elevados, percebe-se que não existe tendência à estabilização, uma vez que para o nível percentual de 50%, em tese o mais próximo dos dados de referência, a razão é inferior a 70%.

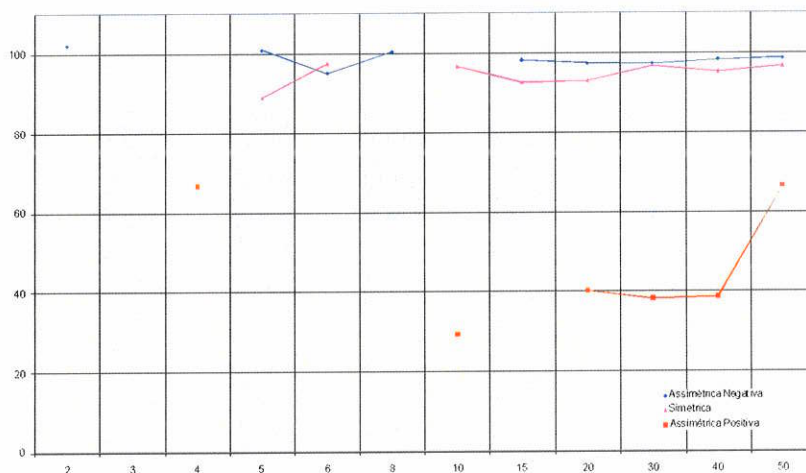


Figura 80: Razão entre o somatório dos espectros de potência de cada harmônico dos dados estimados e a soma das potências dos dados de referência para a amostragem sistemática.

Cabe ressaltar que, nas figuras apresentadas nos sub-itens deste item, os níveis percentuais que não apresentam valores, estão em branco, são aqueles cujos variogramas experimentais não são estruturados e, portanto, não tiveram a estimativa calculada.

Capítulo 5

Discussão dos Resultados

Neste capítulo serão interpretados e discutidos os resultados apresentados no Capítulo 4. Adotar-se-á a mesma seqüência utilizada para a apresentação dos resultados.

5.1 Recobrimento

As variações no tamanho dos subdomínios podem ser interpretadas como sinalizadores da falta de representatividade da amostragem, sobretudo em um depósito de alta variabilidade. Para as três amostragens aplicadas neste estudo, verificou-se que à medida que o tamanho das amostras aumenta, a área de recobrimento, aumenta.

Comparativamente, a amostragem aleatória estratificada é a que apresenta melhor recobrimento em área do domínio, pois com apenas 120 amostras em média, 93% da área foi recoberta. Este percentual de recobrimento, para a amostragem aleatória simples, apenas é atingido quando são coletadas 150 amostras.

Na amostragem sistemática existem dois problemas, o primeiro é que diferentes níveis percentuais apresentam uma mesma área de recobrimento e o segundo é que mesmo para 750 amostras o recobrimento é menor que 90% da área total, observa-se que o recobrimento da amostragem aleatória estratificada para 90 amostras já é, em média, mais elevado que os 90% das 750 amostras.

5.2 Análise Estatística

Dado um conjunto de n amostras, a média desse conjunto é o valor que minimiza a variância (Davis, 2002). Como no caso em estudo, retiram-se n amostras aleatórias de uma população de tamanho N , com $n < N$, então, a variância amostral poderá ser menor ou maior que a variância populacional, dependendo das unidades amostrais selecionadas e, sobretudo, do tipo de distribuição em análise.

5.2.1 Variável Assimétrica Positiva

Para a variável assimétrica positiva, a relação entre a média e a variância amostrais pode ser descrita conforme observação da Figura 81. Se a média amostral for maior que a média populacional, isto indica que uma maior quantidade de valores da cauda direita da distribuição foi amostrada, o que implica em maior valor da variância amostral em relação à populacional. Por outro lado, se a média amostral for menor que a média populacional, a variância amostral tende a ser menor que a variância populacional, pois os valores elevados não foram totalmente amostrados.

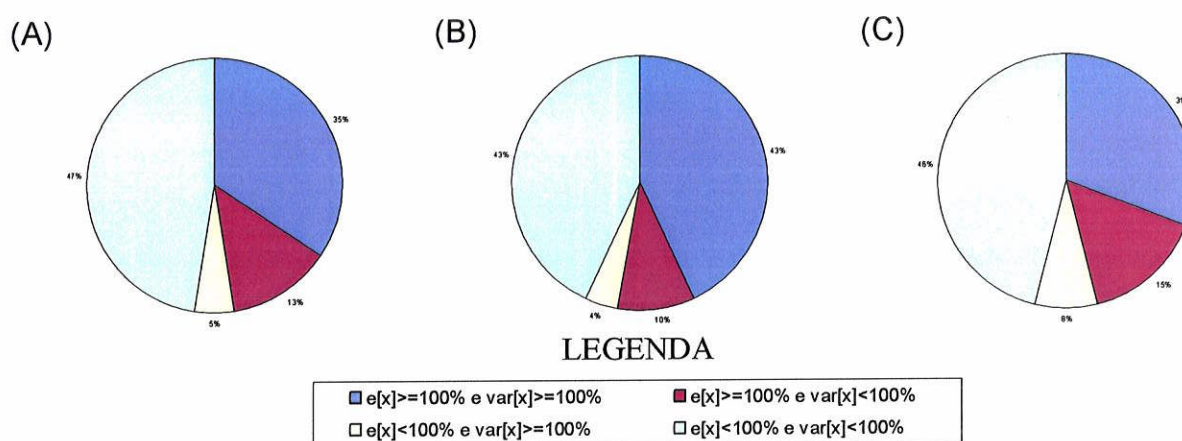


Figura 81: Diagramas em pizza, mostrando as tendências das médias e variâncias, da variável assimétrica positiva, para as três amostragens. (A) Amostragem aleatória simples; (B) Amostragem aleatória estratificada e; (C) Amostragem sistemática.

Tomando por base o apresentado na Figura 81, pode-se afirmar que a amostragem aleatória estratificada gerou amostras mais condizentes com o que era esperado para a variável assimétrica positiva. Há um equilíbrio entre amostras com média amostral menor e com média amostral maior e variâncias amostrais menores e maiores, respectivamente. Este equilíbrio não é observado nas amostras retiradas por amostragem aleatória simples, onde há o predomínio de amostras com médias e variâncias menores que as respectivas populacionais.

5.2.2 Variável Simétrica

Para a variável simétrica, pela forma da distribuição, não há tendência clara, entre as relações de aumento ou diminuição dos valores da média amostral com o aumento ou diminuição da variância amostral (Figura 82).

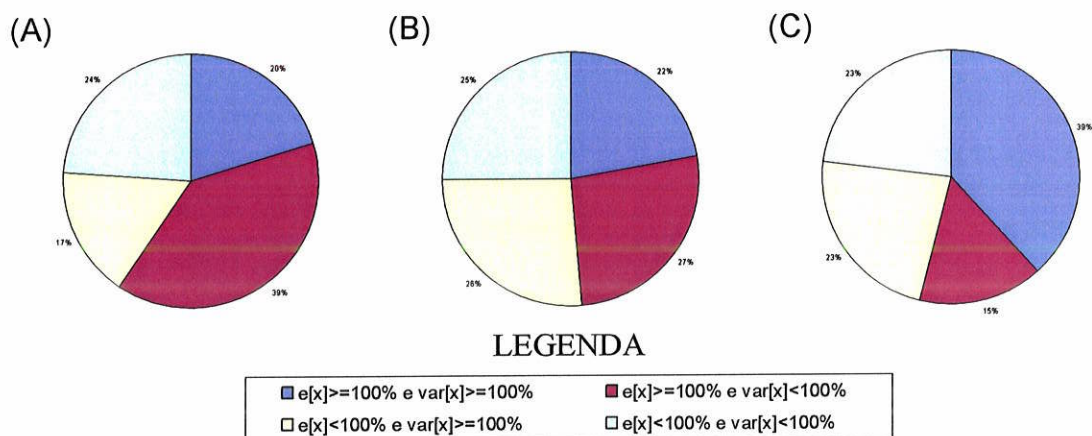


Figura 82: Diagramas em pizza, mostrando as tendências das médias e variâncias, da variável simétrica, para as três amostragens. (A) Amostragem aleatória simples; (B) Amostragem aleatória estratificada e; (C) Amostragem sistemática.

Pela análise da Figura 82, nota-se que novamente a amostragem aleatória estratificada pode ser considerado o método de amostragem mais adequado, pois proporciona uma amostragem equânime dos diversos valores. Na amostragem aleatória simples, há o predomínio da média ser maior que a populacional e a variância menor que a variância populacional.

5.2.3 Variável Assimétrica Negativa

Em uma distribuição assimétrica negativa, a média amostral menor que a média populacional, indica a presença de mais valores baixos na amostra (valores na cauda esquerda), porém como a frequência de valores elevados é alta, a variância amostral se elevará. Por outro lado, se a média amostral for maior que a populacional indica a ausência dos valores responsáveis pela assimetria, logo a variância deverá ser menor que a populacional (Figura 83).

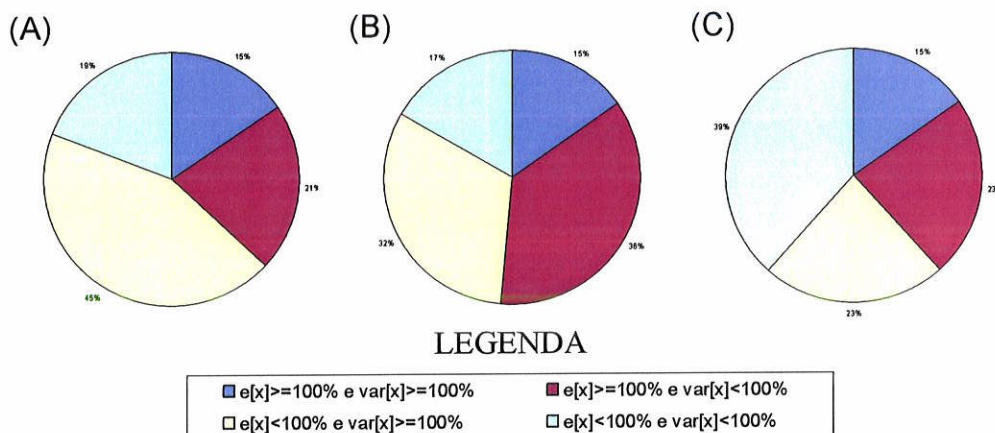


Figura 83: Diagramas em pizza, mostrando as tendências das médias e variâncias, da variável assimétrica negativa, para as três amostragens. (A) Amostragem aleatória simples; (B) Amostragem aleatória estratificada e; (C) Amostragem sistemática.

Neste estudo, a amostragem aleatória estratificada (Figura 83-B) apresenta equilíbrio entre médias elevadas com variâncias baixas e médias baixas com variâncias elevadas, este equilíbrio é o esperado para esta distribuição; confirmando que esta proporciona resultados mais representativos. Já para a amostragem aleatória simples, existe a tendência maior (45%) de amostras com média amostral mais baixa que à populacional e variância mais elevada, indicando enviesamento da amostragem.

5.3 Análise Geoestatística

Inicialmente a convergência dos variogramas experimentais será analisada quanto à assimetria das variáveis em estudo, em cada amostragem testada. Outra análise de convergência será a comparação, de uma mesma variável, entre as três amostragens utilizadas.

Analisando-se os variogramas experimentais médios, derivados de amostras coletadas por amostragem aleatória simples (Figuras 33, 34 e 35), verifica-se que as variáveis simétrica e assimétrica negativa apresentam melhor convergência ao modelo de variograma de referência. Portanto, a distribuição estatística é, também, fator que influencia o cálculo do variograma experimental, pois dados com um mesmo arranjo espacial, mas com diferentes distribuições estatísticas, convergem aos modelos de referência de formas distintas.

Para a amostragem aleatória estratificada, os variogramas experimentais apresentam a mesma tendência verificada na amostragem aleatória simples. Observa-se que para esta amostragem, a variável simétrica apresenta melhor convergência ao modelo de variograma de referência, melhor, inclusive, do que a apresentada pela variável assimétrica negativa que possui menor coeficiente de variação.

Os resultados da amostragem sistemática são limitados, devido à natureza da amostragem, mas, de modo geral, apresentam a mesma tendência verificada anteriormente.

Os variogramas experimentais da variável assimétrica positiva, devido à forma da distribuição, não convergem rapidamente para o modelo de variograma dos dados de referência. Isto pode ser verificado na Figura 33, onde somente os níveis percentuais superiores a 40% aproximam-se do modelo de referência; na Figura 36 onde os níveis percentuais de 20% e 40% convergem e; os Figura 39 os níveis percentuais de 40% e 50%.

Por outro lado, para a variável simétrica, os variogramas experimentais apresentam convergência muito mais rápida, ou seja, em níveis percentuais menores que a variável assimétrica positiva. Esta tendência pode ser observada na Figura 34, onde a convergência não é atingida apenas em níveis percentuais inferiores a 8%; na Figura 37, a convergência já é atingida nos níveis percentuais maiores que 8%; e na Figura 40, a convergência ocorre nos níveis superiores a 10%.

Para a variável assimétrica negativa, os variogramas experimentais, convergem ao modelo de variograma dos dados de referência de modo semelhante à da variável simétrica, diferenciando-se, apenas, por apresentar dispersão em torno do modelo, levemente maior do que a variável simétrica.

O arranjo geométrico proporcionado pela amostragem aleatória estratificada resulta em variogramas experimentais muito mais próximos do modelo de referência. Assim, para um mesmo nível percentual, o arranjo da amostragem aleatória estratificada apresenta melhor distribuição espacial para cálculos de variogramas experimentais que a amostragem aleatória simples. Pela análise da Figura 84-A, verifica-se que para amostras com n pequeno (níveis percentuais menores que 6%), a amostragem aleatória simples apresenta distâncias médias menores, porém a média dos desvios padrão é sempre maior para esta amostragem (Figura 84-B). Nos níveis percentuais maiores do que 6%, as distâncias médias, para as duas amostragens em

questão, são semelhantes, porém os desvios padrão médios são menores na amostragem aleatória estratificada. O aumento nos valores de desvio padrão, na amostragem aleatória simples, reflete o agrupamento de pontos que, com freqüência, é gerado nesta amostragem. Este tipo de arranjo espacial, de pontos amostrais, faz com que as distâncias médias sejam menores, porém com desvios maiores. Por outro lado, a amostragem aleatória estratificada, dada a divisão do domínio em estratos, antes da retirada aleatória de unidades amostrais, garante o melhor espalhamento das unidades amostrais no momento da coleta de dados e, conseqüentemente, melhor distribuição espacial.

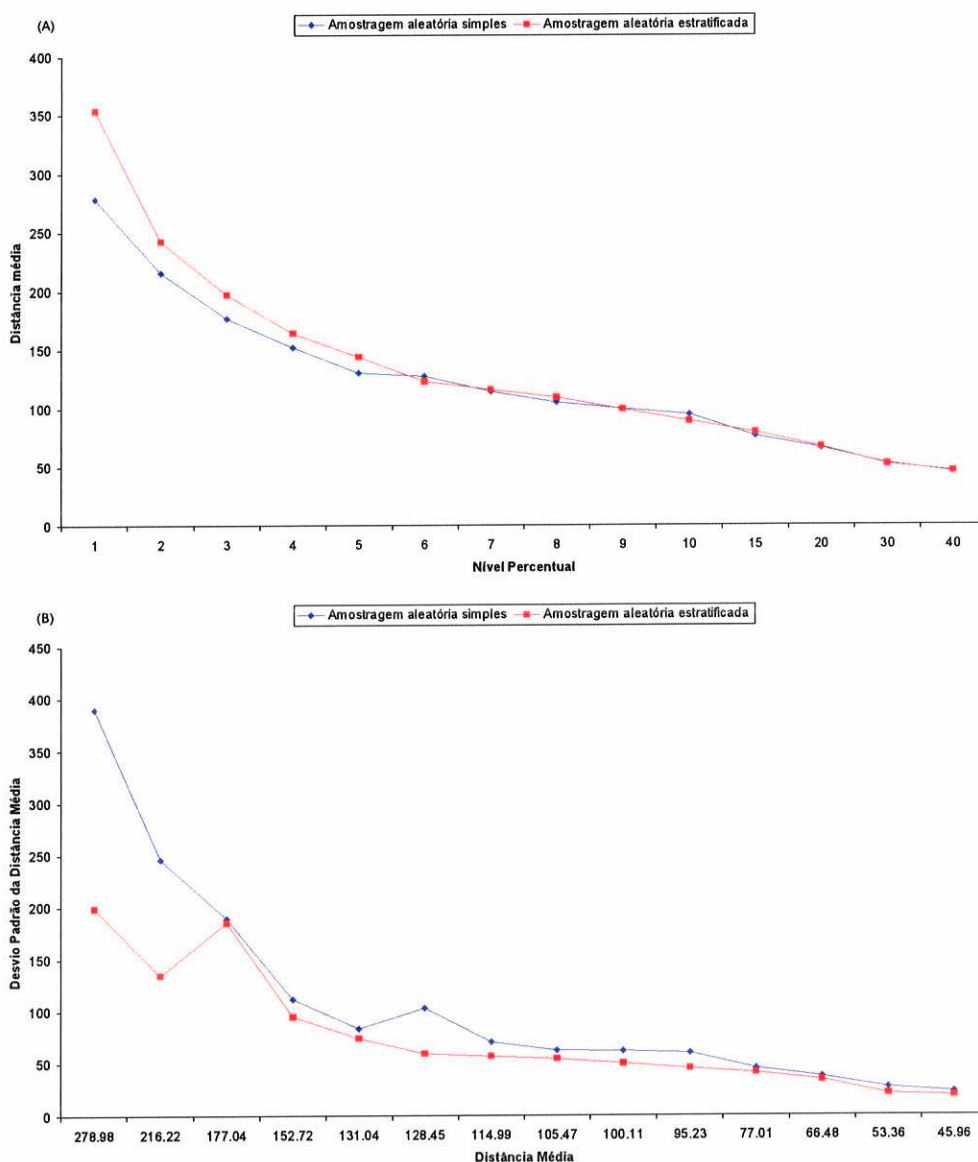


Figura 84: Distância média de separação entre pares de pontos em função do nível percentual (A) e desvios padrão médios em função das distâncias (B).

Outro resultado interessante refere-se aos efeitos pepita, obtidos a partir dos variogramas experimentais, das três variáveis. Pode-se observar que, para a variável assimétrica positiva, mesmo em níveis percentuais elevados e em todas as amostragens, o efeito pepita se dispersa em torno do efeito pepita do modelo de referência. Na variável simétrica, para níveis percentuais elevados, o efeito pepita converge rapidamente ao efeito pepita do modelo de referência. Por fim, para a variável assimétrica negativa, a dispersão dos efeitos pepita tem comportamento semelhante ao da variável simétrica, porém com uma dispersão sutilmente maior.

Estas diferenças de comportamento estão relacionadas à forma da distribuição, ou melhor, ao grau de assimetria que a distribuição apresenta e, também, ao seu sinal. Em variáveis assimétricas positivas, a cauda direita da curva é mais pronunciada e os valores anômalos tendem a ser muito discrepantes em relação à média, elevando a variância desta variável, pois a distribuição é aberta à direita; isto significa que podem ocorrer valores elevados sem uma limitação. A variável simétrica é bi-caudal e a tendência é a ausência de valores discrepantes, logo a variância será baixa. A variável assimétrica negativa é uni-caudal e, embora seja assimétrica, a variância desta variável é baixa, menor do que a variância da variável simétrica, pois a distribuição é fechada a direita, ou seja, não ocorre um valor maior que o máximo possível dessa variável. Portanto, observa-se que, o efeito pepita, além de ser suscetível ao nível de variância, que a distribuição apresenta, depende da simetria da variável para a convergência ao modelo. A observação dos variogramas experimentais, para a variável assimétrica negativa, mostra que, mesmo apresentando menor coeficiente de variação, há um maior espalhamento do efeito pepita.

5.3.1 Amostragem Aleatória Estratificada x Amostragem Sistemática

Ainda, na análise geoestatística dos dados, realizou-se um pequeno exercício para tentar entender melhor o desempenho da amostragem aleatória estratificada sobre as demais. Para tanto, se utilizou um modelo de variograma calculado para uma malha regular de 20 nós em X e 29 nós em Y, totalizando 580 amostras, conforme os variogramas da Figura 85. Esta malha simula uma amostragem sistemática. Sobre ela

far-se-á amostragens, aleatória estratificada, até a reprodução do variograma, conforme os resultados apresentados na Figura 86.

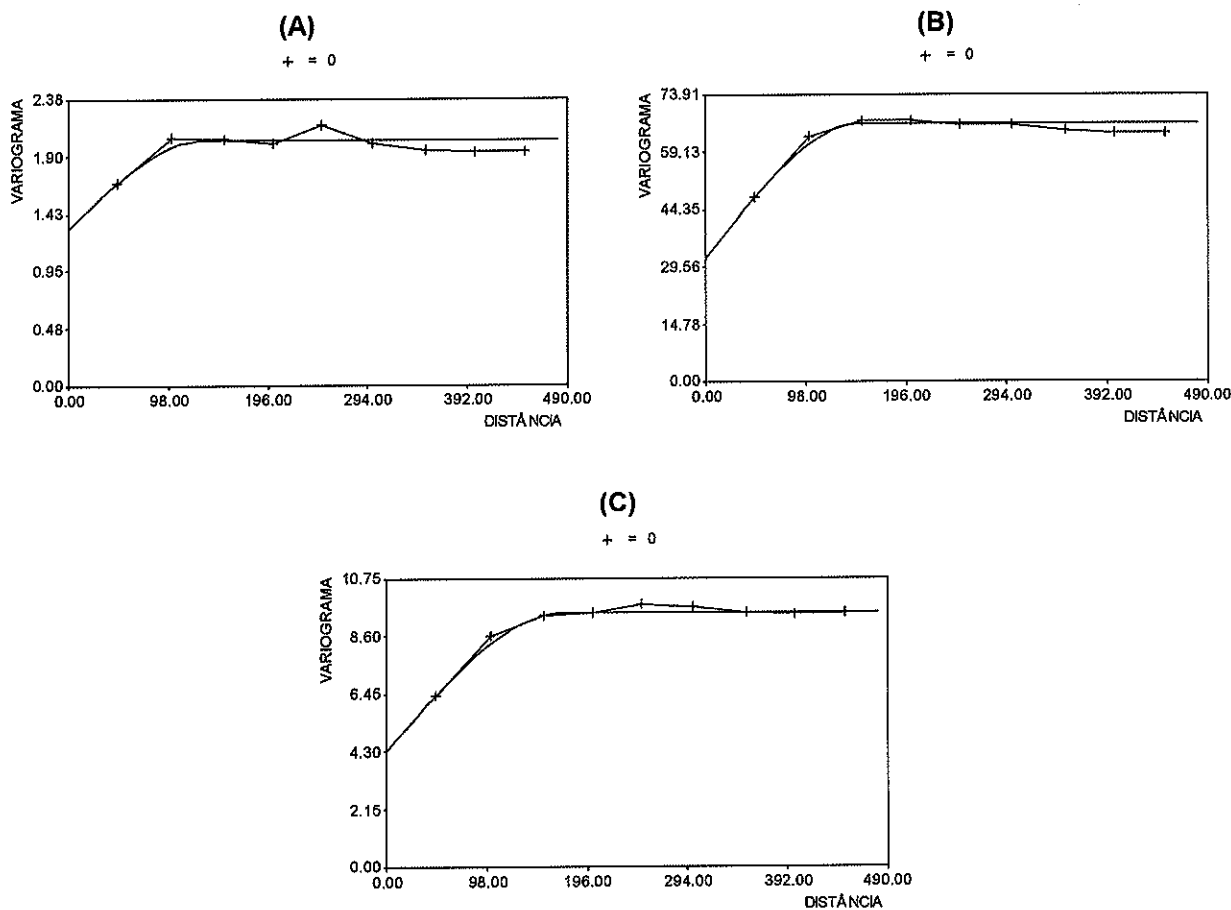


Figura 85: Variogramas experimentais e seus respectivos modelos calculados para uma malha regular de 20 x 29 nós, simulando uma amostragem sistemática. (A) é o variograma da variável assimétrica positiva. (B) é o variograma da variável assimétrica negativa e (C) é o variograma para a variável simétrica.

Assim, conforme os resultados da Figura 86, verifica-se que:

- Para a variável assimétrica positiva o modelo foi reproduzido quando 95% das unidades amostrais foram coletadas (Figura 86-A), ou seja, houve uma economia de 30 unidades amostrais;
- Para a variável assimétrica negativa o modelo foi reproduzido quando 76% do total dos pontos foi coletado. O modelo foi reproduzido, porém com um incremento de 5% do efeito pepita (Figura 86-B), e uma economia de 144 unidades amostrais;

- Para a variável simétrica o modelo foi reproduzido quando foram retiradas apenas 66% de unidades amostrais (Figura 86-C), resultando em uma economia de 204 unidades amostrais.

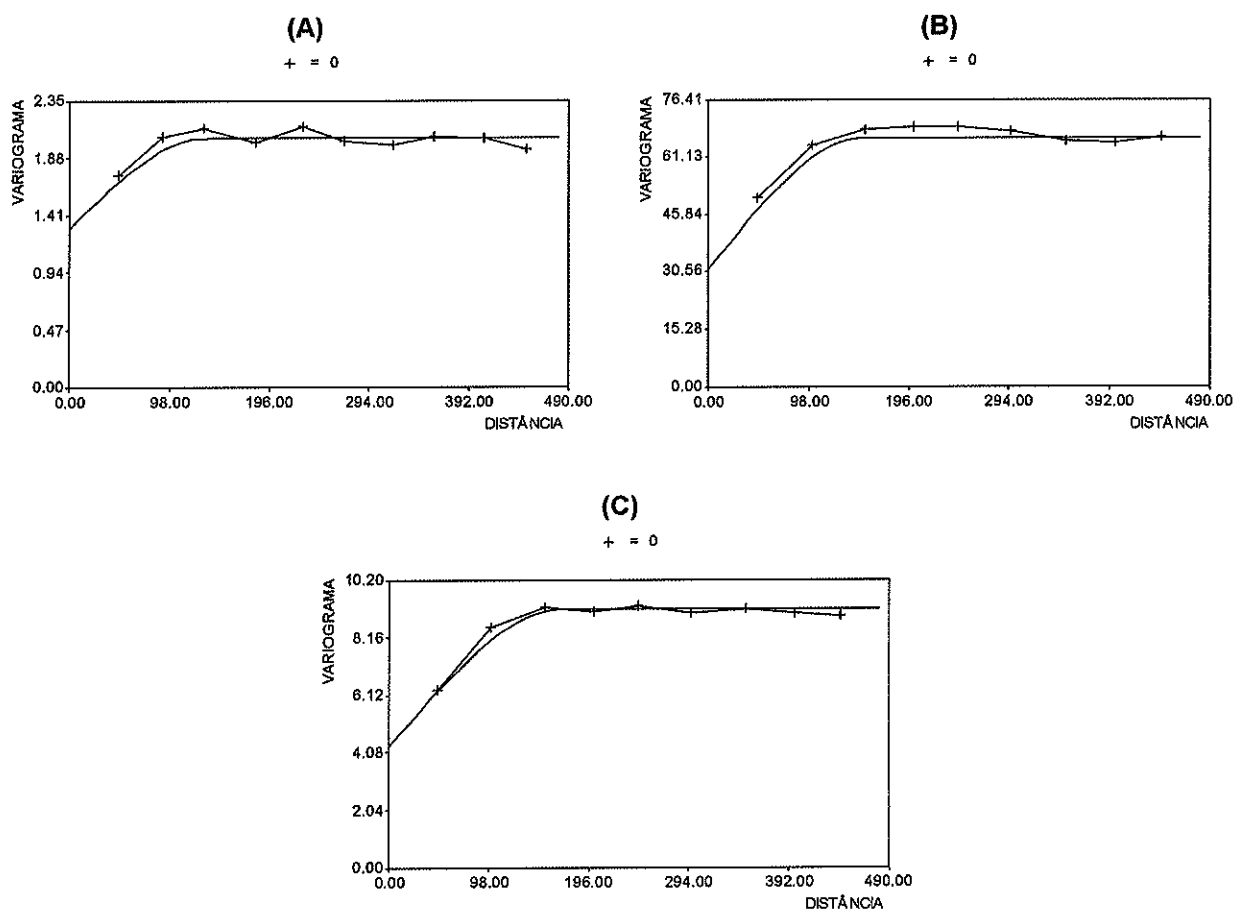


Figura 86: Variogramas experimentais e modelos dos variogramas da Figura 85. (A) variável assimétrica positiva, (B) variável assimétrica negativa e (C) variável simétrica.

5.3.2 Amostragem Aleatória Simples x Amostragem Sistemática

Realizou-se o mesmo procedimento, agora por amostragem aleatória simples, conforme os resultados da Figura 87, na qual observa-se:

- Para a variável assimétrica positiva o modelo foi reproduzido quando 95% dos dados haviam sido coletados (Figura 87-A), representando um economia de 30 unidades amostrais;

- Para a variável assimétrica negativa o modelo foi reproduzido com 80% das unidades amostrais amostradas (Figura 87-B), representado uma economia de 120 unidades amostrais;
- Para a variável simétrica o modelo de variância espacial foi reproduzido com 80% das unidades amostrais sendo sorteadas (Figura 87-C), neste caso houve uma perda de 3% do efeito pepita, porém com uma economia de 120 unidades amostrais.

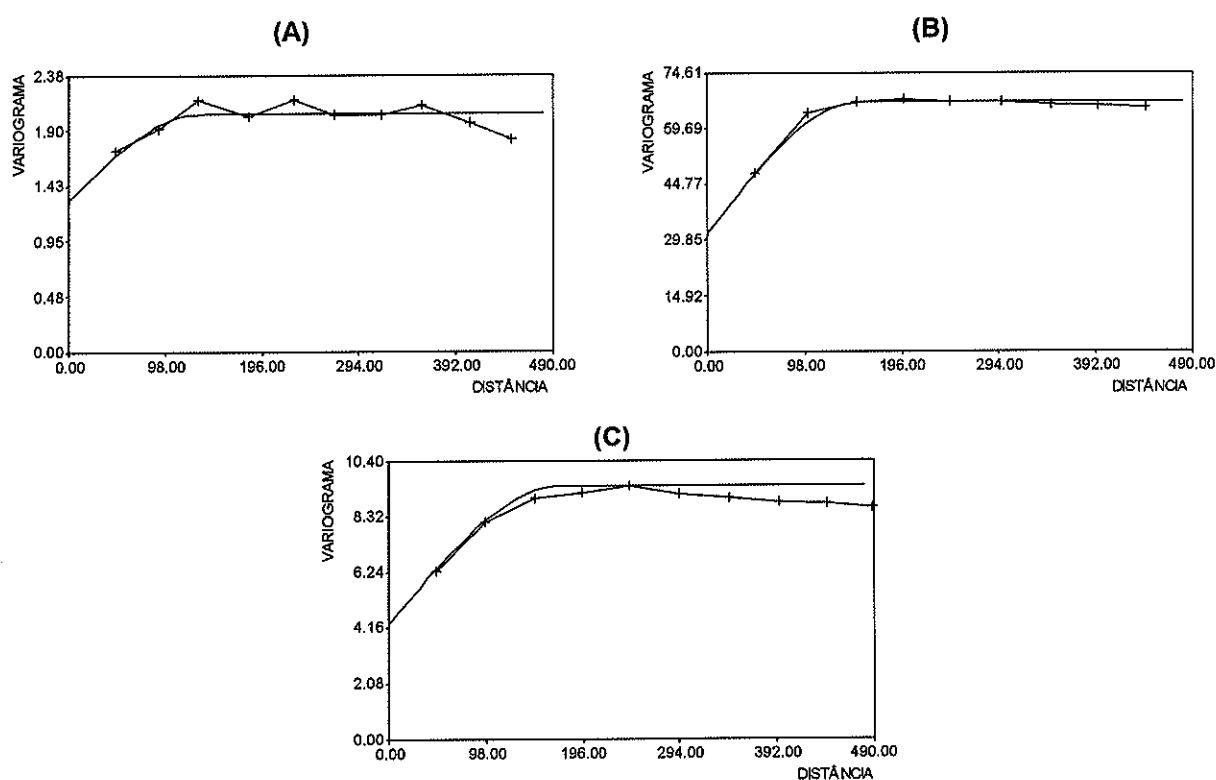


Figura 87: Variogramas experimentais e os modelos de variogramas calculados a partir de pontos de dados gerados por amostragem aleatória simples sobre a malha regular de 20 x 29 nós. (A) variável assimétrica positiva, (B) variável assimétrica negativa e (C) variável simétrica.

Este exercício exemplifica, não apenas a economia direta que se pode obter com estas amostragens, mas, também, como a amostragem aleatória estratificada pode proporcionar resultados sem, contudo, executar toda a amostragem sistemática. Comparando-se, a amostragem aleatória estratificada em relação à amostragem aleatória simples, verifica-se que a primeira é melhor em relação à última, não só pela economia de unidades amostrais, como também pela melhor distribuição espacial proporcionada pela amostragem aleatória estratificada.

O resultado desse exercício pode ser, eventualmente, aplicado a uma campanha de sondagem, planejada sobre uma malha regular, de modo que a seqüência dos furos seja determinada a partir da amostragem aleatória estratificada. Isto significa que, dependendo da variável em estudo, pode-se ter uma boa idéia do variograma experimental sem, contudo, esperar pelo final das sondagens. Além disso, permite calcular os recursos minerais à medida que os trabalhos de pesquisa progridem.

5.4 Superfícies de Tendência

A análise das superfícies de tendência mostra que na amostragem aleatória simples houve a reprodução da tendência dos dados apenas quando o n das amostras é superior a 135 pontos na variável assimétrica positiva, 105 para a variável simétrica e 451 para a variável assimétrica negativa. Os erros RMS calculados confirmam o comportamento das superfícies de grau 1, ou seja, que a tendência regional dos dados de referência é reproduzida esporadicamente. A reprodução esporádica desta tendência pode ser fruto da má distribuição espacial gerada com esta amostragem.

Nesta mesma análise, conduzida nas variáveis amostradas por amostragem aleatória estratificada, a dispersão dos pólos dos planos é menor para todas as variáveis tratadas, se comparada à amostragem anterior, levando a crer que a configuração da malha de pontos obtida por esta amostragem foi mais adequada, e ainda que as distribuições foram melhor amostradas. Os erros quadráticos médios confirmam que a amostragem aleatória estratificada reproduziu melhor as superfícies de tendência de graus 1, 2 e 3. Esta reprodução pode ser relacionada à área amostrada e, também, com a ausência de agrupamentos de dados presentes nas amostras coletadas por amostragem aleatória simples.

A amostragem sistemática não gerou amostras, cujas tendências reproduzissem plenamente as superfícies de tendência das três variáveis dos dados de referência, embora, para grandes amostras, a tendência das superfícies dos três graus tenham se aproximado daquela como mostram os erros quadráticos médios.

5.5 Recursos Geológicos

Para apoiar a discussão de resultados sobre a estimativa de recursos, foram calculados dois índices; um como a média das razões entre os recursos estimados e os recursos reais (\bar{R}_1) e outro pela média das razões entre os recursos estimados e o recurso total (\bar{R}_2), conforme:

$$\bar{R}_1 = \frac{1}{n} \sum_{i=1}^n \left(\frac{RE_i}{RR_i} \right) \quad \text{e} \quad \bar{R}_2 = \frac{1}{n} \sum_{i=1}^n \left(\frac{RE_i}{RT_i} \right)$$

O primeiro índice \bar{R}_1 reflete, em média, a superestimativa ($\bar{R}_1 > 1$) dos recursos estimados em relação aos reais, ou a subestimativa ($\bar{R}_1 < 1$). A Tabela 28 mostra que, à exceção da variável assimétrica positiva, obtida por amostragem aleatória estratificada, todas as estimativas, em média, excedem os recursos reais.

Isto significa que a amostragem produz um viés, nos resultados de cálculos de recursos, por meio da propagação de valores amostrados para regiões vizinhas. A Figura 88 mostra alguns pontos, o efeito de propagação, onde algumas amostras influenciam fortemente os blocos da vizinhança.

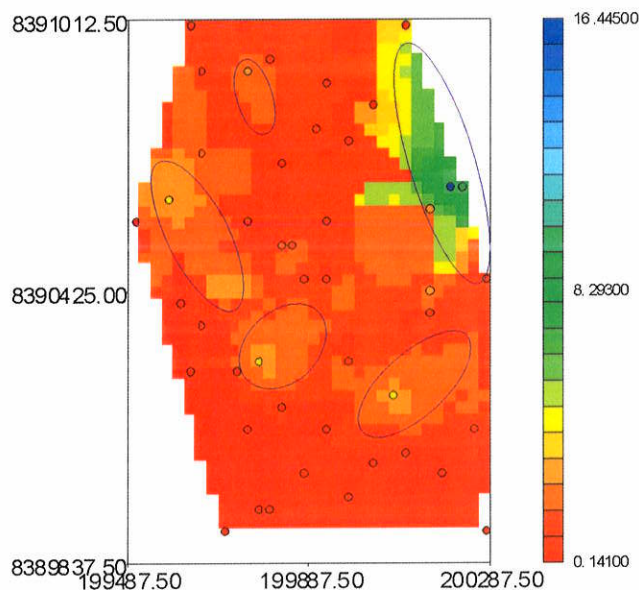


Figura 88: Mapa de pontos em cores proporcionais sobreposto à um mapa de krigagem. As elipses destacam a influência de valores anômalos em áreas sem pontos amostrais.

Tabela 28: Índices médios R_1 .

Amostragem	Assimétrica positiva		Variável Simétrica		Assimétrica Negativa	
	\bar{R}_1	s	\bar{R}_1	s	\bar{R}_1	s
Aleatória simples	1,017	0.2269	1,019	0.0197	1,032	0.1261
Aleatória estratificada	0,970	0.2315	1,012	0.0174	1,018	0.1683
Sistemática	1,058	0.1490	1,040	0.0098	1,108	0.1613

Na verdade, o resultado produzido por amostragens para níveis percentuais baixos, apresenta grande variabilidade. Isto se deve ao pequeno recobrimento do domínio amostrado, bem como à distribuição espacial dos pontos relacionados. Por exemplo, para a variável assimétrica positiva, observa-se que a amostragem aleatória simples (Figura 60) apresenta as maiores variações em relação à amostragem aleatória estratificada (Figura 66). Menor variabilidade significa maior confiabilidade. Uma outra questão a ser analisada refere-se às realizações individuais da amostragem, principalmente para os níveis percentuais baixos. Observe-se que quanto menor o nível percentual, maior a possibilidade de arranjos espaciais diferentes e, conseqüentemente, maior a variabilidade dos resultados. Entre as variáveis estudadas, conclui-se que a variável simétrica é aquela que apresenta os resultados mais próximos do esperado (Figuras 62, 68 e 74). A variável assimétrica negativa, embora tenha um baixo coeficiente de variação (menor que o coeficiente de variação da variável simétrica), apresenta grande variabilidade, muito maior que a variável simétrica, nos resultados da estimativa de recursos (Figuras 64, 70 e 76), pelos motivos apontados anteriormente.

A variável assimétrica positiva apresenta grande variação, devido à característica deste tipo de distribuição. Adiciona-se à ela, os fatores associados às amostragens espaciais.

O segundo índice (\bar{R}_2) mostra, em média, a porcentagem dos recursos estudados em relação aos recursos totais (Tabela 29). Obviamente, como seria de se esperar, todos os índices foram menores que um. Portanto, a superestimativa geral observada anteriormente não constitui um problema, pois os recursos estarão ainda dentro dos máximos possíveis. Cabe ressaltar que, esses índices correspondem aos valores médios alcançados para as realizações individuais.

Tabela 29: Índices médios R2.

Amostragem	Assimétrica positiva		Variável Simétrica		Assimétrica Negativa	
	\bar{R}_1	s	\bar{R}_1	s	\bar{R}_1	s
Aleatória simples	0,915	0.5932	0,882	0.1263	0,883	0.1535
Aleatória estratificada	0,865	0.2531	0,931	0.0672	0,898	0.1746
Sistemática	0,847	0.0743	0,857	0.0707	0,858	0.0740

5.6 Análise Espectral

A análise espectral corrobora com a hipótese que a amostragem aleatória estratificada estimou melhor as populações em estudo. Embora o diagrama da amostragem aleatória simples seja mais homogêneo e aproxime-se mais do espectro dos dados de referência, consideram-se estes resultados menos adequados. O espectro de potência pode ser interpretado como sendo uma medida de variância nos números harmônicos do fenômeno em estudo. Assim sendo, os valores das razões acima de 100% indicam ganho de energia nos espectros de potência, ou seja, aumento da variância. Aqueles valores próximos a 100% indicam que o espectro de potência, ou variância, dos dados de referência foi reproduzido. Neste ponto deve-se levantar uma questão relacionada ao efeito de suavização da krigagem. Pois bem, a krigagem suaviza os valores extremos de uma distribuição, mesmo quando todos os parâmetros utilizados sejam aferidos e adequados. Deste modo, quando os parâmetros definidos para a krigagem o são de tal modo que todos os pontos dentro do domínio populacional sejam estimados com base em algumas poucas unidades amostrais, como foi o caso da estimativa feita nesta análise, o efeito de suavização deverá ser muito grande, pois exceto os pontos condicionantes, se houverem, todos os demais serão estimados próximos à média, ou ainda, com valores semelhantes aos vizinhos mais próximos. Assim sendo, deve-se considerar o melhor resultado aquele onde as razões dos espectros de potência são inferiores a 100%, o que ocorre na amostragem aleatória estratificada.

Capítulo 6

Conclusões

O tipo de amostragem exerce grande influência no recobrimento espacial que as amostras possuirão, considerando-se diferentes valores de n . A amostragem aleatória estratificada é a que apresenta os melhores recobrimentos, mesmo considerando-se amostras pequenas. A amostragem aleatória simples, embora, com resultados semelhantes aos da amostragem aleatória estratificada, resulta em amostras com pontos que tendem a se agrupar em regiões específicas do domínio.

A reprodução mais rápida, ou mais lenta, dos parâmetros estatísticos dependem da forma da distribuição estudada e de sua variância. Amostras retiradas de variáveis assimétricas tendem a ter os parâmetros estatísticos reproduzidos em amostras maiores do que a variável simétrica. Existe, ainda, uma relação entre a média e a variância amostral com o tipo de distribuição. Variáveis assimétricas positivas tendem, com o aumento, ou diminuição, da média amostral a apresentar variância amostral, respectivamente, maior ou menor. Na distribuição simétrica existe um equilíbrio entre o aumento e/ou a diminuição da média amostral com o aumento e/ou diminuição da variância amostral. A variável assimétrica negativa, tende a apresentar comportamento oposto ao da variável assimétrica positiva, ou seja, a tendência é, se houver aumento da média amostral em relação à média populacional, a variância amostral tende a diminuir em relação à variância populacional e vice-versa.

Como em geociências as variáveis analisadas, normalmente, são variáveis regionalizados, os tipos de amostragem, também, desempenham um papel importante na reprodução das estatísticas populacionais. Sob esta ótica, amostragens, que garantam um bom recobrimento espacial do domínio, são mais adequadas, pois resultam em amostras, especialmente, mais representativas. Neste sentido, a amostragem aleatória estratificada mostrou-se mais adequada, independentemente da forma da distribuição estudada.

Neste estudo, conclui-se que, além das condições descritas na literatura, para o cálculo de variogramas experimentais; ou seja, a malha de amostragem, o número de pontos amostrados, a distância média entre pontos e a presença de *outliers*; o tipo de

distribuição estatística tem grande influência nos resultados. Isto é possível de ser observado, nos variogramas experimentais apresentados nos resultados, onde para um mesmo arranjo geométrico de pontos, as variáveis simétrica e assimétrica negativa apresentam comportamentos semelhantes e tendência à convergência em amostras menores.

Como o cálculo, dos variogramas experimentais, depende da distância média entre pontos amostrais, a amostragem aleatória estratificada resulta em amostras onde os variogramas populacionais são reproduzidos mais rapidamente. Conclui-se, ainda, que o efeito pepita é um parâmetro que depende da assimetria da distribuição estatística dos dados, uma vez que, para a variável assimétrica positiva não houve convergência dos efeitos pepita amostrais ao efeito pepita populacional, ao contrário dos resultados obtidos para a variável simétrica. A variável assimétrica negativa apresentou um comportamento intermediário com relação as outras duas variáveis.

A reprodução de tendências regionais, por amostras, está intrinsecamente relacionada a dois fatores; o primeiro é a representatividade estatística das amostras; e o segundo é a representatividade espacial que as amostras possuem, ou seja, para que uma tendência regional seja reproduzida, por uma amostra, é necessário que a maior porção do domínio tenha sido amostrada, não bastando que apenas as estatísticas da população tenham sido bem estimadas.

A quantificação precisa dos recursos geológicos, de uma população, depende da reprodução da textura geológica desta população. O arranjo geométrico, gerado pela amostragem, tem papel fundamental na boa estimativa dos recursos geológicos.

Como pôde ser visto, quanto menor o tamanho da amostra, maior será o número de combinações, de arranjos espaciais de unidades amostrais, possíveis e, por conseguinte, maior será a variação das áreas amostrais. Os recursos geológicos, além de dependerem da reprodução da textura geológica, também dependem da área amostrada no domínio. Deste modo, para amostras pequenas é grande a probabilidade de se amostrar áreas de diferentes tamanhos, logo a variação nos recursos para amostras de tamanhos iguais, nestes níveis percentuais é grande. Neste sentido, a amostragem aleatória estratificada é mais adequada, pois, conforme demonstrado neste trabalho, foi esta amostragem que melhor reproduziu a textura geológica para as três variáveis analisadas e mais rapidamente recobriu, em área, a maior parte do domínio.

A amostragem sistemática, para a reprodução da textura geológica e dos recursos geológicos, dos dados de referência, mostrou-se inadequada, pois para que houvesse a reprodução da textura geológica de referência foi necessário que uma grande quantidade de unidades amostrais fossem coletadas. Em outras palavras, para se obter um variograma estruturado, há necessidade que a abertura da malha regular seja inferior à amplitude do variograma de referência. Isto ocorre apenas em níveis percentuais elevados, ao passo que, para amostragens aleatória simples e estratificada ocorreu em níveis percentuais mais baixos.

A análise espectral, dada a configuração dos parâmetros utilizados, para garantir que todos os pontos do domínio fossem estimados, não apresentou resultados satisfatórios, porém pôde-se concluir que, para a variável assimétrica positiva, esta análise destaca o efeito de suavização da krigagem.

Referências Bibliográficas

- Armstrong M. 1998. *Geoestatística: Avaliação de Jazidas, Planejamento de Lavra e Interface com Análise Econômico-Financeira*. São Paulo, EPUSP. (apostila de curso).
- Barnett V. 1991. *Sample Survey: Principles and Methods*. 2 ed. New York, Edward Arnold Press, 173 p.
- Botter D. A., Paula G. A., Leite J. G., Cordani L. K. 1996. *Noções de Estatística, com Apoio Computacional*. Versão Preliminar. São Paulo, IME - USP, 231p.
- Bourgault G. 1997. Using Non-Gaussian Distributions in Geostatistical Simulations. *Mathematical Geology* **29**(3): 315-334.
- Caers J. 2000. Adding Local Accuracy to Direct Sequential Simulation. *Mathematical Geology* **32**(7): 815-850.
- Caers J. 2001a. Geostatistical Reservoir Modeling Using Statistical Pattern Recognition. *Journal of Petroleum Science and Engineering* **29**: 177-188.
- Caers J., 2001b. Geostatistics for Pattern Recognition to Pattern Reproduction: *In* Nikravesh, M., Aminzadeth, F., Zedeth, L. (eds.) *Soft Computing and Intelligent Data Analysis in Oil Exploration*, Elsevier Publisher, (Unpublished).
- Cochran W. G. 1963. *Técnicas de Amostragem*. Rio de Janeiro, Fundo de Cultura, 555p.

- Conwell P. M., Silliman S. E., Zheng L. 1997. Design of a Piezometer Network for Estimation of the Sample Variogram of Hydraulic Gradient: The Role of the Instrument. *Water Resources Research* **33**(11): 2489-2494.
- Davis J. C. 1986. *Statistics and Data Analysis in Geology*. 2^aed. New York, John Wiley & Sons Inc., 646 p.
- Davis J. C. 2002. *Statistics and Data Analysis in Geology*. 3^a ed. New York, John Wiley & Sons Inc., 638 p.
- Deutsch C. V. & Journel A. G. 1996. *GSLIB: Geostatistical Software Library*. 2^a ed. (Beta Version). Oxford, Oxford University Press, 360 p.
- Deutsch C. V. & Wen X. H. 1998. An Improved Perturbation Mechanism for Simulated Annealing Simulation. *Mathematical Geology* **30**(7): 801-816.
- Goes M. A., Possa M. V., Luz A. B. 1991. *Amostragem de Minérios*. Rio de Janeiro, CTEN, 48p.
- Goovaerts P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford, Oxford University Press, 483 p.
- Gy P. M. 1992. *Sampling of Heterogeneous and Dynamic Material Systems: Theories of Heterogeneity, Sampling, and Homogenizing*. Amsterdam, Elsevier, 653p.
- Harbaugh J. W. & Merriam D. F. 1968. *Computer Applications in Stratigraphic Analysis*. 1^aed. New York, John Wiley & Sons Inc., 282 p.
- Huijbregts C. J. 1975. Regionalized Variables and Quantitative Analysis of Spatial Data. In: Davis, J.C. & McCullagh, M.J. (eds) *Display and Analysis of Spatial data*. New York, John Wiley, pp. 38-53.

- Journel A. G. & Rao S. E. 1997. *Deriving Conditional Distributions From Ordinary Kriging*. 25p. (Relatório SCRF 9)
- Journel A. G. & Huijbregts C.J. 1978. *Mining Geostatistics*. 1^a ed. London, Academic Press, 600p.
- Journel A. G., Kyriakidis P. C., Mao S. 2000. Correcting Smoothing Effect of Estimators: A Spectral Postprocessor. *Mathematical Geology* **32**(7): 787-813.
- Kim Y. C. 1990. *Introductory Geostatistics and Mine Planning*. Tucson, University of Arizona, 212p. (Apostila de curso)
- Krumbein W. C. & Graybill F. A. 1965. *An Introduction to Statistical Models in Geology*. 1^aed. New York, McGraw-Hill Book Company, 475 p.
- Landim P. M. B. 1997. *Análise Estatística de Dados Geológicos*. 1^a ed. São Paulo, Fundação Editora da UNESP, 226p.
- Lloyd C. D. & Atkinson P. M. 1999. Designing Optimal Sampling Configuration with Ordinary Kriging and Indicator Kriging. *In Proceedings of the 4th International Conference on GeoComputation, Virginia*. (http://www.geocomputation.org/1999/065/gc_065.htm).
- Marshak S. & Gautan M. 1998. *Basic Methods of Structural Geology*. 1^a ed. New York, Prentice Hall, 446 p.
- Matheron G. 1971. *The Theory of Regionalized Variables and its Applications*. 211p. (Les Cahiers du Centre de Morphologie Mathématique de Fontainebleau 5)
- Olea R. 1999. *Geostatistics for Engineers and Earth Scientists*. 1^a ed. Massachusetts, Kluwer Academic Publishers, 303 p.

- Olea R. 1991. *Geostatistical Glossary and Multilingual Dictionary*. 1^a ed. Oxford, Oxford University Press, 175p.
- Olea R., Pawlowsky V., 1996, Compensating for Estimation Smoothing in Kriging. *Mathematical Geology* **28**(4): 407-417
- Rayner J. N. 1971. *An Introduction to Spectral Analysis*. 1^a ed. London, Pion Limited, 174 p.
- Russo D. 1984. Design of Optimal Sampling Network for Estimating the Variogram. *Soil Science Society of America Journal* **48**: 708-716
- Silliman S. E. & Berkowitz B. 2000. The Impact of Biased Sampling on the Estimation of the Semivariogram within Fractures Media Containing Multiple Fracture Sets. *Mathematical Geology* **32**(5): 543-560
- Soares A. 2001. Direct Sequential Simulation and Cosimulation. *Mathematical Geology* **33**(8): 911-926
- Srivastava M. 1994. An Overview of Stochastic Methods for Reservoir Characterization, in Yarus J. M. & Chambers R. L. 1994. *Stochastic Modeling and Geostatistics – Principles, Methods and Case Studies*. (AAPG – Computer Applications in Geology 3).
- Warrick A. W. & Myers D. E. 1987. Optimization of Sampling Location for Variograms Calculations. *Water Resources Research* **23**(3): 496-500
- Webster R., Oliver M. A. 1992. Sample Adequately to Estimate Variograms of Soil Properties. *Journal of Soil Science* **43**: 177-192
- Yamamoto J. K. & Rocha M. M. 2001. Conceitos Básicos. in Yamamoto J. K. (ed) 2001. *Avaliação e Classificação de Reservas Minerais*. 1^a ed. São Paulo, EDUSP, 226 p.

- Yamamoto J. K. 1997. Convex_hull – A Pascal Program for Determining the Convex Hull for Planar Sets. *Computers & Geosciences* **23**(7): 725-738
- Yamamoto J. K. 1998. A Review of Numerical Methods for the Interpolation of Geological Data. *Anais da Academia Brasileira de Ciências* **70**(1): 91-116
- Yamamoto J. K. 2000. An Alternative Measure of the Reliability of Ordinary Kriging Estimates. *Mathematical Geology* **32**(4): 489-509
- Yamamoto J. K. 2003. Sistema GeoVisual versão 2.1
- Yamane T. 1967. *Elementary Sampling Theory*. 1^a ed. New York, Prentice-Hall, 405 p.
- Zheng L. & Silliman S. E. 2000. Estimating the Theoretical Semivariogram from Finite Numbers of Measurements. *Water Resources Research* **36**(1): 361-366