

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE GEOCIÊNCIAS

**MAPEAMENTO DA INCERTEZA DE REALIZAÇÕES CONDICIONAIS  
DA SIMULAÇÃO PLURIGAUSSIANA TRUNCADA**

**SANTIAGO DÍAZ LÓPEZ**

Orientador: Prof. Dr. Jorge Kazuo Yamamoto

**DISSERTAÇÃO DE MESTRADO**

Programa de Pós-Graduação em Recursos Minerais e Hidrogeologia

**VERSÃO CORRIGIDA**

SÃO PAULO

2015

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ficha catalográfica preparada pelo Serviço de Biblioteca e Documentação do Instituto de Geociências da Universidade de São Paulo

Diáz López, Santiago

Mapeamento da incerteza de realizações condicionais da simulação pluriGaussiana truncada / Santiago Diáz López. - São Paulo, 2015.

109 p.: il + anexos

Dissertação (Mestrado) : IGc/USP

Orient.: Yamamoto, Jorge Kazuo

1. Geoestatística
2. Simulação pluriGaussiana
3. Mapeamento da Zona de Incerteza I. Título

*Dedico este trabajo a mi familia por su amor, trabajo y sacrificio en todos estos años, que gracias a ustedes y con a su constante ayuda he logrado ser la persona que soy. A mi madre por hacer de mí una mejor persona a través de sus consejos, enseñanza y amor. A mi padre por aconsejarme y darme apoyo incondicional. A mi hermana por estar de mi lado y ser una amiga en mi vida.*

## **AGRADECIMENTOS**

Ao orientador Prof. Dr. Jorge Kazuo Yamamoto por brindar-me com a oportunidade de realizar estudos de pós-graduação. Por sua orientação e pelo apoio recebido ao longo deste mestrado.

À CAPES, pelo auxílio concedido na forma de bolsa de estudos.

Ao Instituto de Geociências da Universidade de São Paulo pela formação acadêmica recebida e pelo apoio logístico de sua estrutura. Aos seus professores pelas orientações oportunas em muitas disciplinas, especialmente ao professor Dr. Marcelo Monteiro da Rocha.

Um profundo agradecimento ao Mario Gonzalez, Lorena e família pelo apoio, amizade e confiança incondicionais que me deram. Agradeço vocês por ajudar-me neste grande êxito, desde o começo até o fim.

À professora Luz Marina por sua amizade, confiança e seu apoio incondicional.

À Lina Maria por sua amizade e seu apoio.

Agradeço a todos aqueles que direta e indiretamente fizeram parte desta nova conquista.

Aos colegas do Laboratório de Informática Geológica e da sala 105 pelos momentos compartilhados.

## RESUMO

LÓPEZ, S. D. Mapeamento da incerteza de realizações condicionais da simulação pluriGaussiana truncada. 2015. 109 f. Dissertação (Mestrado) – Instituto de Geociências, Universidade de São Paulo, São Paulo.

A simulação pluriGaussiana truncada é usada para simulação de dados categóricos, tais como fácies em um contexto geológico. Esse método é uma extensão da Gaussiana truncada, mas que permite a simulação de feições geológicas complexas. Nessa dissertação, a simulação pluriGaussiana truncada foi baseada nas curvas de proporção obtidas da interpolação multiquádrica. Esse método foi testado com uma amostra composta por 24 sondagens obtidas de um modelo geológico sintético. As realizações da simulação pluriGaussiana truncada foram pós-processadas para obter o modelo médio, bem como as incertezas associadas. Os resultados mostraram que as incertezas estão associadas com os contatos litológicos.

**Palavras-chaves:** Geoestatística, Simulação pluriGaussiana, Mapeamento da Zona de Incerteza.

## ABSTRACT

LÓPEZ, S. D. Mapeamento da incerteza de realizações condicionais da simulação pluriGaussiana truncada. 2015. 109f. Dissertação (Mestrado) – Instituto de Geociências, Universidade de São Paulo, São Paulo.

The truncated pluriGaussian simulation is used for categorical data simulation, such as facies in a geological context. This method is an extension of the Gaussian truncated, but it allows simulation of complex geological features. In this dissertation, the method of truncated pluriGaussian simulation was based on proportion curves computed from multiquadric interpolation. This method was tested with a sample composed of 24 drill holes drawn from a synthetic geologic model. Realizations of truncated pluriGaussian simulation are post-processed to get the average model and associated uncertainty as well. Result showed that uncertainties are always associated with geological contacts.

**Keywords:** Geostatistics, truncated pluriGaussian, Mapping of uncertainty zone.

## ÍNDICE DE FIGURAS

Figura 1.1. Ilustração de dois conjuntos disjuntos $F$ e $F$ , respectivamente. ....	18
Figura 1.2. Transformação dos dados originais para um valor padrão normal. ....	22
Figura 1.3. Zona de incerteza definida no intervalo $0.276 < IMQ * x_0; kmax < 0.724$ . .....	27
Figura 2.1 . Realização unidimensional de um campo aleatório associado a um modelo de covariância em 3D. ....	31
Figura 2.2. Ilustração do método de bandas rotativas em duas dimensões. O ponto sobre o plano é projetado em cada linha (quatro linhas) e influenciado o valor simulado. Fonte: (Olea, 1999, p. 161). ....	34
Figura 2.3. Conjunto de 300 direções distribuídas aleatoriamente sobre uma esfera unitária em $R^3$ (sequência de van der Corput). ....	35
Figura 2.4. Fluxograma do método de bandas rotativas condicional de uma função aleatória estacionária. ....	36
Figura 2.5. Mapa de localização de pontos da amostra (esquerda) e o histograma dos dados (direita). ....	37
Figura 2.6. Variograma experimental ajustado a um modelo teórico exponencial. ....	38
Figura 2.7. Localização dos vizinhos próximos ao ponto de coordenadas $x = 30; y =$ $30; z = 0$ . ....	39
Figura 2.8. Um conjunto de quatro direções distribuídas sobre uma esfera unitária em $R^3$ (sequência de Van der Corput). ....	40
Figura 2.9. Realização unidimensional de um campo aleatório associado a um modelo de covariância esférico em 3D. Cada linha gerada é discretizada num número de intervalos e para cada intervalo um sinal (-1 ou +1) é gerado aleatoriamente. ....	42
Figura 2.10. Duas realizações de um campo aleatório simulado via bandas rotativas usando 4 e 1000 linhas, respectivamente. O modelo de covariância consiste em um modelo esférico isotrópico sem efeito pepita. ....	46

Figura 3.1. Curvas de proporção vertical e horizontal (Fonte: Ravenne, 2002).....	48
Figura 3.2. Distribuição lateral das fácies que exibem mudanças na direção horizontal. ...	49
Figura 3.3. Mapa de localização da posição dos poços e as curvas de proporção vertical calculadas para cada área (Fonte: Chihi & De Marsily, 2009, p. 460).....	50
Figura 3.4. Matriz de proporção para cada unidade associada a cada área delimitada (Fonte: Chihi & De Marsily, 2009, p. 461). ....	51
Figura 3.5. Duas subsidências de um reservatório e seus níveis de referência (Fonte: Armstrong <i>et al.</i> , 2003, p. 36).....	53
Figura 3.6. Mapa de localização dos 24 poços em 3D e 2D. Na parte superior se mostra em 3D a posição dos poços e na parte inferior da Figura, círculos cheios se referem aos poços numerados.....	54
Figura 3.7. Mapa de localização dos poços sobre a matriz de proporções.....	55
Figura 3.8. Digrama em pizza das fácies por cada poço e delimitação de cada zona. ....	55
Figura 3.9. Mapa da localização dos poços e o poço a ser interpolado, (■), com indicação dos vizinhos mais próximos. ....	56
Figura 3.10. Mapa de localização de pontos e ponto a ser interpolado, (■), com indicação dos vizinhos mais próximos (dados 3D). ....	58
Figura 3.11. Curva de proporção vertical associada à zona B.....	62
Figura 3.12. Matriz de proporção para cada unidade associada a cada área delimitada Zona A, Zona B e Zona C.....	62
Figura 3.13. A curva contínua (função Gaussiana) é truncada em quatro parâmetros limiares $t_i = 1, \dots, 4$ geradas a partir do conhecimento geológico da distribuição das fácies na área (Fonte: Freulon & de Fouquet, 1991). ....	63
Figura 3.14. Distribuição de uma normal padrão codificada em três tipos de fácies (Fonte: Pyrcz & Deutsch, 2014, p. 237).....	65
Figura 3.15. Representação esquemática em 2D, com duas funções aleatórias Gaussianas e seus valores simulados para os pontos $x_1$ e $x_2$ . ....	66
Figura 3.16. Partição do domínio em retângulos (à esquerda) e seus valores de corte (à direita).....	67

Figura 3.17. Esquemática de duas simulações pluriGaussianas truncadas como diferentes regras de truncamento.....	68
Figura 3.18. Esquema de uma função Gaussiana contínua, no caso não contínuo truncada por três valores de corte. Numa certa localização contatos proibidos são mostrados na ausência da fácies 2 (Fonte: Folle (2009, p. 125)).....	69
Figura 3.19. Regra de truncamento para três fácies e duas funções aleatórias Gaussianas.	69
Figura 3.20. Regra de truncamento baseado na ordem de deposição das fácies para os cinco poços.....	71
Figura 3.21. Localização dos pontos da amostra, no qual os quadros azuis representam valores menores que a mediana (fácies 1) e quadros amarelos, valores maiores que a mediana (fácies 2).....	78
Figura 3.22. Curva de proporção horizontal (esquerda) e uma regra de truncamento (direita). ....	78
Figura 3.23. Proporção da fácies 1 e 2 sob a curva de uma distribuição normal padrão. ...	79
Figura 3.24. Variograma experimental (quadros e linhas ponteadas) e modelado (linhas sólidas) das indicadores na horizontal. ....	80
Figura 3.25. Ilustração do amostrador de Gibbs em 2D.....	82
Figura 3.26. Poço discretizado em quatro amostras de igual comprimento.....	84
Figura 3.27. Distribuição normal padrão, ilustrando a área sob a curva das proporções para as fácies, <b>F1</b> e <b>F2</b> .....	85
Figura 3.28. Fluxograma da simulação plurigaussiana truncada.....	87
Figura 5.1. Curva de proporção vertical das fácies. ....	95
Figura 5.2. Descretização da área de uma distribuição normal padrão, dada por cada parâmetro de corte para cada função aleatória, no caso <b>Z1</b> e <b>Z2</b> .....	97
Figura 5.3. Variogramas das indicadores dos dados experimentais (quadrados e linhas pontilhadas) e modelado (linhas solidas) nas direção vertical e horizontal. ....	98
Figura 5.4. Realizações derivadas da simulação pluriGaussiana truncada. A) realização 1, B) realização 2, C) realização 3 e D) realização 4.....	99

Figura 5.5. (continuaão) Realizaões derivadas da simulaão pluriGaussiana truncada. E) realizaão 5, F) realizaão 6, G) realizaão 7 e H) realizaão 8.....	100
Figura 5.6. (continuaão) Realizaões derivadas da simulaão pluriGaussiana truncada. I) realizaão 9 e J) realizaão 10. ....	101
Figura 5.7. Perfis horizontais na profundidade de $Z = 10$ para cada uma das realizaões selecionadas aleatoriamente (#2, #6, #4 e #10). ....	102
Figura 5.8. Mapeamento da zona de incerteza a partir das 10 realizaões. ....	104

## ÍNDICE DE TABELAS

Tabela 2. 1. Estatísticas descritivas para a amostra em estudo. ....	37
Tabela 2. 2. Pontos de dados vizinhos ao ponto a ser estimado ( $x = 30; y = 30; z = 0$ )....	39
Tabela 2. 3. Direção das linhas na direção de anisotropia. ....	40
Tabela 2. 4. Projeção de cada ponto sobre a linha em sua respectiva direção. ....	40
Tabela 2. 5. Valor aleatório fixo sobre o eixo real. ....	41
Tabela 2. 6. Intervalos associados a cada ponto sobre as linhas. ....	41
Tabela 2. 7. Posição do ponto sobre a função independente no intervalo $-1, 1$ . ....	42
Tabela 2. 8. Valor simulado associado a um modelo de covariância esférico. ....	42
Tabela 2. 9. Valor de cada campo, constante vezes o campo e o valor simulado não condicional. ....	43
Tabela 2. 10. Ponto a simular no grid projetado sobre as quatro linhas. ....	44
Tabela 2. 11. Posição do ponto a simular na grid sobre a função independente no intervalo $-1, 1$ . ....	44
Tabela 2. 12. Valores residuais computados dos pontos amostrados e o valor não condicional. ....	44
Tabela 3. 1. Pontos não amostrados para interpolação de tipo da variável categórica da célula 45 da MP. ....	56
Tabela 3. 2. Codificação binária das fácies para cada ponto coordenado no poço, também as proporções, variância e variância total. ....	57
Tabela 3. 3. Pontos de dados vizinhos codificados ao ponto ser estimado ( <b>135, 90, 100</b> )... ..	58
Tabela 3. 4. Funções indicadoras nos pontos amostrais para interpolação do tipo categórico ao ponto a estimar e resultados. ....	60
Tabela 3. 5. Resultados da interpolação dos tipos no ponto não amostrado. ....	60
Tabela 3. 6. Proporções de todas as fácies nível a nível em percentagens. ....	72
Tabela 3. 7. Parâmetros limiares correspondentes a cada proporção mostrada na tabela 3.6..	73

Tabela 3. 8. Proporções e parâmetros de corte da fácies 1 e fácies 2, respectivamente. ....	79
Tabela 3. 9. Modelo de variograma gaussiano para a função Gaussiana aleatória. ....	80
Tabela 3. 10. Parâmetros de corte das fácies 1 e fácies 2. ....	84
Tabela 3. 11. Interações do método de Gibbs. ....	87
Tabela 4. 1. Proporções do tipo de fácies dos dados amostrados. ....	90
Tabela 5. 1. Parâmetros de corte derivados dos dados amostrais. ....	96
Tabela 5. 2. Parâmetros de modelos de variograma dos campos aleatórios gaussianos usados na simulação pluriGaussiana. ....	97
Tabela 5.3. Proporções computadas dos perfis das realizações #2, #6, #4 e #10, respectivamente. ....	103
Tabela 5. 4. Proporções das fácies após do procedimento do cômputo da zona de incerteza. ....	104

# SUMÁRIO

<b>INTRODUÇÃO .....</b>	<b>15</b>
<b>OBJETIVOS.....</b>	<b>16</b>
<b>CAPÍTULO 1 .....</b>	<b>17</b>
1.1. GEOESTATÍSTICA .....	17
1.1.1. FUNÇÃO INDICADORA E PROPRIEDADES.....	18
1.1.2. COVARIÂNCIA ESPACIAL, VARIOGRAMA E VARIOGRAMA CRUZADO .....	19
1.1.3. VARIOGRAMA E VARIOGRAMA CRUZADO.....	19
1.2. TRANSFORMAÇÃO DE DADOS.....	21
1.2.1. TRANSFORMADA GAUSSIANA.....	21
1.3. KRIGAGEM .....	22
1.3.1. KRIGAGEM SIMPLES.....	23
1.4. EQUAÇÕES MULTIQUÁDRICAS.....	23
1.4.1. ZONA DE INCERTEZA .....	26
<b>CAPÍTULO 2 .....</b>	<b>28</b>
2.1. SIMULAÇÃO ESTOCÁSTICA.....	28
2.2. MAPEAMENTO DE INCERTEZA PARA VARIÁVEIS CATEGÓRICAS DERIVADA DE SIMULAÇÕES ESTOCÁSTICAS.....	29
2.3. MÉTODO DE BANDAS ROTATIVAS.....	29
2.2.1. ERGODICIDADE E NORMALIDADE MULTIVARIADA.....	33
2.2.2. DISTRIBUIÇÃO DAS LINHAS SOBRE A ESFERA.....	34
2.2.3. NÚMERO DE LINHAS USADAS NA SIMULAÇÃO DE CADA ESTRUTURA .....	35
2.2.4. REALIZAÇÕES CONDICIONAIS DE VALORES GAUSSIANOS.....	35
2.4. EXEMPLO DE APLICAÇÃO DO MÉTODO DE BANDAS ROTATIVAS .....	37
<b>CAPÍTULO 3 .....</b>	<b>47</b>
3.1. SIMULAÇÃO PLURIGAUSSIAN TRUNCADA.....	47
3.2. CURVAS DE PROPORÇÃO VERTICAL E HORIZONTAL.....	47
3.1.1. CONSTRUÇÃO DE UMA MATRIZ 3D DE CURVAS DE PROPORÇÕES.....	49

3.1.2.	<i>NÍVEL DE REFERÊNCIA</i> .....	52
3.1.3.	<i>EXEMPLO DE APLICAÇÃO DO CÁLCULO DE UMA MATRIZ DE PROPORÇÃO MEDIANTE AS EQUAÇÕES MULTIVARIÁDAS</i> .....	54
3.3.	TRUNCAMENTOS E PARÂMETROS LIMIARES (THRESHOLDS) .....	63
3.3.1.	<i>EXEMPLOS DE APLICAÇÃO DE TRUNCAMENTO E PARÂMETROS LIMIARES (THRESHOLDS)</i> .....	69
3.4.	CÁLCULO E MODELAGEM DE VARIOGRAMAS EXPERIMENTAIS .....	73
3.4.1.	<i>EXEMPLO DE APLICAÇÃO NA OBTENÇÃO DE UM MODELO DE COVARIÂNCIA APROPRIADO</i> ..	77
3.5.	AMOSTRADOR DE GIBBS .....	81
3.5.1.	<i>VALORES GAUSSIANOS DAS INFORMAÇÕES CATEGÓRICAS</i> .....	81
3.5.2.	<i>EXEMPLO DE APLICAÇÃO PARA GERAR VALORES ALEATÓRIOS GAUSSIANOS EM UM DETERMINADO INTERVALO E UM MODELO CORRETO DE COVARIÂNCIA</i> .....	84
<b>CAPÍTULO 4</b>	.....	<b>88</b>
4.1.	MATERIAIS E MÉTODOS .....	88
4.2.	ETAPAS .....	91
4.2.1.	<i>SIMULAÇÃO PLURIGAUSSIANA TRUNCADA</i> .....	91
4.2.2.	<i>CÁLCULO DA CURVA DE PROPORÇÃO NA DIREÇÃO VERTICAL</i> .....	91
4.2.3.	<i>PARÂMETROS DE CORTE</i> .....	92
4.2.4.	<i>VARIOGRAMA E ANÁLISE ESTRUTURAL</i> .....	93
4.2.5.	<i>VALORES GAUSSIANOS DAS INFORMAÇÕES CATEGÓRICAS</i> .....	93
4.2.6.	<i>MAPEAMENTO DA ZONA DE INCERTEZA DAS SIMULAÇÕES ESTOCÁSTICAS</i> .....	94
<b>CAPÍTULO 5</b>	.....	<b>95</b>
5.1.	RESULTADOS E DISCUSSÃO .....	95
<b>CAPÍTULO 6</b>	.....	<b>105</b>
6.1.	CONCLUSÕES .....	105
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	.....	<b>106</b>
<b>ANEXO A</b> .....		<b>109</b>

## INTRODUÇÃO

A simulação geoestatística é uma ferramenta que foi adotada pela geoestatística para gerar distribuições de uma variável de interesse, quando esta não é suficientemente amostrada. Mas, suas realizações não estão isentas de incertezas, dado que os resultados foram gerados a partir de uma amostra, e esta amostra por si só é uma fonte de incerteza, dada à falta de conhecimento por amostragem insuficiente, bem como a natureza estocástica da variável geológica (Yamamoto et al., 2014, p. 232). O variograma é usado para definir o modelo de correlação espacial, que será empregado para gerar as realizações da simulação estocástica. As simulações podem ser condicionais gerando realizações que honram os pontos amostrais e não condicionais quando isso não acontece. Elas podem ser aplicadas tanto para variáveis contínuas como para variáveis discretas. No caso de variáveis discretas, a simulação pode ser baseada em células (Deutsch, 2002, p. 153).

Existem basicamente dois métodos para simulação baseada em células: simulação sequencial indicadora e a pluriGaussiana truncada (SPGT). Nesse estudo, será considerado o método da simulação pluriGaussiana truncada, cujo algoritmo é mais complexo em relação à simulação indicadora sequencial.

Assim, considerando que existem realizações para cada pixel da malha 2D ou voxel de um modelo 3D, obtém-se em cada localização simulada uma distribuição de probabilidades, que é caracterizada por uma média e variância. Estas estatísticas são válidas para as realizações. A partir delas, pode-se caracterizar um modelo médio e associado a isso, uma zona de incerteza.

Nesse trabalho, foi adotada a proposta de Yamamoto & Landim (2013, p. 172) para pós-processamento das realizações com vistas ao mapeamento das incertezas associadas às realizações da simulação pluriGaussiana truncada.

A abordagem do tema proposto inicia-se no capítulo 1 com conceitos básicos da geoestatística, funções indicadoras e mapeamento da incerteza. No capítulo 2 é descrito o método de bandas rotativas como método de simulação estocástica na estimativa de um atributo espacial contínuo. No capítulo 3, o método da simulação pluriGaussiana truncada é descrito em quatro passos: cômputo de curvas de proporção, parâmetros de truncamento, análise variográfica e o amostrador de Gibbs, este último usado para obtenção de valores gaussianos a partir de valores categóricos. Nesse sentido, o método é detalhado passo a passo

tanto na teoria como na aplicação com exemplos práticos. No capítulo 4 se dá apresentação dos materiais e métodos, seguindo dos resultados e discussão no capítulo 5. Finalmente, no capítulo 6 se apresenta as conclusões.

## **OBJETIVOS**

Estudar aplicação da simulação pluriGaussiana truncada (SPGT) na reprodução da variabilidade de uma variável categórica, bem como o mapeamento da zona de incerteza derivada das realizações independentes da SPGT. Além disso, fazer a implementação do algoritmo das equações multiquádricas no cômputo das curvas de proporção vertical e horizontal.

# CAPÍTULO 1

## 1.1. GEOESTATÍSTICA

A geoestatística é uma parte da estatística “que estuda fenômenos flutuantes no tempo” e/ou espaço, que tem como objetivo o entendimento e modelagem da variabilidade espacial (Deutsch & Journel, 1992, p. 9). Esta ferramenta é baseada na aplicação da teoria das variáveis regionalizadas, que é definida como uma função numérica com distribuição espacial e certa continuidade aparente, que varia de um lugar a outro (Blais & Carlier, 1968 apud Yamamoto & Landim, 2013, p. 24). Desta forma, na geoestatística define-se uma função aleatória como um conjunto de variáveis aleatórias  $Z$  caracterizada com média  $m$ , variância  $S^2$  e uma função de distribuição acumulada (Yamamoto & Landim, 2013, p. 24). Além disso, estas funções aleatórias apresentam certas características tais como: suporte, localização, continuidade e anisotropia.

Segundo Yamamoto & Landim (2013, p. 25), as variáveis aleatórias regionalizadas são subdivididas em variáveis contínuas e discretas. No contexto geológico, uma variável discreta, por exemplo, pode representar a cor de rocha, litologia entre outros.

O objetivo principal da geoestatística é inferir as propriedades do fenômeno espacial em estudo, por meio da interpolação ou estimativa (Yamamoto & Landim, 2013, p. 19). O ponto de partida para esse processo é o cálculo e modelagem do variograma experimental, que descreve a correlação espacial dos dados. A definição de variograma é baseada na hipótese intrínseca (Clark, 1979) e estacionaridade de segunda ordem:

1. Seja  $Z(x)$  uma função aleatória com esperança finita e constante no domínio, ou seja, a média dos incrementos é invariante para qualquer translação de vetor  $h$ .

$$E[Z(x+h) - Z(x)] = 0 \quad (1.1)$$

2. Para qualquer vetor  $h$ , a variância do incremento está definida e é uma função única da distância:

$$Var[Z(x+h) - Z(x)] = 2\gamma(h) \quad (1.2)$$

### 1.1.1. FUNÇÃO INDICADORA E PROPRIEDADES

Seja  $F$  um conjunto aleatório (todos os pontos que pertencem a uma fácies em particular) e  $\bar{F}$  o complemento (todos os pontos que não estão em  $F$ ). Assim, uma função indicadora  $I_F(x)$  é codificada com 0 ou 1 se o ponto  $x$ , está ausente ou presente em  $F$ , respectivamente. A Figura 1.1 ilustra uma área de interesse que contém  $F$  e seu complemento  $\bar{F}$ . Desta forma, a primeira propriedade derivada de uma função indicadora é a união de  $F$  e o complemento, de onde se obtém o espaço completo:

$$I_{F \cup \bar{F}}(x) = I_F(x) + I_{\bar{F}}(x) = 1 \quad (1.3)$$

O valor esperado ou valor médio derivado dos dados experimentais amostrados é o espaço ocupado, ou seja, a proporção. De onde  $P_F(x)$  é a probabilidade que o ponto  $x$  encontra-se sobre  $F$ , tal que:

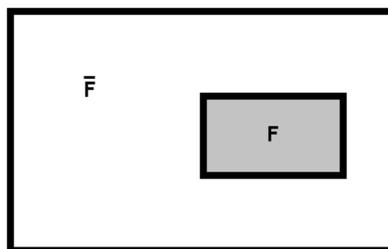
$$P_F(x) = E[I_F(x)] \text{ com } 0 \leq P_F(x) \leq 1 \quad (1.4)$$

Como o valor esperado é linear, temos:

$$P_F(x) + P_{\bar{F}}(x) = 1 \quad (1.5)$$

E a variância computada, com  $[I_F(x)]^2 = I_F(x)$ :

$$Var[I_F(x)] = E[I_F(x)] - \{E[I_F(x)]\}^2 = P_F(x)(1 - P_F(x)) \leq 0.25 \quad (1.6)$$



**Figura 1.1.** Ilustração de dois conjuntos disjuntos  $F$  e  $\bar{F}$ , respectivamente.

### 1.1.2. COVARIÂNCIA ESPACIAL, VARIOGRAMA E VARIOGRAMA CRUZADO

Da mesma forma que na estatística básica, a média, a variância e a covariância são definidas para as variáveis regionalizadas discretas, de acordo com as seguintes relações (Yamamoto & Landim, 2013, p. 33-36):

$$\mu = E\{I_F(x)\} = P_F(x) \quad (1.7)$$

$$Var(I_F(x)) = E\{[I_F(x) - \mu]^2\} \quad (1.8)$$

$$C_F(x, x + h) = E[I_F(x)I_F(x + h)] - \mu^2 \quad (1.9)$$

Além disso, a covariância não centrada de uma variável discreta para dois pontos  $x$  e  $x + h$  separados por uma distância  $h$  é dada por:

$$\begin{aligned} C_F(x, x + h) &= E[I_F(x)I_F(x + h)] = P[I_F(x) = 1 \text{ e } I_F(x + h) = 1] \\ &= P[x \in F \text{ e } x + h \in F] \end{aligned} \quad (1.10)$$

Esta é uma medida de probabilidade onde ambos os pontos se encontram em  $F$ . Similarmente, a covariância centrada  $\sigma_F(x, x + h)$  e a covariância cruzada centrada entre as fácies  $F_i$  e  $F_j$  é denotado por  $\sigma_{F_i F_j}(x, x + h)$  são respectivamente definidas por:

$$\begin{aligned} \sigma_F(x, x + h) &= E\{[I_F(x) - P_F(x)][I_F(x + h) - P_F(x + h)]\} \\ &= C_F(x, x + h) - P_F(x)P_F(x + h) \end{aligned} \quad (1.11)$$

$$\sigma_{F_i F_j}(x, x + h) = E\left\{[I_{F_i}(x) - P_{F_i}(x)][I_{F_j}(x + h) - P_{F_j}(x + h)]\right\} \quad (1.12)$$

### 1.1.3. VARIOGRAMA E VARIOGRAMA CRUZADO

A função variograma das indicadoras, no caso estacionário, é definida como a variância do incremento  $[I_F(x + h) - I_F(x)]$ , conforme a seguinte expressão (Armstrong et al. 2003, p. 11):

$$\begin{aligned} \gamma_F(h) &= \frac{1}{2} Var\{I(x + h) - I(x)\} \\ &= \frac{1}{2} \{E([I_F(x) - I_F(x + h)]^2) - (E[I_F(x) - I_F(x + h)])^2\} \end{aligned} \quad (1.13)$$

O variograma cruzado entre as fácies  $F_i$  e  $F_j$  é definido como:

$$\gamma_{F_i F_j}(x, x+h) = \frac{1}{2} E \left\{ [I_{F_i}(x) - I_{F_i}(x+h)] [I_{F_j}(x) - I_{F_j}(x+h)] \right\} \quad (1.14)$$

Expandindo o produto da equação anterior, dois termos da resolução desaparecem porque  $I_{F_i}$  e  $I_{F_j}$  ambos não podem tomar o valor 1 para um mesmo ponto. Então, a equação anterior é dada por:

$$\gamma_{F_i F_j}(x, x+h) = -\frac{1}{2} \left\{ E [I_{F_i}(x) I_{F_j}(x+h)] + E [I_{F_j}(x) I_{F_i}(x+h)] \right\} \quad (1.15)$$

A seguir, serão apresentadas quatro propriedades dos variogramas das indicadores, e suas demonstrações poderão ser revisadas com detalhe em Armstrong et al. (2003, p. 11-15).

#### **Propriedade 1:**

Dado que o variograma das indicadores tomam unicamente os valores 0 e 1, então o variograma satisfaz:

$$\gamma_F(x, x+h) = \frac{1}{2} \text{Var}[I_F(x) - I_F(x+h)] \leq 0.5 \quad (1.16)$$

#### **Propriedade 2:**

No caso estacionário, o variograma para duas fácies, tem-se:

$$\gamma_F(x, x+h) = \gamma_{\bar{F}}(x, x+h) \quad (1.17)$$

#### **Propriedade 3:**

Seja  $F$  um espaço aleatório, se diz que o variograma depende unicamente do vetor entre dois pontos, mas não de sua posição e similarmente das covariâncias. Nesse sentido, temos:

$$\sigma_F(x, x+h) = \sigma_{\bar{F}}(x, x+h) = -\sigma_{F\bar{F}}(x, x+h) = -\sigma_{\bar{F}F}(x, x+h) \quad (1.18)$$

#### **Propriedade 4:**

Para três fácies A, B e C presentes numa área de interesse e num caso estacionário, temos:

$$\begin{aligned}
\sigma_A(x, x + h) &= -\sigma_{AB}(x, x + h) - \sigma_{AC}(x, x + h) \\
\sigma_B(x, x + h) &= -\sigma_{BA}(x, x + h) - \sigma_{BC}(x, x + h) \\
\sigma_C(x, x + h) &= -\sigma_{CA}(x, x + h) - \sigma_{CB}(x, x + h) \\
\sigma_{AB}(x, x + h) - \sigma_{BA}(x, x + h) &= \sigma_{BC}(x, x + h) - \sigma_{CB}(x, x + h) \\
&= \sigma_{CA}(x, x + h) - \sigma_{AC}(x, x + h)
\end{aligned} \tag{1.19}$$

## 1.2. TRANSFORMAÇÃO DE DADOS

Na geoestatística, algumas ferramentas requerem que a variável regionalizada em estudo, apresente certa distribuição dos dados, geralmente com distribuição normal ou Gaussiana. Nesse sentido, os dados amostrados nem sempre seguem uma distribuição normal. Desta forma, funções lineares ou não lineares são aplicadas na transformação ocorrendo mudanças na média e variância (Yamamoto & Landim, 2013, p. 58). Neste trabalho, foi utilizada a transformada Gaussiana.

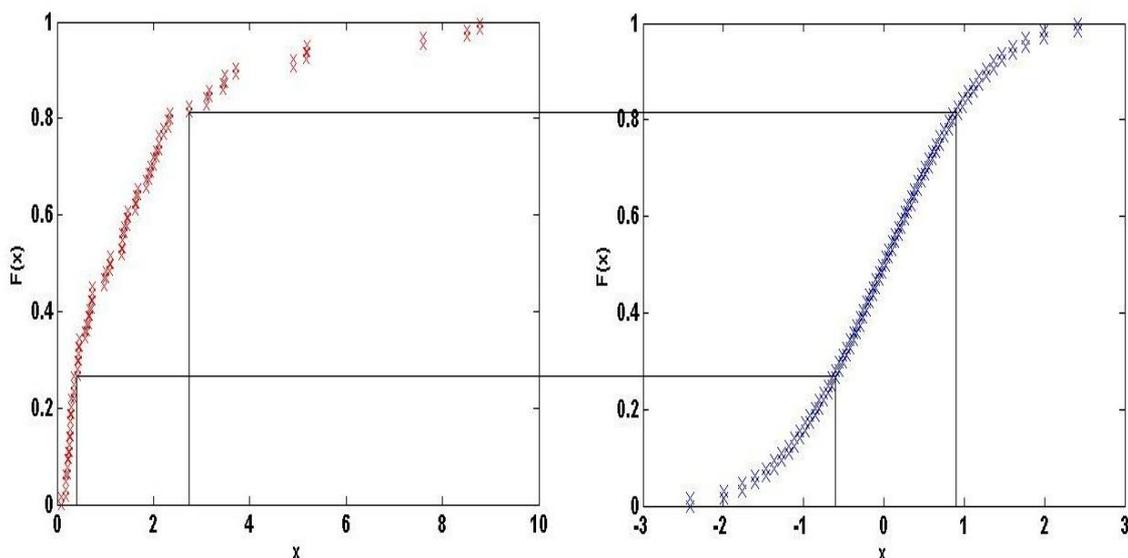
### 1.2.1. TRANSFORMADA GAUSSIANA

Segundo Yamamoto & Landim (2013, p. 58), a transformação dos dados mediante a transformada Gaussiana é baseada na curva teórica da distribuição de Gauss ou distribuição normal. Os mesmos autores resumem o procedimento da transformação em três passos a seguir:

1. Ordenam-se os n-valores de forma crescente, dos quais se geram as  $r(x_i)$  classes ordenadas;
2. Cada classe é dividida pelo número total de amostras mais um (n+1), desta forma obtém-se as proporções de cada classe;
3. Finalmente, para cada valor em (2) e aplicando a função inversa Gaussiana se calculam os escores da distribuição normal padrão:

$$y(x_i) = G^{-1}\left(\frac{r(x_i)}{n+1}\right) \tag{1.20}$$

A Figura 1.2 ilustra graficamente a transformação dos valores originais em uma distribuição normal.



**Figura 1.2.** Transformação dos dados originais para um valor padrão normal.

A distribuição normal tem algumas propriedades que fazem dela uma distribuição interessante e de grande importância na estimativa de modelos geoestatísticos. Por exemplo, a soma de duas ou mais variáveis aleatórias com distribuição normal é normal, resultando na simplicidade de simulação (Leuangthong & Deutsch, 2000, p. 3).

### 1.3. KRIGAGEM

Embora a krigagem tenha sido introduzida para proporcionar estimativas em pontos não amostrados, ela está sendo usada para construir modelos de probabilidade nos pontos não amostrados (Deutsch & Journel, 1992, p. 61). Segundo Yamamoto & Landim (2013, p. 55), a krigagem é um método geoestatístico de estimativa baseado na análise variográfica. Este método de estimativa é conhecido como um estimador linear não enviesado. Além disso, a krigagem é um conjunto de regressões lineares na minimização do erro da variância definido de um modelo apropriado de covariância (Deutsch & Journel, 1992, p. 14). Abrange uma família conhecida, entre outros, krigagem simples, krigagem ordinária etc. (Yamamoto & Landim, 2013, p. 55). Nesta dissertação será usada a krigagem simples.

### 1.3.1. KRIGAGEM SIMPLES

Seja  $x_0$  um valor não amostrado,  $Z^*(x_0)$  um estimador do ponto não amostrado dado pela combinação linear de variáveis aleatória definida como (Deutsch & Journel, 1992, p. 62):

$$Z_{KS}^*(x_0) = m_0 + \sum_{i=1}^n \lambda_i [Z(x_i) - m_i] \quad (1.21)$$

em que  $n$  é o número de pontos vizinhos,  $m_i$  as médias assumidas como conhecidas,  $m_0$  é a média no ponto a estimar e  $\lambda_i$  (com  $i = 1, \dots, n$ ) os ponderadores associados aos  $n$  dados. De acordo com Yamamoto & Landim (2013, p. 63), sob a hipótese estacionaridade de segunda ordem, a função aleatória  $Z$  a média e variância são constantes no domínio. Assim, o estimador da krigagem simples é computado (Olea, 1999, p. 11):

$$Z_{KS}^*(x_0) = m + \sum_{i=1}^n \lambda_i [Z(x_i) - m] \quad (1.22)$$

Porém, os ponderadores  $\lambda_i$  associados a cada ponto vizinho na estimativa de um valor não estimado são dados pela resolução de um sistema de equações matriciais, derivada da minimização do erro da variância mediante as derivadas parciais da variância do erro (Yamamoto & Landim, 2013, p. 64-65). Desta forma, o sistema é escrito como:

$$\begin{bmatrix} Cov(x_1, x_1) & Cov(x_1, x_2) & \cdots & Cov(x_1, x_n) \\ Cov(x_2, x_1) & Cov(x_2, x_2) & \vdots & Cov(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(x_n, x_1) & Cov(x_n, x_2) & \cdots & Cov(x_n, x_n) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} Cov(x_1, x_0) \\ Cov(x_2, x_0) \\ \vdots \\ Cov(x_n, x_0) \end{bmatrix} \quad (1.23)$$

## 1.4. EQUAÇÕES MULTIQUÁDRICAS

A krigagem indicadora seria o método natural para interpolação de funções indicadoras em pontos não amostrados. Entretanto, este método requer um número de modelos de variogramas igual ao número de tipos da variável categórica. Yamamoto et al. (2012, p. 147) afirmaram que a obtenção desses modelos de variogramas é muito difícil para tipos que apresentam poucos pontos de dados, os quais geram poucos pares de pontos com grandes variações estatísticas. Assim, esses autores sugeriram usar as equações multiquádricas, que não dependem dos modelos de variogramas, mas tão somente das funções de base radial.

As equações multiquádricas foram inicialmente introduzidas por Hardy (1971) como um método de interpolação global e logo estendido para interpolações locais, usando apenas os pontos vizinhos mais próximos (Yamamoto & Landim, 2013, p. 108-109). Mais tarde, em trabalhos posteriores as equações multiquádricas foram generalizadas e denominadas como funções de base radial, assim, uma função de base radial é a função variograma calculada e modelada com base nos pontos experimentais (Yamamoto & Landim, 2013, p. 109).

De acordo com Yamamoto (2002, p. 26), as funções de base radial foram desenvolvidas na interpolação de dados não amostrados num espaço 2D, mas podem ser estendidas para dados em 3D, bem como para maiores dimensões. A equação geral em 2D é dada por:

$$Z^*(x_0) = \sum_{i=1}^N C_i [(x_i - x_0)^2 + (y_i - y_0)^2 + C^2]^{\frac{1}{2}} \quad (1.24)$$

em que  $\{C_i = 1, \dots, N\}$  são os coeficientes da equação multiquádrica,  $C^2$  é uma constante positiva e  $N$  é o número total de pontos de dados. Assim, os coeficientes  $C_i$  são computados por meio de uma solução de um sistema de equações multiquádricas lineares, (Yamamoto & Landim, 2013, p. 110).

$$\begin{bmatrix} q_{11} & q_{12} & \dots & q_{1N} \\ q_{21} & q_{22} & \vdots & q_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ q_{N1} & q_{N2} & \dots & q_{NN} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_N \end{bmatrix} = \begin{bmatrix} Z(x_1) \\ Z(x_2) \\ \vdots \\ Z(x_N) \end{bmatrix} \quad (1.25)$$

A equação 1.24 em termos de funções de base radial pode ser escrita, segundo Yamamoto & Landim (2013, p. 30-31), como:

$$Z^*(x_0) = \sum_{i=1}^n C_i \phi(x_i, x_0) \quad (1.26)$$

onde a distância entre o ponto amostrado e o ponto a ser estimado é representado como  $(x_i, x_0)$ , e  $\phi$  é a função de base radial. Deste modo, o  $N$  (número de pontos de dados) é substituído por  $n$  (n pontos vizinhos mais próximos ao ponto a ser estimado). Assim, o método das equações multiquádricas é considerado como um método de interpolação local (Yamamoto & Landim, 2013, p. 111).

Yamamoto (2002, p. 31) resume as funções de base radiais mais usadas, como:

Linear	$\phi(x) =  x $
Cúbica	$\phi(x) =  x ^3$
Multiquádrica generalizada	$\phi(x) = (c +  x ^2)^{\frac{2k+1}{2}}$ para $k = -1, 0, \dots$
Spline	$\phi(x) =  x ^2 \log x $
Gaussiana	$\phi(x) = \exp(-c x ^2)$

em que  $|\cdot|$  é a norma de um vetor em  $R^n$  e  $c$  é uma constante positiva.

A equação 1.26 pode ser escrita de uma forma mais geral com uma constante  $a_0$  adicionada (Madych, 1992, apud Yamamoto & Landim, 2013, p. 111):

$$Z^*(x_0) = \sum_{i=1}^n C_i \phi(x_i, x_0) + a_0 \quad (1.27)$$

Yamamoto (2002, p. 30-32) demonstrou que a equação 1.27 pode ser escrita na forma dual como:

$$Z^*(x_0) = \sum_{i=1}^n W_i Z(x_i) \quad (1.28)$$

Desta forma, os pesos são obtidos da resolução linear de um sistema de equações e com a restrição que  $\sum_{i=1}^n W_i = 1$ , (Yamamoto & Landim, 2013, p. 32):

$$\begin{bmatrix} \phi(x_1, x_1) & \phi(x_1, x_2) & \dots & \phi(x_1, x_n) & 1 \\ \phi(x_2, x_1) & \phi(x_2, x_2) & \vdots & \phi(x_2, x_n) & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi(x_n, x_1) & \phi(x_n, x_2) & \vdots & \phi(x_n, x_n) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \\ \mu \end{bmatrix} = \begin{bmatrix} \phi(x_0, x_1) \\ \phi(x_0, x_2) \\ \vdots \\ \phi(x_0, x_n) \\ 1 \end{bmatrix} \quad (1.29)$$

em que  $\mu$  é o parâmetro de condição incluída de não viés no sistema de equações multiquádricas. Assim, há uma semelhança muito grande entre as funções de base radial e a krigagem ordinária (Yamamoto & Landim, 2013, p. 109).

### 1.4.1. ZONA DE INCERTEZA

Um valor não amostrado mediante as equações multiquádricas é estimado por meio da codificação binária de uma variável categórica de k-tipos derivada de uma função indicadora, como foi definida no item (1.1.1) (Yamamoto et al., 2012, p. 148):

$$I_{MQ}^*(x_0; k) = \sum_{i=1}^n w_i I(x_i; k) \quad (1.30)$$

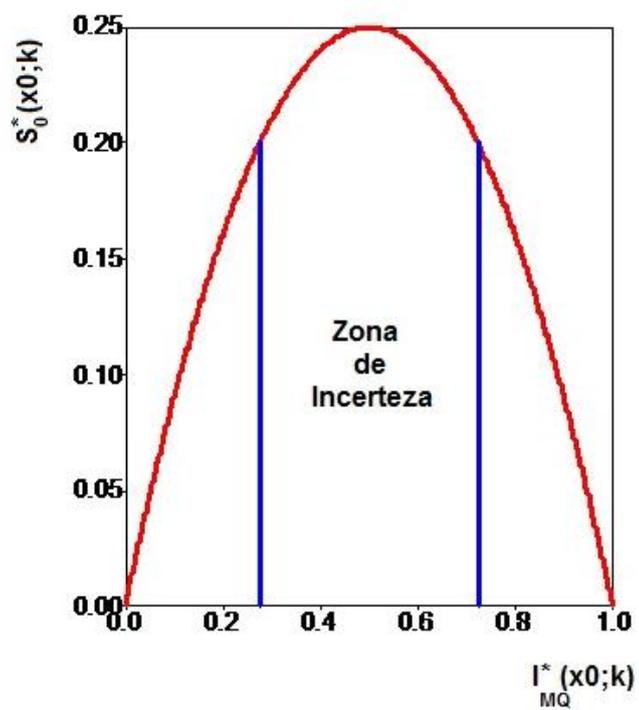
Desta forma, a variância associada ao ponto estimado de um tipo de variável categórica é igual à probabilidade de ser o k-ésimo tipo vezes a probabilidade do complemento, isto é (Yamamoto et al. 2012, p. 147):

$$s_0^2(x_0; k) = I_{MQ}^*(x_0; k)(1 - I_{MQ}^*(x_0; k)) \quad (1.31)$$

Segundo Yamamoto & Landim (2013, p. 115), uma vez calculadas as probabilidades, deve-se retornar o tipo categórico associado ao ponto estimado não amostrado, que é dado pelo valor mais provável, ou seja:

$$I_{MQ}^*(x_0; kmax) = \max(I_{MQ}^*(x_0; k), \text{com } k = 1, \dots, K) \quad (1.32)$$

Ao tipo mais provável associa-se uma incerteza dada pela variância (1.30). Se a variância for maior que 0,20, então o tipo categórico apresenta grande incerteza, ou seja, o ponto interpolado pertence à zona de incerteza, conforme Yamamoto et al. (2012, p. 151), como sendo a região definida por:  $I_{MQ}^*(x_0; kmax) < 0.6$  e  $s_0^2(x_0; kmax) > 0.20$ . O mesmo autor, corrige o intervalo da zona de incerteza de  $I_{MQ}^*(x_0; kmax)$  para 0.72. Na Figura 1.3 mostra-se a zona de incerteza da variância em função da indicadora interpolada, com delimitação da zona de incerteza (Yamamoto & Landim, 2013, p. 116).



**Figura 1.3.** Zona de incerteza definida no intervalo  $0.276 < I_{MQ}^*(x_0; kmax) < 0.724$ .

## CAPÍTULO 2

### 2.1. SIMULAÇÃO ESTOCÁSTICA

Métodos de interpolação de um atributo não amostrado, segundo Yamamoto & Landim (2013, p. 145), suavizam a variabilidade do atributo em estudo, ou seja, valores baixos são superestimados e valores altos subestimados. Nesse sentido, a suavização é inversamente proporcional à quantidade dos dados evidenciando-se duas características importantes (Olea, 1999, p. 141):

1. O histograma dos valores amostrados difere do histograma dos valores estimados,
2. O variograma experimental dos dados originais e o variograma das estimativas são diferentes.

Desta forma, pelos motivos anteriormente mencionados, a simulação estocástica foi adotada pela geoestatística para resolver parcialmente os problemas de suavização, mas não é uma solução perfeita dado que ganha-se em precisão global, mas perde-se em precisão local (Olea, 1999, p. 141).

De acordo com Deutsch & Journel (1992, p. 119), os métodos estocásticos geram realizações equiprováveis de uma variável  $\{Z(x), x \in A\}$  para  $l$  – realizações  $\{Z^l(x), x \in A\}$ . Porém, uma simulação é considerada condicional se as realizações honram os valores de dados referenciados em suas posições originais, isto é:

$$Z^*(x_i) = Z(x_i), \forall l$$

Yamamoto & Landim (2013, p. 147) recomendam analisar minuciosamente os resultados obtidos pelos métodos de simulação estocástica, dado que cada realização gera cenários distintos da realidade.

Métodos de simulação estocástica como o método de bandas rotativas, método de simulação sequencial Gaussiana, simulação direta, simulação indicadora sequencial, entre outros, são bem referenciados na literatura por vários autores (Yamamoto & Landim, 2013, p. 145-173; Olea, 1999, p. 141-158; Deutsch e Journel, 1992, p. 119-191). No presente trabalho, o método de bandas rotativas foi adotado na simulação pluriGaussiana truncada.

## 2.2. MAPEAMENTO DE INCERTEZA PARA VARIÁVEIS CATEGÓRICAS DERIVADA DE SIMULAÇÕES ESTOCÁSTICAS

Além das realizações individuais estarem sujeitas a incertezas, o conjunto das realizações também deverão estar afetados por incertezas (Yamamoto & Landim, 2013, p. 172). Nesse sentido, por meio de uma análise do conjunto de realizações é realizada para determinar o tipo mais provável, bem como a zona de incerteza, que se encontra entre os pontos de contato entre os k-tipos da variável categórica. Determina-se para cada pixel ou célula estimada a proporção do k-ésimo tipo para  $l$ -realizações da simulação estocástica como (Yamamoto & Landim, 2013, p. 172):

$$p(x_0; k) = \sum_{l=1}^l \frac{I^*(x_0; k)}{L} \quad (2.1)$$

Desta forma, computa-se o tipo mais provável como o valor de maior proporção (Yamamoto & Landim, 2013, p. 172), como:

$$p_{\max} = \max[p(x_0; k)] \text{ para } k = 1, \dots, K \quad (2.2)$$

Sendo a variância associada à proporção como:

$$s_k^2(x_0) = p_{\max}[1 - p_{\max}] \quad (2.3)$$

Onde o valor do tipo mais provável para cada pixel ou célula pode ser aceito se o valor estimado encontrar-se-á fora da zona de incerteza definida no item (1.4.1); caso contrário, o ponto estimado pertencente à zona de incerteza (Yamamoto & Landim, 2013, p. 172).

## 2.3. MÉTODO DE BANDAS ROTATIVAS

O método de bandas rotativas foi originalmente introduzido por Matheron (1973), permitindo a simulação de funções aleatórias multivariadas por meio de projeções aleatórias e na transformação do espaço (Olea, 1999, p. 158). Desta forma, a ideia básica do método de bandas rotativas é reduzir a dimensão do espaço atualmente simulado, obtendo-se realizações de funções multidimensionais geradas a partir de uma série de realizações independentes unidimensionais de funções aleatórias com uma estrutura de covariância adequada. De acordo com Emery & Lantuéjoul (2006, p. 1615), o método de bandas rotativas simplifica a simulação de um espaço  $R^d$  (geralmente,  $d = 2$  ou  $3$ ) a um espaço  $R^1$ .

Considere uma variável aleatória  $\{X(t), t \in R^d\}$ , com média zero e função de covariância contínua  $C_x$ , e um vetor aleatório  $U$  com distribuição uniforme sobre uma esfera unitária  $S_d$  em  $R^d$ . Então, para qualquer ponto  $x \in R^d$  (Emery & Lantuéjoul, 2006, p. 1615):

$$\forall x \in R^d, Z(x) = X(\langle x, U \rangle) \quad (2.4)$$

onde  $\langle, \rangle$  é o produto vetorial interno definido em  $R^d$ . O novo campo aleatório  $\{Z(x), x \in R^d\}$  com média zero e covariância isotrópica  $C_Z$ , (1) tal que:

$$C_Z(r) = \int_{S_d} C_X(\langle h, u \rangle) \bar{\omega}_d du \quad (2.5)$$

com  $h$  um vetor em  $R^d$ ,  $r$  o módulo e  $\bar{\omega}_d$  a distribuição uniforme sobre  $S_d$ .

Num caso isotrópico existe uma relação única entre a covariância contínua em  $R^d$  e a covariância contínua em  $R^1$ , segundo Matheron (1973, p. 461-462). Esta propriedade permite substituir a simulação de uma variável aleatória multidimensional com covariância  $C_Z$ , em uma simulação de um campo unidimensional com covariância  $C_X$  (Emery & Lantuéjoul, 2006, p. 1615). A expressão de  $C_X$  em função de  $C_Z$  quando  $d = 2$ , sua solução é integralmente mais complexa (Olea, 1999, p. 159). Desta forma, em um espaço tridimensional (com  $d = 3$ ) e usando coordenadas esféricas, a equação (2.5), tem:

$$C_X(r) = \frac{d}{dr} [r C_Z(r)] \quad (2.6)$$

onde, a covariância  $C_X$  é determinada e associada uma dada covariância isotrópica  $C_Z$  em  $R^3$ .

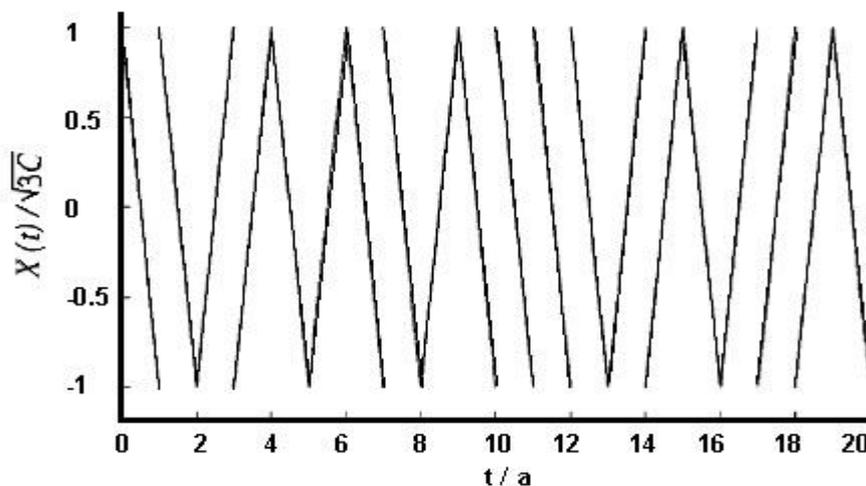
Emery & Lantuéjoul (2006, p. 1618) implementam quinze modelos de covariância em seu programa TBSIM. A seguir, a definição e os passos para obter uma simulação unidimensional são descritos para quatro dos quinze modelos de covariância; modelo de covariância esférico, exponencial, cúbico e gaussiano.

1. Modelo de covariância esférico com patamar  $C$  e amplitude  $a$ :

$$C_Z(r) = C \left( 1 - \frac{3r}{2a} + \frac{r^3}{2a^3} \right) \quad 1_r \leq a \quad (2.7)$$

A variável aleatória unidimensional associada a  $\{X(t), t \in R\}$  é simulada, como (Emery & Lantuéjoul, 2006, p. 1615):

- Sobre o eixo real se fixa um valor aleatório (uniformemente em  $[0,a)$ ) e se divide o eixo em intervalos igualmente espaçados de tamanho  $a$ ;
- Para cada intervalo, uma função linear é traçada com igual probabilidade de ser crescente ou decrescente (Figura 2.1).



**Figura 2.1** . Realização unidimensional de um campo aleatório associado a um modelo de covariância em 3D.

2. Modelo de covariância exponencial com patamar  $C$  teórico e amplitude  $a$ :

$$C_Z(r) = C \exp\left(-\frac{r}{a}\right) \quad (2.8)$$

A simulação de uma variável aleatória unidimensional é realizada por um processo pontual de Poisson em  $R$  (Emery & Lantuéjoul, 2006, p. 1618). Desta forma, o modelo de covariância exponencial é uma mistura de amplitudes de modelos esféricos. Onde:

$$C_Z(r) = \int_r^{+\infty} C \left(1 - \frac{3r}{2u} + \frac{r^3}{2u^2}\right) f(u) du \quad (2.9)$$

com  $f(u)$  uma função de probabilidade em  $R$  a ser determinada. Assim, diferenciando-se com respeito a  $r$  e simplificando a equação (2.7), temos:

$$\forall u > 0, f(u) = \frac{u}{3a^2} \exp\left(-\frac{u}{a}\right) \left(\frac{u}{a} + 1\right) \quad (2.10)$$

De (2.10), a função densidade de probabilidade é derivada de uma variável aleatória Gamma com parâmetros  $(3, a)$  e probabilidades  $\frac{1}{3}$  e  $(2, a)$  com probabilidade  $\frac{1}{3}$ .

Deste modo, a simulação unidimensional de um modelo de covariância exponencial pode ser realizada, por:

- Simular uma variável aleatória  $G$ ;
- Um vetor  $U$  com distribuição aleatória uniforme sobre a esfera  $S_3$  é simulado;
- Para cada componente de  $U$  ou vetor direção é dividido em intervalos de tamanhos  $G$ . Assim, para cada intervalo é atribuído uma função linear aleatória positiva ou negativa, com igual probabilidade (Figura 2.1).

3. Modelo de covariância cúbico com patamar  $C$  e amplitude  $a$ :

$$C_z(r) = C \left(1 - 7 \frac{r^2}{a^2} + \frac{35 r^3}{4 a^3} - \frac{7 r^5}{2 a^5} + \frac{3 r^7}{4 a^7}\right) \quad 1_r \leq a \quad (2.11)$$

A simulação de uma variável aleatória em um espaço unidimensional é obtida de forma similar ao modelo de covariância esférico, mas com uma função cúbica invés de uma linear para cada intervalo.

4. Modelo de covariância gaussiano com patamar  $C$  e amplitude  $a$ :

$$C_z(r) = C \exp\left(-\left(\frac{r}{a}\right)^2\right) \quad (2.12)$$

Este modelo de covariância pode ser simulado aplicando o método espectral contínuo, o qual constitui um caso particular do método de bandas rotativas (Emery & Lantuéjoul, 2006, p. 1619):

$$\forall x \in R^3, \quad Z(x) = \sqrt{2C} \cos(2\pi \langle x, T \rangle + \phi) \quad (2.13)$$

onde  $\phi$  é uma fase uniforme no intervalo  $[0, 2\pi)$  e  $T$  uma frequência distribuída aleatoriamente com média espectral de  $C_Z$ . Portanto,  $T$  é um vetor com distribuição Gaussiana em  $R^3$ .

### 2.2.1. ERGODICIDADE E NORMALIDADE MULTIVARIADA

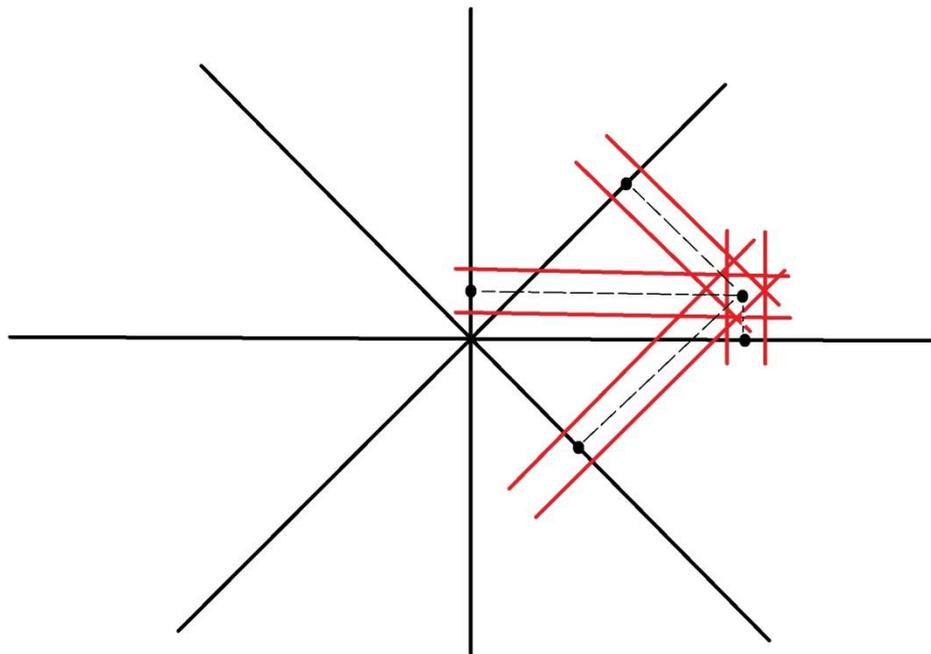
Segundo Chiles e Delfiner (2006, apud Emery e Lantuéjoul, p. 1620), resultados de não-ergodicidade são obtidos por meio da equação (2.4), dado que cada realização exibe anisotropia zonal orientada ao longo do vetor  $U$ , de modo que a covariância experimental não pode ser associada a um modelo teórico isotrópico. Este problema é resolvido por meio de somas independentes de variáveis aleatórias mediante projeções de pontos sobre as linhas orientadas em  $R^3$  (Figura 2.2); assim, uma variável simulada para  $N$  realizações é dada por:

$$Z(x) = \frac{1}{\sqrt{N}} \sum_{i=1}^N X_i(\langle x, U_i \rangle) \quad (2.14)$$

onde  $\{U_i, i = 1, \dots, N\}$  é um vetor independente de direções distribuídas aleatoriamente sobre a esfera  $S_3$  e  $\{X_i, i = 1, \dots, N\}$  são as variáveis aleatórias independentes unidimensionais, com sua respectiva covariância  $C_X$  dada por equação (2.6). Assim, a covariância de  $\{Z(x), x \in R^3\}$  dada cada realização para um vetor de separação  $h$  e modulo  $r$  é:

$$C_Z(r) = \frac{1}{N} \sum_{i=1}^N C_X(\langle h, u_i \rangle) \quad (2.15)$$

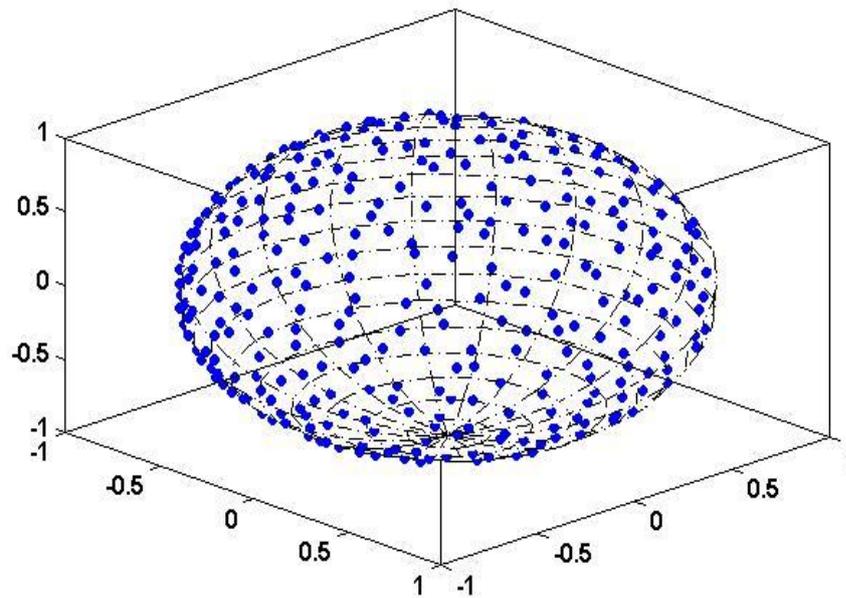
Desta forma, a equação anterior constitui uma aproximação discreta do modelo teórico de covariância em (2.5), sendo reproduzido experimentalmente a integração por simulação de monte carlos, (Emery & Lantuéjoul, 2006, p. 1621).



**Figura 2.2.** Ilustração do método de bandas rotativas em duas dimensões. O ponto sobre o plano é projetado em cada linha (quatro linhas) e influenciado o valor simulado. Fonte: (Olea, 1999, p. 161).

### 2.2.2. DISTRIBUIÇÃO DAS LINHAS SOBRE A ESFERA

Inicialmente, assumimos que o vetor aleatório  $U$  está distribuído uniformemente sobre a esfera  $S_d$  (com  $d = 3$ ) dado que (2.15) é uma aproximação de (2.5), a convergência na prática é dada rapidamente quando as linhas estão regularmente distribuídas sobre a esfera. Segundo Chiles & Delfiner (2006, apud Emery & Lantuéjoul, p. 1621), a distribuição das linhas não necessariamente deve ser regular sobre a esfera, assumindo que o vetor  $U$  é um vetor aleatoriamente distribuído sobre a esfera. Freulon & de Fouquet (1991) examinaram várias sequências para uma série de direções de linhas e mostraram que a propriedade de ergodicidade melhora com a escolha do método de sequências de Van der Corput (Figura 2.3) em lugar de direções uniformes. Desta forma, Emery & Lantuéjoul (2006, p. 1622) programaram uma sub-rotina chamada VDC que define uma rotação aleatória das direções e evita repetir o mesmo conjunto para cada realização.



**Figura 2.3.** Conjunto de 300 direções distribuídas aleatoriamente sobre uma esfera unitária em  $R^3$  (sequência de van der Corput).

### 2.2.3. NÚMERO DE LINHAS USADAS NA SIMULAÇÃO DE CADA ESTRUTURA

Teoricamente é requerido um número infinito de linhas na aproximação de um modelo de covariância, mas na prática alguns autores recomendam variar o número de linhas, dependendo da dimensão do espaço a ser simulado (Olea, 1999, p. 159). Em duas dimensões, Mantoglou & Wilsonm (2006, apud, Emery & Lantuéjoul, p. 1622) sugerem usar 4 a 16 linhas, na simulação de um modelo de covariância. Em três dimensões, Deutsch e Journel (1992, p. 144) usam 15 linhas, mas é uma aproximação insuficiente e várias linhas deveriam ser usadas (Emery & Lantuéjoul, 2006, p. 1622). Desta forma, na aproximação de um modelo de covariância o valor de  $N = 1000$  linhas será considerado neste trabalho.

### 2.2.4. REALIZAÇÕES CONDICIONAIS DE VALORES GAUSSIANOS

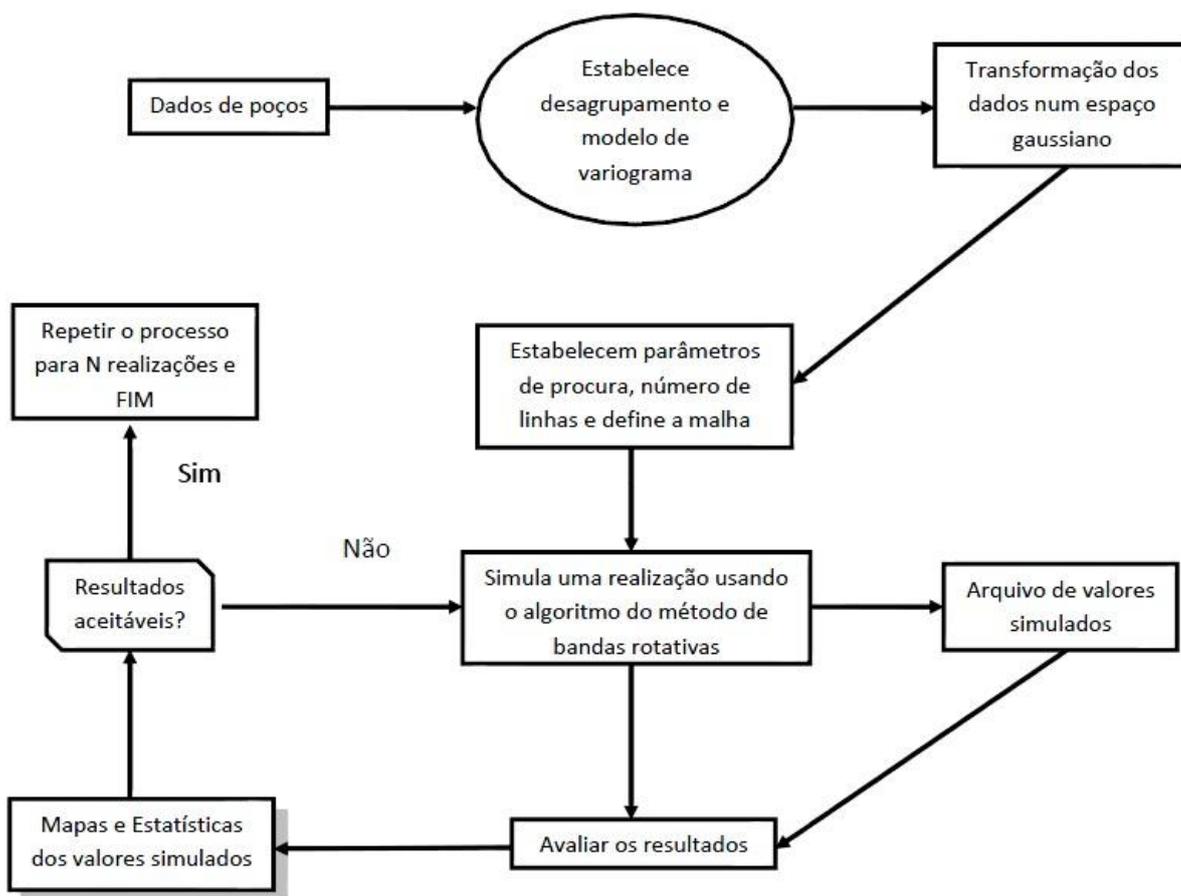
O condicionamento de realizações não condicionais é feito através da krigagem (Emery & Lantuéjoul, 2006, p. 1622). Seja  $\{Z_G(x), x \in R^3\}$  uma simulação não-condicional, então a realização condicional, no caso estacionário, pode ser:

$$\forall x \in R^3, Z_{CS}(x) = Z_S(x) + [Z(x) - Z_S(x)]^{SK} \quad (2.16)$$

O super-índice SK da equação (2.16) é a krigagem simples (com média zero e covariância  $C_Z$ ) de valores dos dados coordenados. De acordo com Emery & Lantuéjoul (2006, p. 1622), uma simulação condicional pode ser construída em quatro passos:

1. Realizações não condicionais via método de bandas rotativas são geradas para cada ponto amostrado  $x$  e cada ponto a ser estimado sobre o grid;
2. Residuais são gerados dos dados amostrados e os dados simulados condicionais coordenados;
3. A krigagem simples é aplicada aos residuais;
4. O resultado é adicionado às realizações não condicional dado por (2.16).

O resumo do método descrito anteriormente encontra-se na figura 2.4.

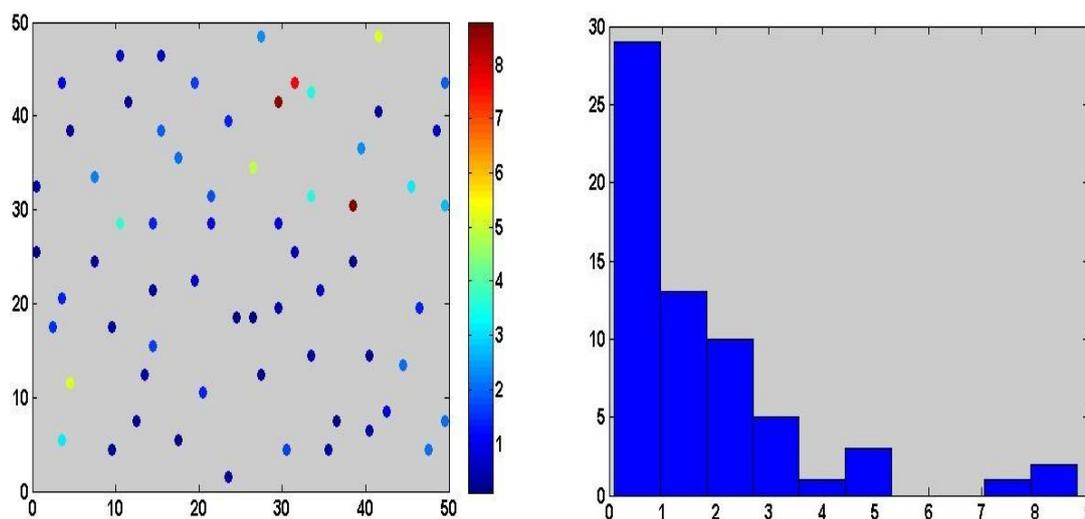


**Figura 2.4.** Fluxograma do método de bandas rotativas condicional de uma função aleatória estacionária.

## 2.4. EXEMPLO DE APLICAÇÃO DO MÉTODO DE BANDAS ROTATIVAS

Para ilustrar o procedimento de aplicação do método de bandas rotativas, foi escolhido um conjunto de dados que apresenta uma distribuição com assimetria positiva com 64 pontos amostrados sobre área (Figura 2.5).

As estatísticas descritivas para a amostra encontram-se resumidas na Tabela 2.1.



**Figura 2.5.** Mapa de localização de pontos da amostra (esquerda) e o histograma dos dados (direita).

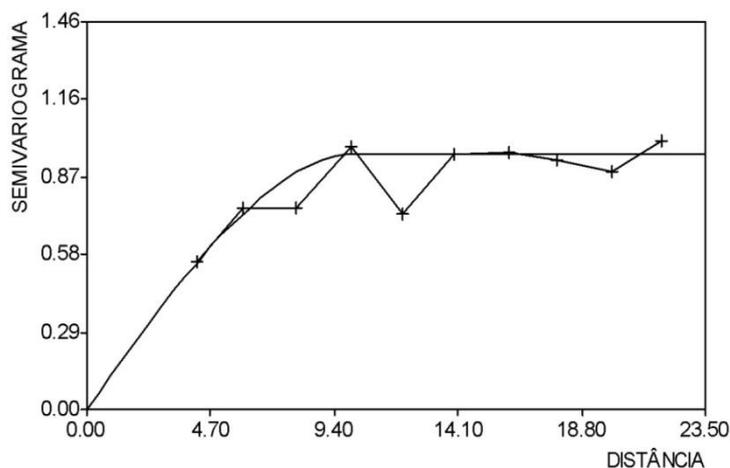
**Tabela 2. 1.** Estatísticas descritivas para a amostra em estudo.

Estatísticas	Amostra
N	64
Média	1,708
Desvio padrão	1,923
Coef. Variação	1,126
Máximo	8,781
Mínimo	0,095
Mediana	1,089

A primeira etapa consiste na transformação da distribuição dos dados a um domínio gaussiano por meio da transformada Gaussiana (Figura 1.2). Desta forma, o novo conjunto de dados fica com média zero e variância igual a um.

Prosseguindo, o segundo passo é obter um modelo de variograma adequado para os valores de dados transformados para aplicar o método de bandas rotativas. O variograma foi calculado com tolerância angular de 90°, ou seja, omnidirecional dado que não existe

diferença alguma da variação espacial em qualquer direção. O resultado encontra-se na Figura 2.6.



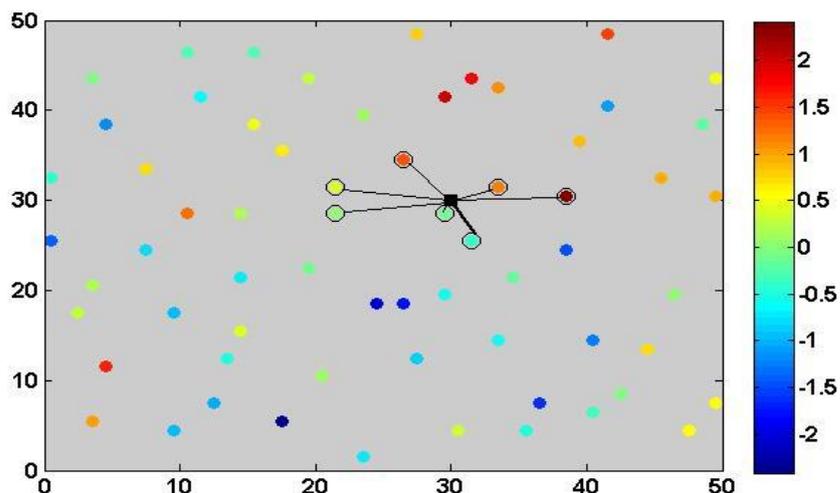
**Figura 2.6.** Variograma experimental ajustado a um modelo teórico exponencial.

O modelo ajustado ao variograma experimental está descrito a seguir:

$$\gamma(h) = 1 * \left( 1,5 \left( \frac{h}{10,32} \right) - 0,5 \left( \frac{h}{10,32} \right)^3 \right) \text{ para } h < 10,32$$

$$\gamma(h) = 1 \text{ para } h \geq 10,32$$

A seguir, o terceiro passo consiste em gerar realizações condicionais dos dados gaussianos via bandas rotativas. Será considerado para o acompanhamento prático da simulação por bandas rotativas o ponto de coordenadas  $(x = 30; y = 30; z = 0)$  do grid. Para esse ponto foram encontrados sete pontos vizinhos pelo critério dos quadrantes (dois pontos mais próximos por quadrantes), como mostra a Figura 2.7 e a Tabela 2.2.



**Figura 2.7.** Localização dos vizinhos próximos ao ponto de coordenadas  $(x = 30; y = 30; z = 0)$ .

**Tabela 2. 2.** Pontos de dados vizinhos ao ponto a ser estimado  $(x = 30; y = 30; z = 0)$ .

Ponto	X	Y	Z	Valor
1	26,50	34,50	0	1.3662
2	21,50	31,50	0	0.4236
3	33,50	31,50	0	1.1892
4	38,50	30,50	0	2.4176
5	29,50	28,50	0	-0.0981
6	21,50	28,50	0	0.0196
7	31,50	25,50	0	-0.3811

Por meio do método de bandas rotativas, quatro linhas foram geradas ( $L_1, L_2, L_3$  e  $L_4$ ) distribuídas aleatoriamente sobre a esfera  $S_3$  (seqüência de Van der Corput) (Figura 2.8). Porém, cada linha é rotacionada:

$$lines = L_i * R'$$

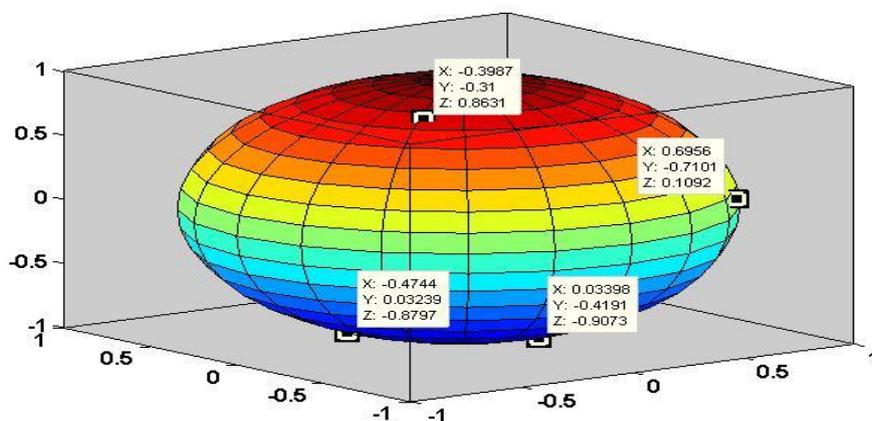
com R a matriz de rotação;

$$R = 1.0e15 * \begin{bmatrix} 0.0 & -0.0 & 0 \\ 0.0 & 0.0 & 0 \\ 0 & 0 & 4.5036 \end{bmatrix}$$

O resultado encontra-se na tabela 2.3.

**Tabela 2. 3.** Direção das linhas na direção de anisotropia.

$u_1 = \text{Lines } 1^*R$	$u_2 = \text{Lines } 2^*R$	$u_3 = \text{Lines } 3^*R$	$u_4 = \text{Lines } 4^*R$
(0 0 4.9170e + 014)	(0 0 -4.0860e + 015)	(0 0 3.8870e + 015)	(0 0 -3.9618e + 015)

**Figura 2.8.** Um conjunto de quatro direções distribuídas sobre uma esfera unitária em  $R^3$  (sequência de Van der Corput).

O próximo passo é projetar perpendicularmente cada ponto da tabela 2.2 sobre as quatro linhas fazendo o uso do produto vetorial em  $R^3$  ( $\langle x, u_i \rangle$  com  $i = 1, 2, 3$  e  $4$ ). O objetivo é gerar uma realização de um campo aleatório  $\{X(t), t \in R\}$  associado a um modelo de covariância unidimensional. Neste caso, um modelo esférico como foi descrito anteriormente. O resultado encontra-se na tabela 2.4.

**Tabela 2. 4.** Projeção de cada ponto sobre a linha em sua respectiva direção.

Ponto	$\langle x, u_1 \rangle$	$\langle x, u_2 \rangle$	$\langle x, u_3 \rangle$	$\langle x, u_4 \rangle$
1	4,2815	1,2280	-0,5540	-1,7227
2	3,7178	1,0082	-0,5894	-1,5641
3	4,5699	1,5112	-0,2173	-1,6030
4	4,8554	1,7173	-0,0225	-1,5718
5	4,0772	1,3333	-0,2217	-1,4477
6	3,5091	0,9980	-0,4698	-1,4218
7	4,0106	1,4070	-0,0401	-1,3119

Dado que o modelo associado ao variograma experimental é esférico, um campo aleatório unidimensional é simulado em duas etapas (equação 2.7). A primeira etapa consiste na construção de um intervalo aleatório no eixo real e dividir o eixo em intervalos equidistantes. Desta forma um valor fixo para cada linha é designado (Tabela 2.5).

**Tabela 2. 5.** Valor aleatório fixo sobre o eixo real.

<b>Valor fixo (offset) para cada linha gerada</b>			
<b>Offset 1</b>	<b>Offset 2</b>	<b>Offset 3</b>	<b>Offset 4</b>
-0,7397	-0,2859	-2,7508	-3,1503

Cada ponto projetado sobre a linha é associado a um intervalo  $[0, a)$ :

$$intervalo = floor(x - offset(i)) + 1 \text{ com } i = 1, 2, 3 \text{ e } 4$$

onde *floor* é uma função de aproximação ao valor menos infinito. Assim, foram definidos oito intervalos para a linha um, três intervalos para a linha dois, cinco intervalos para a linha três e quatro intervalos para a linha quatro. Na tabela 2.6 mostra-se os resultados dos intervalos associados aos pontos da tabela 2.2.

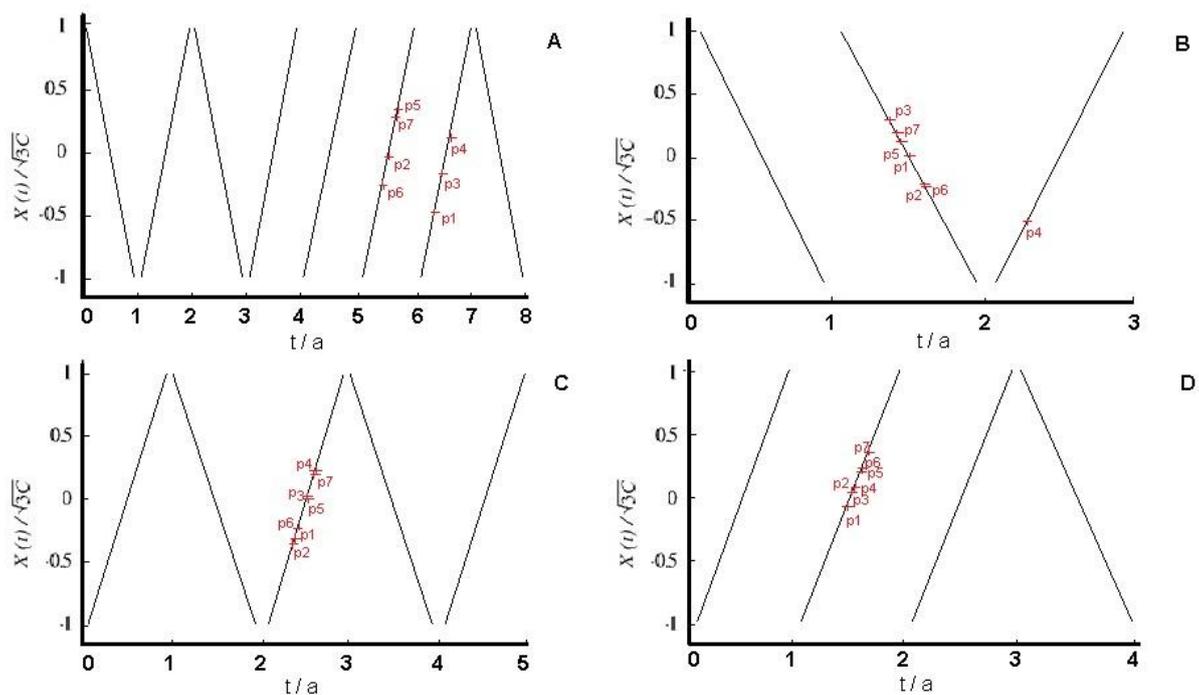
**Tabela 2. 6.** Intervalos associados a cada ponto sobre as linhas.

<b>Intervalo /ponto</b>	<b><math>(x - offset(1)) + 1</math></b>	<b><math>(x - offset(2)) + 1</math></b>	<b><math>(x - offset(3)) + 1</math></b>	<b><math>(x - offset(4)) + 1</math></b>
1	6	2	3	2
2	5	2	3	2
3	6	2	3	2
4	6	3	3	2
5	5	2	3	2
6	5	2	3	2
7	5	2	3	2

Na segunda etapa se gera um sinal independente de um intervalo a outro (+1 ou -1), desenhando uma função linear com igual probabilidade de ser crescente ou decrescente, para cada intervalo em cada linha definida do passo anterior e cada ponto é posicionado por

$$posição = x - offset - intervalo + 0.5$$

como é visto na Figura 2.9 e Tabela 2.7.



**Figura 2.9.** Realização unidimensional de um campo aleatório associado a um modelo de covariância esférico em 3D. Cada linha gerada é discretizada num número de intervalos e para cada intervalo um sinal (-1 ou +1) é gerado aleatoriamente.

**Tabela 2. 7.** Posição do ponto sobre a função independente no intervalo  $[-1, 1]$ .

Ponto	Posição 1	Posição 2	Posição 3	Posição 4
1	-0,4788	0,0139	-0,3031	-0,0724
2	-0,0425	-0,2059	-0,3385	0,0861
3	-0,1904	0,2971	0,0335	0,0473
4	0,0951	-0,4967	0,2284	0,0785
5	0,3169	0,1193	0,0291	0,2026
6	-0,2512	-0,2161	-0,2189	0,2285
7	0,2502	0,1929	0,2107	0,3384

O valor simulado do campo aleatório é dado pelo produto entre os valores da tabela 2.7 e o sinal a onde pertence cada ponto da Figura 2.9 (Tabela 2.8).

**Tabela 2. 8.** Valor simulado associado a um modelo de covariância esférico.

Campo 1	Campo 2	Campo 3	Campo 4
-0.4788	-0.0139	-0.3031	-0.0724
-0.0425	0.2059	-0.3385	0.0861
-0.1904	-0.2971	0.0335	0.0473
0.0951	-0.4967	0.2284	0.0785
0.3169	-0.1193	0.0291	0.2026
-0.2512	0.2161	-0.2189	0.2285
0.2502	-0.1929	0.2107	0.3384

Finalmente, a variável simulada é dada pela soma das realizações independentes. Desta forma, cada linha contribui no ponto a ser simulado dado pela equação (2.14):

$$Z(x) = \frac{1}{\sqrt{4}} \sum_{i=1}^4 X_i(< x, U_i >)$$

com  $N = 4$  linhas. Além disso, o resultado dado de  $Z(x)$  é multiplicado por uma constante  $const = 2 * \sqrt{3C} = 3,4641$ . Na tabela 2.9 encontram-se os valores da tabela 2.8 vezes a constante e o valor simulado não condicional para cada ponto vizinho do ponto a estimar.

**Tabela 2. 9.** Valor de cada campo, constante vezes o campo e o valor simulado não condicional.

Const* campo1	Const* campo2	Const* campo3	Const* campo4	$soma = const * \sum_{i=1}^4 X_i(< x, U_i >)$	$Z(x) = \frac{1}{\sqrt{4}} * soma$
-1.6586	-0.0482	-1.0500	-0.2508	-3.0076	-1.5038
-0.1472	0.7133	-1.1726	0.2983	-0.3082	-0.1542
-0.6596	-1.0292	0.1160	0.1639	-1.4089	-0.7045
0.3294	-1.7206	0.7912	0.2719	-0.3281	-0.1641
1.0978	-0.4133	0.1008	0.7018	1.4871	0.7435
-0.8702	0.7486	-0.7583	0.7915	-0.0884	-0.0443
0.8667	-0.6682	0.7299	1.1723	2.1007	1.0504

Da mesma maneira, para simular o ponto com coordenadas  $(x = 30; y = 30; z = 0)$ , por meio do método de bandas rotativas, uma simulação não condicional deve ser feita. Assim, o primeiro passo é projetar o ponto sobre cada linha (Tabela 2.10). Os valores fixos para dividir cada linha são os mesmos da tabela 2.5 e, desta forma, o ponto associado a um intervalo (5, 2, 3 e 2) é posicionado (tabela 2.11, Figura 2.9):

**Tabela 2. 10.** Ponto a simular no grid projetado sobre as quatro linhas.

Ponto	$\langle x, u_1 \rangle$	$\langle x, u_2 \rangle$	$\langle x, u_3 \rangle$	$\langle x, u_4 \rangle$
1	4,2171	1,3594	-0,2660	-1,5206

**Tabela 2. 11.** Posição do ponto a simular na grid sobre a função independente no intervalo  $(-1, 1)$ .

Ponto	Posição 1	Posição 2	Posição 3	Posição 4
1	0,4567	0,1453	-0,0152	0,1298

Onde o valor não condicional do ponto a estimar é gerado multiplicando os valores da tabela 2.11 pelo sinal atribuído a cada intervalo (Figura 2.9) e da soma independentes de cada campo aleatório (equação 2.14). Além disso, o valor é multiplicado pela constante  $const = 3,4641$ . Portanto, o valor simulado não condicional do ponto a estimar sobre o grid é:

$$Z(x) = \frac{1}{\sqrt{4}} * ((0.4567 * 1) + (0.1453 * -1) + (-0.0151 * 1) + (0.1298 * 1)) * 3.4641$$

$$= 0,7378$$

A seguir, deve-se calcular os residuais dos pontos amostrados. Nesse sentido, os residuais são calculados para os pontos amostrados vizinhos da coordenadas  $(x = 30; y = 30; z = 0)$ . Deste modo, o valor residual é dado por:

$$residual = x - Z(x)$$

Para efetuar o cálculo dos residuais (pontos vizinhos amostrados), os valores da tabela 2.2 para cada ponto e os valores não condicionais da tabela 2.9 são considerados. O resultado é mostrado na tabela 2.12.

**Tabela 2. 12.** Valores residuais computados dos pontos amostrados e o valor não condicional.

Ponto	X	Y	Z	Valor	Valor não condicional	Residual
1	26,50	34,50	0	1.3662	-1.5038	2.8700
2	21,50	31,50	0	0.4236	-0.1542	0.5778
3	33,50	31,50	0	1.1892	-0.7045	1.8937
4	38,50	30,50	0	2.4176	-0.1641	2.5817
5	29,50	28,50	0	-0.0981	0.7435	-0.8416
6	21,50	28,50	0	0.0196	-0.0443	0.0639
7	31,50	25,50	0	-0.3811	1.0504	-1.4315

Prosseguindo os cálculos, o seguinte passo é realizar a krigagem simples dos valores residuais dos dados coordenados próximos ao ponto a estimar. Assim, monta-se o sistema de equações de krigagem simples, conforme a equação (1.23):

$$\begin{bmatrix} 1.0000 & 0.2245 & 0.0785 & 0 & 0.1447 & 0.0667 & 0 \\ 0.2245 & 1.0000 & 0 & 0 & 0.0303 & 0.5635 & 0 \\ 0.0785 & 0 & 1.0000 & 0.3014 & 0.3125 & 0 & 0.1778 \\ 0 & 0 & 0.3014 & 1.0000 & 0.0089 & 0 & 0.0279 \\ 0.1447 & 0.0303 & 0.3125 & 0.0089 & 1.0000 & 0.0560 & 0.4826 \\ 0.0667 & 0.5635 & 0 & 0 & 0.0560 & 1.0000 & 0 \\ 0 & 0 & 0.1778 & 0.0279 & 0.4826 & 0 & 1.0000 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \\ \lambda_6 \\ \lambda_7 \end{bmatrix} = \begin{bmatrix} 0.2375 \\ 0.0268 \\ 0.4564 \\ 0.0315 \\ 0.7648 \\ 0.0268 \\ 0.3419 \end{bmatrix}$$

Resolvendo o sistema, os ponderadores da krigagem simples são obtidos:

$$\begin{aligned} \lambda_1 &= 0.1226; \lambda_2 = -0.0151; \lambda_3 = 0.2536; \lambda_4 = -0.0502; \lambda_5 = 0.6850; \lambda_6 \\ &= -0.0112; \lambda_7 = -0.0324 \end{aligned}$$

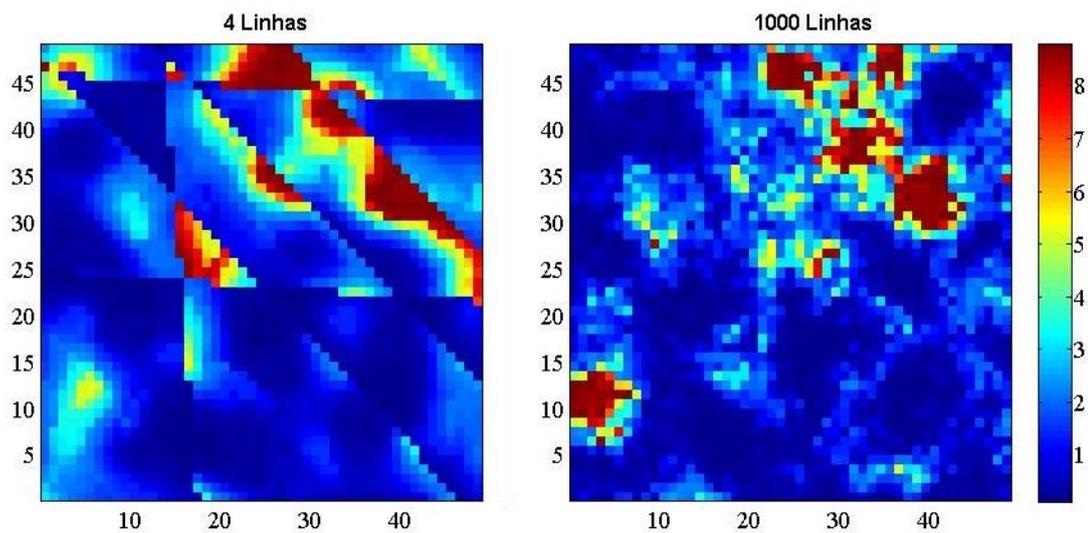
Aplicando-se os ponderadores na estimativa do ponto de coordenadas ( $x = 30; y = 30; z = 0$ ) na equação (2.16), obtém-se o valor condicional:

$$\begin{aligned} Z_{CS}(x) &= 0.7378 \\ &+ [(0.1226 * 2.8700) + (-0.0151 * 0.5778) + (0.2536 * 1.8937) \\ &+ (-0.0502 * 2.5817) + (0.6850 * -0.8416) + (-0.0112 * 0.0639) \\ &+ (-0.0324 * -1.4315)] = 0.9007 \end{aligned}$$

O valor  $Z_{CS}(x) = 0.9007$  deve ser transformado ao domínio original mediante uma função inversa. Assim, o valor estimado na escala original é:

$$Z_{CS}(x) = 2.6346$$

Finalmente, a Figura 2.10 mostra-se uma só realização para as quatro linhas geradas no começo do exemplo. Desta forma, o uso de quatro linhas ( $N = 4$ ) claramente evidencia uma deficiência na aproximação do modelo, com a presença de artefatos. Porém, tais artefatos não são perceptíveis quando o número de linhas é aumentado para mil linhas ( $N = 1000$ ).



**Figura 2.10.** Duas realizações de um campo aleatório simulado via bandas rotativas usando 4 e 1000 linhas, respectivamente. O modelo de covariância consiste em um modelo esférico isotrópico sem efeito pepita.

## CAPÍTULO 3

### 3.1. SIMULAÇÃO PLURIGAUSSIAN TRUNCADA

Segundo Armstrong et al. (2003) a simulação plurigaussiana é realizada em quatro passos: tipo de modelo, estimativa de valores parâmetros, gerar valores gaussianos para cada ponto amostrado e finalmente simular valores condicionados.

### 3.2. CURVAS DE PROPORÇÃO VERTICAL E HORIZONTAL

Na década de 1980 foram introduzidas as curvas de proporção que auxilia na quantificação dos dados geológicos em percentagem do espaço ocupado por todas as fácies, sendo de grande importância na interpretação de uma sequência estratigráfica e de correlações (Ravenne, 2002, p. 318).

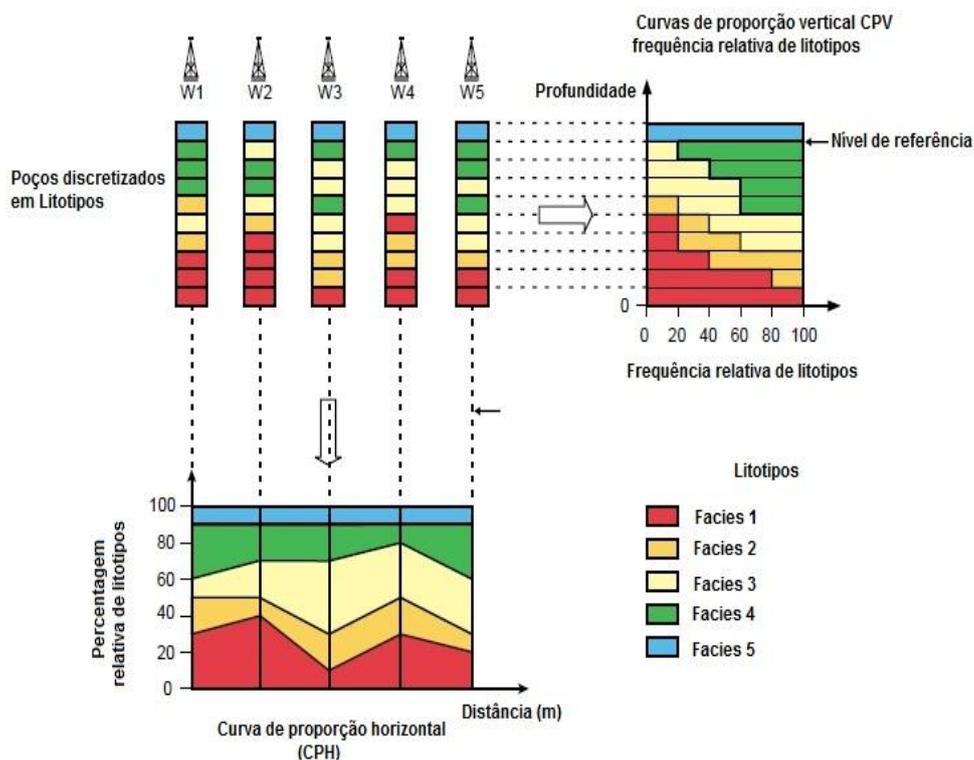
Segundo Armstrong *et al.* (2003, p. 27), as curvas de proporção são calculadas dos dados experimentais sendo o primeiro passo na simulação Gaussiana ou pluriGaussiana truncada. Assim, as curvas de proporção são uma visualização gráfica da distribuição espacial das fácies representadas em um histograma de fácies, calculadas usualmente na direção vertical e horizontal. No caso estacionário, onde as proporções são constantes, estas são calculadas da percentagem de todos os dados pertencentes a todas as fácies (Armstrong *et al.* 2003, p. 27).

Nesse sentido, as curvas de proporção fornecem informações do ambiente de deposição e por meio delas pode-se definir intervalos de fácies dominantes e informações que ajudam na modelagem do reservatório (Ravenne, 2002, p. 319).

Segundo Ravenne (2002, p. 318) e Braga (1998, p. 44), as curvas de proporções verticais e horizontais obtêm-se de uma sequência de deposição, dependendo de uma ordem das fácies (1, 2, ..., *etc.*) e mediante uma escolha de um nível de referência. As curvas de proporção vertical (CPV) são calculadas a partir da média calculada nível a nível (ou plano a plano) paralela a um nível de referência (datum), preferencialmente um marco cronoestratigráfico. As curvas de proporção são representadas graficamente por um

histograma acumulado, onde no eixo horizontal se mostram as proporções de todas as fácies em função da profundidade.

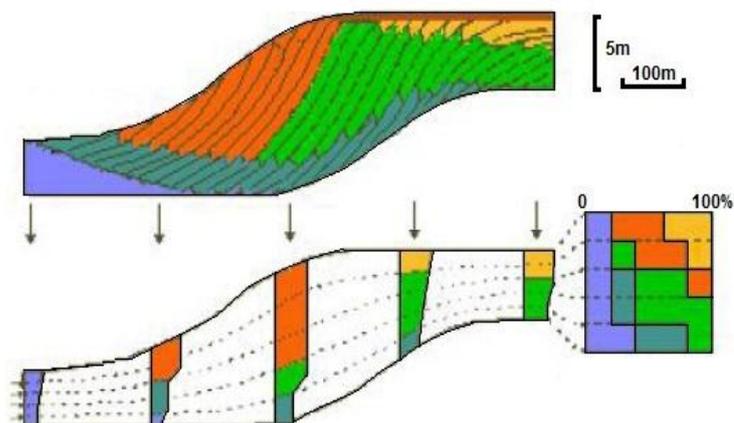
As curvas de proporção horizontal (CPH) são uma projeção sobre o eixo horizontal da proporção de cada fácies poço a poço, apresentando-se as percentagem das fácies numa determinada direção. Desta forma, estas curvas exibem a evolução lateral das fácies.



**Figura 3.1.** Curvas de proporção vertical e horizontal (Fonte: Ravenne, 2002).

A Figura 3.1 ilustra o cálculo de proporções nas direções vertical e horizontal mediante cinco poços fictícios discretizados em dez intervalos a partir de um nível de referência (topo) e um ordenamento das mesmas. Neste sentido, a estimativa das proporções é feita desde a primeira camada que tem 100% da fácies 5, e assim sucessivamente continuando com a ordem das fácies até a fácies 1, verificando se que existe uma variação na distribuição das litologias. No cálculo das curvas de proporção horizontal, não há nenhum posicionamento preferencial das fácies, considerando-se a distribuição de fácies como estacionárias. Em outras áreas e como se aprecia na Figura 3.2, a distribuição lateral das fácies exibem mudanças na direção horizontal.

Nesse caso, na presença de variações laterais da distribuição das fácies, ou seja, mudanças laterais das proporções poço a poço (domínio não estacionário), deve-se construir uma matriz 3D de curvas de proporções (Armstrong *et al.*, 2003, p. 27).



**Figura 3.2.** Distribuição lateral das fácies que exibem mudanças na direção horizontal.

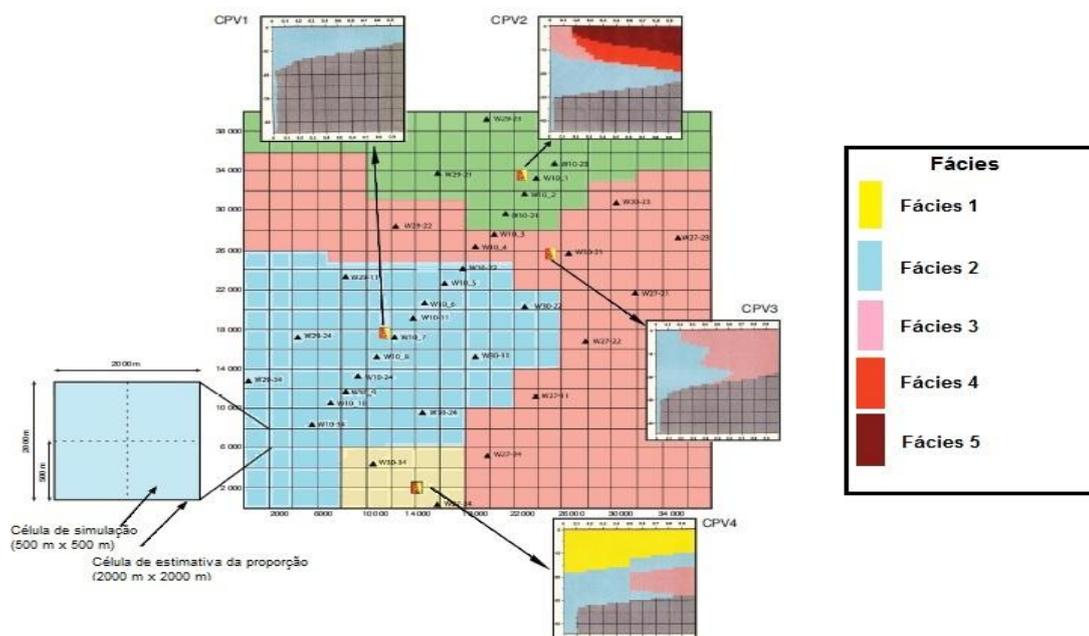
### 3.1.1. CONSTRUÇÃO DE UMA MATRIZ 3D DE CURVAS DE PROPORÇÕES

A variação na direção vertical e na direção horizontal das proporções das fácies requerem a construção de uma matriz de proporções 3D (Braga, 1998, p. 77). Esta matriz de proporção representa um conjunto de curvas de proporções relacionadas a uma célula de um grid 2D que ajuda na quantificação da variação das fácies. Segundo Braga (1998, p. 77) e Chihi & De Marsily (2009, p. 459), os dois métodos mais aplicados na estimativa de uma curva de proporção verticais são: matriz de proporção associada a uma CPV calculada por área, estimativa das proporções a partir das curvas de proporção vertical. Sendo que o primeiro método será adotado nesta dissertação.

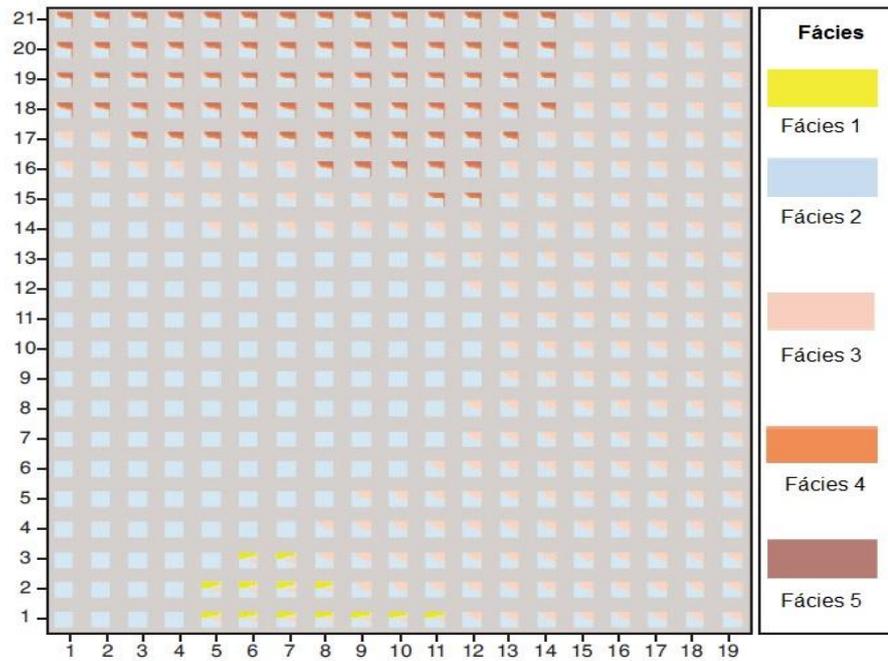
### 3.1.1.1 MATRIZ DE PROPORÇÃO CALCULADA POR ÁREAS

O primeiro passo no cômputo é a definição da matriz de proporções (MP) de fácies sobre a malha de simulação. Segundo Braga (1998, p. 78), esta matriz de proporção deve ser subdividida em zonas a partir de uma série de informações como interpretações geológicas, mapas geológicos derivados de um análise de sísmica das fácies e conhecimentos geológicos, que ajudam na delimitação destas áreas e na construção da matriz de proporções associadas a uma determinada curva de proporção vertical.

De acordo com Chihi & De Marsily (2009, p. 459) e Braga (1998, p. 78), para cada zona delimitada uma curva de proporção vertical é calculada das informações dos poços localmente localizados na zona considerada (ver Figura 3.3). Consequentemente, para cada zona ( $Z_i$ ) computada das curvas de proporção dos poços presentes na área delimitada, cada célula da matriz de proporção é preenchida com sua respectiva curva de proporção estimada (Figura 3.4).



**Figura 3.3.** Mapa de localização da posição dos poços e as curvas de proporção vertical calculadas para cada área (Fonte: Chihi & De Marsily, 2009, p. 460).



**Figura 3.4.** Matriz de proporção para cada unidade associada a cada área delimitada (Fonte: Chihi & De Marsily, 2009, p. 461).

Vários autores como Braga (1998, p. 78), Geovariances (2015, p. 4), Chihi & De Marsily (2009, p. 461), Armstrong *et al.* (2003, p. 32), Remacre *et al.* (2008, pág. 82, sugerem que o cálculo das curvas de proporções pode ser feito via média aritmética ou a partir do uso de interpoladores locais como: o método da krigagem ordinária, o quadrado inverso da distância e da krigagem da média).

Nesse sentido, o método de interpolação local das equações multiquádricas como estimador de variáveis categóricas foi introduzido na aproximação das curvas de proporção verticais na construção da matriz de proporções. Assim, para cada centro da célula 3D da matriz de proporção (MP) do grid, a fácies é interpolada nível a nível, através das equações multiquádricas, com vizinhança única, das fácies disponíveis. Deste modo, a fácies  $k$  pode ser estimada, no ponto  $X_0$  pela seguinte expressão:

$$Z^*(x_0) = \sum_{i=1}^{NP} W_i Z_i(x_i) \quad (3.1)$$

Essa equação tem como condição de restrição  $\sum_{i=1}^{NP} W_i = 1$ . Com NP igual ao número de poços e  $W_i$  os pesos que são obtidos da resolução de um sistema de equações (Yamamoto, 2002, p. 32):

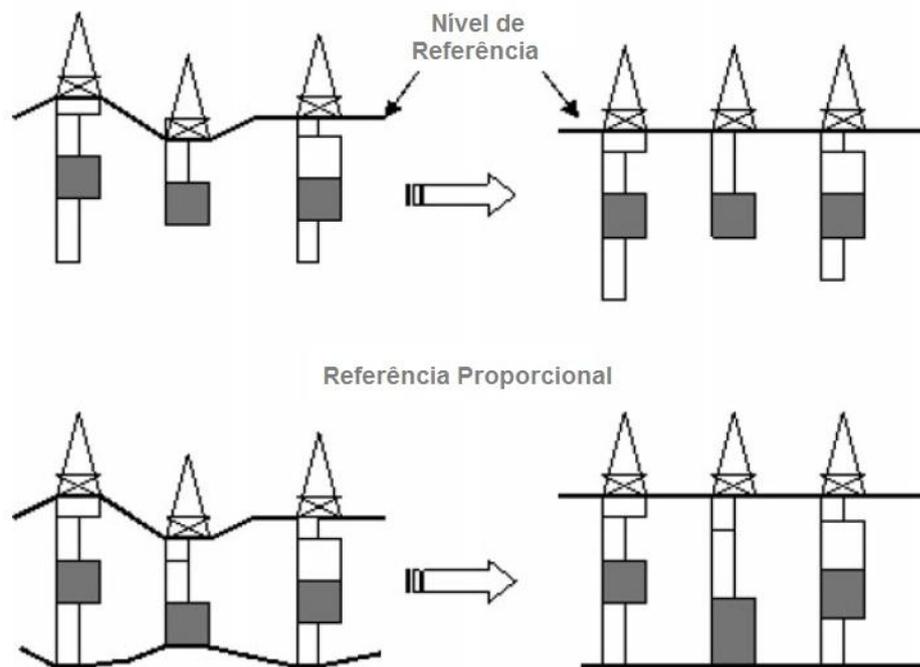
$$\begin{bmatrix} \phi(x_1 - x_1) & \phi(x_1 - x_2) & \cdots & \phi(x_1 - x_n) & 1 \\ \phi(x_2 - x_1) & \phi(x_2 - x_2) & \cdots & \phi(x_2 - x_n) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \phi(x_n - x_1) & \phi(x_n - x_2) & \cdots & \phi(x_n - x_n) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \\ \mu \end{bmatrix} = \begin{bmatrix} \phi(x_0 - x_1) \\ \phi(x_0 - x_2) \\ \vdots \\ \phi(x_0 - x_n) \\ 1 \end{bmatrix} \quad (3.2)$$

em que  $\mu$  é o parâmetro adicional de não viés no sistema de equações multiquádricas.

### 3.1.2. NÍVEL DE REFERÊNCIA

De acordo com Braga (1998, p. 44) e Armstrong et al. (2003, p. 27), o sucesso na aplicação das curvas de proporções depende da escolha de um nível de referência ou marco cronoestratigráfico.

O nível de referência é de muita importância nos resultados da simulação e no cômputo da curva de proporção vertical, esta é uma linha de tempo, preferencialmente marco cronoestratigráfico, não é uma inconformidade ou discordância (Armstrong *et al.*, 2003, p. 35). Nesse caso, em reservatórios de petróleo este marco ou nível de referência é utilizado para restaurar a geometria do depósito no momento da deposição, assim, o reservatório é nivelado mediante o nível de referência (Armstrong *et al.*, 2003, p. 35), como mostra a parte superior da Figura 3.5.



**Figura 3.5.** Duas subsidências de um reservatório e seus níveis de referência (Fonte: Armstrong *et al.*, 2003, p. 36).

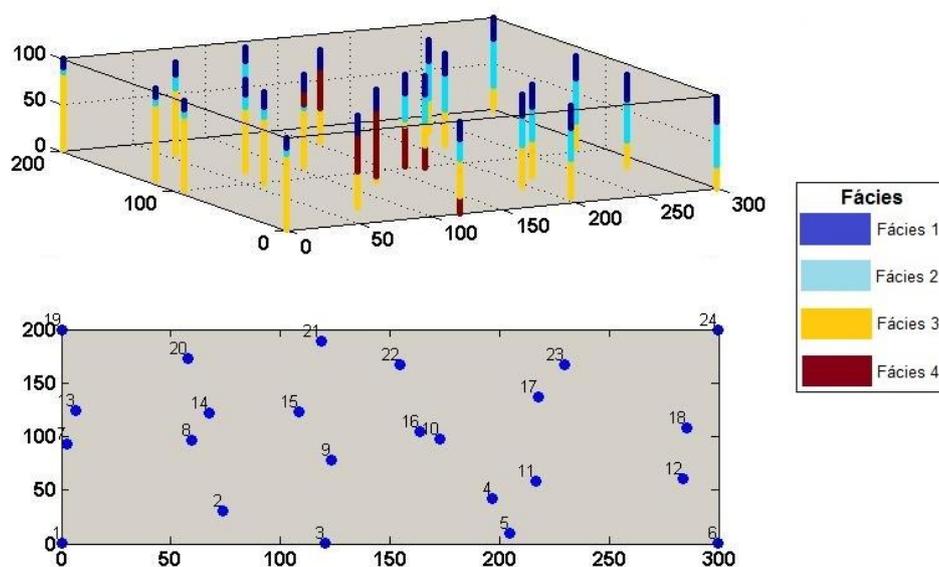
Segundo Armstrong *et al.* (2003, p. 35), o nível de referência usualmente é escolhido na parte superior ou inferior do depósito. O mesmo autor faz referência que a importância na escolha de um nível de referência, testando o impacto como três possíveis níveis de referência no cálculo de uma curva de proporção vertical e na simulação. A discussão extensiva sobre os resultados podem ser vistos em Armstrong *et al.* (2003, p. 35-37).

No cálculo de uma curva de proporção vertical com diferentes níveis, o complemento da proporção faltante deve ser realizado por meio de (Geovariâncas, 2015, p. 3):

1. Níveis não amostrados ou vazios devem ser preenchidos a partir de informações dadas por uma média aritmética estimada camadas amostradas;
2. Ou para cada nível não amostrado, este pode ser preenchido por meio de um conhecimento prévio geológico da área.

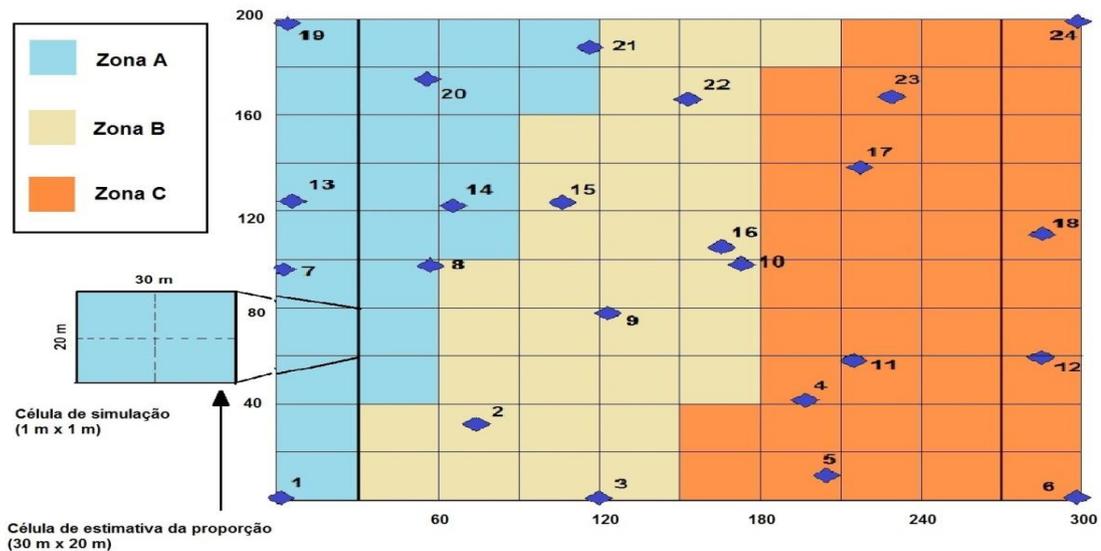
### 3.1.3. EXEMPLO DE APLICAÇÃO DO CÁLCULO DE UMA MATRIZ DE PROPORÇÃO MEDIANTE AS EQUAÇÕES MULTIQUÁDRICAS

Para ilustrar o procedimento no cálculo de uma matriz de proporção aplicando o método das equações multiquádricas, considere-se amostra usada por Yamamoto et al. (2015), composta por 24 poços obtidos por amostragem aleatória estratificada de um modelo geológico sintético (Figura 3.6).



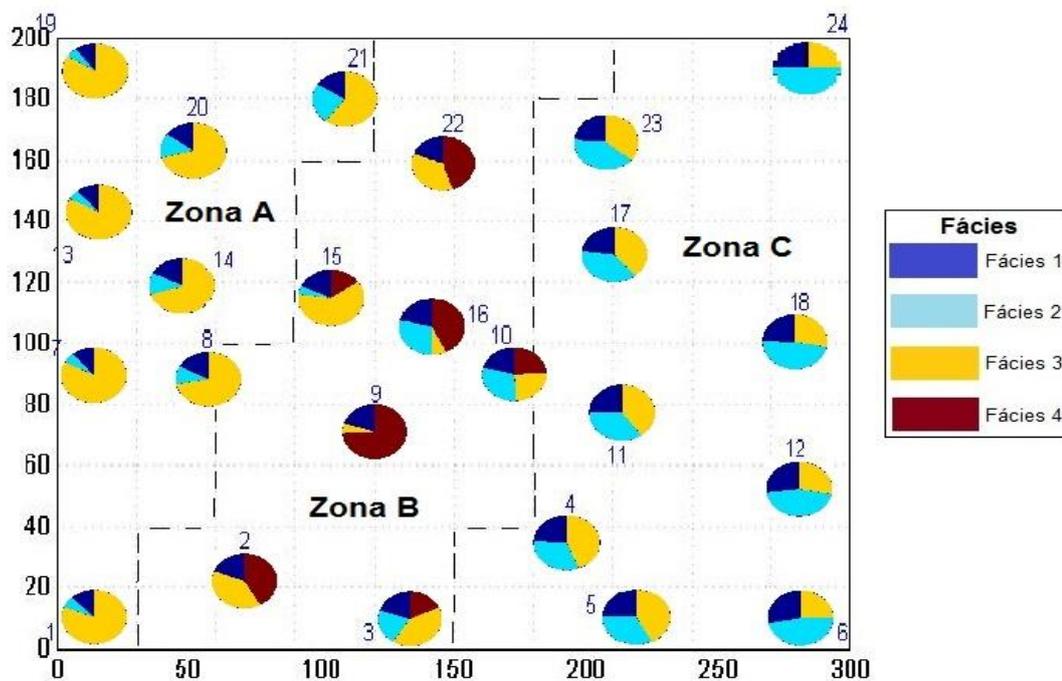
**Figura 3.6.** Mapa de localização dos 24 poços em 3D e 2D. Na parte superior se mostra em 3D a posição dos poços e na parte inferior da Figura, círculos cheios se referem aos poços numerados.

O primeiro passo consiste na definição de uma matriz de proporções de fácies. Esta matriz é construída sobre um grid 2D, com um tamanho de célula de 30 x 20 (Figura 3.7), com tamanho maior que a grid de simulação (1x1).



**Figura 3.7.** Mapa de localização dos poços sobre a matriz de proporções.

A segunda etapa se refere à delimitação de zonas a partir da distribuição espacial das fácies na área, com ajuda de um diagrama em pizza das frequências acumuladas das fácies para cada poço, como é visto na Figura 3.8. Nesse sentido, três zonas foram delimitadas da distribuição das fácies; Zona A, Zona B e Zona C como é visto nas Figuras 3.7 e 3.8.

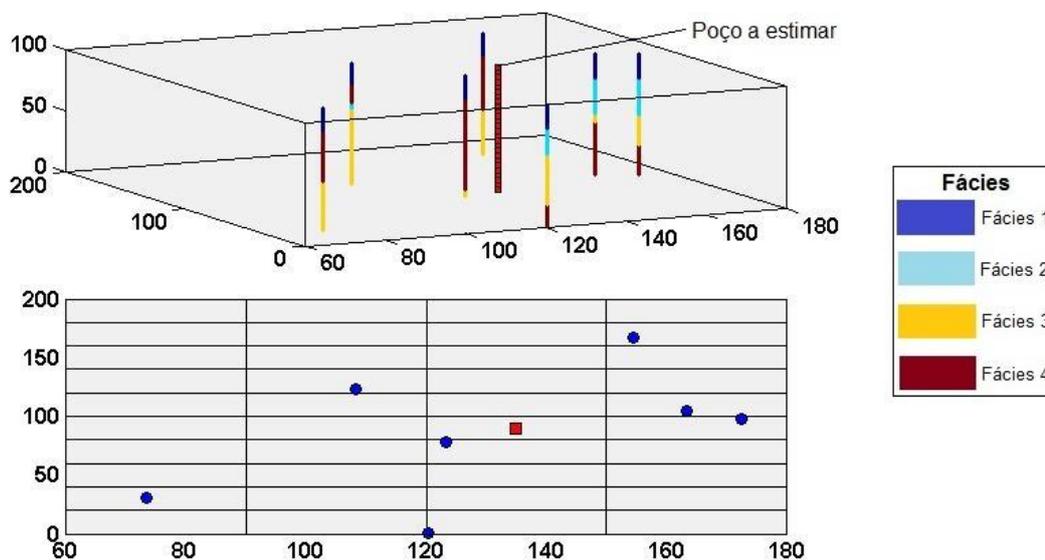


**Figura 3.8.** Diagrama em pizza das fácies por cada poço e delimitação de cada zona.

A seguir, um nível de referência e uma ordem das fácies devem ser considerados a partir das informações geológicas e interpretação dos dados.

De acordo com Yamamoto et al. (2015), o modelo geológico sintético de dados é composto por duas unidades sedimentares (formações Teresina e Piramboia) e um dique associado a vulcanismo que atravessa de forma transversal estas rochas sedimentares. Desta forma, pode se inferir um nível de referência para a interpolação e uma ordem das fácies, sendo a Fácies 1 o nível de referência e uma ordem como; Fácies 1 até a Fácies 4 (Figura 3.6).

O passo seguinte no cálculo da matriz de proporções se dá pela aplicação das equações multiquádricas na aproximação da CPV. Para fins de ilustração foi escolhida a zona B (Figura 3.7), célula 45 da MP (Figura 3.9). A Tabela 3.1 mostra as coordenadas do poço discretizado em 26 unidades ou pontos a estimar.



**Figura 3.9.** Mapa da localização dos poços e o poço a ser interpolado, (■), com indicação dos vizinhos mais próximos.

**Tabela 3. 1.** Pontos não amostrados para interpolação de tipo da variável categórica da célula 45 da MP.

Ponto	X (m)	Y (m)	Z (m)
1	135	90	100
2	135	90	96
3	135	90	92
...	...	...	...
26	135	90	0

A primeira etapa é a transformação de variáveis por meio da codificação binária (Tabela 3.2), dada por (Yamamoto & Landim, 2013, p. 62):

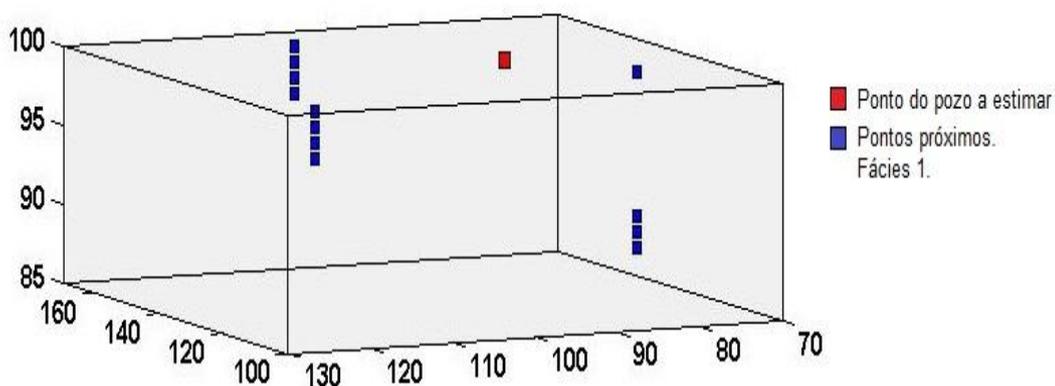
$$I(x, k) = \begin{cases} 1 & \text{se } Z(x) = \text{tipo } k \\ 0 & \text{se } Z(x) \neq \text{tipo } k \end{cases}$$

A codificação binária é mutuamente exclusiva e completa (Yamamoto et al., 2012, p. 147). Trata-se em atribuir probabilidades à ocorrência das fácies. Assim, os resultados da interpolação multiquádrica são também probabilidades. A Tabela 3.2 mostra a codificação binária, bem como também as proporções calculadas, variância e variância total.

**Tabela 3. 2.** Codificação binária das fácies para cada ponto coordenado no poço, também as proporções, variância e variância total.

Nº de Poço	Coordenadas			Tipo de variável categórica			
	X (m)	Y (m)	Z (m)	$I(x_i, k)$ = fácies 1)	$I(x_i, k)$ = fácies 2)	$I(x_i, k)$ = fácies 3)	$I(x_i, k)$ = fácies 4)
1	0,5	0,5	99,50	1	0	0	0
1	0,5	0,5	98,50	1	0	0	0
...	...	...	...	...	...	...	...
1	0,5	0,5	0,5	0	0	1	0
2	73,50	30,50	99,50	1	0	0	0
2	73,50	30,50	98,50	1	0	0	0
...	...	...	...	...	...	...	...
2	73,50	30,50	0,5	0	0	1	0
...	...	...	...	...	...	...	...
24	299,50	199,50	99,50	1	0	0	0
24	299,50	199,50	99,50	1	0	0	0
...	...	...	...	...	...	...	...
24	299,50	199,50	0,50	0	0	1	0
Proporções				0,19	0,23	0,47	0,11
Variância				0,15	0,18	0,25	0,10
Variância Total					0,68		

Para ilustrar passo a passo a interpolação de um ponto não amostrado de tipos de variável categórica, consideramos a primeira coordenada do poço a estimar (135, 90, 100) (Tabela 3.1), que tem doze vizinhos mais próximos, como indicado na Figura 3.10.



**Figura 3.10.** Mapa de localização de pontos e ponto a ser interpolado, (■), com indicação dos vizinhos mais próximos (dados 3D).

A procura dos vizinhos mais próximos foi realizada com base em os quatro pontos mais próximos por quadrante, dentro de um raio de 70 m, onde se encontrou doze pontos vizinhos, conforme a Tabela 3.3.

**Tabela 3. 3.** Pontos de dados vizinhos codificados ao ponto ser estimado (135, 90, 100).

X (m)	Y (m)	Z (m)	$I(x_i, k = \text{fácies 1})$	$I(x_i, k = \text{fácies 2})$	$I(x_i, k = \text{fácies 3})$	$I(x_i, k = \text{fácies 4})$
108,50	123,50	99,50	1	0	0	0
108,50	123,50	98,50	1	0	0	0
108,50	123,50	97,50	1	0	0	0
108,50	123,50	96,50	1	0	0	0
163,50	104,50	99,50	1	0	0	0
163,50	104,50	98,50	1	0	0	0
163,50	104,50	97,50	1	0	0	0
163,50	104,50	96,50	1	0	0	0
123,50	78,50	99,50	1	0	0	0
123,50	78,50	90,50	1	0	0	0
123,50	78,50	89,50	1	0	0	0
123,50	78,50	88,50	1	0	0	0

Com os pontos da Tabela 3.3 pode-se organizar o sistema de equações multiquádricas conforme a equação (1.28):

$$\begin{bmatrix}
 0 & 1,00 & 2,00 & 3,00 & 58,18 & 58,19 & 58,22 & 58,26 & 47,43 & 48,28 & 48,47 & 48,69 \\
 1,00 & 0 & 1,00 & 2,00 & 58,19 & 58,18 & 58,19 & 58,22 & 47,44 & 48,10 & 48,28 & 48,47 \\
 2,00 & 1,00 & 0 & 1,00 & 58,22 & 58,19 & 58,18 & 58,19 & 47,47 & 47,94 & 48,10 & 48,28 \\
 3,00 & 2,00 & 1,00 & 0 & 58,26 & 58,22 & 58,19 & 58,18 & 47,52 & 47,81 & 47,94 & 48,10 \\
 58,18 & 58,19 & 58,22 & 58,26 & 0 & 1,00 & 2,00 & 3,00 & 47,70 & 48,54 & 48,74 & 48,95 \\
 58,19 & 58,18 & 58,19 & 58,22 & 1,00 & 0 & 1,00 & 2,00 & 47,71 & 48,37 & 48,54 & 48,74 \\
 58,22 & 58,19 & 58,18 & 58,19 & 2,00 & 1,00 & 0 & 1,00 & 47,74 & 48,21 & 48,37 & 48,54 \\
 58,26 & 58,22 & 58,19 & 58,18 & 3,00 & 2,00 & 1,00 & 0 & 47,80 & 48,08 & 48,21 & 48,37 \\
 47,43 & 47,44 & 47,47 & 47,52 & 47,70 & 47,71 & 47,74 & 47,80 & 0 & 9,00 & 10,00 & 11,00 \\
 48,28 & 48,10 & 47,94 & 47,81 & 48,54 & 48,37 & 48,21 & 48,08 & 9,00 & 0 & 1,00 & 2,00 \\
 48,47 & 48,28 & 48,10 & 47,94 & 48,74 & 48,54 & 48,37 & 48,21 & 10,00 & 1,00 & 0 & 1,00 \\
 48,69 & 48,47 & 48,28 & 48,10 & 48,95 & 48,74 & 48,54 & 48,37 & 11,00 & 2,00 & 1,00 & 0 \\
 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
 & & & & & & & & & & & 0
 \end{bmatrix}
 \begin{bmatrix}
 W_1 \\
 W_2 \\
 W_3 \\
 W_4 \\
 W_5 \\
 W_6 \\
 W_7 \\
 W_8 \\
 W_9 \\
 W_{10} \\
 W_{11} \\
 W_{12} \\
 \mu
 \end{bmatrix}
 =
 \begin{bmatrix}
 42,71 \\
 42,74 \\
 42,78 \\
 42,78 \\
 42,85 \\
 31,98 \\
 32,01 \\
 32,07 \\
 32,16 \\
 16,27 \\
 18,83 \\
 19,35 \\
 19,91 \\
 1
 \end{bmatrix}$$

Sendo a constante do sistema de equações multiquádricas zero, o núcleo multiquadrático torna-se a distância euclidiana entre pontos. Por exemplo, a distância entre os pontos 1 e 2 na Tabela 3.3 é:

$$d(x_1, x_2) = \sqrt{(108,50 - 108,50)^2 + (123,50 - 123,50)^2 + (99,50 - 98,50)^2} = 1$$

E a distância entre o ponto 1 e o  $(x_0)$  o ponto a ser estimado é:

$$d(x_1, x_0) = \sqrt{(135 - 108,50)^2 + (90 - 123,50)^2 + (100 - 99,50)^2} = 42,71$$

Resolvendo o sistema de equações, obtêm-se os coeficientes:

$$\begin{aligned}
 W_1 &= 0,070; W_2 = 0,0028; W_3 = 0,0027; W_4 = 0,0129; W_5 = 0,1735; W_6 = 0,0084; W_7 \\
 &= 0,0083; W_8 = 0,0953; W_9 = 0,4389; W_{10} = 0,1000; W_{11} = 0,0145; W_{12} \\
 &= 0,0676; \mu = -3,5862
 \end{aligned}$$

Dessa forma, pode se interpolar cada ponto do poço não amostrado fazendo o uso dos coeficientes  $(W_i, i = 1, \dots, 12)$  e a equação 1.28 (Tabela 3.4).

**Tabela 3. 4.** Funções indicadoras nos pontos amostrais para interpolação do tipo categórico ao ponto a estimar e resultados.

$W_i$	$I(x_i; k = \text{fácies 1})$	$I(x_i; k = \text{fácies 2})$	$I(x_i; k = \text{fácies 3})$	$I(x_i; k = \text{fácies 4})$
0,0700	1	0	0	0
0,0028	1	0	0	0
0,0027	1	0	0	0
0,0129	1	0	0	0
0,1735	1	0	0	0
0,0084	1	0	0	0
0,0083	1	0	0	0
0,0953	1	0	0	0
0,4389	1	0	0	0
0,1000	1	0	0	0
0,0145	1	0	0	0
0,0676	1	0	0	0
$I^*(x_0; k)$	1,00	0	0	0
$S_0^2(x_0; k)$	0,00	0	0	0

Segundo Yamamoto & Landim (2013, p. 115), uma vez calculadas as probabilidades (Tabela 3.4) o tipo categórico do ponto não amostrado é dado pelo valor mais provável:

$$I_{MQ}^*(x_0; k) = \max(I_{MQ}^*(x_0; k), k = 1, \dots, k)$$

Desta forma, na tabela 3.4 o valor mais provável é *fácies 1* com um valor de 1,00 sendo o tipo interpolado para o ponto  $x_0$  com coordenadas (135,90,100). Além disso, o tipo categórico é aceitável quando não está na zona de incerteza descrita por Yamamoto & Landim (2013, p. 116). A zona de incerteza esta definida no intervalo  $0,276 < I_{MQ}^*(x_0; k) < 0,724$  (Figura 1.3).

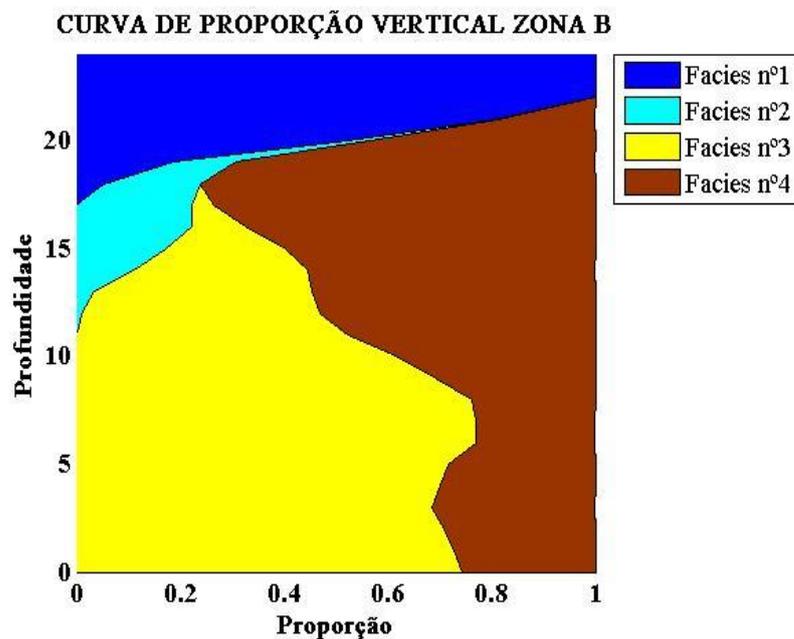
Como base nisso, pode-se atribuir o tipo categórico interpolado ou não na obtenção das curvas de proporção verticais (Tabela 3.5). Por exemplo, aos pontos 15, 16 e 17 foram atribuídos os tipos *fácies 4*, respectivamente, enquanto os pontos 5, 6 e 26 caíram na zona de incerteza, alta variância.

**Tabela 3. 5.** Resultados da interpolação dos tipos no ponto não amostrado.

Ponto	Estatísticas	$k = \text{fácies 1}$	$k = \text{fácies 2}$	$k = \text{fácies 3}$	$k = \text{fácies 4}$
1	$I_{MQ}^*(x_0; k)$	1,00	0,00	0,00	0,00
	$S_0^2(x_0; k)$	0,00	0,00	0,00	0,00
2	$I_{MQ}^*(x_0; k)$	1,00	0,00	0,00	0,00
	$S_0^2(x_0; k)$	0,00	0,00	0,00	0,00
3	$I_{MQ}^*(x_0; k)$	1,00	0,00	0,00	0,00

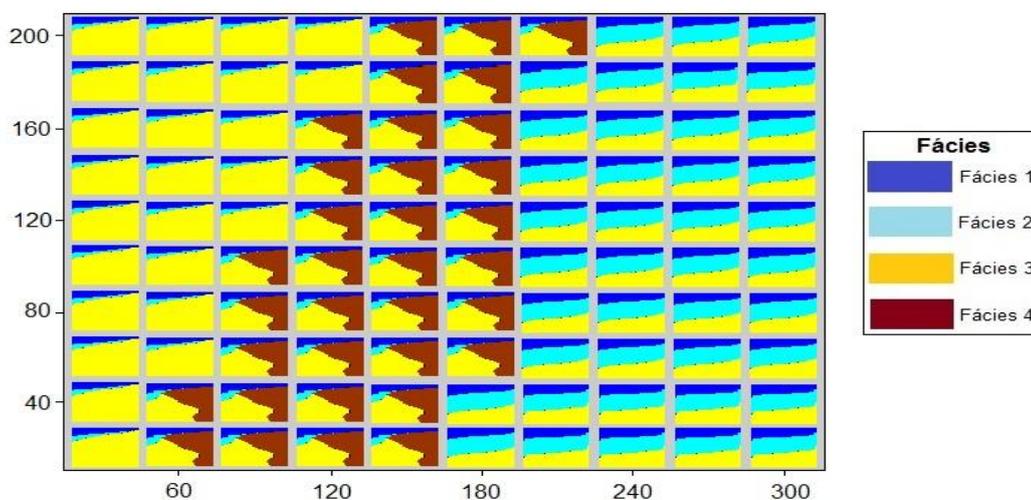
	$s_0^2(x_0; k)$	0,00	0,00	0,00	0,00
4	$I_{MQ}^*(x_0; k)$	0,86	0,00	0,00	0,14
	$s_0^2(x_0; k)$	0,12	0,00	0,00	0,12
5	$I_{MQ}^*(x_0; k)$	0,68	0,08	0,00	0,24
	$s_0^2(x_0; k)$	0,22	0,07	0,00	0,18
6	$I_{MQ}^*(x_0; k)$	0,58	0,12	0,00	0,30
	$s_0^2(x_0; k)$	0,24	0,10	0,00	0,21
7	$I_{MQ}^*(x_0; k)$	0,42	0,18	0,00	0,40
	$s_0^2(x_0; k)$	0,24	0,15	0,00	0,24
8	$I_{MQ}^*(x_0; k)$	0,27	0,24	0,00	0,49
	$s_0^2(x_0; k)$	0,20	0,18	0,00	0,25
9	$I_{MQ}^*(x_0; k)$	0,00	0,33	0,00	0,67
	$s_0^2(x_0; k)$	0,00	0,22	0,00	0,22
10	$I_{MQ}^*(x_0; k)$	0,00	0,30	0,04	0,66
	$s_0^2(x_0; k)$	0,00	0,21	0,04	0,22
11	$I_{MQ}^*(x_0; k)$	0,00	0,33	0,05	0,62
	$s_0^2(x_0; k)$	0,00	0,22	0,05	0,23
12	$I_{MQ}^*(x_0; k)$	0,00	0,29	0,09	0,62
	$s_0^2(x_0; k)$	0,00	0,20	0,08	0,23
13	$I_{MQ}^*(x_0; k)$	0,00	0,16	0,22	0,62
	$s_0^2(x_0; k)$	0,00	0,13	0,17	0,23
14	$I_{MQ}^*(x_0; k)$	0,00	0,14	0,15	0,71
	$s_0^2(x_0; k)$	0,00	0,12	0,13	0,20
15	$I_{MQ}^*(x_0; k)$	0,00	0,10	0,14	0,76
	$s_0^2(x_0; k)$	0,00	0,09	0,12	0,18
16	$I_{MQ}^*(x_0; k)$	0,00	0,06	0,14	0,80
	$s_0^2(x_0; k)$	0,00	0,06	0,12	0,16
17	$I_{MQ}^*(x_0; k)$	0,00	0,00	0,19	0,81
	$s_0^2(x_0; k)$	0,00	0,00	0,15	0,15
18	$I_{MQ}^*(x_0; k)$	0,00	0,00	0,09	0,91
	$s_0^2(x_0; k)$	0,00	0,00	0,08	0,08
19	$I_{MQ}^*(x_0; k)$	0,00	0,00	0,09	0,91
	$s_0^2(x_0; k)$	0,00	0,00	0,08	0,08
20	$I_{MQ}^*(x_0; k)$	0,00	0,00	0,09	0,91
	$s_0^2(x_0; k)$	0,00	0,00	0,08	0,08
21	$I_{MQ}^*(x_0; k)$	0,00	0,00	0,09	0,91
	$s_0^2(x_0; k)$	0,00	0,00	0,08	0,08
22	$I_{MQ}^*(x_0; k)$	0,00	0,00	0,23	0,77
	$s_0^2(x_0; k)$	0,00	0,00	0,18	0,18
23	$I_{MQ}^*(x_0; k)$	0,00	0,00	0,32	0,68
	$s_0^2(x_0; k)$	0,00	0,00	0,22	0,22
24	$I_{MQ}^*(x_0; k)$	0,00	0,00	0,43	0,57
	$s_0^2(x_0; k)$	0,00	0,00	0,24	0,24
25	$I_{MQ}^*(x_0; k)$	0,00	0,00	0,54	0,46
	$s_0^2(x_0; k)$	0,00	0,00	0,25	0,25
26	$I_{MQ}^*(x_0; k)$	0,00	0,00	0,63	0,37
	$s_0^2(x_0; k)$	0,00	0,00	0,23	0,23

Assim, uma curva de proporção vertical para zona B (Figura 3.7) é interpolada das informações dos poços pertencentes por meio de equações multiquádricas, Figura 3.11.



**Figura 3.11.** Curva de proporção vertical associada à zona B.

Finalmente, para cada zona delimitada uma curva de proporção é estimada das informações dos poços localmente localizados e associada à zona considerada. Desta forma, uma matriz de curva de proporção é obtida (Figura 3.12).

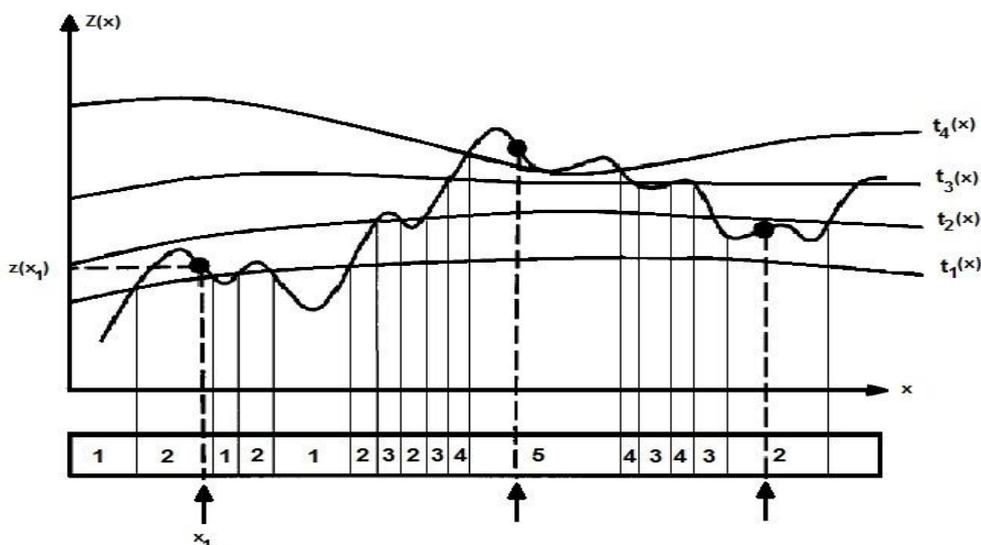


**Figura 3.12.** Matriz de proporção para cada unidade associada a cada área delimitada Zona A, Zona B e Zona C.

### 3.3. TRUNCAMENTOS E PARÂMETROS LIMIARES (THRESHOLDS)

A ideia principal da simulação Gaussiana truncada é de gerar classes de fácies mediante truncamentos de uma função aleatória Gaussiana contínua estacionária, com base em informações geológicas de fácies previamente definidas (Figura 3.13) (Pyrcz & Deutsch, 2014, p. 237). Neste sentido, a simulação pluriGaussiana truncada como uma extensão da Gaussiana é baseada em truncamentos de funções contínuas Gaussianas (Galli et al. 1994, p. 223).

Por exemplo, num plano 2D deseja-se simular duas fácies,  $F_1$  e  $F_2$ , a partir da transformação de uma variável Gaussiana simulada em litologias, intuitivamente temos: se o valor numérico simulado é menor que  $t$ , então obtemos a fácies  $F_1$ ; em outro caso, tem-se a fácies  $F_2$ . Este valor  $t$  é chamado de parâmetro limiar (threshold), segundo Armstrong et al. (2003, p. 42).



**Figura 3.13.** A curva contínua (função Gaussiana) é truncada em quatro parâmetros limiares  $t_i = 1, \dots, 4$  geradas a partir do conhecimento geológico da distribuição das fácies na área (Fonte: Freulon & de Fouquet, 1991).

No exemplo anterior, pode-se definir o parâmetro de corte mediante suas funções indicadoras para uma função contínua Gaussiana  $Z(x)$  simulada, ou seja, mediante uma codificação na presença ou ausência de qualquer ponto  $x$  sobre o domínio  $D$ , com funções indicadoras  $I_{F_1}(x)$  e  $I_{F_2}(x)$ , para as fácies  $F_1$  e  $F_2$ , respectivamente. Onde, em termos matemáticos, a transformada usada é:

Se  $x \in F_1$ , então

$$I_{F_1}(x) = 1 \text{ se e somente se } -\infty \leq Z(x) < t_1 \quad (3.3)$$

Em outro caso, se  $x \in F_2$  temos:

$$I_{F_2}(x) = 1 \text{ se e somente se } t_1 < Z(x) \quad (3.4)$$

Conseqüentemente, num caso geral considerando N-fácies ( $F_1, F_2, \dots, F_N$ ) presentes no espaço, i-variáveis de fácies são definidas entre intervalos especificados por valores de corte associadas a uma função aleatória Gaussiana normal  $Z(x)$ , ou seja:

$$F_i = \{x \in R^d, t_{i-1} \leq Z(x) < t_i\} \quad (3.5)$$

Nesse sentido, para N possíveis valores de fácies devem ser definidos N-1 parâmetros de corte e são considerados numa forma ascendente (Armstrong et al. 2003, p. 43):

$$t_1 \leq t_2 \leq \dots \leq t_{i-1} \leq t_i \leq t_{i+1} \leq \dots \leq t_{N-1} \quad (3.6)$$

Os dados experimentais fornecem as proporções para todas as fácies, que são as probabilidades de que ocorram. Assim, para uma particular fácies  $F_i$  no ponto  $x$ , a probabilidade de obter a fácies  $F_i$  nesse ponto  $x$ , em termos de probabilidade é:

$$P_{F_i}(x) = P(x = F_i) = E\{I_{F_i}(x)\} \quad (3.7)$$

com,

$$I_{F_i}(x) = 1 \Leftrightarrow t_{i-1} \leq Z(x) < t_i \quad (3.8)$$

de onde,

$$P_{F_i}(x) = P(t_{i-1} \leq Z(x) < t_i) = P(-\infty < Z(x) < t_i) - P(-\infty < Z(x) < t_{i-1}) \quad (3.9)$$

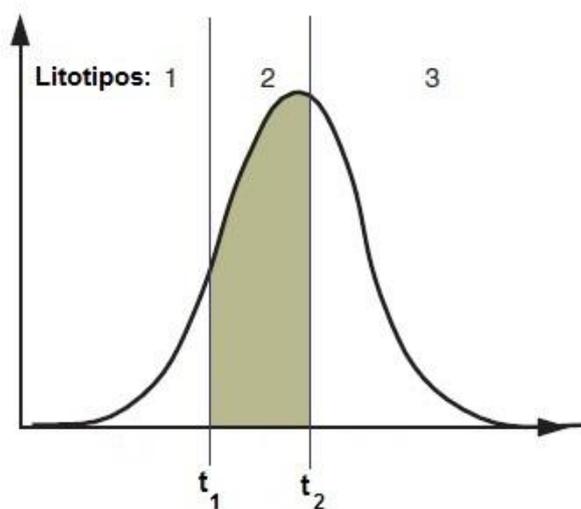
$$P_{F_i}(x) = G(t_i) - G(t_{i-1})$$

onde  $G(t)$  é a distribuição acumulada de uma normal padrão,  $P_{F_i}$  é a probabilidade de obter a fácies  $F_i$  e  $E\{\cdot\}$  é a esperança matemática de um dado valor (Armstrong et al., 2003, p. 45). Assim, os valores limiares são estimados das proporções acumuladas das fácies mediante a função inversa Gaussiana, como:

$$t_i = G^{-1} \left( \sum_{i=1}^N P_{F_i} \right) \quad (3.10)$$

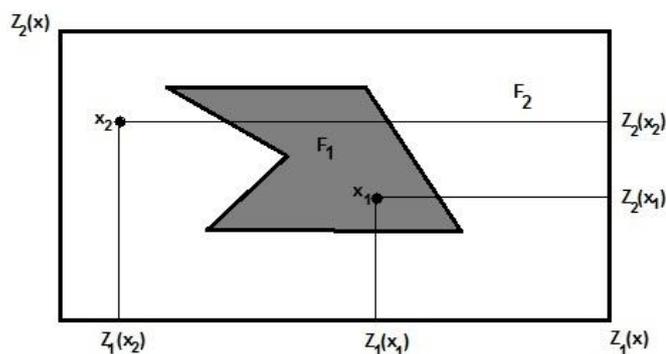
De acordo com Armstrong et al. (2003, p. 47), o primeiro e o último valor de corte são:  $t_0 = -\infty$  e  $t_N = \infty$ , porém,  $N+1$  valores de corte são considerados. Além disso, para o caso não estacionário, se a  $i$ -enésima litologia não está presente, então  $t_{i-1} = t_i$ , segundo os mesmos autores.

Contudo, os parâmetros limiaries, segundo Pyrcz & Deutsch (2014, p. 237), são uma partição ou discretização da área de uma normal padrão, ou seja, as proporções de cada fácies é uma área sob a curva de uma distribuição normal padrão (Figura 3.14).



**Figura 3.14.** Distribuição de uma normal padrão codificada em três tipos de fácies (Fonte: Pyrcz & Deutsch, 2014, p. 237).

No método da simulação pluriGaussiana truncada, funções Gaussianas são simuladas e após, valores categóricos das fácies são obtidas por truncamentos das  $p$  combinações das funções aleatórias Gaussianas geradas (Galli et al. 1994, p. 223). Armstrong et al. (2003, p. 48) ilustram com um exemplo em um plano 2D mediante duas funções aleatórias Gaussianas por meio da partição do domínio em subdomínios  $F_1$  e  $F_2$  (Figura 3.15).



**Figura 3.15.** Representação esquemática em 2D, com duas funções aleatórias Gaussianas e seus valores simulados para os pontos  $x_1$  e  $x_2$ .

A Figura 3.15 apresenta a partição do domínio em duas fácies (branco e cinza). No ponto  $x_1$ , o par de valores simulados  $Z_1(x_1)$  e  $Z_2(x_1)$  pertence à litologia  $F_1$ ; portanto, ao ponto  $x_1$  o valor categórico atribuído é a fácies cinza. No ponto  $x_2$ , o valor atribuído a partir dos valores simulados  $Z_1(x_2)$  e  $Z_2(x_2)$  é a fácies branca. Embora as funções Gaussianas sejam válidas no intervalo  $-\infty$  até  $\infty$ , para uma simplificação gráfica serão representadas por um retângulo.

Em forma geral, sejam  $D_1, D_2, \dots, D_N$  a partição do domínio em subdomínios referentes a  $i$ -variáveis de fácies  $F_i$  e  $Z(x)$  um vetor de  $N$  variáveis aleatórias Gaussianas (independentes ou dependentes) distribuídas normalmente com média 0 e variância 1, tal que:

$$\forall x \in R^d, I_{F_i}(x) = i \Leftrightarrow (Z_1(x), Z_2(x), \dots, Z_N(x)) \in D_i \quad (3.11)$$

Assim, na simulação pluriGaussiana truncada por meio das proporções das fácies  $F_i$  para qualquer ponto  $x$ , a probabilidade de obter a fácies  $F_i$  é descrito como (Armstrong et al. 2003, p. 49):

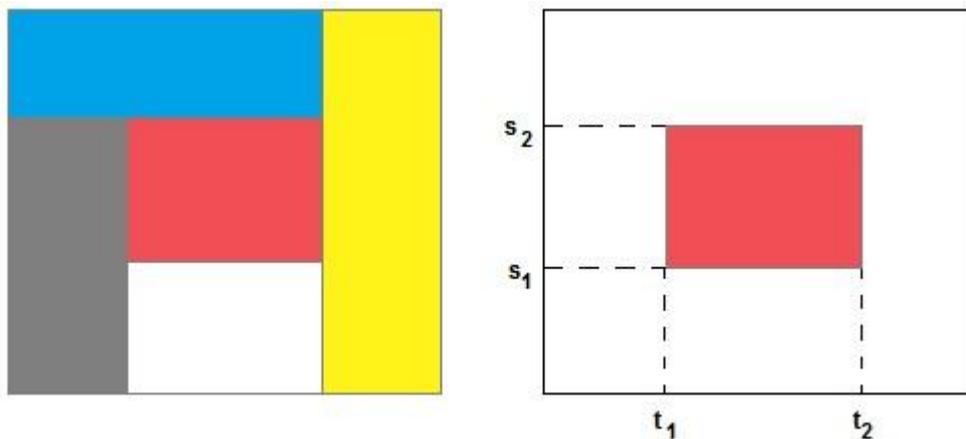
$$P_{F_i}(x) = E\{I_{F_i}(x)\} \quad (3.12)$$

Deste modo, por (3.11) temos:

$$P_{F_i}(x) = P\{Z_1(x), Z_2(x), \dots, Z_N(x)\} = \int_{D_i} g_{\Sigma}(z_1, z_2, \dots, z_N) dz_1 dz_2 \dots dz_N \quad (3.13)$$

onde  $g_{\Sigma}(z_1, z_2, \dots, z_N)$  é a função densidade Gaussiana N-variada com média 0 e variância 1, e  $\Sigma$  é a matriz de correlação entre as variáveis  $z_1, z_2, \dots, z_N$ . No caso onde  $\Sigma$  e  $D_i$  são desconhecidos, a estimativa a partir das proporções  $P_{F_i}$ , é impossível dado que há infinitas soluções, isto é, existem contatos complexos entre subdomínios. Logo, algumas restrições devem ser impostas, em consequência, não falaremos mais de limiares (Armstrong et al. 2003, p. 49).

De acordo com Armstrong et al. (2003, p. 50), para resolver este tipo de problema o espaço gaussiano é particionado em retângulos (duas funções Gaussianas) ou em paralelepípedos retangulares com mais de duas funções Gaussianas. Por exemplo, numa área de interesse (Figura 3.16), mostra-se uma divisão do domínio em cinco fâcies em retângulos para duas funções aleatórias Gaussianas. Os valores de corte são definidos a partir das projeções sobre os eixos de coordenadas Gaussianas.

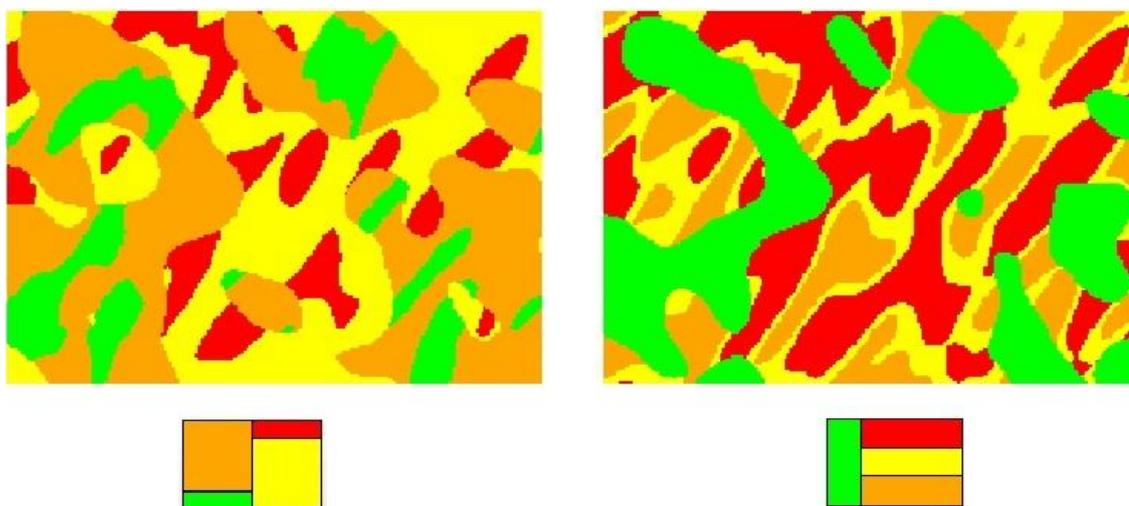


**Figura 3.16.** Partição do domínio em retângulos (à esquerda) e seus valores de corte (à direita).

Sendo  $N$  o número de funções aleatórias Gaussianas, o número de valores de corte para o domínio dividido em retângulos é  $2^N$  (alguns podem ser infinitos), além disso, conhecendo  $2^N - 1$  deles, a equação (3.13) pode ser invertida numericamente e inferir o último valor de corte desconhecido.

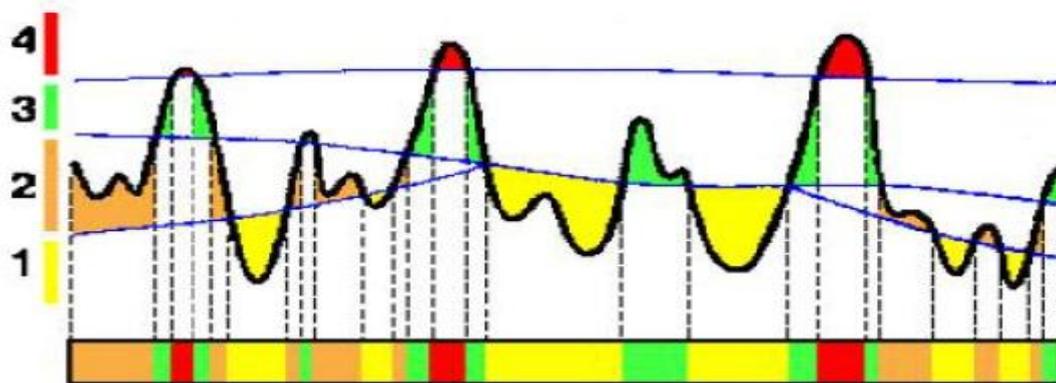
Vários autores, tais como Armstrong et al. (2003, p. 50) e Dowd et al. (2003, p. 125), afirmam que mesmo fazendo o uso de uma partição do espaço em retângulos, existem diversas combinações possíveis. Por exemplo, para três fácies com duas funções Gaussianas há duas possibilidades. Com quatro fácies e duas funções Gaussianas têm-se dez combinações possíveis. Neste caso, se houve um aumento no número de fácies, diversas partições do domínio podem gerar as mesmas configurações nas imagens simuladas.

Deste modo, Emery (2007, p. 1190) e Dowd et al. (2003, p. 125) sugerem que deve ser adotado um critério na escolha da partição do domínio, mediante as relações espaciais entre fácies, estabelecendo contatos permitidos e proibidos, e reproduzindo uma ordem cronológica das fácies, regra de truncamento. No caso de não estacionaridade as proporções de fácies, algumas podem desaparecer localmente variando o diagrama de partição de forma a impedir contatos proibidos. Esta regra ou partição é simbolizada por uma bandeira sob ou ao lado de cada realização (Figura 3.17).



**Figura 3.17.** Esquemática de duas simulações pluriGaussianas truncadas como diferentes regras de truncamento.

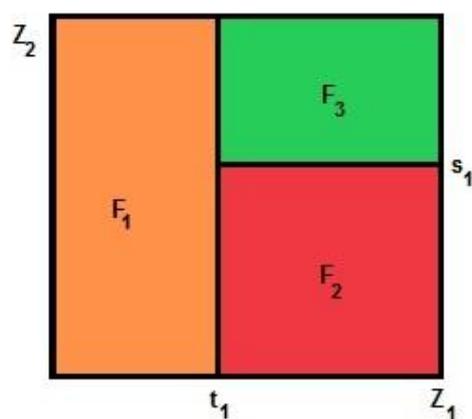
Contudo, em algumas simulações contatos não permitidos estão presentes, sendo evidenciado este tipo de problema por várias razões; no caso estacionário e não estacionário mostra-se quando a função Gaussiana é descontínua ou a discretização é muito grossa. Além disso, no caso não estacionário, onde há mudanças nas proporções das fácies que variam bruscamente em níveis consecutivos, contatos não presentes na regra de truncamento são vistos (Armstrong et al. 2003, p. 54-55) (Figura 3.18).



**Figura 3.18.** Esquema de uma função Gaussiana contínua, no caso não contínuo truncada por três valores de corte. Numa certa localização contatos proibidos são mostrados na ausência da fácies 2 (Fonte: Folle (2009, p. 125)).

### 3.3.1. EXEMPLOS DE APLICAÇÃO DE TRUNCAMENTO E PARÂMETROS LIMIARES (THRESHOLDS)

A seguir apresentam-se dois exemplos da aplicação de truncamento e cálculo de parâmetros de corte. O primeiro exemplo começa com um exercício mostrando passo a passo o cálculo de parâmetro de corte a partir de uma regra de truncamento o qual foi dividido em três fácies  $F_1$ ,  $F_2$  e  $F_3$  para duas funções Gaussianas contínuas  $Z_1(x)$  e  $Z_2(x)$  com média 0 e variância 1 e independentes (Figura 3.19).



**Figura 3.19.** Regra de truncamento para três fácies e duas funções aleatórias Gaussianas.

Em todos os casos  $t_0 = -\infty$  e  $t_N = \infty$  para a primeira função Gaussiana  $Z_1(x)$  e  $s_0 = -\infty$  e  $s_N = \infty$  para a segunda função Gaussiana  $Z_2(x)$ . Da equação (3.5) pode-se inferir que qualquer fácies pode ser obtida discretizando a função aleatória Gaussiana, ou seja:

$$\begin{aligned} F_1 &= \{x; Z_1(x) \leq t_1\} \\ F_2 &= \{x; Z_1(x) > t_1, Z_2(x) \leq s_1\} \\ F_3 &= \{x; Z_1(x) > t_1, Z_2(x) > s_1\} \end{aligned}$$

Nesse caso, para cada função Gaussiana  $N - 1$  valores de corte devem ser estimados. Assim, a probabilidade de obter certa fácies num ponto  $x$  é dada pela equação (3.9), isto é:

$$\begin{aligned} P_{F_1}(x) &= P\{Z_1(x) \leq t_1\} \\ t_1 &= G^{-1}(P_{F_1}(x)) \\ P_{F_2}(x) &= P\{Z_1(x) > t_1, Z_2(x) \leq s_1\} \\ P_{F_2}(x) &= P\{Y_1 > t_1, (\rho Y_1 + \sqrt{1 - \rho^2} Y_2) \leq s_1\} \end{aligned}$$

Como  $Z_1(x)$  e  $Z_2(x)$  são duas variáveis Gaussianas independentes e usando a densidade Gaussiana monovariada  $g(x)$ , a expressão anterior pode ser escrita como:

$$P_{F_2}(x) = P\{Y_1 > t_1, Y_2 \leq s_1\}$$

$$P_{F_2}(x) = \int_{t_1}^{\infty} \int_{-\infty}^{s_1} g(u)g(v)du dv$$

Ou usando a função de distribuição Gaussiana  $G(x)$ :

$$P_{F_2}(x) = \int_{t_1}^{\infty} (G(s_2) - G(-\infty)) g(v)dv$$

Para um melhor entendimento na aplicação das equações desenvolvidas anteriormente, são especificadas as proporções das fácies sendo constantes no domínio, assim, sejam  $P_{F_1} = 30\%$  e  $P_{F_2} = 25\%$ .

Usando (3.10), temos:

$$t_1 = -0.524$$

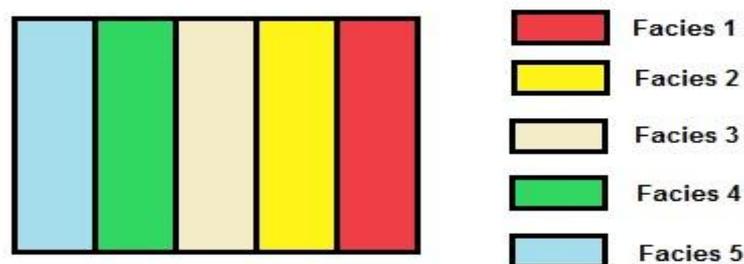
De (3.10) e fazendo o uso da independência entre as Gaussianas, pode se inferir:

$$\begin{aligned}
 P_{F_2} &= P\{Y_1 > t_1, Y_2 \leq s_1\} \\
 P_{F_2} &= P\{Y_1 > t_1\}P\{Y_2 \leq s_1\} \\
 P_{F_2} &= (1 - P_{F_1})G(s_1) \\
 s_1 &= G^{-1}\left(\frac{P_{F_2}}{(1 - P_{F_1})}\right)
 \end{aligned}$$

Assim,  $s_1 = -0.366$ .

Para ilustrar o procedimento do cálculo de valores de corte num espaço não estacionário, foi considerado o exemplo do item 3.2 (cálculo das proporções) a Figura 3.1. Neste exercício, sobre uma malha regular se tem cinco poços fictícios com alguns testemunhos ao longo do poço. Além disso, as cinco fácies são apreciadas de igual comprimento. No caso onde as fácies são estacionárias os valores limiares permanecem constantes, caso contrário, variam no espaço, segundo Armstrong et al. (2003, p. 47).

O primeiro passo na estimativa de valores limiares consiste na identificação do número de fácies presentes no domínio e a relação espacial entre contatos das fácies, com os quais define-se uma regra de truncamento ver Figura 3.20.



**Figura 3.20.** Regra de truncamento baseado na ordem de deposição das fácies para os cinco poços.

Segundo Armstrong et al. (2003, p. 6), numa deposição sequencial das fácies somente uma função aleatória Gaussiana deve ser simulada, a qual é denominada simulação Gaussiana truncada.

No segundo passo, a partir do nível de referência faz-se a numeração de cada nível e a proporção de todas as fácies pertencentes ao nível (Tabela 3.6). No exemplo foram numerados 10 níveis de forma ascendente desde a parte superior (nível de referência) até a parte inferior como se mostra na Tabela 3.6.

**Tabela 3. 6.** Proporções de todas as fácies nível a nível em percentagens.

Nível	$F_5$	$F_4$	$F_3$	$F_2$	$F_1$
1	100%	0%	0%	0%	0%
2	0%	80%	20%	0%	0%
3	0%	40%	60%	0%	0%
4	0%	60%	40%	0%	0%
5	0%	40%	40%	20%	0%
6	0%	0%	60%	20%	20%
7	0%	0%	40%	40%	20%
8	0%	0%	0%	60%	40%
9	0%	0%	0%	20%	80%
10	0%	0%	0%	0%	100%

Com base na percentagem das proporções na tabela 3.6 são calculados os valores de corte nível a nível, fazendo uso das proporções acumuladas e a inversa da função Gaussiana equação (3.10):

$$t_0 = -\infty$$

$$t_1 = G^{-1}(P_5) = G^{-1}(100\%) = \infty$$

$$t_2 = G^{-1}(P_5 + P_{F_4}) = G^{-1}(100\% + 0\%) = \infty$$

$$t_3 = G^{-1}(P_{F_5} + P_{F_4} + P_{F_3}) = G^{-1}(100\% + 0\% + 0\%) = \infty$$

$$t_4 = G^{-1}(P_{F_5} + P_{F_4} + P_{F_3} + P_{F_2}) = G^{-1}(100\% + 0\% + 0\% + 0\%) = \infty$$

$$t_5 = G^{-1}(P_{F_5} + P_{F_4} + P_{F_3} + P_{F_2} + P_{F_1}) = G^{-1}(100\% + 0\% + 0\% + 0\% + 0\%) = \infty.$$

O resultado anterior na resolução da equação (3.10) mostra o primeiro nível. Assim, para cada nível desde o 1 até 10 deve se calcular os parâmetros de corte (tabela 3.7).

**Tabela 3. 7.** Parâmetros limiaries correspondentes a cada proporção mostrada na tabela 3.6.

Nível	$t_0$	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$
1	$-\infty$	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$
2	$-\infty$	$-\infty$	0.84	$\infty$	$\infty$	$\infty$
3	$-\infty$	$-\infty$	-0.25	$\infty$	$\infty$	$\infty$
4	$-\infty$	$-\infty$	0.25	$\infty$	$\infty$	$\infty$
5	$-\infty$	$-\infty$	-0.25	0.84	$\infty$	$\infty$
6	$-\infty$	$-\infty$	$-\infty$	0.25	0.84	$\infty$
7	$-\infty$	$-\infty$	$-\infty$	-0.25	0.84	$\infty$
8	$-\infty$	$-\infty$	$-\infty$	$-\infty$	0.25	$\infty$
9	$-\infty$	$-\infty$	$-\infty$	$-\infty$	-0.84	$\infty$
10	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$-\infty$	$\infty$

### 3.4. CÁLCULO E MODELAGEM DE VARIOGRAMAS EXPERIMENTAIS

O passo seguinte na simulação de variáveis categóricas mediante o método da Gaussiana ou pluriGaussiana truncada é a determinação de um ou mais modelos apropriados de variograma de valores Gaussianos para gerar campos aleatórios gaussianos (Kyriakidis, Deutsch, & Grant, 1999, p. 161-162).

Segundo Armstrong et al. (2003, p. 58), Le Loc'h e Galli (1997, p. 217) e Dowd et al. (2003, p. 127), teoricamente existe uma relação entre o variograma de indicadores e o variograma cruzado para cada diferente fácies com a covariância de uma função aleatória Gaussiana:

$$C_F(x, x + h) = E[I_F(x)I_F(x + h)] = P[x \in F \text{ e } x + h \in F] \quad (3.14)$$

e

$$C_{F_i F_j}(x, x + h) = E[I_{F_i}(x)I_{F_j}(x + h)] = P[x \in F_i \text{ e } x + h \in F_j] \quad (3.15)$$

Assim, e de acordo com Le Loc'h e Galli (1997, p. 215), a partir do conhecimento do coeficiente de correlação, os valores de corte e modelos dos variogramas Gaussianos podem-se computar os variogramas e variogramas cruzados das funções indicadoras das fácies no caso estacionário e não estacionário.

Primeiramente, no caso estacionário por meio da simulação Gaussiana truncada para a fácies  $F_i$  com  $t_{i-1}$  e  $t_i$  parâmetros de corte e uma relação entre o variograma equação (1.13) (definição de variograma) pode-se inferir:

$$\begin{aligned} \gamma_{F_i}(x, x+h) = & \frac{1}{2} \{P[t_{i-1} \leq Z(x) < t_i] + P[t_{i-1} \leq Z(x+h) < t_i]\} \\ & - P[t_{i-1} \leq Z(x) < t_i \text{ e } t_{i-1} \leq Z(x+h) < t_i] \end{aligned} \quad (3.16)$$

da equação (3.14), temos:

$$\gamma_F(x, x+h) = p_{F_i} - C_F(x, x+h) \quad (3.17)$$

ou

$$\gamma_F(x, x+h) = p_{F_i} - \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} g_{\rho(h)}(u, v) du dv \quad (3.18)$$

onde  $p_{F_i}$  é a proporção da fácies  $F_i$  e  $g_{\rho(h)}$  é a densidade de probabilidade de uma função bivariada Gaussiana padrão com coeficiente de correlação  $\rho(h)$ . Analogamente, o variograma cruzado derivado da covariância não centrada é dado por:

$$\gamma_{F_i F_j}(x, x+h) = -\frac{1}{2} \left\{ E \left[ I_{F_i}(x) I_{F_j}(x+h) \right] + \left[ I_{F_j}(x) I_{F_i}(x+h) \right] \right\} \quad (3.19)$$

assim, da equação (3.15):

$$\gamma_{F_i F_j}(x, x+h) = -\frac{1}{2} \left\{ C_{F_i F_j}(x, x+h) + C_{F_j F_i}(x, x+h) \right\} \quad (3.20)$$

ou

$$\gamma_{F_i F_j}(x, x+h) = - \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} g_{\rho(h)}(u, v) du dv \quad (3.21)$$

com  $g_{\rho(h)}$  a densidade de probabilidade de uma função bivariada Gaussiana padrão com coeficiente de correlação  $\rho(h)$ .

As equações anteriores foram calculadas para um espaço estacionário. Assim, para o caso não estacionário com variações nos parâmetros de corte para uma distância  $h$ , o modelo de variograma é dado por (Armstrong et al., 2003, p. 64):

$$\begin{aligned} \gamma_{F_i} \left( I_{F_i}(x), I_{F_i}(x+h) \right) &= \frac{1}{2} \{ P[t_{i-1}(x) \leq Z(x) < t_i(x)] \\ &+ P[t_{i-1}(x+h) \leq Z(x+h) < t_i(x+h)] \} \\ &- P[t_{i-1}(x) \leq Z(x) < t_i(x) \text{ e } t_{i-1}(x+h) \leq Z(x+h) \\ &< t_i(x+h)] \end{aligned} \quad (3.22)$$

onde  $t_{i-1}$  e  $t_i$  são os valores de corte da fácies  $F_i$ . Desta maneira, a expressão anterior (3.22) pode ser escrita como (equação (1.10)):

$$\gamma_{F_i} \left( I_{F_i}(x), I_{F_i}(x+h) \right) = \frac{1}{2} \{ p_{F_i}(x) + p_{F_i}(x+h) - C_{F_i} \left( I_{F_i}(x), I_{F_i}(x+h) \right) \} \quad (3.23)$$

com  $p_{F_i}(x)$  e  $p_{F_i}(x+h)$  as proporções da fácies  $F_i$  para os pontos  $x$  e  $x+h$  separados por uma distância  $h$ . Do mesmo modo, o modelo de variograma cruzado é dado por:

$$\begin{aligned} \gamma_{F_i F_j} \left( I_{F_i}(x), I_{F_j}(x+h) \right) &= -\frac{1}{2} \{ C_{F_i F_j} \left( I_{F_i}(x), I_{F_j}(x+h) \right) + C_{F_j F_i} \left( I_{F_j}(x), I_{F_i}(x+h) \right) \} \end{aligned} \quad (3.24)$$

No caso da simulação pluriGaussiana truncada, os variogramas experimentais diretos e cruzados das indicadoras são calculados exatamente como os passos anteriores do método gaussiano truncado e derivados da covariância não centrada (Armstrong et al. (2003, p. 58) e Emery (2007, p. 1190)):

$$\gamma_{F_i F_j}(x, x+h) = C_{F_i F_j}(0) - \frac{\{C_{F_i F_j}(x, x+h) + C_{F_j F_i}(x, x+h)\}}{2} \quad (3.25)$$

com

$$C_{F_i F_j}(x, x+h) = Prob(Y(x) \in D_i, Y(x+h) \in D_j) \quad (3.26)$$

Para um caso estacionário, da equação 3.26 com  $D_i$  o domínio de um espaço gaussiano.

Segundo Emery (2007, p. 1190), no caso das funções aleatórias Gaussianas serem mutuamente independentes, a equação (3.26) é previamente descrita pelo produto das probabilidades, da seguinte forma:

$$\begin{aligned} & \prod_{i=1}^m Prob(Y_i(x) \in [t_{i-1}, t_i], Y_i(x+h) \in [s_{i-1}, s_i]) \\ & = \prod_{i=1}^m \int_{t_{i-1}}^{t_i} \int_{s_{i-1}}^{s_i} g_{\rho_i(h)}(u, v) du dv \end{aligned} \quad (3.27)$$

onde  $g_{p(h)}$  é a densidade de probabilidade de uma função Gaussiana bivariada padrão com coeficiente de correlação  $\rho$ .

Vários autores, entre eles, Dowd et al. (2003, p. 131), Galli et al. (1994, p. 221) e Emery (2007, p. 1190) afirmaram que a equação (3.27) usualmente é computada por integração numérica.

Segundo Kyriakidis et al. (1999, p. 162) e Pyrcz & Deutsch (2014, p. 238), como os dados disponíveis são as indicadoras das fácies e um modelo de covariância derivado da função aleatória Gaussiana, torna-se necessário a geração de um campo gaussiano. No caso de duas fácies separadas unicamente pelo valor de corte  $t_i$ , há uma relação direta entre o modelo de covariância normal e o modelo de covariância das indicadoras. Esta relação direta é dada unicamente no caso de duas fácies.

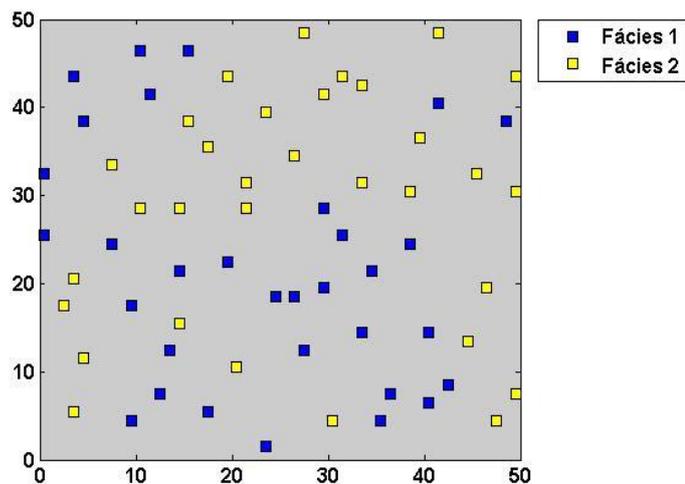
Recentemente, em aplicações da simulação pluriGaussiana truncada (Emery ,2007, p. 1190; Emery & Gonzalez ,2007, p. 32) uma análise variográfica, a partir de realizações iterativas baseadas no método de tentativa e erro, foi usada na seleção de um modelo de covariância normal e covariância cruzada (Le Loc'h & Galli, 1997, p. 217).

Finalmente, o ajuste de um modelo de variograma consiste na comparação gráfica de um modelo escolhido de covariância normal e o modelo experimental de variograma, (Le Loc'h & Galli, 1997, p. 217). Esta seleção baseada na estratégia de tentativa e erro pode ser resumida em três passos (Emery & Gonzalez, 2007, p. 33):

1. Modelos de covariâncias para cada função aleatória Gaussiana são propostos;
2. Obter os modelos de variogramas de cada fácies e variogramas cruzados das indicadoras;
3. Comparar esses modelos com seus variogramas experimentais diretos e cruzados, até o ajuste ser satisfatório.

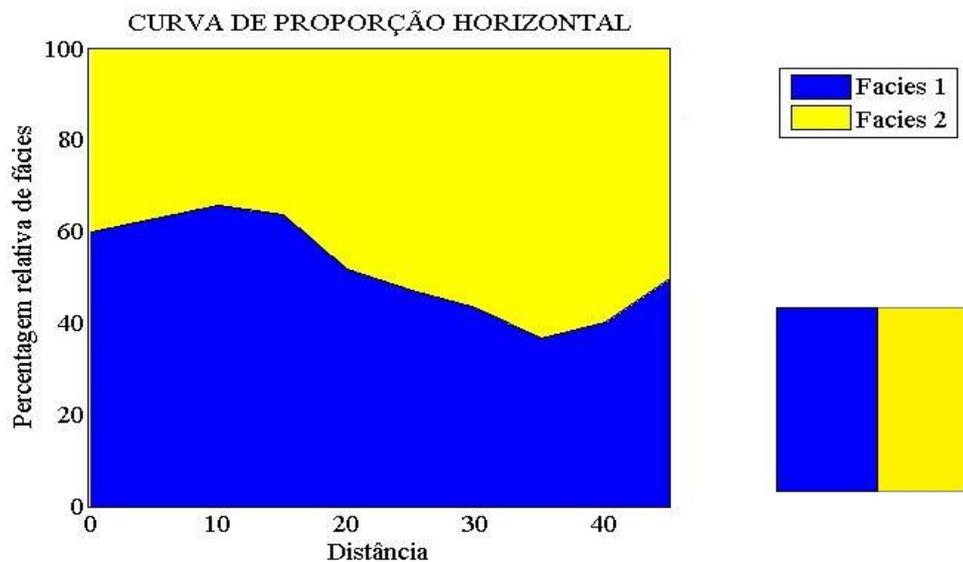
#### ***3.4.1. EXEMPLO DE APLICAÇÃO NA OBTENÇÃO DE UM MODELO DE COVARIÂNCIA APROPRIADO***

A seguir se apresenta um exemplo de aplicação na ilustração de seleção de um modelo de covariância apropriado baseado na estratégia de tentativa e erro, para o qual foi escolhido o Arquivo 12, Anexo B disponível em Yamamoto & Landim (2013, p. 201). O primeiro passo é a codificação binária dos dados por meio de um valor de corte, no caso a mediana, com forme o mapa de localização da Figura 3.21.



**Figura 3.21.** Localização dos pontos da amostra, no qual os quadros azuis representam valores menores que a mediana (fácies 1) e quadros amarelos, valores maiores que a mediana (fácies 2).

A seguir, uma ordem de fácies é definida representada por uma bandeira ou regra de truncamento, ilustrando o tipo de contato entre fácies e o cálculo das curvas de proporção no eixo horizontal é realizado como foi descrito no cômputo de uma curva de proporção (Figura 3.22):



**Figura 3.22.** Curva de proporção horizontal (esquerda) e uma regra de truncamento (direita).

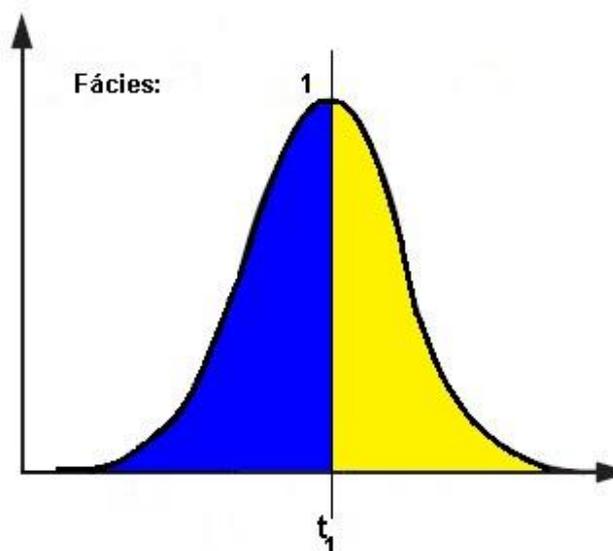
A curva de proporção na direção horizontal (Figura 3.22) computada das informações de um banco de dados 2D mostra que não existe nenhum posicionamento preferencial das fácies, ou seja, considerando-se a distribuição das fácies como estacionárias na área.

Nesse sentido, as proporções na área são consideradas constantes em todo o domínio. Assim, procede-se a calcular o parâmetro de corte  $t_1$  dados pela equação (3.10), que separa unicamente as fácies  $F_1$  e  $F_2$  (Tabela 3.8).

**Tabela 3. 8.** Proporções e parâmetros de corte da fácies 1 e fácies 2, respectivamente.

Proporção		Valor de corte		
$p_{F_1}$	$p_{F_2}$	$t_0$	$t_1$	$t_2$
0,50	0,50	$-\infty$	0	$\infty$

Segundo Pyrcz & Deutsch (2014, p. 237), os valores de corte são uma discretização da área de uma distribuição normal padrão (Figura 3.23).



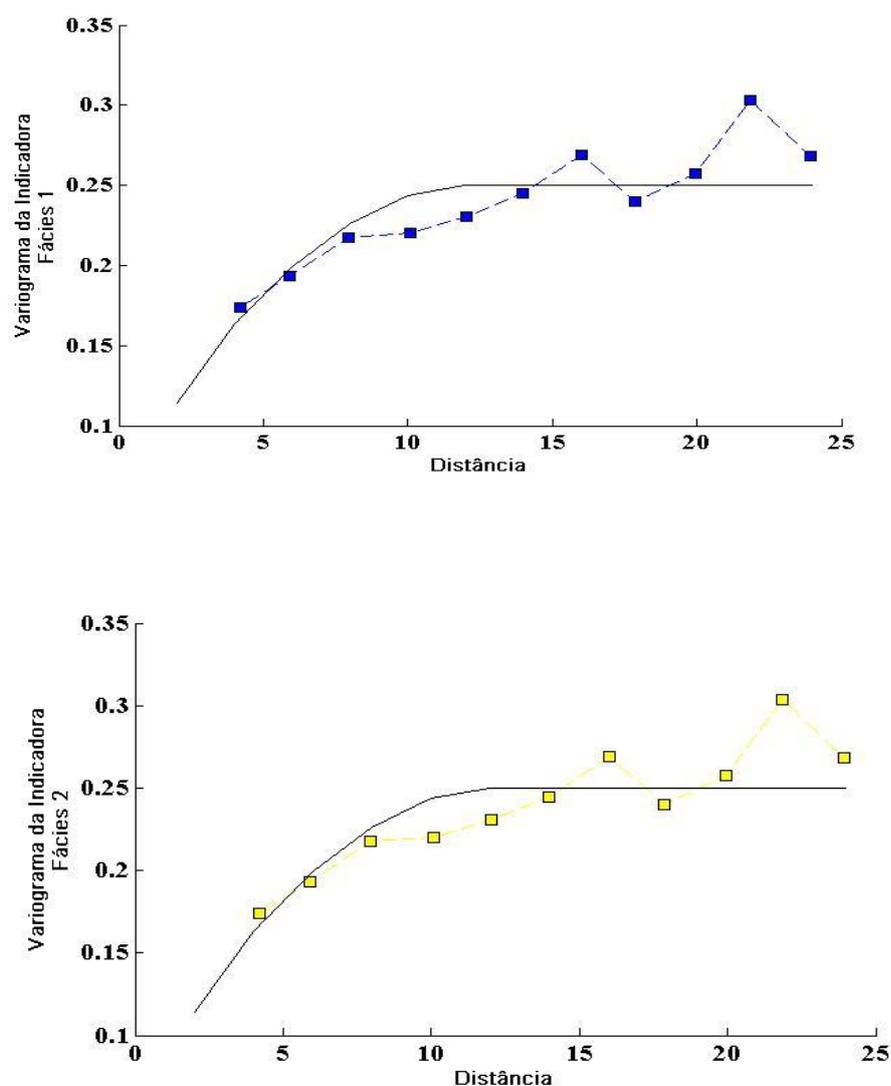
**Figura 3.23.** Proporção da fácies 1 e 2 sob a curva de uma distribuição normal padrão.

Com base nessa curva e a regra de truncamento (Figura 3.22) se define uma única função aleatória Gaussiana, ou seja, um só campo aleatório gaussiano deve ser simulado e um modelo apropriado de covariância da função Gaussiana deve ser proposto baseado na técnica de tentativa e erro realizado com o programa VMODEL\_IND.M desenvolvido por Emery (2007).

Um variograma esférico é usado para ajustar dois simples variogramas para cada indicadora de fácies. A Figura 3.24 mostra o ajuste do variograma e os parâmetros do modelo indicado na tabela 3.9. Podemos observar que os variogramas das indicadoras tem um comportamento linear perto da origem.

**Tabela 3. 9.** Modelo de variograma gaussiano para a função Gaussiana aleatória.

	Tipo de modelo	Patamar	Rango Horizontal
Primeira Função Aleatória	Esférico	0,04	12
	Esférico	0,96	12



**Figura 3.24.** Variograma experimental (quadros e linhas ponteadas) e modelado (linhas sólidas) das indicadoras na horizontal.

### **3.5. AMOSTRADOR DE GIBBS**

O objetivo final da simulação Gaussiana e pluriGaussiana truncada é condicionar os dados. A primeira etapa é a transformação das informações de tipos de fácies, dados categóricos, em valores gaussianos com um correto modelo de covariância e pertencentes a um intervalo com o objetivo de reproduzir a continuidade espacial de um campo gaussiano  $Z$ , segundo Emery (2007, p. 1191) e Armstrong et al. (2003, p. 81).

Uma vez gerado o conjunto de valores Gaussianos, a segunda etapa consiste em obter realizações condicionais de funções Gaussianas contínuas  $Z$ , por meio de qualquer método de simulação (Armstrong et al., 2003, p. 77); por exemplo: bandas rotativas, simulação Gaussiana sequencial, LU decomposição etc. (Yamamoto & Landim (2013, p. 149-159); Olea (1999, p. 141-158); Pyrcz & Deutsch (2014)). Finalmente, as realizações contínuas de funções Gaussianas são transformadas em tipos categóricos de fácies por meio de uma regra de truncamento de acordo com Emery (2007, p. 1191).

#### **3.5.1. VALORES GAUSSIANOS DAS INFORMAÇÕES CATEGÓRICAS**

As informações originais dos dados experimentais de tipos categóricos (fácies), de um depósito ou reservatório, devem ser transformadas em valores gaussianos. Através de um método de aceitação e rejeição na aproximação baseada em um processo sequencial, os valores gaussianos para cada ponto amostrado na área de interesse são gerados para os tipos de fácies e a seguir simulados em valores Gaussianos (Dowd et al., 2003, p. 132). Segundo Emery (2007, p. 1192), este método falha na contabilidade da informação e é unicamente uma aproximação. Além disso, o valor Gaussiano gerado não leva em consideração outras informações próximas.

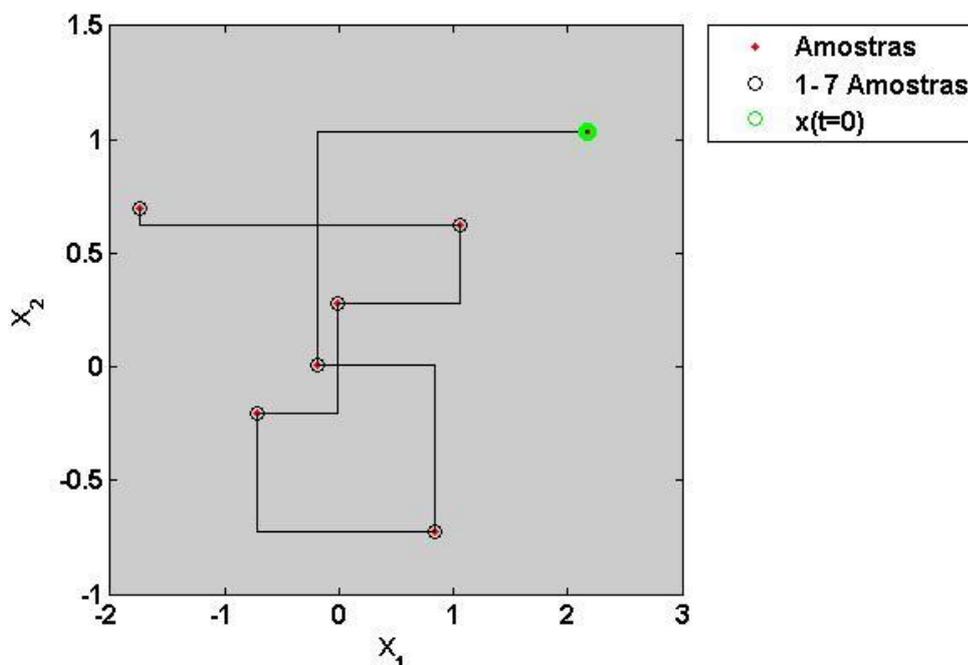
Em outros métodos de aceitação e rejeição como decomposição LU, para gerar valores Gaussianos com um modelo de covariância correto e rejeitar aqueles valores que se encontram fora de um específica intervalo, a taxa de rejeição é um fator limitante no processo, (Armstrong et al. 2003, p. 82).

Porém, estes valores gaussianos não podem ser gerados diretamente da simulação de uma distribuição Gaussiana, dado que para um aumento de intervalos a dificuldade da simulação aumenta (Armstrong et al. 2003, p. 81). Desta forma, um método indireto deve ser empregado na aproximação, que é o amostrador de Gibbs.

O método de Gibbs é uma técnica de interação baseada nas propriedades básicas de uma cadeia de Markov. Foi inicialmente proposta por Geman & Geman (1992), no processamento de imagens digitais. De acordo com Casella & George (1992, p. 167), é uma técnica para gerar variáveis de uma distribuição (marginal) indiretamente, ou seja, mediante uma sequência de valores gerados que permite aproximar uma distribuição marginal. O código de interação pode ser definido, como:

1. Gerar um vetor inicial,  $X^0 = (X_1^0, \dots, X_p^0)$ ;
2. A transição de um estado a outro, é realizado a seguir:
  - a. Gerar um número aleatório entre 1 e p. E seja i o número gerado;
  - b. Considerar a variável aleatória (V.A.) X com densidade de probabilidade  $P(X_i = x | X_j = X_j^t, i \neq j)$ . Gerar uma V.A. com a densidade descrita anteriormente. Seja  $X = x$  a V.A. gerada, então  $X^{t+1} = X_1^t, \dots, X_{i-1}^t, x, X_{i+1}^t, \dots, X_n^t$ .
3. Repetir o passo 2.a.

A Figura 3.25 ilustra um exemplo do algoritmo do amostrador de Gibbs em 2D.



**Figura 3.25.** Ilustração do amostrador de Gibbs em 2D.

Freulon (1992) e Freulon & de Fouquet (1993) adaptaram o método de Gibbs para truncar uma simulação Gaussiana (Emery, 2007, p. 1192) e (Armstrong et al., 2003, p. 82). Segundo Emery (2007, p. 1192), a adaptação do método a um caso de pluriGaussiana é uma generalização do algoritmo a seguir:

Seja  $\{x_\alpha, \alpha \in A\}$  um conjunto de informações localizadas com  $\{i_\alpha, \alpha \in A\}$  o número de fácies para cada localização. Então, para  $\alpha \in A$  um vetor aleatório independente  $Z(x_\alpha)$  que pertence a um domínio  $D_{i_\alpha}$ . O processo iterativo é dado por:

1. Um índice com distribuição aleatória uniforme  $j$  é selecionado em  $A$ . Com  $A' = A - \{j\}$ ;
2. A distribuição condicional de  $Z(x_j)$  é calculada a partir dos valores vizinhos  $\{Z(x_\alpha), \alpha \in A'\}$ . Portanto, um campo aleatório  $Z$  de valores Gaussianos independentes é construído e a distribuição condicional de  $Z_k(x_j)$  (com  $k = 1, \dots, m$ ) são Gaussianos, com média igual à krigagem simples estimada de  $\{Z_k(x_\alpha), \alpha \in A'\}$  e variância igual à variância da krigagem simples. Na estimativa dos pesos da krigagem uma única vizinhança de busca é utilizada;
3. Da distribuição condicional  $Z_k(x_j)$  um novo vetor aleatório  $Z'(x_j)$  é simulado;
4. Se  $Z'(x_j)$  pertence a  $D_{i_j}$ , então  $Z(x_j)$  é substituído por  $Z'(x_j)$ ;
5. Retornar ao passo 1, até completar o número de interações.

Segundo Armstrong et al. (2003, pp. 86 -90) e Emery (2007, p. 1192), a convergência do algoritmo descrito anteriormente é um resultado de muita importância, tendo em vista a distribuição que desejamos simular, dependendo da correlação entre os valores gaussianos e o patamar. Alguns resultados experimentais e teóricos podem ser encontrados em Armstrong et al. (2003, pp. 86 -90).

De acordo com Emery (2007, p. 1192), na prática a convergência é garantida da atualização de cada dado em uma determinada interação. Além disso, segundo mesmo autor no caso de um efeito pepita significativo (efeito pepita puro), o amostrador de Gibbs converge em uma única interação.

### 3.5.2. EXEMPLO DE APLICAÇÃO PARA GERAR VALORES ALEATÓRIOS GAUSSIANOS EM UM DETERMINADO INTERVALO E UM MODELO CORRETO DE COVARIÂNCIA

Para entender como valores gaussianos são gerados pelo método de aceitação e rejeição de Gibbs, suponhamos que ao longo de um poço discretizado em quatro unidades com igual comprimento e unicamente duas fácies  $F_1$  e  $F_2$ , com proporções de 50% e 50%, respectivamente. No caso estacionário, com um modelo de covariância exponencial com patamar igual a um e um parâmetro de alcance  $a = 6$  (Figura 3.26).



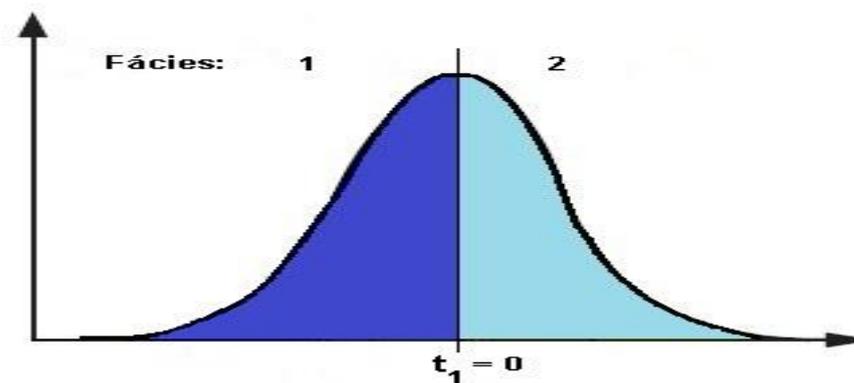
**Figura 3.26.** Poço discretizado em quatro amostras de igual comprimento.

A seguir, os parâmetros de corte são estimados das proporções acumuladas de fácies, equação (3.10), por meio da função inversa Gaussiana e  $N+1$  valores de corte são obtidos para  $N$  tipos de fácies, tabela 3.10.

**Tabela 3. 10.** Parâmetros de corte das fácies 1 e fácies 2.

$t_0$	$t_1$	$t_2$
$-\infty$	0	$\infty$

Dado que as proporções são uma área sob a curva de uma distribuição normal padrão a Figura 3.27 apresenta essas áreas. Desta maneira, intervalos específicos são definidos para cada tipo de fácies. Assim, a fácies 1 pode se encontrar entre  $(-\infty, t_1]$  e a fácies 2 entre  $(t_1, \infty)$ .



**Figura 3.27.** Distribuição normal padrão, ilustrando a área sob a curva das proporções para as fácies,  $F_1$  e  $F_2$ .

Tipos categóricos de fácies para cada amostra no poço são transformados em valores Gaussianos contínuos. Primeiro, geramos um valor aleatório com distribuição Gaussiana  $x$ , se o ponto pertence ao intervalo específico da fácies  $F_i$  (com  $i = 1$  e  $2$ ) a ser transformado, então o ponto é aceito, caso contrário é rejeitado.

Por exemplo, para a primeira amostra do poço que corresponde à fácies 1, geramos um ponto aleatório  $x = -1,4023$ . Dado que  $x$  pertence ao intervalo da fácies 1 que é  $(-\infty, t_1]$ . Então, o ponto é aceito. Este procedimento é realizado para cada amostra ao longo do poço dando como intervalo apropriado inicial  $(-1.4023, -1.4224, 0.8404, 0.1001)$ .

Desta forma, o processo iterativo do amostrador de Gibbs é realizado. A primeira etapa consiste em selecionar um índice  $j$  gerado aleatoriamente com distribuição uniforme. Para o caso prático, um caminho uniforme aleatório foi implementado  $j = (2, 3, 1$  e  $4)$ .

No segundo passo se aplica a krigagem simples para cada ponto amostrado ao longo do poço dado pelo caminho do passo anterior. Por exemplo, ao valor da amostra dois, a krigagem é usada para os outros três pontos de dados, o novo valor estimado com média e variância obtidas por krigagem simples é aceito ou rejeitado. Se o valor é aceito, o valor anterior é atualizado ao novo valor estimado. Logo, o valor a seguir é o ponto três dado pelo caminho gerado e a krigagem é usada para os valores de pontos anteriores e o novo valor atualizado. Posteriormente, depois de completar o ponto três, a krigagem simples é aplicada sucessivamente para cada ponto da amostra, seguindo o caminho gerado. Desta forma, cada ponto do poço é atualizado por krigagem simples, e as quatro interações são completadas.

**Ponto N° 2**

O valor estimado da krigagem para  $Z(x_2)$ , baseado no valor inicial  $(-1.4023, -1.4224, 0.8404, 0.1001)$  para os outros três pontos é:

$$Z_{ks}(x_2) = (-1.4023 * 0.4981) + (0.8404 * 0.5060) + (0.1001 * -0.0081) = -0.2741$$

Suponhamos que o valor aleatório normal padrão é  $-0.3386$  e o valor da variância é  $0.4099$ . Logo o valor atualizado é:

$$Z(x_2) = -0.2741 + (0.4099 * -0.3386) = -0.4129$$

onde o valor estimado é aceito e o valor anterior é atualizado a seu novo valor  $Z(x_2) = -0.4129$ .

**Ponto N° 3**

O valor estimado da krigagem para  $Z(x_3)$ , baseado nos valores iniciais de  $Z(x_1)$  e  $Z(x_4)$ , e o valor atualizado de  $Z(x_2)$  é:

$$Z_{ks}(x_3) = (-1.4023 * -0.0081) + (-0.4129 * 0.5060) + (0.1001 * 0.4981) = -0.1477$$

E o valor gerado de resíduo igual a  $0.8799$ , então o valor atualizado de  $Z(x_3)$  é:

$$Z(x_3) = -0.1477 + (0.4099 * 0.8799) = 0.2130$$

**Ponto N° 1**

O processo para calcular o valor estimado para  $Z(x_1)$  é o mesmo, aplicando a krigagem simples:

$$Z_{ks}(x_1) = (-0.4129 * 0.8929) + (0.2130 * -0.0146) + (0.1001 * -0.0695) = -0.3787$$

E o um valor gerado de  $-1.0278$ , então o valor atualizado de  $Z(x_1)$  é  $-0.9428$ .

**Ponto N° 4**

Os mesmos passos são realizados para estimar o valor de  $Z(x_4)$  via krigagem simples:

$$Z_{ks}(x_4) = (-0.3787 * -0.0695) + (-0.4129 * -0.0146) + (0.2130 * 0.8929) = 0.2225$$

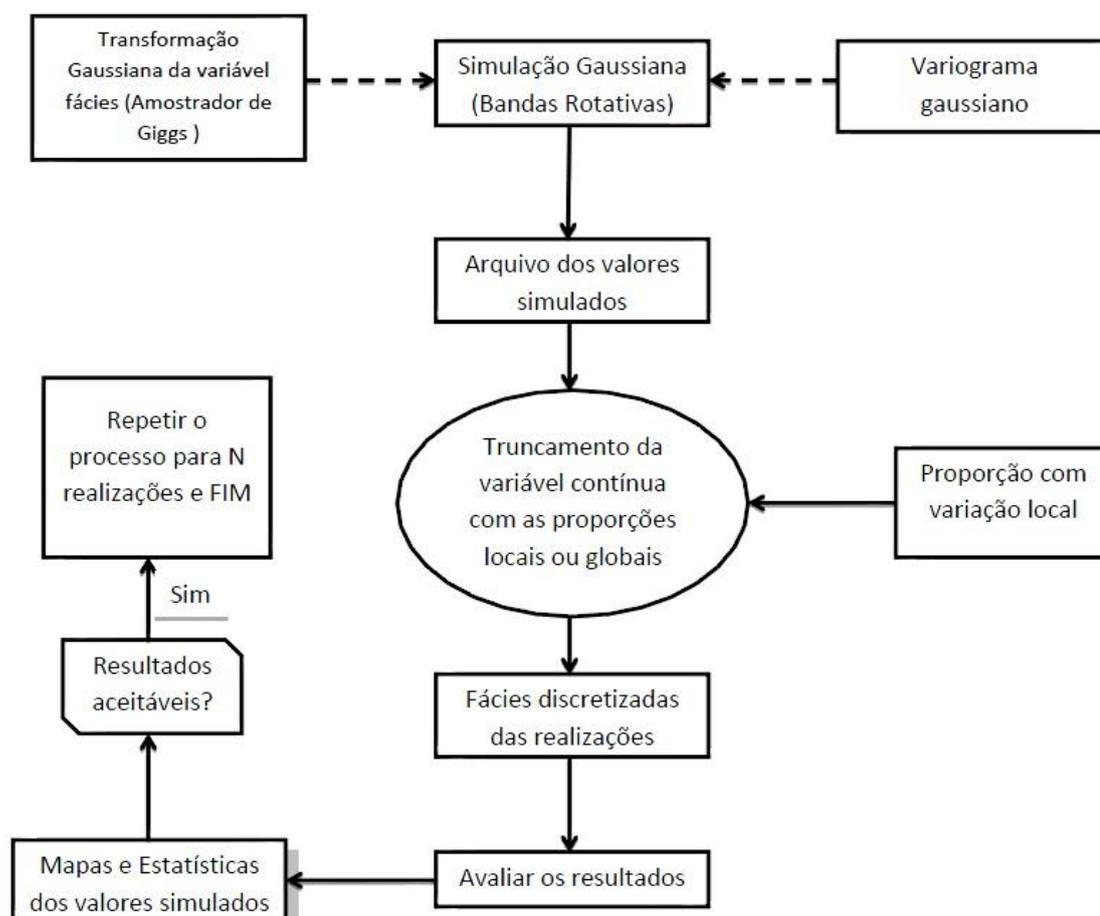
Se o valor gerado é  $0.1221$ , então o valor atualizado dado para  $Z(x_4)$  é  $0.2895$ .

A tabela 3.11 mostra os resultados das primeiras quatro iterações do método de Gibbs. Esta atualização é realizada em iterações repetidas, geralmente o número de iterações é igual a cem, mil ou cinco mil iterações.

**Tabela 3. 11.** Interações do método de Gibbs.

	Ponto	Inicial	Iter. 1	Iter. 2	Iter. 3	Iter. 4	
	1	F <sub>1</sub>	-1.4023	-1.4023	-1.4023	-0.9428	-0.9428
	2	F <sub>1</sub>	-1.4224	-0.2741	-0.2741	-0.2741	-0.2741
	3	F <sub>2</sub>	+0.8404	+0.8404	+0.2130	+0.2130	+0.2130
	4	F <sub>2</sub>	+0.1001	+0.1001	+0.1001	+0.1001	+0.2895

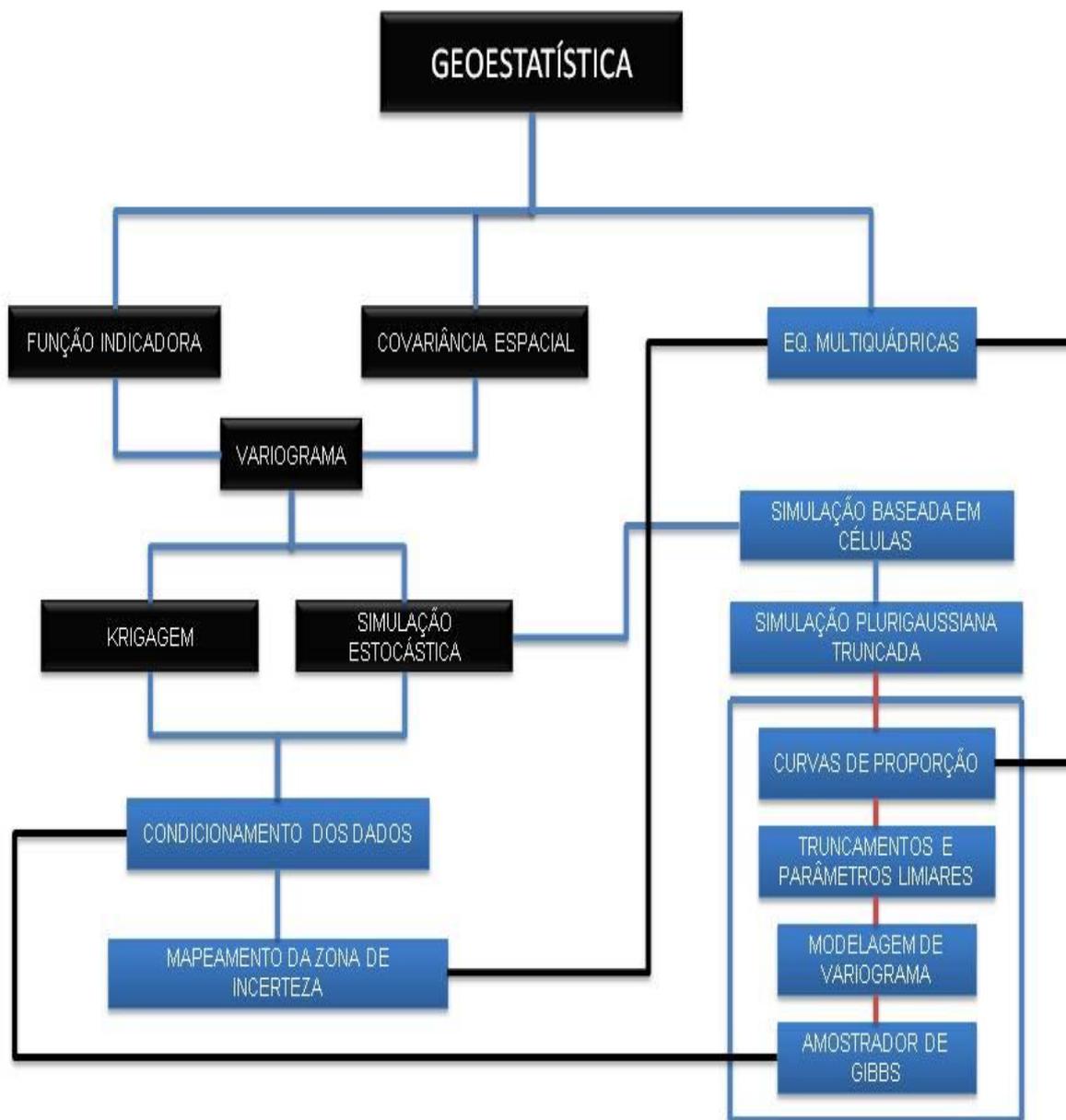
Finalmente, a simulação pluriGaussiana truncada é resumida como se ilustra na figura 3.28.



**Figura 3.28.** Fluxograma da simulação pluriGaussiana truncada.

# CAPÍTULO 4

## 4.1. MATERIAIS E MÉTODOS

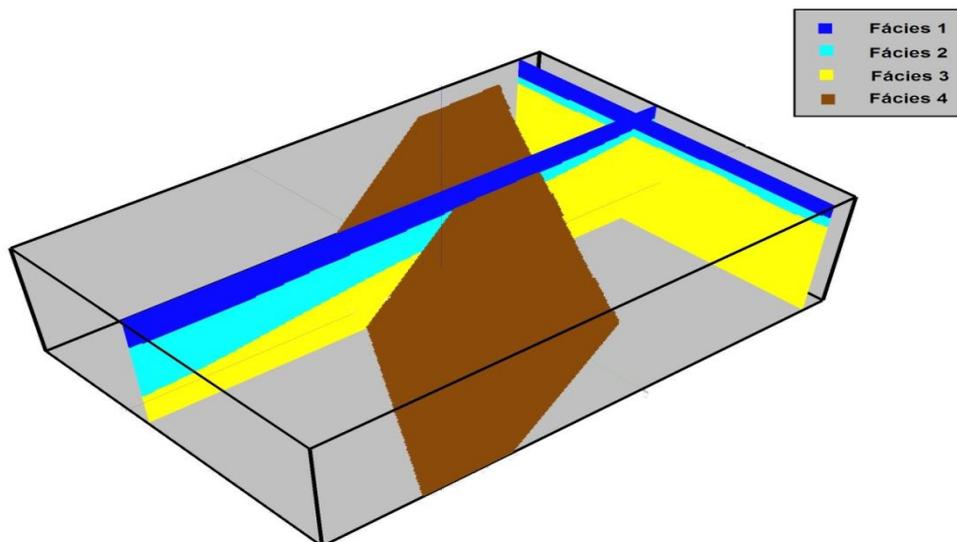


**Figura 4. 1.** Métodos utilizados no desenvolvimento do projeto. A cor preta referencia aqueles métodos a onde foi feita uma breve introdução da teoria básica e em Azul começando com a simulação baseada em células foi feita uma introdução teórica como exemplos de aplicação.

O ponto de partida é o variograma das indicadoras, que define um modelo de correlação de covariância que será utilizado na aplicação da simulação estocástica na reprodução de um atributo não amostrado, uma variável categórica. Este tipo de simulação é baseado em células, que reproduz o atributo discreto, por exemplo, as fácies. Existem dois métodos baseados em células, simulação sequencial indicadora e simulação pluriGaussiana truncada. A simulação pluriGaussiana truncada é realizada em quatro passos, o primeiro passo se refere ao cálculo das curvas de proporção, tanto na vertical como na horizontal. O segundo passo consiste na definição dos parâmetros de truncamento e limites derivados do conhecimento geológico entre as fácies. O terceiro passo é a modelagem de variograma associado a cada campo a simular por meio da tentativa e erro. E, finalmente, o amostrador de Gibbs. Sendo o amostrador de Gibbs o passo mais importante da simulação pluriGaussiana truncada. Dado que, para cada ponto discreto no domínio é simulado um valor contínuo de uma distribuição Gaussiana. Assim, como o objetivo é conhecer as características do fenômeno em estudo mediante a interpolação é aplicado o método de bandas rotativas. O método de bandas rotativas gera simulações não condicionais dos dados, que eventualmente são condicionados mediante o método da krigagem simples. Desta forma, os dados condicionados no campo contínuos são transformados ao domínio original, variável discreta, por meio da regra de truncamento. Finalmente, o mapeamento da zona de incerteza é derivado das realizações condicionais da simulação pluriGaussiana truncada, aplicando o método proposto por Yamamoto (2002, p. 151), ver o resumo na Figura 4.1.

Para testar a metodologia da simulação pluriGaussiana truncada foi considerado um modelo geológico sintético. Esse modelo é composto por duas unidades sedimentares (formações Teresina e Piramboia) e um dique de diabase associado o vulcanismo da formação Serra Geral, que corta as fácies de forma transversal na direção de N30E (azimute) e mergulho 60SE. Este modelo sintético simula uma situação geológica no domo de Anhembi, comum na Bacia de Paraná (Estado de São Paulo, Brasil).

O modelo sintético de dados é composto por 300 blocos no eixo X, 200 blocos no eixo Y e 100 blocos no eixo Z, com o total de 6.000.000 de blocos como se ilustra na Figura 4.2, com dimensões de  $1\text{ m} \times 1\text{ m} \times 1\text{ m}$ . As quatro fácies dentro do modelo sintético foram definidas como Fácies 1, Fácies 2, Fácies 3 e Fácies 4.



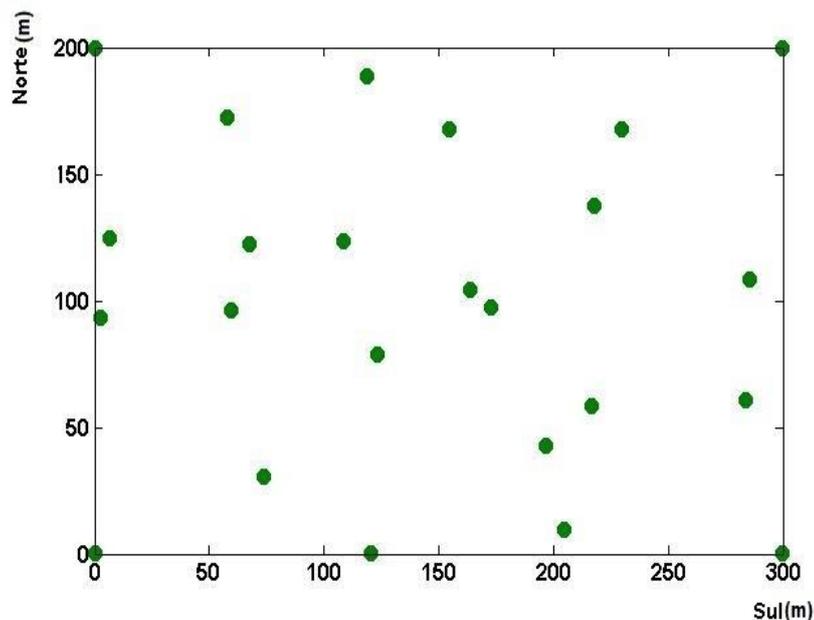
**Figura 4. 2.** Modelo geológico sintético que mostra o dique diabásico que corta de forma transversal as fácies.

Desta forma, uma amostragem aleatória estratificada de 24 poços foi realizada no modelo sintético (Figura 4.3). Assim, unicamente as informações dos 24 poços são avaliadas, com um total de 2400 amostras totalizadas de todos os poços. As proporções para cada tipo de fácies são apresentadas na Tabela 4.1.

**Tabela 4. 1.** Proporções do tipo de fácies dos dados amostrados.

Tipo	Número de dados	Porcentagem
Fácies 1	470	19.58
Fácies 2	542	22.58
Fácies 3	1127	46.96
Fácies 4	261	10.88
Total	2400	100.00

Esses dados serão usados para gerar 10 realizações da simulação pluriGaussiana truncada. O algoritmo a ser utilizado foi descrito no capítulo 3. Além disso, como um dos objetivos é proporcionar o mapeamento da zona de incerteza para cada realização da simulação pluriGaussiana truncada, os pixels simulados que apresentem uma variância maior ou igual a 0,20 serão consideradas pertencentes à zona de incerteza.



**Figura 4. 3.** Mapa de localização dos pontos amostrados do modelo sintético.

As simulações foram realizadas no Software Matlab a partir do código desenvolvido por Emery (2007), que pode ser obtido no site <http://www.sciencedirect.com/science/article/pii/S009830040700088X> e o código para o cálculo das incerteza encontra-se no anexo A.

## 4.2. ETAPAS

### 4.2.1. SIMULAÇÃO PLURIGAUSSIANA TRUNCADA

### 4.2.2. CÁLCULO DA CURVA DE PROPORÇÃO NA DIREÇÃO VERTICAL

O primeiro passo no cômputo de uma curva de proporção vertical por meio das equações multiquádricas é a definição de uma malha sobre a área a simular. Desta forma, se considerou no eixo  $X = 150\text{ m}$ ,  $Y = 100\text{ m}$  e  $Z = 50\text{ m}$ , com tamanhos de blocos igual a  $2\text{ m} \times 2\text{ m} \times 2\text{ m}$ . A procura dos vizinhos mais próximos com base nos quatro pontos mais próximos por quadrante foi realizada dentro de um raio de 70. Isto, com o objetivo de conhecer o comportamento vertical das fácies levando em consideração a influência dos poços e a localização dos mesmos.

Porém, o sucesso na aplicação das curvas de proporções depende da escolha de um nível de referência ou marco cronoestratigráfico. Desta forma, o nível de referência escolhido para o conjunto de dados sintéticos foi o topo, ou seja, a fácies de referência foi a fácies 1.

#### 4.2.3. PARÂMETROS DE CORTE

A partir do estudo das relações espaciais entre as fácies, se estabelece contatos permitidos e proibidos com o objetivo de adotar um critério na escolha da partição do domínio (Emery (2007, p. 1190) e Dowd et al. (2003, p. 125)). No presente caso, quatro fácies foram definidas, honrando os contatos entre os domínios. A fácies 4 (o dique) esta em contato com todas as outras fácies. Além disso, a fácies 3 está em contato com a fácies 2, mas não com a fácies 1. Esta situação motiva a definir duas funções aleatórias Gaussianas  $Z_1$  e  $Z_2$  na reprodução do campo em estudo e definir a seguinte regra de truncamento (Figura 4.4):

- Para  $x$  que pertence à fácies 4

$$\Leftrightarrow Z_1(x) < t_1$$

- Para  $x$  que pertence à fácies 3

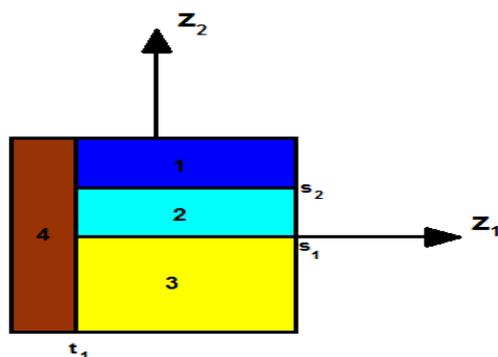
$$\Leftrightarrow \begin{cases} t_1 \leq Z_1(x) \\ Z_2(x) < s_1 \end{cases}$$

- Para  $x$  que pertence à fácies 2

$$\Leftrightarrow \begin{cases} t_1 \leq Z_1(x) \\ s_1 \leq Z_2(x) < s_2 \end{cases}$$

- Para  $x$  que pertence à fácies 1

$$\Leftrightarrow \begin{cases} t_1 \leq Z_1(x) \\ s_2 \leq Z_2(x) \end{cases}$$



**Figura 4. 4.** Regra de truncamento do campo em estudo.

#### **4.2.4. VARIOGRAMA E ANÁLISE ESTRUTURAL**

A partir dos parâmetros inferidos no item anterior (parâmetros de corte e regra de truncamento), um modelo de variograma de covariância normal deve ser ajustado para cada função Gaussiana  $Z_1$  e  $Z_2$ , baseado na estratégia de tentativa e erro aos variogramas de indicadoras.

#### **4.2.5. VALORES GAUSSIANOS DAS INFORMAÇÕES CATEGÓRICAS**

A seguir, os dados categóricos das fácies para cada tipo devem ser transformados em valores gaussianos por meio de um correto modelo de covariância definido na etapa anterior e um intervalo apropriado definido pelos parâmetros de corte, com o objetivo de reproduzir a continuidade espacial de um campo gaussiano. A transformada é dada pelo amostrador de gibbs com base em 5000 mil realizações. Uma vez gerado este conjunto de valores Gaussianos, as realizações condicionais das funções Gaussianas contínuas, mediante o método de bandas rotativas, são obtidas. Desta forma, foram geradas 1000 linhas sobre uma esfera unitária. Finalmente, as realizações contínuas de funções Gaussianas são transformadas em tipos categóricos de fácies por meio da regra de truncamento definida para o conjunto de dados sintéticos (Figura 4.4).

#### **4.2.6. MAPEAMENTO DA ZONA DE INCERTEZA DAS SIMULAÇÕES ESTOCÁSTICAS**

Na etapa final foi analisado o conjunto das 10 realizações para determinar a zona de incerteza associada aos 4-tipos de variáveis categóricas (Yamamoto & Landim, 2013, p. 172). Desta forma, para cada célula a estimar a proporção do k-ésimo tipo para as 10 realizações da simulação estocástica é dada pela equação (2.1), onde o tipo mais provável é o valor de maior proporção (equação (2.2)):

$$p_{\max} = \max[p(x_0; k)] \text{ para } k = 1, \dots, 4$$

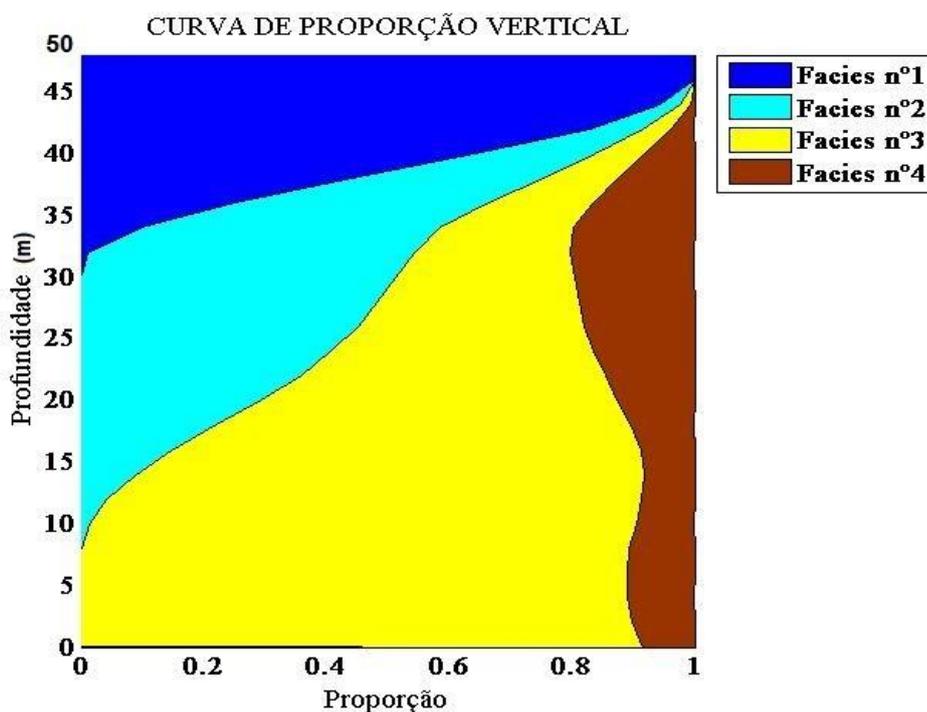
Nesse sentido, o valor do tipo mais provável é atribuído se a variância for menor que 0,20, caso contrário o valor simulado está na zona de incerteza (Yamamoto & Landim, 2013, p. 172).

## CAPÍTULO 5

### 5.1. RESULTADOS E DISCUSSÃO

Nesse capítulo são apresentados principalmente os resultados obtidos da simulação pluriGaussiana truncada e o mapeamento da zona de incerteza conforme (Yamamoto & Landim, 2013, p. 172).

A curva de proporção vertical foi calculada nível a nível assumindo-se que todos os poços estão referenciados a um mesmo horizonte (Fácies 1). A estimativa foi feita a partir do método das equações multiquádricas levando em consideração o posicionamento e influência dos poços dentro de um raio de 70 m. Desta forma, eventos prováveis e não prováveis foram considerados na estimativa da curva de proporção como se vê na Figura 5.1.



**Figura 5.1.** Curva de proporção vertical das fácies.

Na Figura 5.1 se observa que para cada nível na direção vertical no eixo X mostrem as proporções estimadas. Por exemplo, no nível cem há 100% de fácies 1, no nível vinte e cinco há aproximadamente 40% de fácies 2, 40% de fácies 3 e 20% de fácies 4. Desta forma, se

verifica que existe uma variação na distribuição das fácies, considerando-se a distribuição das fácies como não estacionárias.

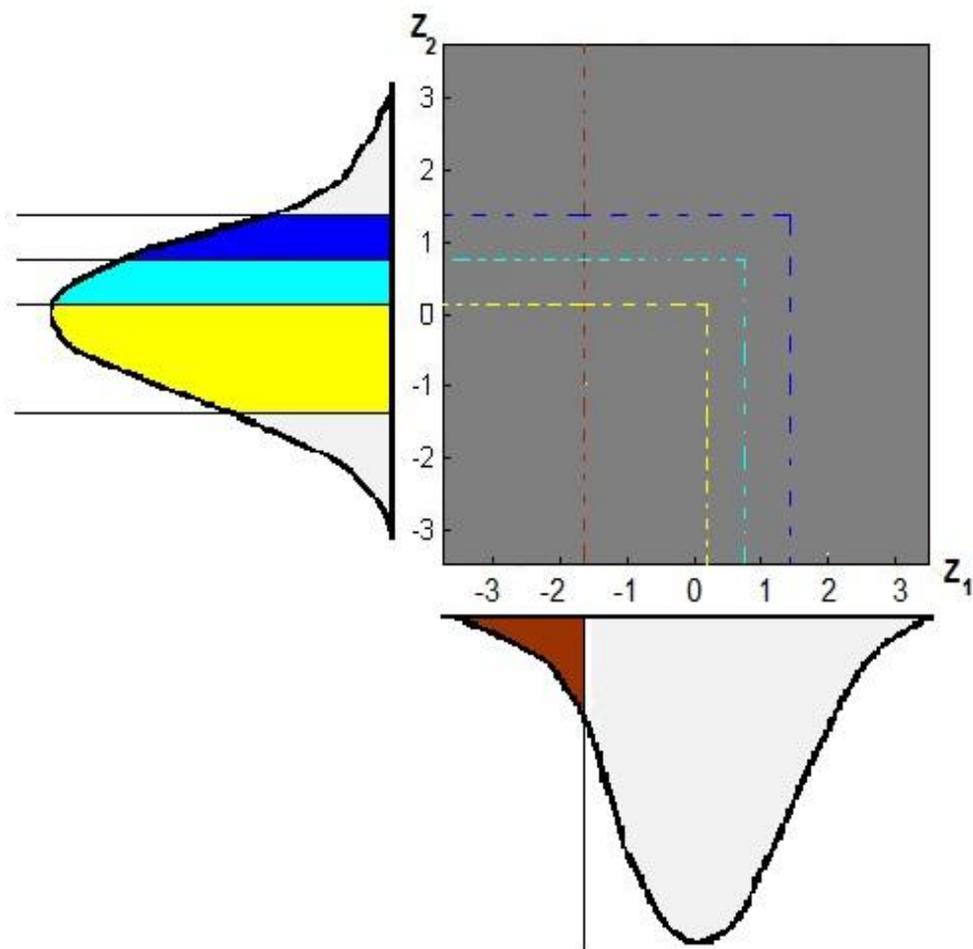
Os parâmetros de corte  $t_1$ ,  $s_1$  e  $s_2$  são estimados das proporções globais do domínio em consideração. Porém, por meio das proporções dos dados experimentais na tabela 4.1, computam-se os parâmetros de corte (Tabela 5.1):

**Tabela 5. 1.** Parâmetros de corte derivados dos dados amostrais.

Parâmetros de corte		
$Z_1$	$Z_2$	
$t_1$	$s_1$	$s_2$
-1.2816	0.0557	0.7647

Estes parâmetros de corte derivados de uma regra de truncamento dividem a área sob a curva de uma distribuição normal padrão como se observa na Figura 5.2. Esta divisão é uma discretização da área por meio das duas funções aleatórias Gaussianas. Na primeira função Gaussiana, a área sob a curva da distribuição está discretizada em duas partes; a primeira área corresponde à fácies 4 com 10% de proporção e o restante às fácies 1, 2 e 3. A seguir, a segunda função Gaussiana é discretizada pelos parâmetros de corte  $s_1$  e  $s_2$ . Desta forma, se aprecia que a área com cores (amarelo, Cian e azul) corresponde ao complemento, no caso 90 % restante das fácies.

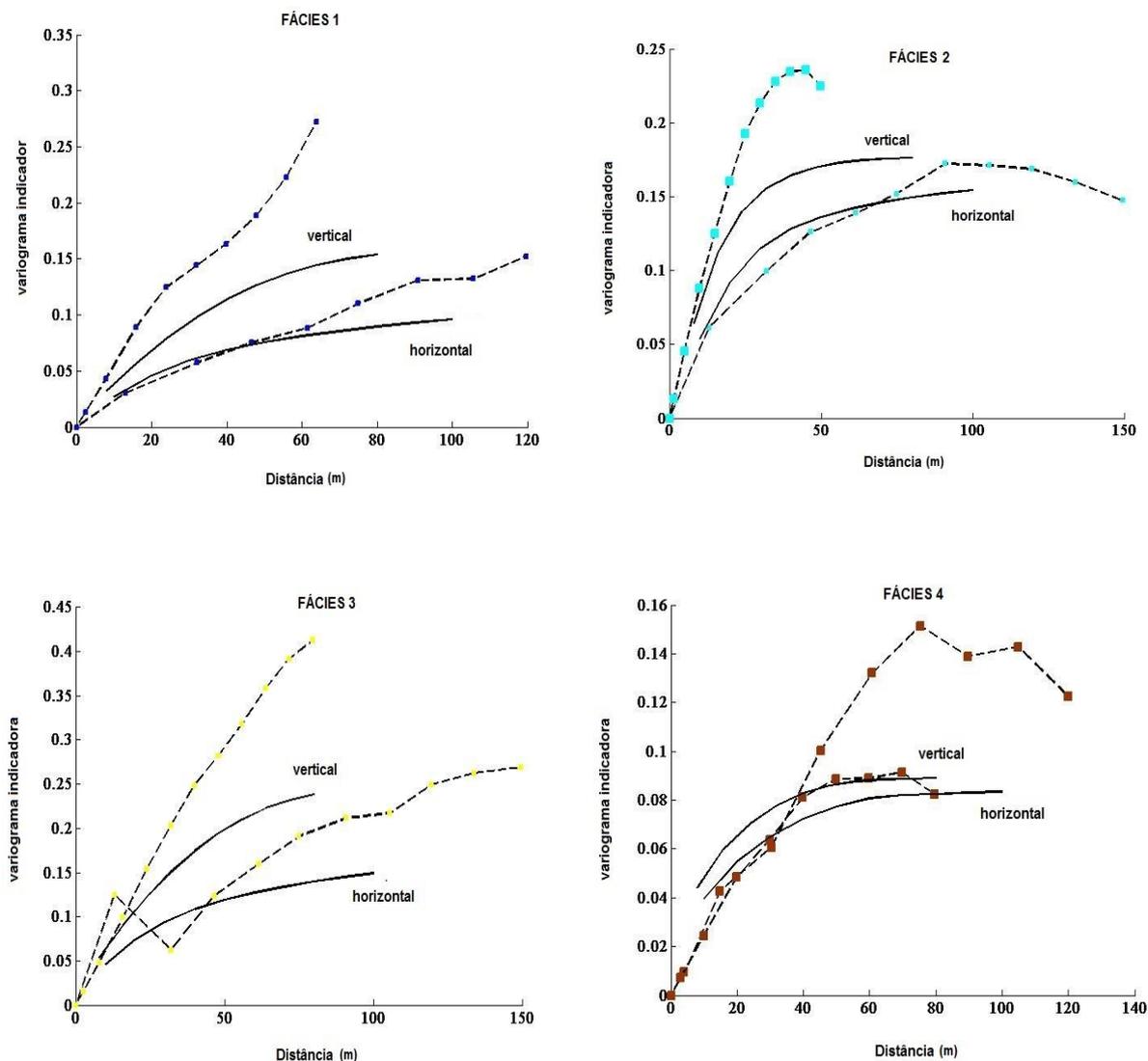
Prosseguindo, os variogramas derivados dos parâmetros de corte e da regra de truncamento por meio da estratégia de tentativa e erro (Emery & González 2007, p. 33) são apresentados na Figura 5.3 e os parâmetros correspondentes aos modelos de variogramas na tabela 5.2.



**Figura 5.2.** Descretização da área de uma distribuição normal padrão, dada por cada parâmetro de corte para cada função aleatória, no caso  $Z_1$  e  $Z_2$ .

**Tabela 5. 2.** Parâmetros de modelos de variograma dos campos aleatórios gaussianos usados na simulação pluriGaussiana.

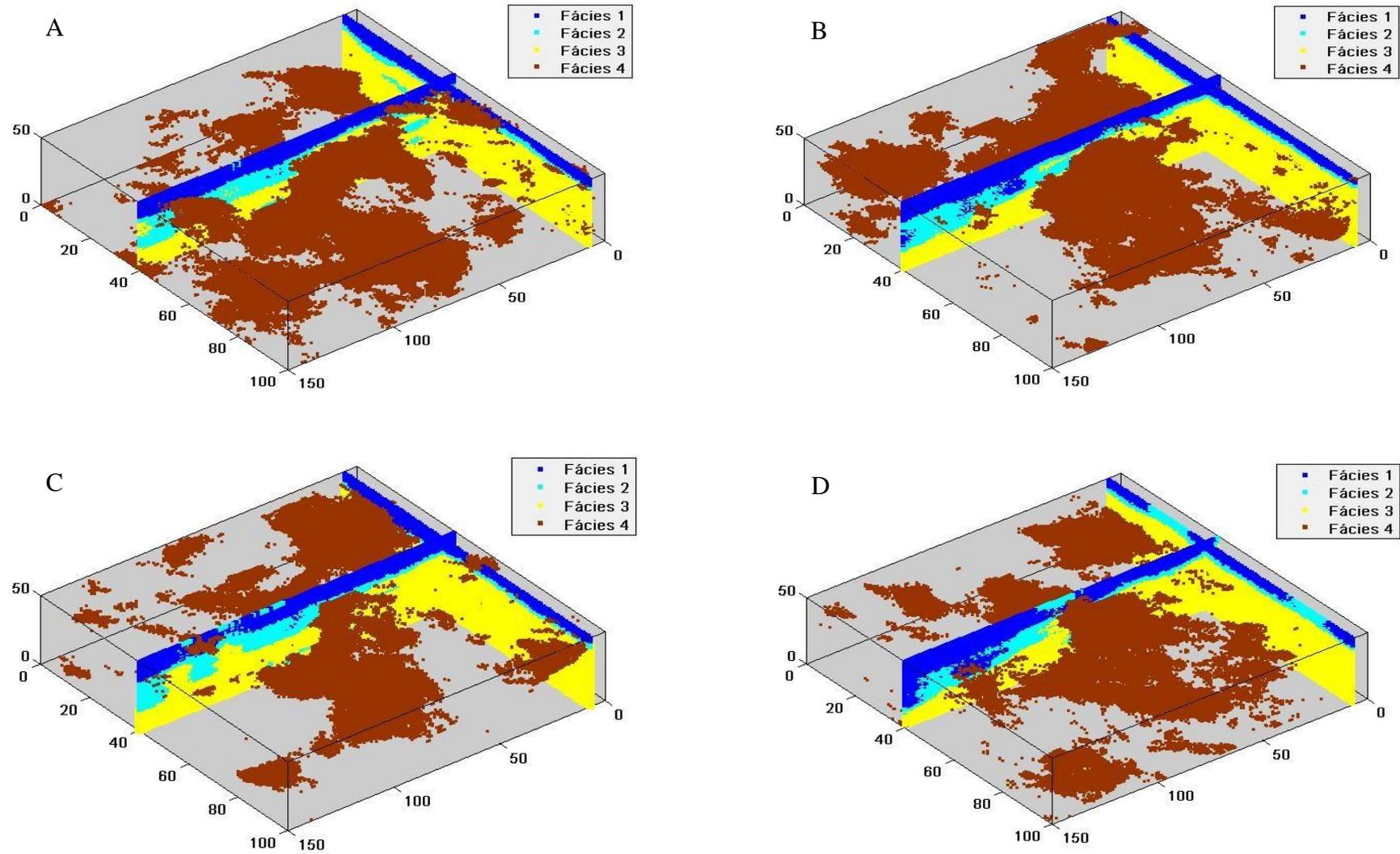
	Tipo de Variograma	Patamar	Amplitude horizontal	Amplitude vertical	Efeito pepita
Primeira função Gaussiana	Exponencial	0.03	50	25	0
	Esférico	0.67	70	60	
Segunda função Gaussiana	Exponencial	0.30	200	35	0
	Gaussiano	0.10	27	54	
	Gaussiano	0.60	1e20	50	
	Gaussiano	0.30	90	50	



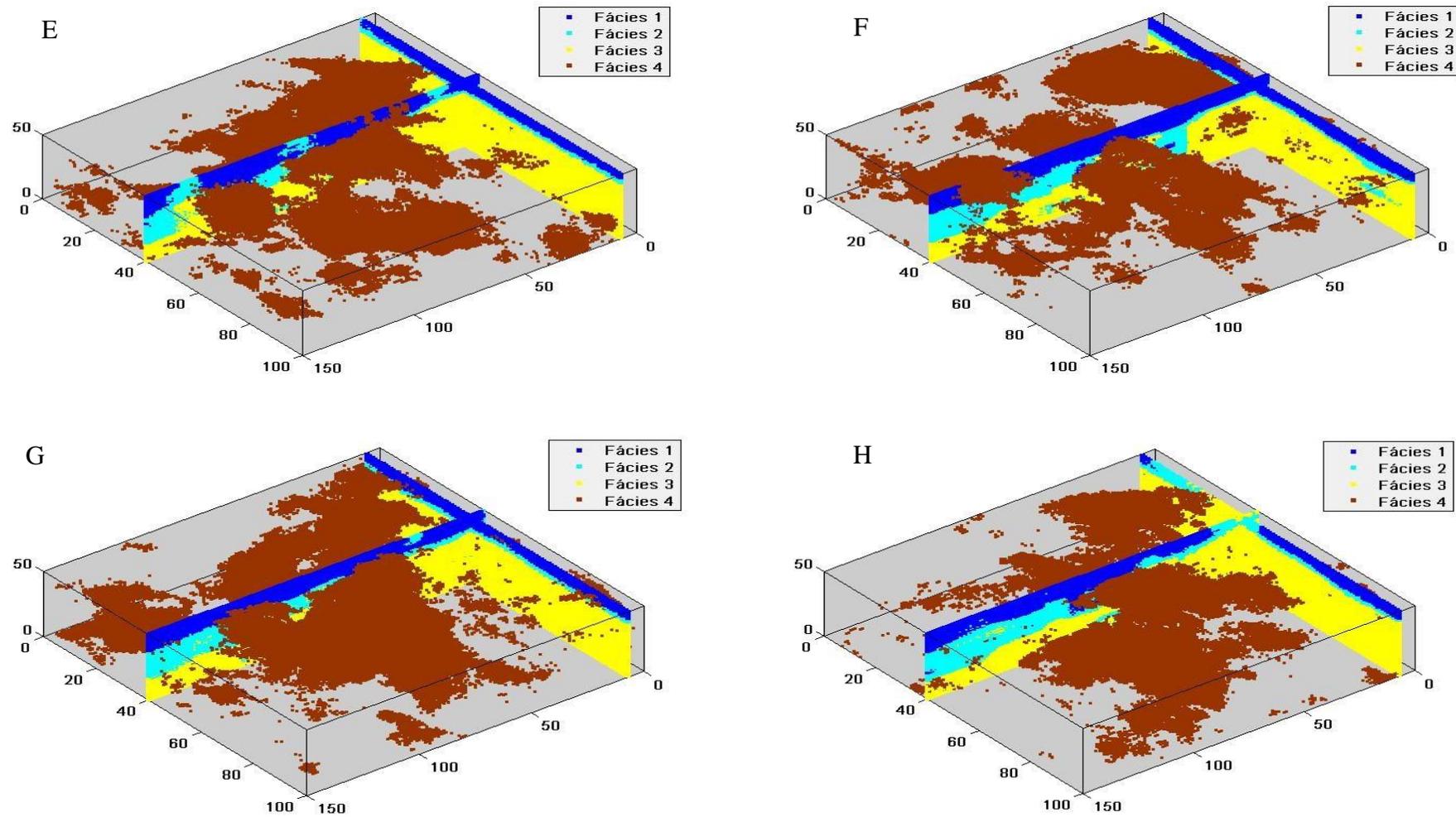
**Figura 5.3.** Variogramas das indicadoras dos dados experimentais (quadrados e linhas pontilhadas) e modelado (linhas solidas) nas direção vertical e horizontal.

Nos variogramas indicadoras da Figura 5.3 se observa que perto da origem, os variogramas têm uma tendência linear. Lembrando-se que  $p$  é a proporção do domínio espacial, então o patamar do variograma da indicadora é  $p(1 - p)$ .

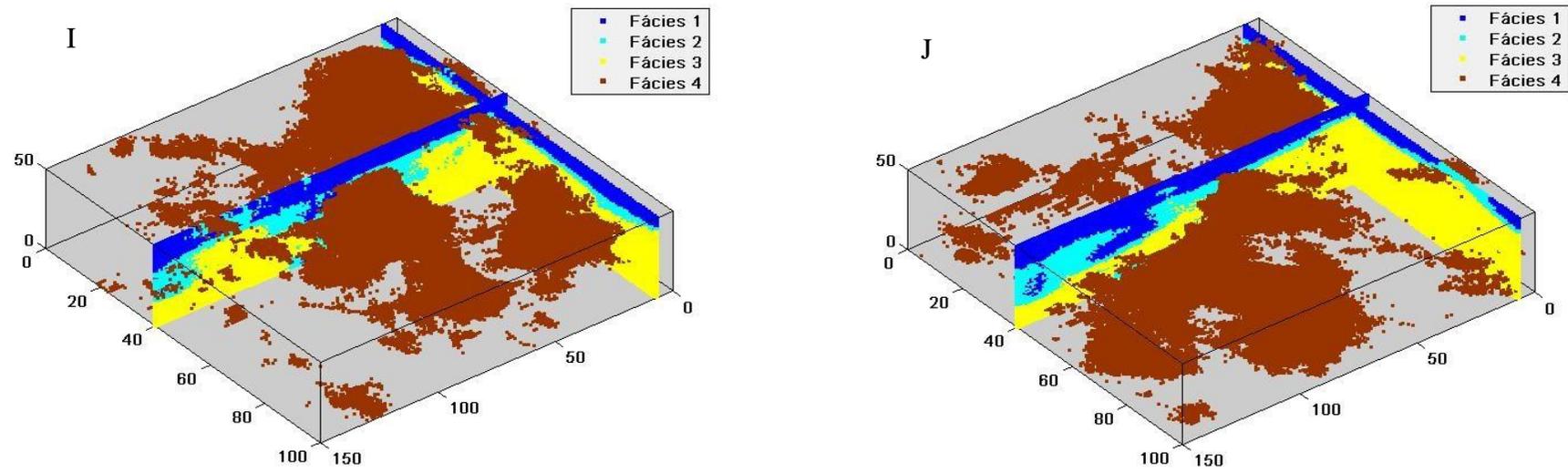
O programa PLURISIM.M foi usado para criar realizações condicionais das fácies presentes no modelo sintético. A Figura 5.4 ilustra as dez simulações condicionais dos dados. Desta forma, para cada uma das quatro figuras apresentadas, toma-se um plano horizontal na profundidade de  $Z = 10$  e as proporções associadas para cada imagem são calculadas (Figura 5.7 e Tabela 5.3).



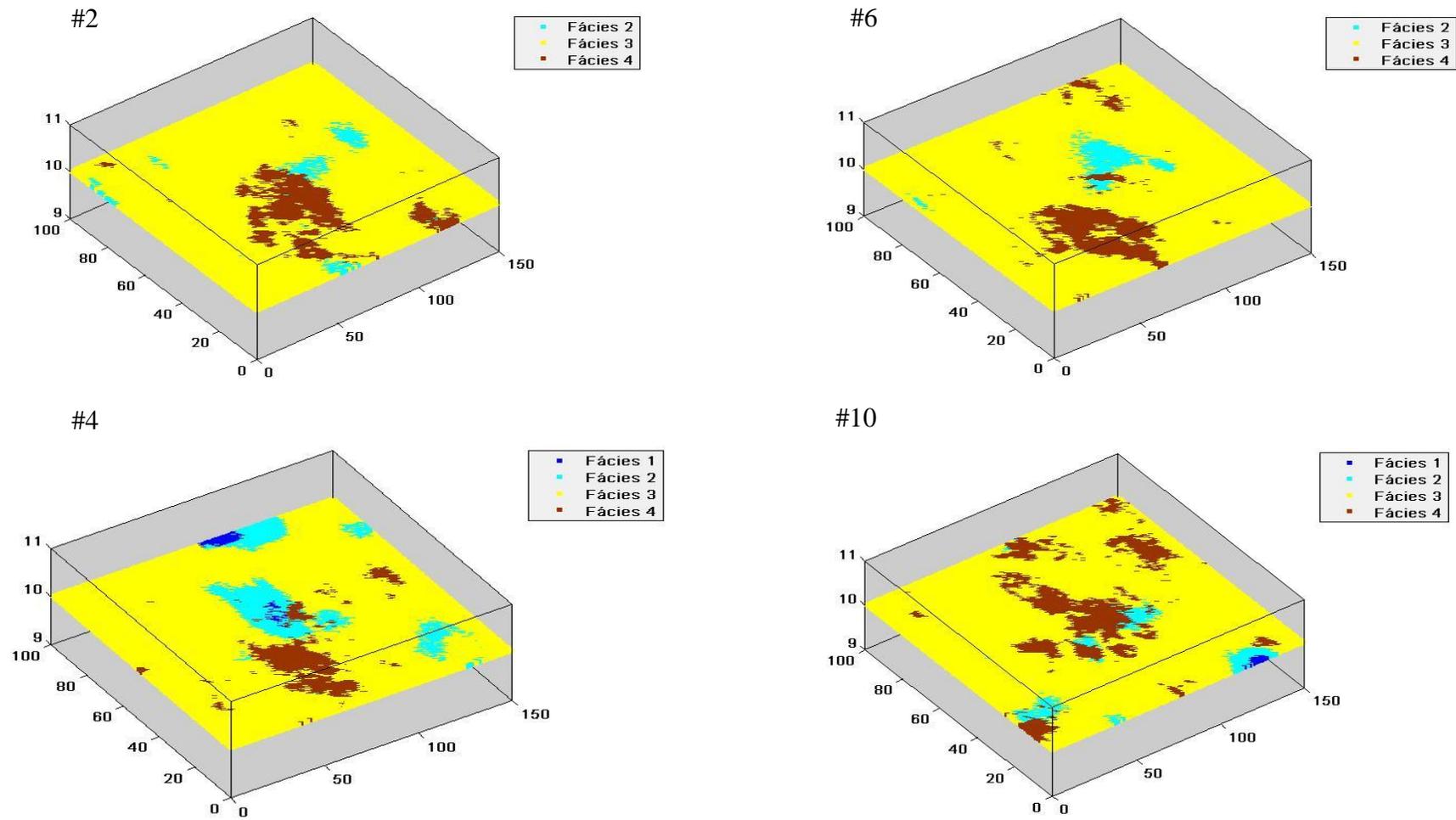
**Figura 5.4.** Realizações derivadas da simulação pluriGaussiana truncada. A) realização 1, B) realização 2, C) realização 3 e D) realização 4.



**Figura 5.5. (continuação)** Realizações derivadas da simulação pluriGaussiana truncada. E) realização 5, F) realização 6, G) realização 7 e H) realização 8.



**Figura 5.6. (continuação)** Realizações derivadas da simulação pluriGaussiana truncada. I) realização 9 e J) realização 10.



**Figura 5.7.** Perfis horizontais na profundidade de  $Z = 10$  para cada uma das realizações selecionadas aleatoriamente (#2, #6, #4 e #10).

**Tabela 5.3.** Proporções computadas dos perfis das realizações #2, #6, #4 e #10, respectivamente.

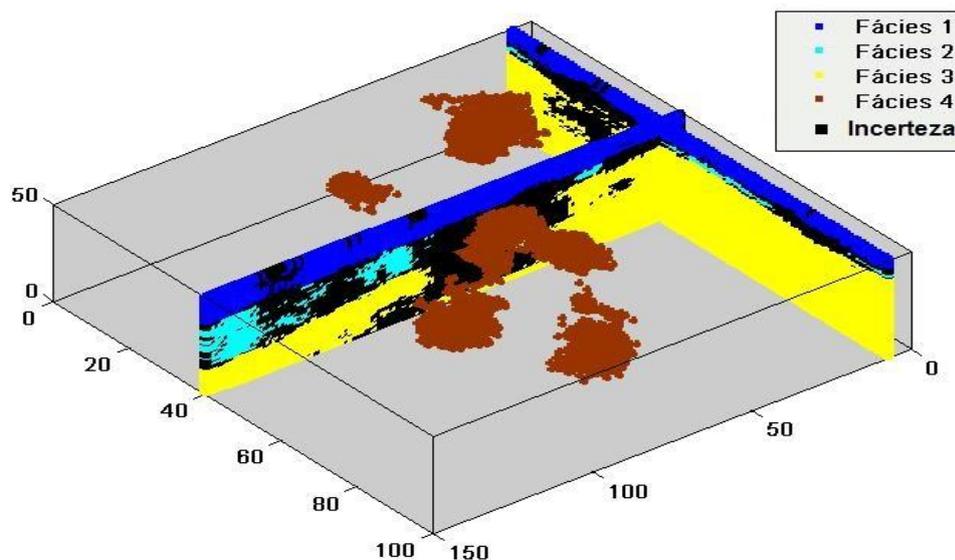
Tipo	Realização							
	#2		#6		#4		#10	
	Nº dados	Percentage m	Nº dados	Percentage m	Nº dados	Percentage m	Nº dados	Percentage m
Fácies 1	0	0	120	0.80	0	0	23	0.15
Fácies 2	399	2.66	1404	9.36	516	3.44	520	3.47
Fácies 3	13036	86.91	12468	83.12	13080	87.20	12411	82.74
Fácies 4	1565	10.43	1008	6.72	1404	9.36	2046	13.64
Total	15000	100.00	15000	100.00	15000	100.00	15000	100.00

Para todas as proporções do perfil da (tabela 5.3) se verifica que o valor simulado na profundidade  $Z = 10$  está muito próximo do valor estimado nas curvas de proporções na Figura 5.1.

A simulação pluriGaussiana truncada é um método fácil de aplicação na prática, leva em consideração diversos tipos de informação como: contato entre os domínios (regra de truncamento), perfil de fácies em função da profundidade (curvas de proporção vertical), continuidade geológica (variogramas das indicadoras) e as informações dos dados (dados condicionais).

A partir das 10 realizações por meio da simulação pluriGaussiana truncada, se determina a zona de incerteza associada aos 4-tipos de variáveis categóricas como se vê na Figura 5.5.

Pode se observar que na Figura 5.8, a zona de incerteza encontra-se entre os contatos das fácies. O dique (Fácies 4) não é bem reproduzido, mas os pontos simulados (Figura 5.8) são reduzidos àqueles que tem maior probabilidade de pertencer à fácies 4. Além disso, a Figura 5.8 mostra a incerteza aumenta se os contatos se encontram nos contatos proibidos entre fácies.



**Figura 5.8.** Mapeamento da zona de incerteza a partir das 10 realizações.

Na tabela 5.4, se apresentam os resultados das proporções obtidas para todas as fácies após procedimento do cômputo da zona de incerteza como também a proporção desta zona.

**Tabela 5. 4.** Proporções das fácies após do procedimento do cômputo da zona de incerteza.

<b>Proporção em percentagem</b>				
<b>Fácies 1</b>	<b>Fácies 2</b>	<b>Fácies 3</b>	<b>Fácies 4</b>	<b>ZI</b>
15.89%	7.80%	42.41%	0.82%	33.08%

## CAPÍTULO 6

### 6.1. CONCLUSÕES

A simulação pluriGaussiana truncada é um método estocástico para variáveis categóricas que reproduzem a variabilidade do domínio ( fácies), que honra as restrições impostas como contatos e ordem das fácies, em contraste à simulação indicadora sequencial. É uma ferramenta complexa de aplicação, mas muito útil na prática. A simulação pluriGaussiana truncada fornece a descrição probabilística de um domínio geológico, reservatório de petróleo ou uma formação sedimentar, como as hidrofácies de um aquífero, tipos de solo etc. Este tipo de simulação contribui no controle de variáveis quantitativas de interesse como a porosidade, permeabilidade, concentração etc. A partir das realizações independentes geradas pelo método estudado, possibilitou-se o mapeamento da zona de incerteza derivadas das L realizações. Esta zona de incerteza está associada aos contatos entre as fácies, diminuindo o campo simulado aqueles valores que tem maior probabilidade de ser obtidos, bem como uma zona de incerteza.

As curvas de proporção computadas por meio das equações multiquádricas ajudam na diminuição do cálculo de parâmetros na estimativa desta ferramenta de grande importância, dado que reflete a distribuição das fácies que ajudam na caracterização do modelo em estudo. Ademais, leva em consideração a influência dos poços e sua localização, bem como pontos amostrados que pertencem à zona de incerteza.

As curvas de proporções são estimadas dos dados amostrados existentes e é utilizada no método da pluriGaussiana truncada como uma imagem de “treinamento” que ajuda na interpretação e verificação dos valores simulados.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ARMSTRONG, M. et al. Plurigaussian simulations in Geosciences. Springer, 2003. 149 p.
- BRAGA, M. S. S. Construção de matrizes de proporção de fácies para a modelagem geológica 3D de reservatórios deltaicos da Bacia do Recôncavo (BA). Universidade Estadual de Campinas. Campinas-São Paulo, p. 115. 1998.
- CASELLA, G.; GEORGE, E. I. Explaining the Gibbs Sampler. JSTOR, v. 46, n. 3, p. 167-174, Agosto 1992.
- CHIHAI, H.; DE MARSILY, G. Simulating Non-Stationary seismic facies distribution in a prograding shelf environment. Oil & Gas Science and Technology - Rev. 2FP, v. 64, n. 4, p. 17, 2009.
- CLARK, I. Practical geostatistics. Elsevier Science & Technology, 1979. 141 p.
- DEUTSCH, C. V. Geostatistical reservoir modeling. OXFORD UNIVERSITY PRESS, 2002. 376 p.
- DEUTSCH, C. V.; JOURNEL, A. G. GSLIB: Geoestatistical software library and user's guide. [S.l.]: Oxford University Press, 1992. 340p p.
- DOWD, P. A.; PARDO-IGÚZQUIZA, E.; XU, C. Plurigau: a computer program for simulating spatial facies using the truncated plurigaussian method. Computers & Geosciences, v. 29, n. 2, p. 123-141, Março 2003.
- EMERY, X. Simulation of geological domains using the plurigaussian model: New developments and computer programs. Computers & Geosciences, v. 33, n. 9, p. 1189-1201, Setembro 2007.
- EMERY, X.; GONZALEZ, K. E. Incorporating the uncertainty on geological boundaries into mineral resources evaluation. Computer & Geosciences, v. 69, n. 1, p. 29-38, Janeiro 2007.
- EMERY, X.; LANTUÉJOUL, C. TBSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. Computers & Geosciences, v. 32, n. 10, p. 1615-1628, Dezembro 2006.
- FOLLE, D. Análise e aplicações da geoestatística no contexto geológico-geotécnico urbano. Universidade Federal do Rio Grande do Sul. Porto Alegre, RS, p. 232. 2009.

FREULON, X. Conditionnement du modèle gaussien par des inégalités ou des randomisées. ENSMP. p. 168. 1992.

FREULON, X.; DE FOUQUET, C. Remarques sur la pratique des bandes tournantes à trois dimensions. Cahiers de géostatistique, Ecole des Mines de Paris, Fontainebleau, p. 101-117, 1991.

FREULON, X.; DE FOUQUET, C. Conditioning a Gaussian model with inequalities. In: Soares ed. Dordrecht: Kluwer, v. 1, n. 92, p. 201-212, 1993.

GALLI, A. et al. The pros and cons of the truncated Gaussian method. Kluwer Academic Publishers, v. 7, p. 217-233, 1994.

GEOVARIANCES. Compute 3D facies proportions with isatis. p. 1-4.

JOURNEL, A. G. Non-parametric estimation of spatial distribution. Math. Geology, v. 15, p. 445-468, 1983.

KYRIAKIDIS, P. C.; DEUTSCH, C. V.; GRANT, M. L. Calculation fo the normal score variogram used for truncated Gaussian lithofacies simulation: Theory and FORTRAN code. Computers & Geosciences, v. 25, n. 2, p. 161-169, Março 1999.

LE LOC'H, G.; GALLI, A. Truncated plurigaussian method: Theoretical and practical points of view. In: BAAF, E. Y. Geostatistics Wollongong &96. Springer Science & Business Media, 1997, v. 1, 1997. p. 1333.

LEUANGTHONG, O. K. D.; DEUTSCH, C. V. Stepwise conditional transformation for simplified cosimulation of reservoir properties. CCG Paper 2000-106, University of Alberta, Edmonton, Canada, 2000.

MATHERON, G. The intrinsic random functions and their applications. Advances in Applied Probability 5, p. 439-468, 1973.

OLEA, R. A. Geostatistics for Engineers and Earth Scientists. Boston: Kluwer Academic Plublishers, 1999. 303 p.

PYRCZ, M. J.; DEUTSCH, C. V. Geostatistical Reservoir Modeling. 2<sup>a</sup> Edição. ed. [S.l.]: Oxford University Press, 2014. 433 p.

RAVENNE, C. Stratigraphy and Oil: A Review. Part 2. Characterization of reservoirs and sequence stratigraphy: Quantification and Modeling. Oil & Gas Science and Technology, v. 57, n. 4, p. 311-340, 2002.

REMACRE, A. Z.; NORMANDO, M. N.; SANCEVERO, S. S. Krigagem das proporções utilizando a krigagem da média: uma ferramenta auxiliar na modelagem de reservatórios. *Revista Brasileira de Geociências*, v. 38, n. 1, p. 82-87, Março 2008.

YAMAMOTO, J. K. Ore reserve estimation using radial basis functions. *Revista do Instituto Geológico, São Paulo*, v. 23, n. 1, p. 25-38, 2002.

YAMAMOTO, J. K. et al. Mapping an uncertainty zone between interpolated types of a categorical variable. *Computers & Geosciences*, v. 40, p. 146-152, Setembro 2012.

YAMAMOTO, J. K. et al. Post-processing for uncertainty reduction in computed 3D geological models. *Tectonophysics*, v. 633, p. 232-245, 2014.

YAMAMOTO, J. K. et al. Post-processing of sequential indicator simulation realizations for modeling geologic bodies. *Computational Geosciences*, p. 1-10, Janeiro 2015.

YAMAMOTO, J. K.; LANDIM, P. M. B. *Geoestatística conceitos e aplicações*. São Paulo: Oficina de Textos, 2013. 215 p.

## ANEXO A

```

% Calculo da incerteza em Matlab.

load 'plurisim_RT.out'; - Arquivo de entrada

nx = 150; ny = 100; nz = 50;

nfacies = 4;

indicadora = zeros(nx*ny*nz,nfacies); - indicadora, tamanho igual a nx*ny* nz

L = nrealiz; - número de realizações

prob = zeros(nx*ny*nz,1);

for i = 1:nx*ny*nz

    for j = 1:nfacies

        indicadora(i, j) = sum(plurisim_RT(i, :) == j);

    end

    maxl = max(indicadora(i,:));

    p_maxl = maxl/L;

    var = (1 - p_maxl)*p_maxl;

    index = find(indicadora(i, :) == maxl);

        n = length(index);

        if var < 0.20

            if length(index) < 2

                prob(i) = find(indicadora(i,:) == maxl);

            end

        else

            prob(i) = 5;

        end

    end

end

```