

ANA BEATRIZ TOZZO MARTINS

**ANÁLISE GEOESTATÍSTICA DE DADOS
COMPOSICIONAIS**

**CURITIBA
FEVEREIRO 2010**

ANA BEATRIZ TOZZO MARTINS

**ANÁLISE GEOESTATÍSTICA DE DADOS
COMPOSICIONAIS**

Tese apresentada ao Curso de Pós-graduação em Métodos Numéricos em Engenharia do Setor de Tecnologia do Centro de Estudos de Engenharia Civil Professor Inaldo Ayres Vieira da Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutor em Ciências.

Orientador: Prof. PhD. Paulo Justiniano Ribeiro Jr.

**CURITIBA
FEVEREIRO 2010**

TERMO DE APROVAÇÃO

ANA BEATRIZ TOZZO MARTINS

ANÁLISE GEOESTATÍSTICA DE DADOS COMPOSICIONAIS

Tese aprovada como requisito parcial para obtenção do grau de Doutora em Ciências, pelo Programa de Pós-Graduação em Métodos Numéricos em Engenharia do Setor de Tecnologia do Centro de Estudos de Engenharia Civil Professor Inaldo Ayres Vieira da Universidade Federal do Paraná, pela seguinte banca examinadora:

Prof. PhD. Paulo Justiniano Ribeiro Jr.
Universidade Federal do Paraná

Prof. Dra. Silvia Shimakura
Universidade Federal do Paraná

Prof. Dr. Antonio Carlos Andrade
Gonçalves
Universidade Estadual de Maringá

Prof. Dra. Alexandra Schmidt
Universidade Federal do Rio de Janeiro

Prof. Dr. Joel Mauricio Correa da Rosa
Universidade Federal do Paraná

Curitiba, 20 de fevereiro de 2010

A meus pais, Joaquim e Edila e irmãs Luisa e Maria do Carmo.

AGRADECIMENTOS

Agradeço a Jesus que é o caminho, a verdade e a vida (João, 14-6), que me levou a caminhar este tempo todo dedicado aos estudos, descobrindo a cada dia na sua verdade o que eu deveria escrever e dando-me a graça da vida.

Agradeço ao Prof. Paulo pela orientação, dedicação na formação de seus alunos e empenho em fazer do LEG um excelente ambiente de pesquisa.

Aos meus pais, Joaquim e Edila, Luisa e Maria do Carmo (irmãs), Marcos e Júnior (cunhados), Caio, Ana Luisa, Camila e Luana (sobrinhos), pelas orações e apoio recebido durante este período.

A Terezinha Guedes pela amizade, apoio e disponibilidade em me ajudar.

A amiga de fé Alaíde Brita e sua mãe pelas orações.

Ao Edson A.A. Silva pela amizade, paciência nos meus primeiros passos com Linux, R e geostatística, e leitura do trabalho com preciosas contribuições.

Ao Wagner H. Bonat, pela paciência, companheirismo nos estudos e parceria em trabalhos.

À Vanderly Janeiro, Robson Rossi, Gledson Picharski, Bruno Fonseca, Melissa Souza, Prof. Raul Prado, Elias Krainski que contribuíram de alguma forma na elaboração da tese.

As amigas de curso Bernadete Brandão, Vanessa Ales, Rosângela Villwock e Roberta Suero pelos momentos difíceis e alegres que passamos juntas.

Aos colegas do LEG e CESEC.

A Maristela pela paciência e carinho com que nos trata.

Aos professores do PPGMNE.

Aos colegas do Departamento de Estatística da UEM.

Ao Prof. Antônio Carlos (DAG/UEM) por disponibilizar dados de sua tese enriquecendo a pesquisa.

À esposa do Prof. Paulo, Prof. Silvia e seus filhos Paulo Neto e Luca pelos momentos em que o privei de seu convívio familiar.

À Capes - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior por seu apoio financeiro.

Agradecimento à CAPES pelo apoio financeiro. Esse trabalho foi parcialmente financiado pela FINEP projeto CT-INFRA/UFPR.

Sumário

Lista de Figuras	vii
Lista de Tabelas	x
Lista de Siglas	xii
Resumo	xiii
Abstract	xiv
1 INTRODUÇÃO	1
2 REVISÃO DA LITERATURA	5
2.1 MODELO GEOESTATÍSTICO GAUSSIANO	5
2.1.1 Definição do Modelo	5
2.1.2 Componentes do Modelo.....	7
2.1.3 Estimaco dos Parâmetros do Modelo	10
2.2 PREDIÇÃO LINEAR ESPACIAL	15
2.2.1 Conceitos de Predico	15
2.2.2 Krigagem	18
2.3 INFERÊNCIA BAYESIANA PARA PREDIÇÃO ESPACIAL	20
2.3.1 Distribuico Preditiva Bayesiana	20
2.3.2 Cadeias de Markov.....	22
2.3.3 Intervalo de Credibilidade.....	26
2.4 MODELO MULTIVARIADO	27
2.4.1 Definição do Modelo Multivariado.....	29
2.4.2 Cokrigagem	31
2.5 DADOS COMPOSICIONAIS	31

2.5.1	Composição Regionalizada	33
2.5.2	Base Regionalizada	34
2.5.3	Representação Gráfica	35
2.5.4	Subcomposição Regionalizada	37
2.5.5	Amalgamação e Partição Regionalizada	38
2.5.6	Transformação Logística Modificada	39
2.5.7	Perturbação e Potência	42
2.5.8	Estatísticas Descritivas e Domínio de Confiança Para Dados Composicionais	43
2.5.9	Estacionariedade	44
2.5.10	Estrutura de Covariância Espacial	44
2.5.11	Estrutura de Covariância Espacial Intrínscica	46
3	MÉTODOS	47
3.1	Modelo Geoestatístico Composicional	47
3.1.1	Estimação dos Parâmetros do Modelo Geoestatístico Composicional	49
3.1.2	Incerteza Sobre os Parâmetros do Modelo	51
3.1.3	Transformação de Volta Dos Valores Preditos Para o Espaço Amostral Simplex	52
3.2	Inferência Bayesiana Para o Modelo Geoestatístico Composicional	56
3.3	Dados Simulados	59
3.4	Intervalos de Cobertura para os Dados Simulados	68
4	RESULTADOS	71
4.1	ANÁLISE DE FRAÇÕES GRANULOMÉTRICAS DE UM SOLO	72
4.2	INFERÊNCIA BAYESIANA DE DADOS COMPOSICIONAIS	77
4.2.1	Avaliação da Convergência	78
5	CONCLUSÕES E SUGESTÕES DE TRABALHOS FUTUROS	87
	Referências Bibliográficas	89
	Anexo A – Pacote geoComp	96
	Anexo B – Help do pacote geoComp	114
	agl	114

as.geoComp	115
classificaSolo	116
classSolo	118
cokri.bayes	119
cokrigagem	120
geoComp-internal	122
geoComp-package	123
infbayes	124
logarit.vero	126
mec	128
monta.V	131
pivo	133
posteriori	135
resul.mh	137
tab.coord.pol	138
volta.cokri	139
volta.quad	141
Anexo C – Código fonte R das análises estatísticas	145

Lista de Figuras

Figura 1.1	Triângulo textural.	2
Figura 1.2	Diagrama de classificação textural do solo.	2
Figura 2.1	Situações relacionando estacionariedade e isotropia.	8
Figura 2.2	(a) Composições de 3 partes como raios partindo da origem em \mathbb{R}_+^3 ; (b) O $\underline{p} = (p_a, p_b, p_c)\%$ representado no simplex \mathbb{S}^3	35
Figura 2.3	O dado composicional representado no simplex \mathbb{S}^3 e no \mathbb{R}^2	36
Figura 2.4	Diagrama ternário para dados do Lago Ártico incluindo o centro da distribuição e região 2—desvios-padrão de confiança.	37
Figura 2.5	Desenho que representa pontos em um diagrama ternário com mesmas percentagens (sobre a linha horizontal) do componente A e percentagens iguais (sobre a linha vertical) dos componentes B e C.	37
Figura 2.6	Interpretação geométrica da formação da subcomposição $X_{12}(\underline{x})$ da composição $\underline{X}(\underline{x})$: (a) em \mathbb{R}_+^3 ; (b) em \mathbb{S}^3	38
Figura 3.1	Distribuição das localizações no quadrado unitário.	59
Figura 3.2	Distribuição de X_1 , X_2 e X_3 e diagrama ternário das composições para a primeira configuração.	60
Figura 3.3	Distribuição das log-razão e correspondente diagrama de dispersão para a primeira configuração.	61
Figura 3.4	Mapas das percentagens de X_1 , X_2 e X_3 para dados da configuração 1.	62
Figura 3.5	Valores observados versus preditos de X_1 , X_2 e X_3 para a configuração 1.	62

Figura 3.6	Distribuição de X_1 , X_2 e X_3 e diagrama ternário das composições para a segunda configuração.	63
Figura 3.7	Distribuição das log-razão e correspondente diagrama de dispersão para a segunda configuração.	64
Figura 3.8	Mapas das porcentagens de X_1 , X_2 e X_3 para dados da configuração 2.	64
Figura 3.9	Valores observados versus preditos de X_1 , X_2 e X_3 para a configuração 2.	65
Figura 3.10	Distribuição de X_1 , X_2 e X_3 e diagrama ternário das composições para a terceira configuração.	66
Figura 3.11	Distribuição das log-razão e correspondente diagrama de dispersão para a terceira configuração.	66
Figura 3.12	Mapas das porcentagens de X_1 , X_2 e X_3 para dados da configuração 3.	67
Figura 3.13	Valores observados versus preditos de X_1 , X_2 e X_3 para a configuração 3.	67
Figura 3.14	Distribuições das estimativas dos parâmetros para 1000 simulações e amostras de tamanho 300.	70
Figura 3.15	Boxplot das estimativas dos parâmetros para 1000 simulações e amostras de tamanho 300.	70
Figura 4.1	Foto aérea do campo experimental de irrigação da ESALQ-USP com área de estudo correspondente ao quadrante irrigado por um sistema pivô-central.	71
Figura 4.2	Primeiro (círculos), segundo (triângulos), terceiro (diagonais verticais) e quarto (diagonais cruzadas) quartil da percentagem de areia (à esquerda), silte (meio) e argila (à direita) nas localizações amostrais.	72
Figura 4.3	Diagrama de classificação textural.	72
Figura 4.4	Distribuição de areia, silte e argila e diagrama ternário das composições.	74
Figura 4.5	Distribuição das log-razão e correspondente diagrama de dispersão.	75

Figura 4.6	Mapas das percentagens de areia, silte e argila obtidos por quadratura de Gauss-Hermite (A-C) e por simulação (D-F).	76
Figura 4.7	Porcentagens de areia, silte e argila obtidos por quadratura de Gauss-Hermite versus simulação.	77
Figura 4.8	Valores máximos e mínimos obtidos por simulação.	78
Figura 4.9	Log-verossimilhanças perfilhadas para ϕ e ρ	79
Figura 4.10	Deviance para ϕ com linhas de corte de 83,5% e 60% de confiança.	80
Figura 4.11	Deviance para ρ com linhas de corte de 95% de confiança.	81
Figura 4.12	Valores preditos de areia (à esquerda), silte (centro) e argila (à direita) por inferência bayesiana.	81
Figura 4.13	Autocorrelações dos parâmetros à <i>posteriori</i>	82
Figura 4.14	Trajetória das posterioris dos parâmetros.	83
Figura 4.15	Densidades estimadas para as posterioris dos parâmetros com respectivos intervalos de alta densidade à posteriori a 95%.	84
Figura 4.16	Diagnóstico de convergência para o Teste de Geweke.	85
Figura 4.17	Diagnóstico de convergência para o Teste de Gelman e Rubin.	86

Lista de Tabelas

Tabela 3.1	Estimativas, erros padrão e intervalos de confiança para a primeira configuração, pelo método delta via método de otimização “L-BFGS-B”. . . .	61
Tabela 3.2	Estimativas, erros padrão e intervalos de confiança para a segunda configuração, pelo método delta via método de otimização “L-BFGS-B”. . . .	64
Tabela 3.3	Estimativas, erros padrão e intervalos de confiança para a terceira configuração, pelo método delta via método de otimização “L-BFGS-B”. . . .	67
Tabela 3.4	Parâmetros, média e intervalo de 95% de confiança das estimativas dos parâmetros, e cobertura (%) dos intervalos via método delta para 1000 simulações de amostras de tamanho 300.	69
Tabela 4.1	Estatísticas descritivas dos percentuais de areia, silte e argila.	73
Tabela 4.2	Estimativas, erros padrão e intervalos de confiança pelo método Delta via método de otimização “L-BFGS-B”.	75
Tabela 4.3	Médias dos percentuais máximos e mínimos de areia, silte e argila obtidos por simulação.	75
Tabela 4.4	Esperança das 1200 simulações da distribuição à <i>posteriori</i> de θ e intervalos de 95% de credibilidade obtidos por inferência bayesiana considerando-se 12000 simulações, <i>burn-in</i> = 1000 e salto= 10.	79
Tabela 4.5	Autocorrelações das posteriores dos parâmetros.	80
Tabela 4.6	Correlações das posteriores dos parâmetros.	81
Tabela 4.7	Limites dos intervalos de 95% de credibilidade de alta densidade à <i>posteriori</i>	82
Tabela 4.8	Valores de z_G do teste de Geweke para as <i>posterioris</i> dos parâmetros. . . .	82

Tabela 4.9 Fatores de redução de escala potencial do teste de Gelman e Rubin. . . 83

Lista de Siglas

ALR	- Additive logratio (corresponde a transformação razão log-aditiva)
AGL	- Additive generalized logistic (corresponde a transformação logística generalizada aditiva)
CLR	- Center logratio (corresponde a transformação razão log-centrada)
EM	- Erro de Medida
EQM	- Erro de Predição Quadrático Médio
ESALQ	- Escola Superior de Agricultura " Luiz de Queiroz"
LR	- Log-Razão
MCMC	- Cadeias de Markov de Monte Carlo
USP	- Universidade de São Paulo
VME	- Variação de Pequena ou Micro Escala

Resumo

Palavras-chave: Geoestatística multivariada, geoestatística bayesiana, verossimilhança, agricultura de precisão, dados composicionais.

Abstract

Key-words: Multivariate Geostatistics, Bayesian Geostatistics, Likelihood, precision agriculture, compositional data.

1 INTRODUÇÃO

Os diversos processos naturais relacionados ao desenvolvimento e à produção vegetal dependem, em grande extensão, das características do solo relacionadas com a capacidade de armazenar e infiltrar água, com a capacidade de reter e disponibilizar nutrientes, com a resistência mecânica à penetração e com a capacidade de trocas gasosas no seu espaço poroso devido a sua amplitude e geometria. Além destas características fundamentais do solo, diversas outras podem exercer influência nestes processos e todas estão intimamente relacionadas com a proporção de seus elementos constituintes. Esta composição caracteriza a textura do solo, expressa por meio da proporção relativa das frações granulométricas, tradicionalmente caracterizadas como areia, silte e argila. Por ser uma característica física de grande importância e pouco susceptível a mudanças ao longo do tempo, a análise granulométrica é uma das primeiras e principais ações a serem conduzidas, para se inferir a respeito das futuras estratégias de manejo. Assim, o conhecimento da proporção segundo a qual as frações granulométricas se apresentam no solo contribui, inclusive, para a sua classificação, tornando-se informação de relevância para o sistema agrícola.

Com o aperfeiçoamento das técnicas de produção, com o avanço do conhecimento científico em torno do agrossistema e com o aprimoramento das tecnologias envolvidas, notadamente em relação ao maquinário empregado, tem-se buscado frequentemente a compreensão de que o sistema agrícola apresenta uma distribuição espacial heterogênea das suas características. Esta distribuição não pode ser ignorada quando se busca a otimização do processo de produção, para se atingir metas fixadas em decorrência de conceitos ambientais e econômicos.

Descrever a distribuição espacial dos valores das frações granulométricas do solo consiste em uma base de importância primordial para a descrição espacial das suas características relevantes para o processo de produção. Estudos como o de Gonçalves (1997), Gonçalves, Folegatti e Mata (2001), Eguchi, Silva e Oliveira (2002), por exemplo, buscam a descrição da distribuição espacial dos valores de areia, silte e argila do solo, utilizando técnicas estatísticas, com o emprego de métodos geoestatísticos. Isto tem possibilitado a descrição da distribuição dos valores de cada fração granulométrica, de forma independente.

Como se trata de uma proporção relativa entre as frações granulométricas, a análise textural do solo sempre revela um total de 100%, para o somatório destes componentes. Esta soma pode ser representada por um ponto $P = (80, 10, 10)$ correspondente aos valores percentuais de areia, silte e argila, no interior de um triângulo formado pelos

três elementos (Figura 1.1) e usado na classificação textural do solo.

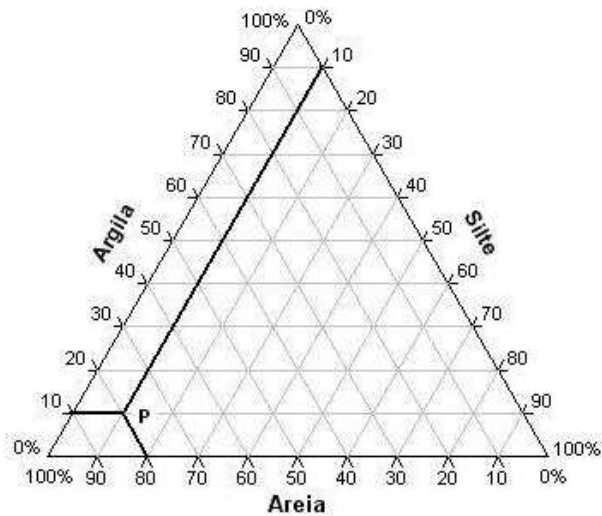


Figura 1.1: Triângulo textural.
FONTE: A autora (2010).

Este triângulo é dividido em áreas conforme Lemos e Santos (1996) *apud* Reichardt e Timm (2004) e de acordo com a localização do ponto nestas áreas tem-se uma classificação para o solo como pode ser visto na Figura 1.2.

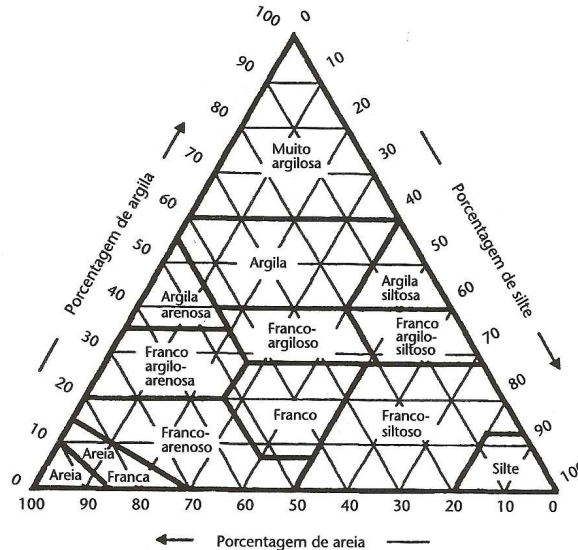


Figura 1.2: Diagrama de classificação textural do solo.
FONTE: Lemos e Santos (1996) *apud* Reichardt e Timm (2004).

Assim, areia, silte e argila podem ser tratados como dados composicionais, para os quais o somatório é uma constante o que implica em variações proporcionais dos mesmos. A partir desta abordagem, a descrição da distribuição espacial das três frações granulométricas pode ser feita levando-se em conta a relação entre elas enriquecendo potencialmente a análise.

A estatística espacial tem se apresentado como uma área de grande importância, para as mais diversas aplicações, inclusive em ciências agrárias particularmente, dentro do

contexto da denominada “Agricultura de precisão”. Em Agricultura de precisão citam-se, por exemplo, os trabalhos de Molin (1997a), Molin (1997b), Lake, Bock e Goode (1997), Srinivasan (2006) e outros. Trabalhos de grande relevância têm sido desenvolvidos na área de estatística espacial, como os de Matheron (1963), Cressie (1993), Bailey e Gatrell (1995), Banerjee, Carlin e Gelfand (2004), Schabenberger e Gotway (2005), etc. Na área de geoestatística com enfoque para a agricultura de precisão cita-se o trabalho de Silva (2000). Surgiram ainda trabalhos em geoestatística como os de Diggle, Tawn e Moyeed (1998), Schabenberger e Pierce (2001), Diggle e Ribeiro Jr. (2007) e Silva (2008) que diferem da linha tradicional no sentido de que a análise é baseada em modelos que induzem uma estrutura de covariância. A partir destes modelos é possível aplicar métodos clássicos de inferência baseados em verossimilhança que permitem produzir estimativas mais eficientes dos parâmetros e avaliar a incerteza em predições espaciais.

Trabalhos realizados por Aitchison (1986) em análise de dados composicionais apresentam uma metodologia adequada para analisar dados caracterizados por se apresentarem em forma de proporções complementares. A partir dos anos 2000, estes tipos de dados são analisados considerando-se a espacialização das variáveis, mas ainda sob a abordagem geoestatística tradicional. Odeh, Tood e Triantafilis (2003) mostram que, sem considerar que a matriz de covariância associada ao modelo composicional seja definida positiva e que os valores interpolados satisfaçam a restrição “soma um”, a interpolação espacial de dados de frações de partículas de solo produzem incerteza e valores interpolados irreais. Pawlowsky-Glahn e Olea (2004) satisfazem estas exigências mas não adotam declaração explícita de modelo e não consideram predição espacial bayesiana. Lark e Bishop (2007) fazem um estudo sobre cokrigagem de frações de partículas do solo concluindo que a predição pode ser feita através da cokrigagem razão log-aditiva (ALR) que considera o logaritmo das razões dos componentes e que esta apresenta vantagens em relação à cokrigagem sem transformação que considera apenas as razões dos componentes. Concluem ainda que existem vantagens se a transformação de volta das predições para a escala original, das composições, são calculadas por quadratura de Gauss-Hermite para aproximar a esperança condicional. Tjelmeland e Lund (2003) modelam dados composicionais espaciais com metodologia bayesiana sem adotar forma explícita para a função de covariância e sem fazer predição espacial, já Obage (2007) faz inferência bayesiana para dados composicionais sem espacialização, considerando as transformações ALR e Box-Cox e assumindo uma distribuição gaussiana multivariada para erros correlacionados.

O objetivo geral deste trabalho foi propor e implementar um modelo geoestatístico para dados espaciais composicionais utilizando estruturas multivariadas, como componentes do modelo especificados por função de correlação, e desenvolvendo métodos de inferência baseadas em verossimilhança e sob o enfoque bayesiano.

Os objetivos específicos foram:

- Desenvolver uma forma alternativa à proposta de Pawlowsky-Glahn e Olea (2004) baseada na declaração explícita de modelos;
- construir um modelo em que a dependência espacial e entre variáveis seja considerada na obtenção de uma função de covariância válida;
- derivar métodos de inferência baseados em verossimilhança para a estimação dos parâmetros desconhecidos do modelo;

- aplicar métodos bayesianos para a inferência dos parâmetros do modelo;
- desenvolver rotinas computacionais para análise de dados composicionais;
- aplicar a metodologia proposta em um conjunto de dados de solo elaborando mapas temáticos de modelos composicionais em estudo de caso.

O trabalho foi dividido em cinco capítulos. No Capítulo 2 apresenta-se uma revisão de literatura sobre geoestatística, teoria de dados composicionais e inferência bayesiana com a aplicação da metodologia proposta em três conjuntos de dados composicionais simulados. Além disso foi feito um estudo, por simulação, para a determinação dos intervalos de cobertura dos parâmetros do modelo. No Capítulo 3 propõe-se o modelo geoestatístico bivariado para dados composicionais, desenvolve-se a teoria para a estimação dos parâmetros pelo método clássico e sob o paradigma de inferência bayesiana. No Capítulo 4 apresenta-se os resultados da aplicação da metodologia proposta em um conjunto de dados reais de areia, silte e argila tanto pelo método clássico quanto por inferência bayesiana. O Capítulo 5 finaliza o trabalho com a conclusão e sugestões de trabalhos futuros. Em anexo, apresenta-se o manual e as funções do pacote estatístico geoComp desenvolvido para este trabalho bem como os scripts das análises realizadas.

2 REVISÃO DA LITERATURA

2.1 MODELO GEOESTATÍSTICO GAUSSIANO

2.1.1 Definição do Modelo

Hoel, Port e Stone (1972) definem um processo estocástico como qualquer coleção de variáveis aleatórias $S(\underline{x})$, $\underline{x} \in \mathbb{R}^d$, definidas num espaço de probabilidade comum e com d o número de entradas do vetor de localização \underline{x} . A caracterização das propriedades de processos estocásticos requer uma forma de expressar a distribuição conjunta das variáveis aleatórias $S(\underline{x}_1), S(\underline{x}_2), \dots, S(\underline{x}_n)$ (FALEIROS; YONEYAMA, 2002). Em particular, para $d = 2$, um processo espacial gaussiano $S = (S(\underline{x}) : \underline{x} \in \mathbb{R}^2)$ é um processo estocástico com a propriedade de que, para qualquer coleção de localizações $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$, com $\underline{x}_i \in \mathbb{R}^2$, $S = (S(\underline{x}_1), S(\underline{x}_2), \dots, S(\underline{x}_n))$ tem uma distribuição conjunta Gaussiana multivariada e fica completamente especificado pela função média, e pela matriz de covariância cujos elementos correspondem a função $Cov(S(\underline{x}_i), S(\underline{x}_j))$ (DIGGLE; RIBEIRO JR., 2007). Schabenberger e Pierce (2001) destacam que a função de distribuição acumulada é aquela de uma distribuição gaussiana n -variada

$$F(S(\underline{x})) = P(S(\underline{x}_1) < s_1, \dots, S(\underline{x}_n) < s_n).$$

Segundo Diggle, Ribeiro Jr e Christensen (2003) a geoestatística é um ramo da estatística espacial na qual os dados consistem de mensurações y_1, y_2, \dots, y_n obtidas nas localizações $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ amostradas em uma região $A \subset \mathbb{R}^2$ espacialmente contínua. As mensurações estão relacionadas a um fenômeno espacial que pode ser tratado como a realização de um processo estocástico $S(\underline{x}); \underline{x} \in \mathbb{R}^2$, denominado sinal que, em geral, não é diretamente observável. O valor observado y_i é uma realização de $Y(\underline{x}_i)$ que é uma versão ruidosa de $S(\underline{x}_i)$ no caso gaussiano. O delineamento amostral será de forma que $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ sejam fixos ou estocasticamente independentes de $Y(\underline{x}_1), Y(\underline{x}_2), \dots, Y(\underline{x}_n)$.

De acordo com Matérn (1960) *apud* Diggle e Ribeiro Jr. (2007), a teoria de amostragem espacial mostra que, sob suposições típicas de modelagem, as propriedades espaciais são mais eficientemente estimadas por um delineamento regular nos quais as localizações formam uma grade regular sobre a região em estudo. Ainda, delineamentos em grade são geralmente eficientes para predição espacial. Se não existir relação entre

a escolha da coordenada \underline{x} e o atributo y , que será o caso deste trabalho, a amostragem será denominada não preferencial, caso contrário, será denominada preferencial (DIGGLE; MENEZES; SU, 2010).

A terminologia “geoestatística baseada em modelos” foi introduzida por Diggle, Tawn e Moyeed (1998) e se caracteriza pela declaração explícita e completa e utilização de métodos de inferência estatística baseadas na verossimilhança aplicada a problemas geoestatísticos. Assim, um modelo geoestatístico é a especificação da distribuição conjunta $[Y(\underline{x}), S(\underline{x})]$ que usualmente é especificada na forma fatorada como

$$[S(\underline{x}), Y(\underline{x})] = [S(\underline{x})][Y(\underline{x})|S(\underline{x})].$$

Em particular, esse modelo não especifica a distribuição de $\{\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n\}$ (delineamento amostral) o qual será assumido independente de $S(\underline{x})$ e de $Y(\underline{x})$.

De acordo com Diggle, Ribeiro Jr e Christensen (2003), no modelo geoestatístico gaussiano estacionário $S(\underline{x})$ é um processo gaussiano estacionário, com $E(S(\underline{x}_i)) = \underline{\mu}$; $Var(S(\underline{x}_i)) = \sigma^2$ e função de correlação $\rho(u_{ij}) = Corr(S(\underline{x}_i), S(\underline{x}_j))$ em que $u_{ij} = \|\underline{x}_i - \underline{x}_j\|$, $i, j = 1, \dots, n$ e $Y(\underline{x}_i)$ o valor observado na localização \underline{x}_i . Essa função de correlação é definida como:

$$\rho(u) = \frac{Cov(S(\underline{x}_i), S(\underline{x}_j))}{\sigma^2},$$

que é simétrica em u , ou seja, $\rho(u) = \rho(-u)$.

Os autores definem como um modelo geoestatístico plausível aquele em que a distribuição de $Y(\underline{x}_i)$, $i = 1, 2, \dots, n$ condicionada a distribuição de S , $S(\cdot)$, é gaussiana com média $S(\underline{x}_i)$ e variância τ^2 e $Y(\underline{x}_i)$ são mutuamente independentes condicionados em $S(\cdot)$, e que pode então ser escrito na forma:

$$Y(\underline{x}_i) = S(\underline{x}_i) + Z(\underline{x}_i) \quad i = 1, \dots, n, \quad (2.1)$$

em que $Z(\underline{x}_i) \sim N(\mathbf{0}; \tau^2)$ são erros aleatórios independentes de $S(\underline{x})$. Integrando-se a distribuição conjunta de $Y(\underline{x})$ e $S(\underline{x})$ com relação a $S(\underline{x})$, esse modelo pode ser representado também como:

$$\underline{Y}(\underline{x}) \sim N_n(\underline{\mu}\mathbf{1}; \sigma^2\mathbf{R} + \tau^2\mathbf{I})$$

em que $\mathbf{1}$ é um vetor com n elementos iguais a 1, \mathbf{R} é uma matriz de ordem $n \times n$ cujos elementos são as correlações $\rho(u_{ij})$ e \mathbf{I} é a matriz identidade de ordem $n \times n$.

Com a reparametrização $\mathbf{v}^2 = \frac{\tau^2}{\sigma^2}$, segue que:

$$Var(\underline{Y}(\underline{x})) = \sigma^2\mathbf{R} + \tau^2\mathbf{I} = \sigma^2 \left(\mathbf{R} + \frac{\tau^2}{\sigma^2}\mathbf{I} \right) = \sigma^2\mathbf{V}$$

onde

$$\mathbf{V} = \mathbf{R} + v^2\mathbf{I}, \quad (2.2)$$

e a distribuição de $[Y(\underline{x})]$ é escrita na forma

$$Y(\underline{x}) \sim N_n(\underline{\mu}\mathbf{1}; \sigma^2\mathbf{V}).$$

2.1.2 Componentes do Modelo

Segundo Schabenberger e Gotway (2005) um processo estocástico é estacionário se a distribuição espacial de $S(\underline{x})$ é invariante sob translação das coordenadas. Schabenberger e Pierce (2001) afirmam que geometricamente isto implica que a distribuição espacial é invariante sob rotação e estiramento do sistema de localização das amostras. Bailey e Gatrell (1995) afirmam que isto acontece se as propriedades estatísticas são independentes da localização absoluta na região de estudo $A \subset \mathbb{R}^2$. Isto implica que a média e a variância são constantes na região A e não dependem da localização \underline{x} . Implica também que a covariância $Cov(S(\underline{x}_i), S(\underline{x}_j))$, $i \neq j$, dependerá somente das localizações relativas destes dois pontos, da distância u que as separa e da direção entre elas, e não de sua localização absoluta na região A . Desta forma, este processo pode ser pensado como o equivalente espacial de uma amostra aleatória em estatística clássica que dá origem a variáveis aleatórias independentes com a mesma média e dispersão (SCHABENBERGER; PIERCE, 2001). Ainda segundo estes autores, a função covariância é denominada isotrópica na ausência de dependência da direção, ou melhor, quando a função de covariância depende somente da distância absoluta, neste trabalho a distância euclidiana, entre os pares de pontos. O processo espacial é isotrópico se, em acréscimo à estacionariedade, a covariância depender somente da distância entre dois pontos, e não da direção nos quais estão separados (BAILEY; GATRELL, 1995).

A Figura 2.1 ilustra uma situação onde os pares de pontos $(\underline{x}_1, \underline{x}_2)$ e $(\underline{x}_3, \underline{x}_4)$ representam a condição estacionariedade porque \underline{x}_1 e \underline{x}_2 estão separados pela mesma distância que \underline{x}_3 e \underline{x}_4 , têm a mesma direção, estão em locais diferentes em A , e $Cov(\underline{x}_1, \underline{x}_2) = Cov(\underline{x}_3, \underline{x}_4)$. Pode-se observar que os pares $(\underline{x}_1, \underline{x}_2)$ e $(\underline{x}_5, \underline{x}_6)$ estão separados pela mesma distância mas estão em direções diferentes. Neste caso, ao considerar covariâncias diferentes, $Cov(\underline{x}_1, \underline{x}_2) \neq Cov(\underline{x}_5, \underline{x}_6)$, teria-se estacionariedade se, ao rotacionar o par $(\underline{x}_5, \underline{x}_6)$ colocando-o na mesma direção de $\underline{x}_1, \underline{x}_2$, as covariâncias passassem a ser iguais. Por outro lado, se as covariâncias fossem iguais teria-se isotropia pois, em acréscimo a estacionariedade, não haveria necessidade de rotacionar para que tivessem a mesma covariância.

De acordo com Diggle, Ribeiro Jr e Christensen (2003), a especificação da função de correlação, $\rho(u)$, determina a suavidade do processo $S(\underline{x})$. A descrição matemática formal da suavidade de uma superfície espacial é dada por seu grau de diferenciabilidade.

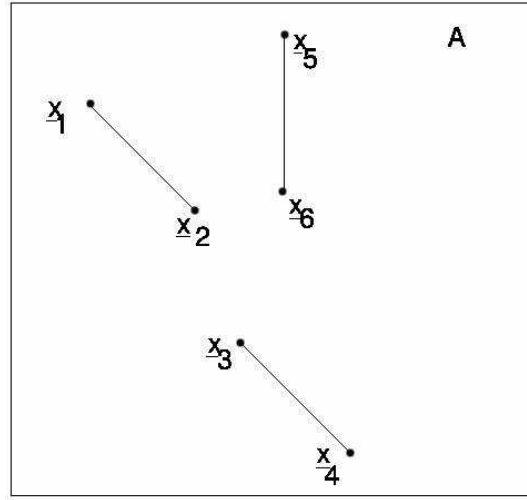


Figura 2.1: Situações relacionando estacionariedade e isotropia.
 FONTE: A autora (2010).

$S(\underline{x})$ é quadrado-médio contínuo se $\lim_{u \rightarrow 0} E(\{S(x_i) - S(x_j)\}^2) = 0$ para todo \underline{x} . Da mesma forma, é quadrado-médio diferenciável se existe um processo $S'(\underline{x})$ tal que

$$\lim_{u \rightarrow 0} E \left(\left\{ \frac{S(x_i) - S(x_j)}{u} - S'(x_i) \right\}^2 \right) = 0.$$

Então, a diferenciabilidade quadrado-médio de $S(\underline{x})$ está diretamente relacionada com a diferenciabilidade de sua função de covariância através do resultado que diz que se $S(\underline{x})$ é um processo gaussiano estacionário com função de correlação $\rho(u), u \in \mathbb{R}$, então $S(\underline{x})$ é quadrado-médio contínuo se, e somente se, $\rho(u)$ é contínuo em $u = 0$ e é k vezes quadrado-médio diferenciável se, e somente se, $\rho(u)$ é ao menos $2k$ vezes diferenciável em $u = 0$. A demonstração deste resultado pode ser encontrada em Stein (1999).

A função de covariância está associada à função de correlação. Na estrutura do modelo proposto, a família Matérn de funções de correlação apresentada em Diggle e Ribeiro Jr. (2007) é uma importante função paramétrica de correlação com expressão analítica dada por

$$\rho(u, k, \phi) = \frac{1}{2^{k-1} \Gamma(k)} \left(\frac{u}{\phi} \right)^k K_k \left(\frac{u}{\phi} \right)$$

em que $K_k(\cdot)$ denota a função de Bessel modificada de ordem k , $\phi > 0$ é um parâmetro de escala, associado à extensão de dependência espacial (alcance) e $k > 0$ é um parâmetro de forma que determina a suavidade analítica do processo $S(\underline{x})$, interpretado como uma medida da diferenciabilidade do processo. Especificamente, $S(\underline{x})$ é $\lceil k - 1 \rceil$ vezes quadrado médio diferenciável, onde o símbolo “ $\lceil \cdot \rceil$ ” é denominado *ceiling* e “ $\lceil k \rceil$ ” significa o menor inteiro maior ou igual a k . Devido a dificuldade de identificação de todos os parâmetros do modelo, na prática, os valores de k podem ser escolhidos dentre o conjunto de valores $\{0, 5; 1, 5; 2, 5\}$ correspondendo, respectivamente, à não diferenciabilidade, um processo estocástico uma e duas vezes diferenciável na origem. Assim, para diferentes valores de k pode-se obter diferentes valores de amplitude prática do modelo que é a distância u_0

no qual $\rho(u_0) = \alpha$, onde α é um valor tão pequeno quanto o pesquisador determinar, usualmente 0,05. Neste caso, para o conjunto estabelecido acima, os valores de u_0 são aproximadamente $\{3\phi; 4,75\phi; 5,92\phi\}$, respectivamente. Nesta família, fazendo $k = 0,5$ obtém-se a função de correlação exponencial $\rho(u) = \exp\left(-\frac{u}{\phi}\right)$, $u_0 = 3\phi$ e $\lim_{k \rightarrow \infty} \rho(u) = \exp\left\{-\left(\frac{u}{\phi}\right)^2\right\}$ obtendo-se a função de correlação gaussiana para a qual $u_0 \simeq \sqrt{3}\phi$.

Em textos de geoestatística tais como Isaaks e Srisvastava (1989), Kitanidis (1997), Goovaerts (1997, Seção 4.2.3), Chilès e Delfiner (1999), Schabenberger e Pierce (2001), Diggle e Ribeiro Jr. (2007) e outros, pode-se encontrar outras funções de correlação. Neste trabalho, sem perda de generalidade, será adotada a função de correlação exponencial ($k = 0,5$).

A função de correlação, por sua vez, está associada à matriz de covariância. Isaaks e Srisvastava (1989) afirmam que a matriz de covariância deve ser definida positiva como uma garantia de que a variância de qualquer variável aleatória formada pela combinação linear ponderada de outras variáveis aleatórias será positiva.

Quando não existe estacionariedade na média a situação mais comum é que $\mu(\underline{x})$, denominada superfície de tendência, seja escrita como um modelo de regressão polinomial usando potências e produtos cruzados das coordenadas cartesianas de \underline{x} como variáveis explicativas. Diggle e Ribeiro Jr. (2007) afirmam que superfícies de tendência linear e quadrática podem fornecer descritores empíricos úteis da tendência espacial não explicada, mas superfícies de ordem maior devem ser evitadas porque tendências mais complexas podem ser melhor descritas através do componente estocástico do modelo.

A especificação de $\mu(\underline{x})$ pode também ser feita em função de outras variáveis explicativas e, neste caso, a média se associa a uma “tendência externa”.

Uma superfície de tendência linear pode ser escrita como

$$\mu(\underline{x}) = \beta_0 + \sum_{j=1}^p \beta_j d_j(\underline{x})$$

em que $d_j(\underline{x})$ são variáveis explicativas espaciais, dependentes ou não das coordenadas. Neste trabalho será considerada uma superfície de tendência constante ($\mu(\underline{x}) = \mu$), embora uma superfície de tendência linear possa ser adotada.

Outro componente implícito do modelo é o “efeito pepita” que é um termo usado para representar a variância τ^2 da variável $Z(x_i)$ em (2.1). Esta variância pode ser dividida em dois componentes como

$$\tau^2 = \text{EM} + \text{VME} \tag{2.3}$$

em que EM significa erro de medida e VME variação de pequena ou micro escala, que representa uma variação não capturada pelo processo $S(\underline{x})$ e que ocorre em distâncias menores do que a menor distância observada entre duas localizações. Em muitas situações, a distinção dos dois componentes não é possível, mas conforme Equação (2.3) pode-se determinar o valor de VME se o valor de EM for conhecido. No entanto, isto só será

possível se houver mais de uma observação de Y na mesma localização. Neste trabalho o efeito pepita não será considerado.

Quando os dados observados $Y(\underline{x})$ não seguem uma distribuição gaussiana ou em problemas de não estacionariedade da variância, o valor do erro quadrático médio mínimo do preditor é afetado de forma a se obter aproximações ruins. Em muitos casos, através de transformações, é possível que os dados passem a seguir uma distribuição gaussiana.

Box e Cox (1964) propõem a família de transformações

$$Y^*(\underline{x}) = \begin{cases} \frac{Y(\underline{x})^\lambda - 1}{\lambda} & , \lambda \neq 0 \\ \log Y(\underline{x}) & , \lambda = 0 \end{cases}$$

em que o parâmetro λ introduz flexibilidade ao modelo. Alguns valores interpretáveis para esse parâmetro são:

$$\begin{aligned} \lambda = 1,0 & : \text{ sem transformação} \\ \lambda = 0,5 & : \text{ transformação raiz quadrada} \\ \lambda = -1,0 & : \text{ transformação recíproco.} \end{aligned}$$

Pode-se ainda modelar explicitamente os dados $Y(\underline{x})$ e, neste caso, Diggle, Tawn e Moyeed (1998) propõem modelos para distribuições não gaussianas de Y na família exponencial.

2.1.3 Estimação dos Parâmetros do Modelo

Os métodos usuais na estimação de parâmetros de um modelo gaussiano linear são o método dos mínimos quadrados e o método da máxima verossimilhança.

De acordo com Diggle e Ribeiro Jr. (2007), na estimação dos parâmetros da superfície de tendência linear pelo método dos mínimos quadrados, a média do modelo, $E(Y(\underline{x}))$, é assumida como

$$\mu(\underline{x}) = \beta_0 + \sum_{j=1}^p \beta_j d_j(\underline{x}),$$

em que $d_j(\underline{x})$, $j = 1, \dots, p$ são covariáveis espaciais e β_j , $j = 0, \dots, p$ são os parâmetros da regressão linear.

Usando o método dos mínimos quadrados ordinários, as estimativas $\tilde{\beta}_j$ são aquelas que minimizam a soma dos quadrados dos resíduos, e que sob o modelo dado pela Equação (2.1) incluindo a tendência

$$Y(\underline{x}_i) = \mu(\underline{x}_i) + S(\underline{x}_i) + Z(\underline{x}_i) \quad i = 1, \dots, n, \quad (2.4)$$

é dada por

$$\text{SQR}(\boldsymbol{\beta}) = \sum_{i=1}^n Z^2(x_i) = \sum_{i=1}^n (Y(x_i) - \mu(x_i) - S(x_i))^2.$$

Para estimar corretamente os parâmetros da média, os parâmetros da correlação deveriam ser conhecidos mas, geralmente, não o são. Por outro lado, para estimar os da correlação, seriam necessários os da média que tampouco são conhecidos. O que se faz então é considerar, inicialmente, o modelo $Y(x_i) = \mu(x_i) + Z(x_i)$ que ignora a correlação e cuja soma de quadrados dos resíduos é dada por

$$\text{SQR}(\boldsymbol{\beta}) = \sum_{i=1}^n Z^2(x_i) = \sum_{i=1}^n (Y(x_i) - \mu(x_i))^2.$$

Matricialmente, a equação de regressão linear múltipla é

$$\underline{Y}(\underline{x}) = \mathbf{D}\underline{\beta}$$

sendo \mathbf{D} uma matriz $n \times p$ de covariáveis, $\underline{\tilde{\beta}}$ o vetor dos parâmetros da regressão e a aplicação do método dos mínimos quadrados resultará no estimador

$$\underline{\tilde{\beta}} = (\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'\underline{Y}(\underline{x}).$$

Assumindo que o modelo para a média seja especificado corretamente, as estimativas $\underline{\tilde{\beta}}_j$ são não viesadas e consistentes independentemente da estrutura de covariância. Com essas estimativas pode-se calcular o vetor dos resíduos

$$\underline{Z}(\underline{x}) = \underline{Y}(\underline{x}) - \mathbf{D}\underline{\tilde{\beta}}$$

com elementos $Z(x_i)$, $i = 1, \dots, n$ que substituídos na Equação (2.4) permitem a obtenção das estimativas dos parâmetros de $S(x)$, e por sua vez a obtenção do estimador $\hat{\mathbf{V}}$ para \mathbf{V} apresentado na Equação (2.2). Com este estimador é possível obter uma estimativa de $\underline{\beta}$ mais eficiente dada pela estimativa de mínimos quadrados generalizados conforme a equação

$$\underline{\hat{\beta}} = (\mathbf{D}'\mathbf{V}^{-1}\mathbf{D})^{-1}\mathbf{D}'\mathbf{V}^{-1}\underline{Y}(\underline{x}), \quad (2.5)$$

substituindo-se nesta, \mathbf{V} por $\hat{\mathbf{V}}$. Este procedimento é repetido até a obtenção da convergência.

A estimativa $\underline{\hat{\beta}}$ é não viesada e tem a menor variância entre todas as estimativas lineares não viesadas.

Por outro lado, utilizando o método da máxima verossimilhança pode-se obter consistentemente estimativas dos parâmetros de média e covariância maximizando a função de log-verossimilhança. No caso de $\underline{Y}(\underline{x})$ apresentar distribuição gaussiana, $\underline{\hat{\beta}}$ dado

em (2.5) coincide com a estimativa de máxima verossimilhança.

Admitindo uma superfície de tendência polinomial para $\underline{\mu}(\underline{x})$, tem-se

$$\underline{Y}(\underline{x}) \sim N_n(\underline{D}\underline{\beta}; \sigma^2\mathbf{R} + \tau^2\mathbf{I}) \quad (2.6)$$

onde \mathbf{D} é uma matriz $n \times p$ de covariáveis; $\underline{\beta}$ é o vetor de parâmetros da regressão correspondente e a matriz de correlação \mathbf{R} depende de ϕ e da distância, no caso o parâmetro da função de correlação exponencial. A função de verossimilhança de $\underline{Y}(\underline{x})$ é dada por:

$$L(\underline{Y}(\underline{x})) = (2\pi)^{-\frac{n}{2}} |\sigma^2\mathbf{R} + \tau^2\mathbf{I}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\underline{Y}(\underline{x}) - \underline{D}\underline{\beta})' (\sigma^2\mathbf{R} + \tau^2\mathbf{I})^{-1} (\underline{Y}(\underline{x}) - \underline{D}\underline{\beta}) \right\},$$

e a função de log-verossimilhança é

$$\begin{aligned} l(\underline{\beta}, \tau^2, \sigma^2, \phi) &= \ln \left[(2\pi)^{-\frac{n}{2}} \right] - \ln(|\sigma^2\mathbf{R} + \tau^2\mathbf{I}|^{\frac{1}{2}}) - \frac{1}{2} (\underline{Y}(\underline{x}) - \underline{D}\underline{\beta})' (\sigma^2\mathbf{R} + \tau^2\mathbf{I})^{-1} (\underline{Y}(\underline{x}) - \underline{D}\underline{\beta}) \\ &= -\frac{1}{2} \left\{ n \ln(2\pi) + \ln(|\sigma^2\mathbf{R} + \tau^2\mathbf{I}|) + (\underline{Y}(\underline{x}) - \underline{D}\underline{\beta})' (\sigma^2\mathbf{R} + \tau^2\mathbf{I})^{-1} (\underline{Y}(\underline{x}) - \underline{D}\underline{\beta}) \right\} \end{aligned} \quad (2.7)$$

Para proceder a maximização de (2.7), considera-se $\underline{Y}(\underline{x}) \sim N_n(\underline{D}\underline{\beta}; \sigma^2\mathbf{V})$ e os seguintes resultados da álgebra matricial:

$$\frac{\partial}{\partial \underline{X}} (\underline{A}\underline{X}) = \underline{A}' \quad (2.8)$$

$$\frac{\partial}{\partial \underline{X}} (\underline{X}'\underline{A}\underline{X}) = 2\underline{A}\underline{X} \quad (2.9)$$

em que \underline{A} é uma matriz quadrada de ordem $n \times n$ e \underline{X} um vetor de ordem $n \times 1$.

Desta forma, a Equação (2.7) pode ser reescrita como

$$l(\underline{\beta}, \sigma^2) = -\frac{1}{2} \left\{ n \ln(2\pi) + \ln(|\sigma^2\mathbf{V}|) + (\underline{Y}(\underline{x}) - \underline{D}\underline{\beta})' (\sigma^2\mathbf{V})^{-1} (\underline{Y}(\underline{x}) - \underline{D}\underline{\beta}) \right\}. \quad (2.10)$$

Nessa equação,

$$\begin{aligned} (\underline{Y}(\underline{x}) - \underline{D}\underline{\beta})' (\sigma^2\mathbf{V})^{-1} (\underline{Y}(\underline{x}) - \underline{D}\underline{\beta}) &= \underline{Y}'(\underline{x}) (\sigma^2\mathbf{V})^{-1} \underline{Y}(\underline{x}) - \underline{Y}'(\underline{x}) (\sigma^2\mathbf{V})^{-1} \underline{D}\underline{\beta} - \\ &\quad \underline{\beta}' \underline{D}' (\sigma^2\mathbf{V})^{-1} \underline{Y}(\underline{x}) + \underline{\beta}' \underline{D}' (\sigma^2\mathbf{V})^{-1} \underline{D}\underline{\beta} \end{aligned}$$

de modo que permite expressar (2.10) como:

$$\begin{aligned} l(\underline{\beta}, \sigma^2) &= -\frac{1}{2} \left\{ n \ln(2\pi) + \ln(|\sigma^2\mathbf{V}|) + \frac{1}{\sigma^2} \left[\underline{Y}(\underline{x})' \mathbf{V}^{-1} \underline{Y}(\underline{x}) - \right. \right. \\ &\quad \left. \left. 2\underline{Y}'(\underline{x}) \mathbf{V}^{-1} \underline{D}\underline{\beta} + \underline{\beta}' (\underline{D}' \mathbf{V}^{-1} \underline{D}) \underline{\beta} \right] \right\}. \end{aligned} \quad (2.11)$$

Aplicando os resultados (2.8) e (2.9) em (2.11) tem-se:

$$\frac{\partial}{\partial \underline{\beta}} l(\underline{\beta}, \sigma^2) = -\frac{1}{\sigma^2} \left[-(\underline{Y}'(\underline{x})\mathbf{V}^{-1}\mathbf{D})' + (\mathbf{D}'\mathbf{V}^{-1}\mathbf{D})\underline{\beta} \right].$$

Fazendo

$$\frac{1}{\sigma^2} \left(\mathbf{D}'\mathbf{V}^{-1}\underline{Y}(\underline{x}) - \mathbf{D}'\mathbf{V}^{-1}\mathbf{D}\hat{\underline{\beta}} \right) = 0,$$

vem:

$$\hat{\underline{\beta}} = (\mathbf{D}'\mathbf{V}^{-1}\mathbf{D})^{-1}\mathbf{D}'\mathbf{V}^{-1}\underline{Y}(\underline{x}). \quad (2.12)$$

Nesta equação, $\hat{\underline{\beta}}$ coincide com a estimativa de mínimos quadrados generalizada e depende somente de ϕ e \mathbf{v}^2 . Reescrevendo (2.10) e usando o fato de que $|\text{Var}(\underline{Y}(\underline{x}))| = |\sigma^2\mathbf{V}| = (\sigma^2)^n|\mathbf{V}|$ tem-se

$$l(\underline{\beta}, \sigma^2) = -\frac{1}{2} \left\{ n \ln(2\pi) + n \ln(\sigma^2) + \ln|\mathbf{V}| + \frac{1}{\sigma^2} (\underline{Y}(\underline{x}) - \mathbf{D}\underline{\beta})' \mathbf{V}^{-1} (\underline{Y}(\underline{x}) - \mathbf{D}\underline{\beta}) \right\}. \quad (2.13)$$

Derivando (2.13) em relação a σ^2 vem

$$\frac{\partial}{\partial \sigma^2} l(\underline{\beta}, \sigma^2) = -\frac{1}{2} \left\{ \frac{n}{\sigma^2} + (\underline{Y}(\underline{x}) - \mathbf{D}\underline{\beta})' \mathbf{V}^{-1} (\underline{Y}(\underline{x}) - \mathbf{D}\underline{\beta}) \left(-\frac{1}{(\sigma^2)^2} \right) \right\},$$

e igualando este termo a zero obtém-se a expressão do estimador

$$\hat{\sigma}^2 = n^{-1} \left[(\underline{Y}(\underline{x}) - \mathbf{D}\hat{\underline{\beta}})' \mathbf{V}^{-1} (\underline{Y}(\underline{x}) - \mathbf{D}\hat{\underline{\beta}}) \right] \quad (2.14)$$

No ponto de máximo, $\underline{\beta} = \hat{\underline{\beta}}$ e $\sigma^2 = \hat{\sigma}^2$ de forma que substituindo as Equações (2.12) e (2.14) na Equação (2.10) obtém-se a log-verossimilhança concentrada

$$l_0(\phi, \mathbf{v}^2) = -\frac{1}{2} \left\{ n \ln(2\pi) + n \ln(\hat{\sigma}^2) + \ln|\mathbf{V}| + n \right\} \quad (2.15)$$

que recebe as constantes $\underline{Y}(\underline{x})$, a matriz \mathbf{D} cujos elementos são os valores das covariáveis e a matriz \mathbf{V} . Para obter a estimativa dos parâmetros a Equação (2.15) deve ser otimizada numericamente com relação a ϕ e \mathbf{v} . Com o valor obtido para $\hat{\phi}$ e $\hat{\mathbf{v}}^2$ obtém-se a matriz $\hat{\mathbf{V}}$, que substituída na Equação (2.12) permite a obtenção de $\hat{\underline{\beta}}$ e, conseqüentemente, $\hat{\sigma}^2$ de (2.14). Desta forma, substituindo-se $\hat{\mathbf{v}}^2$ e $\hat{\sigma}^2$ em $\mathbf{v}^2 = \tau^2/\sigma^2$ obtém-se $\hat{\tau}^2$.

Quando os dados $\underline{Y}(\underline{x})$ sofrem uma transformação $\underline{Y}^*(\underline{x})$ como sugerido na Subseção 2.1.2, aplicando-se o método Jacobiano obtém-se a log-verossimilhança que é uma

função dos parâmetros e de $\underline{Y}(\underline{x})$:

$$l(\underline{\beta}, \sigma^2, \phi, \mathbf{v}^2, \lambda) = (\lambda - 1) \sum_{i=1}^n \ln(y_i(\underline{x})) - 0,5 \{n \ln(2\pi) + \ln|\sigma^2 \mathbf{V}(\phi, \mathbf{v}^2)| + \\ (\underline{Y}^*(\underline{x}) - \mathbf{D}\underline{\beta})' [\sigma^2 \mathbf{V}(\phi, \mathbf{v}^2)]^{-1} (\underline{Y}^*(\underline{x}) - \mathbf{D}\underline{\beta})\},$$

que é otimizável no sentido de se obter estimativas de máxima verossimilhança dos parâmetros do modelo. A transformação exige um método para transformar de volta os valores $\underline{Y}^*(\underline{x})$ para a escala original e isto pode ser feito através de formas analíticas como também por simulação.

Para a avaliação da incerteza sobre os parâmetros do modelo pode-se aplicar o método Delta, útil para aproximar a distribuição de alguma função das estatísticas (DEGROOT; SCHERVISH, 2002). O método pode ser visto, inclusive, como um procedimento de aproximação quadrática da função de verossimilhança e é dado pelo seguinte resultado (PAWITAN, 2001):

Teorema 2.1 : *Seja $\hat{\theta}$ uma estimativa de θ baseada numa amostra de tamanho n tal que*

$$(\hat{\theta} - \theta) \rightarrow N(0; \frac{\sigma^2}{n}).$$

Então, para qualquer função $h(\cdot)$ que é diferenciável ao redor de θ e $h'(\hat{\theta}) \neq 0$, tem-se

$$(h(\hat{\theta}) - h(\theta)) \rightarrow N(0; \frac{\sigma^2 |h'(\theta)|^2}{n}).$$

Pode-se dizer então que $h(\hat{\theta})$ é aproximadamente gaussiana com média $h(\theta)$ e variância $|h'(\theta)|^2 \text{Var}(\hat{\theta})$.

A obtenção dos intervalos de confiança pode ser baseada na função deviance definida por

$$D(\theta) = 2 \ln \left(\frac{L(\hat{\theta})}{L(\theta)} \right).$$

Se θ é o valor verdadeiro do parâmetro, assintoticamente,

$$D(\theta) = 2[l(\hat{\theta}) - l(\theta)] \sim \chi_{1,(1-\alpha)}^2. \quad (2.16)$$

Assim, o intervalo para θ baseado na deviance (IC_D) será a região

$$\text{IC}_D = \{\theta; D(\theta) - D(\hat{\theta})\} < \chi_{1,(1-\alpha)}^2. \quad (2.17)$$

Lindsey (2001) e Pawitan (2001), por exemplo, apresentam ainda outra forma de se obter a incerteza sobre os parâmetros através da verossimilhança perfilhada. Pawitan

(2001) afirma que existem casos, por exemplo, onde o interesse pode estar em um subconjunto de parâmetros. Por exemplo em um modelo gaussiano, o interesse pode estar em $\boldsymbol{\mu}$, enquanto $\boldsymbol{\sigma}$ é um parâmetro *nuisance* que torna o modelo capaz de acomodar a variabilidade dos dados. Torna-se necessário então “concentrar” a verossimilhança em um único parâmetro eliminando-se o parâmetro *nuisance*, mas considerando-se a incerteza devido à este parâmetro.

O procedimento para eliminar o parâmetro *nuisance* é substituí-lo por sua estimativa de máxima verossimilhança em cada valor fixado do parâmetro de interesse, resultando na função de verossimilhança perfilhada. Dessa verossimilhança constróem-se os intervalos de confiança.

No caso de um vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$, dada a função de verossimilhança conjunta $L(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$ a verossimilhança perfilhada de $\boldsymbol{\theta}_1$ é definida como

$$L(\boldsymbol{\theta}_1) = \max_{\boldsymbol{\theta}_2} L(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$$

onde a maximização é realizada em um valor de $\boldsymbol{\theta}_1$ fixado. Pawitan (2001) observa que no valor fixado de $\boldsymbol{\theta}_1$ a estimativa de máxima verossimilhança de $\boldsymbol{\theta}_2$ é geralmente uma função de $\boldsymbol{\theta}_1$,

$$L(\boldsymbol{\theta}_1) = L(\boldsymbol{\theta}_1, \hat{\boldsymbol{\theta}}_{2\boldsymbol{\theta}_1}). \quad (2.18)$$

Alternativamente à verossimilhança perfilhada, existem outras formas de eliminar parâmetros *nuisance* em inferência estatística que são os métodos da verossimilhança condicional, marginal e verossimilhança perfilhada modificada descritos, por exemplo, em Lindsey (2001) e Pawitan (2001).

Finalmente, a avaliação da incerteza sobre os parâmetros também pode ser feita mediante a aplicação de métodos bayesianos, conforme será apresentado adiante.

2.2 PREDIÇÃO LINEAR ESPACIAL

2.2.1 Conceitos de Predição

Considere o vetor $\underline{S}(\underline{x}) = (S(x_1), S(x_2), \dots, S(x_n))'$ com distribuição multivariada

$$\underline{S}(\underline{x}) \sim N_n(\boldsymbol{\mu}\mathbf{1}; \boldsymbol{\sigma}^2\mathbf{R}),$$

em que \mathbf{R} é uma matriz de ordem $n \times n$ com elementos r_{ij} definidos pela função de correlação $\boldsymbol{\rho}(\|x_i - x_j\|)$ e

$$\underline{Y}(\underline{x}) \sim N_n(\boldsymbol{\mu}\mathbf{1}; \boldsymbol{\sigma}^2\mathbf{V}).$$

De acordo com a teoria apresentada em Diggle e Ribeiro Jr. (2007), o interesse está na predição do processo estacionário $S(\underline{x})$ em uma localização onde $\underline{Y}(\underline{x})$ não foi observado. O vetor $\underline{Y}(\underline{x})$ tem como elementos variáveis aleatórias cujos valores são observados e T é uma variável aleatória cujo valor será predito a partir do vetor observado $\underline{Y}(\underline{x})$.

Um preditor pontual de T é uma função qualquer de $\underline{Y}(\underline{x})$ representada por

$$\hat{T} = t(\underline{Y}(\underline{x})),$$

que tem erro quadrático médio (EQM) de predição definido como

$$\text{EQM}(\hat{T}) = E((T - \hat{T})^2).$$

O $\text{EQM}(\hat{T})$ assume o valor mínimo quando $\hat{T} = E(T|\underline{Y}(\underline{x}))$ pois

$$\begin{aligned} E((T - \hat{T})^2) &= E_{\underline{Y}(\underline{x})}(E_T((T - \hat{T})^2|\underline{Y}(\underline{x}))) \\ &= E_{\underline{Y}(\underline{x})}(\text{Var}_T((T - \hat{T})|\underline{Y}(\underline{x})) + \{E_T((T - \hat{T})|\underline{Y}(\underline{x}))\}^2), \end{aligned} \quad (2.19)$$

em que os subscritos nos dois operadores da esperança indicam que as esperanças são calculadas com relação a $\underline{Y}(\underline{x})$ e T , respectivamente.

Dado que $\text{Var}_T(\hat{T}|\underline{Y}(\underline{x}))$ e $\text{Cov}_T(T|\underline{Y}(\underline{x}), \hat{T}|\underline{Y}(\underline{x}))$ são zero, já que condicionado em $\underline{Y}(\underline{x})$, \hat{T} que é função de $\underline{Y}(\underline{x})$ é constante, tem-se

$$\begin{aligned} \text{Var}_T((T - \hat{T})|\underline{Y}(\underline{x})) &= \text{Var}_T(T|\underline{Y}(\underline{x})) + \text{Var}_T(\hat{T}|\underline{Y}(\underline{x})) - 2\text{Cov}_T(T|\underline{Y}(\underline{x}), \hat{T}|\underline{Y}(\underline{x})) \\ &= \text{Var}_T(T|\underline{Y}(\underline{x})) \end{aligned}$$

e

$$\begin{aligned} E_T((T - \hat{T})|\underline{Y}(\underline{x})) &= E(T|\underline{Y}(\underline{x})) - E(\hat{T}|\underline{Y}(\underline{x})) \\ &= E(T|\underline{Y}(\underline{x})) - \hat{T}, \end{aligned}$$

que substituídas em (2.19) fornecem

$$E((T - \hat{T})^2) = E_{\underline{Y}(\underline{x})}(\text{Var}_T(T|\underline{Y}(\underline{x})) + \{E(T|\underline{Y}(\underline{x})) - \hat{T}\}^2). \quad (2.20)$$

Da Equação (2.20) obtém-se o erro quadrático médio de \hat{T}

$$E((T - \hat{T})^2) = E_{\underline{Y}(\underline{x})}(\text{Var}_T(T|\underline{Y}(\underline{x}))), \quad (2.21)$$

quando $\hat{T} = E(T|\underline{Y}(\underline{x}))$.

Nota-se também que

$$\begin{aligned} E((T - \hat{T})^2) &= \text{Var}(T - \hat{T}) + \{E(T - \hat{T})\}^2 \\ &= \text{Var}(T) + \text{Var}(\hat{T}) - 2\text{Cov}(T, \hat{T}) + \{E(T) - \hat{T}\}^2 \\ &= \text{Var}(T) - 2\text{Cov}(T, \hat{T}). \end{aligned}$$

Então,

$$\text{Var}(T) = E((T - \hat{T})^2) + 2\text{Cov}(T, \hat{T})$$

e, conseqüentemente,

$$E((T - \hat{T})^2) \leq \text{Var}(T)$$

se T e $\underline{Y}(\underline{x})$ são independentes.

Se $S(\underline{x})$ for um processo gaussiano estacionário, os dados $\underline{Y}(\underline{x})$ gerados por um modelo gaussiano estacionário e se T é igual a $S(\underline{x})$, o vetor $(T, \underline{Y}(\underline{x})) = (S(\underline{x}), \underline{Y}(\underline{x}))$ tem distribuição gaussiana multivariada e a distribuição condicional de T dado $\underline{Y}(\underline{x})$ (MOOD; GRAYBILL; BOES, 1974) é também gaussiana com média

$$\mu_{T|\underline{Y}(\underline{x})} = \mu_T + \Sigma_{T\underline{Y}(\underline{x})} \Sigma_{\underline{Y}(\underline{x})\underline{Y}(\underline{x})}^{-1} (\underline{Y}(\underline{x}) - \mu_{\underline{Y}(\underline{x})}),$$

e variância

$$\Sigma_{T|\underline{Y}(\underline{x})} = \Sigma_{TT} - (\underline{x}) \Sigma_{T\underline{Y}(\underline{x})} \Sigma_{\underline{Y}(\underline{x})\underline{Y}(\underline{x})}^{-1} \Sigma_{\underline{Y}(\underline{x})T}.$$

Logo, $[T, \underline{Y}(\underline{x})]$ é gaussiana multivariada com média $\mu \underline{1}$ e matriz de covariância

$$\begin{bmatrix} \sigma^2 & \sigma^2 \underline{r}' \\ \sigma^2 \underline{r} & \sigma^2 \mathbf{V} \end{bmatrix}.$$

em que \underline{r} é um vetor com elementos $r_i = \rho(\|\underline{x} - \underline{x}_i\|)$.

Desta forma, com $\Sigma_{T\underline{Y}(\underline{x})} = \sigma^2 \underline{r}'$ e $\Sigma_{\underline{Y}(\underline{x})\underline{Y}(\underline{x})} = \sigma^2 \mathbf{V}$ o preditor do erro quadrado médio mínimo para $S(\underline{x})$ é:

$$\hat{S}(\underline{x}) = \mu + \underline{r}' \mathbf{V}^{-1} (\underline{Y}(\underline{x}) - \mu \underline{1})$$

com variância de predição

$$\text{Var}(S(\underline{x})|\underline{Y}(\underline{x})) = \sigma^2 (1 - \underline{r}' \mathbf{V}^{-1} \underline{r}).$$

Como a variância de predição não depende de $\underline{Y}(\underline{x})$, da Equação (2.21) tem-se $E((S(\underline{x}) - \hat{S}(\underline{x}))^2) = \text{Var}(S(\underline{x})|\underline{Y}(\underline{x}))$.

Ao escrever o preditor de S em termos de $\hat{S}(\underline{x}_0)$ onde \underline{x}_0 é a localização de predição, pode-se observar que $\underline{r}' \mathbf{V}^{-1}$ nada mais é do que uma combinação linear da média μ e de

$Y(\underline{x}_i)$ de modo que

$$\begin{aligned}\hat{S}(\underline{x}_0) &= \mu + \sum_{i=1}^n a_i(\underline{x}_0)(Y(\underline{x}_i) - \mu) \\ &= \mu + \sum_{i=1}^n a_i(\underline{x}_0)Y(\underline{x}_i) - \sum_{i=1}^n a_i(\underline{x}_0)\mu \\ &= \left\{1 - \sum_{i=1}^n a_i(\underline{x}_0)\right\}\mu + \sum_{i=1}^n a_i(\underline{x}_0)Y(\underline{x}_i)\end{aligned}$$

onde $a_1(\underline{x}_0), a_2(\underline{x}_0), \dots, a_n(\underline{x}_0)$, são denominados pesos de predição.

Segundo Diggle, Ribeiro Jr e Christensen (2003), em muitas aplicações, o foco inferencial pode não estar em $S(\underline{x}_0)$, mas em alguma outra propriedade do processo tal como a média ou valor máximo sobre a área de estudo ou subárea desta, por exemplo. Primeiramente, os autores consideram T qualquer funcional linear de $S(\underline{x})$, ou seja,

$$T = \int_A a(\underline{x})S(\underline{x}) d\underline{x}$$

para alguma função peso $a(\underline{x})$. Como já visto, sob o modelo gaussiano, $[T, \underline{Y}(\underline{x})]$ é gaussiana multivariada e $[T|\underline{Y}(\underline{x}) = \underline{y}(\underline{x})]$ é gaussiana univariada se T é um escalar. A média é dada por

$$E(T|\underline{Y}(\underline{x})) = \int_A a(\underline{x})E(S(\underline{x})|\underline{Y}(\underline{x})) d\underline{x},$$

que resulta

$$\hat{T} = \int_A a(\underline{x})\hat{S}(\underline{x}) d\underline{x}.$$

A variância de $T|\underline{Y}(\underline{x})$ é

$$Var(T|\underline{Y}(\underline{x})) = \int_A \int_A a(\underline{x})a(\underline{x}')Cov(S(\underline{x}), S(\underline{x}')) d\underline{x}d\underline{x}'.$$

Em outras palavras, os autores afirmam que dada a superfície predita $\hat{S}(\underline{x})$, é razoável calcular qualquer propriedade linear desta superfície e usar o resultado como o preditor para a propriedade linear correspondente da superfície verdadeira $S(\underline{x})$. Isto não será válido para propriedades não lineares.

2.2.2 Krigagem

Banerjee, Carlin e Gelfand (2004) colocam que o problema é de predição espacial ótima: dado as observações de um processo estocástico $\underline{Y}(\underline{x}) = (Y(\underline{x}_1), Y(\underline{x}_2), \dots, Y(\underline{x}_n))'$

deseja-se prever a variável $Y(\underline{x})$ em uma localização não observada. Em outras palavras, deseja-se encontrar o melhor preditor do valor de $Y(\underline{x}_0)$ baseado nas observações $y(\underline{x})$ de $Y(\underline{x})$.

Como visto na Subseção 2.2.1, o melhor preditor é o que apresenta o menor erro quadrático médio para $T = S(\underline{x}_0)$ e é dado por

$$\hat{T} = \mu + \underline{r}'\mathbf{V}^{-1}(\underline{Y}(\underline{x}) - \mu\mathbf{1}) \quad (2.22)$$

com variância de predição

$$\text{Var}(T|\underline{Y}(\underline{x})) = \sigma^2(1 - \underline{r}'\mathbf{V}^{-1}\underline{r}). \quad (2.23)$$

Observa-se que devido aos parâmetros do modelo serem quantidades desconhecidas, a estimativa $\hat{\sigma}^2$ obtidas a partir de (2.12) e (2.14) é substituída em (2.22) e (2.23). O preditor \hat{T} é então linear nos dados e este método é conhecido como krigagem simples quando um valor arbitrário ou conhecido de μ é especificado.

No método conhecido como krigagem ordinária, o parâmetro média é tratado como desconhecido e os da covariância são conhecidos. Assim, o preditor é escrito como a combinação linear

$$\hat{T} = \hat{S}(\underline{x}) = \sum_{i=1}^n a_i(\underline{x})Y(x_i)$$

onde $a_i(\underline{x})$, os pesos de krigagem, satisfazem $\sum_{i=1}^n a_i(\underline{x}) = 1$ para qualquer localização de predição. De forma equivalente, em (2.22), substitui-se a média μ pelo estimador de mínimos quadrados generalizados

$$\hat{\mu} = (\mathbf{1}'\mathbf{V}^{-1}\mathbf{1})^{-1}\mathbf{1}'\mathbf{V}^{-1}\underline{Y}(\underline{x})$$

de onde segue que

$$\hat{T} = (\mathbf{1}'\mathbf{V}^{-1}\mathbf{1})^{-1}\mathbf{1}'\mathbf{V}^{-1}\underline{Y}(\underline{x}) + \underline{r}'\mathbf{V}^{-1} [\underline{Y}(\underline{x}) - (\mathbf{1}'\mathbf{V}^{-1}\mathbf{1})^{-1}\mathbf{1}'\mathbf{V}^{-1}\underline{Y}(\underline{x})].$$

Quando uma transformação nos dados originais é feita com o objetivo de que estes passem a seguir uma distribuição gaussiana, em geral os dados transformados são escritos como $\underline{Y}^*(\underline{x}) = h_\lambda(\underline{Y}(\underline{x}))$ onde $h(\cdot)$ é uma função de transformação. Para diferentes especificações de valores de λ tem-se modelos e resultados diferentes e as predições são consideradas como as médias das predições obtidas através dos modelos correspondentes aos diferentes valores de λ . Para transformar de volta os dados $\underline{Y}^*(\underline{x})$ para a escala original, seria necessário calcular a função inversa $h_\lambda^{-1}(\cdot)$ para cada valor de λ e calcular as médias obtendo-se valores aproximados. No caso da função logarítmica o resultado é exato e $h(\cdot) = \ln(\cdot)$ implica $h^{-1}(\cdot) = \exp(\cdot)$.

Supondo, então, $T(\underline{x}) = \exp\{\mu + S(\underline{x})\}$, pode-se escrever

$$T(\underline{x}) = \exp\{\mu\} + \exp\{S(\underline{x})\} = \exp\{\mu\} + T_0(\underline{x}).$$

A distribuição de $S(\underline{x})$ dado $\underline{Y}^*(\underline{x})$ é gaussiana univariada com média $\hat{S}(\underline{x})$ e variância $v(\underline{x})$ dadas por (2.22) e (2.23), respectivamente, substituindo-se $\underline{Y}(\underline{x})$ por $\underline{Y}^*(\underline{x})$. Então, a função geratriz de momentos de $S(\underline{x})$ é

$$\psi_{S(\underline{x})}(a) = E\left(e^{aS(\underline{x})}\right) = e^{a\hat{S}(\underline{x}) + \frac{1}{2}a^2v(\underline{x})} \quad a \in \mathbb{R},$$

a qual para $a = 1$, resulta em

$$\hat{T}_0(\underline{x}) = E(T_0(\underline{x})) = e^{\hat{S}(\underline{x}) + \frac{v(\underline{x})}{2}},$$

e para $a = 2$ em

$$E((T_0(\underline{x}))^2) = e^{2\hat{S}(\underline{x}) + 2v(\underline{x})},$$

de onde se obtém a variância de predição:

$$\text{Var}(T_0(\underline{x})|\underline{Y}^*) = e^{2\hat{S}(\underline{x}) + v(\underline{x})} [e^{v(\underline{x})} - 1].$$

Como já mencionado anteriormente, a transformação de volta dos valores $\underline{Y}^*(\underline{x})$ para a escala original pode, ainda, ser feita por simulação como será descrito adiante.

2.3 INFERÊNCIA BAYESIANA PARA PREDIÇÃO ESPACIAL

2.3.1 Distribuição Preditiva Bayesiana

Na inferência bayesiana os parâmetros do modelo são considerados variáveis aleatórias. Na estimação pontual desses parâmetros é possível incorporar formalmente nas predições, essas incertezas. Para isto, considera-se um vetor de parâmetros $\underline{\theta}$ e um vetor aleatório $\underline{Y}(\underline{x})$ com distribuição de probabilidade determinada pela função $P(\underline{Y}(\underline{x})|\underline{\theta})$. Ao desprezar o termo constante da função $P(\underline{Y}(\underline{x})|\underline{\theta})$, a função de verossimilhança $L(\underline{\theta}; \underline{Y}(\underline{x}))$ pode ser escrita como

$$L(\underline{\theta}; \underline{Y}(\underline{x})) \propto |\sigma^2\mathbf{R} + \tau^2\mathbf{I}|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}(\underline{Y}(\underline{x}) - \mathbf{D}\underline{\beta})'(\sigma^2\mathbf{R} + \tau^2\mathbf{I})^{-1}(\underline{Y}(\underline{x}) - \mathbf{D}\underline{\beta})\right\}. \quad (2.24)$$

De acordo com Gilks, Richardson e Spiegelhalter (1996) e Gamerman e Lopes (2006), para se fazer inferência é necessário construir a distribuição conjunta dos vetores aleatórios $\underline{Y}(\underline{x})$ e $\underline{\theta}$ em duas partes: uma distribuição a priori $P(\underline{\theta})$ e uma função de verossimilhança $P(\underline{Y}(\underline{x})|\underline{\theta})$, ou seja:

$$P(\underline{Y}(\underline{x}), \underline{\theta}) = P(\underline{Y}(\underline{x})|\underline{\theta})P(\underline{\theta}).$$

Gill (2002) acrescenta que a expressão $P(\underline{\theta})$ é uma afirmação formalizada na forma de uma distribuição de probabilidade do conhecimento prévio sobre $\underline{\theta}$ antes de se observar os dados. A idéia básica é especificar uma distribuição a *priori* para $\underline{\theta}$ que descreva o que se sabe em termos probabilísticos sobre $\underline{\theta}$ e, portanto, especificando premeditadamente informação a *priori* e de incertezas.

O Teorema de Bayes é aplicado para determinar a distribuição de $\underline{\theta}$ condicionada aos dados $\underline{Y}(\underline{x})$:

$$P(\underline{\theta}|\underline{Y}(\underline{x})) = \frac{P(\underline{\theta}, \underline{Y}(\underline{x}))}{P(\underline{Y}(\underline{x}))} = \frac{P(\underline{Y}(\underline{x})|\underline{\theta})P(\underline{\theta})}{P(\underline{Y}(\underline{x}))} = \frac{P(\underline{Y}(\underline{x})|\underline{\theta})P(\underline{\theta})}{\int P(\underline{\theta}, \underline{Y}(\underline{x}))d\underline{\theta}}, \quad (2.25)$$

e é chamada de distribuição à *posteriori* de $\underline{\theta}$. Como o denominador da expressão não depende de $\underline{\theta}$, pode ser considerado uma constante normalizadora, e ao desprezá-la a Equação (2.25) pode ser escrita como:

$$P(\underline{\theta}|\underline{Y}(\underline{x})) \propto P(\underline{Y}(\underline{x})|\underline{\theta})P(\underline{\theta}). \quad (2.26)$$

Para Gill (2002) o lado direito da Equação (2.26) implica que a inferência para $\underline{\theta}$ é um compromisso entre informação a *priori* e a informação fornecida pelos dados, enquanto o lado esquerdo fornece a distribuição atualizada para $\underline{\theta}$ após condicionar aos dados.

Em alguns casos é possível derivar analiticamente uma distribuição a *posteriori* de modo que se apresente como uma distribuição conhecida, caso contrário, utilizam-se métodos numéricos computacionalmente intensivos. Mas, como afirma Gamerman e Lopes (2006), a obtenção da distribuição a *posteriori* é um passo importante, mas não o final. É preciso ser capaz de se extrair informações significativas desta distribuição e isto está relacionado com a avaliação de resumos estatísticos tais como média, mediana ou moda, ou intervalos de probabilidade. Estas quantidades podem ser expressas em termos de esperanças a *posteriori* de funções de $\underline{\theta}$ (GILKS; RICHARDSON; SPIEGELHALTER, 1996). A esperança a *posteriori* de uma função $g(\underline{\theta})$ é

$$E(g(\underline{\theta})|\underline{Y}(\underline{x})) = \frac{\int g(\underline{\theta})P(\underline{\theta})P(\underline{Y}(\underline{x})|\underline{\theta})d\underline{\theta}}{\int P(\underline{\theta})P(\underline{Y}(\underline{x})|\underline{\theta})d\underline{\theta}}, \quad (2.27)$$

ou

$$E(g(\underline{\theta})|\underline{Y}(\underline{x})) \propto \int f(\underline{\theta})P(\underline{\theta})P(\underline{Y}(\underline{x})|\underline{\theta})d\underline{\theta}.$$

definidas no domínio de variação das *priori*.

A resolução numérica destas integrais incluem aproximações analíticas tais como aproximação de Laplace (Nott, Fielding e Leonte (2009), Gelman et al. (2003), Gilks, Richardson e Spiegelhalter (1996)), aproximação numérica como por exemplo quadratura de Gauss-Hermite (Paulino, Turkman e Murteira (2003), Gamerman e Lopes (2006)) e integração Monte Carlo, incluindo cadeias de Markov de Monte Carlo - MCMC (Gamerman e Lopes (2006), Rue e Held (2005), Lee (2004), Gelman et al. (2003), Gill (2002), Gilks, Richardson e Spiegelhalter (1996)). Neste trabalho foram adotados os dois últimos procedimentos.

Quanto às *priori*, Ribeiro Jr e Diggle (1999) mencionam que a escolha é uma questão delicada em inferência bayesiana. *Priori* que levam a uma *posteriori* da mesma família de distribuições são chamadas *priori conjugadas*. Essas *priori* podem ser computacionalmente convenientes mas não deveriam ser escolhidas somente por isso. Dois casos extremos para a escolha da *priori* são: quando os parâmetros são perfeitamente conhecidos as *priori* podem ser vistas como distribuições degeneradas nos valores dos parâmetros; quando o conhecimento da *priori* sobre os parâmetros é vaga podem ser adotadas *priori* não informativas, *flats* ou impróprias.

Por outro lado, de acordo com Ribeiro Jr e Diggle (1999) a base da predição bayesiana é a distribuição preditiva $P(\underline{S}(\underline{x})|\underline{Y}(\underline{x}))$. Esta distribuição leva em consideração a incerteza sobre os parâmetros calculando, por exemplo, a média da distribuição condicional $P(\underline{S}(\underline{x})|\underline{Y}(\underline{x}), \underline{\theta})$, sobre o espaço dos parâmetros, com pesos dados pela distribuição à *posteriori* dos parâmetros do modelo $P(\underline{\theta}|\underline{Y}(\underline{x}))$:

$$\begin{aligned} P(\underline{S}(\underline{x})|\underline{Y}(\underline{x})) &= \int P(\underline{S}(\underline{x}), \underline{\theta}|\underline{Y}(\underline{x})) d\underline{\theta} \\ &= \int P(\underline{S}(\underline{x})|\underline{Y}, \underline{\theta}(\underline{x}))P(\underline{\theta}|\underline{Y}(\underline{x})) d\underline{\theta}. \end{aligned}$$

Cabe ressaltar que podem ser calculadas outras estatísticas de interesse ou mais apropriadas, como a mediana ou moda, a partir da distribuição preditiva.

2.3.2 Cadeias de Markov

A teoria apresentada neste tópico segue àquelas apresentadas em Gelman et al. (2003). Para esses autores uma *cadeia de Markov* é uma sequência de variáveis aleatórias $\underline{\theta}_1, \underline{\theta}_2, \dots$ para os quais, para qualquer t , a distribuição de $\underline{\theta}_t$ dado os $\underline{\theta}$ anteriores depende somente do anterior, $\underline{\theta}_{t-1}$.

A simulação de uma *cadeia de Markov* (*cadeia de Markov de Monte Carlo* ou MCMC) é definida como um método baseado na amostragem de valores de $\underline{\theta}$ de distribuições aproximadas, corrigindo-os para uma melhor aproximação da distribuição posteriori, $P(\underline{\theta}|\underline{Y}(\underline{x}))$. Os valores de $\underline{\theta}$ são amostrados sequencialmente, com a distribuição dependente do último valor amostrado. Estes valores amostrados formam então uma *cadeia de Markov*.

O MCMC é usado quando não é possível, ou não é computacionalmente eficiente, amostrar $\underline{\theta}_t$ diretamente de $P(\underline{\theta}|\underline{Y}(\underline{x}))$. Amostra-se iterativamente de forma tal que a cada passo do processo espera-se que os valores sejam amostrados de uma distribuição que seja a mais próxima possível de $P(\underline{\theta}|\underline{Y}(\underline{x}))$.

O princípio básico do MCMC (GILL, 2002) é que, se uma cadeia iterativa de valores consecutivos, gerados computacionalmente, pode ser construído com cuidado suficiente, estimativas empíricas podem ser obtidas dos últimos valores da cadeia. Em geral, inicia-se a cadeia após um período descarte de iterações, denominado período de aquecimento (*burn-in*) da cadeia deixando-a “correr” até que se aproxime da distribuição

estacionária.

As técnicas MCMC utilizadas neste trabalho foram o algoritmo Metropolis-Hastings e amostrador de Gibbs. Para o algoritmo Metropolis-Hastings (GILKS; RICHARDSON; SPIEGELHALTER, 1996), em cada estado $\underline{\theta}^{(t)}$ da cadeia, o próximo estado $\underline{\theta}^{(t+1)}$ é escolhido amostrando-se um valor $\underline{\theta}'$ de uma distribuição proposta $q(\cdot|\underline{\theta}^{(t)})$ que pode depender do valor atual $\underline{\theta}^{(t)}$. O valor $\underline{\theta}'$ será aceito com probabilidade $\alpha(\underline{\theta}^{(t)}, \underline{\theta}')$, em que

$$\alpha(\underline{\theta}, \underline{\theta}') = \min \left(1, \frac{P(\underline{Y}(\underline{x})|\underline{\theta}')P(\underline{\theta}')q(\underline{\theta}|\underline{\theta}')}{P(\underline{Y}(\underline{x})|\underline{\theta})P(\underline{\theta})q(\underline{\theta}'|\underline{\theta})} \right).$$

Se o novo valor for aceito, o próximo estado torna-se $\underline{\theta}^{(t+1)} = \underline{\theta}'$. Caso contrário, a cadeia “não se move”, ou seja, $\underline{\theta}^{(t+1)} = \underline{\theta}^{(t)}$.

Ehlers (2006), Gilks, Richardson e Spiegelhalter (1996), Gamerman e Lopes (2006) desenvolveram os seguintes passos de um algoritmo com este propósito de aceitar/rejeitar valores:

1. Escolher o número n de etapas (número de simulações ou número de passos da cadeia);
2. iniciar o contador de iterações t , fazendo $t = 1$;
3. especificar um valor inicial para $\underline{\theta}^{(t)}$;
4. gerar um valor $\underline{\theta}'$ de uma distribuição proposta $q(\underline{\theta}'|\underline{\theta})$. Esta distribuição deve ser escolhida com cautela para garantir a eficiência do algoritmo;
5. gerar $u \sim U(0;1)$ (*priori*);
6. determinar a probabilidade de aceitação $\alpha(\underline{\theta}; \underline{\theta}')$ dada por:

$$\alpha(\underline{\theta}; \underline{\theta}') = \min \left\{ 1; \frac{P(\underline{Y}|\underline{\theta}')P(\underline{\theta}')q(\underline{\theta}|\underline{\theta}')}{P(\underline{Y}|\underline{\theta})P(\underline{\theta})q(\underline{\theta}'|\underline{\theta})} \right\}$$

onde $P(\underline{\theta}|\underline{Y})$, a distribuição a *posteriori* é a distribuição de interesse;

7. se $u < \alpha(\underline{\theta}; \underline{\theta}')$, o valor de $\underline{\theta}'$ é aceito e $\underline{\theta}^{(t)} = \underline{\theta}'$. Caso contrário, $\underline{\theta}'$ é rejeitado e $\underline{\theta}^{(t)} = \underline{\theta}^{(t-1)}$;
8. se $t = n$ (é o último passo) encerrar o processo, senão:
9. incrementar o contador t fazendo $t = t + 1$.
10. voltar para o passo 4;
11. FIM.

Segundo Ehlers (2006) no amostrador de Gibbs, as transições de estado são feitas de acordo com as distribuições condicionais completas $P(\theta_i|\underline{\theta}_{-i})$, obtidas a partir da

distribuição conjunta como:

$$P(\theta_i | \underline{\theta}_{-i}) = \frac{P(\underline{\theta})}{\int P(\underline{\theta}) d\theta_i} \quad (2.28)$$

onde $\underline{\theta}_{-i} = \mathbf{1}_{-i}\underline{\theta}$, com $\mathbf{1}_{-i}$ uma matriz identidade excluída a i -ésima linha e $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_p)'$, $i = 1, 2, \dots, p$, sendo p neste caso o número de parâmetros a serem estimados, ou seja:

$$\theta_{-i} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_{i-1} \\ \theta_{i+1} \\ \vdots \\ \theta_p \end{bmatrix}_{(p-1) \times 1} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{bmatrix}_{(p-1) \times p} \cdot \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \vdots \\ \theta_p \end{bmatrix}_{p \times 1}$$

Em outras palavras, pode-se dizer que a distribuição condicional completa é a distribuição da i -ésima componente de $\underline{\theta}$ condicionada em todas as outras componentes e a Equação (2.28) pode ser escrita como $P(\underline{\theta}) = P(\theta_i | \underline{\theta}_{-i})P_{\theta_{-i}}(\underline{\theta}_{-i})$.

Desta forma, deseja-se gerar uma amostra de $P(\underline{\theta})$ onde a transição se dá da seguinte forma:

- i. Especificar valores para um vetor inicial $\underline{\theta}^{(0)}$;
- ii. calcular a probabilidade de θ_1 condicionada aos θ' s restantes com os valores especificados inicialmente;
- iii. nas transições de θ_{r-1} para θ_r , $r = 2, 3, \dots, p$, calcular, para θ_r , a probabilidade $P(\theta_r | \underline{\theta}_{-r})$ substituindo em $\underline{\theta}_{-r}$ o valor de θ_{r-1} calculado no passo anterior.

Paulino, Turkman e Murteira (2003) afirmam que este procedimento é o adotado pelo algoritmo de Metropolis-Hastings com $q(\underline{\theta}; \underline{\theta}')$ e $q(\underline{\theta}'; \underline{\theta})$ positivas se, e somente se, $\underline{\theta}$ e $\underline{\theta}'$ diferirem no máximo em uma coordenada.

Se as distribuições condicionais completas forem conhecidas, então o algoritmo será desenvolvido seguindo os passos:

1. Iniciar o contador de iterações em $t = 0$ e definir o número n de iterações;
2. especificar valores iniciais para $\underline{\theta}^{(0)}$;
3. avançar t fazendo $t = t + 1$ e obter $\underline{\theta}^{(t)}$ a partir de $\underline{\theta}^{(t-1)}$ por geração sucessiva de valores como:

$$\begin{aligned} \theta_1^{(t)} &\sim P(\theta_1 | \theta_2^{(t-1)}, \theta_3^{(t-1)}, \theta_4^{(t-1)}, \dots, \theta_p^{(t-1)}) \\ \theta_2^{(t)} &\sim P(\theta_2 | \theta_1^{(t)}, \theta_3^{(t-1)}, \theta_4^{(t-1)}, \dots, \theta_p^{(t-1)}) \\ \theta_3^{(t)} &\sim P(\theta_3 | \theta_1^{(t)}, \theta_2^{(t)}, \theta_4^{(t-1)}, \dots, \theta_p^{(t-1)}) \\ &\vdots \\ \theta_p^{(t)} &\sim P(\theta_p | \theta_1^{(t)}, \theta_2^{(t)}, \theta_3^{(t)}, \dots, \theta_{p-1}^{(t)}) \end{aligned}$$

4. Se $t = n$, encerrar;
5. Voltar a etapa 3.

Uma vez que métodos de simulação de cadeias de *Markov* são utilizados, faz-se necessário um estudo sobre a convergência das cadeias. Diagnósticos de convergência são discutidos em Gilks, Richardson e Spiegelhalter (1996), Gamerman e Lopes (2006), Gill (2002), Carlin e Louis (2009), entre outros.

Para Gilks, Richardson e Spiegelhalter (1996), a tarefa de monitorar a convergência é estimar o quanto a inferência baseada em simulações de *cadeia de Markov* diferem da distribuição posteriori e o método adotado, inspirado na análise de variância, é formar uma super e uma subestimativa da variância da distribuição posteriori, com a propriedade que as estimativas serão grosseiramente iguais em convergência, mas não antes.

Análise de convergência pode ser feita, por exemplo, calculando-se as autocorrelações dos parâmetros à *posteriori* (GILL, 2002). Decréscimos nas autocorrelações à medida que as defasagens aumentam são indicativos de convergência da cadeia.

Gráficos da trajetória da cadeia também são utilizados na avaliação. Espera-se que a série ocorra de forma estável em torno da linha média. Observa-se a não convergência quando a trajetória se afasta seguindo uma tendência que não seja a linha média.

Os gráficos das densidades estimadas dos parâmetros à posteriori auxiliam no sentido de verificação de multimodalidade, indicativo de não convergência da cadeia.

O teste de Geweke apresentados em Gill (2002) e Gamerman e Lopes (2006), por exemplo, baseia-se na comparação de alguma proporção do início da cadeia após o período de aquecimento com alguma proporção do fim da cadeia. É um teste de diferença de médias usando uma aproximação assintótica para o erro padrão da diferença.

Considere uma cadeia com $m + n$ iterações, $\theta_1, \dots, \theta_{m+n}$, uma função de interesse $f()$ e as médias

$$\bar{f}_b = \frac{1}{n_b} \sum_{i=m+1}^{m+n_b} f(\theta_i) \quad \text{e} \quad \bar{f}_a = \frac{1}{n_a} \sum_{i=m+n-n_a+1}^{m+n} f(\theta_i)$$

onde $n_b + n_a < n$. Se m é o comprimento do período de aquecimento, então \bar{f}_a e \bar{f}_b são as médias ergódicas (não sensíveis às condições iniciais) no fim e começo do período de convergência e deveriam se comportar de forma similar. À medida que n aumenta e as razões n_a e n_b permanecem fixas então

$$z_G = \frac{\bar{f}_a - \bar{f}_b}{\sqrt{\hat{V}_{\text{ar}}(f_a) + \hat{V}_{\text{ar}}(f_b)}} \rightarrow N(0; 1)$$

É sugerido o uso de $n_b = 0,1n$ e $n_a = 0,5n$ e estimadores da densidade espectral são usadas para as variâncias. Por fim, valores grandes para z_G indicam falta de convergência mas valores pequenos não implicam convergência. Uma preocupação deve ocorrer para valores maiores que 2 em termos absolutos.

Outra forma de avaliação da convergência apresentada pelos mesmos autores é o

diagnóstico de uma sequência de cadeias, denominado Diagnóstico de Sequência Múltipla de Gelman e Rubin. Este diagnóstico é baseado na comparação de um conjunto de cadeias com diferentes pontos que são super dispersos com relação à distribuição posteriori. Baseia-se também na teoria da aproximação gaussiana para as posteriores marginais usando testes como ANOVA e diagnóstico com distribuição *t-Student*.

Para a realização deste teste considera-se m cadeias paralelas de comprimento n :

$$\begin{array}{cccc} \theta_{(1)}^{[1]} & , & \theta_{(1)}^{[2]} & , \dots , & \theta_{(1)}^{[n]} \\ \theta_{(2)}^{[1]} & , & \theta_{(2)}^{[2]} & , \dots , & \theta_{(2)}^{[n]} \\ \vdots & & & & \\ \theta_{(m)}^{[1]} & , & \theta_{(m)}^{[2]} & , \dots , & \theta_{(m)}^{[n]} \end{array}$$

As variâncias entre cadeias, V_E , e dentro das cadeias, V_D , são dadas por

$$V_E = \frac{n}{m-1} \sum_{j=1}^m (\bar{\theta}_{(j)} - \bar{\bar{\theta}})^2 \quad \text{e} \quad V_D = \frac{1}{m(n-1)} \sum_{j=1}^m \sum_{i=1}^n (\bar{\theta}_{(j)}^{[i]} - \bar{\theta}_{(j)})^2$$

onde $\bar{\bar{\theta}}$ é a média das médias das cadeias já que todas tem o mesmo comprimento e $\bar{\theta}_{(j)}$ é a média dos n valores da j -ésima cadeia. Sob convergência, todos os mn valores são selecionados da posteriori e a variância $\text{Var}(\theta)$ pode ser estimada por

$$\hat{V}\text{ar}(\theta) = \left(1 + \frac{1}{n}\right) V_D + \left(\frac{1}{n}\right) V_E.$$

Se as cadeias ainda não convergiram, os valores iniciais estarão influenciando as trajetórias. Devido às suas superdispersão, os valores forçarão $\hat{V}\text{ar}(\theta)$ a superestimar $\text{Var}(\theta)$. Por outro lado, após a convergência, V_D tenderá a subestimar $\text{Var}(\theta)$ porque cada cadeia ainda não terá atravessado todo o espaço. Desta forma, um indicador de convergência é o fator *redução de escala estimado*:

$$\sqrt{\hat{F}_{RE}} = \sqrt{\frac{\hat{V}\text{ar}(\theta)}{V_D}}.$$

De acordo com a literatura, valores de $\sqrt{\hat{F}_{RE}}$ próximos a 1, mais precisamente inferiores a 1,2 são aceitáveis para convergência.

2.3.3 Intervalo de Credibilidade

As distribuições de probabilidade a *posteriori* dos parâmetros no vetor θ contêm todas as informações sobre estes parâmetros. Resumos estatísticos destas distribuições

podem ser calculados como média, mediana e moda representando medidas de locação e a variação pode ser resumida calculando-se o desvio padrão e quantis. Segundo Gelman et al. (2003) as médias, por exemplo, são as esperanças a *posteriori* dos parâmetros, as modas podem ser interpretadas como os únicos valores “mais prováveis”, dado os dados e o modelo, mas além disso, é importante descrever a incerteza a *posteriori*. Para isto, os autores apresentam os quantis das distribuições a *posteriori* ou um intervalo central da probabilidade a *posteriori*, que corresponde no caso de um intervalo de $100(1 - \alpha)\%$, ao intervalo dos valores acima e abaixo do qual encontra-se exatamente $100(\alpha/2)\%$ da probabilidade a *posteriori*. Estes intervalos, denominados intervalo central a *posteriori* ou intervalo de credibilidade são o análogo a intervalo de confiança na teoria clássica de inferência. Carlin e Louis (2009) definem um intervalo de credibilidade de $100(1 - \alpha)\%$ para $\underline{\theta}$ como um subconjunto C de Θ tal que

$$P(C|Y) = \int_C P(\underline{\theta}|Y)d\underline{\theta} \geq 1 - \alpha.$$

onde o sinal “ \geq ” é usado para considerar conjuntos discretos nos quais pode não ser possível a obtenção de um intervalo de cobertura de exatamente $1 - \alpha$ de probabilidade.

Gelman et al. (2003) apresentam uma outra maneira de resumir a incerteza a *posteriori* que é calcular uma região de alta densidade a *posteriori*. Esta região corresponde a região de valores que contém $100(1 - \alpha)\%$ da probabilidade a *posteriori*, tem a característica que a densidade dentro da região nunca é menor que a densidade fora da região e será idêntica ao intervalo a *posteriori* se a distribuição a *posteriori* é unimodal e simétrica. Carlin e Louis (2009) definem um intervalo de credibilidade de *mais alta densidade a posteriori* como

$$C = \underline{\theta} \in \Theta : P(\underline{\theta}|Y) \geq k(\alpha)$$

onde $k(\alpha)$ é a maior constante satisfazendo

$$P(C|Y) \geq 1 - \alpha.$$

2.4 MODELO MULTIVARIADO

Ao se realizar uma pesquisa, os dados coletados por amostragem em localizações espaciais, frequentemente são multivariados, ou seja, mais de uma variável é mensurada em cada localização. As técnicas multivariadas encontradas em Johnson e Wichern (1998) e Reis (1997), por exemplo, são usadas segundo Bailey e Gatrell (1995) para fins de redução dos dados e exploração do espaço do atributo multidimensional, com o objetivo de

identificar um número pequeno de sub-dimensões de interesse dado por combinações dos atributos, que podem então ser examinados sob uma perspectiva espacial, explorando padrões espaciais e relacionamentos, ou para uso em classificação e discriminação espacial. Wackernagel (1998) introduz geoestatística de uma perspectiva multivariada, Schabenberger e Pierce (2001) apresentam teoria de modelos estatísticos para dados espaciais uni e multivariados com enfoque para plantas e ciências do solo e Silva (2000), Silva et al. (2003) apresentam estudos sobre a variabilidade espacial de atributos do solo que são importantes para a agricultura de precisão.

Da perspectiva de Diggle e Ribeiro Jr. (2007), procura-se descrever a distribuição espacial conjunta das variáveis; a distribuição condicional de uma variável resposta de interesse dado uma ou mais covariáveis referenciadas espacialmente, ou ainda, como em Bognola et al. (2008), modelos multivariados são utilizados quando a variável de interesse é de difícil obtenção ou de custo elevado mas existe uma variável secundária correlacionada com esta, de fácil obtenção de modo a combinar poucas mensurações da variável de interesse com um número maior de mensurações da outra.

Em se tratando de mais de uma única variável resposta, considera-se o processo estocástico $\underline{Y}(\underline{x}) = (Y_1(\underline{x}), \dots, Y_p(\underline{x}))'$ de dimensão p , $\underline{x} \in \mathbb{R}^2$, como sendo o valor observado da variável resposta associada a localização \underline{x}_i da forma $\underline{y}(\underline{x}_i) = (y_1(\underline{x}_i), \dots, y_p(\underline{x}_i))'$. Desta maneira, de acordo com Diggle e Ribeiro Jr. (2007), para duas localizações \underline{x} e \underline{x}' , a função de covariância de $\underline{Y}(\underline{x})$ será uma função matriz $\Gamma(\underline{x}, \underline{x}')$ dada por

$$\Gamma(\underline{x}, \underline{x}') = \begin{bmatrix} \text{Cov}(Y_1(\underline{x}), Y_1(\underline{x}')) & \text{Cov}(Y_1(\underline{x}), Y_2(\underline{x}')) & \cdots & \text{Cov}(Y_1(\underline{x}), Y_p(\underline{x}')) \\ \text{Cov}(Y_1(\underline{x}), Y_2(\underline{x}')) & \text{Cov}(Y_2(\underline{x}), Y_2(\underline{x}')) & \cdots & \text{Cov}(Y_2(\underline{x}), Y_p(\underline{x}')) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(Y_1(\underline{x}), Y_p(\underline{x}')) & \text{Cov}(Y_2(\underline{x}), Y_p(\underline{x}')) & \cdots & \text{Cov}(Y_p(\underline{x}), Y_p(\underline{x}')) \end{bmatrix},$$

denominada matriz de covariância cruzada em que $\gamma_{jk}(\underline{x}, \underline{x}') = \gamma_{kj}(\underline{x}, \underline{x}')$, $k, j = 1, \dots, p$, ou seja, é uma matriz simétrica.

Agora, se $\underline{Y}(\underline{x})$ for um processo estacionário, $\gamma_{jj}(\underline{x}, \underline{x}) = \text{Var}(Y_j(\underline{x})) = \sigma_j^2$ e para $j \neq k$, $\gamma_{kj}(\underline{x}, \underline{x}')$ depende somente de $u = \|\underline{x} - \underline{x}'\|$, de modo que:

$$\Gamma(\underline{x}, \underline{x}') = \begin{bmatrix} \sigma_1^2 & \text{Cov}(Y_1(\underline{x}), Y_2(\underline{x}')) & \cdots & \text{Cov}(Y_1(\underline{x}), Y_p(\underline{x}')) \\ \text{Cov}(Y_2(\underline{x}), Y_1(\underline{x}')) & \sigma_2^2 & \cdots & \text{Cov}(Y_2(\underline{x}), Y_p(\underline{x}')) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(Y_p(\underline{x}), Y_1(\underline{x}')) & \text{Cov}(Y_p(\underline{x}), Y_2(\underline{x}')) & \cdots & \sigma_p^2 \end{bmatrix}.$$

A função de correlação de $\underline{Y}(\underline{x})$ é definida então como a função $\mathbf{R}(u)$ uma matriz dada por:

$$\mathbf{R}(u) = \begin{bmatrix} \rho_1(u) & \frac{\gamma_{12}(u)}{\sigma_1 \sigma_2} & \cdots & \frac{\gamma_{1p}(u)}{\sigma_1 \sigma_p} \\ \frac{\gamma_{21}(u)}{\sigma_2 \sigma_1} & \rho_2(u) & \cdots & \frac{\gamma_{2p}(u)}{\sigma_2 \sigma_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\gamma_{p1}(u)}{\sigma_p \sigma_1} & \frac{\gamma_{p2}(u)}{\sigma_p \sigma_2} & \cdots & \rho_p(u) \end{bmatrix},$$

em que para $j = k$, as funções $\rho_{jj}(u)$ são funções de correlação do processo univariado de

$Y_j(\underline{x})$ e satisfazem $\rho_{jk}(u) = \rho_{kj}(-u)$. Para $j \neq k$ as funções $\rho_{jk}(u)$, funções de correlação cruzadas de $\underline{Y}(\underline{x})$, não são necessariamente simétricas mas satisfazem $\rho_{jk}(u) = \rho_{kj}(-u)$.

O grande problema encontrado na modelagem de dados multivariados é garantir que a matriz de covariância $\Gamma(\underline{x}, \underline{x}')$ seja definida positiva, $\forall \underline{x}, \underline{x}' \in \mathbb{R}^2$. Isto pode ser garantido mediante a construção de combinações lineares de componentes independentes cuja descrição podem ser encontradas em Diggle e Ribeiro Jr. (2007), Schmidt e Gelfand (2003), Banerjee, Carlin e Gelfand (2004), Schmidt e Sansó (2006). O modelo de correio-linearização linear é apresentado em em Goovaerts (1997, Seção 4.2.3), Wackernagel (1998), Schmidt e Sansó (2006). Procedimentos baseados em separabilidade, médias móveis, convolução e correio-linearização, são descritos em Banerjee, Carlin e Gelfand (2004).

2.4.1 Definição do Modelo Multivariado

Considere que as variáveis Y_1, Y_2, \dots, Y_p , sigam qualquer distribuição e que $\underline{S}(\underline{x}) = (S_1(\underline{x}), S_2(\underline{x}), \dots, S_n(\underline{x}))$ seja um processo gaussiano estacionário não observado (latente). Muitas vezes a distribuição conjunta de probabilidade de $\underline{Y}(\underline{x}) = (Y_1(\underline{x}), Y_2(\underline{x}), \dots, Y_p(\underline{x}))'$ não é conhecida, mas ao escrever

$$[\underline{Y}(\underline{x}), \underline{S}(\underline{x})] = [\underline{S}(\underline{x})][\underline{Y}(\underline{x})|\underline{S}(\underline{x})] \quad (2.29)$$

e, considerando como exemplo dados bivariados, $\underline{Y}(\underline{x}) = (Y_1(\underline{x}), Y_2(\underline{x}))'$, observa-se que para um conjunto finito de pontos, $\underline{S}(\underline{x})$ é gaussiana multivariada e $\underline{Y}(\underline{x})|\underline{S}(\underline{x})$ é um produto de densidades univariadas, já que $Y_1(\underline{x})$ e $Y_2(\underline{x})$ que inicialmente são dependentes, dado o conhecimento de $S_1(\underline{x})$ e $S_2(\underline{x})$, tornam-se independentes. Logo, com um processo gaussiano multivariado latente $\underline{S}(\underline{x})$ e uma independência condicional ($\underline{Y}(\underline{x})|\underline{S}(\underline{x})$) é possível construir um modelo multivariado para $\underline{Y}(\underline{x})$ sem a necessidade de que este tenha distribuição gaussiana, bastando para isso integrar (2.29) com relação a $\underline{S}(\underline{x})$ obtendo-se a distribuição marginal de $\underline{Y}(\underline{x})$

$$[\underline{Y}(\underline{x})] = \int [\underline{Y}(\underline{x}), \underline{S}(\underline{x})] d\underline{S}(\underline{x}) = \int [\underline{S}(\underline{x})][\underline{Y}(\underline{x})|\underline{S}(\underline{x})] d\underline{S}(\underline{x}).$$

Portanto este produto, integrado em relação a $\underline{S}(\underline{x})$, resulta numa distribuição de probabilidade que é o objetivo da modelagem. No caso em que $[\underline{Y}(\underline{x})|\underline{S}(\underline{x})]$ é assumida gaussiana, $[\underline{Y}(\underline{x})]$ será gaussiana multivariada.

Seguindo a proposta de Diggle e Ribeiro Jr. (2007), Bognola et al. (2008), no caso bivariado o modelo é dado por:

$$\begin{cases} Y_1(x_i) = \mu_1(x_i) + S_1(x_i) + Z_1(x_i) \\ Y_2(x_i) = \mu_2(x_i) + S_2(x_i) + Z_2(x_i) \end{cases} \quad (2.30)$$

em que as respostas

$$\underline{Y}(\underline{x}) = (Y_1(\underline{x}), Y_2(\underline{x}))' = (Y_1(\underline{x}_1), Y_1(\underline{x}_2), \dots, Y_1(\underline{x}_{n_1}), Y_2(\underline{x}_1), Y_2(\underline{x}_2), \dots, Y_2(\underline{x}_{n_2}))'$$

são medidas nas localizações $\underline{x}_i, \underline{x}_{i'}$, $i, i' = 1, \dots, n_j$, $j = 1, 2$, $\underline{S}(\underline{x}) = (S_1(\underline{x}), S_2(\underline{x}))$ é considerado um processo gaussiano estacionário com média $\underline{\mu} = (0, 0)'$, variâncias $\sigma_j^2 = \text{Var}(S_j(\underline{x}))$ e estrutura de correlação dada por

$$\begin{bmatrix} \text{Corr}(S_1(\underline{x}), S_1(\underline{x} - \underline{u})) & \text{Corr}(S_1(\underline{x}), S_2(\underline{x} - \underline{u})) \\ \text{Corr}(S_1(\underline{x}), S_2(\underline{x} - \underline{u})) & \text{Corr}(S_2(\underline{x}), S_2(\underline{x} - \underline{u})) \end{bmatrix},$$

e $Z_j \sim N(0; \tau^2)$.

É assumido para $Y_1(\underline{x}_i)$ que $\mu_1(\underline{x}_i)$ é o efeito fixo e o efeito aleatório $S_1(\underline{x}_i)$ é decomposto em dois outros efeitos aleatórios: $S_{01}^*(\underline{x}_i)$ comum a $Y_1(\underline{x}_i)$ e $Y_2(\underline{x}_{i'})$, e $S_1^*(\underline{x}_i)$ específico a $Y_1(\underline{x}_i)$. O mesmo ocorre para $Y_2(\underline{x}_{i'})$. Mas, como em geral $Y_j(\underline{x}_i)$ são medidas em escalas diferentes, é necessário fazer uma padronização que é dada por $\sigma_{0j}U_0(\underline{x}_i)$ e $\sigma_j U_j(\underline{x}_i)$ sendo $U_0(\underline{x}_i)$ e $U_j(\underline{x}_i)$ efeitos aleatórios padronizados que têm correlação espacial, logo adimensionais, com as unidades preservadas nas constantes padronizadoras σ_{0j} e σ_j . Estes efeitos apresentam distribuição gaussiana com vetor de médias iguais a zero e matriz de covariância com variâncias unitárias na diagonal principal e covariâncias cruzadas dadas pela função de correlação adotada. Desta forma, o modelo bivariado pode ser reescrito como

$$\begin{cases} Y_1(\underline{x}_i) &= \mu_1(\underline{x}_i) + \sigma_{01}U_0(\underline{x}_i) + \sigma_1 U_1(\underline{x}_i) + Z_1(\underline{x}_i) \\ Y_2(\underline{x}_{i'}) &= \mu_2(\underline{x}_{i'}) + \sigma_{02}U_0(\underline{x}_{i'}) + \sigma_2 U_2(\underline{x}_{i'}) + Z_2(\underline{x}_{i'}) \end{cases} \quad (2.31)$$

É possível calcular a correlação entre Y_1 em uma localização \underline{x}_i e Y_1 em uma outra localização $\underline{x}_{i'}$, $i, i' : 1, \dots, n_1$ através do modelo espacial dado pela Equação (2.31), primeira linha, em que $U_0(\underline{x}_i)$ e $U_1(\underline{x}_i)$ contribuem para este cálculo. Da mesma forma é possível calcular a correlação considerando-se Y_2 , segunda linha do modelo (2.31). Também, a correlação de Y_1 em uma localização \underline{x}_i e Y_2 em uma localização $\underline{x}_{i'}$ decorre do modelo (2.31) e, neste caso, apenas $U_0(\underline{x}_i)$ e $U_0(\underline{x}_{i'})$ contribuem para o cálculo já que os outros termos são independentes.

Considerando-se apenas duas localizações (\underline{x}_1 e \underline{x}_2) a matriz Σ será dada por:

$$\Sigma = \begin{bmatrix} \sigma_{01}^2 \rho_0(\underline{x}_1, \underline{x}_1) + \sigma_1^2 \rho_1(\underline{x}_1, \underline{x}_1) & \sigma_{01}^2 \rho_0(\underline{x}_1, \underline{x}_2) + \sigma_1^2 \rho_1(\underline{x}_1, \underline{x}_2) & \sigma_{01} \sigma_{02} \rho_0(\underline{x}_1, \underline{x}_1) & \sigma_{01} \sigma_{02} \rho_0(\underline{x}_1, \underline{x}_2) \\ \sigma_{01}^2 \rho_0(\underline{x}_2, \underline{x}_1) + \sigma_1^2 \rho_1(\underline{x}_2, \underline{x}_1) & \sigma_{01}^2 \rho_0(\underline{x}_2, \underline{x}_2) + \sigma_1^2 \rho_1(\underline{x}_2, \underline{x}_2) & \sigma_{01} \sigma_{02} \rho_0(\underline{x}_2, \underline{x}_1) & \sigma_{01} \sigma_{02} \rho_0(\underline{x}_2, \underline{x}_2) \\ \sigma_{01} \sigma_{02} \rho_0(\underline{x}_1, \underline{x}_1) & \sigma_{01} \sigma_{02} \rho_0(\underline{x}_1, \underline{x}_2) & \sigma_{02}^2 \rho_0(\underline{x}_1, \underline{x}_1) + \sigma_2^2 \rho_2(\underline{x}_1, \underline{x}_1) & \sigma_{02}^2 \rho_0(\underline{x}_1, \underline{x}_2) + \sigma_2^2 \rho_2(\underline{x}_1, \underline{x}_2) \\ \sigma_{01} \sigma_{02} \rho_0(\underline{x}_2, \underline{x}_1) & \sigma_{01} \sigma_{02} \rho_0(\underline{x}_2, \underline{x}_2) & \sigma_{02}^2 \rho_0(\underline{x}_2, \underline{x}_1) + \sigma_2^2 \rho_2(\underline{x}_2, \underline{x}_1) & \sigma_{02}^2 \rho_0(\underline{x}_2, \underline{x}_2) + \sigma_2^2 \rho_2(\underline{x}_2, \underline{x}_2) \end{bmatrix}.$$

A estimação assim como a avaliação da incerteza sobre os parâmetros do modelo segue numa extensão da apresentada na Subseção 2.1.3.

2.4.2 Cokrigagem

A predição espacial no contexto de mais de uma variável segue numa extensão da teoria apresentada na Subsecção 2.2.2. Considera-se a predição espacial de $Y_0(\underline{x})$ em localizações não amostradas $\underline{x}_0 = (x_{10}, x_{20}, \dots, x_{n_20})$. O vetor de valores esperados correspondentes às variáveis Y_1 e Y_2 para todas as localizações de predição e a matriz de covariância são dadas pelo seguinte resultado da distribuição gaussiana multivariada:

Teorema 2.2 *Seja $\underline{Y} = (Y_0, \underline{Y})'$ um vetor bivariado com distribuição gaussiana multivariada conjunta com vetor de médias $\underline{\mu} = (\mu_{Y_0}, \underline{\mu}_Y)'$ e matriz de covariância*

$$\Sigma = \begin{bmatrix} \Sigma_{Y_0Y_0} & \Sigma_{Y_0Y} \\ \Sigma_{Y_0Y} & \Sigma_{YY} \end{bmatrix}$$

isto é, $\underline{Y} \sim N(\underline{\mu}; \Sigma)$. Então, a distribuição condicional de Y_0 dado \underline{Y} é também gaussiana multivariada,

$$Y_0|\underline{Y} \sim N(\underline{\mu}_{Y_0|\underline{Y}}; \Sigma_{Y_0|\underline{Y}})$$

onde

$$\underline{\mu}_{Y_0|\underline{Y}} = \underline{\mu}_{Y_0} + \Sigma_{Y_0Y} \Sigma_{YY}^{-1} (\underline{Y} - \underline{\mu}_Y) \quad e \quad \Sigma_{Y_0|\underline{Y}} = \Sigma_{Y_0Y_0} - \Sigma_{Y_0Y} \Sigma_{YY}^{-1} \Sigma_{YY_0}. \quad (2.32)$$

Sendo desconhecidos os valores de $\underline{\mu}_{Y_0}$, estes são substituídos pelo vetor de médias estimadas obtidas no processo de otimização. A matriz Σ , de onde se extrai as matrizes $\Sigma_{Y_0Y_0}$, Σ_{Y_0Y} , Σ_{YY_0} e Σ_{YY} , é calculada substituindo-se os valores estimados, restantes no processo de otimização, na matriz \mathbf{V} multiplicada por $\hat{\sigma}_1^2$.

2.5 DADOS COMPOSICIONAIS

Os dados são ditos composicionais na medida em que registram informação sobre frequências relativas associadas com diferentes componentes de um sistema, por exemplo, proporções associadas com diferentes nutrientes (BUTLER; GLASBEY, 2008).

A análise de dados composicionais foi sistematizada por Aitchison (1982). Tais dados consistem de vetores \underline{X} de proporções de algum “todo” e apresentam variabilidade de vetor para vetor. Cada vetor é denominado uma composição e os componentes de qualquer composição de B partes (X_1, X_2, \dots, X_B) devem satisfazer as exigências de que

cada componente é não negativo:

$$X_1 \geq 0, \dots, X_B \geq 0,$$

e que a soma de todos os componentes é 1:

$$X_1 + X_2 + \dots + X_B = 1. \quad (2.33)$$

Aitchison (1986) apresenta duas maneiras de se determinar uma composição. Na primeira, a composição fica completamente especificada pelos componentes de um subvetor de $(B-1)$ -partes onde $(X_1, X_2, \dots, X_{B-1})$,

$$X_B = 1 - X_1 - X_2 - \dots - X_{B-1},$$

o que significa que uma composição de B -partes é um vetor $(B-1)$ -dimensional podendo ser representado em algum conjunto $(B-1)$ -dimensional. Na segunda, especificando $B-1$ razões r_i dadas por:

$$r_i = X_i/X_B \quad i = 1, \dots, B-1.$$

A composição é determinada por

$$\begin{aligned} X_i &= r_i/(r_1 + \dots + r_{B-1} + 1) \quad i = 1, \dots, B-1. \\ X_B &= 1/(r_1 + \dots + r_{B-1} + 1). \end{aligned}$$

Aitchison e Greenacre (2002) apresentam três maneiras equivalentes de se considerar razões dentro de uma composição:

- a) as $\frac{1}{2}B(B-1)$ razões X_i/X_j entre os pares de componentes, assumindo $i < j$ ao selecionar o par;
- b) as $B-1$ razões X_i/X_B entre os primeiros $B-1$ componentes;
- c) as B razões $X_i/g(\underline{X})$ entre os componentes e sua média geométrica $g(\underline{X}) = \sqrt[B]{X_1 X_2 \dots X_B}$.

Neste trabalho serão consideradas razões como as do item (b).

As aplicações com dados composicionais são frequentemente encontrados em áreas como geologia, agricultura, biologia, literatura, meio ambiente, estatística médica, economia, dentre outras (LABUS, 2005). As características principais de um conjunto de dados composicionais segundo Reyment e Savazzi (1999) *apud* Labus (2005) são:

- podem ser representados na forma de uma matriz;
- cada linha da matriz corresponde a uma única unidade observacional ou experimental e soma 1 no caso de proporções, ou 100% no caso de percentagem, embora algumas vezes outra constante pode ser encontrada devido à alguma manipulação por parte do pesquisador;

- cada coluna da matriz representa uma única parte (variável);
- os coeficientes de correlação mudam se uma das variáveis (partes) é excluída da matriz de dados e as linhas somam 1 ou 100 novamente. O mesmo ocorre se um novo componente é adicionado.

Esta última propriedade significa que alterar uma ou mais variáveis do conjunto de dados pode ter um efeito numericamente significativo nas correlações entre as variáveis restantes.

Pawlowsky-Glahn e Olea (2004) e Tolosana-Delgado, Otero e Pawlowsky-Glahn (2005) apontam que os dados composicionais apresentam um efeito de correlação espúria o que significa que a aplicação dos métodos estatísticos padrão podem levar a resultados inconsistentes. Isto significa que as covariâncias estão sujeitas a controles não estocásticos, isto é, sofrem distorções devido à restrição da soma totalizar 1 levando à interpretação errônea da estrutura de covariância espacial. Um outro problema, levantado por Labus (2005), é que é uma consequência da Equação (2.33) o fato de que os componentes não são independentes e isto, segundo Pawlowsky-Glahn e Olea (2004), implica em singularidade da matriz de covariância de uma composição, excluindo por exemplo, o uso de técnicas de estimação como cokrigagem de todos os componentes.

Pawlowsky-Glahn e Olea (2004) relatam que os problemas com correlação espacial espúria e singularidade da matriz de covariância são relacionados à suposição básica de que o espaço amostral é irrestrito, uma suposição implícita na análise estatística de corregeionalizações e, a suposição de que a distribuição do erro de estimação em cada ponto da região amostral é gaussiana.

Desta forma surge a necessidade de expandir adequadamente os métodos de análise de dados composicionais para acomodar estruturas espaciais.

2.5.1 Composição Regionalizada

Uma composição regionalizada é um vetor função aleatória $\underline{X}(\underline{x})$ com $\underline{x} \in A \subset \mathbb{R}^n$ que satisfaz as seguintes condições em cada localização \underline{x} (PAWLOWSKY-GLAHN; OLEA, 2004):

- i) $X_1(\underline{x}) > 0, X_2(\underline{x}) > 0, \dots, X_B(\underline{x}) > 0$;
- ii) $X_1(\underline{x}) + X_2(\underline{x}) + \dots + X_B(\underline{x}) = 1$.

O caso $X_i(\underline{x}) = 0, i = 1, \dots, B$, foi excluído para evitar complicação desnecessária na apresentação e procedimentos para tratar esta situação pode ser encontrada em Pawlowsky-Glahn e Olea (2004, Seção 6.1).

O espaço amostral natural para uma composição é o simplex. McBratney, De Gruijter e Brus (1992) *apud* Odeh, Tood e Triantafilis (2003) definem um simplex como uma representação geométrica do espaço de atributos, onde uma composição de B partes

é representada por um número mínimo de vértices em um espaço de um dado número de dimensões. Em outras palavras, o espaço amostral para $\underline{X}(\underline{x})$ é o simplex- B unitário embutido no espaço real B -dimensional \mathbb{R}^B , dado por:

$$\mathbb{S}^B = \{\underline{X}(\underline{x}) \in \mathbb{R}^B; X_i(\underline{x}) > 0, i = 1, \dots, B; \underline{1}'\underline{X}(\underline{x}) = 1\},$$

sendo $\underline{1}'$ um vetor com B elementos iguais a 1.

2.5.2 Base Regionalizada

Uma base regionalizada, $\underline{W}(\underline{x})$, $\underline{x} \in A \subset \mathbb{R}^n$, é um vetor função aleatória cujos componentes são todos positivos e medidos na mesma escala (PAWLOWSKY-GLAHN; OLEA, 2004). Isto implica que o espaço amostral de uma base de B -partes é dado por:

$$\mathbb{R}_+^B = \{\underline{W}(\underline{x}) \in \mathbb{R}^B; W_i(\underline{x}) > 0, i = 1, \dots, B\}.$$

A transformação de uma base regionalizada em uma composição regionalizada se dá através do operador fechamento, \mathcal{C} como:

$$\begin{aligned} \mathcal{C} : \mathbb{R}_+^B &\longrightarrow \mathbb{S}^B \\ \underline{W}(\underline{x}) &\longrightarrow \mathcal{C}(\underline{W}(\underline{x})) = \frac{\underline{W}(\underline{x})}{\underline{1}'\underline{W}(\underline{x})}. \end{aligned}$$

O operador fechamento garante que o vetor resultante seja uma composição.

Barceló-Vidal, Martín-Fernández e Pawlowsky-Glahn (2001) apresentam uma interpretação geométrica para o operador fechamento \mathcal{C} . Considerando como exemplo uma composição com três componentes, pode-se entender \mathcal{C} como a intersecção do raio partindo da origem através de $\underline{W}(\underline{x})$ e o hiperplano de \mathbb{R}^3 definido pela equação $W_1(\underline{x}) + W_2(\underline{x}) + W_3(\underline{x}) = 1$, como na Figura 2.2(a). O conjunto de todos estes pontos é o simplex regular dado por:

$$\mathbb{S}^3 = \{(W_1(\underline{x}), W_2(\underline{x}), W_3(\underline{x}))'; W_1(\underline{x}) > 0, W_2(\underline{x}) > 0, W_3(\underline{x}) > 0; W_1(\underline{x}) + W_2(\underline{x}) + W_3(\underline{x}) = 1\}.$$

De acordo com Pawlowsky-Glahn e Olea (2004), de qualquer vetor aleatório $\underline{W}(\underline{x}) = (W_1(\underline{x}), W_2(\underline{x}), \dots, W_B(\underline{x}))'$ cujos elementos são positivos e que não seja uma composição pode-se obter uma, dividindo-se os elementos individuais pela soma de todos os elementos, ou seja:

$$X_i(\underline{x}) = \frac{W_i(\underline{x})}{W_1(\underline{x}) + W_2(\underline{x}) + \dots + W_B(\underline{x})}.$$

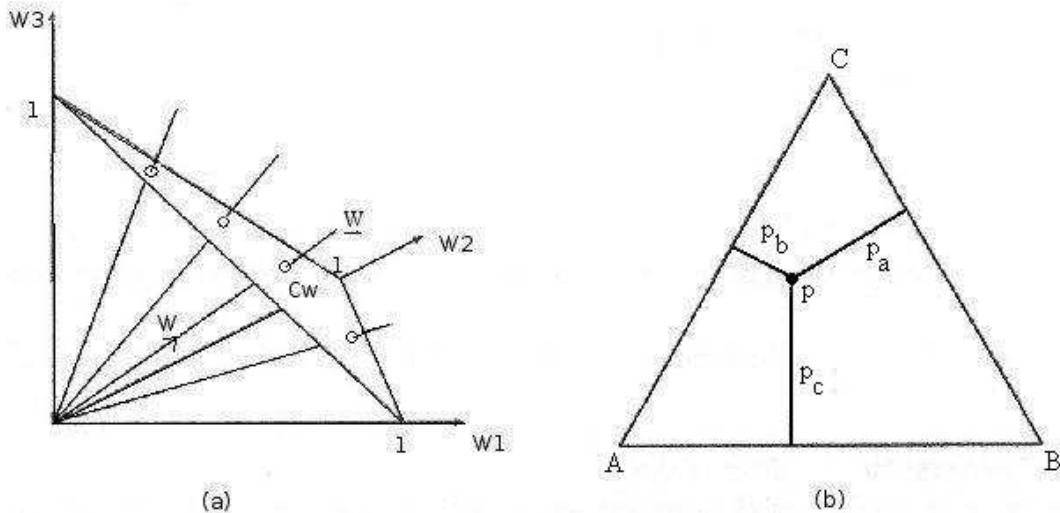


Figura 2.2: (a) Composições de 3 partes como raios partindo da origem em \mathbb{R}_+^3 ; (b) O $\underline{p} = (p_a, p_b, p_c)\%$ representado no simplex \mathbb{S}^3 .

FONTE: Adaptado de Barceló-Vidal, Martín-Fernández e Pawlowsky-Glahn (2001).

NOTA: Em particular, nesta figura, $\underline{W}(x) = \{kW(x) : k \in \mathbb{R}^+\}$.

A composição resultante é $\underline{X}(x) = (X_1(x), X_2(x), \dots, X_B(x))'$ e a relação $\frac{W_i(x)}{W_j(x)} = \frac{X_i(x)}{X_j(x)}$ é válida para quaisquer índices $i, j = 1, 2, \dots, B$, para $W_j(x)$ e $X_j(x) \neq 0$, enfatizando-se que o dado composicional contém informação somente sobre magnitudes relativas e não absolutas.

2.5.3 Representação Gráfica

O diagrama ternário, triângulo cujos vértices representam os três componentes da composição (BUTLER; GLASBEY, 2008) é uma técnica gráfica utilizada na compreensão de dados composicionais e adotada neste trabalho. Outras técnicas gráficas como Diagramas de Variação, Diagramas Ternários, Diagramas de Dispersão de Razão e Razão Log, Gráfico Bivariado Composicional e Diagrama de Dispersão das Coordenadas podem ser encontradas, por exemplo, em Aitchison e Egozcue (2005).

De acordo com Pawlowsky-Glahn, Egozcue e Tolosana-Delgado (2007), para composições de três componentes, o diagrama ternário é uma forma de visualização do simplex que é um triângulo equilátero com lados de comprimento unitário e com vértices A , B e C como na Figura 2.2(b), ou seja, é uma representação equivalente. Neste triângulo, qualquer ponto $\underline{p} = (p_a, p_b, p_c)$ é marcado a uma distância p_a do lado oposto ao vértice A , a uma distância p_b do lado oposto ao vértice B e a uma distância p_c do lado oposto ao vértice C . A tripla (p_a, p_b, p_c) é denominada coordenadas baricêntricas de \underline{p} mas o que interessa são as expressões das coordenadas dos vértices e das composições num sistema de coordenadas cartesianas bidimensional, $(\underline{u}; \underline{v})$.

Para determinar as coordenadas $(\underline{u}; \underline{v})$ correspondentes a composição $\underline{X} = (35; 13; 52)\%$, por exemplo, considere a Figura 2.3. No simplex, os vértices do triângulo tem coordenadas $A = (100, 0, 0)\%$, $B = (0, 100, 0)\%$ e $C = (0, 0, 100)\%$. Sejam $(\underline{u}_0; \underline{v}_0)$ as coordenadas do vértice A , com escala escolhida de forma a centrar o triângulo no sistema de coordenadas cartesianas. Então, $B = (\underline{u}_0 + 1; \underline{v}_0)$. Como o triângulo tem

lados de comprimento 1, pelo teorema de Pitágoras a altura deste é igual a $\sqrt{3}/2$, logo, $C = (u_0 + 0,5; v_0 + \sqrt{3}/2)$. As coordenadas $(u; v)$ são obtidas como uma combinação linear convexa das coordenadas dos vértices como:

$$(u, v) = (X_1A + X_2B + X_3C).$$

Desta forma, considerando $(u_0; v_0) = (0,2; 0,2)$ e para a composição $\underline{X} = (35; 13; 52)$, as correspondentes coordenadas no sistema de coordenadas bidimensional é:

$$(u, v) = \frac{1}{100}(X_1(u_0; v_0) + X_2(u_0 + 1; v_0) + X_3(u_0 + 0,5; v_0 + \sqrt{3}/2)) = (0,59; 0,65).$$

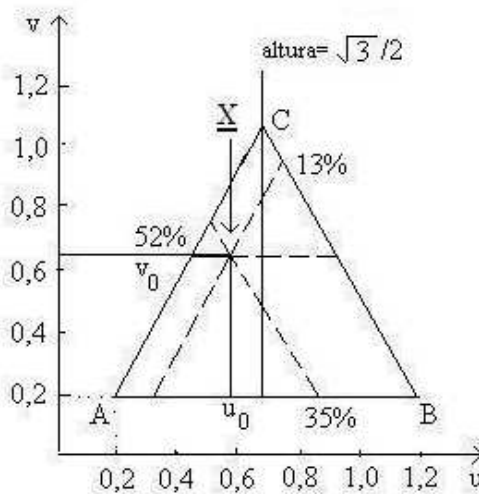


Figura 2.3: O dado composicional representado no simplex S^3 e no \mathbb{R}^2 .
 FONTE: Adaptado de Pawlowsky-Glahn, Egozcue e Tolosana-Delgado (2007).

A Figura 2.4 mostra um diagrama ternário, baseado em dados de areia, silte e argila, apresentados em Aitchison (1986). Na interpretação da figura, pontos localizados próximo a um vértice indicam altas proporções do componente associado àquele vértice, enquanto pontos localizados no centro do triângulo tem proporções iguais para todos os três componentes. Por exemplo, nesta amostra observa-se que os menores valores ocorreram para o componentes argila. O ponto vermelho representa o centro da distribuição conforme Subseção 2.5.8.

Segundo Boogaart (2005) e como se pode observar pela Figura 2.5, todos os pontos localizados sobre uma linha paralela ao eixo oposto a um vértice, A por exemplo, têm a mesma percentagem do componente associado a este vértice. Esta percentagem é igual a distância da linha ao eixo BC sobre a distância do vértice A ao eixo BC. Por outro lado, todos os pontos em uma linha reta partindo de um dos vértices, A por exemplo, têm percentagens iguais dos outros dois componentes e são representadas pelo ponto onde a linha cruza o eixo oposto, BC.

Além disso, de acordo com Aitchison (1986), para avaliar a extensão da variabilidade da razão entre dois componentes, deve-se traçar linhas retas partindo do vértice complementar e passando por todos os pontos. A amplitude da intersecção destas linhas com o eixo oposto ao vértice complementar resulta na extensão da variabilidade.

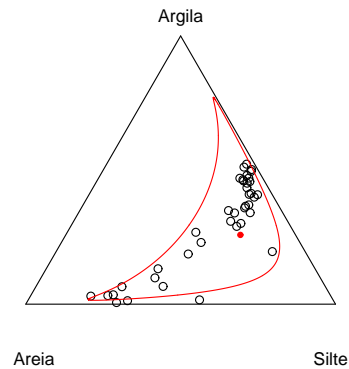


Figura 2.4: Diagrama ternário para dados do Lago Ártico incluindo o centro da distribuição e região 2–desvios-padrão de confiança.
 FONTE: Aitchison (1986).

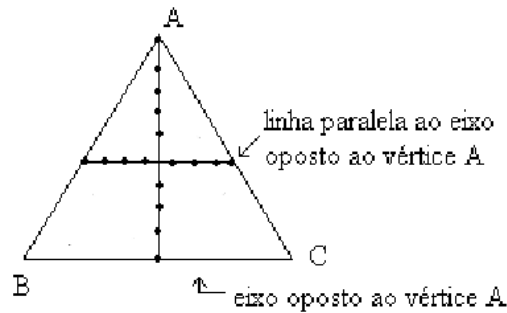


Figura 2.5: Desenho que representa pontos em um diagrama ternário com mesmas percentagens (sobre a linha horizontal) do componente A e percentagens iguais (sobre a linha vertical) dos componentes B e C.
 FONTE: A autora (2010).

Por exemplo, na Figura 2.4, verifica-se que a variabilidade da razão Areia/Silte é maior que a da razão Argila/Silte, pois a amplitude observada no lado Areia-Silte é maior que a observada no lado Argila-Silte. Interpretações semelhantes são feitas para os outros lados.

2.5.4 Subcomposição Regionalizada

Uma subcomposição regionalizada é um subconjunto de uma composição regionalizada. Assim, considerando-se $\underline{X}(\underline{x})$ uma composição regionalizada de B partes e $s \subset \{1, 2, \dots, B\}$ de forma a tornar $\underline{X}_s(\underline{x})$ um subvetor cujos elementos são os componentes de $\underline{X}(\underline{x})$ correspondentes às partes em s , então uma subcomposição regionalizada será:

$$\mathcal{C}(\underline{X}_s(\underline{x})) = \frac{\underline{X}_s(\underline{x})}{\mathbf{1}'\underline{X}_s(\underline{x})}.$$

Segundo Aitchison (1986), Aitchison e Greenacre (2002), Aitchison e Egozcue (2005) o conceito de subcomposição é importante no que se refere a coerência subcomposicional. Isto significa que, se um pesquisador, tendo acesso à uma composição, também fizer inferência a partir de uma subcomposição desta e, se um outro, com acesso somente a uma subcomposição mas com partes comuns à do primeiro pesquisador fizer inferência a partir desta, os dois resultados deverão coincidir.

As correlações momento-produto e análise de componentes principais baseadas em covariâncias calculadas no dado composicional em linha, não tem coerência subcomposicional, mas a característica importante de uma subcomposição é que esta preserva relações de razões. Daí, se $\underline{s}(x) = \mathcal{C}(\underline{X}_s(x))$, então $\frac{s_i(x)}{s_j(x)} = \frac{X_i(x)}{X_j(x)}$, $i, j \in s$. Observa-se que as razões são formadas com os componentes do vetor de dado composicional (linha) através das colunas da matriz de dados.

Da mesma forma que na Subseção 2.5.2, Barceló-Vidal, Martín-Fernández e Pawlowsky-Glahn (2001) apresentam uma interpretação geométrica para uma subcomposição. Novamente, considerando como exemplo uma composição de 3-partes, a formação de uma subcomposição corresponde à projeção ortogonal do raio associado a $\underline{X}(x)$ em \mathbb{R}^+ em um subespaço que é gerado pelos eixos coordenados das partes selecionadas para formar a subcomposição (Figura 2.6).

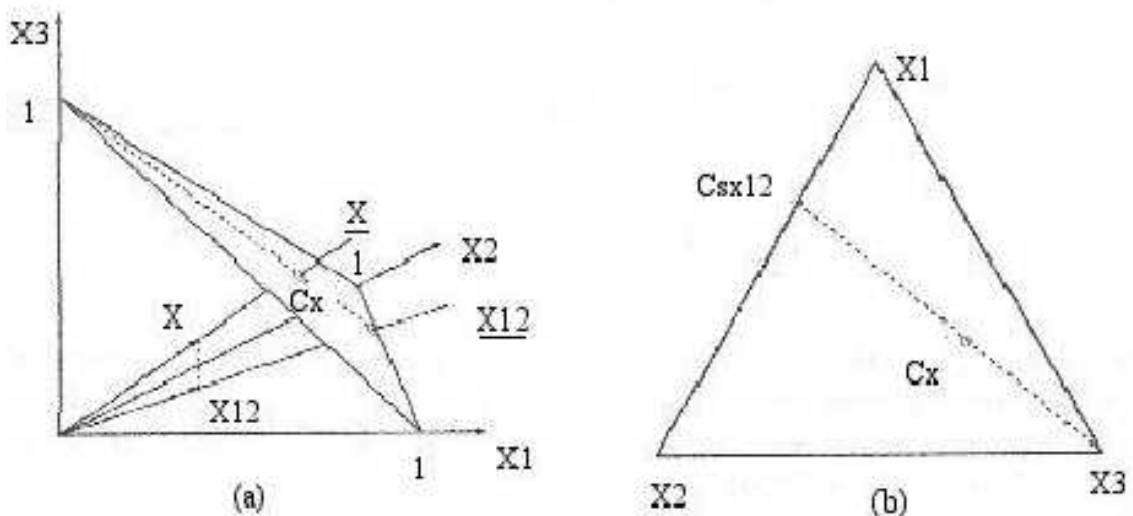


Figura 2.6: Interpretação geométrica da formação da subcomposição $\underline{X}_{12}(x)$ da composição $\underline{X}(x)$: (a) em \mathbb{R}^3_+ ; (b) em \mathbb{S}^3 .

FONTE: Adaptado de Barceló-Vidal, Martín-Fernández e Pawlowsky-Glahn (2001).

NOTA: Em particular, nesta figura, $\underline{X}(x) = \{kX(x) : k \in \mathbb{R}^+\}$.

2.5.5 Amalgamação e Partição Regionalizada

Seja $\underline{X}(x) = (X_1(x), X_2(x), \dots, X_B(x))'$ uma composição regionalizada de B partes que é dividida em C ($C \leq B$) subconjuntos mutuamente exclusivos. Seja $\underline{X}_i(x)$, $i = 1, \dots, C$,

o vetor cujos componentes são os elementos do i -ésimo subconjunto. Pawlowsky-Glahn e Olea (2004) definem amalgamação regionalizada como a composição regionalizada com C componentes em que os componentes de cada um dos C subconjuntos são somados:

$$\underline{A}(\underline{x}) = (A_1(\underline{x}), A_2(\underline{x}), \dots, A_C(\underline{x}))',$$

com $A_i(\underline{x}) = \underline{1}'\underline{X}_i(\underline{x})$.

Ao considerar a amalgamação regionalizada $\underline{A}(\underline{x})$ juntamente com as subcomposições regionalizadas, $s_i(\underline{x}) = \mathcal{C}(\underline{X}_i(\underline{x}))$, $i = 1, \dots, C$:

$$P_C(\underline{X}(\underline{x})) = (\underline{s}_1(\underline{x}), \underline{s}_2(\underline{x}), \dots, \underline{s}_C(\underline{x}); \underline{A}(\underline{x})),$$

tem-se definida a partição regionalizada de ordem C de $\underline{X}(\underline{x})$. Assim, as informações contidas nos subvetores são preservadas em uma partição.

Segundo Pawlowsky-Glahn e Olea (2004) uma amalgamação regionalizada é entendida como uma mistura e combinação de vários componentes. Aitchison (1986) exemplifica a idéia no contexto de uma enquete sobre gastos de uma dona de casa onde a composição $\underline{X}(\underline{x})$ de proporções das despesas é formada pelos nove componentes a seguir:

$X_1(\underline{x})$: Gêneros alimentícios	$X_6(\underline{x})$: Bens duráveis
$X_2(\underline{x})$: Habitação	$X_7(\underline{x})$: Bens diversos
$X_3(\underline{x})$: Combustível e luz	$X_8(\underline{x})$: Transporte e veículos
$X_4(\underline{x})$: Cigarro e álcool	$X_9(\underline{x})$: Serviços,
$X_5(\underline{x})$: Vestuário e calçado	

enquanto o interesse pode estar em grupos de mercadorias, ou seja, na amalgamação $\underline{A}(\underline{x}) = (A_1(\underline{x}), A_2(\underline{x}), A_3(\underline{x}), A_4(\underline{x}))$, por exemplo, onde

- $A_1(\underline{x}) = (X_2(\underline{x}) + X_3(\underline{x}))$
- $A_2(\underline{x}) = (X_1(\underline{x}) + X_4(\underline{x}))$
- $A_3(\underline{x}) = (X_5(\underline{x}) + X_6(\underline{x}) + X_7(\underline{x}))$
- $A_4(\underline{x}) = (X_8(\underline{x}) + X_9(\underline{x}))$

2.5.6 Transformação Logística Modificada

Graf (2006) expõe de forma clara que a restrição de que a soma dos valores dos componentes é a unidade implica que existe, necessariamente, uma correlação negativa entre os componentes e isto faz com que as correlações não sejam diretamente interpretáveis. Para relaxar ou evitar esta restrição Aitchison (1986) propôs transformações que generalizam a transformação logística para um vetor composicional de duas partes. O autor afirma que a transformação log-razão é viável para investigar a importância ou irrelevância de componentes individuais, por fornecer uma estrutura de dependência correta e

interpretável para descrever os padrões reais de variabilidade composicional, que permita a investigação coerente da variabilidade subcomposicional. A natureza essencial de uma composição é que as magnitudes relativas dos componentes são as unidades relevantes sob estudo. Estas magnitudes relativas (ou razões) implicam em tratabilidade e interpretação estatística (AITCHISON, 1999). Algumas transformações logísticas, como a aditiva, a aditiva modificada, a multiplicativa e a híbrida podem ser encontradas em Aitchison (1982), Aitchison et al. (2000) e Odeh, Tood e Triantafilis (2003), por exemplo.

No sentido de composições regionalizadas, Pawlowsky-Glahn, Olea e Davis (1995), Pawlowsky-Glahn e Olea (2004) definem a transformação razão log-aditiva (ALR) como a função

$$\begin{aligned} \text{alr} : \mathbb{S}^B &\longrightarrow \mathbb{R}^{B-1} \\ \underline{X}(x) &\longrightarrow \text{alr}(\underline{X}(x)) = \left(\ln \left(\frac{X_1(x)}{X_B(x)} \right), \dots, \ln \left(\frac{X_{B-1}(x)}{X_B(x)} \right) \right)' \end{aligned}$$

Pode-se observar que para $B = 2$, a transformação ALR torna-se a transformação logística convencional. Nota-se também que se a composição possuir três componentes, após a transformação a composição passará a ter dois componentes: Um vetor bivariado. Então, um vetor auxiliar, $\underline{X}^*(x) = (X_1(x) \ X_1(x) \dots X_{B-1}(x) \ 0)$ é introduzido para definir a transformação inversa, denominada transformação logística generalizada aditiva (AGL), que traz os dados de volta ao simplex:

$$\begin{aligned} \text{agl} : \mathbb{R}^{B-1} &\longrightarrow \mathbb{S}^B \\ \underline{Y}(x) &\longrightarrow \underline{X}(x) = \mathcal{C} \left(\left(\exp \left\{ \ln \left(\frac{X_1(x)}{X_B(x)} \right) \right\}, \dots, \exp\{0\} \right) \right)' \end{aligned}$$

Se a restrição de soma constante é $c \neq 0$ os dados no espaço simplex são obtidos fazendo-se:

$$\underline{X}(x) = c \cdot \text{agl}(\underline{Y}(x)).$$

Outra transformação viável é a transformação razão log centrada (CLR) definida como

$$\begin{aligned} \text{clr} : \mathbb{R}_+^B &\longrightarrow \mathbb{R}^B \\ \underline{W}(x) &\longrightarrow \text{clr}(\underline{W}(x)) = \ln \left(\frac{\underline{W}(x)}{g(\underline{W}(x))} \right) \end{aligned}$$

ou

$$\begin{aligned} \text{clr} : \mathbb{S}^B &\longrightarrow \mathbb{R}^B \\ \underline{X}(x) &\longrightarrow \text{clr}(\underline{X}(x)) = \ln \left(\frac{\underline{X}(x)}{g(\underline{X}(x))} \right) \end{aligned}$$

onde $g(\underline{W}(x)) = \sqrt[B]{\prod_{i=1}^B W_i(x)}$ é a média geométrica dos componentes da base regionalizada $\underline{W}(x)$ e $g(\underline{X}(x)) = \sqrt[B]{\prod_{i=1}^B X_i(x)}$ é a média geométrica dos componentes da composição regionalizada $\underline{X}(x)$.

Existe teste para validar a suposição de normalidade logística das composições log-razões e este é apresentado em Aitchison (1986), mas seu uso é limitado e insuficientes porque supõe a independência das observações o que não ocorre neste trabalho. O teste baseia-se em Andrews, Gnanadesikan e Warner (1973) e nas recomendações de Stephens (1974) para o uso de testes da função de distribuição empírica, nas formas de Anderson-Darling, Cramer-von Mises e Watson. Considera-se todas as $B - 1$ distribuições marginais univariadas das composições log-razão.

Para a realização do teste considera-se que, para a j -ésima distribuição marginal da composição log-razão aditiva, as observações são

$$y_{ij} = \ln \frac{x_{ij}}{x_{iB}}, \quad i = 1, \dots, n; \quad j = 1, 2.$$

Com $\hat{\mu}_j$ e $\hat{\sigma}_j^2$, as estimativas da média e variância marginais calcula-se $(y_{ij} - \hat{\mu}_j) / \sqrt{\hat{\sigma}_j^2}$ obtendo-se os valores z_r , $r = 1, \dots, N$ onde N é o número de observações e, em seguida, calcula-se $\Phi(z_r)$, onde $\Phi(\cdot)$ é a função de distribuição de uma distribuição $N \sim (0; 1)$. Os valores de z_r são arranjados em ordem crescente de magnitude e calculam-se as seguintes estatísticas teste para a função de distribuição empírica:

Anderson-Darling:	$Q_A = -\frac{1}{N} \sum_{r=1}^N (2r-1) \{\ln(z_r) + \ln(1+z_{N+1-r})\} - N$
Cramer-von Mises:	$Q_C = \sum_{r=1}^N \left[\frac{z_r - (2r-1)}{2N} \right]^2 + \frac{1}{12N}$
Watson:	$Q_W = Q_C - N \left(\bar{z} - \frac{1}{2} \right)^2$ onde $\bar{z} = \frac{1}{N} \sum_{r=1}^N z_r$

A idéia do teste é que, se as observações são normalmente distribuídas, então z_r deve ser aproximadamente a estatística de ordem de uma distribuição uniforme sobre o intervalo $(0, 1)$.

Para avaliar desvios da normalidade logística aditiva, compara-se as estatísticas da função de distribuição empírica modificada

Anderson-Darling:	$Q_A \left[1 + \frac{4}{N} + \frac{25}{N^2} \right]$
Cramer-von Mises:	$Q_C \left[1 + \frac{1}{2N} \right]$
Watson:	$Q_W \left[1 + \frac{1}{2N} \right]$

com valores críticos 0,787 (Anderson-Darling), 0,126 (Cramer-von Mises) e 0,116 (Watson) com nível de significância de 5% (STEPHENS, 1974).

A regra de decisão é a de que valores da estatística teste inferiores ao valor crítico evidenciam desvios não significativos da normalidade logística aditiva.

2.5.7 Perturbação e Potência

Os conceitos de perturbação e potência são válidos tanto no contexto espacial como não espacial e são definidos omitindo-se o índice espacial “(x)”. Perturbação no simplex para Eynatten, Barceló-Vidal e Pawlowsky-Glahn (2003) é uma operação que pode ser usada para descrever numericamente mudanças em uma composição e a combinação de perturbação e transformação potência fornecem um método para a análise de processos lineares composicionais no simplex. Para Aitchison (1986), Tolosana-Delgado, Otero e Pawlowsky-Glahn (2005), Aitchison e Egozcue (2005), por exemplo, estas operações definem uma estrutura de espaço vetorial, de dimensão $B - 1$ no simplex, com a perturbação como uma operação comutativa e a potência como um produto externo. A perturbação, representada pelo símbolo \oplus , corresponde a multiplicar composições componente a componente e dividir cada um pela soma de todos para se obter soma igual a 1, ou seja,

$$\underline{X}_1 \oplus \underline{X}_2 = (X_{11}, X_{12}, \dots, X_{1B}) \oplus (X_{21}, X_{22}, \dots, X_{2B}) = \mathcal{C}(X_{11}X_{21}, X_{12}X_{22}, \dots, X_{1B}X_{2B}).$$

É a operação análoga a translação no espaço real. A potência, representada por \odot , é a análoga a multiplicação por um escalar no espaço real:

$$\alpha \odot (X_{11}, X_{12}, \dots, X_{1B}) = \mathcal{C}(X_{11}^\alpha, X_{12}^\alpha, \dots, X_{1B}^\alpha).$$

Consequentemente, tem-se o vetor de diferenças composicionais,

$$\underline{X}_1 \ominus \underline{X}_2 = \underline{X}_1 \oplus (-1 \odot \underline{X}_2).$$

Acrescentando às operações um produto interno,

$$\langle \underline{X}_1, \underline{X}_2 \rangle = \sum_{i=1}^B \ln \left(\frac{X_{1i}}{g(\underline{X}_1)} \right) \ln \left(\frac{X_{2i}}{g(\underline{X}_2)} \right)$$

tem-se uma estrutura de espaço Euclidiano real para o simplex onde $g(\underline{X}_1) = \sqrt[B]{\prod_{j=1}^B X_{1j}}$ é a média geométrica. Este produto interno induz uma distância (uma medida que pode ser entendida, por exemplo, como grau de alteração) no simplex, denominada distância de Aitchison, usada para calcular a distância ou diferença entre duas composições e útil para entender a variabilidade dentro de um conjunto de dados:

$$d(\underline{X}_1, \underline{X}_2) = \sqrt{\sum_{i=1}^B \left(\ln \left(\frac{X_{1i}}{g(\underline{X}_1)} \right) - \ln \left(\frac{X_{2i}}{g(\underline{X}_2)} \right) \right)^2}$$

Com isto, tem-se a geometria de Aitchison do simplex.

2.5.8 Estatísticas Descritivas e Domínio de Confiança Para Dados Composicionais

A média aritmética não é uma medida representativa em se tratando de dados composicionais. Pawlowsky-Glahn e Olea (2004) sugeriram como medida de tendência central ou centro da distribuição o fechamento da média geométrica $\text{cen}(\underline{X})$ que, para uma amostra de n composições com B componentes, é dada por:

$$\text{cen}(\underline{X}) = \frac{1}{g_s} [g(X_1) \quad g(X_2) \quad \dots \quad g(X_B)]', \quad (2.34)$$

em que $g(X_i)$ é a média geométrica do i -ésimo componente como definido na Subseção 2.5.6 e

$$g_s = g(X_1) + g(X_2) + \dots + g(X_B).$$

Pelo mesmo motivo da Subseção 2.5.7, o índice espacial é omitido.

Os autores afirmam que nem o desvio padrão nem a variância são medidas de dispersão adequadas quando o problema é considerado do ponto de vista multivariado. Mas, consideram que uma medida de dispersão é importante para comparações entre conjuntos de dados e uma medida da variabilidade total como uma medida de dispersão ao redor de uma medida do centro da distribuição faz sentido. Esta medida considera as $\frac{1}{2}(B-1)B$ variâncias das log-razões e é definida por:

$$\text{Totvar}(\underline{X}) = \frac{1}{B} \sum_{i < j} \text{Var} \left(\ln \left(\frac{X_i}{X_j} \right) \right). \quad (2.35)$$

Considerando que $\ln(X_i/X_i) = 0$, $i = 1, \dots, B$ e que $\text{Var}(\ln(X_i/X_j)) = \text{Var}(\ln(X_j/X_i))$, $i = 1, \dots, B-1; j = i+1, \dots, B$, pode-se dispor as variâncias das log-razões na forma de uma matriz de ordem $B \times B$, denominada matriz variação, permitindo uma forma simples de visualizar a variabilidade composicional.

Sob a suposição de que $\left(\ln \left(\frac{X_1}{X_B} \right), \ln \left(\frac{X_2}{X_B} \right), \dots, \ln \left(\frac{X_{B-1}}{X_B} \right) \right)$ tem distribuição assintoticamente gaussiana (GRAF, 2006; AITCHISON, 1986) o domínio de confiança para $\underline{X} = (X_1, X_2, \dots, X_B)$ é um subconjunto do simplex \mathbb{S}^B dado por:

$$B_{1-\alpha}(\underline{X}) = \left\{ \underline{X} \in \mathbb{S}^B \mid \left(\ln \left(\frac{\underline{X}_{-B}}{\underline{X}_B} \right) - \underline{\mu} \right)' \Sigma^{-1} \left(\ln \left(\frac{\underline{X}_{-B}}{\underline{X}_B} \right) - \underline{\mu} \right) \leq \chi_{B-1; 1-\alpha}^2 \right\},$$

em que $\chi_{B-1; 1-\alpha}^2$ é o quantil $(1-\alpha)$ da distribuição qui-quadrado com $(B-1)$ graus de liberdade; $\underline{\mu}$ é o vetor $(B-1) \times 1$ de médias aritméticas das log-razões e Σ é a matriz de covariâncias $(B-1) \times (B-1)$ das log-razões.

2.5.9 Estacionariedade

Seja $\underline{X}(\underline{x})$ um vetor função aleatória, $\underline{x} \in A \subset \mathbb{R}^n$, e f uma função (ex: logaritmo) de modo que $f(\underline{X}(\underline{x}))$ também é um vetor função aleatória. De acordo com Pawlowsky-Glahn e Olea (2004), $\underline{X}(\underline{x})$ é uma função estacionária de 2^a ordem se $f(\underline{X}(\underline{x}))$ é estacionária de 2^a ordem, isto é, se $f(\underline{X}(\underline{x}))$ satisfaz as condições:

- a) o vetor de valores esperados $E\left(f(\underline{X}(\underline{x}))\right) = \boldsymbol{\mu}$ existe e não depende de \underline{x} ;
- b) a matriz função de covariância

$$\text{Cov}\left(f(\underline{X}(\underline{x}_1)), f(\underline{X}(\underline{x}_2))\right) = \boldsymbol{\Sigma}(\underline{x}_2 - \underline{x}_1)$$

existe e não depende de \underline{x}_1 , \underline{x}_2 , mas somente da diferença $\underline{h} = \underline{x}_2 - \underline{x}_1$.

Além disso, $\underline{X}(\underline{x})$ é uma função intríntrica se $f(\underline{X}(\underline{x}))$ satisfaz:

- a) o vetor de valores esperados $E\left(f(\underline{X}(\underline{x}))\right) = \boldsymbol{\mu}$ existe e não depende de \underline{x} ;
- b) a matriz função de covariância

$$\text{Cov}\left(f(\underline{X}(\underline{x}_2)) - f(\underline{X}(\underline{x}_1))\right) = \boldsymbol{\Gamma}(\underline{x}_2 - \underline{x}_1)$$

existe e não depende de \underline{x}_1 , \underline{x}_2 , mas somente da diferença $\underline{h} = \underline{x}_2 - \underline{x}_1$.

Nota-se que $\boldsymbol{\Sigma}(\underline{x}_2 - \underline{x}_1)$ e $\boldsymbol{\Gamma}(\underline{x}_2 - \underline{x}_1)$ são as matrizes função de covariância cruzada e de variogramas e variogramas cruzados.

Acrescenta-se, ainda, que $\underline{X}(\underline{x})$ é log razão estacionária de 2^a ordem (LR estacionária) se o conjunto das log razões entre todos os pares é estacionário de 2^a ordem e, respectivamente, é log razão intríntrica (LR intríntrica) se o conjunto das log razões entre todos os pares é intríntrico.

2.5.10 Estrutura de Covariância Espacial

Como citado anteriormente, o interesse deste trabalho está na metodologia de geoestatística aplicada à dados composicionais. Em se tratando de geoestatística, tem-se a existência de dependência espacial entre os locais de observação. Por outro lado, observa-se que sempre existirá correlação entre os componentes de uma composição de forma que a estrutura de covariância é essencial na modelagem.

Pawlowsky-Glahn e Olea (2004) definem a estrutura de covariância espacial de

uma composição regionalizada $\underline{X}(\underline{x})$ como o conjunto de funções

$$\sigma_{ij\cdot kl}(\underline{h}) = Cov \left(\ln \left(\frac{X_i(\underline{x})}{X_k(\underline{x})} \right), \ln \left(\frac{X_j(\underline{x} + \underline{h})}{X_l(\underline{x} + \underline{h})} \right) \right), \quad i, j, k, l \in \{1, 2, \dots, B\}$$

para $\underline{x}, \underline{x} + \underline{h} \in A$.

Neste caso, as seguintes propriedades são válidas considerando-se $i, j, k, l, m, n \in \{1, 2, \dots, B\}$:

- a) em geral, $\sigma_{ij\cdot kl}(\underline{h}) \neq \sigma_{ji\cdot lk}(\underline{h})$;
- b) em geral, $\sigma_{ij\cdot kl}(\underline{h}) \neq \sigma_{ij\cdot kl}(-\underline{h})$;
- c) $\sigma_{ij\cdot kl}(\underline{h}) = \sigma_{ji\cdot lk}(-\underline{h})$;
- d) $\sigma_{ij\cdot kl}(\underline{h}) = -\sigma_{kj\cdot il}(\underline{h}) = \sigma_{kl\cdot ij}(\underline{h}) = -\sigma_{il\cdot kj}(\underline{h})$;
- e) $\sigma_{ij\cdot il}(\underline{h}) = \sigma_{ij\cdot kj}(\underline{h}) = \sigma_{ij\cdot ij}(\underline{h}) = \sigma_{ii\cdot ii}(\underline{h}) = 0$;
- f) $\sigma_{ij\cdot kl}(\underline{h}) = \sigma_{ij\cdot mn}(\underline{h}) + \sigma_{in\cdot ml}(\underline{h}) + \sigma_{mj\cdot kn}(\underline{h}) + \sigma_{mn\cdot kl}(\underline{h})$.

Do exposto, segue que para quaisquer dois componentes $X_i(\underline{x}), X_j(\underline{x})$, $i, j \in \{1, 2, \dots, B\}$ de uma composição $\underline{X}(\underline{x})$:

- a) a autocovariância das log razões, denominada LR autocovariância é

$$\tau_{i\cdot j}(\underline{h}) = \sigma_{ii\cdot jj}(\underline{h})$$

que são os elementos da matriz de LR autocovariâncias (Matriz Variação), $\mathbf{T}(\underline{h})$, de dimensão $B \times B$.

- b) a ALR covariância cruzada é a função

$$\sigma_{ij}(\underline{h}) = \sigma_{ij\cdot BB}(\underline{h})$$

que são os elementos da matriz de ALR covariâncias cruzadas, $\mathbf{\Sigma}(\underline{h})$, de dimensão $(B-1) \times (B-1)$.

- c) a CLR covariância cruzada é a função

$$\xi_{ij}(\underline{h}) = Cov \left(\ln \left(\frac{X_i(\underline{x})}{g(\underline{X}(\underline{x}))} \right), \ln \left(\frac{X_j(\underline{x} + \underline{h})}{g(\underline{X}(\underline{x} + \underline{h}))} \right) \right),$$

que são os elementos da matriz de CLR covariâncias cruzadas, $\mathbf{\Xi}(\underline{h})$, de dimensão $B \times B$.

2.5.11 Estrutura de Covariância Espacial Intrínica

Em continuação ao item anterior, a estrutura de covariância espacial intrínica de uma composição regionalizada $\underline{X}(\underline{x})$ é definida como o conjunto de funções

$$V_{ij \cdot kl}(\underline{h}) = \frac{1}{2} \text{Cov} \left(\ln \left(\frac{X_i(\underline{x})}{X_k(\underline{x})} \right) - \ln \left(\frac{X_i(\underline{x} + \underline{h})}{X_k(\underline{x} + \underline{h})} \right), \ln \left(\frac{X_j(\underline{x})}{X_l(\underline{x})} \right) - \ln \left(\frac{X_j(\underline{x} + \underline{h})}{X_l(\underline{x} + \underline{h})} \right) \right),$$

para $i, j, k, l \in \{1, 2, \dots, B\}$.

Como propriedades, apresenta-se:

- a) $V_{ij \cdot kl}(0) = 0$;
- b) $V_{ij \cdot kl}(\underline{h}) = V_{ji \cdot lk}(\underline{h}) = V_{ij \cdot kl}(-\underline{h})$;
- c) $V_{ij \cdot il}(\underline{h}) = V_{ij \cdot kj}(\underline{h}) = V_{ij \cdot ij}(\underline{h}) = V_{ii \cdot ii}(\underline{h}) = 0$;
- d) $V_{ij \cdot kl}(\underline{h}) = -V_{kj \cdot il}(\underline{h}) = V_{kl \cdot ij}(\underline{h}) = -V_{il \cdot kj}(\underline{h})$;
- e) $|V_{ij \cdot kl}(\underline{h})| \leq \sqrt{V_{ii \cdot kk}(\underline{h})} \sqrt{V_{jj \cdot ll}(\underline{h})}$ (Desigualdade de Cauchy-Schwarz).

Agora, em se tratando de variograma, e considerando-se quaisquer dois componentes $X_i(\underline{x})$, $X_j(\underline{x})$, $i, j \in \{1, 2, \dots, B\}$ de uma composição $\underline{X}(\underline{x})$ tem-se:

- a) o variograma da log razão, LR variograma é dado por

$$V_{i \cdot j}(\underline{h}) = V_{ii \cdot jj}(\underline{h})$$

que são os elementos da matriz de LR variogramas (Matriz Variação Intrínica), $\Gamma(\underline{h})$, de dimensão $B \times B$.

- b) o ALR variograma cruzado é a função

$$\psi_{ij}(\underline{h}) = V_{ij \cdot BB}(\underline{h})$$

que são os elementos da matriz de ALR covariâncias cruzadas intrínicas, $\Psi(\underline{h})$, de dimensão $(B-1) \times (B-1)$.

- c) a CLR covariância cruzada é a função

$$\delta_{ij}(\underline{h}) = \text{Cov} \left(\ln \left(\frac{X_i(\underline{x})}{g(\underline{X}(\underline{x}))} \right) - \ln \left(\frac{X_i(\underline{x} + \underline{h})}{g(\underline{X}(\underline{x} + \underline{h}))} \right), \ln \left(\frac{X_j(\underline{x})}{g(\underline{X}(\underline{x}))} \right) - \ln \left(\frac{X_j(\underline{x} + \underline{h})}{g(\underline{X}(\underline{x} + \underline{h}))} \right) \right),$$

que são os elementos da matriz de CLR covariâncias cruzadas, $\Delta(\underline{h})$, de dimensão $B \times B$.

3 MÉTODOS

Como visto no capítulo anterior, a literatura contempla trabalhos sobre a teoria de dados composicionais e geoestatística. Aitchison (1986) apresenta a teoria de dados composicionais considerando a independência das observações; Pawlowsky-Glahn e Olea (2004) fazem análise geoestatística de dados composicionais seguindo a teoria clássica da geoestatística e Lark e Bishop (2007) apresentam um estudo sobre cokrigagem de frações de partículas do solo concluindo que a predição de dados composicionais pode ser feita por cokrigagem alr com transformação de volta calculadas por quadratura de Gauss-Hermite para aproximar a esperança condicional. Por outro lado, Diggle e Ribeiro Jr. (2007), Schmidt e Sansó (2006) e outros trabalham com modelagem geoestatística uni e multivariada. Sob o paradigma bayesiano, tem-se os trabalhos de Obage (2007) que faz inferência bayesiana de dados composicionais sem considerar o efeito espacial e Tjelmeland e Lund (2003) que o consideram mas não adotam uma forma explícita para a função de covariância e não fazem predição espacial. Então, o que se apresenta a seguir é o desenvolvimento de uma metodologia para análise geoestatística de dados composicionais que compreende a definição do modelo para composições de três componentes, estimação dos parâmetros via verossimilhança e predição espacial tanto clássica como sob o paradigma bayesiano.

Todo o trabalho foi realizado utilizando recursos de *software* livre em ambiente operacional GNU/Linux; no ambiente estatístico R (R development Core Team, 2008), utilizando o pacote *geoR* (RIBEIRO JR.; DIGGLE, 2001), *compositions* (BOOGAART; TOLOSANA; BREN, 2009), *statmod* (SMYTH; HU; DUNN, 2009), *MCMCpack* (MARTIN; QUINN; PARK, 2009), *coda* (PLUMMER et al., 2006) e rotinas desenvolvidas especificamente para o desenvolvimento deste trabalho.

3.1 Modelo Geoestatístico Composicional

Para $\underline{X} = (X_1, \dots, X_B)'$ sendo uma composição com B componentes e $\underline{Y} = \left(\ln \left(\frac{X_1}{X_B} \right), \dots, \ln \left(\frac{X_{B-1}}{X_B} \right) \right)'$ um vetor com $B - 1$ elementos, o modelo geoestatístico com componente comum pode ser obtido seguindo a formulação dada em Diggle e Ri-

beiro Jr. (2007). Neste trabalho, considerou-se composições de apenas 3 componentes, $X_1 = \text{Areia}$, $X_2 = \text{Silte}$ e $X_3 = \text{Argila}$.

A partir do modelo geoestatístico bivariado com componente espacial comum (2.30) reescrito como (2.31), Seção 2.4, propõe-se o seguinte modelo geoestatístico bivariado composicional:

$$\begin{cases} Y_1(\underline{x}_i) &= \mu_1(\underline{x}_i) + S_1(\underline{x}_i) + Z_1(\underline{x}_i) \\ Y_2(\underline{x}_{i'}) &= \mu_2(\underline{x}_{i'}) + S_2(\underline{x}_{i'}) + Z_2(\underline{x}_{i'}) \end{cases}$$

em que $\underline{x}_i, \underline{x}_{i'} \in \mathbb{R}^2$, são as localizações amostrais $i, i' = 1, \dots, n_1$, onde n_1 é o tamanho da amostra; $Y_1 = \ln(X_1/X_3)$, $Y_2 = \ln(X_2/X_3)$ são as variáveis resposta do modelo de modo que $\underline{Y}_{n \times 1} = (Y_1(\underline{x}_1), Y_2(\underline{x}_1), \dots, Y_1(\underline{x}_{n_1}), Y_2(\underline{x}_{n_1}))'$; $S_j(\underline{x}) \sim N(0; \sigma_j^2)$ e $Z_j(\underline{x}) \sim N(0; \tau_j^2)$, $j = 1, 2$. Neste modelo, os efeitos aleatórios com estrutura espacial S_1 e S_2 foram substituídos por um efeito aleatório padronizado U . Assumiu-se que este efeito tem distribuição Gaussiana multivariada com vetor de médias iguais a zero e matriz de variância/covariância, com variâncias unitárias e covariâncias dadas pela função de correlação exponencial, ρ_U . Esta função é caracterizada pelo parâmetro de alcance, ϕ , que controla o decaimento da correlação como função da separação espacial entre duas localizações. No modelo bivariado geral as unidades de medida são preservadas nas constantes padronizadoras σ_1 e σ_2 , enquanto que no contexto considerado aqui são adimensionais. Os efeitos aleatórios Z_1 e Z_2 capturam a variabilidade não espacial incluindo a correlação, ρ , induzida pela estrutura composicional e o modelo pode então ser reescrito como:

$$\begin{cases} Y_1(\underline{x}_i) &= \mu_1(\underline{x}_i) + \sigma_1 U(\underline{x}_i; \phi) + Z_1(\underline{x}_i) \\ Y_2(\underline{x}_{i'}) &= \mu_2(\underline{x}_{i'}) + \sigma_2 U(\underline{x}_{i'}; \phi) + Z_2(\underline{x}_{i'}) \end{cases}$$

Sendo assim, $\underline{Y} \sim N(\underline{\mu}; \underline{\Sigma})$, com a matriz de covariâncias $\underline{\Sigma}$ composta pelos elementos

$$\begin{aligned} \text{Cov}(Y_1(\underline{x}_i); Y_1(\underline{x}_i)) &= \sigma_1^2 + \tau_1^2 & \text{Cov}(Y_1(\underline{x}_i); Y_1(\underline{x}_{i'})) &= \sigma_1^2 \rho_U(\underline{x}_i; \underline{x}_{i'}) \\ \text{Cov}(Y_2(\underline{x}_i); Y_2(\underline{x}_i)) &= \sigma_2^2 + \tau_2^2 & \text{Cov}(Y_2(\underline{x}_i); Y_2(\underline{x}_{i'})) &= \sigma_2^2 \rho_U(\underline{x}_i; \underline{x}_{i'}) \end{aligned}$$

e

$$\text{Cov}(Y_1(\underline{x}_i); Y_2(\underline{x}_{i'})) = \sigma_1 \sigma_2 I_2(i, i') + \tau_1 \tau_2 I_3(i, i')$$

com as funções indicadoras I_2 e I_3 definidas como:

$$I_2(i, i') = \begin{cases} 1 & , \text{ se } i = i' \\ \rho_U(\underline{x}_i; \underline{x}_{i'}) & , \text{ se } i \neq i' \end{cases} \quad I_3(i, i') = \begin{cases} \rho & , \text{ se } i = i' \\ 0 & , \text{ se } i \neq i' \end{cases}$$

ou seja,

$$\Sigma = \begin{bmatrix} \sigma_1^2 + \tau_1^2 & \sigma_1\sigma_2 + \tau_1\tau_2\rho & \sigma_1^2\rho_U(x_1;x_2) & \sigma_1\sigma_2\rho_U(x_1;x_2) & \cdots & \sigma_1^2\rho_U(x_1;x_{n_1}) & \sigma_1\sigma_2\rho_U(x_1;x_{n_1}) \\ \sigma_1\sigma_2 + \tau_1\tau_2\rho & \sigma_2^2 + \tau_2^2 & \sigma_1\sigma_2\rho_U(x_1;x_2) & \sigma_2^2\rho_U(x_1;x_2) & \cdots & \sigma_1\sigma_2\rho_U(x_1;x_{n_1}) & \sigma_2^2\rho_U(x_1;x_{n_1}) \\ \sigma_1^2\rho_U(x_2;x_1) & \sigma_1\sigma_2\rho_U(x_2;x_1) & \sigma_1^2 + \tau_1^2 & \sigma_1\sigma_2 + \tau_1\tau_2\rho & \cdots & \sigma_1^2\rho_U(x_2;x_{n_1}) & \sigma_1\sigma_2\rho_U(x_2;x_{n_1}) \\ \sigma_1\sigma_2\rho_U(x_2;x_1) & \sigma_2^2\rho_U(x_2;x_1) & \sigma_1\sigma_2 + \tau_1\tau_2\rho & \sigma_2^2 + \tau_2^2 & \cdots & \sigma_1\sigma_2\rho_U(x_2;x_{n_1}) & \sigma_2^2\rho_U(x_2;x_{n_1}) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_1^2\rho_U(x_{n_1};x_1) & \sigma_1\sigma_2\rho_U(x_{n_1};x_1) & \sigma_1^2\rho_U(x_{n_1};x_2) & \sigma_1\sigma_2\rho_U(x_{n_1};x_2) & \cdots & \sigma_1^2 + \tau_1^2 & \sigma_1\sigma_2 + \tau_1\tau_2\rho \\ \sigma_1\sigma_2\rho_U(x_{n_1};x_1) & \sigma_2^2\rho_U(x_{n_1};x_1) & \sigma_1\sigma_2\rho_U(x_{n_1};x_2) & \sigma_2^2\rho_U(x_{n_1};x_2) & \cdots & \sigma_1\sigma_2 + \tau_1\tau_2\rho & \sigma_2^2 + \tau_2^2 \end{bmatrix}.$$

3.1.1 Estimação dos Parâmetros do Modelo Geoestatístico Composicional

A inferência sobre o vetor de parâmetros $\underline{\theta} = (\mu_1, \mu_2, \sigma_1, \sigma_2, \tau_1, \tau_2, \phi, \rho)'$ é feita usando a teoria da verossimilhança cuja função de verossimilhança é

$$L(\underline{\theta}; \underline{Y}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\underline{Y} - \underline{\mu}_{\underline{Y}})' \Sigma^{-1} (\underline{Y} - \underline{\mu}_{\underline{Y}}) \right\}.$$

Fazendo-se a reparametrização: $\eta = \sigma_2/\sigma_1$; $v_1 = \tau_1/\sigma_1$; $v_2 = \tau_2/\sigma_1$, pode-se escrever

$$\Sigma = \sigma_1^2 \mathbf{R} + \tau_1^2 \mathbf{I}_b = \sigma_1^2 \mathbf{V},$$

em que

$$\mathbf{R} = \begin{bmatrix} \rho_U(x_1, x_1) & \eta\rho_U(x_1, x_1) & \rho_U(x_1, x_2) & \eta\rho_U(x_1, x_2) & \cdots & \rho_U(x_1, x_{n_1}) & \eta\rho_U(x_1, x_{n_1}) \\ \eta\rho_U(x_1, x_1) & \eta^2\rho_U(x_1, x_1) & \eta\rho_U(x_1, x_2) & \eta^2\rho_U(x_1, x_2) & \cdots & \eta^2\rho_U(x_1, x_{n_1}) & \eta^2\rho_U(x_1, x_{n_1}) \\ \rho_U(x_2, x_1) & \eta\rho_U(x_2, x_1) & \rho_U(x_2, x_2) & \eta\rho_U(x_2, x_2) & \cdots & \rho_U(x_2, x_{n_1}) & \eta\rho_U(x_2, x_{n_1}) \\ \eta\rho_U(x_2, x_1) & \eta^2\rho_U(x_2, x_1) & \eta\rho_U(x_2, x_2) & \eta^2\rho_U(x_2, x_2) & \cdots & \eta^2\rho_U(x_2, x_{n_1}) & \eta^2\rho_U(x_2, x_{n_1}) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_U(x_{n_1}, x_1) & \eta\rho_U(x_{n_1}, x_1) & \rho_U(x_{n_1}, x_2) & \eta\rho_U(x_{n_1}, x_2) & \cdots & \rho_U(x_{n_1}, x_{n_1}) & \eta\rho_U(x_{n_1}, x_{n_1}) \\ \eta\rho_U(x_{n_1}, x_1) & \eta^2\rho_U(x_{n_1}, x_1) & \eta\rho_U(x_{n_1}, x_2) & \eta^2\rho_U(x_{n_1}, x_2) & \cdots & \eta^2\rho_U(x_{n_1}, x_{n_1}) & \eta^2\rho_U(x_{n_1}, x_{n_1}) \end{bmatrix};$$

$$\mathbf{I}_b = \begin{bmatrix} v_1^2 & v_1 v_2 \rho & 0 & 0 & \cdots & 0 & 0 \\ v_1 v_2 \rho & v_2^2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & v_1^2 & v_1 v_2 \rho & \cdots & 0 & 0 \\ 0 & 0 & v_1 v_2 \rho & v_2^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & v_1^2 & v_1 v_2 \rho \\ 0 & 0 & 0 & 0 & \cdots & v_1 v_2 \rho & v_2^2 \end{bmatrix}$$

e, consequentemente,

$$\mathbf{V} = \begin{bmatrix} 1 + v_1^2 & \eta + v_1 v_2 \rho & \rho_U(x_1, x_2) & \eta \rho_U(x_1, x_2) & \cdots & \rho_U(x_1, x_{n_1}) & \eta \rho_U(x_1, x_{n_1}) \\ \eta + v_1 v_2 \rho & \eta^2 + v_2^2 & \eta \rho_U(x_1, x_2) & \eta^2 \rho_U(x_1, x_2) & \cdots & \eta \rho_U(x_1, x_{n_1}) & \eta^2 \rho_U(x_1, x_{n_1}) \\ \rho_U(x_2, x_1) & \eta \rho_U(x_2, x_1) & 1 + v_1^2 & \eta + v_1 v_2 \rho & \cdots & \rho_U(x_2, x_{n_1}) & \eta \rho_U(x_2, x_{n_1}) \\ \eta \rho_U(x_2, x_1) & \eta^2 \rho_U(x_2, x_1) & \eta + v_1 v_2 \rho & \eta^2 + v_2^2 & \cdots & \rho_U(x_2, x_{n_1}) & \eta^2 \rho_U(x_2, x_{n_1}) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_U(x_{n_1}, x_1) & \eta \rho_U(x_{n_1}, x_1) & \rho_U(x_{n_1}, x_2) & \eta \rho_U(x_{n_1}, x_2) & \cdots & 1 + v_1^2 & \eta + v_1 v_2 \rho \\ \eta \rho_U(x_{n_1}, x_1) & \eta^2 \rho_U(x_{n_1}, x_1) & \eta \rho_U(x_{n_1}, x_2) & \eta^2 \rho_U(x_{n_1}, x_2) & \cdots & \eta + v_1 v_2 \rho & \eta^2 + v_2^2 \end{bmatrix}.$$

A função de log-verossimilhança reparametrizada é dada por

$$l(\underline{\theta}; \underline{Y}) = -\frac{1}{2} \left(n \ln(2\pi) + 2n \ln(\sigma_1) + \ln(|\mathbf{V}|) + \frac{1}{\sigma_1^2} Qe \right). \quad (3.1)$$

Considerando $\underline{\mu}_Y = \mathbf{D}\underline{\mu}$ em que \mathbf{D} é a matriz do delineamento de ordem $n \times 2$, tem-se

$$Qe = (\underline{Y} - \underline{\mu}_Y)' \mathbf{V}^{-1} (\underline{Y} - \underline{\mu}_Y) = \underline{Y}' \mathbf{V}^{-1} \underline{Y} - 2(\underline{Y}' \mathbf{V}^{-1} \mathbf{D}) \underline{\mu} + \underline{\mu}' (\mathbf{D}' \mathbf{V}^{-1} \mathbf{D}) \underline{\mu}.$$

Expressões analíticas fechadas podem ser obtidas para os estimadores de máxima verossimilhança de $\underline{\mu} = (\mu_1, \mu_2)'$ e σ_1 diferenciando a função (3.1) em relação aos respectivos parâmetros e estes são dados por

$$\hat{\underline{\mu}} = (\mathbf{D}' \mathbf{V}^{-1} \mathbf{D})^{-1} (\mathbf{D}' \mathbf{V}^{-1} \underline{Y}) \quad \text{e} \quad \hat{\sigma}_1 = \sqrt{\hat{Q}e/n}. \quad (3.2)$$

É importante notar que $\hat{Q}e$ pode ser escrita como

$$\hat{Q}e = \underline{Y}' \mathbf{V}^{-1} \underline{Y} - (\underline{Y}' \mathbf{V}^{-1} \mathbf{D}) (\mathbf{D}' \mathbf{V}^{-1} \mathbf{D})^{-1} (\mathbf{D}' \mathbf{V}^{-1} \underline{Y}).$$

Ao substituir as expressões (3.2) em (3.1) obtém-se a função de log-verossimilhança concentrada

$$l(\underline{\theta}^*; \underline{Y}) = -\frac{1}{2} \left[\ln(|\mathbf{V}|) + n \left(\ln(2\pi) + \ln(\hat{Q}e) - \ln(n) + 1 \right) \right],$$

que é uma função do vetor de parâmetros desconhecidos $\underline{\theta}^* = (\eta, v_1, v_2, \phi, \rho)'$, e pode ser maximizada numericamente.

Os algoritmos de otimização empregados no processo de maximização foram “L-BFGS-B”, método de Byrd et al. (1995) que permite informar os limites inferior e superior de busca no espaço paramétrico; “Nelder-Mead”, uma implementação do método de Nelder e Mead (1965); “Gradiente Conjugado”, baseado no método de Fletcher e Reeves (1964) e “BFGS”, um método quasi-Newton. Todos já estão implementados no ambiente estatístico R (R Development Core Team, 2008).

3.1.2 Incerteza Sobre os Parâmetros do Modelo

Do processo de maximização obtém-se $\hat{\underline{\theta}}^* = (\hat{\eta}, \hat{v}_1, \hat{v}_2, \hat{\phi}, \hat{\rho})'$ e as respectivas variâncias através da matriz Hessiana numérica dada pela derivada segunda do logaritmo da função de verossimilhança em relação aos parâmetros em $\underline{\theta}^*$. A matriz Informação de Fisher observada é definida como o negativo da matriz Hessiana e é dada por

$$\mathbf{I}_F(\hat{\underline{\theta}}^*) = -\frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \hat{\underline{\theta}}^* \partial (\hat{\underline{\theta}}^*)'} = -\begin{bmatrix} \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \eta^2} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \eta \partial v_1} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \eta \partial v_2} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \eta \partial \phi} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \eta \partial \rho} \\ \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial v_1 \partial \eta} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial v_1^2} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial v_1 \partial v_2} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial v_1 \partial \phi} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial v_1 \partial \rho} \\ \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial v_2 \partial \eta} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial v_2 \partial v_1} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial v_2^2} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial v_2 \partial \phi} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial v_2 \partial \rho} \\ \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \phi \partial \eta} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \phi \partial v_1} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \phi \partial v_2} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \phi^2} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \phi \partial \rho} \\ \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \rho \partial \eta} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \rho \partial v_1} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \rho \partial v_2} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \rho \partial \phi} & \frac{\partial^2 l(\hat{\underline{\theta}}^*)}{\partial \rho^2} \end{bmatrix}.$$

Para se obter $\hat{\mu}_1$, $\hat{\mu}_2$, e $\hat{\sigma}_1$, basta substituir $\hat{\underline{\theta}}^*$ nas equações em (3.2).

Como o interesse está na obtenção de $\hat{\underline{\theta}} = (\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\tau}_1, \hat{\tau}_2, \hat{\phi}, \hat{\rho})'$ e suas respectivas variâncias, o método Delta (Subsecção 2.1.3) é aplicado para obter uma aproximação da distribuição de $\hat{\underline{\theta}}$ e maiores detalhes podem ser encontrados em DeGroot e Schervish (2002), Cox e Hinkley (1974), Azzalini (1996) e Pawitan (2001). Assintoticamente, a distribuição de $\hat{\underline{\theta}}$ será aproximadamente multivariada gaussiana com vetor de médias $\hat{\underline{\theta}} = g(\hat{\underline{\theta}}^*)$ e variância

$$\text{Var}(\hat{\underline{\theta}}) \geq \nabla g(\hat{\underline{\theta}}^*)' \mathbf{I}_{Fe}(\hat{\underline{\theta}}^*)^{-1} \nabla g(\hat{\underline{\theta}}^*),$$

em que \mathbf{I}_{Fe} é a matriz Informação de Fisher esperada e

$$\nabla g(\hat{\underline{\theta}}^*) = \left(\frac{\partial g(\hat{\underline{\theta}}^*)}{\partial \eta}, \frac{\partial g(\hat{\underline{\theta}}^*)}{\partial v_1}, \frac{\partial g(\hat{\underline{\theta}}^*)}{\partial v_2}, \frac{\partial g(\hat{\underline{\theta}}^*)}{\partial \phi}, \frac{\partial g(\hat{\underline{\theta}}^*)}{\partial \rho} \right)'$$

é a função escore $U(\hat{\underline{\theta}}^*)$. Assim, dado que:

- $\eta = \frac{\sigma_2}{\sigma_1} \Rightarrow \sigma_2 = \eta \sigma_1$, se $g(\hat{\underline{\theta}}^*) = \sigma_2$, então $\frac{\partial g(\hat{\underline{\theta}}^*)}{\partial \eta} = (\sigma_1, 0, 0, 0, 0)'$;
- $v_1 = \frac{\tau_1}{\sigma_1} \Rightarrow \tau_1 = v_1 \sigma_1$, se $g(\hat{\underline{\theta}}^*) = \tau_1$, então $\frac{\partial g(\hat{\underline{\theta}}^*)}{\partial v_1} = (0, \sigma_1, 0, 0, 0)'$;
- $v_2 = \frac{\tau_2}{\sigma_1} \Rightarrow \tau_2 = v_2 \sigma_1$, se $g(\hat{\underline{\theta}}^*) = \tau_2$, então $\frac{\partial g(\hat{\underline{\theta}}^*)}{\partial v_2} = (0, 0, \sigma_1, 0, 0)'$;
- $\phi = \phi$, se $g(\hat{\underline{\theta}}^*) = \phi$, então $\frac{\partial g(\hat{\underline{\theta}}^*)}{\partial \phi} = (0, 0, 0, 1, 0)'$;
- $\rho = \rho$, se $g(\hat{\underline{\theta}}^*) = \rho$, então $\frac{\partial g(\hat{\underline{\theta}}^*)}{\partial \rho} = (0, 0, 0, 0, 1)'$,

tem-se

$$\nabla g(\hat{\theta}^*) = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & 0 \\ 0 & \sigma_1 & 0 & 0 & 0 \\ 0 & 0 & \sigma_1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

e a matriz Informação de Fisher esperada para $\hat{\theta}^*$, baseada nos dados \underline{Y} , é substituída pela matriz $\mathbf{I}_F(\hat{\theta}^*)$ que é assintoticamente equivalente, de modo que

$$\text{Var}(\hat{\theta}) \geq \nabla g(\hat{\theta}^*)' \mathbf{I}_F(\hat{\theta}^*)^{-1} \nabla g(\hat{\theta}^*).$$

Para encontrar as variâncias para $\hat{\underline{\mu}}$ e $\hat{\sigma}_1$, através da função (3.1), obtém-se

$$\mathbf{I}_F(\underline{\mu}) = -\frac{\partial^2 l(\underline{\theta})}{\partial \underline{\mu}^2} = \frac{1}{\sigma_1^2} (\mathbf{D}' \mathbf{V}^{-1} \mathbf{D})' \quad \text{e} \quad \mathbf{I}_F(\sigma_1) = -\frac{\partial^2 l(\underline{\theta})}{\partial \sigma_1^2} = -\frac{n}{\sigma_1} + \frac{3Qe}{\sigma_1^3},$$

e portanto,

$$\begin{aligned} \text{Var}(\hat{\underline{\mu}}) &= \mathbf{I}_F(\hat{\underline{\mu}})^{-1} = \hat{\sigma}_1^2 (\mathbf{D}' \hat{\mathbf{V}}^{-1} \mathbf{D})^{-1} \\ \text{Var}(\hat{\sigma}_1) &= \mathbf{I}_F(\hat{\sigma}_1)^{-1} = \frac{\hat{\sigma}_1^3}{3\hat{Q}e - n\hat{\sigma}_1}. \end{aligned}$$

Na continuação do trabalho, a predição espacial, ou melhor, a cokrigagem foi feita como explicado e conforme o Sistema de Equações (2.32) na Subseção 2.4.2.

3.1.3 Transformação de Volta Dos Valores Preditos Para o Espaço Amostral Simplex

Uma vez que a transformação ALR foi aplicada aos dados originais e o procedimento de estimação e cokrigagem foi realizada com os dados transformados em \mathbb{R}^2 , deve-se fazer a transformação de volta do vetor de médias e da matriz de covariância para o espaço amostral original, o simplex \mathbb{S}^3 , como descrito em Pawlowsky-Glahn e Olea (2004).

O objetivo é calcular para cada localização uma estimativa de

$$\underline{\mu}_X = E(\underline{X}) = \int_{\mathbb{S}^B} \underline{X} f(\underline{X}) d\underline{X} \quad (3.3)$$

e

$$\Sigma_X = \text{Cov}(\underline{X}, \underline{X}) = \int_{\mathbb{S}^B} (\underline{X} - \underline{\mu}_X)(\underline{X} - \underline{\mu}_X)' f(\underline{X}) d\underline{X}. \quad (3.4)$$

Sabe-se que

$$f(\underline{Y}) = (2\pi)^{-\frac{B-1}{2}} |\Sigma_{\underline{Y}}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\underline{Y} - \underline{\mu}_{\underline{Y}})' \Sigma_{\underline{Y}}^{-1} (\underline{Y} - \underline{\mu}_{\underline{Y}}) \right\},$$

segue uma distribuição Gaussiana multivariada e $\underline{Y} = \text{alr}(\underline{X})$. Assim, para voltar a escala original dos dados, $\underline{X} = \text{agl}(\underline{Y})$, é usual utilizar o método do jacobiano (JAMES, 2004) cuja transformação é dada por

$$J_{\text{alr}(\underline{X})} = \left| \frac{\partial \underline{Y}}{\partial \underline{X}} \right| = \left(\prod_{i=1}^B X_i \right)^{-1} \quad (3.5)$$

de modo que $f(\underline{X})$ se torna

$$f(\underline{X}) = (2\pi)^{-\frac{B-1}{2}} |\Sigma_{\underline{Y}}|^{-\frac{1}{2}} \left(\prod_{i=1}^B X_i \right)^{-1} \exp \left\{ -\frac{1}{2} (\text{alr}(\underline{X}) - \underline{\mu}_{\underline{Y}})' \Sigma_{\underline{Y}}^{-1} (\text{alr}(\underline{X}) - \underline{\mu}_{\underline{Y}}) \right\}.$$

Aitchison (1986), Pawlowsky-Glahn e Olea (2004) afirmam que para se resolver as integrais (3.3) e (3.4) estas devem ser expressas como

$$\underline{\mu}_{\underline{X}} = \int_{\mathbb{R}^{B-1}} g_1(\underline{Z}) f(-\underline{Z}'\underline{Z}) d\underline{Z} \quad (3.6)$$

e

$$\Sigma_{\underline{X}} = \int_{\mathbb{R}^{B-1}} g_2(\underline{Z}) f(-\underline{Z}'\underline{Z}) d\underline{Z} \quad (3.7)$$

em que \underline{Z} é a transformação

$$\underline{Z} = \frac{\text{alr}(\underline{X}) - \underline{\mu}_{\underline{Y}}}{\sqrt{2}\mathbf{R}'} = \frac{1}{\sqrt{2}} (\mathbf{R}')^{-1} (\text{alr}(\underline{X}) - \underline{\mu}_{\underline{Y}}), \quad (3.8)$$

e \mathbf{R} é a decomposição Cholesky (BURDEN; FAIRES, 2003) de $\Sigma_{\underline{Y}}$, uma matriz triangular superior.

Desta forma, as integrais (3.6) e (3.7) podem, então, ser aproximadas pela integração de Gauss-Hermite multivariada de ordem k :

$$\int_{\mathbb{R}^{B-1}} g(\underline{Z}) f(-\underline{Z}'\underline{Z}) d\underline{Z} \approx \sum_{i_1=1}^k \sum_{i_2=1}^k \cdots \sum_{i_{B-1}=1}^k \omega_{i_1} \omega_{i_2} \cdots \omega_{i_{B-1}} g(Z_{i_1}, Z_{i_2}, \dots, Z_{i_{B-1}}), \quad (3.9)$$

em que os pesos $\omega_{i_1} \omega_{i_2} \cdots \omega_{i_{B-1}}$ e as abscissas $Z_{i_1}, Z_{i_2}, \dots, Z_{i_{B-1}}$ são conhecidos e seus valores podem ser encontrados, por exemplo, em Abramowitz e Stegun (1972, p. 924). Segundo Gamerman (1997) ordens de quadratura de 6 a 8 são suficientes para aproximar a integral.

Explicitando o procedimento, observa-se que a integral (3.3) pode ser reescrita como

$$\underline{\mu}_X = \int_{\mathbb{S}^B} g(\underline{X}) \left((2\pi)^{-\frac{B-1}{2}} |\Sigma_Y|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\text{alr}(\underline{X}) - \underline{\mu}_Y)' \Sigma_Y^{-1} (\text{alr}(\underline{X}) - \underline{\mu}_Y) \right\} \right) d\underline{X},$$

onde

$$g(\underline{X}) = \underline{X} \left(\prod_{i=1}^B X_i \right)^{-1}. \quad (3.10)$$

Considere que $\Sigma_Y^{-1} = (\mathbf{R}'\mathbf{R})^{-1} = \mathbf{R}^{-1}(\mathbf{R}')^{-1} = \mathbf{R}^{-1}(\mathbf{R}^{-1})'$. Da transformação (3.8), tem-se

$$\underline{X} = \text{agl}(\underline{\mu}_Y + \sqrt{2}\mathbf{R}'\underline{Z}) = \text{agl}(\underline{Y})$$

e de (3.5),

$$\partial \underline{X} = \left(\prod_{i=1}^B X_i \right) \partial \underline{Y}. \quad (3.11)$$

Além disso,

$$J = \left| \frac{\partial \underline{Y}}{\partial \underline{Z}} \right| = \left| \frac{\partial(\underline{\mu}_Y + \sqrt{2}\mathbf{R}'\underline{Z})}{\partial \underline{Z}} \right| = \left| \sqrt{2}\mathbf{R}' \right| = (\sqrt{2})^{B-1} |\mathbf{R}'|;$$

sendo \mathbf{R} a decomposição Cholesky de Σ_Y ,

$$|\mathbf{R}'| = |\mathbf{R}| = |\Sigma_Y|^{\frac{1}{2}},$$

e

$$J = \left| \frac{\partial \underline{Y}}{\partial \underline{Z}} \right| = 2^{\frac{B-1}{2}} |\Sigma_Y|^{\frac{1}{2}} \Rightarrow \partial \underline{Y} = 2^{\frac{B-1}{2}} |\Sigma_Y|^{\frac{1}{2}} \partial \underline{Z}. \quad (3.12)$$

Substituindo a equação (3.12) em (3.11) tem-se

$$\partial \underline{X} = \left(\prod_{i=1}^B X_i \right) 2^{\frac{B-1}{2}} |\Sigma_Y|^{\frac{1}{2}} \partial \underline{Z} \Rightarrow \partial \underline{Z} = 2^{-\frac{B-1}{2}} \left(\prod_{i=1}^B X_i \right)^{-1} |\Sigma_Y|^{-\frac{1}{2}} \partial \underline{X}.$$

Por outro lado, de (3.10) tem-se $g(\text{agl}(\underline{\mu}_Y + \sqrt{2}\mathbf{R}'\underline{Z})) = \text{agl}(\underline{\mu}_Y + \sqrt{2}\mathbf{R}'\underline{Z}) \left(\prod_{i=1}^B X_i \right)^{-1}$ que substituída em (3.6) produz

$$\underline{\mu}_X = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \pi^{-\frac{B-1}{2}} \text{agl}(\underline{\mu}_Y + \sqrt{2}\mathbf{R}'\underline{Z}) \exp\{\underline{Z}'\underline{Z}\} d\underline{Z} \quad (3.13)$$

com $g_1(\underline{Z}) = \pi^{-\frac{B-1}{2}} \text{agl}(\underline{\mu}_Y + \sqrt{2}\mathbf{R}'\underline{Z})$. Logo, a aproximação de Gauss-Hermite (3.9) de ordem 3, por exemplo, para $\underline{\mu}_X$ é

$$\begin{aligned} \underline{\mu}_X &\approx \sum_{i_1=1}^3 \sum_{i_2=1}^3 \omega_{i_1} \omega_{i_2} g_1(\underline{Z}_{i_1}, \underline{Z}_{i_2}) \\ &\approx \omega_1^2 g_1(\underline{Z}_1, \underline{Z}_1) + \omega_1 \omega_2 [g_1(\underline{Z}_1, \underline{Z}_2) + g_1(\underline{Z}_2, \underline{Z}_1)] + \omega_1 \omega_3 [g_1(\underline{Z}_1, \underline{Z}_3) + g_1(\underline{Z}_3, \underline{Z}_1)] + \\ &\quad + \omega_2^2 g_1(\underline{Z}_2, \underline{Z}_2) + \omega_2 \omega_3 [g_1(\underline{Z}_2, \underline{Z}_3) + g_1(\underline{Z}_3, \underline{Z}_2)] + \omega_3^2 g_1(\underline{Z}_3, \underline{Z}_3). \end{aligned}$$

Lembrando que $\underline{\mu}_X$ representa a média da composição em cada localização, esta é um vetor trivariado. Então na função g_1 , $\underline{\mu}_Y$ é um vetor bivariado extraído de $\underline{\mu}_{Y_0|Y}$, cujo primeiro elemento corresponde à média da cokrigagem para a variável Y_1 , o segundo para a variável Y_2 e \mathbf{R}' de ordem 2×2 será a correspondente matriz de variância/covariâncias de Y_1, Y_2 , extraída do bloco diagonal da matriz $\Sigma_{Y_0|Y}$. Assim, em cada localização a resolução de (3.13) implica na resolução de 3 integrais.

O mesmo procedimento é aplicado para a obtenção de Σ_X dada pela integral (3.7), agora reescrita como

$$\Sigma_X = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \pi^{-\frac{B-1}{2}} \left(\text{agl}(\underline{\mu}_Y + \sqrt{2}\mathbf{R}'\underline{Z}) - \underline{\mu}_X \right) \left(\text{agl}(\underline{\mu}_Y + \sqrt{2}\mathbf{R}'\underline{Z}') - \underline{\mu}_X \right)' \exp\{\underline{Z}'\underline{Z}\} d\underline{Z} \quad (3.14)$$

com g_2 dada por $g_2(\underline{Z}) = \pi^{-\frac{B-1}{2}} \left(\text{agl}(\underline{\mu}_Y + \sqrt{2}\mathbf{R}'\underline{Z}) - \underline{\mu}_X \right) \left(\text{agl}(\underline{\mu}_Y + \sqrt{2}\mathbf{R}'\underline{Z}) - \underline{\mu}_X \right)'$ e, neste caso, sendo Σ_X uma matriz de ordem 3×3 , a resolução da integral (3.14) para cada localização implica na resolução de 9 integrais. Portanto, a estimação do vetor de médias e variâncias em n_2 localizações implica na resolução de $n_2 \times 12$ integrais.

Finalizando, uma análise geoestatística de dados composicionais pode ser feita seguindo os passos:

- a. escrever o modelo geoestatístico bivariado para dados composicionais;
- b. escrever a função de verossimilhança do modelo;
- c. maximizar a função de verossimilhança, obtendo-se as estimativas dos parâmetros;
- d. fazer a cokrigagem obtendo-se o vetor de valores esperados preditos e a matriz de covariância no \mathbb{R}^2 ;
- e. utilizar a integração de Gauss-Hermite ou fazer simulação para obter os valores preditos na escala original, \mathbb{S}^3 ;
- f. construir mapas dos valores esperados preditos para cada componente da composição.

No item (e), no caso de simulação, utiliza-se os seguintes passos para cada localização:

- e.1 gerar dados de uma distribuição gaussiana multivariada com vetor de médias e matriz de covariância iguais ao vetor de médias e matriz de covariância obtidos por cokrigagem (item (d));
- e.2 aplicar a transformação AGL no vetor de valores esperados obtidos no item (e.1);
- e.6 calcular a média de cada componente ou funcionais destes.

Neste trabalho, os funcionais calculados foram os valores máximos e mínimos obtidos em 1000 simulações, em cada localização, com os quais foram contruídos mapas.

3.2 Inferência Bayesiana Para o Modelo Geoestatístico Composicional

Uma análise geoestatística de dados composicionais também pode ser feita por inferência bayesiana. Como visto na Subseção 2.3.1, sob o paradigma bayesiano o vetor de parâmetros $\underline{\theta} = (\underline{\mu}, \sigma_1, \sigma_2, \tau_1, \tau_2, \phi, \rho)'$ é considerado vetor aleatório, a função de log-verossimilhança do modelo é a mesma dada na Equação (2.24) e para se fazer inferência bayesiana é preciso construir a distribuição conjunta dos vetores aleatórios \underline{Y} e $\underline{\theta}$ que substituída na Equação (2.25) fornece a distribuição à posteriori para $\underline{\theta}$

$$P(\underline{\theta}|\underline{Y}) = \frac{L(\underline{\theta}, \underline{Y})P(\underline{\theta})}{P(\underline{Y})} \propto L(\underline{\theta}, \underline{Y})P(\underline{\theta})$$

ou

$$P(\underline{\mu}, \sigma_1, \sigma_2, \tau_1, \tau_2, \phi, \rho|\underline{Y}) \propto P(\underline{Y}|\underline{\mu}, \sigma_1, \sigma_2, \tau_1, \tau_2, \phi, \rho)P(\underline{\mu}, \sigma_1, \sigma_2, \tau_1, \tau_2, \phi, \rho).$$

Observa-se que para o modelo geoestatístico bivariado composicional não é possível derivar analiticamente as distribuições marginais a *posteriori* para todos os parâmetros do modelo. Neste caso, considerando a reparametrização apresentada na Subseção 3.1.1 com $\underline{\theta}^* = (\eta, v_1, v_2, \phi, \rho)$ e supondo $(\underline{\mu}, \sigma_1^2)$ e $\underline{\theta}^*$ independentes pode-se escrever

$$P(\underline{\mu}, \sigma_1^2, \underline{\theta}^*|\underline{Y}) \propto P(\underline{Y}|\underline{\mu}, \sigma_1^2, \underline{\theta}^*)P(\underline{\mu}, \sigma_1^2)P(\underline{\theta}^*), \quad (3.15)$$

e então é possível obter expressões fechadas para as distribuições marginais para $\underline{\mu}$ e σ_1^2 . Para

isto, observa-se que integrando ambos os lados da Equação (3.15) em relação à $\underline{\theta}^*$ tem-se

$$\int_{\underline{\theta}^*} P(\underline{\mu}, \sigma_1^2, \underline{\theta}^* | \underline{Y}) d\underline{\theta}^* \propto \int_{\underline{\theta}^*} P(\underline{Y} | \underline{\mu}, \sigma_1^2, \underline{\theta}^*) P(\underline{\mu}, \sigma_1^2) P(\underline{\theta}^*) d\underline{\theta}^*.$$

e supondo $1/\sigma_1^2$ como *priori* para $P(\underline{\mu}, \sigma_1^2)$ obtém-se

$$P(\underline{\mu}, \sigma_1^2 | \underline{Y}) \propto P(\underline{Y} | \underline{\mu}, \sigma_1^2, \underline{\theta}^*) \frac{1}{\sigma_1^2} \int_{\underline{\theta}^*} P(\underline{\theta}^*) d\underline{\theta}^*. \quad (3.16)$$

Não sendo possível derivar analiticamente a distribuição a *posteriori* para $\underline{\theta}^*$ e tampouco calcular analiticamente a integral em (3.16), esta é resolvida numericamente usando o algoritmo de Metropolis-Hastings. As *prioris* utilizadas para η , \mathbf{v}_1 e \mathbf{v}_2 foram lognormais com parâmetros correpondentes às razões das estimativas de máxima verossimilhança obtidas no processo de maximização. Para ϕ utilizou-se uma distribuição Gama com parâmetros (66; 1) em que o valor 66 é a estimativa de máxima verossimilhança de ϕ e para ρ uma *priori flat*. Os valores estimados obtidos foram então substituídos em (3.16) de modo que:

$$\begin{aligned} P(\underline{\mu}, \sigma_1^2 | \underline{\theta}^*, \underline{Y}) &\propto (\sigma_1^2)^{-1} P(\underline{Y} | \underline{\mu}, \sigma_1^2, \underline{\theta}^*) \\ &\propto (\sigma_1^2)^{-(\frac{n}{2}+1)} \exp \left\{ -\frac{1}{2} (\underline{Y} - \underline{\mu}_Y)' \mathbf{V}^{-1} (\underline{Y} - \underline{\mu}_Y) \right\}. \end{aligned}$$

Da definição de probabilidade condicional tem-se

$$P(\underline{\mu} | \sigma_1^2, \underline{\theta}^*, \underline{Y}) = \frac{P(\underline{Y} | \underline{\mu}, \sigma_1^2, \underline{\theta}^*) P(\underline{\mu} | \sigma_1^2, \underline{\theta}^*) P(\sigma_1^2, \underline{\theta}^*)}{P(\underline{Y} | \sigma_1^2, \underline{\theta}^*) P(\sigma_1^2, \underline{\theta}^*)},$$

e conseqüentemente,

$$P(\underline{\mu} | \sigma_1^2, \underline{\theta}^*, \underline{Y}) \propto P(\underline{\mu} | \sigma_1^2, \underline{\theta}^*) P(\underline{Y} | \underline{\mu}, \sigma_1^2, \underline{\theta}^*).$$

Assim, supondo uma *priori flat* para $P(\underline{\mu} | \sigma_1^2, \underline{\theta}^*)$ vem

$$\begin{aligned} P(\underline{\mu} | \sigma_1^2, \underline{\theta}^*, \underline{Y}) &\propto P(\underline{Y} | \underline{\mu}, \sigma_1^2, \underline{\theta}^*) \\ &\propto (\sigma_1^2)^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2\sigma_1^2} (\underline{\mu} - \hat{\underline{\mu}})' (\mathbf{D}' \mathbf{V}^{-1} \mathbf{D}) (\underline{\mu} - \hat{\underline{\mu}})' \right\} \times \\ &\quad \exp \left\{ -\frac{1}{2\sigma_1^2} (\underline{Y} - \mathbf{D} \hat{\underline{\mu}})' \mathbf{V}^{-1} (\underline{Y} - \mathbf{D} \hat{\underline{\mu}}) \right\} \\ &\propto (\sigma_1^2)^{-\frac{n}{2}} \exp \left\{ -\frac{1}{2\sigma_1^2} (\underline{\mu} - \hat{\underline{\mu}})' (\mathbf{D}' \mathbf{V}^{-1} \mathbf{D}) (\underline{\mu} - \hat{\underline{\mu}})' \right\}, \end{aligned}$$

e a distribuição a *posteriori* marginal para $\underline{\mu} | \sigma_1^2, \underline{\theta}^*, \underline{Y}$ é

$$[\underline{\mu} | \sigma_1^2, \underline{\theta}^*, \underline{Y}] \sim N(\hat{\underline{\mu}}; \sigma_1^2 (\mathbf{D}' \mathbf{V}^{-1} \mathbf{D})^{-1}). \quad (3.17)$$

Como

$$P(\underline{\mu}, \sigma_1^2 | \underline{\theta}^*, \underline{Y}) = P(\underline{\mu} | \sigma_1^2, \underline{\theta}^*, \underline{Y}) P(\sigma_1^2 | \underline{\theta}^*, \underline{Y}) \Rightarrow P(\sigma_1^2 | \underline{\theta}^*, \underline{Y}) = \frac{P(\underline{\mu}, \sigma_1^2 | \underline{\theta}^*, \underline{Y})}{P(\underline{\mu} | \sigma_1^2, \underline{\theta}^*, \underline{Y})}, \quad (3.18)$$

substituindo (3.17) em (3.18) tem-se:

$$\begin{aligned} P(\sigma_1^2 | \underline{\theta}^*, \underline{Y}) &\propto (\sigma_1^2)^{-\left(\frac{n-n_{\underline{\mu}}}{2}+1\right)} \exp \left\{ -\frac{1}{2\sigma_1^2} (\underline{Y} - \mathbf{D}\hat{\underline{\mu}})' \mathbf{V}^{-1} (\underline{Y} - \mathbf{D}\hat{\underline{\mu}}) \right\} \\ &\propto (\sigma_1^2)^{-\left(\frac{n-n_{\underline{\mu}}}{2}+1\right)} \exp \left\{ -\frac{1}{2\sigma_1^2} n \hat{\sigma}_1^2 \right\} \\ &\propto (\sigma_1^2)^{-\left(\frac{n-n_{\underline{\mu}}}{2}+1\right)} \exp \left\{ -\frac{1}{2\sigma_1^2} \frac{(n-n_{\underline{\mu}})}{(n-n_{\underline{\mu}})} n \hat{\sigma}_1^2 \right\} \\ &\propto (\sigma_1^2)^{-\left(\frac{n-n_{\underline{\mu}}}{2}+1\right)} \exp \left\{ -\frac{1}{2\sigma_1^2} (n-n_{\underline{\mu}}) S^2 \right\}, \end{aligned}$$

e a distribuição marginal a *posteriori* para $\sigma_1^2 | \underline{\theta}^*, \underline{Y}$ é

$$[\sigma_1^2 | \underline{\theta}^*, \underline{Y}] \sim \chi_{\text{Sinv}}^2(n - n_{\underline{\mu}}; S^2), \quad (3.19)$$

onde χ_{Sinv}^2 é a distribuição qui-quadrado escalonada inversa e

$$S^2 = \frac{n}{n - n_{\underline{\mu}}} \hat{\sigma}_1^2 = \frac{(\underline{Y} - \mathbf{D}\hat{\underline{\mu}})' \mathbf{V}^{-1} (\underline{Y} - \mathbf{D}\hat{\underline{\mu}})}{n - n_{\underline{\mu}}}. \quad (3.20)$$

Por fim, com o desenvolvimento descrito pode-se resumir os passos para uma análise bayesiana de dados composicionais espacializados, segundo o modelo proposto, da seguinte forma:

- a. derivar analiticamente as distribuições a *posteriori* para $\underline{\mu}$ e σ_1^2 ;
- b. executar o algoritmo Metropolis-Hastings no vetor $\underline{\theta}^*$;
- c. com o vetor $\underline{\theta}^*$ estimado, calcular a matriz de covariância \mathbf{V} do modelo;
- d. calcular $\hat{\underline{\mu}}$ dado em (3.2);
- e. calcular S^2 dado em (3.20);
- f. executar um passo Gibbs que significa amostrar um valor de $\sigma_1^2 | \underline{\theta}^*, \underline{Y}$ da distribuição (3.19);
- g. com $\hat{\underline{\mu}}$ e $\text{Var}(\underline{\mu} | \sigma_1^2, \underline{\theta}^*, \underline{Y})$ executar um passo Gibbs amostrando um valor de $\underline{\mu} | \sigma_1^2, \underline{\theta}^*, \underline{Y}$ de (3.17);
- h. repetir o procedimento até o número de simulações desejado;
- i. fazer a cokrigagem com cada conjunto de parâmetros simulados;

- j. gerar uma amostra de uma distribuição normal multivariada com cada resultado obtido no item (i), vetor de médias e matriz de covariância;
- k. construir mapa de predição para cada componente da composição.

3.3 Dados Simulados

Como um estudo da metodologia proposta na Seção 3.1 analisou-se três conjuntos de dados simulados, cada um com 100 valores de X_1 , X_2 e X_3 cujas localizações em um quadrado unitário encontram-se representadas na Figura 3.1.

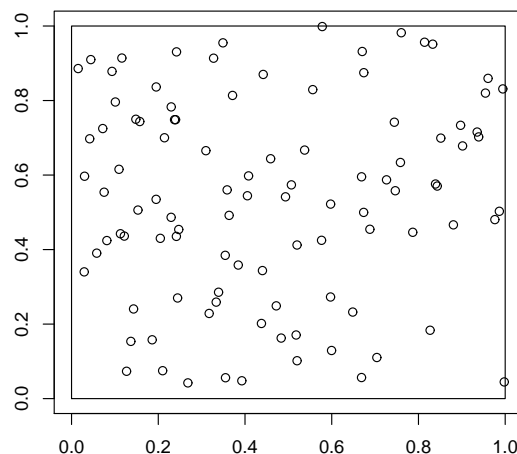


Figura 3.1: Distribuição das localizações no quadrado unitário.
FONTE: A autora (2010).

O primeiro conjunto, denominado configuração 1, foi gerado a partir do vetor de parâmetros $\underline{\theta}=(\mu_1, \mu_2, \sigma_1, \sigma_2, \tau_1, \tau_2, \phi, \rho) = (-0,2; -0,5; 1; 1,5; 0,3; 0,3; 0,6; 0,9)$. Atribuindo-se valores próximos e de mesmo sinal para as médias, fez-se com que a dispersão de pontos se situasse na parte central do triângulo. O fato de estarem concentrados deu-se pelo alto valor de ρ : quanto menor, mais espalhados. O fato das variâncias dos efeitos espaciais serem próximos com $\sigma_1 < \sigma_2$ fez com que os pontos se aproximassem mais do vértice X_2 do que do X_1 e, por último, valores iguais para as variâncias τ_1 e τ_2 justificou os pontos no interior do diagrama.

Em cada uma das localizações tem-se os valores de X_1 , X_2 e X_3 e as 100 composições representadas por círculos em um diagrama ternário podem ser vistas na Figura 3.2. Composições

próximas a um vértice têm altas proporções do componente correspondente àquele vértice, de modo que esta figura indica que as maiores percentagens na amostra ocorreram para componente X_3 . Também, pode-se observar, por exemplo, que a variabilidade da razão X_3/X_1 foi maior que a variabilidade da razão X_2/X_1 , já que a amplitude da intersecção das linhas partindo do vértice X_2 , passando por todos os pontos, com o lado $\overline{X_1X_3}$ é maior que a amplitude da intersecção das linhas com o lado $\overline{X_1X_2}$. Por outro lado, a variabilidade da razão X_1/X_3 , X_2/X_3 foram similares. Além disso, uma região de confiança de 4-desvios-padrão contemplou todas as composições e o centro da distribuição dado pelo fechamento da média geométrica dos três componentes está representado pelo ponto vermelho no diagrama. Pelos histogramas nota-se que as percentagens dos três componentes não sugerem uma distribuição gaussiana.

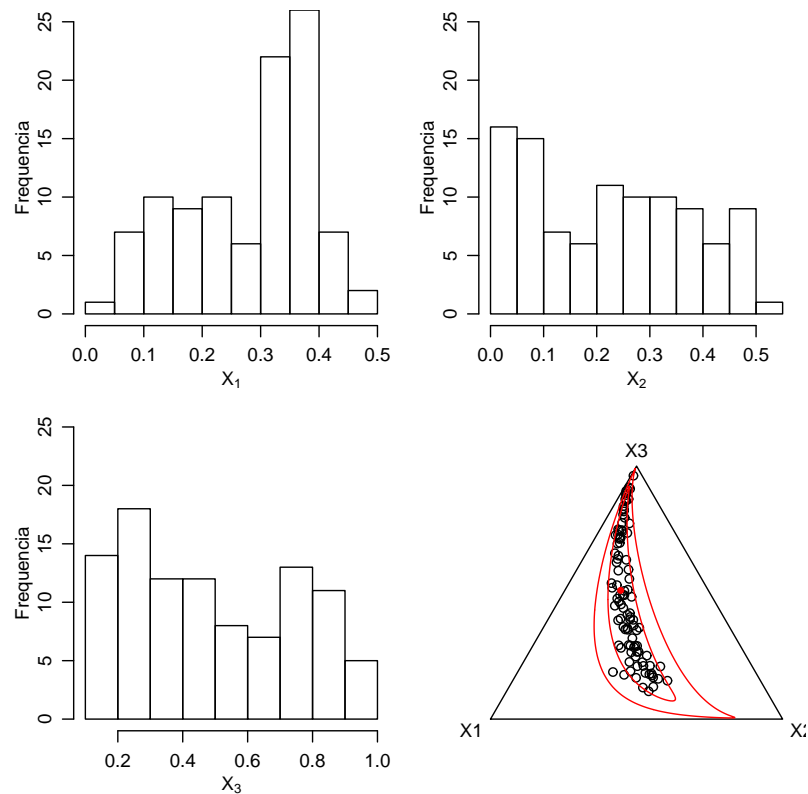


Figura 3.2: Distribuição de X_1 , X_2 e X_3 e diagrama ternário das composições para a primeira configuração.

FONTE: A autora (2010).

Porém, ao fazer a transformação ALR nos dados simulados obteve-se as variáveis $Y_1 = \ln(X_1/X_3)$ e $Y_2 = \ln(X_2/X_3)$ e os dados transformados passaram a seguir distribuição gaussiana como pode ser visto pela Figura 3.3. O diagrama de dispersão, por sua vez, mostra pontos alinhados em forma linear crescente indicando uma possível correlação linear positiva.

Em seguida, fez-se o ajuste do modelo. Dentre os métodos de otimização testados adotou-se o método L-BFGS-B por não apresentar problemas de convergência e as estimativas

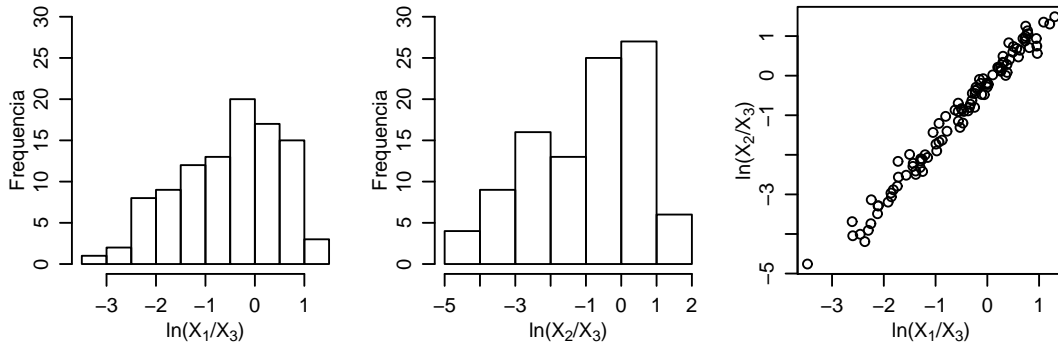


Figura 3.3: Distribuição das log-razão e correspondente diagrama de dispersão para a primeira configuração.

FONTE: A autora (2010).

Tabela 3.1: Estimativas, erros padrão e intervalos de confiança para a primeira configuração, pelo método delta via método de otimização “L-BFGS-B”.

Parâmetro	Valor	Estimativa	Erro Padrão	LI. Delta	LS. Delta
μ_1	-0,2	-0,9925955	1,60663318	-2,5992287	0,6140377
μ_2	-0,5	-1,6890643	1,98310148	-3,6721658	0,2940372
σ_1	1	1,1530662	0,10190384	1,0511623	1,2549700
σ_2	1,5	1,7581358	0,05412534	1,7040104	1,8122611
τ_1	0,3	0,4004140	0,02972426	0,3706897	0,4301382
τ_2	0,3	0,4274941	0,04045833	0,3870357	0,4679524
ϕ	0,6	0,9530621	0,51284603	0,4402160	1,4659081
ρ	0,9	0,9573298	0,02626424	0,9310656	0,9835941

FONTE: A autora (2010).

obtidas para os parâmetros bem como os respectivos intervalos de confiança calculados pelo método Delta são apresentados na Tabela 3.1. Observa-se, em geral, estimativas não muito próximas aos valores atribuídos aos parâmetros mas com mesmos sinais e que apenas três de oito intervalos contêm os valores atribuídos.

No processo de otimização os valores iniciais para o vetor de médias foram considerados como as médias dos valores observados para Y_1 e Y_2 . Metade da variância calculada para Y_1 foi atribuída para o efeito espacial e a outra metade para o efeito composicional. Da mesma forma, procedeu-se com Y_2 . O valor inicial para ρ foi calculado como o coeficiente de correlação de Pearson e $\phi = \min + 0,2(\max - \min)$, onde “min” e “max” foram, respectivamente, a menor e maior distância entre duas localizações. Então, estes valores indicaram uma log-verossimilhança concentrada de $-301,35132$. Após o processo de otimização, a log-verossimilhança maximizada foi $59,64009$.

Como resultado da cokrigagem obteve-se, para cada localização, o vetor de médias e a matriz de covariância no espaço \mathbb{R}^2 . A volta dos valores preditos para o simplex \mathbb{S}^3 foi feita

em uma grade constituída de 1156 pontos usando aproximação de Gauss-Hermite com ordem de quadratura $k = 7$. Na Figura 3.4 tem-se os mapas dos três componentes conforme a configuração 1 e, assim como revelou o diagrama ternário, as maiores proporções ocorreram para o componente X_3 . A Figura 3.5, por sua vez, mostra boa concordância entre os valores observados e preditos para os três componentes.

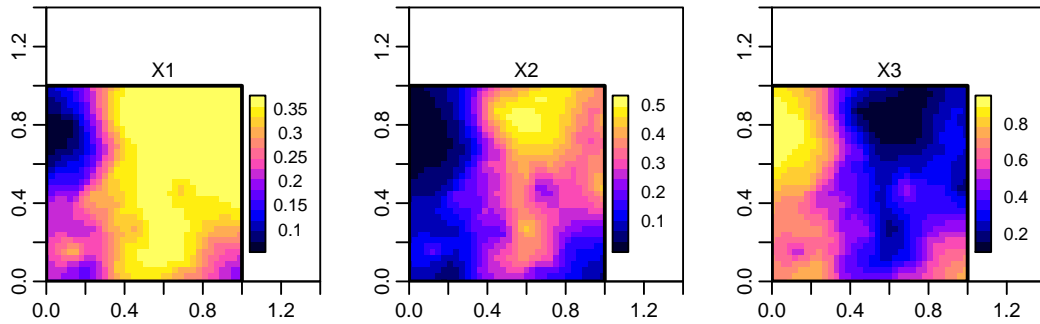


Figura 3.4: Mapas das percentagens de X_1 , X_2 e X_3 para dados da configuração 1. FONTE: A autora (2010).

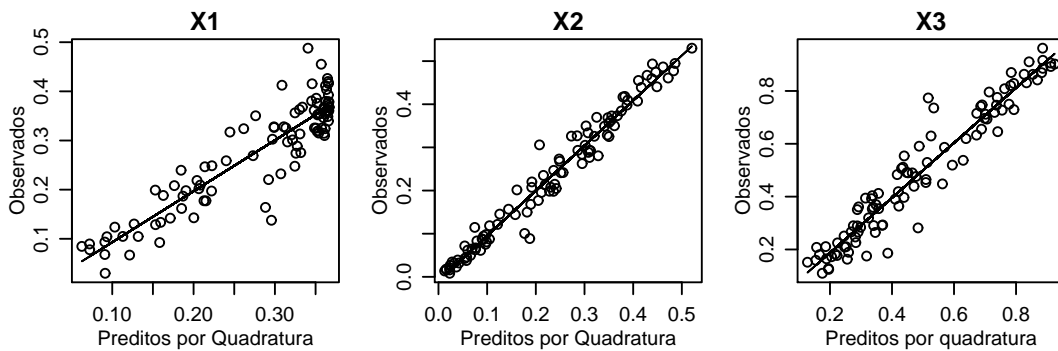


Figura 3.5: Valores observados versus preditos de X_1 , X_2 e X_3 para a configuração 1. FONTE: A autora (2010).

O segundo conjunto de dados, configuração 2, foi gerado considerando-se o vetor de parâmetros $(\mu_1, \mu_2, \sigma_1, \sigma_2, \tau_1, \tau_2, \phi, \rho) = (1; 1; 1, 2; 1, 5; 0, 9; 1; 0, 6; 0, 5)$. Assim como na configuração 1, em relação às médias, buscou-se fazer com que a dispersão de pontos continuasse na parte central do diagrama; a proximidade dos pontos ao vértice X_3 justificada pelo baixo valor de ϕ e o espalhamento pelo baixo valor de ρ . Também aqui, $\sigma_1 < \sigma_2$ fez com que os pontos se aproximassem mais de X_2 do que de X_1 e a diferença entre τ_1 e τ_2 forçou os pontos a se aproximarem mais do lado $\overline{X_1 X_2}$, com maior aproximação para X_2 por τ_1 ser o menor valor.

Para este vetor de parâmetros a configuração das composições no diagrama ternário da Figura 3.6 apresentou-se mais espalhada em relação à anterior com maiores percentagens

relativas aos componentes X_2 e X_3 . Ao mesmo tempo, observa-se muitas composições com percentagens muito baixas de X_3 o que se confirma pelo respectivo histograma.

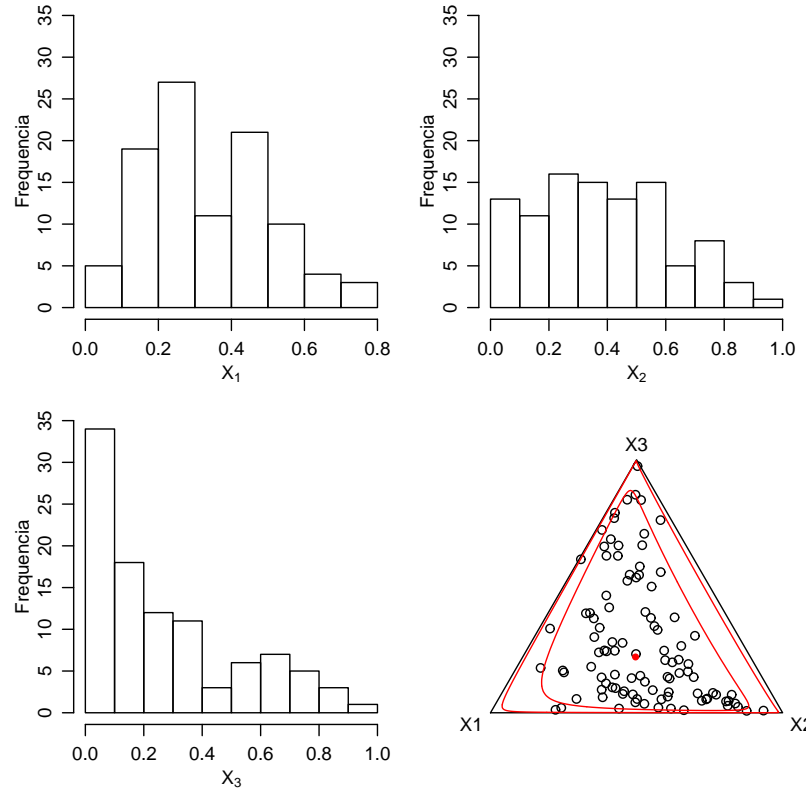


Figura 3.6: Distribuição de X_1 , X_2 e X_3 e diagrama ternário das composições para a segunda configuração.

FONTE: A autora (2010).

A Figura 3.7 mostra que os dados transformados também seguem uma distribuição gaussiana e a dispersão de pontos no diagrama de dispersão continua de forma linear crescente mas com um espalhamento maior em relação ao da configuração 1.

Na Tabela 3.2 tem-se os resultados do ajuste do modelo para a configuração 2 e o que se observa é que as estimativas permanecem com mesmos sinais; para as médias não são boas mas os respectivos intervalos de confiança contêm os valores atribuídos para os parâmetros. Somente o intervalo para σ_2 não conteve o valor do parâmetro apesar de estar próximo ao limite inferior. Os valores das log-verossimilhanças concentrada e otimizada foram iguais a $-330,3923$ e $298,63124$, respectivamente.

Na Figura 3.8 tem-se os mapas onde se observam poucas mudanças em relação aos mapas obtidos para a configuração 1 e a Figura 3.9 mostra que concordância já não é tão boa para X_1 quanto para X_2 e X_3 .

A última configuração deste estudo caracterizou-se pelo conjunto de parâmetros

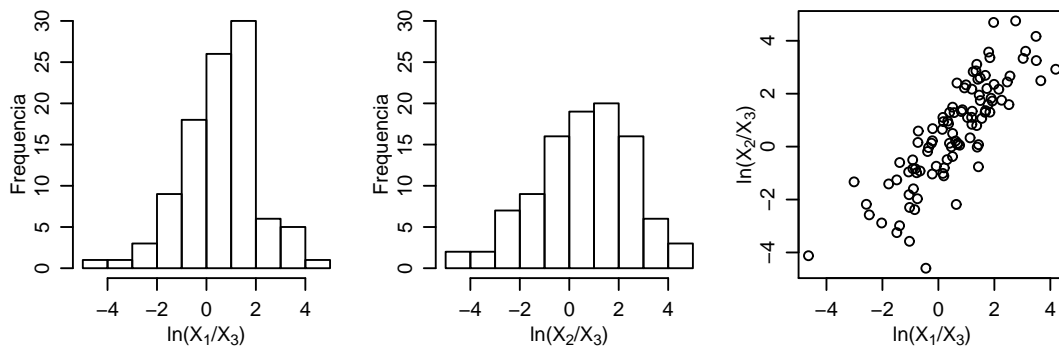


Figura 3.7: Distribuição das log-razão e correspondente diagrama de dispersão para a segunda configuração.

FONTE: A autora (2010).

Tabela 3.2: Estimativas, erros padrão e intervalos de confiança para a segunda configuração, pelo método delta via método de otimização “L-BFGS-B”.

Parâmetro	Valor	Estimativa	Erro Padrão	LI. Delta	LS. Delta
μ_1	1	0,3883751	1,4633944	-1,07501928	1,8517695
μ_2	1	0,2844116	1,7558488	-1,47143723	2,0402604
σ_1	1,2	1,2584047	0,1046900	1,15371472	1,3630947
σ_2	1,5	1,8313026	0,2648620	1,56644065	2,0961646
τ_1	0,9	0,9824042	0,1806521	0,80175207	1,1630563
τ_2	1	1,0052126	0,2132289	0,79198365	1,2184415
ϕ	0,6	0,5236357	0,4430230	0,08061266	0,9666587
ρ	0,5	0,5401549	0,1562457	0,38390917	0,6964006

FONTE: A autora (2010).

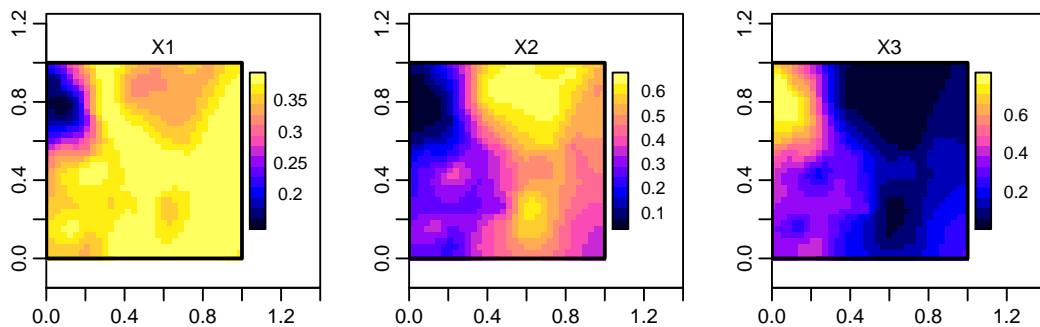


Figura 3.8: Mapas das porcentagens de X1, X2 e X3 para dados da configuração 2.

FONTE: A autora (2010).

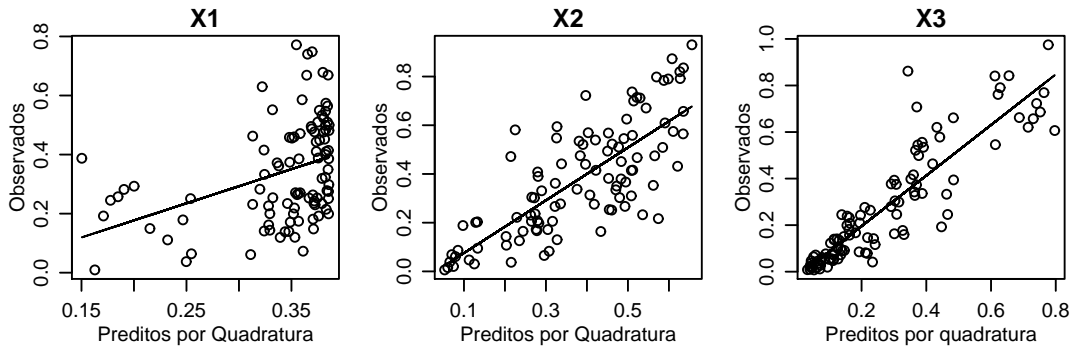


Figura 3.9: Valores observados versus preditos de X_1 , X_2 e X_3 para a configuração 2. FONTE: A autora (2010).

$(\mu_1, \mu_2, \sigma_1, \sigma_2, \tau_1, \tau_2, \phi, \rho) = (-0,2; -1; 0,45; 0,13; 0,3; 0,3; 0,6; 0,95)$. O baixo valor absoluto de μ_1 fez com que os dados se aproximassem do lado esquerdo do diagrama com proximidade a X_1 ; caso contrário, a aproximação seria para o lado direito. Também neste caso, o baixo valor de ϕ aproximou os pontos ao vértice X_3 enquanto o alto valor de ρ tornou-os concentrados novamente. Como nesta configuração $\sigma_1 > \sigma_2$, os pontos se afastaram do vértice X_2 e continuaram no interior do diagrama pelos valores iguais e baixos para τ_1 e τ_2 .

Como pode ser visto na Figura 3.10 tem-se baixos percentuais do componente X_2 , e as distribuições de X_1 , X_2 e X_3 se aproximam de uma distribuição gaussiana.

De acordo com a Figura 3.11, os dados transformados continuam apresentando distribuição gaussiana e os pontos no diagrama de dispersão também apresentam-se em forma linear crescente mas com espalhamento maior que o da configuração 1.

Pela Tabela 3.3, também neste caso, as estimativas conservaram os mesmos sinais dos valores verdadeiros dos parâmetros e observa-se que apenas os intervalos de confiança para as variâncias dos efeitos aleatórios Z_1 e Z_2 não conteve o valor atribuído para os parâmetros τ_1 e τ_2 . Neste caso, o valor da log-verossimilhança concentrada foi $-92,32910$ e o da log-verossimilhança otimizada foi $14,34055$.

Na Figura 3.12 tem-se os mapas referentes à configuração 3 podendo-se verificar uma mudança das predições de X_1 e X_2 em relação às outras configurações. Neste caso, a pior concordância entre os valores observados e os preditos ocorreu para o segundo componente (Figura 3.13).

Analisando o comportamento do modelo nas três configurações, de forma geral, observa-se que o componente X_1 apresentou-se o mais similar dentre as configurações com percentuais de médios para baixos. Com relação a X_2 , a configuração 2 evidenciou a maior variação nas porcentagens e as maiores porcentagens para X_3 apareceram na configuração 1.

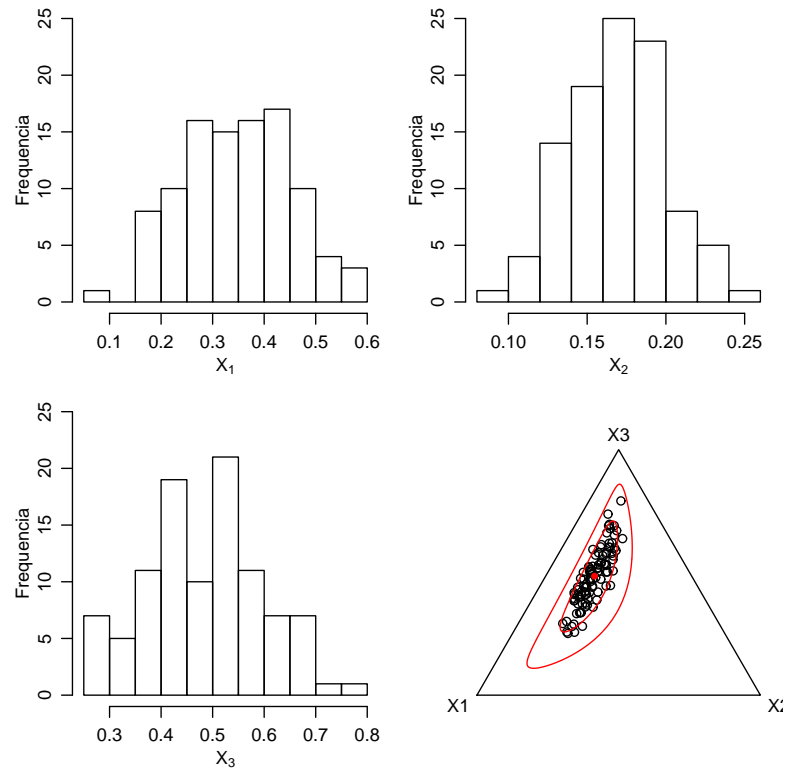


Figura 3.10: Distribuição de X_1 , X_2 e X_3 e diagrama ternário das composições para a terceira configuração.

FONTE: A autora (2010).

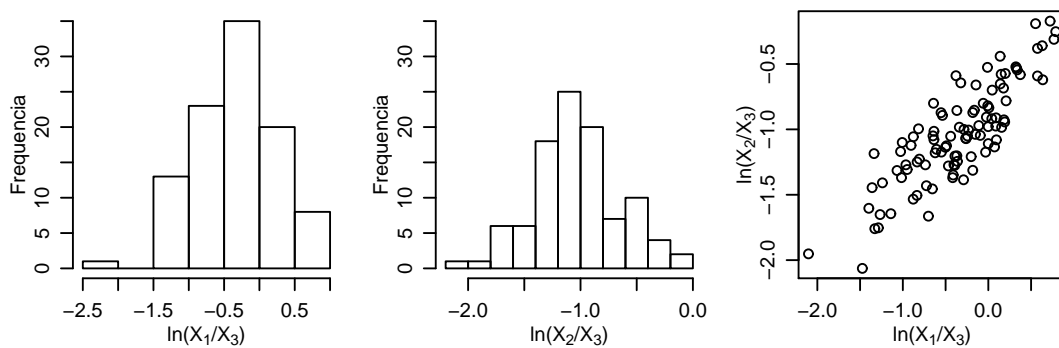


Figura 3.11: Distribuição das log-razão e correspondente diagrama de dispersão para a terceira configuração.

FONTE: A autora (2010).

Tabela 3.3: Estimativas, erros padrão e intervalos de confiança para a terceira configuração, pelo método delta via método de otimização “L-BFGS-B”.

Parâmetro	Valor	Estimativa	Erro Padrão	LI. Delta	LS. Delta
μ_1	-0,2	-0,5896071	0,69119506	-1,2808021	0,1015880
μ_2	-1	-1,1146150	0,40214445	-1,5167595	-0,7124706
σ_1	0,45	0,5305950	0,09559040	0,4350046	0,6261854
σ_2	0,13	0,1765225	0,05364894	0,1228735	0,2301714
τ_1	0,3	0,3488848	0,03599187	0,3128929	0,3848766
τ_2	0,3	0,3446375	0,03074784	0,3138897	0,3753853
ϕ	0,6	0,7403627	0,52040659	0,2199561	1,2607693
ρ	0,95	0,9383066	0,03744746	0,9008592	0,9757541

FONTE: A autora (2010).

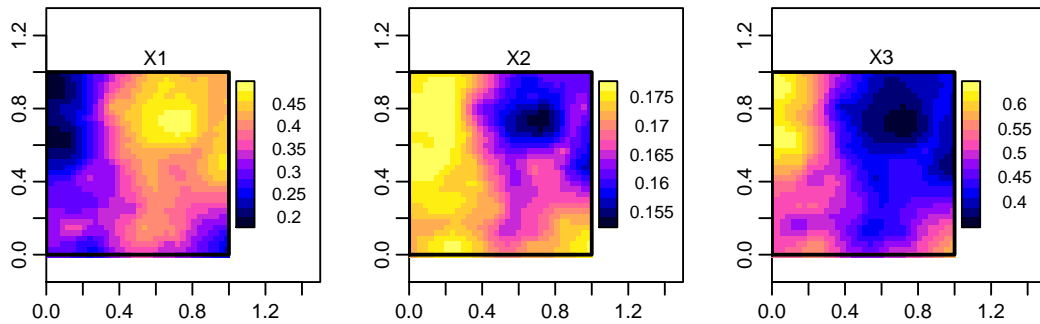


Figura 3.12: Mapas das porcentagens de X_1 , X_2 e X_3 para dados da configuração 3.
FONTE: A autora (2010).

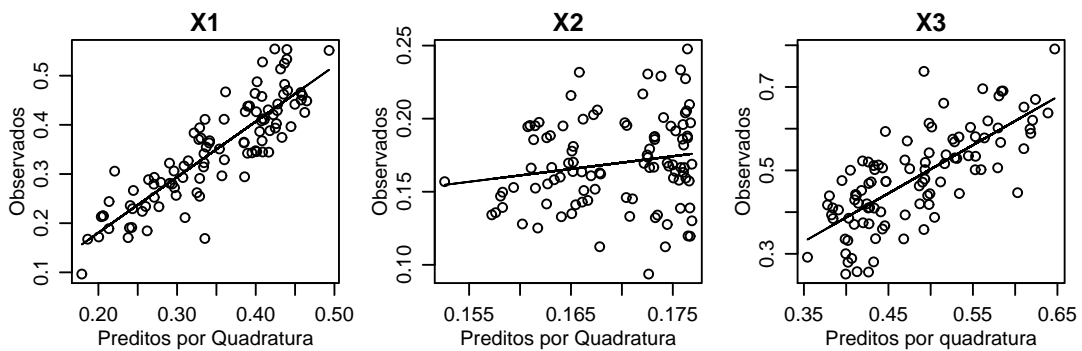


Figura 3.13: Valores observados versus preditos de X_1 , X_2 e X_3 para a configuração 3.
FONTE: A autora (2010).

A configuração 1 foi a que mais se aproximou de uma forma linear, foi a que apresentou as piores estimativas, enquanto que a configuração 3 que esperaria-se pior não resultou a pior em termos de qualidade de ajuste. As melhores estimativas foram obtidas com os dados da configuração 2. Os intervalos construídos via aproximação quadrática pelo método Delta não foram bons devido a baixa cobertura, principalmente para os parâmetros de variância.

Em relação aos mapas de valores preditos considerando-se as três configurações, ressalta-se que estes confirmam os resultados observados no diagrama ternário. As duas primeiras se mostram mais semelhantes por componente dado a dispersão dos pontos na parte central do triângulo. As mudanças observadas nos mapas correspondentes à configuração 3 se justificam pelo deslocamento da dispersão dos pontos para o lado esquerdo implicando em baixos valores de X_2 . As maiores concordâncias entre valores observados e preditos resultaram da configuração 1.

Observa-se que se a volta dos valores preditos para o simplex for feita por simulação, para cada localização, os passos a serem seguidos são:

- a. gerar dados bivariados de uma distribuição gaussiana padrão;
- b. fazer a decomposição Cholesky da matriz de covariância obtida por cokrigagem;
- c. aplicar a transformação linear, vetor de médias da cokrigagem somada ao produto dos dados bivariados obtidos no item (a) pela matriz Cholesky obtida no item (b);
- d. aplicar a transformação AGL nos valores esperados obtidos no item (c);
- e. construir os mapas de médias de predição para cada componente.

3.4 Intervalos de Cobertura para os Dados Simulados

Na avaliação da incerteza sobre os parâmetros do modelo fez-se um estudo de simulação para verificar a cobertura dos intervalos de confiança. O vetor de parâmetros considerado foi $\underline{\theta} = (\mu_1, \mu_2, \sigma_1, \sigma_2, \tau_1, \tau_2, \phi, \rho) = (50; 70; 2,5; 3,5; 2,5; 3,5; 0,6; 0,5)'$.

Inicialmente foram geradas 300 amostras de dados bivariados do modelo proposto sendo o efeito aleatório composicional gerado de uma normal bivariada com vetor de médias iguais a zero e matriz de covariância calculada com os devidos valores em $\underline{\theta}$. O campo aleatório gaussiano foi gerado com média zero, variância um, parâmetro de alcance ϕ dado em $\underline{\theta}$, e função de covariância exponencial.

Tabela 3.4: Parâmetros, média e intervalo de 95% de confiança das estimativas dos parâmetros, e cobertura (%) dos intervalos via método delta para 1000 simulações de amostras de tamanho 300.

Parâmetro	Valor	Média	LI.Clássico	LS.Clássico	Cobertura
μ_1	50	50,0418	49,9469	50,1367	69,2
μ_2	70	70,0634	69,9307	70,1962	63,4
σ_1	2,5	2,0259	1,9898	2,0621	5,4
σ_2	3,5	2,8411	2,7899	2,8924	24,2
τ_1	2,5	2,4660	2,4576	2,4743	94,6
τ_2	3,5	3,4570	3,4453	3,4687	94,1
ϕ	0,6	0,3869	0,3665	0,4073	43,1
ρ	0,5	0,4888	0,4856	0,4921	91,6

Em seguida, foi realizado o processo de estimação (algoritmo L-BFGS-B) obtendo-se as estimativas dos parâmetros e foram construídos os respectivos intervalos de confiança via método Delta. Este procedimento foi repetido num total de 1000 simulações.

Para cada parâmetro verificou-se quantos dos respectivos 1000 intervalos contiveram o seu verdadeiro valor e dividindo-se o resultado por 1000, obteve-se a cobertura dos intervalos. Pela Tabela 3.4, observa-se que a cobertura dos intervalos foi baixa exceto para τ_1 , τ_2 e ρ , o que sugere a construção por verossimilhança perfilhada ou inferência bayesiana. Tem-se também, as médias das 1000 estimativas com respectivos intervalos de 95% confiança construídos pelo método clássico.

Nas Figuras 3.14 e 3.15 pode-se observar as distribuições e os boxplots das estimativas.

Com estes estudos de simulação foi possível verificar a viabilidade do modelo e a partir daí passou-se a análise de um conjunto de dados reais a ser apresentado em seguida.

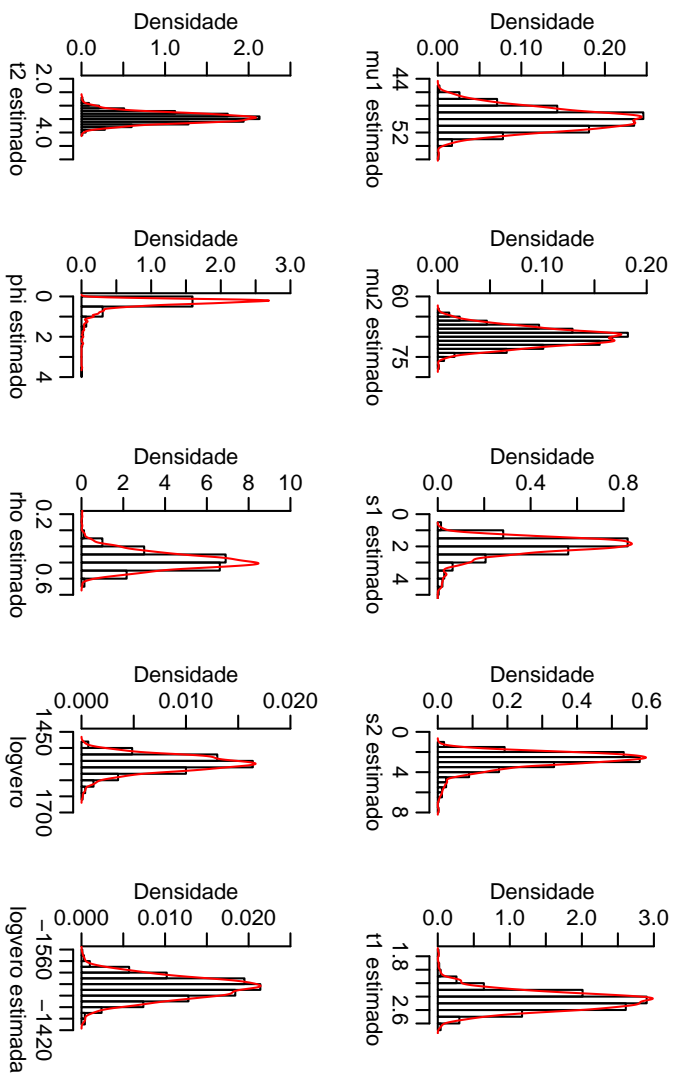


Figura 3.14: Distribuições das estimativas dos parâmetros para 1000 simulações e amostras de tamanho 300.

FONTE: A autora (2010).

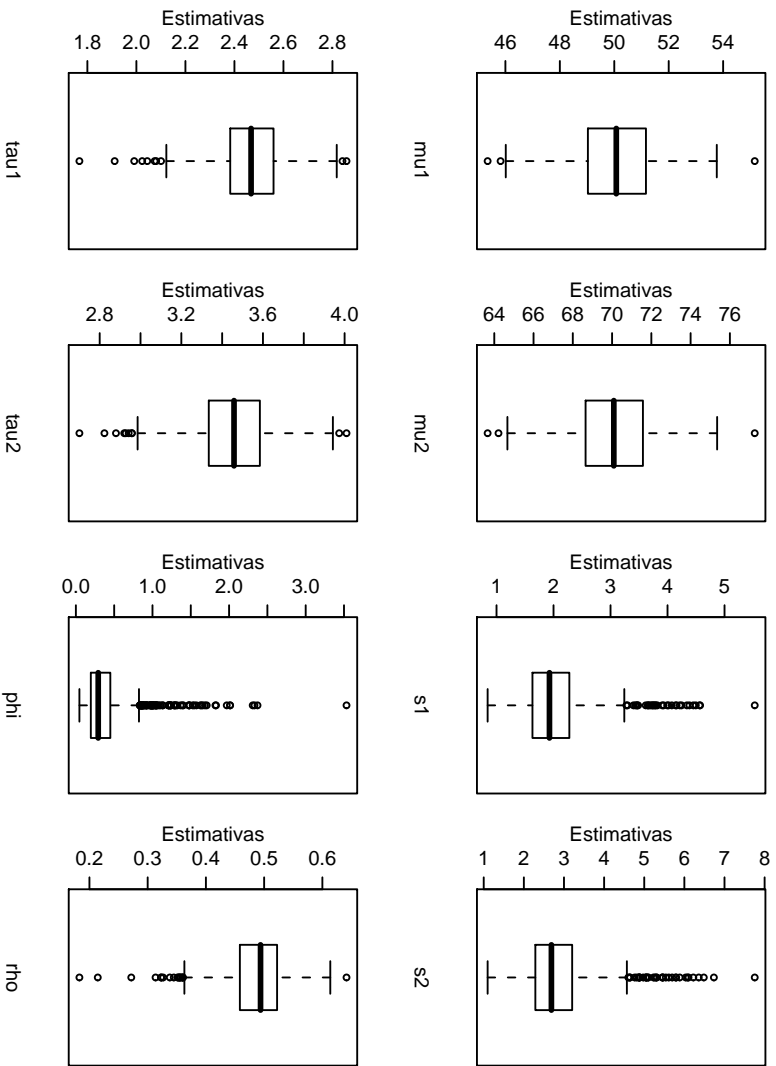


Figura 3.15: Boxplot das estimativas dos parâmetros para 1000 simulações e amostras de tamanho 300.

FONTE: A autora (2010).

4 RESULTADOS

Analisou-se neste estudo de caso um conjunto de dados obtidos a partir do trabalho de Gonçalves (1997) conduzido em uma área irrigada por pivô-central na Fazenda Areão, Figura 4.1, pertencente ao campus da Escola Superior de Agricultura - “Luiz de Queiroz” (ESALQ-USP). Nela foi demarcado um quadrante na porção mais elevada (topo da encosta) no qual foram obtidas 76 amostras de solo na profundidade entre 0 e 0,20m em uma malha regular quadrada de amostragem, de lado igual a 20 metros. Em cada amostra foram medidos os valores das frações granulométricas, de areia, silte e argila.

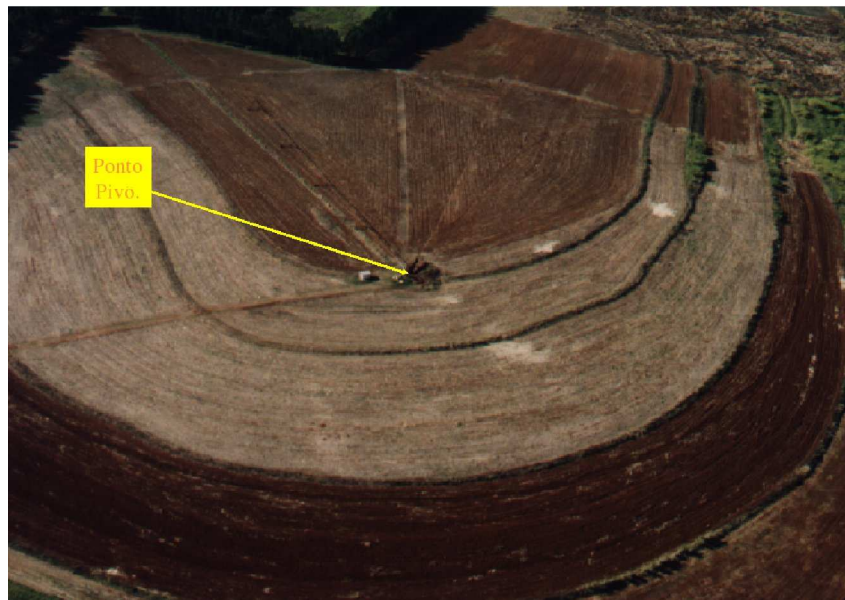


Figura 4.1: Foto aérea do campo experimental de irrigação da ESALQ-USP com área de estudo correspondente ao quadrante irrigado por um sistema pivô-central.
FONTE: Gonçalves (1997).

4.1 ANÁLISE DE FRAÇÕES GRANULOMÉTRICAS DE UM SOLO

Nos gráficos da Figura 4.2 tem-se representado o quadrante da área com as localizações amostradas. Círculos, triângulos, diagonais verticais e cruzadas correspondem aos quartis das distribuições dos valores das frações granulométricas cujos valores se encontram na Tabela 4.1. A coordenada mínima foi igual a (0,0) e máxima igual a (180,180). A disposição no diagrama

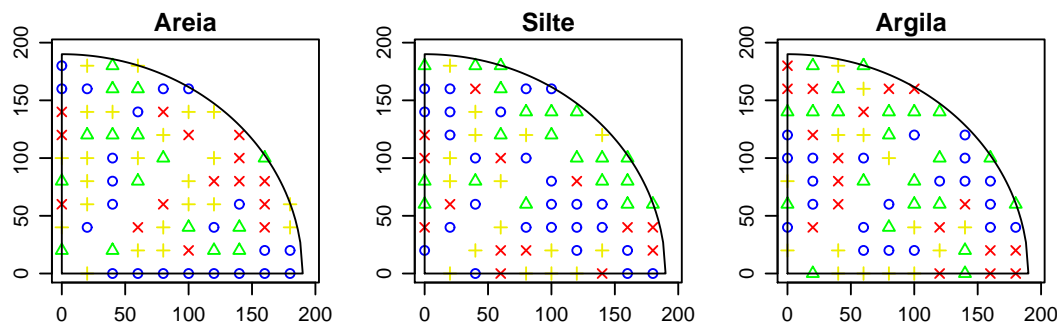


Figura 4.2: Primeiro (círculos), segundo (triângulos), terceiro (diagonais verticais) e quarto (diagonais cruzadas) quartil da percentagem de areia (à esquerda), silte (meio) e argila (à direita) nas localizações amostrais.

FONTE: A autora (2010).

de classificação textural pode ser visto na Figura 4.3. Pode-se observar ainda que, segundo esta classificação, o solo varia de muito argiloso para franco-argiloso com a maior frequência em argiloso.

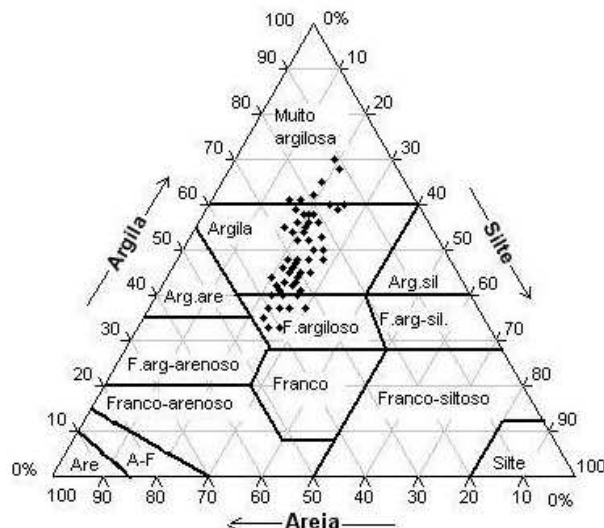


Figura 4.3: Diagrama de classificação textural.

FONTE: A autora (2010).

Os resultados apresentados na Tabela 4.1 revelam que argila apresentou os maiores valores enquanto silte os menores, exceto para o valor mínimo, já areia apresentou o maior

Tabela 4.1: Estatísticas descritivas dos percentuais de areia, silte e argila.

Estatística	Areia	Silte	Argila
Mínimo	11	15	33
Quartil 1	24	21	42
Mediana	29,50	23	48
Média	28,32	23	48,68
Quartil 3	33	25	55
Máximo	42	30	70
Desvio padrão	7,04	3,03	8,57
Variância	49,5	9,17	73,4
Coef. de Variação	0,25	0,13	0,16

FONTE: A autora (2010).

coeficiente de variação indicando maior variabilidade. Em acréscimo às informações reveladas na Figura 4.2, tem-se os histogramas na Figura 4.4 sugerindo uma distribuição aproximadamente gaussiana dos conteúdos. Para a construção do diagrama ternário, considerou-se a classe “acom” do pacote *compositions* em que os dados são analisados na geometria relativa. Observa-se que nas composições amostradas, o silte é o componente que se apresenta em menor percentagem e uma região de confiança de 4-desvios padrão contempla todas as amostras. O ponto vermelho no diagrama representa o centro da distribuição, o fechamento das médias geométricas das composições conforme Equação (2.34) e é igual ao vetor (0,2786;0,2324;0,4890). As variâncias das log-razões encontram-se na seguinte matriz variação:

$$\text{Matriz.Varição} = \begin{bmatrix} 0 & 0,0727 & 0,2012 \\ 0,0727 & 0 & 0,0799 \\ 0,2012 & 0,0799 & 0 \end{bmatrix}.$$

Observa-se que a maior variabilidade ocorreu para $\ln(\text{Areia}/\text{Argila}) = \ln(\text{Argila}/\text{Areia}) = 0,2012$ enquanto $\ln(\text{Areia}/\text{Silte}) = 0,0727$ e $\ln(\text{Silte}/\text{Argila}) = 0,0799$. A variabilidade composicional dada pela Equação (2.35) é igual a:

$$\begin{aligned} \text{Totvar}(\underline{X}) &= \frac{1}{B} \sum_{i < j} \text{Var} \left(\ln \left(\frac{X_i}{X_j} \right) \right) \\ &= \frac{1}{B} \left[\sum_{\substack{i=1 \\ j=2,3}} \text{Var} \left(\ln \left(\frac{X_i}{X_j} \right) \right) + \sum_{\substack{i=2 \\ j=3}} \text{Var} \left(\ln \left(\frac{X_i}{X_j} \right) \right) \right] \\ &= \frac{1}{3} \left[\text{Var} \left(\ln \left(\frac{X_1}{X_2} \right) \right) + \text{Var} \left(\ln \left(\frac{X_1}{X_3} \right) \right) + \text{Var} \left(\ln \left(\frac{X_2}{X_3} \right) \right) \right] \\ &= \frac{1}{3} (0,0727 + 0,2012 + 0,0799) \\ &= 0,1179. \end{aligned}$$

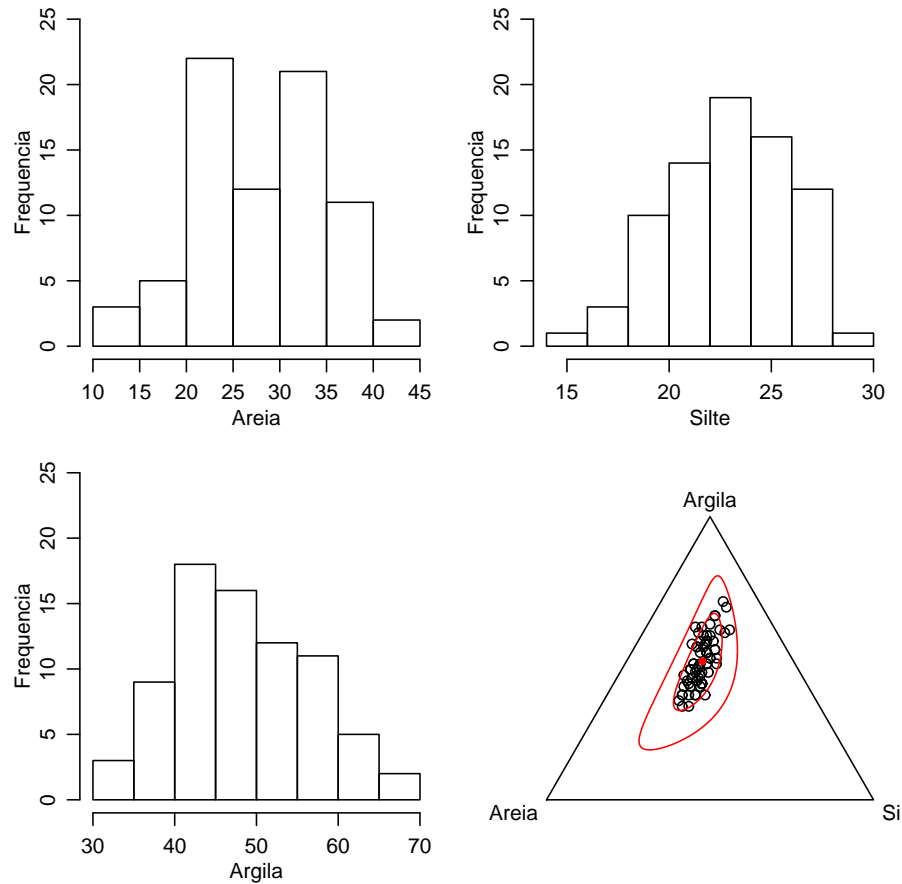


Figura 4.4: Distribuição de areia, silte e argila e diagrama ternário das composições. FONTE: A autora (2010).

Ao fazer a transformação ALR nos dados originais obteve-se as variáveis $Y_1 = \ln\left(\frac{\text{Areia}}{\text{Argila}}\right)$, $Y_2 = \ln\left(\frac{\text{Silte}}{\text{Argila}}\right)$ cujas distribuições (Figura 4.5) apresentam-se aproximadamente gaussianas e o diagrama de dispersão evidencia uma possível correlação linear positiva.

Dentre os métodos de otimização testados adotou-se o método L-BFGS-B, por não apresentar problemas de convergência e, para este método, as estimativas obtidas para os parâmetros do modelo proposto são dados na Tabela 4.2.

O valor da log-verossimilhança concentrada foi de $-42,5541$, calculado considerando-se as estimativas obtidas através da amostra com ρ dado pelo coeficiente de correlação de Pearson e $\phi = \min + 0.2(\max - \min)$, onde “min” e “max” são, respectivamente, a menor e maior distância entre duas localizações e o valor da log-verossimilhança otimizada foi $0,8292$.

A Figura 4.6 apresenta o mapa de predição para os três componentes considerando-se a volta dos valores preditos em \mathbb{R}^2 para o simplex S^3 por aproximação de Gauss-Hermite (gráficos (A), (B) e (C)) e por simulação (gráficos (D), (E) e (F)). Foram realizadas 1000 simulações e

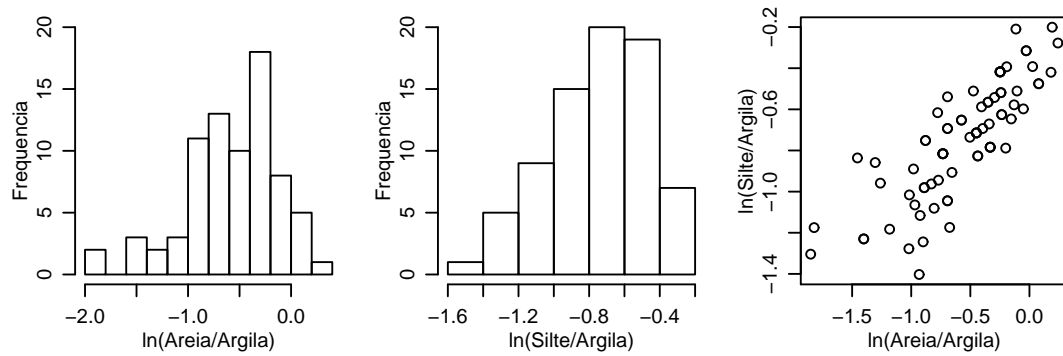


Figura 4.5: Distribuição das log-razão e correspondente diagrama de dispersão.
FONTE: A autora (2010).

Tabela 4.2: Estimativas, erros padrão e intervalos de confiança pelo método Delta via método de otimização “L-BFGS-B”.

Parâmetros	Estimativas	Erro Padrão	LI. Delta	LS. Delta
μ_1	-0,7598748	0,44029631	-1,20017113	-0,3195785
μ_2	-0,7941109	0,22857184	-1,02268277	-0,5655391
σ_1	0,4504836	0,12080170	0,32968188	0,5712853
σ_2	0,1152904	0,05153622	0,06375416	0,1668266
τ_1	0,2870196	0,04510491	0,24191470	0,3321245
τ_2	0,2669836	0,03142454	0,23555903	0,2984081
ϕ	66,9084688	42,55386352	24,35460526	109,4623323
ρ	0,9544251	0,06803953	0,88638554	1,0224646

FONTE: A autora (2010).

no caso da aproximação de Gauss-Hermite, a ordem de quadratura foi igual a 7. A grade foi constituída de 2500 pontos.

Dados os resultados mostrados na Figura 4.6 conclui-se que os mapas de predição pelos dois métodos são muito similares e os componentes areia e argila se complementam na área de estudo. Uma comparação dos métodos é apresentada na Figura 4.7.

Com os resultados obtidos por simulação, da transformação de volta dos valores preditos para o simplex, calculou-se os percentuais máximos e mínimos das 1000 simulações em cada localização, para cada componente e os mapas podem ser vistos na Figura 4.8. A Tabela 4.3 apresenta as médias dos percentuais máximos e mínimos encontrados em cada localização.

Tabela 4.3: Médias dos percentuais máximos e mínimos de areia, silte e argila obtidos por simulação.

Componente	Máximo	Mínimo
Areia	0,70773023	0,04044566
Silte	0,88974020	0,08607770
Argila	0,46024885	0,05946616

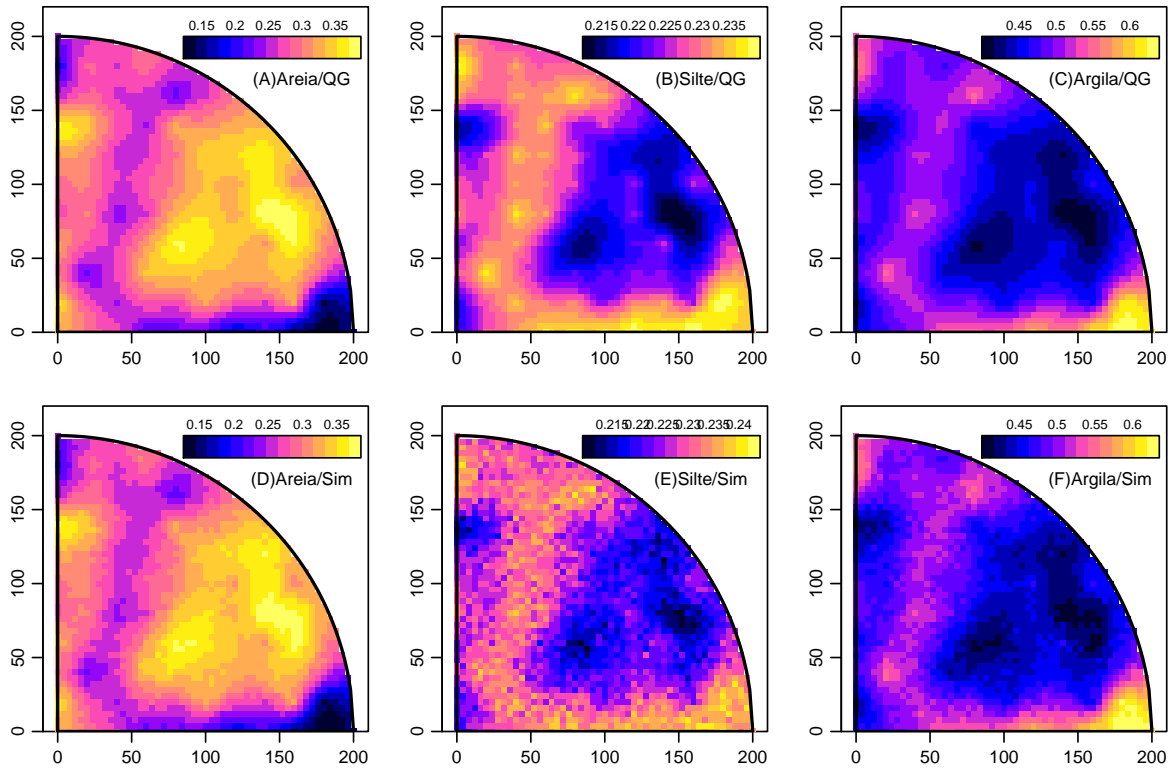


Figura 4.6: Mapas das percentagens de areia, silte e argila obtidos por quadratura de Gauss-Hermite (A-C) e por simulação (D-F).
 FONTE: A autora (2010).

Como visto na Subseção 3.3, os estudos de simulação mostraram que na avaliação da incerteza sobre os parâmetros do modelo, os intervalos de confiança não são bons e uma alternativa foi a construção de intervalos baseados na verossimilhança perfilhada. Para exemplificar, os gráficos da log-verossimilhança perfilhada para os parâmetros ϕ e ρ são apresentados na Figura 4.9.

Os intervalos de confiança são obtidos a partir da deviance conforme Subseção 2.1.3 Equação (2.17). Os gráficos da deviance para ϕ e ρ são apresentados na Figura 4.11. Primeiramente, observa-se que o nível de confiança máximo para a obtenção de um intervalo para ϕ é de 83,5%. Mas o limite superior ultrapassa o limite da área em estudo que é de 180m. Isso significa que o modelo foi difícil de ser identificado pois a informação contida nos dados não foi suficiente, ou seja, os dados o identificam mas com muita incerteza levando a resultados possivelmente instáveis. Para se obter um intervalo cujos limites pertençam à área, isto é, para se fazer inferências razoáveis para ϕ com as informações disponíveis é necessário um nível de 56,6% de confiança e os resultados são dados nas equações em (4.1).

Com relação a ρ , o espaço paramétrico impõe um truncamento no intervalo. O limite inferior é 95% menos plausível do que o máximo, $\hat{\rho}$, mas não é possível ter a mesma plausibilidade

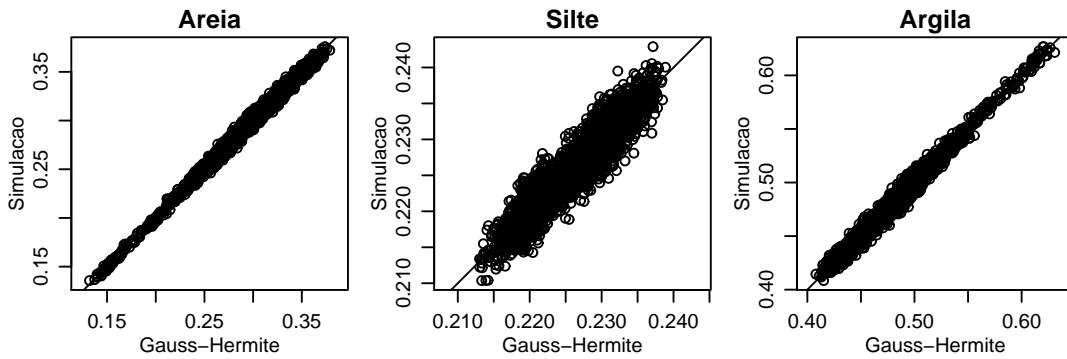


Figura 4.7: Porcentagens de areia, silte e argila obtidos por quadratura de Gauss-Hermite versus simulação.

FONTE: A autora (2010).

para o limite superior.

$$\begin{aligned}
 IC_{\phi} &= \{\phi; D(\phi) - D(\hat{\phi})\} < \chi_{1;0,566}^2 & \text{ou} & 37,19 < \phi < 178,90 \\
 &= \{\phi; D(\phi) - D(\hat{\phi})\} < \chi_{1;0,835}^2 & \text{ou} & 26,57 < \phi < 522,53 \\
 IC_{\rho} &= \{\rho; D(\rho) - D(\hat{\rho})\} < \chi_{1;0,95}^2 & \text{ou} & 0,84 < \rho < 0,98.
 \end{aligned} \tag{4.1}$$

4.2 INFERÊNCIA BAYESIANA DE DADOS COM- POSICIONAIS

Os resultados através de métodos bayesianos foram obtidos utilizando-se o algoritmo Metropolis-Hastings. Foram executadas 12000 simulações, o período de aquecimento foi de 1000 simulações usando um salto igual a 10 e uma sintonia (*tune*) igual a (0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5; 0,5). Com isto as cadeias foram formadas por 1200 simulações.

As estimativas dos parâmetros bem como os respectivos intervalos de credibilidade são apresentados na Tabela 4.4.

Ao comparar as estimativas na Tabela 4.4 com as apresentadas na Tabela 4.2, observa-se que são aproximadamente iguais com a maior discrepância para ϕ com intervalo de credibilidade mais preciso.

Os mapas dos valores preditos de areia, silte e argila por inferência bayesiana são apresentados na Figura 4.12, cujo padrão geral assemelha-se aos obtidos por inferência clássica (Figura 4.6).

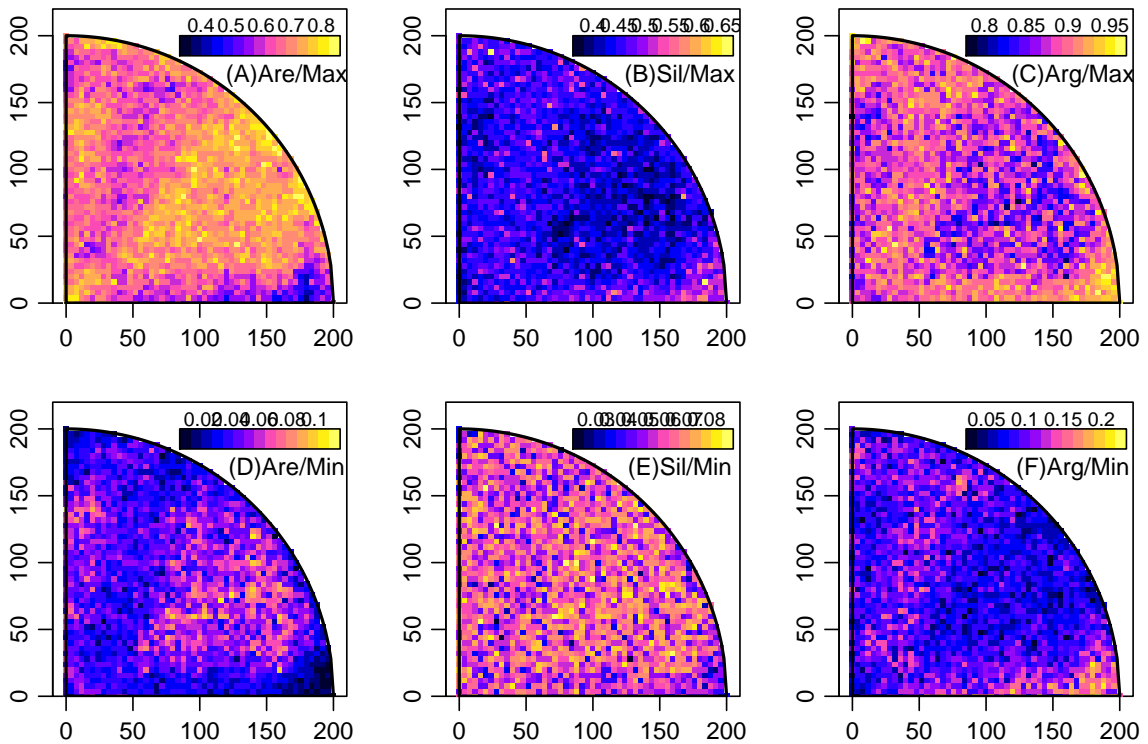


Figura 4.8: Valores máximos e mínimos obtidos por simulação.

4.2.1 Avaliação da Convergência

Numa avaliação dos resultados obtidos por inferência bayesiana apresenta-se a seguir diagnósticos de convergência das cadeias simuladas.

Na Tabela 4.5 tem-se as correlações *a posteriori* dos parâmetros estimados dentro da cadeia. Observa-se que o acréscimo das correlações para μ_1 e μ_2 até a defasagem 50 são um indicativo de baixa mistura e possivelmente ausência de convergência individual. Para os outros parâmetros o decréscimo até a defasagem 50 mostra uma característica típica de convergência.

Quanto aos gráficos da Figura 4.13, as autocorrelações decrescem com o aumento da defasagem e mesmo com valores maiores ocorrendo para a *posteriori* de ρ , estes parecem não indicar problemas de convergência.

As altas correlações entre os parâmetros da cadeia mostradas na Tabela 4.6 tendem a indicar baixa convergência. Pode-se observar na Figura 4.14 que as cadeias parecem razoavelmente estáveis como na Figura 4.13.

Os gráficos da Figura 4.15 são densidades estimadas de cada parâmetro cuja multimodalidade seria um indicativo de não convergência e isto não aparece nestes casos. Um diagnóstico de convergência adicional é comparar os intervalos de mais alta densidade *a posteriori*, Tabela 4.7,

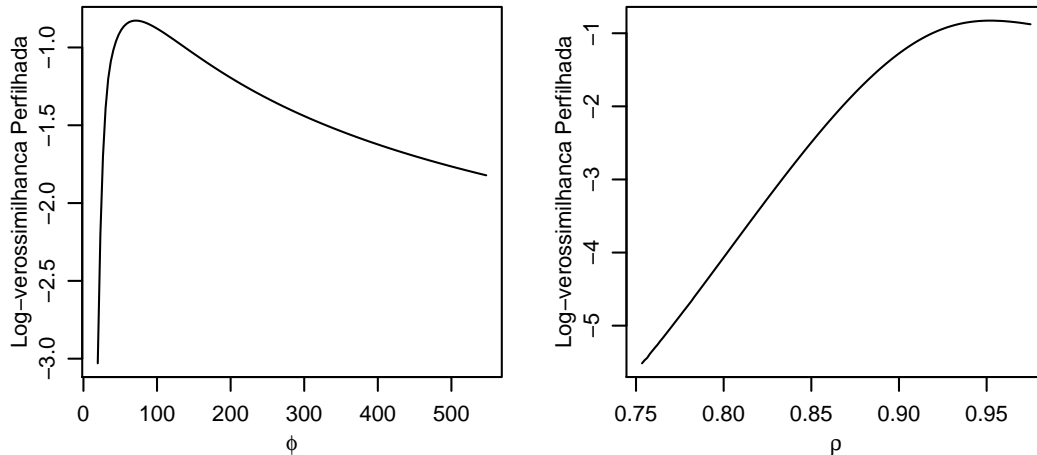


Figura 4.9: Log-verossimilhanças perfilhadas para ϕ e ρ .
 FONTE: A autora (2010).

Tabela 4.4: Esperança das 1200 simulações da distribuição à *posteriori* de θ e intervalos de 95% de credibilidade obtidos por inferência bayesiana considerando-se 12000 simulações, *burn-in*= 1000 e salto= 10.

Parâmetros	Estimativas	LI. Delta	LS. Delta
μ_1	-0,7643910	-1,1951539	-0,3326118
μ_2	-0,7946658	-0,9350368	-0,6740090
σ_1	0,4427920	0,3409507	0,5573306
σ_2	0,1126342	0,0592708	0,1804150
τ_1	0,2871070	0,2220835	0,3567449
τ_2	0,2658107	0,2255757	0,3099205
ϕ	65,727842	51,5340474	84,4576982
ρ	0,9314342	0,8252095	0,9924404

FONTE: A autora (2010).

que são construídos com todos os valores da cadeia e o centro de cada distribuição à *posteriori*. Os intervalos estão representados por segmentos de reta acima dos *rug plot*. Grandes discrepâncias são indicativos que os últimos valores não são típicos da amostra completa. Novamente, parece que os dados não indicam total falta de convergência.

Foi realizado o teste de Geweke apresentado na Subseção 2.3.2, considerando uma fração de 90% do início da cadeia e 50% do fim da cadeia. Pode-se observar que todos os valores apresentados na Tabela 4.8 são inferiores a 2 o que implica na não preocupação com a convergência das cadeias.

A Tabela 4.9 apresenta as estimativas pontuais e quantis de 95% de credibilidade para os fatores de redução de escala potencial do diagnóstico de convergência de sequências múltiplas de Gelman e Rubin. Como os valores resultaram inferiores a 1.2, pode-se concluir pela convergência

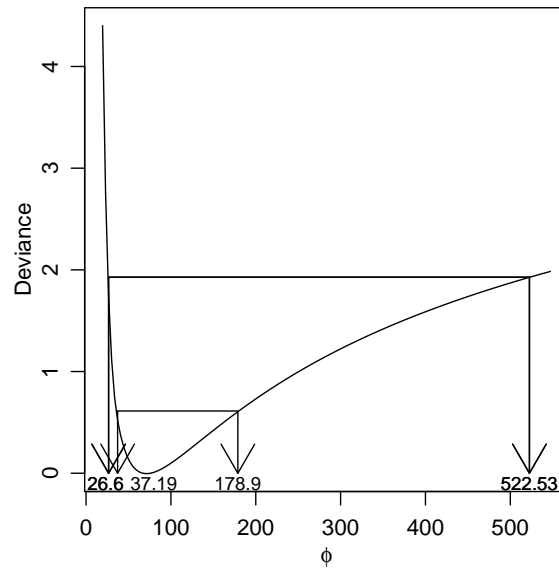


Figura 4.10: Deviance para ϕ com linhas de corte de 83,5% e 60% de confiança.
 FONTE: A autora (2010).

Tabela 4.5: Autocorrelações das posteriores dos parâmetros.

Def.	μ_1	μ_2	σ_1	σ_2	τ_1	τ_2	ϕ	ρ
0	1	1	1	1	1	1	1	1
1	-0,0141	0,0404	0,4519	0,5281	0,5318	0,3816	0,0793	0,7653
5	-0,0529	-0,0384	0,1004	0,0829	0,2145	0,2445	0,0233	0,3823
10	-0,0212	-0,0199	-0,0197	0,0306	0,0656	0,1019	0,0271	0,1506
50	0,0437	0,0413	0,0267	-0,0389	-0,0339	0,0137	-0,0106	-0,0733

FONTE: A autora (2010).

das cadeias. Os resultados estão representados graficamente na Figura 4.17.

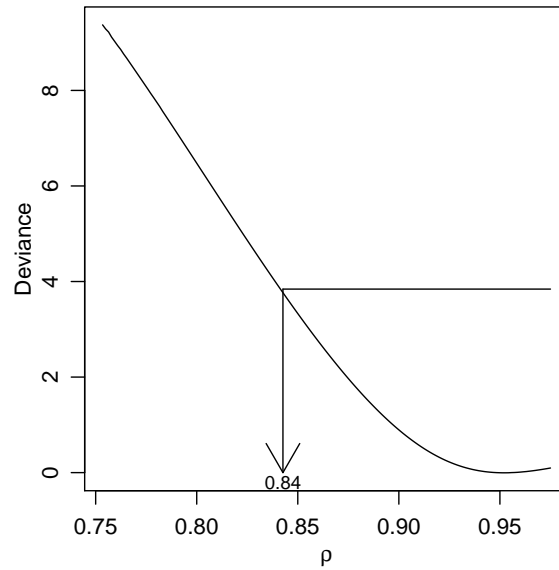


Figura 4.11: Deviance para ρ com linhas de corte de 95% de confiança.
FONTE: A autora (2010).

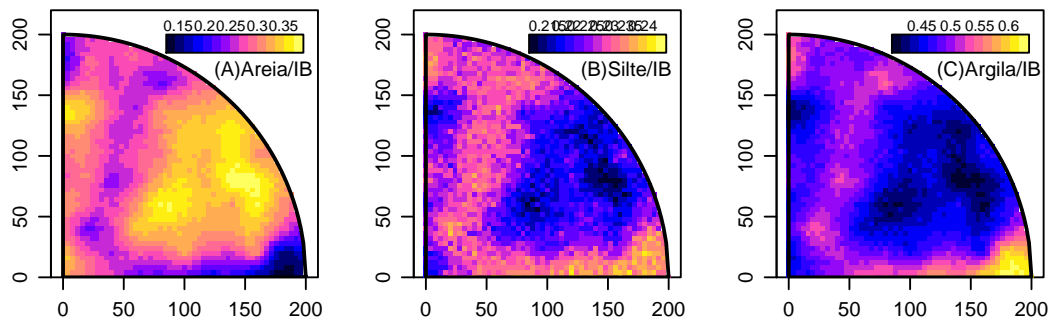


Figura 4.12: Valores preditos de areia (à esquerda), silte (centro) e argila (à direita) por inferência bayesiana.
FONTE: A autora (2010).

Tabela 4.6: Correlações das posteriores dos parâmetros.

Posteriori	μ_1	μ_2	σ_1	σ_2	τ_1	τ_2	ϕ	ρ
μ_1	1							
μ_2	0,8992	1						
σ_1	-0,0690	-0,1158	1					
σ_2	-0,0330	-0,1835	0,6905	1				
τ_1	0,0297	0,0686	-0,1959	-0,2540	1			
τ_2	0,0028	0,0313	0,1158	-0,0101	0,6101	1		
ϕ	-0,0995	-0,0587	0,2063	0,0456	0,1430	0,0797	1	
ρ	-0,0403	0,0105	0,2838	-0,0814	0,00321	0,3599	-0,1521	1

FONTE: A autora (2010).

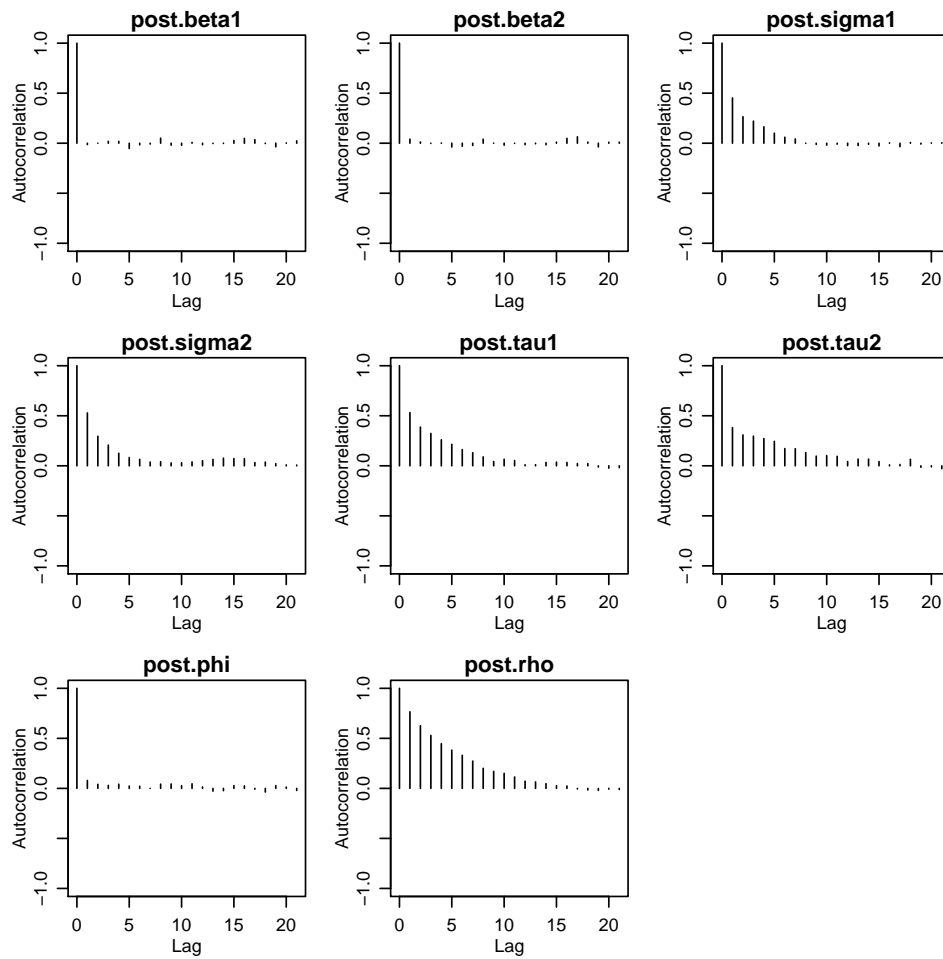


Figura 4.13: Autocorrelações dos parâmetros à *posteriori*.
 FONTE: A autora (2010).

Tabela 4.7: Limites dos intervalos de 95% de credibilidade de alta densidade à posteriori.

Posteriori	μ_1	μ_2	σ_1	σ_2	τ_1	τ_2	ϕ	ρ
LI	-1,1743	-0,9198	0,3413	0,0547	0,2226	0,2222	51,3383	0,8529
LS	-0,3170	-0,6681	0,5581	0,1747	0,3572	0,3048	83,7790	0,9994

FONTE: A autora (2010).

Tabela 4.8: Valores de z_G do teste de Geweke para as posteriors dos parâmetros.

Posteriori	μ_1	μ_2	σ_1	σ_2	τ_1	τ_2	ϕ	ρ
z_G	-1,0748	-0,4615	-0,3654	-0,6429	0,2827	0,3602	0,5379	0,2413

FONTE: A autora (2010).

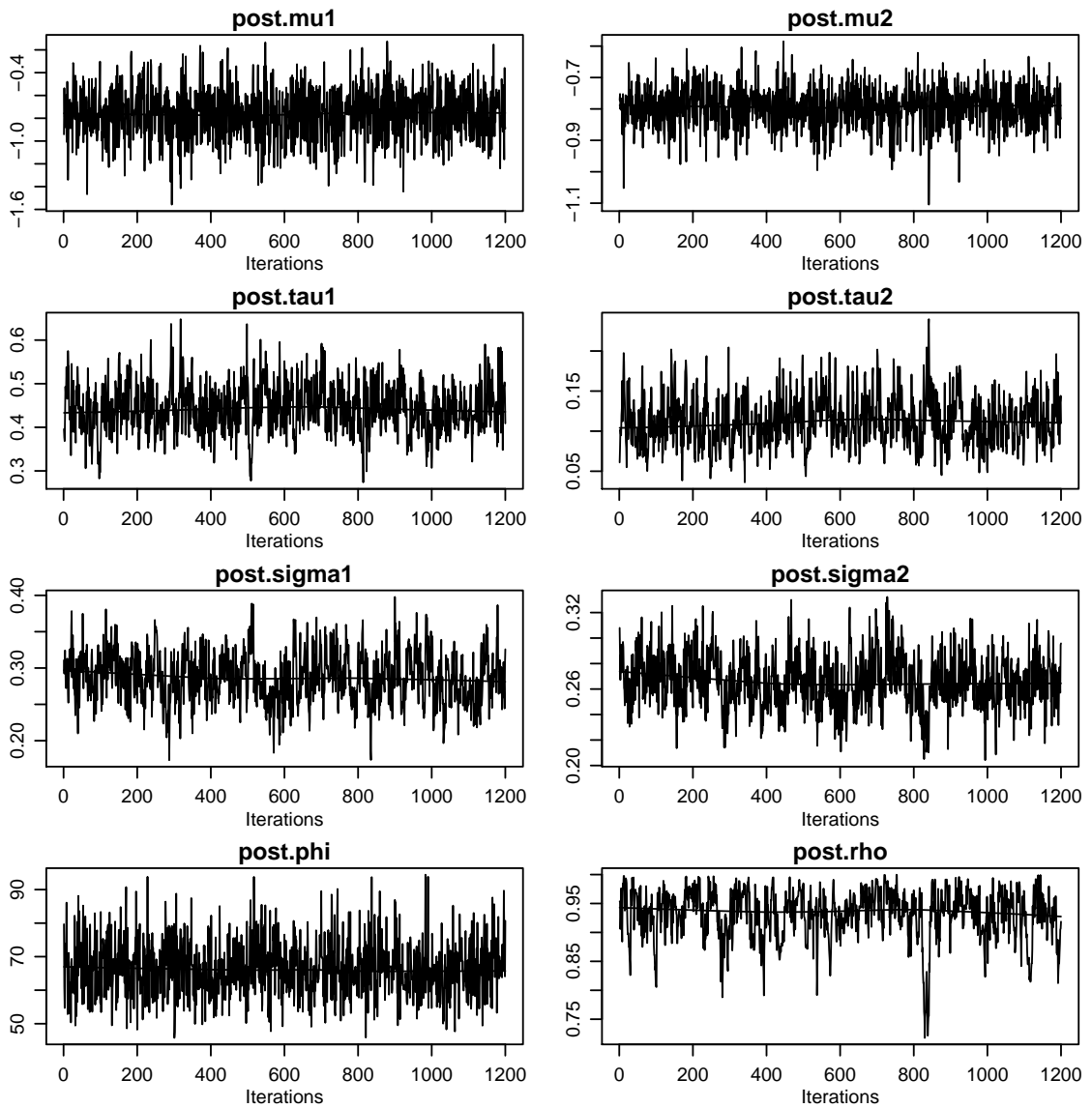


Figura 4.14: Trajetória das posteriores dos parâmetros.
 FONTE: A autora (2010).

Tabela 4.9: Fatores de redução de escala potencial do teste de Gelman e Rubin.

Posteriori	μ_1	μ_2	σ_1	σ_2	τ_1	τ_2	ϕ	ρ
Est.Pontual	1	1,01	1	1	1	1	1	1,01
Quantil 95%	1,02	1,02	1	1	1	1	1	1,01

FONTE: A autora (2010).

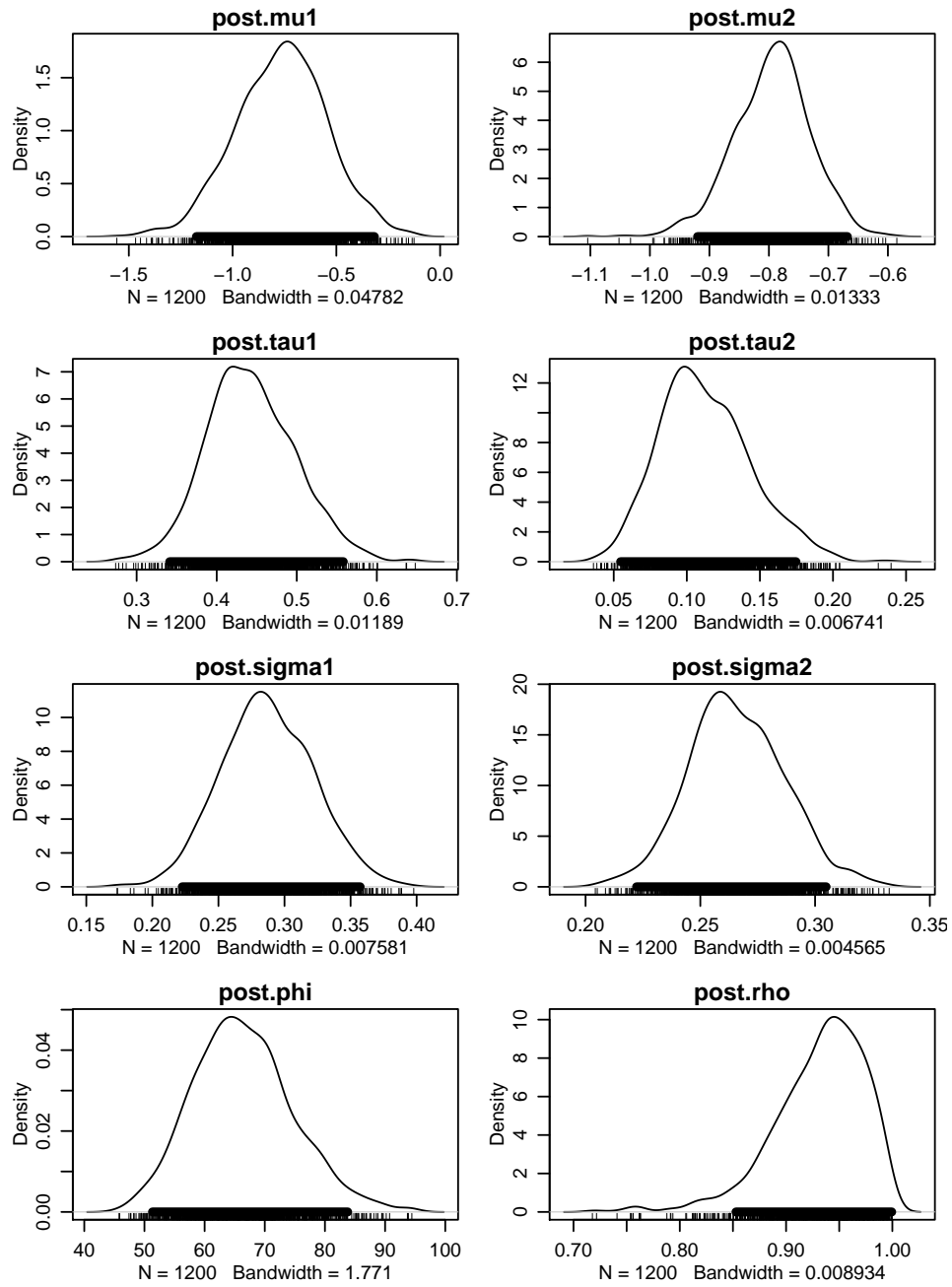


Figura 4.15: Densidades estimadas para as posteriores dos parâmetros com respectivos intervalos de alta densidade à posteriori a 95%.
 FONTE: A autora (2010).

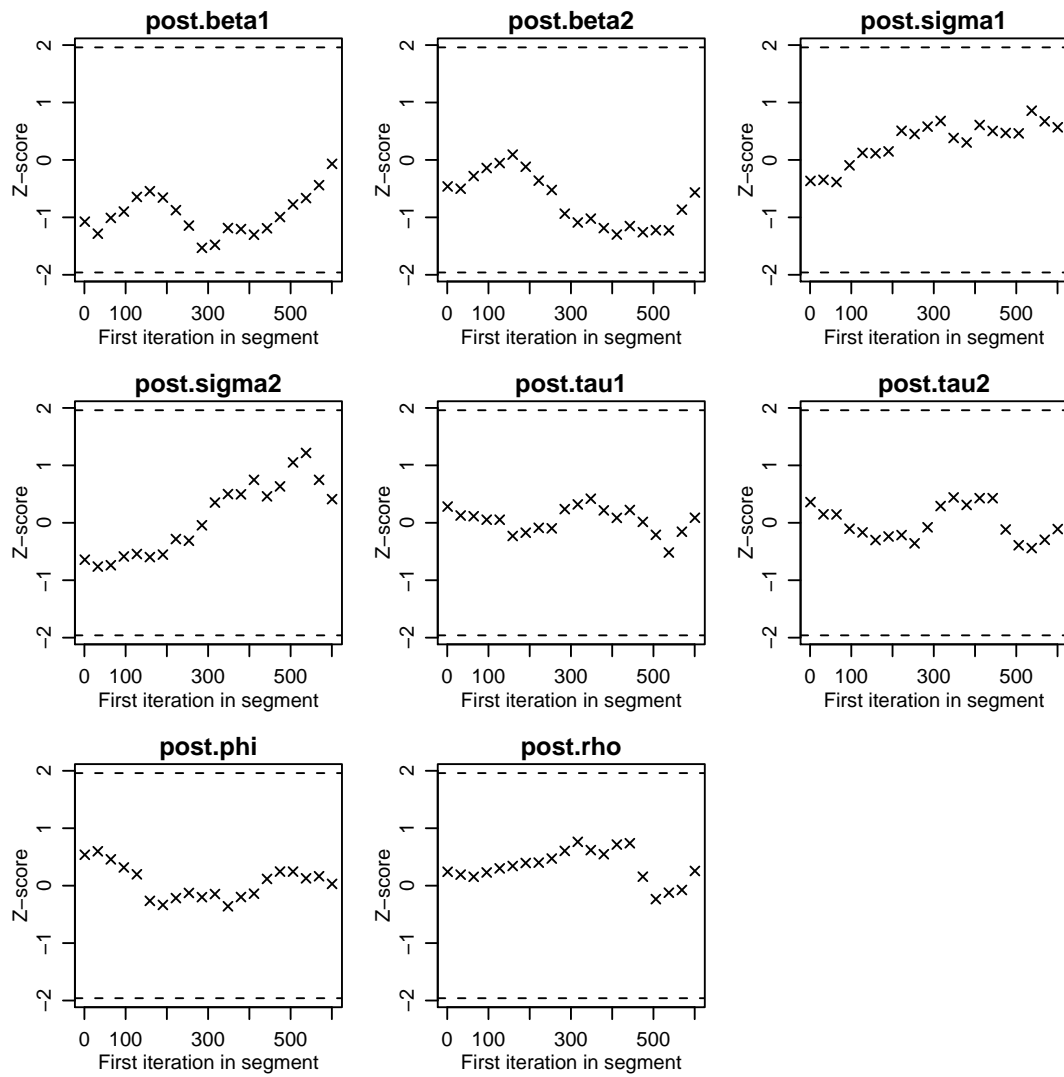


Figura 4.16: Diagnóstico de convergência para o Teste de Geweke.
 FONTE: A autora (2010).

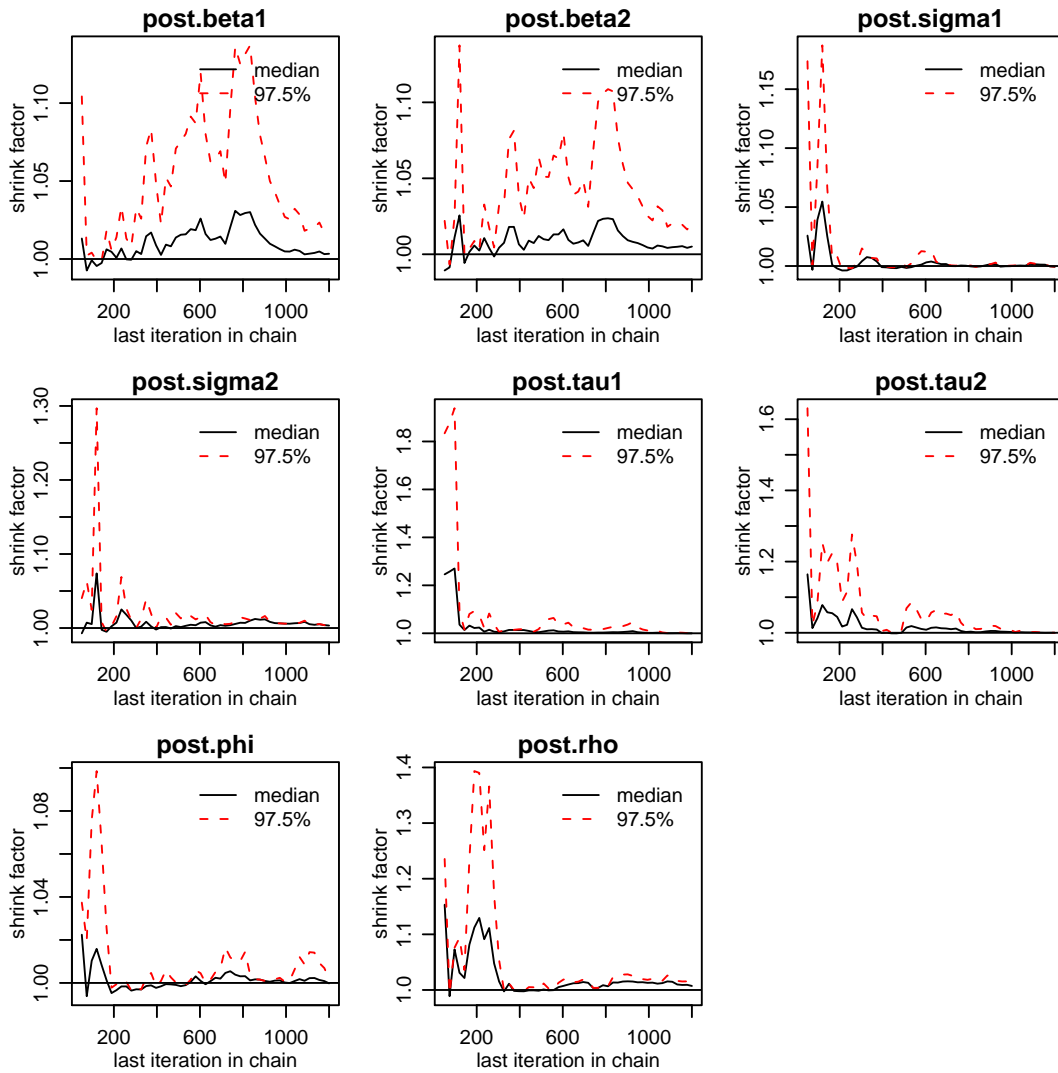


Figura 4.17: Diagnóstico de convergência para o Teste de Gelman e Rubin.
 FONTE: A autora (2010).

5 CONCLUSÕES E SUGESTÕES DE TRABALHOS FUTUROS

O modelo proposto foi capaz de capturar variações espaciais, induzidas pelas composições e não estruturadas, e os procedimentos adotados permitiram a construção de mapas de areia, silte e argila por uma metodologia que implicitamente garante a restrição de que as frações somem 1, não só nos pontos observados como nos pontos preditos.

A metodologia desenvolvida foi baseada na teoria de verossimilhança e a declaração explícita do modelo permitiu que fossem feitas inferências sobre parâmetros de forma usual e também inferência bayesiana a fim de se considerar nas predições a incerteza associada à estimação dos parâmetros do modelo.

A transformação de volta do vetor de médias e da matriz de covariância para o simplex \mathbb{S}^3 por quadratura de Gauss-Hermite de ordem igual a 7 não diferiram dos obtidos com ordem 20.

Os mapas de valores preditos de areia, silte e argila construídos de forma usual, tanto com a volta para o simplex por quadratura de Gauss-Hermite como por simulação, e por inferência bayesiana apresentaram padrão geral semelhante com diferenças nos detalhes.

Os tempos computacionais requeridos para a elaboração dos mapas de predição foram de 480 horas para dados reais quando da utilização da quadratura de Gauss-Hermite, 168 horas considerando a volta para o simplex por simulação e 168 horas por inferência bayesiana.

Como recomendação para trabalhos futuros que tratem do assunto em questão sugere-se:

1. aplicar a metodologia proposta considerando-se outras funções de correlação, como por exemplo, a Mátern;
2. investigar alternativas para computação mais eficiente;

3. considerar a especificação do modelo para o caso de maiores números de componentes;
4. considerar outras formas de especificação do modelo multivariado como por exemplo o apresentado em Schmidt e Sansó (2006).

Referências Bibliográficas

- ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of mathematical functions with formulas, graphs, and mathematical tables**. Washington: Milton Abramowitz and Irene A. Stegun, 1972, p. 924.
- AITCHISON, J. The statistical analysis of compositional data. **Royal Statistical Society, Series B**, v. 44, p. 139–177, 1982.
- AITCHISON, J. (Ed.). **The Statistical analysis of compositional data**. First. New Jersey: The Blackburn Press, 1986.
- AITCHISON, J. Logratios and natural laws in compositional data analysis. **Mathematical Geology**, v. 31, p. 563–580, 1999. Subject Collection: Earth and Environment Science.
- AITCHISON, J. et al. Logratios analysis and compositional distance. **Mathematical Geology**, v. 32, p. 563–580, 2000.
- AITCHISON, J.; EGOZCUE, J. J. Compositional data analysis: Where are we and where should we be heading? **Mathematical Geology**, v. 37, p. 829–850, October 2005. Subject Collection Earth and Environmental Science.
- AITCHISON, J.; GREENACRE, M. Biplot of compositional data. **Journal of the Royal Statistical Society, Series C**, v. 51, p. 375–392, October 2002.
- ANDREWS, D. F.; GNANADESIKAN, R.; WARNER, J. L. Methods for assessing multivariate normality. In: **Multivariate Analysis Survey Research Methods Section, ASA**. New York: [s.n.], 1973.
- AZZALINI, A. (Ed.). **Statistical Inference Based on the Likelihood**. London: Chapman and Hall, 1996.
- BAILEY, T. C.; GATRELL, A. C. **Interactive spatial data analysis**. Harlow: Longman, 1995.
- BANERJEE, S.; CARLIN, B. P.; GELFAND, A. E. **Hierarchical modelling and analysis for spatial data**. Boca Raton: Chapman and Hall, 2004.
- BARCELÓ-VIDAL, C.; MARTÍN-FERNÁNDEZ, J. A.; PAWLOWSKY-GLAHN, V. Mathematical foundations of compositional data analysis. 2001. Acesso em 08/06/08. Disponível em: <http://ima.udg.edu/~jamf/carles_martin_vera_cancun.pdf>.

- BOGNOLA, I. A. et al. Modelagem uni e bivariada da variabilidade espacial de rendimento de pinus taeda l. **Floresta**, v. 38, p. 373–385, 2008.
- BOOGAART, K. G. v. d. Using the r package “compositions”. p. 1–17, June 2005. Acesso em 09/02/08. Disponível em: <<http://www.stat.boogaart.de/compositions>>.
- BOOGAART, K. G. v. d.; TOLOSANA, R.; BREN, M. compositions package. Mar 2009. Acesso em 06/11/09. Disponível em: <<http://cran.r-project.org/web/packages/compositions>>.
- BOX, G.; COX, D. An analysis of transformation. **Journal of the Royal Statistical Society, Series B**, p. 211–252, 1964.
- BURDEN, R. L.; FAIRES, J. D. **Análise Numérica**. São Paulo: Pioneira Thomson Learning, 2003.
- BUTLER, A.; GLASBEY, C. A latent gaussian model for compositional data with zeros. **Journal of the Royal Statistical Society, Series C, (in press)**, 2008. Disponível em: <http://www.bioss.ac.uk/staff/adam/documents/ButlerGlasbey_resubmit.pdf>.
- BYRD, R. H. et al. A limited memory algorithm for bound constrained optimization. **SIAM J. Scientific Computing**, v. 16, p. 1190–1208, 1995.
- CARLIN, B. P.; LOUIS, T. A. **Bayesian Methods for Data Analysis**. Third. Boca Raton: Chapman and Hall, 2009.
- CHILÈS, J. P.; DELFINER, P. (Ed.). **Geostatistics: modeling spatial uncertainty**. New York: John Wiley & Sons, 1999.
- COX, D. R.; HINKLEY, D. V. **Theoretical Statistics**. London: Chapman & Hall, 1974.
- CRESSIE, N. A. C. (Ed.). **Statistics spatial data**. New York: Wiley, 1993.
- DEGROOT, M. H.; SCHERVISH, M. J. **Probability and Statistics**. Third. Boston: Addison Wesley, 2002.
- DIGGLE, P.; MENEZES, R.; SU, T.-l. Geostatistical inference under preferential sampling. **Journal of the Royal Statistical Society: Series C**, v. 59, 2010.
- DIGGLE, P. J.; RIBEIRO JR., P. J. **Model-based geostatistics**. USA: Springer Series in Statistics, 2007.
- DIGGLE, P. J.; RIBEIRO JR, P. J.; CHRISTENSEN, O. F. An introduction to model-based geostatistics. In: MØLLER, J. (Ed.). **Spatial Statistics and Computational Methods**. New York: Springer, 2003. p. 43–86.

DIGGLE, P. J.; TAWN, J. A.; MOYEED, R. A. Model-based geostatistics. **Applied Statistics**, v. 47, n. 3, p. 299–350, 1998. Disponível em: <citeseer.ist.psu.edu/diggle98modelbased.html>.

EGUCHI, E. S.; SILVA, E. L.; OLIVEIRA, M. S. Variabilidade espacial da textura e da densidade de partículas em um solo aluvial no município de lavras, mg. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 6, p. 242–246, 2002.

EHLERS, R. S. **Introdução à Inferência Bayesiana**. Curitiba: [s.n.], 2006. Disponível em: <<http://www.leg.ufpr.br/~ehlers>>.

EYNATTEN, H. v.; BARCELÓ-VIDAL, C.; PAWLOWSKY-GLAHN, V. Modelling compositional change: the example of chemical weathering of granitoid rocks. **Mathematical Geology**, v. 35, p. 231–251, April 2003. Subject Collection Earth and Environmental Science.

FALEIROS, A. C.; YONEYAMA, T. (Ed.). **Teoria Matemática de Sistemas**. São Paulo: Editora Arte & Ciência, 2002.

FLETCHER, R.; REEVES, C. M. Function minimization by conjugate gradients. **Computer Journal**, v. 7, p. 148–154, 1964.

GAMERMAN, D. **Markov chain Monte Carlo: stochastic simulation for Bayesian inference**. 1.ed.. ed. Londres: Chapman & Hall/CRC, 1997.

GAMERMAN, D.; LOPES, H. F. **Markov chain Monte Carlo: stochastic simulation for Bayesian inference**. 2a. ed. Londres: Chapman & Hall/CRC, 2006.

GELMAN, A. et al. (Ed.). **Bayesian Data Analysis**. Second. Boca Raton: Chapman and Hall, 2003.

GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. **Markov Chain Monte Carlo in Practice**. London: Chapman and Hall, 1996.

GILL, J. **Bayesian Methods For The Social And Behavioral Sciences**. London: Chapman and Hall, 2002.

GONÇALVES, A. C. A. (Ed.). **Variabilidade espacial de propriedades físicas do solo para fins de manejo da irrigação**. Tese (Doutorado em Agronomia) - Escola Superior de Agricultura “Luiz de Queiroz”. Universidade de São Paulo: Piracicaba, 1997.

GONÇALVES, A. C. A.; FOLEGATTI, M. V.; MATA, J. D. V. Análise exploratória e geoestatística da variabilidade de propriedades físicas de um Argissolo Vermelho. **Acta Scientiarum**, v. 23, p. 1149–1157, June 2001.

GOOVAERTS, P. **Geostatistics for Natural Resources Evaluation**. New York: Oxford University Press, 1997, Seção 4.2.3.

- GRAF, M. Precision of compositional data in a stratified two-stage cluster sample: comparison of the swiss earnings structure survey 2002 and 2004. In: **Survey Research Methods Section, ASA**. [S.l.: s.n.], 2006. p. 3066–3072.
- HOEL, P. G.; PORT, S. C.; STONE, C. J. (Ed.). **Introduction to Stochastic Processes**. Boston: Houghton Mifflin Company, 1972.
- ISAAKS, E. H.; SRISVASTAVA, R. M. **An introduction to applied geostatistics**. New York: Oxford University Press, 1989.
- JAMES, B. R. (Ed.). **Probabilidade: um curso em nível intermediário**. 3.ed.. ed. IMPA: Rio de Janeiro, 2004.
- JOHNSON, R. A.; WICHERN, D. W. **Applied statistical analysis**. Fourth. USA: Prentice Hall, 1998.
- KITANIDIS, P. **Introduction to geostatistics: applications in hydrogeology**. Third. New York: Cambridge University Press, 1997.
- LABUS, M. Compositional data analysis as a toll for interpretation of rock porosity parameters. **Geological Quarterly**, v. 49, p. 347–354, 2005.
- LAKE, J. V.; BOCK, G. R.; GOODE, J. A. (Ed.). **Precision agriculture: spatial and temporal variability of environmental quality (norvatis foundation symposia)**. USA: John wiley & Sons Ltd, 1997.
- LARK, R. M.; BISHOP, T. F. A. Cokriging particle size fractions of the soil. **European Journal of Soil Science**, v. 58, p. 763–774, june 2007.
- LEE, P. M. **Bayesian Statistics**. Third. New York: Arnold, 2004.
- LEMOS, R. C.; SANTOS, R. D. **Manual de descrição e coleta de solo no campo**. 3.ed.. ed. Campinas: Sociedade Brasileira de Ciência do Solo, 1996.
- LINDSEY, J. K. **Parametric statistical Inference**. New York: Oxford University Press, Inc., 2001.
- MARTIN, A. D.; QUINN, K. M.; PARK, J. H. Markov chain monte carlo (mcmc) package. p. 1–103, September 2009. Acesso em 01/10/09. Disponível em: <<http://mcmcpack.wustl.edu>>.
- MATÉRN, B. **Spatial Variation**. Stockholm, 1960.
- MATHERON, G. Principles of geostatistics. **Economic Geology**, v. 58, p. 1246–1266, 1963.
- MCBRATNEY, A. B.; DE GRUIJTER, J. J.; BRUS, D. J. Spatial prediction and mapping of continuous soil classes. **Geoderma**, p. 39–64, 1992.

- MOLIN, J. P. Agricultura de precisão. parte 1: o que é e estado da arte em sensoriamento. **Engenharia Agrícola**, v. 17, p. 97–107, 1997.
- MOLIN, J. P. Agricultura de precisão. parte 2: diagnóstico, aplicação localizada e considerações econômicas. **Engenharia Agrícola**, v. 17, p. 108–121, 1997.
- MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. (Ed.). **Introduction to the Theory of Statistics**. Third. USA: McGraw-Hill, 1974.
- NELDER, J. A.; MEAD, R. A simplex algorithm for function minimization. **Computer Journal**, v. 7, p. 308–313, 1965.
- NOTT, D. J.; FIELDING, M.; LEONTE, D. On a generalization of the Laplace approximation. **Statistics and Probability Letters**, v. 79, p. 1397–1403, 2009.
- OBAGE, S. C. (Ed.). **Uma análise bayesiana para dados composicionais**. Dissertação (Mestrado em Estatística). Universidade Federal de São Carlos: São Carlos, 2007.
- ODEH, I. O. A.; TOOD, A. J.; TRIANTAFILIS, J. Spatial prediction of soil particle-size fractions as compositional data. **Soil Science**, v. 168, n. 7, p. 501–515, July 2003.
- PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: Fundação Calouste Gulbenkian, 2003.
- PAWITAN, Y. (Ed.). **In All Likelihood**. New York: Oxford University Press, Inc., 2001.
- PAWLOWSKY-GLAHN, V.; EGOZCUE, J. J.; TOLOSANA-DELGADO, R. **Lecture Notes on Compositional Data Analysis**. [s.n.], 2007. Disponível em: <<http://www.sediment.uni-goettingen.de/staff/tolosana/extra/CoDa.pdf>>.
- PAWLOWSKY-GLAHN, V.; OLEA, R. A. (Ed.). **Geostatistical Analysis of Compositional Data**. New York: Oxford University Press, Inc., 2004. (Studies in Mathematical Geology;7).
- PAWLOWSKY-GLAHN, V.; OLEA, R. A. (Ed.). **Geostatistical Analysis of Compositional Data**. New York: Oxford University Press, Inc., 2004, Seção 6.1. (Studies in Mathematical Geology;7).
- PAWLOWSKY-GLAHN, V.; OLEA, R. A.; DAVIS, J. C. Estimation of regionalized compositions: a comparison of three methods. **Mathematical Geology**, v. 27, n. 1, p. 105–127, 1995.
- PLUMMER, M. et al. CODA: Convergence diagnosis and output analysis for MCMC. **R News**, v. 6, n. 1, p. 7–11, March 2006. Disponível em: <<http://CRAN.R-project.org/doc/Rnews/>>.
- REICHARDT, K.; TIMM, L. C. **Solo, planta e atmosfera: conceitos, processos e aplicações**. Barueri, SP: Manole, 2004.

- REIS, E. **Estatística Multivariada Aplicada**. Lisboa: Edições Sílabo, LDA, 1997.
- REYMENT, R. A.; SAVAZZI, E. (Ed.). **Aspects of Multivariate Statistical Analysis in Geology**. [S.l.]: Elsevier, 1999.
- RIBEIRO JR, P. J.; DIGGLE, P. J. **Bayesian inference in Gaussian model-based geostatistics**. Lancaster, 1999. Relatório Técnico.
- RIBEIRO JR., P. J.; DIGGLE, P. J. geoR: a package from geostatistical analysis. **R-NEWS**, v. 1, p. 15–18, June 2001. ISSN 1609-3631. Disponível em: <<http://cran.R-project.org/doc/Rnews>>.
- RUE, H.; HELD, L. (Ed.). **Gaussian Markov Random Fields**. Boca Raton: Chapman and Hall, 2005.
- SCHABENBERGER, O.; GOTWAY, C. A. **Statistical Methods for Spatial Data Analysis**. Boca Raton: Chapman and Hall, 2005.
- SCHABENBERGER, O.; PIERCE, F. J. (Ed.). **Contemporary Statistical Models for the Plant and Soil Sciences**. Boca Raton: Taylor and Francis, 2001.
- SCHMIDT, A. M.; GELFAND, A. E. A bayesian coregionalization approach for multivariate pollutant data. **Journal of Geophysical Research**, v. 108, p. 10–1–10–8, 2003.
- SCHMIDT, A. M.; SANSÓ, B. Modelagem bayesiana da estrutura de covariância de processos espaciais e espaço temporais. In: *17 SINAPE e ABE-Associação Brasileira de Estatística*. Caxambu: Associação Brasileira de Estatística, 2006. **Minicurso**.
- SILVA, E. A. A. (Ed.). **Aplicação de um estimador robusto na análise de variabilidade espacial de parâmetros de agricultura de precisão**. Tese (Mestrado em engenharia agrícola). Universidade Estadual do Oeste do Paraná: Cascavel, 2000.
- SILVA, E. A. A. (Ed.). **Aplicação de métodos geoestatísticos multivariados em problemas de mapeamento de variáveis do sistema solo-planta**. Tese (Doutorado em ciências). Universidade de Federal de Curitiba: Curitiba, 2008.
- SILVA, E. A. A. et al. Um estimador robusto e o semivariograma cruzado na análise de variabilidade espacial de atributos de solo e planta. **Acta Scientiarum**, v. 25, p. 365–371, 2003.
- SMYTH, G.; HU, Y.; DUNN, P. statmod package. September 2009. Acesso em 06/11/09. Disponível em: <<http://cran.r-project.org/web/packages/statmod>>.
- SRINIVASAN, A. (Ed.). **Handbook of precision agriculture: principles and applications**. New York: The Haworth of Precision Agriculture, 2006.

STEIN, M. L. **Interpolation of Spatial Data: Some Theory for Kriging**. New York: Springer Verlag, 1999.

STEPHENS, M. A. EDF statistics for goodness of fit and some comparisons. **Journal of American Statistical Association**, n. 69, p. 730–737, 1974.

TJELMELAND, H.; LUND, K. V. Bayesian modelling of spatial compositional data. **Journal of Applied Statistics**, v. 30, p. 87–100, 2003.

TOLOSANA-DELGADO, R.; OTERO, N.; PAWLOWSKY-GLAHN, V. Some basic concepts of compositional geometry. **Mathematical Geology**, v. 37, n. 7, p. 563–580, October 2005.

WACKERNAGEL, H. (Ed.). **Multivariate Geostatistics: An Introduction with Applications**. Second. Germany: Springer, 1998.

ANEXO A – Pacote geoComp

Pacote geoComp

pivo.R:

```
pivo <-
structure(list(X = c(20, 40, 60, 80, 100, 120, 140, 160, 180,
0, 40, 60, 80, 100, 120, 140, 160, 180, 0, 20, 60, 80, 100, 120,
140, 160, 180, 0, 20, 40, 80, 100, 120, 140, 160, 180, 0, 20,
40, 60, 100, 120, 140, 160, 0, 20, 40, 60, 80, 120, 140, 160,
0, 20, 40, 60, 80, 100, 140, 0, 20, 40, 60, 80, 100, 120, 0,
20, 40, 60, 80, 100, 0, 20, 40, 60), Y = c(0, 0, 0, 0, 0, 0,
0, 0, 0, 20, 20, 20, 20, 20, 20, 20, 20, 20, 20, 40, 40, 40, 40,
40, 40, 40, 40, 40, 60, 60, 60, 60, 60, 60, 60, 60, 60, 60, 80, 80,
80, 80, 80, 80, 80, 80, 100, 100, 100, 100, 100, 100, 100, 100,
120, 120, 120, 120, 120, 120, 120, 120, 140, 140, 140, 140, 140, 140,
140, 160, 160, 160, 160, 160, 160, 160, 180, 180, 180, 180),DenSoil=c(1.69,
1.58, 1.44, 1.5, 1.58, 1.45, 1.69, 1.49, 1.38, 1.56, 1.68, 1.62,
1.67, 1.64, 1.52, 1.66, 1.55, 1.51, 1.69, 1.49, 1.69, 1.55, 1.53,
1.48, 1.58, 1.65, 1.6, 1.63, 1.69, 1.56, 1.67, 1.67, 1.51, 1.59,
1.61, 1.51, 1.56, 1.58, 1.41, 1.54, 1.63, 1.6, 1.69, 1.54, 1.71,
1.71, 1.61, 1.64, 1.56, 1.6, 1.63, 1.64, 1.7, 1.52, 1.64, 1.69,
1.66, 1.7, 1.74, 1.63, 1.67, 1.72, 1.69, 1.64, 1.61, 1.69, 1.49,
1.6, 1.75, 1.65, 1.48, 1.64, 1.58, 1.68, 1.64, 1.65),DenPar=c(2.63,
2.87, 2.55, 2.57, 2.56, 2.69, 2.62, 2.7, 2.71, 2.76, 2.57, 2.68,
2.57, 2.41, 2.69, 2.68, 2.62, 2.73, 2.52, 2.49, 2.4, 2.71, 2.89,
2.78, 2.83, 2.66, 2.69, 2.56, 2.41, 2.55, 2.38, 2.59, 2.56, 2.84,
2.67, 2.55, 2.59, 2.74, 2.6, 2.58, 2.74, 2.53, 2.6, 2.51, 2.56,
2.65, 2.71, 2.48, 2.57, 2.58, 2.64, 2.56, 2.63, 2.67, 2.64, 2.61,
2.53, 2.55, 2.54, 2.34, 2.45, 2.51, 2.71, 2.48, 2.65, 2.6, 2.72,
```



```

2.58, 2.55, 2.44, 2.56, 2.54, 2.57, 2.55, 2.56, 2.7), PoroTot=c(35.61,
45.08, 43.51, 41.48, 38.31, 46.05, 35.43, 44.93, 49.19, 43.56,
34.46, 39.69, 34.98, 31.92, 43.44, 38.07, 40.77, 44.58, 32.88,
40.34, 29.62, 42.84, 47.1, 46.76, 44.22, 38.09, 40.55, 36.35,
29.97, 38.75, 29.64, 35.64, 41.04, 43.88, 39.68, 40.91, 39.9,
42.16, 45.61, 40.49, 40.54, 36.91, 34.92, 38.61, 33.17, 35.59,
40.66, 33.97, 39.11, 37.99, 38.26, 35.98, 35.31, 43.23, 37.74,
35.3, 34.55, 33.4, 31.47, 30.41, 31.71, 31.3, 37.77, 33.91, 39.25,
35.12, 45.37, 38.06, 31.55, 32.37, 42.2, 35.46, 38.64, 34.03,
35.75, 38.82), Areia = c(31, 24, 23, 22, 22, 16, 24, 11, 11,
28, 25, 32, 33, 36, 25, 27, 24, 14, 33, 16, 38, 31, 27, 24, 27,
40, 32, 34, 32, 23, 40, 33, 31, 22, 38, 30, 25, 33, 21, 27, 33,
36, 42, 40, 32, 33, 22, 32, 27, 31, 36, 29, 36, 25, 25, 25, 31,
36, 42, 36, 33, 30, 21, 34, 30, 32, 19, 23, 28, 25, 16, 23, 17,
32, 25, 30), Silte = c(25, 21, 27, 25, 25, 25, 28, 19, 21, 17,
25, 27, 30, 24, 21, 25, 15, 26, 27, 19, 25, 21, 19, 17, 19, 27,
27, 23, 27, 21, 23, 21, 21, 17, 22, 23, 23, 25, 23, 25, 21, 27,
23, 23, 27, 25, 20, 27, 21, 23, 22, 23, 27, 19, 25, 23, 25, 23,
25, 20, 21, 25, 21, 23, 23, 23, 19, 19, 27, 23, 19, 21, 23, 25,
23, 23), Argila = c(44, 55, 50, 53, 53, 59, 48, 70, 68, 55, 50,
41, 37, 40, 54, 48, 61, 60, 40, 65, 37, 48, 54, 59, 54, 33, 41,
43, 41, 56, 37, 46, 48, 61, 40, 47, 52, 42, 56, 48, 46, 37, 35,
37, 41, 42, 58, 41, 52, 46, 42, 48, 37, 56, 50, 52, 44, 41, 33,
44, 46, 45, 58, 43, 47, 45, 62, 58, 45, 52, 65, 56, 60, 43, 52,
47)), .Names = c("X", "Y", "DenSoil", "DenPar", "PoroTot", "Areia",
"Silte", "Argila"), row.names = c(NA, 76L), class = "data.frame")

```

Função `as.geoComp.R`:

```

`as.geoComp` <-
function(original, # Um data frame contendo Areia Silte Argila X Y
          perc=TRUE # Se está ou não em porcentagem
        ){
  if(perc == TRUE){
    malha <- transform(original, Areia=(Areia/100), Silte=(Silte/100),
                      Argila=(Argila/100))
  }
  malha <- original
  malha <- transform(malha, y1=log(Areia/Argila), y2=log(Silte/Argila))

```

```

seq11 <- seq(1,length(malha$y1)*2,by=2)
seq22 <- seq(2,length(malha$y2)*2,by=2)
y <- c()
y[seq11] <- malha$y1
y[seq22] <- malha$y2
saida <- list()
saida[[1]] <- malha[c(6,7)]
saida[[2]] <- data.frame(y)
saida[[3]] <- data.frame(original[4],original[5])
names(saida[[1]]) <- c("Y1","Y2")
names(saida[[2]]) <- c("Y")
names(saida[[3]]) <- c("Coord.X","Coord.Y")
class(saida) <- "geoComp"
return(saida)
}

```

Função **monta.V.R**:

```

`monta.V` <-
function(theta1,dados.comp){
  y <- dados.comp[[2]][[1]]
  eta <- theta1[1]
  nu1 <- theta1[2]
  nu2 <- theta1[3]
  phi <- theta1[4]
  rho <- theta1[5]
  ## Calculando os elementos da matriz de correlação espacial
  distancia <- unname(as.matrix(dist(dados.comp[[3]],diag=TRUE,upper=TRUE)))
  correla <- exp(-distancia/phi)
  ## Organizando a matriz de correlação espacial com eltos y1 e y2
  ## intercalados. Ou seja, calculando R1.
  coluna1 <- rep(c(1, eta),length(y)/2)
  coluna2 <- rep(c(eta,eta^2),length(y)/2)
  seq1 <- seq(1,length(y),by=2)
  seq2 <- seq(2,length(y),by=2)
  ERRE1 <- matrix(ncol=length(y),nrow=length(y))
  ERRE1[,c(seq1)] <- coluna1
  ERRE1[,c(seq2)] <- coluna2
  ## Organizando a matriz de correlação espacial com eltos y1 e y2

```

```

## intercalados.
ERRE2 <- matrix(nrow=length(y),ncol=length(y))
for(i in 1:length(y)/2){
ERRE2[seq1[i],] <- rep(correla[i,],each=2)
ERRE2[seq2[i],] <- rep(correla[i,],each=2)
}
## Calculando a matriz R
ERRE <- ERRE1*ERRE2
## Calculando a matriz de covariância composicional
Ib1 <- matrix(c(nu1^2,nu1*nu2*rho,nu1*nu2*rho,nu2^2),nc=2)
Ib <- kronecker(diag(length(y)/2),Ib1)
## Matriz de covariância espacial composicional
V <- ERRE + Ib
return(V)
}

```

Função **logarit.vero.R**:

```

`logarit.vero` <-
function(theta1,dados.comp,print.pars=FALSE){
  if(theta1[1] <= 0) return(.Machine$double.xmax^0.5)
  if(theta1[2] <= 0) return(.Machine$double.xmax^0.5)
  if(theta1[3] <= 0) return(.Machine$double.xmax^0.5)
  if(theta1[4] < 0) return(.Machine$double.xmax^0.5)
  if(abs(theta1[5]) > 1) return(.Machine$double.xmax^0.5)
  ## Pegando o Y
  y <- dados.comp[[2]][[1]]
  X <- cbind(rep(1:0,length=length(y)),rep(0:1,length=length(y)))
  V <- monta.V(theta1, dados.comp = dados.comp)
  ldetV <- determinant(V,log=TRUE)$modulus[1]
  ## mu não eficiente:
  ## mu <- drop(solve(crossprod(X,solve(V,X)))*%crossprod(X,solve(V,y)))
  ## mu eficiente:
  R <- chol(V)
  back.x <- backsolve(R,X,upper.tri=TRUE,transpose=TRUE)
  back.y <- backsolve(R,y,upper.tri=TRUE,transpose=TRUE)
  mu <- as.numeric(solve(crossprod(back.x,back.x))*%
                    crossprod(back.x,back.y))
  ## Qe não eficiente:

```

```

## Qe <- drop(crossprod(y,solve(V,y))-2*crossprod(y,solve(V,X%*%mu))
              +crossprod(mu,crossprod(X,solve(V,X%*%mu))))
## Qe eficiente:
desvio <- y-X%*%mu
back.desvio <- backsolve(R,desvio,upper.tri=TRUE,transpose=TRUE)
Qe <- as.numeric(crossprod(back.desvio,back.desvio))
if(Qe < 0) return(Qe=.Machine$double.xmax^0.5)
n <- length(y)
s1 <- sqrt(Qe/n)
# drop diminui a dimensão do objeto
ll <- drop(-0.5*(n*log(2*pi)+n*log(s1^2)+ ldetV +n ))
if(print.pars) print(c(theta1,ll))
return(-ll)
}

```

Função **mec.R**:

```

`mec` <-
function(dados.comp, metodo="L-BFGS-B",print.pars=FALSE, alpha=0.975){
  if(class(dados.comp) != "geoComp") stop("A classe do objeto deve ser
                                          geoComp")

  y1 <- dados.comp[[1]]$Y1
  y2 <- dados.comp[[1]]$Y2
  coords <- dados.comp[[3]]
  X <- cbind(rep(1:0,length=length(y1)*2),rep(0:1,length=length(y1)*2))
  ## Calculando os valores iniciais
  mu1 <- mean(y1)
  mu2 <- mean(y2)
  var_y1 <- var(y1)
  s1 <- var_y1/2
  tau1 <- s1
  var_y2 <- var(y2)
  s2 <- var_y2/2
  tau2 <- s2
  dim <- range(dist(coords))
  ## Um "chute" para phi
  phi <- dim[1] + 0.2*(dim[2]-dim[1])
  rho <- cor(y1,y2)
  ## Pegando o Y

```

```

y <- dados.comp[[2]][[1]]
## Reparametrizando
eta <- s2/s1
nu1 <- tau1/s1
nu2 <- tau2/s1
theta1 <- c(eta,nu1,nu2,phi,rho)
## Calculando a log-vero para o chute inicial
ll <- logarit.vero(theta1,dados.comp)
## Otimizando
if(metodo== "L-BFGS-B"){
  estim <- optim(theta1,logarit.vero,dados.comp = dados.comp,
                hessian=TRUE,method=metodo,print.pars=print.pars,
                lower=c(1e-32,1e-32,1e-32,1e-32,-1),
                upper=c(Inf,Inf,Inf,Inf,1))
  ifelse(estim$convergence[1] == 0, print("O algoritmo convergiu"),
        print("Existe problemas na convergencia"))
}
if(metodo=="Nelder-Mead" | metodo=="CG" | metodo=="BFGS"){
  estim <- optim(theta1,logarit.vero,dados.comp = dados.comp,hessian=TRUE,
                method=metodo,print.pars=print.pars)
  ifelse(estim$convergence[1] == 0, print("O algoritmo convergiu"),
        print("Existe problemas na convergencia"))
}
## Voltando aos valores iniciais
thetaest <- estim$par
Vest <- monta.V(thetaest,dados.comp)
ldetVest <- determinant(Vest,log=TRUE)$modulus[1]
## muest não eficiente:
## muest<- drop(solve(crossprod(X,solve(Vest,X)))*%
               crossprod(X,solve(Vest,y)))
## muest eficiente:
R.est <- chol(Vest)
back.x.est <- backsolve(R.est,X,upper.tri=TRUE,transpose=TRUE)
back.y.est <- backsolve(R.est,y,upper.tri=TRUE,transpose=TRUE)
invXVX <- solve(crossprod(back.x.est,back.x.est))
muest <- as.numeric(invXVX*%crossprod(back.x.est,back.y.est))
## Qeest não eficiente:
## Qeest <- drop(crossprod(y,solve(Vest,y))-

```

```

                2*crossprod(y,solve(Vest,X%%muest))+
                crossprod(muest,crossprod(X,solve(Vest,X%%muest))))
## Qeest eficiente:
desvio.est <- y-X%%muest
back.desvio.est <- backsolve(R.est,desvio.est,upper.tri=TRUE,
                             transpose=TRUE)
Qeest <- as.numeric(crossprod(back.desvio.est,back.desvio.est))
n <- length(y)
s1est <- sqrt(Qeest/n)
llest <- drop(-0.5*(n*log(2*pi)+n*log(s1est^2)+ ldetVest +n ))
s2est <- thetaest[1]*s1est
tau1est <- thetaest[2]*s1est
tau2est <- thetaest[3]*s1est
## Calculando a variância para mu e s1: 1/Io(mu) - 1/Io(s1)
## V.g.mu nao eficiente
## V.g.mu <- (s1est)^2*solve(crossprod(X,solve(Vest,X)))
## V.g.mu eficiente
V.g.mu <- (s1est)^2*invXVX
V.g.s1 <- (s1est)^3/(3*Qeest-n*s1est)
## Calculando o gradiente
g<- matrix(c(s1est,0,0,0,0,0,s1est,0,0,0,0,0,s1est,
             0,0,0,0,0,1,0,0,0,0,0,1),nr=5)
## Calculando a variância para s2,tau1,tau2,phi e rho - usa hessiana
V.g <- crossprod(g,diag(1/diag(estim$hessian))%*%g)
## Preparando a saída
para <- c("mu1","mu2","s1","s2","tau1","tau2","phi","rho")
estima <- c(muest[1],muest[2],s1est,s2est,tau1est,tau2est,
           thetaest[4],thetaest[5])
errosdpd <- qnorm(alpha)*sqrt(c(V.g.mu[1],V.g.mu[2],V.g.s1,diag(V.g)))
## Intervalo de confiança pelo método Delta
ic.min <- estima - errosdpd
ic.max <- estima + errosdpd
retorna <- list()
retorna[1][[1]] <- data.frame(para,estima,errosdpd,ic.min,ic.max)
names(retorna[1][[1]]) <- c("Parametros","Estimativas","Erro Padrao",
                          "LI.Delta","LS.Delta")
retorna[2][[1]] <-c(l1,llest)
names(retorna[2][[1]]) <- c("LogLik Start", "LogLik Optim")

```

```

    return(retorna)
}

```

Função **cokrigagem.R**:

```

`cokrigagem` <-
function(estim.par, ## Parâmetros da otimização
        dim.gride, ## Dimensão da grade para predição
        locations,
        dados.comp ## Uma lista com dados e localizações amostradas
        ){
  if(missing(locations)){
    rang.coord.x <- range(dados.comp[[3]][[1]])
    rang.coord.y <- range(dados.comp[[3]][[1]])
    c.x <- seq(rang.coord.x[1],rang.coord.x[2],length=dim.gride)
    c.y <- seq(rang.coord.y[1],rang.coord.y[2],length=dim.gride)
    locations <- expand.grid(c.x,c.y)
  }
  estim.par <- as.matrix(estim.par)
  n.local <- length(locations[,1])*2
  dad <- list()
  dad[[2]] <- data.frame(c(rep(0,length=n.local),dados.comp[[2]][[1]]))
  dad[[3]] <- as.data.frame(rbind(as.matrix(locations),
                                as.matrix(dados.comp[[3]])))
  ## Reparametrizando
  theta1ck <- c(estim.par[4]/estim.par[3],estim.par[5]/estim.par[3],
               estim.par[6]/estim.par[3],estim.par[7],estim.par[8])
  ## Montando a matriz de covariância
  sigmack <- monta.V(theta1ck,dad)
  n.local.mais <- n.local +1
  dim.matriz <- dim(sigmack)[1]
  S00 <- sigmack[c(1:n.local),c(1:n.local)]
  S0y <- sigmack[c(1:n.local),c(n.local.mais:dim.matriz)]
  Sy0 <- sigmack[c(n.local.mais:dim.matriz),c(1:n.local)]
  Syy <- sigmack[c(n.local.mais:dim.matriz),c(n.local.mais:dim.matriz)]
  muy <- rep(c(estim.par[1],estim.par[2]),length(dados.comp[[1]]$Y1))
  ## Como a média de Y0 é desconhecida, substituímos pela média geral estimada
  muy0 <- rep(c(estim.par[1],estim.par[2]),n.local/2)
  SoyInvSyy <- S0y %*% solve(Syy)

```

```

## Calculando a média de cokrigagem estimada:
## muck.est não eficiente:
## muck.est <- muy0 + S0y %*% solve(Syy, (dados[[2]][[1]]-muy))
muck.est <- muy0 + SoyInvSyy %*% (dados[[2]][[1]]-muy)
## Calculando a variancia de cokrigagem estimada:
## vck.est não eficiente:
## vck.est <- S00 - S0y %*% solve(Syy,Sy0)
## vck.est eficiente:
vck.est <- S00 - SoyInvSyy %*% Sy0
res.cokri <- list()
res.cokri[[1]] <- muck.est
res.cokri[[2]] <- vck.est
if(missing(locations)) attr(res.cokri, "locations") <- locations
return(res.cokri)
}

```

Função **agl.R**:

```

`agl` <-
function(dados){
    acre <- exp(data.frame(dados[1],dados[2],0))
    compos <- acre/rowSums(acre)
    return(compos)
}

```

Função **volta.quad.R**:

```

`volta.quad` <-
function(med.cov,          # Média e covariância de cokrigagem.
        n.pontos=7,      # Ordem da quadratura de Gauss-Hermite.
        Variancia=FALSE # Calcula ou não a variância.
        ){
    mu.ck <- med.cov[[1]]
    sigma.ck <- med.cov[[2]]
    desvio.pd <- sqrt(diag(sigma.ck))
    ic.mu <- data.frame(mu.ck-qnorm(0.975)*desvio.pd,mu.ck+
                        qnorm(0.975)*desvio.pd)
    names(ic.mu) <- c("L.Minimo","L.Maximo")
    g <- function(Y,mu,R,pos){

```



```

arg.agl <- mu+sqrt(2)*t(R)%*%Y
g1 <- t(pi^-1*agl(arg.agl))
g2 <- pi^-1*(t(as.matrix(agl(arg.agl)))-g1)%*%
          t(t(as.matrix(agl(arg.agl)))-g1)
resultado <- cbind(g1,g2)
return(resultado[pos])
}

## Calculando a quadratura de Gauss-Hermite
quad.gauss.comp <- function(mu,R,func=g,pos,np=n.pontos){
  AbcPeso <- gauss.quad(np,kind='hermite')
  Y <- AbcPeso$nodes
  peso<- AbcPeso$weights
  soma <- 0
  for(i in 1:np){
    for(j in 1:np){
      soma<-peso[i]*peso[j]*func(Y=c(Y[i],Y[j]),mu=mu,R=R,
          pos=pos)+soma
    }
  }
  return(soma)
}

## Calculando a quadratura de Gauss-Hermite para o vetor de médias
n.linhas <- length(mu.ck)/2
compo.md <- matrix(ncol=3,nrow=n.linhas)
seq1 <- seq(1,n.linhas*2,by=2)
seq2 <- seq(2,n.linhas*2,by=2)
for(i in 1:n.linhas){
  sigma.ck <- med.cov[[2]][seq1[i]:seq2[i],seq1[i]:seq2[i]]
  mu.ck <- med.cov[[1]][seq1[i]:seq2[i]]
  cp <- quad.gauss.comp(mu=mu.ck,R=chol(sigma.ck),func=g,pos=c(1:3),
          np=n.pontos)
  compo.md[i,] <- cp
}
#rowSums(compo.md)

## Calculando a quadratura de Gauss-Hermite para a matriz
## matriz de covariâncias
if(Variancia == TRUE){
  compo.var <- matrix(ncol=3,nrow=n.linhas)

```

```

for(i in 1:n.linhas){
  sigma.ck <- med.cov[[2]][seq1[i]:seq2[i],seq1[i]:seq2[i]]
  mu.ck <- med.cov[[1]][seq1[i]:seq2[i]]
  cp <- quad.gauss.comp(mu=mu.ck,R=chol(sigma.ck),func=g,pos=c(4,8,12),
                        np=n.pontos)

  compo.var[i,] <- cp
  }}
retorna <- list()
retorna[[1]] <- compo.md
if(Variancia == TRUE){
  retorna[[2]] <- compo.var}
return(retorna)
}

```

Função `volta.cokri.R`

```

`volta.cokri` <-
function(mat.cokri, num.simu, retorna.tudo=FALSE, int.conf=0.95){
  nlinhas <- dim(mat.cokri[[1]])[1]
  compos1 <- data.frame(matrix(nrow=nlinhas/2,ncol=3))
  compos <- data.frame(matrix(nrow=nlinhas/2,ncol=3))
  for(i in 1:num.simu){
    g <- mvnorm(n=1, mat.cokri[[1]],mat.cokri[[2]])
    seq1 <- seq(1,nlinhas,by=2)
    seq2 <- seq(2,nlinhas,by=2)
    y1 <- g[seq1]
    y2 <- g[seq2]
    gerado <- data.frame(y1,y2)
    compos <- agl(gerado)
    compos1 <- cbind(compos,compos1)
  }
  compos2 <- as.matrix(compos1)
  dim.vetor <- num.simu*3
  sy1 <- seq(1,dim.vetor,by=3)
  sy2 <- seq(2,dim.vetor,by=3)
  sy3 <- seq(3,dim.vetor,by=3)
  amostra1 <- as.matrix(compos2[,sy1],ncol=num.simu)
  amostra2 <- as.matrix(compos2[,sy2],ncol=num.simu)
  amostra3 <- as.matrix(compos2[,sy3],ncol=num.simu)
}

```

```

if(retorna.tudo == TRUE){
  retorna <- list()
  retorna[[1]] <- amostra1
  retorna[[2]] <- amostra2
  retorna[[3]] <- amostra3
  return(retorna)
}
if(retorna.tudo == FALSE){
  med1 <- apply(amostra1,1,mean)
  med2 <- apply(amostra2,1,mean)
  med3 <- apply(amostra3,1,mean)
  q1 <- t(apply(amostra1,1,quantile,prob=c(1-int.conf,int.conf)))
  q2 <- t(apply(amostra2,1,quantile,prob=c(1-int.conf,int.conf)))
  q3 <- t(apply(amostra3,1,quantile,prob=c(1-int.conf,int.conf)))
  quantis <- cbind(q1,q2,q3)
  resultado <- list()
  resultado$preditos <- data.frame(med1,med2,med3)
  names(resultado$preditos) <- c("Areia","Silte","Argila")
  resultado$intervalo <- data.frame(quantis)
  names(resultado$intervalo) <- c("LI Areia", "LS Areia",
    "LI Silte", "LS Silte", "LI Argila", "LS Argila")
  return(resultado)
}
}

```

Arquivo **posteriori.R**:

```

## theta1=c(eta,nu1,nu2,phi,rho) obtidos da amostra.
## mediahiper=medias para c(eta,nu1,nu2,phi) e correspondentes
## desvios padrão.
`posteriori` <-
function(theta1,dados.comp,mediahiper,print.pars=TRUE){
  if(theta1[1] <= 0)      return(-Inf)
  if(theta1[2] <= 0)      return(-Inf)
  if(theta1[3] <= 0)      return(-Inf)
  if(theta1[4] <= 0)      return(-Inf)
  if(abs(theta1[5]) > 1) return(-Inf)
  ## Pegando o Y
  y <- dados.comp[[2]][[1]]

```

```

X <- cbind(rep(1:0,length=length(y)),rep(0:1,length=length(y)))
V <- monta.V(theta1, dados.comp = dados.comp)
ldetV <- determinant(V,log=TRUE)$modulus[1]
## mu não eficiente:
## mu <- drop(solve(crossprod(X,solve(V,X)))*%crossprod(X,solve(V,y)))
## mu eficiente:
R <- chol(V)
back.x <- backsolve(R,X,upper.tri=TRUE,transpose=TRUE)
back.y <- backsolve(R,y,upper.tri=TRUE,transpose=TRUE)
mu <- as.numeric(solve(crossprod(back.x,back.x))*%
                  crossprod(back.x,back.y))
## Qe não eficiente:
## Qe <- drop(crossprod(y,solve(V,y))-2*crossprod(y,solve(V,X%*%mu))+
##           crossprod(mu,crossprod(X,solve(V,X%*%mu))))
## Qe eficiente:
desvio <- y-X%*%mu
back.desvio <- backsolve(R,desvio,upper.tri=TRUE,transpose=TRUE)
Qe <- as.numeric(crossprod(back.desvio,back.desvio))
if(Qe < 0) return(Qe=.Machine$double.xmax^0.5)
n <- length(y)
s1 <- sqrt(Qe/n)
ll <- drop(-0.5*(n*log(2*pi)+n*log(s1^2)+ldetV+n))+
        dlnorm(theta1[1],log(0.24),sd=0.3,log=TRUE)+
        dlnorm(theta1[2],log(0.63),sd=0.3,log=TRUE)+
        dlnorm(theta1[3],log(0.59),sd=0.3,log=TRUE)+
        dgamma(theta1[4],66,1,log=TRUE)
## Usuário entrando com os valores dos hiperparâmetros
if(length(mediahiper)==8){
  ll <- drop(-0.5*(n*log(2*pi)+n*log(s1^2)+ldetV+n))+
          dnorm(theta1[1],mediahiper[1],sd=mediahiper[5],log=TRUE)+
          dnorm(theta1[2],mediahiper[2],sd=mediahiper[6],log=TRUE)+
          dnorm(theta1[3],mediahiper[3],sd=mediahiper[7],log=TRUE)+
          dgamma(theta1[4],mediahiper[4],mediahiper[8],log=TRUE)
}
if(print.pars) print(c(theta1,ll))
return(ll)
}

```

Arquivo `funcaoInfBayes.R`:

```

`infbayes` <-
function(dados,tune,queima,salto,nsim,verbose=TRUE,mediahiper){
  y1 <- dados[[1]]$Y1
  y2 <- dados[[1]]$Y2
  coords <- dados[[3]]
  mu1 <- mean(y1)
  mu2 <- mean(y2)
  var_y1 <- var(y1)
  s1 <- var_y1/2
  tau1 <- s1
  var_y2 <- var(y2)
  s2 <- var_y2/2
  tau2 <- s2
  dim <- range(dist(coords))
  phi <- dim[1] + 0.2*(dim[2]-dim[1])
  rho <- cor(y1,y2)
  eta <- s2/s1
  nu1 <- tau1/s1
  nu2 <- tau2/s1
  theta1 <- c(eta,nu1,nu2,phi,rho)
  if(verbose==TRUE)print("Otimizando a log-verossimilhanca")
  inicial.optim <- optim(theta1, logarit.vero, dados.comp=dados,
                        method="L-BFGS-B",
                        lower=c(1e-32,1e-32,1e-32,1e-32,-1),
                        upper=c(Inf,Inf,Inf,Inf,1),hessian=TRUE)
  print(inicial.optim$par)
  inv.hessiana <-solve(inicial.optim$hessian)
  sd.5 <- sqrt(c(inv.hessiana[1,1],inv.hessiana[2,2],inv.hessiana[3,3],
                inv.hessiana[4,4],inv.hessiana[5,5]))
  tune <- tune
  diag.tune <- diag(c(tune))
  mc.V <- diag.tune %*% inv.hessiana %*% diag.tune
  if(verbose==TRUE)print("Rodando o MCMC")
  theta1.inicial <- inicial.optim$par
  if(length(mediahiper)==0){
    mh <- MCMCmetrop1R(posteriori,theta1.inicial,

```

```

        mediahiper=c(theta1.inicial[1:4],sd.5[1:4]),
        dados.comp=dados,burnin=queima,mcmc=nsim,thin=salto,
        V=mc.V)
}
if(length(mediahiper)==8){
  mh1 <- MCMCmetrop1R(posteriori,theta1.inicial,
    mediahiper=c(mediahiper[1:4],mediahiper[5:8]),
    dados.comp=dados,burnin=queima,mcmc=nsim,thin=salto,
    V=mc.V)
}
#plot(mh)
saida.mh <-< mh
n <- length(dados[[2]]$Y)
p <- length(dados[[1]])
D <- cbind(rep(1:0,length=n),rep(0:1,length=n))
sigma <- c()
tamanho <- dim(mh)[1]
sigma.media <- matrix(ncol=3,nrow=tamanho)
for(i in 1:tamanho){
  V <- monta.V(mh[i,],dados)
  ## beta não eficiente
  ## inv.V <- solve(V)
  ## beta <- solve(t(D)%*%inv.V%*%D)%*%(t(D)%*%inv.V%*%dados[[2]]$Y)
  ## beta eficiente
  Rbayes <- chol(V)
  back.bayes1 <- backsolve(Rbayes,D,upper.tri=TRUE,transpose=TRUE)
  back.bayes2 <- backsolve(Rbayes,dados[[2]]$Y,upper.tri=TRUE,
    transpose=TRUE)
  inv.tDinvVD <- solve(crossprod(back.bayes1,back.bayes1))
  beta <- inv.tDinvVD%*%crossprod(back.bayes1,back.bayes2)
  ## s2 não eficiente:
  ## s2 <- 1/(n-p)*t(dados[[2]]$Y-D%*%beta)%*%inv.V%*%
    (dados[[2]]$Y-D%*%beta)
  ## s2 eficiente:
  desvio.bayes <- dados[[2]]$Y-D%*%beta
  back.desvio.bayes <- backsolve(Rbayes,desvio.bayes,
    upper.tri=TRUE,transpose=TRUE)
  s2 <- 1/(n-p)*crossprod(back.desvio.bayes,back.desvio.bayes)
}

```

```

sigma[i] <- rinvchisq(1,n-p,s2)
## var.beta não eficiente:
## var.beta <- as.numeric(sigma[i])*solve(t(D)%*%inv.V%*%D)
## var.beta eficiente:
var.beta <- as.numeric(sigma[i])*inv.tDinvVD
sigma.media[i,] <- c(mvrnorm(1,mu=beta,Sigma=var.beta),sqrt(sigma[i]))
print(i)
}
retorna <- list()
retorna[[1]] <- data.frame(sigma.media)
names(retorna[[1]]) <- c('beta1','beta2','sigma1')
retorna[[2]] <- data.frame(mh)
names(retorna[[2]]) <- c('eta','nu1','nu2','phi','rho')
retorna[[3]] <- data.frame(inicial.optim$par)
return(retorna)
}

```

Arquivo **resultado.mh.R**:

```

`resul.mh` <-
function(resultado){
  post.beta1 <- resultado[[1]]$beta1
  post.beta2 <- resultado[[1]]$beta2
  post.sigma1 <- resultado[[1]]$sigma1
  post.sigma2 <- post.sigma1*resultado[[2]]$eta
  post.tau1 <- post.sigma1*resultado[[2]]$nu1
  post.tau2 <- post.sigma1*resultado[[2]]$nu2
  post.phi <- resultado[[2]]$phi
  post.rho <- resultado[[2]]$rho
  saida <- data.frame(post.beta1,post.beta2,post.sigma1,post.sigma2,
                     post.tau1,post.tau2,post.phi,post.rho)
  par(mfrow=c(2,4))
  plot(density(post.beta1),main="Posteriori mu1")
  plot(density(post.beta2),main="Posteriori mu2")
  plot(density(post.sigma1),main="Posteriori s1")
  plot(density(post.sigma2),main="Posteriori s2")
  plot(density(post.tau1),main="Posteriori t1")
  plot(density(post.tau2),main="Posteriori t2")
}

```

```

plot(density(post.phi),main="Posteriori phi")
plot(density(post.rho),main="Posteriori rho")
m1<-c(quantile(post.beta1,prob=0.025),mean(post.beta1),
      quantile(post.beta1,prob=0.975))
m2<-c(quantile(post.beta2,prob=0.025),mean(post.beta2),
      quantile(post.beta2,prob=0.975))
s1<-c(quantile(post.sigma1,prob=0.025),mean(post.sigma1),
      quantile(post.sigma1,prob=0.975))
s2<-c(quantile(post.sigma2,prob=0.025),mean(post.sigma2),
      quantile(post.sigma2,prob=0.975))
t1<-c(quantile(post.tau1,prob=0.025),mean(post.tau1),
      quantile(post.tau1,prob=0.975))
t2<-c(quantile(post.tau2,prob=0.025),mean(post.tau2),
      quantile(post.tau2,prob=0.975))
phi<-c(quantile(post.phi,prob=0.025),median(post.phi),
      quantile(post.phi,prob=0.975))
rho<-c(quantile(post.rho,prob=0.025),mean(post.rho),
      quantile(post.rho,prob=0.975))
tabela <- t(data.frame(m1,m2,s1,s2,t1,t2,phi,rho))
estimativa.optim <- estima.bayes[[3]]
retorna <- list()
retorna[[1]] <- saida
retorna[[2]] <- tabela
retorna[[3]] <- estimativa.optim
return(retorna)
}

```

Arquivo **cokri.bayes.R**:

```

`cokri.bayes` <-
function(esti.par,locations,dados.comp){
  nloc <- nrow(locations)*2
  seq.y1 <- seq(1, nloc, by=2)
  seq.y2 <- seq(2, nloc, by=2)
  res <- rowMeans(sapply(1:nrow(esti.par), function(i) {
    print(i)
    temporario <- cokrigagem(esti.par[i,], locations=locations,
                           dados.comp=dados.comp)
    amos.bivar <- mvrnorm(1, mu=temporario[[1]], Sigma=temporario[[2]])
  }

```



```
    acre <- exp(cbind(amos.bivar[seq.y1], amos.bivar[seq.y2], 0))
    return(as.vector(acre/rowSums(acre)))
  ))
  retorna <- list()
  retorna$espe.bayes <- matrix(res, ncol=3)
  retorna$locations <- locations
  return(retorna)
}
```

ANEXO B – Help do pacote geoComp

`agl` *Transformação alr inversa.*

Description

A função transforma um vetor bivariado, de dados `alr` transformados, em um vetor trivariado, uma composição.

Usage

```
agl(dados)
```

Arguments

`dados` Vetor bivariado de dados `alr` transformados.

Value

A função retorna uma composição.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

Examples

```
x<-c(0.5,0.3)
agl(x)
```

<code>as.geoComp</code>	<i>Transforma um data frame em um objeto da classe geoComp.</i>
-------------------------	---

Description

A função faz a transformação `alr` dos dados e cria um objeto da classe `geoComp`.

Usage

```
as.geoComp(original, perc = TRUE)
```

Arguments

<code>original</code>	Data frame contendo os três componentes.
<code>perc</code>	Lógico, indicando se os dados estão expressos em valores percentuais ('perc = TRUE') ou se os dados estão em valores decimais ('perc = FALSE'). O default é 'TRUE'.

Details

As composições são preparadas para análise geoestatística bivariada.

Value

A função retorna uma lista contendo os seguintes elementos:

<code>comp1</code>	Data frame contendo os dados bivariados <code>alr</code> transformados.
<code>comp2</code>	Vetor contendo os dados bivariados <code>alr</code> transformados arranjados de forma intercalada.
<code>comp3</code>	Data frame contendo a localização espacial. Geralmente coordenadas Euclidianas num plano.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

- Aitchison, J. (1986). *The Statistical analysis of compositional data*, The Blackburn Press, New Jersey.
- Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H. (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n.3, 456-477.
- Pawlowsky-Glahn, V., and Olea, R. A. (2004). *Geostatistical Analysis of Compositional Data*, Oxford University Press, Inc., New York.

Examples

```
### Carregando pacote
require(geoComp)
### Carregando os dados
data(pivo)
### Selecionando as colunas referentes aos Componentes 1, 2 e 3,
### Coordenada X e Coordenada Y
dados <- pivo[,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados <- as.geoComp(dados)
```

`classificaSolo` *Classificação de solo.*

Description

A função faz a classificação de solo em classes texturais segundo Reichardt e Timm (2004).

Usage

```
classificaSolo(dados, class.dados = TRUE, plot.dados = TRUE, newClass = TRUE,
               new.comp, plot.new = TRUE)
```

Arguments

<code>dados</code>	Data.frame contendo 3 colunas correspondentes aos percentuais de areia, silte e argila.
<code>class.dados</code>	Lógico, indicando se a função classifica as composições de acordo com a classe textural a que pertencem (<code>'class.dados = TRUE'</code>) ou não (<code>'class.dados = FALSE'</code>). O default é <code>'TRUE'</code> .
<code>plot.dados</code>	Lógico, indicando se a função plota o diagrama de classificação textural incluindo as amostras (<code>'plot.dados = TRUE'</code>) ou não (<code>'plot.dados = FALSE'</code>). O default é <code>'TRUE'</code> .
<code>newClass</code>	Lógico, indicando se a função deve fazer a classificação de um novo vetor (<code>'newClass = TRUE'</code>) ou não (<code>'new.class = FALSE'</code>). O default é <code>'TRUE'</code> .
<code>new.comp</code>	Vetor trivariado correspondente a composição a ser classificada.
<code>plot.new</code>	Lógico, indicando se a função deve plotar o diagrama de classificação textural incluindo a nova composição (<code>'plot.new = TRUE'</code>) ou não (<code>'plot.new = FALSE'</code>). O default é <code>'TRUE'</code> .

Value

A função retorna a classificação de uma amostra, o diagrama de classificação textural incluindo as amostras e a nova composição e a classificação da nova composição.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Reichardt, K. and Timm, L. C.(2004). *Solo, Planta e Atmosfera: Conceitos, Processos e Aplicações*. Manole, Barueri, SP.

Examples

```
require(geoR)
require(geoComp)
require(sp)
data(pivo)
```

```

data(tab.coord.pol)
saida <- classificaSolo(dados=pivo[,c(6,7,8)],class.dados=TRUE,
                        plot.dados=TRUE,newClass=TRUE,
                        new.comp=c(20,10,70),plot.new=TRUE)

```

<code>classSolo</code>	<i>Classificação de uma composição.</i>
------------------------	---

Description

A função faz a classificação de uma composição de acordo com classificação textural apresentada em Reichardt e Timm (2004).

Usage

```
classSolo(x, poligono2D)
```

Arguments

<code>x</code>	Vetor correspondente a uma composição.
<code>poligono2D</code>	Lista contendo as coordenadas dos vértices dos polígonos no diagrama de classificação textural.

Value

A função retorna a classificação de uma composição.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Reichardt, K. and Timm, L. C.(2004). *Solo, Planta e Atmosfera: Conceitos, Processos e Aplicações*. Manole, Barueri, SP.

`cokri.bayes` *Predição espacial bayesiana.*

Description

A função realiza o processo de predição espacial com métodos bayesianos.

Usage

```
cokri.bayes(esti.par, locations, dados.comp)
```

Arguments

`esti.par` Matriz de parâmetros resultante do processo de estimação bayesiana usando o algoritmo Metropolis-Hastings.

`locations` Data frame com as localizações de predição.

`dados.comp` Objeto da classe `geoComp`.

Value

A função retorna uma lista com os seguintes elementos:

`comp1` matriz dos componentes preditos.

`comp2` data frame com as localizações de predição.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

Diggle, P. J., and Ribeiro, Jr., P. J. (2007). *Model-based geostatistics*. Springer Series in Statistics, USA.

Examples

```

### Carregando pacotes
require(geoComp)
require(MCMCpack)
require(geoR)
require(MASS)
### Carregando os dados
data(pivo)
### Selecionando as colunas referentes aos Componentes 1, 2 e 3,
### Coordenada X e Coordenada Y
dados <- pivo[,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados <- as.geoComp(dados)
### Fazendo estimação Bayesiana
estima.bayes <- infbayes(dados,tune=c(0.5,0.5,0.5,0.5,0.5),queima=1,salto=5,
                        nsim=20,mediahiper=c())
summary(estima.bayes[[2]])
### Organizando os resultados
res.mh <- resul.mh(resultado=estima.bayes)
res.mh.ic <- data.frame(res.mh[[2]])
names(res.mh.ic) <- c('2.5%', 'media', '97.5%')
metro.bayes <- data.frame(res.mh[[1]])
### Construindo a borda da área de predição
bor <- cbind(c(0,seq(0,200,l=100),0),c(0,sqrt(200^2-seq(0,200,l=100)^2),0))
### Construindo uma grade de predição
gr <- pred_grid(bor, by=70)
### Fazendo predição espacial
cokri.bayes <- cokri.bayes(estim.par=metro.bayes,locations=gr,
                          dados.comp=dados)

```

cokrigagem

Predição espacial dos dados alr transformados.

Description

A função calcula o vetor de valores esperados preditos e a respectiva matriz de covariância para os dados `alr` transformados utilizando os resultados da distribuição Gaussiana multi-

variada.

Usage

```
cokrigagem(estim.par, dim.gride, locations, dados.comp)
```

Arguments

`estim.par` Vetor das estimativas dos parâmetros calculados através da função `mec`.

`dim.gride` Vetor numérico com o valor da dimensão da grade de predição.

`locations` Data frame contendo as coordenadas das localizações de predição.

`dados.comp` Um objeto da classe `geoComp`.

Details

A predição espacial de Y_0 em localizações não amostradas $x_0 = (x_{10}, x_{20}, \dots, x_{n_20})$ é realizada segundo resultados da distribuição Gaussiana multivariada. O vetor de valores esperados nas n_2 localizações é dado por

$$\mu_{Y_0|Y} = \mu_{Y_0} + \Sigma_{Y_0Y} \Sigma_{YY}^{-1} (Y - \mu_Y)$$

e a matriz de variância/covariância é dada por

$$\Sigma_{Y_0|Y} = \Sigma_{Y_0Y_0} - \Sigma_{Y_0Y} \Sigma_{YY}^{-1} \Sigma_{YY_0}.$$

Value

A função retorna uma lista com os seguintes resultados:

`comp1` Vetor dos valores esperados preditos.

`comp2` Matriz de covariância predita.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Diggle, P. J., and Ribeiro, Jr., P. J., (2007). *Model-based geostatistics*. Springer Series in Statistics, USA.

Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

Examples

```
### carregando pacotes
require(geoComp)
require(MASS)
require(statmod)
require(geoR)
### Carregando os dados
data(pivo)
### Selecionando as colunas referentes aos componentes 1, 2 e 3,
### Coordenada X e Coordenada Y
dados <- pivo[,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados <- as.geoComp(dados)
### Construindo a borda da área de predição
bor <- cbind(c(0,seq(0,200,l=100),0),c(0,sqrt(200^2-seq(0,200,l=100)^2),0))
### Construindo uma grade de predição
gr <- pred_grid(bor, by=70)
### Ajustando o modelo
estima <- mec(dados)
### Obtendo os valores preditos no R^2
md.cov.ck <- cokrigagem(estima[[1]]$Estimativas, loc=gr,dados.comp=dados)
```

geoComp-internal *Funções internas do pacote geoComp.*

Description

Estas são funções internamente chamadas por outras funções no pacote **geoComp** e não precisam ser chamadas pelo usuário.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

geoComp-package *Análise geoestatística de dados composicionais.*

Description

A função ajusta um modelo geoestatístico bivariado para dados composicionais com três componentes, tanto pelo método clássico como por inferência bayesiana.

Details

Package: geoComp
Type: Package
Version: 1.0
Date: 2009-12-01
License: What license is it under?
LazyLoad: yes

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H. (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

<code>infbayes</code>	<i>Ajuste do modelo por inferência bayesiana para dados composicionais.</i>
-----------------------	---

Description

A função ajusta um modelo para dados composicionais por métodos Bayesianos. Utiliza-se o algoritmo Metropolis-Hastings, função `MCMCmetrop1R` do pacote **MCMCpack** para a estimação dos parâmetros η , ν_1 , ν_2 , ϕ e ρ .

Usage

```
infbayes(dados, tune, queima, salto, nsim, verbose = TRUE, mediahiper)
```

Arguments

<code>dados</code>	Objeto da classe <code>geoComp</code> .
<code>tune</code>	Vetor com 5 elementos iguais aos limites de variâncias para η , ν_1 , ν_2 , ϕ e ρ .
<code>queima</code>	Vetor numérico correspondente ao número de simulações a serem descartadas como período de aquecimento da cadeia.
<code>salto</code>	Vetor numérico que indica de quanto em quanto os dados são armazenados.
<code>nsim</code>	Vetor numérico correspondente ao número de simulações.
<code>verbose</code>	Lógico, indicando se a mensagem de que a log-verossimilhança está sendo otimizada deve aparecer na tela ('verbose = TRUE') ou não ('verbose = FALSE'). O default é 'TRUE'.
<code>mediahiper</code>	Vetor com 8 elementos correspondentes aos valores dos hiperparâmetros das prioris log-normais para η , ν_1 , ν_2 e Gama para ϕ . Os três primeiros correspondem as médias para η , ν_1 , ν_2 , o quarto é o hiperparâmetro α para ϕ , os três seguintes correspondem aos desvios-padrão para η , ν_1 , ν_2 e o último, hiperparâmetro β para ϕ . O default é $(\log(24), \log(0.63), \log(0.59), 66, 0.3, 0.3, 0.3, 1)$.

Details

A distribuição a posteriori marginal para $\underline{\mu} | \sigma_1^2, \theta^*, \underline{Y}$ é

$$[\underline{\mu} | \sigma_1^2, \theta^*, \underline{Y}] \sim N(\hat{\underline{\mu}}; \sigma_1^2 (D'V^{-1}D)^{-1}),$$

onde $\underline{\hat{\mu}}$ é dado em `mec.R` e para $\sigma_1^2 | \theta^*, Y$ é

$$[\sigma_1^2 | \theta^*, Y] \sim \chi_{Simv}^2(n - n_{\underline{\mu}}; S^2),$$

onde χ_{Simv}^2 é a distribuição qui-quadrado escalonada inversa e

$$S^2 = \frac{n}{n - n_{\underline{\mu}}} \hat{\sigma}_1^2 = \frac{(\underline{Y} - D\underline{\hat{\mu}})'V^{-1}(\underline{Y} - D\underline{\hat{\mu}})}{n - n_{\underline{\mu}}}.$$

Value

A função retorna uma lista com os seguintes elementos:

- `comp1` data frame contendo os valores a posteriori de β_1 , β_2 e σ_1 .
- `comp2` data frame contendo os valores a posteriori de η , v_1 , v_2 , ϕ e ρ .

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

- Diggle, P. J., and Ribeiro, Jr., P. J., (2007). *Model-based geostatistics*. Springer Series in Statistics, USA.
- Gamerman, D., and Lopes, H. F., (2006). *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. 2.ed. Chapman and Hall/CRC, Londres.
- Gelman, A., and Carlin, J. B., and Stern, H. S., and Rubin, D. B., (2003). *Bayesian Data Analysis*. 2.ed. Chapman and Hall, Boca Raton.
- Gilks, W. R., and Richardson, S., and Spiegelhalter, D. J., (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London.
- Gill, J., (2000). *Bayesian Methods For The Social And Behavioral Sciences*. Chapman and Hall, London.
- Lee, P. M., (2004). *Bayesian Statistics*. 3.ed. Arnold, New York.
- Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

Ribeiro, Jr, P. J. and Diggle, P. J., (1999). *Bayesian inference in Gaussian model-based geostatistics*. ST-99-08., Dept. of Maths and Stats, Lancaster University, Lancaster, Technical Report.

Rue, H., and Held, L., (2005). *Gaussian Markov Random Fields*. Chapman and Hall, Boca Raton.

Examples

```
### Carregando pacotes
require(geoR)
require(geoComp)
require(MCMCpack)
require(MASS)
### Carregando os dados
data(pivo)
### Selecionando as colunas referentes aos Componentes 1, 2 e 3,
### Coordenada X e Coordenada Y
dados <- pivo[,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados <- as.geoComp(dados)
### Fazendo o ajuste do modelo por métodos Bayesianos
estima.bayes <- infbayes(dados,tune=c(0.5,0.5,0.5,0.5,0.5),queima=1,salto=5,
                        nsim=20,mediahiper=c())
```

logarit.vero *Log-Verossimilhança.*

Description

A função calcula a log-verossimilhança do modelo.

Usage

```
logarit.vero(theta1, dados.comp, print.pars = FALSE)
```

Arguments

- `theta1` Vetor com valores para os parâmetros η , v_1 , v_2 , ϕ e ρ .
- `dados.comp` Objeto da classe `geoComp`.
- `print.pars` Lógico. Se 'TRUE', indica que os parâmetros e o valor do negativo da log-verossimilhança (a menos de uma constante) são printadas na tela cada vez que a função a ser minimizada for chamada.

Details

A função de verossimilhança do modelo é

$$L(\theta; \underline{Y}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2}(\underline{Y} - \underline{\mu}_Y)' \Sigma^{-1} (\underline{Y} - \underline{\mu}_Y)\right\},$$

com vetor de parâmetros $\theta = (\underline{\mu}, \sigma_1, \sigma_2, \tau_1, \tau_2, \phi, \rho)$. A função de log-verossimilhança com a reparametrização $\eta = \frac{\sigma_2}{\sigma_1}$, $v_1 = \frac{\tau_1}{\sigma_1}$ e $v_2 = \frac{\tau_2}{\sigma_1}$ é

$$l(\theta; \underline{Y}) = -\frac{1}{2}(n \ln(2\pi) + 2n \ln(\sigma_1) + \ln(|V|) + \frac{1}{\sigma_1^2} Qe).$$

com $Qe = (\underline{Y} - \underline{\mu}_Y)' V^{-1} (\underline{Y} - \underline{\mu}_Y)$ e $\theta^* = (\eta, v_1, v_2, \phi, \rho)$.

Value

A função retorna o negativo do valor da log-verossimilhança do modelo.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

Examples

```
### Carregando pacote
require(geoComp)
```

```

### Carregando dados
data(pivo)
### Selecionando dados
dados <- pivo[1:15,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados.comp <- as.geoComp(dados)
y1 <- dados.comp[[1]]$Y1
y2 <- dados.comp[[1]]$Y2
coords <- dados.comp[[3]]
### Calculando os valores iniciais
mu1 <- mean(y1)
mu2 <- mean(y2)
var_y1 <- var(y1)
s1 <- var_y1/2
tau1 <- s1
var_y2 <- var(y2)
s2 <- var_y2/2
tau2 <- s2
dim <- range(dist(coords))
phi <- dim[1] + 0.2*(dim[2]-dim[1])
rho <- cor(y1,y2)
y <- dados.comp[[2]][[1]]
### Reparametrizando
eta <- s2/s1
nu1 <- tau1/s1
nu2 <- tau2/s1
theta1 <- c(eta,nu1,nu2,phi,rho)
### Calculando a matriz de covariância
V <- monta.V(theta1,dados.comp)
### Calculando o logaritmo da vero. para o theta1 inicial
ll <- logarit.vero(theta1,dados.comp)

```


Description

A função ajusta um modelo espacial baseado em verossimilhança para dados composicionais.

Usage

```
mec(dados.comp, metodo = "L-BFGS-B", print.pars = FALSE, alpha = 0.975)
```

Arguments

<code>dados.comp</code>	Objeto da classe <code>geoComp</code> .
<code>metodo</code>	Algoritmo utilizado na função <code>optim</code> . Outras opções: “Nelder-Mead”, “CG” e “BFGS”.
<code>print.pars</code>	Lógico. Se ‘TRUE’, indica que os parâmetros e o valor do negativo da log-verossimilhança (a menos de uma constante) são printadas na tela cada vez que a função a ser minimizada for chamada. O default é ‘FALSE’
<code>alpha</code>	Vetor numérico correspondente ao nível crítico utilizado no cálculo do intervalo de confiança.

Details

Esta função estima os parâmetros do modelo de campo aleatório Gaussiano especificado como:

$$Y_1(x) = \mu_1(x) + S_1(x) + Z_1(x)$$

$$Y_2(x) = \mu_2(x) + S_2(x) + Z_2(x)$$

onde

- x define uma localização espacial. Geralmente coordenadas Euclidianas num plano.
- Y_1 e Y_2 são as variáveis observadas na localização x e constituem o vetor Y .
- $\underline{\mu}(x) = D\underline{\mu}$ é o componente média do modelo (tendência).
- $S(x)$ é um processo Gaussiano estacionário com variância σ^2 (patamar) e uma função de correlação exponencial com parâmetro de alcance ϕ .
- $Z(x)$ é o efeito composicional.

A função de verossimilhança do modelo é

$$L(\theta; Y) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\left\{-\frac{1}{2} (Y - \underline{\mu}_Y)' \Sigma^{-1} (Y - \underline{\mu}_Y)\right\}.$$

Fazendo-se a reparametrização: $\eta = \frac{\sigma_2}{\sigma_1}$; $\nu_1 = \frac{\tau_1}{\sigma_1}$; $\nu_2 = \frac{\tau_2}{\sigma_1}$, pode-se escrever

$$\Sigma = \sigma_1^2 R + \tau_1^2 I_b = \sigma_1^2 V,$$

e a função de log-verossimilhança reparametrizada é dada por

$$l(\theta; \underline{Y}) = -(1/2)(n \ln(2\pi) + 2n \ln(\sigma_1) + \ln(|V|) + (1/\sigma_1^2) Qe).$$

onde $Qe = (\underline{Y} - \underline{\mu}_{\underline{Y}})' V^{-1} (\underline{Y} - \underline{\mu}_{\underline{Y}})$. Os estimadores de máxima verossimilhança são:

$$\hat{\underline{\mu}} = (D' V^{-1} D)^{-1} (D' V^{-1} \underline{Y})$$

e

$$\hat{\sigma}_1 = \sqrt{\hat{Q}e/n}.$$

As variâncias dos estimadores obtidos pelo método Delta são:

$$Var(\hat{\underline{\mu}}) = \hat{\sigma}_1^2 (D' \hat{V}^{-1} D)^{-1}$$

e

$$Var(\hat{\sigma}_1) = (\hat{\sigma}_1^3 / (3\hat{Q}e - n\hat{\sigma}_1)).$$

As estimativas para os parâmetros σ^2 , τ^1 , τ^2 , ϕ e ρ são obtidas numericamente.

Value

A função retorna uma lista com os dois seguintes objetos:

- | | |
|--------------------|---|
| <code>comp1</code> | estimativas dos parâmetros, erros-padrão e intervalos de 95% de confiança via método Delta. |
| <code>comp2</code> | negativo da log-verossimilhança inicial e otimizada. |

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,

Ana Beatriz Tozzo Martins <abtmartins@uem.br>,

Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

- Azzalini, A., (1996). *Statistical Inference Based on the Likelihood*. Chapman and Hall, London.
- Cox, D. R., and Hinkley, D. V., (1974). *Theoretical Statistics*. Chapman and Hall, London.
- DeGroot, M. H. and Schervish, M. J. (2002). *Probability and Statistics*. 3.ed., Addison Wesley, Boston.
- Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.
- Pawitan, Y., (2001). *In All Likelihood*. Oxford University Press, Inc, New York.

Examples

```
### Carregando pacote
require(geoComp)
### Carregando os dados
data(pivo)
### Selecionando as colunas referentes aos Componentes 1, 2 e 3,
### Coordenada X e Coordenada Y
dados <- pivo[,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados <- as.geoComp(dados)
### Fazendo o ajuste do modelo
estima <- mec(dados)
```

monta.V *Matriz de covariância.*

Description

A função constrói a matriz de covariância do modelo bivariado onde as variáveis são dispostas de forma intercalada.

Usage

```
monta.V(theta1, dados.comp)
```

Arguments

`theta1` Vetor com valores para os parâmetros η , v_1 , v_2 , ϕ e ρ .
`dados.comp` Objeto da classe `geoComp`.

Value

A função retorna a matriz de covariância do modelo.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

Examples

```
### Carregando pacote
require(geoComp)
### Carregando dados
data(pivo)
### Selecionando dados
dados <- pivo[1:15,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados.comp <- as.geoComp(dados)
y1 <- dados.comp[[1]]$Y1
y2 <- dados.comp[[1]]$Y2
coords <- dados.comp[[3]]
### Calculando os valores iniciais
mu1 <- mean(y1)
mu2 <- mean(y2)
var_y1 <- var(y1)
s1 <- var_y1/2
tau1 <- s1
var_y2 <- var(y2)
```

```

s2 <- var_y2/2
tau2 <- s2
dim <- range(dist(coords))
phi <- dim[1] + 0.2*(dim[2]-dim[1])
rho <- cor(y1,y2)
y <- dados.comp[[2]][[1]]
### Reparametrizando
eta <- s2/s1
nu1 <- tau1/s1
nu2 <- tau2/s1
theta1 <- c(eta,nu1,nu2,phi,rho)
### Calculando a matriz de covariância
V <-monta.V(theta1,dados.comp)

```

pivo

Propriedades do solo irrigado por um sistema de pivo.

Description

Propriedades do solo irrigado por um sistema de pivo.

Usage

```
data(pivo)
```

Format

Um data frame contendo 76 observações das seguintes 8 variáveis:

Xum vetor numérico com as coordenadas X das observações.

Yum vetor numérico com as coordenadas Y das observações.

DenSoilnumérico, densidade do solo g/cm^3 .

DenParnumérico, densidade da partícula (%).

Porototnumérico, porosidade total (%).

Areianumérico, fração do conteúdo de areia (%).

Siltenumérico, fração do conteúdo de silte (%).

Argilanumérico, fração do conteúdo de argila (%).

Details

A densidade do solo é definida como:

$$DenSoil = \frac{m_{ss}}{V_T}$$

onde m_{ss} é a massa de solo seco e V_T é o volume total.

$$V_T = V_s + V_p,$$

sendo V_s o volume de sólidos e V_p o volume de poros. A densidade da partícula é definida como:

$$DenPar = \frac{m_{ss}}{V_s}$$

e a porosidade total é como:

$$Porotot = \frac{V_p}{V_T}.$$

Source

Gonçalves, A. C. A., (1997). *Variabilidade espacial de propriedades físicas do solo para fins de manejo da irrigação*. 119p., Tese (Doutorado em Agronomia) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba.

References

Gonçalves, A. C. A., and Folegatti, M. V., and Mata, J. D. V., (2001). Análises exploratória e geoestatística da variabilidade de propriedades físicas de um Argissolo Vermelho. *Acta Scientiarum* **23**, n. 5, 1149-1157.

Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

Examples

```
### Carregando pacotes
require(compositions)
require(geoComp)
### Lendo um arquivos de dados
data(pivo)
### Selecionado os dados composicionais
dados <- pivo[,c(6,7,8)]
```

```

### Escala representada na classe acomp onde os dados são analisados
### na geometria relativa
comp <- acomp(dados)
### Construindo um diagrama ternário
plot(comp)
### Incluindo o centro da distribuição
plot(mean(comp),add=TRUE,pch=20,col='red')
### Incluindo as regiões 2 e 4-desvios-padrão de confiança
ellipses(mean(comp),var(comp),col='red',r=2)
ellipses(mean(comp),var(comp),col='red',r=4)

```

posteriori

Posteriori do modelo bayesiano para dados composicionais espaciais.

Description

A função calcula a posteriori do modelo geoestatístico bivariado para dados composicionais.

Usage

```
posteriori(theta1, dados.comp, mediahiper, print.pars = TRUE)
```

Arguments

<code>theta1</code>	Vetor cujos elementos são os valores para os parâmetros η , v_1 , v_2 , ϕ e ρ .
<code>dados.comp</code>	Objeto da classe <code>geoComp</code> .
<code>mediahiper</code>	Vetor com 8 elementos correspondentes aos valores dos hiperparâmetros das prioris log-normais para η , v_1 , v_2 e Gama para ϕ . Os três primeiros correspondem as médias para η , v_1 , v_2 , o quarto é o hiperparâmetro α para ϕ , os três seguintes correspondem aos desvios-padrão para η , v_1 , v_2 e o último, hiperparâmetro β para ϕ . O default é $(\log(24), \log(0.63), \log(0.59), 66, 0.3, 0.3, 0.3, 1)$.
<code>print.pars</code>	Lógico, indicando se a função retorna o vetor <code>theta1</code> e o valor da posteriori (<code>'print.pars = TRUE'</code>) ou apenas o valor da posteriori (<code>'print.pars = FALSE'</code>).

Details

posteriori=vero x priori(η) x priori(v_1) x priori(v_2) x priori(ϕ) x 1

Value

A função retorna o valor da posteriori e o vetor `theta1`.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

Examples

```
### Carregando pacotes
require(geoComp)
require(MCMCpack)
require(MASS)
### Carregando os dados
data(pivo)
### Selecionando as colunas referentes aos Componentes 1, 2 e 3,
### Coordenada X e Coordenada Y
dados <- pivo[,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados <- as.geoComp(dados)
### Calculando os valores iniciais
y1 <- dados[[1]]$Y1
y2 <- dados[[1]]$Y2
coords <- dados[[3]]
mu1 <- mean(y1)
mu2 <- mean(y2)
var_y1 <- var(y1)
s1 <- var_y1/2
```



```

tau1 <- s1
var_y2 <- var(y2)
s2 <- var_y2/2
tau2 <- s2
dim <- range(dist(coords))
phi <- dim[1] + 0.2*(dim[2]-dim[1])
rho <- cor(y1,y2)
eta <- s2/s1
nu1 <- tau1/s1
nu2 <- tau2/s1
theta1 <- c(eta,nu1,nu2,phi,rho)
### Calculando a posteriori
posteriori(theta1,dados.comp=dados,mediahiper=c(),print.pars=TRUE)

```

resul.mh

Resultados da inferência Bayesiana.

Description

A função organiza os resultados obtidos através do algoritmo Metropolis-Hastings e calcula os intervalos de 95% de credibilidade.

Usage

```
resul.mh(resultado)
```

Arguments

resultado Lista contendo os resultados obtidos através da função `infbayes`.

Value

A função retorna uma lista contendo os seguintes elementos:

- comp1** data frame contendo as posteriores para β_1 , β_2 , σ_1 , σ_2 , τ_1 , τ_2 , ϕ e ρ .
- comp2** data frame contendo as esperanças a posteriori e intervalos de 95% de credibilidade para os parâmetros do modelo.
- comp3** data frame contendo as estimativas obtidas no processo de otimização.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

Examples

```
### Carregando pacotes
require(geor)
require(geoComp)
require(MCMCpack)
require(MASS)
### Carregando os dados
data(pivo)
### Selecionando as colunas referentes aos Componentes 1, 2 e 3,
### Coordenada X e Coordenada Y
dados <- pivo[,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados <- as.geoComp(dados)
### Fazendo o ajuste do modelo por métodos Bayesianos
estima.bayes <- infbayes(dados,tune=c(0.5,0.5,0.5,0.5,0.5),queima=1,salto=5,
                        nsim=20,mediahiper=c())
summary(estima.bayes[[2]])
### Organizando os resultados
res.mh <- resul.mh(resultado=estima.bayes)
res.mh.ic <- data.frame(res.mh[[2]])
names(res.mh.ic) <- c('2.5%', 'media', '97.5%')
```

Description

Coordenadas dos vértices no triângulo de classificação textural.

Usage

```
data(tab.coord.pol)
```

Format

Um data frame contendo 69 observações das seguintes 4 variáveis.

Areia numérico, fração do conteúdo de areia (%).

Silte numérico, fração do conteúdo de silte (%).

Argila numérico, fração do conteúdo de argila (%).

Poligonoum vetor numérico.

Examples

```
data(tab.coord.pol)
## maybe str(tab.coord.pol) ; plot(tab.coord.pol) ...
```

`volta.cokri` *Transformação de volta para o simplex por simulação.*

Description

A função transforma o vetor de valores esperados e matriz de covariância preditos no R^2 para o simplex utilizando simulação.

Usage

```
volta.cokri(mat.cokri, num.simu, retorna.tudo = FALSE, int.conf = 0.95)
```

Arguments

- `mat.cokri` Uma lista cujo primeiro elemento contém o vetor de valores esperados preditos e segundo elemento contém a matriz de covariância predita.
- `num.simu` Vetor numérico correspondente ao número de simulações.
- `retorna.tudo` Lógico, indicando que se a função deve retornar, separadamente, os valores esperados preditos por simulação para cada componente (`'retorna.tudo = TRUE'`) ou a função deve calcular as médias do número de simulações para cada localização e para cada componente os intervalos de 95% de confiança (`'retorna.tudo = FALSE'`).
- `int.conf` Vetor numérico correspondente ao nível de confiança.

Details

Os dados são simulados de uma distribuição Gaussiana multivariada com vetor de médias e matriz de covariância dados em `mat.cokri` e em seguida é aplicada a função `agl` transformando-os em composições.

Value

Se `retorna.tudo=TRUE`, a função retorna uma lista com os três primeiros elementos a seguir e, se `retorna.tudo=FALSE`, a função retorna uma lista com os dois últimos:

- `comp1` Valores esperados preditos simulados para cada localização de predição do componente 1.
- `comp2` Valores esperados preditos simulados para cada localização de predição do componente 2.
- `comp3` Valores esperados preditos simulados para cada localização de predição do componente 3.
- `comp4` Um data frame com 3 colunas correspondentes as médias do número de simulações para cada localização e para cada componente.
- `comp5` Um data frame contendo os limites dos intervalos de confiança.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.

Diggle, P. J., and Ribeiro, Jr., P. J., (2007). *Model-based geostatistics*. Springer Series in Statistics, USA.

Examples

```
### carregando pacotes
require(geoComp)
require(MASS)
require(geoR)
### Carregando os dados
data(pivo)
### Selecionando as colunas referentes aos componentes 1, 2 e 3,
### Coordenada X e Coordenada Y
dados <- pivo[,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados <- as.geoComp(dados)
### Construindo a borda da área de predição
bor <- cbind(c(0,seq(0,200,l=100),0),c(0,sqrt(200^2-seq(0,200,l=100)^2),0))
### Construindo uma grade de predição
gr <- pred_grid(bor, by=70)
### Ajustando o modelo
estima <- mec(dados)
### Obtendo os valores preditos no R^2
md.cov.ck <- cokrigagem(estima[[1]]$Estimativas, loc=gr,dados.comp=dados)
### Obtendo os valores preditos no simplex por simulação
preditos.simu <- volta.cokri(md.cov.ck,num.simu=10,int.conf=0.95)
preditos.simu <- data.frame(preditos.simu[[1]])
preditos.simu.ic <- data.frame(preditos.simu[[2]])
```

volta.quad

Transformação de volta para o simplex por quadratura de Gauss-Hermite.

Description

A função transforma o vetor de valores esperados e matriz de covariância preditos no R^2 para o simplex utilizando quadratura Gaussiana.

Usage

```
volta.quad(med.cov, n.pontos = 7, Variancia = FALSE)
```

Arguments

<code>med.cov</code>	Uma lista cujo primeiro elemento contém o vetor de valores esperados preditos e segundo elemento contém a matriz de covariância predita.
<code>n.pontos</code>	Vetor numérico correspondente a ordem da quadratura de Gauss-Hermite.
<code>Variancia</code>	Lógico, indicando se a função não deve calcular a matriz de covariância no espaço amostral simplex ('Variancia = FALSE') ou se deve calcular ('Variancia = TRUEf')

Details

O objetivo é calcular para cada localização uma estimativa de

$$\mu_X = \int_{S^B} X f(X) dX$$

e

$$\Sigma_X = \int_{S^B} (X - \mu_X)(X - \mu_X)' f(X) dX.$$

De acordo com Aitchison (1986), Pawlowsky-Glahn e Olea (2004) as integrais são expressas como

$$\mu_X = \int_{R^{B-1}} g_1(Z) f(-Z'Z) dZ$$

e

$$\Sigma_X = \int_{R^{B-1}} g_2(Z) f(-Z'Z) dZ$$

em que Z é a transformação

$$Z = \frac{1}{\sqrt{2}}(R')^{-1}(alr(X - \mu_Y));$$

R é a decomposição Cholesky de Σ_Y , uma matriz triangular superior;

$$g_1(Z) = \pi^{-\frac{B-1}{2}} agl(\mu_Y + \sqrt{2}R'Z).$$

e

$$g_2(\mathbf{Z}) = \pi^{-\frac{B-1}{2}} (\text{agl}(\boldsymbol{\mu}_Y + \sqrt{2}\mathbf{R}'\mathbf{Z}) - \boldsymbol{\mu}_X)(\text{agl}(\boldsymbol{\mu}_Y + \sqrt{2}\mathbf{R}'\mathbf{Z}) - \boldsymbol{\mu}_X)'$$

Desta forma, as integrais são aproximadas pela integração de Gauss-Hermite multivariada de ordem k :

$$\int_{\mathbb{R}^{B-1}} g(\mathbf{Z})f(-\mathbf{Z}'\mathbf{Z})d\mathbf{Z} \approx \sum_{i_1=1}^k \sum_{i_2=1}^k \cdots \sum_{i_{B-1}=1}^k \boldsymbol{\omega}_{i_1} \boldsymbol{\omega}_{i_2} \cdots \boldsymbol{\omega}_{i_{B-1}} g(\mathbf{Z}_{i_1}, \mathbf{Z}_{i_2}, \dots, \mathbf{Z}_{i_{B-1}}),$$

em que os pesos $\boldsymbol{\omega}_{i_1} \boldsymbol{\omega}_{i_2} \cdots \boldsymbol{\omega}_{i_{B-1}}$ e as abscissas $\mathbf{Z}_{i_1}, \mathbf{Z}_{i_2}, \dots, \mathbf{Z}_{i_{B-1}}$ podem ser encontrados, por exemplo, em Abramowitz e Stegun (1992) ou calculados usando a função `gauss.quad` do pacote `statmod`. As aproximações resultam no vetor de valores esperados preditos e na matriz de covariância predita no simplex.

Value

A função retorna uma lista com os seguintes elementos:

- `comp1` Valores esperados preditos dos componentes 1, 2 e 3 no simplex.
- `comp2` Matriz de covariância predita dos componentes 1, 2 e 3 no simplex.

Author(s)

Wagner Hugo Bonat <wagner@leg.ufpr.br>,
 Ana Beatriz Tozzo Martins <abtmartins@uem.br>,
 Paulo Justiniano Ribeiro Jr. <paulojus@leg.ufpr.br>.

References

- Abramowitz, M., and Stegun, I. A., (1972). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Milton Abramowitz and Irene A. Stegun, Washington.
- Aitchison, J., (1986). *The Statistical analysis of compositional data*. The Blackburn Press, New Jersey.
- Lark, R. M, and Bishop, T. F. A., (2007). Cokriging particle size fractions of the soil. *European Journal of Soil Science*, june, **58**, 763-774.
- Martins, A. B. T., and Ribeiro Jr, P. J., and Bonat, W. H., (2009). Um modelo geoestatístico para dados composicionais. *Revista Brasileira de Biometria* **27**, n. 3, 456-477.
- Paulino, C. D., and Turkman, M. A. A., and Murteira, B., (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian, Lisboa.

Pawłowsky-Glahn, V., and Olea, R. A., (2004). *Geostatistical Analysis of Compositional Data*. Oxford University Press, Inc, New York.

Examples

```
### carregando pacotes
require(geoComp)
require(MASS)
require(statmod)
require(geoR)
### Carregando os dados
data(pivo)
### Selecionando as colunas referentes aos Componentes 1, 2 e 3,
### Coordenada X e Coordenada Y
dados <- pivo[,c(6,7,8,1,2)]
### Transformando o objeto "dados" para a classe geoComp
dados <- as.geoComp(dados)
### Construindo a borda da área de predição
bor <- cbind(c(0,seq(0,200,l=100),0),c(0,sqrt(200^2-seq(0,200,l=100)^2),0))
### Construindo uma grade de predição
gr <- pred_grid(bor, by=70)
### Ajustando o modelo
estima <- mec(dados)
### Obtendo os valores preditos no R^2
md.cov.ck <- cokrigagem(estima[[1]]$Estimativas, loc=gr,dados.comp=dados)
### Obtendo os valores preditos no simplex por quadratura Gaussiana
preditos.gh <- volta.quad(md.cov.ck,n.pontos=6,Variancia=FALSE)
preditos.gh <- data.frame(preditos.gh)
```


ANEXO C – Código fonte R das análises estatísticas

Scripts para o estudo de simulação:

Arquivo **simula.R**

```
## Função para simular um conjunto de dados
simu.dados <- function(theta,n.simul){
  mu1 <- theta[1]
  mu2 <- theta[2]
  s1 <- theta[3]
  s2 <- theta[4]
  tau1 <- theta[5]
  tau2 <- theta[6]
  phi <- theta[7]
  rho <- theta[8]

  ## Gerando erros de uma normal bivariada com mu=0,
  ## variâncias tau1, tau2 e correlação rho
  v <- matrix(c(tau1^2,tau1*tau2*rho,tau1*tau2*rho,tau2^2),2,2)
  Sc <- kronecker(diag(n.simul/2),v)
  set.seed(3333)
  e <- mvrnorm(n.simul,mu=c(0,0),v)
  ## Gerando um campo aleatório Gaussiano com média 0, variancia 1 e
  ## parametros de alcance phi
  set.seed(3333)
  U <- grf(n.simul,grid='irreg',cov.model='exp',cov.pars=c(1,phi),
          mean=0,messages=FALSE)
  S1 <- s1*U$data
  S2 <- s2*U$data
  ## Calculando e Organizando os dados
  y1 <- as.geodata(data.frame(U$coords[,1],U$coords[,2],mu1+S1+e[,1]))
  y2 <- as.geodata(data.frame(U$coords[,1],U$coords[,2],mu2+S2+e[,2]))
```

```

seq1 <- seq(1,length(y1$data)*2,by=2)
seq2 <- seq(2,length(y2$data)*2,by=2)
y <- c()
y[seq1] <- y1$data
y[seq2] <- y2$data
coords <- data.frame(U$coords[,1],U$coords[,2])
## Preparando a saída
retorna <- list()
retorna[1][[1]] <- y1$data
retorna[2][[1]] <- y2$data
retorna[3][[1]] <- y
retorna[4][[1]] <- coords
return(retorna)
}

```

Arquivo branco1.R

```

## Script rodado para a 1a configuração de diagrama ternário
## Gera valores preditos por quadratura e simulação pra fazer mapas
require(MASS)
require(geoR)
require(geoComp)
require(compositions)
require(statmod)
source("simula.R")
## Função que gera diagrama ternário
gerater <- function(theta,n.simul=100){
  dados <- simu.dados(theta,n.simul=n.simul)
  y1 <- dados[1][[1]]
  y2 <- dados[2][[1]]
  y12 <- data.frame(y1,y2)
  comp <- agl(y12)
  comp1 <- acomp(comp)
  names(comp1) <- c("Areia", "Silte", "Argila")
  plot(comp1)
  plot(mean(comp1),add=T,pch=20,col='red')
  ellipses(mean(comp1),var(comp1),col='red',r=2)
  ellipses(mean(comp1),var(comp1),col='red',r=4)
  retorna <- data.frame(comp,dados[[4]][,1],dados[[4]][,2])
}

```

```

    return(retorna)
}
##### 1a configuração #####
theta <- c(-0.2,-0.5,1,1.5,0.3,0.3,0.6,0.9)
set.seed(1333)
dados1 <- gerater(theta,n.simul=100)
names(dados1) <- c("Areia","Silte","Argila","Coord.X", "Coord.Y")
write.table(dados1,"dad_conf1.txt")
dados1 <- as.geoComp(dados1)
write.table(dados1[[3]],"points1.txt")
estima <- mec(dados1)
write.table(estima[[1]],"estima1par.txt")
write.table(estima[[2]],"estima1log.txt")
borda <- cbind(c(0,0,1,1,0),c(0,1,1,0,0))
gr <- pred_grid(borda, by=0.03)
source("cokrigagem.R")
md.cov.ck <- cokrigagem(estima[[1]]$Estimativas,loc=gr,
                        dados.comp=dados1)
preditos.gh <- volta.quad(md.cov.ck,n.pontos=7,Variancia=FALSE)
write.table(preditos.gh,"dt1predgh.txt")
preditos.simu <- volta.cokri(md.cov.ck,num.simu=150,int.conf=0.95)
write.table(preditos.simu,"dt1predsimu.txt")

```

Arquivo **branco2.R**:

```

## Script rodado para a 2a configuração de diagrama ternário
## Gera valores preditos por quadratura e simulação pra fazer mapas
require(MASS)
require(geoR)
require(geoComp)
require(compositions)
require(statmod)
source("simula.R")
## Função que gera diagrama ternário
gerater <- function(theta,n.simul=100){
  dados <- simu.dados(theta,n.simul=n.simul)
  y1 <- dados[1][[1]]
  y2 <- dados[2][[1]]
  y12 <- data.frame(y1,y2)

```

```

comp <- agl(y12)
comp1 <- acomp(comp)
names(comp1) <- c("Areia", "Silte", "Argila")
plot(comp1)
plot(mean(comp1),add=T,pch=20,col='red')
ellipses(mean(comp1),var(comp1),col='red',r=2)
ellipses(mean(comp1),var(comp1),col='red',r=4)
retorna <- data.frame(comp,dados[[4]][,1],dados[[4]][,2])
return(retorna)
}
##### 2a configuração #####
theta <- c(1,1,1.2,1.5,0.9,1,0.6,0.5)
set.seed(1333)
dados1 <- gerater(theta,n.simul=100)
names(dados1) <- c("Areia","Silte","Argila","Coord.X", "Coord.Y")
#write.table(dados1,"dad_conf2.txt")
dados1 <- as.geoComp(dados1)
#write.table(dados1[[3]],"points2.txt")
estima <- mec(dados1)
write.table(estima[[1]],"estima2par.txt")
write.table(estima[[2]],"estima2log.txt")
borda <- cbind(c(0,0,1,1,0),c(0,1,1,0,0))
gr <- pred_grid(borda, by=0.03)
source("cokrigagem.R")
md.cov.ck <- cokrigagem(estima[[1]]$Estimativas,loc=gr,
                        dados.comp=dados1)
preditos.gh <- volta.quad(md.cov.ck,n.pontos=7,Variancia=FALSE)
write.table(preditos.gh,"dt2predgh.txt")
preditos.simu <- volta.cokri(md.cov.ck,num.simu=150,int.conf=0.95)
write.table(preditos.simu,"dt2predsimu.txt")

```

Arquivo branco3.R

```

## Script rodado para a 3a configuração de diagrama ternário
## Gera valores preditos por quadratura e simulação pra fazer mapas
require(MASS)
require(geoR)
require(geoComp)
require(compositions)

```

```

require(statmod)
source("simula.R")
## Função que gera diagrama ternário
gerater <- function(theta,n.simul=100){
  dados <- simu.dados(theta,n.simul=n.simul)
  y1 <- dados[1][[1]]
  y2 <- dados[2][[1]]
  y12 <- data.frame(y1,y2)
  comp <- agl(y12)
  comp1 <- acomp(comp)
  names(comp1) <- c("Areia", "Silte", "Argila")
  plot(comp1)
  plot(mean(comp1),add=T,pch=20,col='red')
  ellipses(mean(comp1),var(comp1),col='red',r=2)
  ellipses(mean(comp1),var(comp1),col='red',r=4)
  retorna <- data.frame(comp,dados[[4]][,1],dados[[4]][,2])
  return(retorna)
}
##### 3a configuração #####
theta <- c(-0.2,-1,0.45,0.13,0.3,0.3,0.6,0.95)
set.seed(1333)
dados1 <- gerater(theta,n.simul=100)
names(dados1) <- c("Areia","Silte","Argila","Coord.X", "Coord.Y")
write.table(dados1,"dad_conf3.txt")
dados1 <- as.geoComp(dados1)
write.table(dados1[[3]],"points3.txt")
estima <- mec(dados1)
write.table(estima[[1]],"estima3par.txt")
write.table(estima[[2]],"estima3log.txt")
borda <- cbind(c(0,0,1,1,0),c(0,1,1,0,0))
gr <- pred_grid(borda, by=0.03)
source("cokrigagem.R")
md.cov.ck <- cokrigagem(estima[[1]]$Estimativas,loc=gr,
                        dados.comp=dados1)
preditos.gh <- volta.quad(md.cov.ck,n.pontos=7,Variancia=FALSE)
write.table(preditos.gh,"dt3predgh.txt")
preditos.simu <- volta.cokri(md.cov.ck,num.simu=150,int.conf=0.95)
write.table(preditos.simu,"dt3predsimu.txt")

```

Scripts para o estudo dos intervalos de cobertura:

Arquivo **simula.R**

```
## Função para simular um conjunto de dados
simu.dados <- function(theta,n.simul){
  mu1 <- theta[1]
  mu2 <- theta[2]
  s1 <- theta[3]
  s2 <- theta[4]
  tau1 <- theta[5]
  tau2 <- theta[6]
  phi <- theta[7]
  rho <- theta[8]

  ## Gerando erros de uma normal bivariada com mu=0,
  ## variâncias tau1, tau2 e correlação rho
  v <- matrix(c(tau1^2,tau1*tau2*rho,tau1*tau2*rho,tau2^2),2,2)
  Sc <- kronecker(diag(n.simul/2),v)
  e <- mvrnorm(n.simul,mu=c(0,0),v)

  ## Gerando um campo aleatório Gaussiano com média 0, variância 1
  ## e parâmetros de alcance phi
  U <- grf(n.simul,grid='irreg',cov.model='exp',cov.pars=c(1,phi),
          mean=0,messages=FALSE)

  S1 <- s1*U$data
  S2 <- s2*U$data

  ## Calculando e Organiznado os dados
  y1 <- as.geodata(data.frame(U$coords[,1],U$coords[,2],mu1+S1+e[,1]))
  y2 <- as.geodata(data.frame(U$coords[,1],U$coords[,2],mu2+S2+e[,2]))
  seq1 <- seq(1,length(y1$data)*2,by=2)
  seq2 <- seq(2,length(y2$data)*2,by=2)
  y <- c()
  y[seq1] <- y1$data
  y[seq2] <- y2$data
  coords <- data.frame(U$coords[,1],U$coords[,2])

  ## Preparando a saída
  retorna <- list()
  retorna[1][[1]] <- y1$data
  retorna[2][[1]] <- y2$data
  retorna[3][[1]] <- y
}
```

```

  retorna[4][[1]] <- coords
  return(retorna)
}

```

Arquivo milSimulacoes.R

```

simulando <- function(theta,n.simu=100,n.amostra=100,
                      metodo="L-BFGS-B"){
  MATRIZ <- matrix(nrow=n.simu,ncol=27)
  dados <- list()
  assign("dados", dados, envir = .GlobalEnv)
  X <- cbind(rep(1:0,length=n.amostra*2),rep(0:1,length=n.amostra*2))
  assign("X", X, envir = .GlobalEnv)
  y <- c()
  assign("y", y, envir = .GlobalEnv)
  for( i in 1:n.simu){
    sim <- simu.dados(theta,n.simul=n.amostra)# n.simul=n. de
                                                # observações simuladas

    y1 <- sim[1][[1]]
    y2 <- sim[2][[1]]
    y <- sim[3][[1]]
    assign("y",y, envir= .GlobalEnv)
    coords <- sim[4][[1]]
    dados[1][[1]] <- sim[3][[1]]
    dados[2][[1]] <- sim[4][[1]]
    assign("dados", dados, envir = .GlobalEnv)
    estimado <- try(mec(y1,y2,X,coords=coords,metodo=metodo,
                      otimizador="optim"))
    recurso <- list()
    recurso[[1]] <- data.frame(matrix(ncol=5,nrow=8))
    recurso[[2]] <- c(1,1,1)
    if(class(estimado) == "try-error"){
      est <- recurso
    }
    if(class(estimado) != "try-error"){
      est <- estimado
    }
    print(est)
    t1 <- as.matrix(est[[1]][c(2,4,5)])[1,]
  }
}

```

```

t2 <- as.matrix(est[[1]][c(2,4,5)])[2,]
t3 <- as.matrix(est[[1]][c(2,4,5)])[3,]
t4 <- as.matrix(est[[1]][c(2,4,5)])[4,]
t5 <- as.matrix(est[[1]][c(2,4,5)])[5,]
t6 <- as.matrix(est[[1]][c(2,4,5)])[6,]
t7 <- as.matrix(est[[1]][c(2,4,5)])[7,]
t8 <- as.matrix(est[[1]][c(2,4,5)])[8,]
linha <- c(matrix(c(t1,t2,t3,t4,t5,t6,t7,t8,est[[2]][1],est[[2]][2],
                    est[[2]][3]),nrow=1))
MATRIZ[i,] <- linha
}
MATRIZ <- data.frame(MATRIZ)
names(MATRIZ) <- c("mu1","Lmin.mu1","Lmax.mu1","mu2","Lmin.mu2",
                  "Lmax.mu2","s1","Lmin.s1","Lmax.s1","s2" ,
                  "Lmin.s2","Lmax.s2","t1","Lmin.t1","Lmax.t1",
                  "t2","Lmin.t2","Lmax.t2" ,"phi","Lmin.phi",
                  "Lmax.phi","rho","Lmin.rho","Lmax.rho","logvero",
                  "logvero.est","conv")
return(MATRIZ)
}

```

Arquivo **LBFGBS.R**

```

## Carregando pacotes
require(MASS)
require(geoR)
require(geoComp)
## Carregando funções
source("simula.R")
source("milSimulacoes.R")
## Usando a função simulando
theta <- c(50,70,2.5,3.5,2.5,3.5,0.6,0.5)
LBFGBS <- simulando(theta,n.simu=1000,n.amostra=300,metodo="L-BFGS-B")
write.table(LBFGBS,"EstudoSimns1000na300.txt")

```

Gráficos referentes aos intervalos de cobertura

```

## Gráficos de densidades à posteriori
dad.sim.na300ns1000 <- read.table('EstudoSimns1000na300.txt')

```



```

par(mfrow=c(2,5),mar=c(3.5,3.5,1.5,1.5),mgp=c(1.6,0.6,0))
hist(dad.sim.na300ns1000$mu1,prob=T,main=' ',xlim=c(44,56),
      ylim=c(0,0.25),xlab=('mu1 estimado'),ylab=('Densidade'))
lines(density(dad.sim.na300ns1000$mu1),col=2)
hist(dad.sim.na300ns1000$mu2,prob=T,main=' ',xlim=c(60,80),
      ylim=c(0,0.20),xlab=('mu2 estimado'),ylab=('Densidade'))
lines(density(dad.sim.na300ns1000$mu2),col=2)
hist(dad.sim.na300ns1000$s1,prob=T,main=' ',xlim=c(0,5),
      ylim=c(0,0.9),xlab=('s1 estimado'),ylab=('Densidade'))
lines(density(dad.sim.na300ns1000$s1),col=2)
hist(dad.sim.na300ns1000$s2,prob=T,main=' ',xlim=c(0,8),
      ylim=c(0,0.6),xlab=('s2 estimado'),ylab=('Densidade'))
lines(density(dad.sim.na300ns1000$s2),col=2)
hist(dad.sim.na300ns1000$t1,prob=T,main=' ',xlab=('t1 estimado'),
      ylab=('Densidade'))
lines(density(dad.sim.na300ns1000$t1),col=2)
hist(dad.sim.na300ns1000$t2,prob=T,main=' ',xlim=c(2,5),
      ylim=c(0,2.5),xlab=('t2 estimado'),ylab=('Densidade'))
lines(density(dad.sim.na300ns1000$t2),col=2)
hist(dad.sim.na300ns1000$phi,prob=T,main=' ',ylim=c(0,3),
      xlab=('phi estimado'),ylab=('Densidade'))
lines(density(dad.sim.na300ns1000$phi),col=2)
hist(dad.sim.na300ns1000$rho,prob=T,main=' ',xlim=c(0.2,0.7),
      ylim=c(0,10),xlab=('rho estimado'),ylab=('Densidade'))
lines(density(dad.sim.na300ns1000$rho),col=2)
hist(dad.sim.na300ns1000$logvero,prob=T,main=' ',xlim=c(1450,1700),
      ylim=c(0,0.02),xlab=('logvero'),ylab=('Densidade'))
lines(density(dad.sim.na300ns1000$logvero),col=2)
hist(dad.sim.na300ns1000$logvero.est,prob=T,main=' ',xlim=c(-1560,-1420),
      ylim=c(0,0.025),xlab=('logvero estimada'),ylab=('Densidade'))
lines(density(dad.sim.na300ns1000$logvero.est),col=2)
## Boxplots das posteriores dos parâmetros
par(mfrow=c(2,4),mar=c(3.5, 3.5, 1.5, 0.5), mgp=c(1.8,.8, 0))
boxplot(dad.sim.na300ns1000$mu1,main='mu1',ylab="Estimativas")
boxplot(dad.sim.na300ns1000$mu2,main='mu2',ylab="Estimativas")
boxplot(dad.sim.na300ns1000$s1,main='s1',ylab="Estimativas")
boxplot(dad.sim.na300ns1000$s2,main='s2',ylab="Estimativas")
boxplot(dad.sim.na300ns1000$t1,main='tau1',ylab="Estimativas")

```

```

boxplot(dad.sim.na300ns1000$t2,main='tau2',ylab="Estimativas")
boxplot(dad.sim.na300ns1000$phi,main='phi',ylab="Estimativas")
boxplot(dad.sim.na300ns1000$rho,main='rho',ylab="Estimativas")

```

Script para a obtenção dos valores preditos utilizando os dados do pivo:

```

require(geoComp)
require(MASS)
require(statmod)
require(geoR)
data(pivo)
dados <- pivo[,c(6,7,8,1,2)]
dados <- as.geoComp(dados)
bor <- cbind(c(0,seq(0,200,l=100),0),
             c(0,sqrt(200^2-seq(0,200,l=100)^2),0))
estima <- mec(dados)
gr <- pred_grid(bor, by=4)
md.cov.ck <- cokrigagem(estima[[1]]$Estimativas, loc=gr,
                      dados.comp=dados)
preditos.gh <- volta.quad(md.cov.ck,n.pontos=7,Variancia=FALSE)
preditos.gh <- data.frame(preditos.gh)
write.table(preditos.gh,"predGH.txt")
preditos.simu <- volta.cokri(md.cov.ck,num.simu=1000,int.conf=0.95)
preditos.simu <- data.frame(preditos.simu[[1]])
preditos.simu.ic <- data.frame(preditos.simu[[2]])
write.table(preditos.simu,"predsimu.txt")
write.table(preditos.simu.ic,"predsimuic.txt")

```

Scripts para o estudo da Verossimilhança perfilhada:

Perfilhada para ϕ

Arquivo **monta.V.phi.R**

```

## theta1=(eta,nu1,nu2,rho)
monta.V.phi <-function(theta1, phi, dados.comp){
  y <- dados.comp[[2]][[1]]
  eta <- theta1[1]
  nu1 <- theta1[2]

```

```

nu2 <- theta1[3]
rho <- theta1[4]
phi <- phi
## Calculando os elementos da matriz de correlação espacial
distancia <- unname(as.matrix(dist(dados.comp[[3]],diag=TRUE,
                                upper=TRUE)))
correla <- exp(-distancia/phi)
## Organizando a matriz de correlação espacial com eltos y1 e y2
## intercalados. Ou seja, calculando R1
coluna1 <- rep(c(1, eta),length(y)/2)
coluna2 <- rep(c(eta,eta^2),length(y)/2)
seq1 <- seq(1,length(y),by=2)
seq2 <- seq(2,length(y),by=2)
ERRE1 <- matrix(ncol=length(y),nrow=length(y))
ERRE1[,c(seq1)] <- coluna1
ERRE1[,c(seq2)] <- coluna2
## Organizando a matriz de correlação espacial com eltos y1 e y2
## intercalados
ERRE2 <- matrix(nrow=length(y),ncol=length(y))
for(i in 1:length(y)/2){
  ERRE2[seq1[i],] <- rep(correla[i,],each=2)
  ERRE2[seq2[i],] <- rep(correla[i,],each=2)
}
## Calculando a matriz R
ERRE <- ERRE1*ERRE2
## Calculando a matriz de covariância composicional
Ib1 <- matrix(c(nu1^2,nu1*nu2*rho,nu1*nu2*rho,nu2^2),nc=2)
Ib <- kronecker(diag(length(y)/2),Ib1)
## Matriz de covariância espacial composicional
V <- ERRE + Ib
return(V)
}

```

Arquivo **log.vero.phi.R**

```

## theta1=(eta,nu1,nu2,rho)
log.vero.phi <-function(theta1,phi,dados.comp,print.pars=F){
  if(theta1[1] <= 0)      return(.Machine$double.xmax^0.5)
  if(theta1[2] <= 0)      return(.Machine$double.xmax^0.5)

```

```

if(theta1[3] <= 0)      return(.Machine$double.xmax^0.5)
if(abs(theta1[4]) > 1) return(.Machine$double.xmax^0.5)
y <- dados.comp[[2]][[1]]
X <- cbind(rep(1:0,length=length(y)),rep(0:1,length=length(y)))
V <- monta.V.phi(theta1=theta1,phi=phi ,dados.comp = dados.comp)
ldetV <- determinant(V,log=TRUE)$modulus[1]
mu <- drop(solve(crossprod(X,solve(V,X))) %*%crossprod(X,solve(V,y)))
Qe <- drop(crossprod(y,solve(V,y))-2*crossprod(y,solve(V,X%*%mu))+
           crossprod(mu,crossprod(X,solve(V,X%*%mu))))
if(Qe < 0) return(Qe=.Machine$double.xmax^0.5)
n <- length(y)
s1 <- sqrt(Qe/n)
## drop diminui a dimensão do objeto
ll <- drop(-0.5*(n*log(2*pi)+n*log(s1^2)+ ldetV + (1/(s1^2))*Qe))
if(print.pars) print(c(theta1,ll))
return(-ll)
}

```

Arquivo **perf.phi.R**

```

## Obtendo valores iniciais para theta1 = (eta, nu1, nu2, rho)
perf.phi <- function(dados.comp,tamanho, # Tamanho vetor phi pra perfilhar
                    min.delta, # Qto % menor que o mínimo dado pelo delta
                    max.delta, # Qto % maior que o máximo
                    alpha){
var_y1 <- var(dados.comp[[1]][1])
s1 <- var_y1/2
tau1 <- s1
var_y2 <- var(dados.comp[[1]][2])
s2 <- var_y2/2
tau2 <- s2
eta <- s2/s1
nu1 <- tau1/s1
nu2 <- tau2/s1
rho <- cor(dados[[1]][1],dados[[1]][2])
theta1 <- c(eta,nu1,nu2,rho)
## Fazendo a maximização
source("monta.V.phi.R")
source("log.vero.phi.R")

```

```

estima <- mec(dados.comp=dados.comp)
## ic_i.delta.phi = limite inferior para phi pelo método delta
## ic_s.delta.phi = limite superior para phi pelo método delta
ic_i.delta.phi <- estima[[1]]$LI.Delta[7]
ic_s.delta.phi <- estima[[1]]$LS.Delta[7]
phi <- seq(ic_i.delta.phi*min.delta,ic_s.delta.phi*max.delta,l=tamanho)
lista.res <- list()
for(i in 1:length(phi)){
  teste <- optim(theta1, log.vero.phi, phi=phi[i],
                dados.comp=dados.comp, method="L-BFGS-B",
                lower=c(1e-32,1e-32,1e-32,-1),
                upper=c(Inf,Inf,Inf,1), hessian=TRUE)
  ## Todos os resultados para cada valor de phi
  lista.res[i][[1]] <- teste
}
perf.phi <- c()
for(i in 1:length(phi)){
  perf.phi[i] <- lista.res[[i]][2]$value
}
## Gráfico da perfilhada: phi versus log-verossimilhança
#postscript('perf.phi.eps')
#plot(phi,perf.phi,ylab='Log-verossimilhanca Perfilhada')
#dev.off()
maximo <- -estima[[2]][2]
deviance.phi <- 2*(perf.phi-maximo)
# plot(phi,deviance.phi,ylab='Deviance',type="l")
# abline(h=qchisq(0.835,df=1))
val.phi.perf.dev <- data.frame(phi,perf.phi,deviance.phi)
intervalo.phi <- val.phi.perf.dev[which(val.phi.perf.dev$deviance.phi<
qchisq(alpha,df=1)),]$phi
retorna <- list()
retorna[[1]] <- val.phi.perf.dev
retorna[[2]] <- intervalo.phi
return(retorna)
}

```

Arquivo **script.perfilhada.phi.R**

```
rm(list=ls(all=T))
```

```

#par.ori <- par(no.readonly=TRUE)
#require(compositions)
###require(geoComp)

## Carregando o arquivo de dados do geoComp
source('pivo.R')
###data(pivo)
#head(pivo)
pivo

## Selecionando os componentes e as coordenadas
dados <- pivo[,c(6,7,8,1,2)]
#head(dados)

## Convertendo o arquivo para o formato geoComp
## Transformacao alr ( denominador=argila )
source('as.geoComp.R')
dados <- as.geoComp(dados)

source('monta.V.R')
source('log.vero.R')
source('mec.R')
dados.comp <- dados
estima <- mec(dados)

## Fazendo a perfilhada para phi com alpha=0.566
source('perf.phi.R')
perfilhada.phi <-perf.phi(dados,tamanho=150,min.delta=0.8,max.delta=5,
                        alpha=0.566)

perfilhada.phi
## Arquivo contendo valores de phi, perf.phi e deviance.phi
perfilhada.phi.1 <- perfilhada.phi[[1]]
write.table(perfilhada.phi.1,'dataframePerfphi0566p150.txt')
## Arquivo contendo todos os valores para os quais
## deviance<qchisq(alpha,df=1))
perfilhada.phi.2 <- perfilhada.phi[[2]]
write.table(perfilhada.phi.2 , 'IC.Perf.phi.0566p150.txt')

```

```

## Fazendo a perfilhada para phi com alpha=0.835
perfilhada.phi.3 <-perf.phi(dados,tamanho=150,min.delta=0.8,max.delta=5,
                           alpha=0.835)
## Arquivo contendo valores de phi, perf.phi e deviance.phi
perfilhada.phi.4 <- perfilhada.phi.3[[1]]
write.table(perfilhada.phi.4,'dataframePerfphi0835p150.txt')
## Arquivo contendo todos os valores para os quais
## deviance<qchisq(alpha,df=1))
perfilhada.phi.5 <- perfilhada.phi.3[[2]]
write.table(perfilhada.phi.5 , 'IC.Perf.phi.0835p150.txt')

## Lendo os resultados
dad.per.phi.0566 <- read.table('dataframePerfphi0566p150.txt')

dad.ic.phi.0566 <- read.table('IC.Perf.phi.0566p150.txt')
dad.ic.phi.0835 <- read.table('IC.Perf.phi.0835p150.txt')

## Organizando os intervalos
dad.ic.phi.0566 <- data.frame(dad.ic.phi.0566[1,],dad.ic.phi.0566[41,])
names(dad.ic.phi.0566) <- c('Li.phi.0566','Ls.phi.0566')
dad.ic.phi.0835 <- data.frame(dad.ic.phi.0835[1,],dad.ic.phi.0835[141,])
names(dad.ic.phi.0835) <- c('Li.phi.0835','Ls.phi.0835')

## Gráficos das perfilhadas: phi versus log-verossimilhanca
postscript('perfphi0566.eps')
par(mfrow=c(1,1),mar=c(3,3,1.5,1.5),mgp=c(1.8,0.8,0))
plot(dad.per.phi.0566$phi,-dad.per.phi.0566$perf.phi,xlab=expression(phi),
      ylab="Log-verossimilhanca Perfilhada",type='l')
dev.off()

#postscript('perfphi0835.eps')
#plot(dad.per.phi.0835$phi,-dad.per.phi.0835$perf.phi,xlab=expression(phi),
      ylab="Log-verossimilhanca Perfilhada",type='l')
#dev.off()

# Gráficos das Deviances: phi versus Deviance
postscript('deviancephi060e0835.eps')
par(mfrow=c(1,1),mar=c(3,3,1.5,1.5),mgp=c(1.8,0.8,0))

```

```

plot(dad.per.phi.0566$phi,dad.per.phi.0566$deviance.phi,xlab=expression(phi),
      ylab='Deviance',type="l")

segments(x0=dad.ic.phi.0566$Li.phi.0566,x1=dad.ic.phi.0566$Ls.phi.0566,
          y0=qchisq(0.566,df=1),y1=qchisq(0.566,df=1))

arrows(x0=dad.ic.phi.0566$Li.phi.0566,y0=qchisq(0.566,df=1),
        x1=dad.ic.phi.0566$Li.phi.0566,y1=0)
arrows(x0=dad.ic.phi.0566$Ls.phi.0566,y0=qchisq(0.566,df=1),
        x1=dad.ic.phi.0566$Ls.phi.0566,y1=0)
text(x=70,y=-0.1,label=as.character(round(dad.ic.phi.0566$Li.phi.0566,2)),
      cex=0.8)
text(x=dad.ic.phi.0566$Ls.phi.0566,y=-0.1,
      label=as.character(round(dad.ic.phi.0566$Ls.phi.0566,2)),cex=0.8)

segments(x0=dad.ic.phi.0835$Li.phi.0835,x1=dad.ic.phi.0835$Ls.phi.0835,
          y0=qchisq(0.835,df=1),y1=qchisq(0.835,df=1))
arrows(x0=dad.ic.phi.0835$Li.phi.0835,y0=qchisq(0.835,df=1),
        x1=dad.ic.phi.0835$Li.phi.0835,y1=0)
arrows(x0=dad.ic.phi.0835$Ls.phi.0835,y0=qchisq(0.835,df=1),
        x1=dad.ic.phi.0835$Ls.phi.0835,y1=0)
text(x=23,y=-0.1,label=as.character(round(dad.ic.phi.0835$Li.phi.0835,1)),
      cex=0.8)
text(x=dad.ic.phi.0835$Ls.phi.0835,y=-0.1,
      label=as.character(round(dad.ic.phi.0835$Ls.phi.0835,2)),cex=0.8)
dev.off()

```

Perfilhada para ρ

Arquivo **monta.V.rho.R**

```

## theta1=(eta,nu1,nu2,phi)
monta.V.rho <-function(theta1, rho, dados.comp){
  y <- dados.comp[[2]][[1]]
  eta <- theta1[1]
  nu1 <- theta1[2]
  nu2 <- theta1[3]
  phi <- theta1[4]
  rho <- rho
  ## Calculando os elementos da matriz de correlação espacial

```



```

distancia <- unname(as.matrix(dist(dados.comp[[3]],diag=TRUE,
                                upper=TRUE)))
correla <- exp(-distancia/phi)
## Organizando a matriz de correlação espacial com eltos y1 e y2
## intercalados. Ou seja, calculando R1
coluna1 <- rep(c(1, eta),length(y)/2)
coluna2 <- rep(c(eta,eta^2),length(y)/2)
seq1 <- seq(1,length(y),by=2)
seq2 <- seq(2,length(y),by=2)
ERRE1 <- matrix(ncol=length(y),nrow=length(y))
ERRE1[,c(seq1)] <- coluna1
ERRE1[,c(seq2)] <- coluna2
## Organizando a matriz de correlação espacial com eltos y1 e y2
## intercalados
ERRE2 <- matrix(nrow=length(y),ncol=length(y))
for(i in 1:length(y)/2){
  ERRE2[seq1[i],] <- rep(correla[i,],each=2)
  ERRE2[seq2[i],] <- rep(correla[i,],each=2)
}
## Calculando a matriz R
ERRE <- ERRE1*ERRE2
## Calculando a matriz de covariância composicional
Ib1 <- matrix(c(nu1^2,nu1*nu2*rho,nu1*nu2*rho,nu2^2),nc=2)
Ib <- kronecker(diag(length(y)/2),Ib1)
## Matriz de covariância espacial composicional
V <- ERRE + Ib
return(V)
}

```

Arquivo **log.vero.rho.R**

```

## theta1=(eta,nu1,nu2,phi)
log.vero.rho <-function(theta1,rho,dados.comp,print.pars=F){
  if(theta1[1] <= 0)      return(.Machine$double.xmax^0.5)
  if(theta1[2] <= 0)      return(.Machine$double.xmax^0.5)
  if(theta1[3] <= 0)      return(.Machine$double.xmax^0.5)
  if(theta1[4] < 0)       return(.Machine$double.xmax^0.5)
  y <- dados.comp[[2]][[1]]
  X <- cbind(rep(1:0,length=length(y)),rep(0:1,length=length(y)))

```

```

V <- monta.V.rho(theta1=theta1,rho=rho ,dados.comp = dados.comp)
ldetV <- determinant(V,log=TRUE)$modulus[1]
mu <- drop(solve(crossprod(X,solve(V,X))) %*%crossprod(X,solve(V,y)))
Qe <- drop(crossprod(y,solve(V,y))-2*crossprod(y,solve(V,X%*%mu))+
           crossprod(mu,crossprod(X,solve(V,X%*%mu))))
if(Qe < 0) return(Qe=.Machine$double.xmax^0.5)
n <- length(y)
s1 <- sqrt(Qe/n)
ll <- drop(-0.5*(n*log(2*pi)+n*log(s1^2)+ ldetV + (1/(s1^2))*Qe))
if(print.pars) print(c(theta1,ll))
return(-ll)
}

```

Arquivo **perf.rho.R**

```

## Obtendo valores iniciais para theta1 = (eta, nu1, nu2, phi)
perf.rho <- function(dados.comp,
                    tamanho,# Tamanho do vetor rho p/ perfilhar
                    min.delta, # Qto % menor que o mínimo dado pelo delta
                    max.delta, # Qto % maior que o máximo
                    alpha){
var_y1 <- var(dados.comp[[1]][1])
s1 <- var_y1/2
tau1 <- s1
var_y2 <- var(dados.comp[[1]][2])
s2 <- var_y2/2
tau2 <- s2
dim <- range(dist(dados[[3]]))
phi <- dim[1]+0.2*(dim[2]-dim[1])
eta <- s2/s1
nu1 <- tau1/s1
nu2 <- tau2/s1
theta1 <- c(eta,nu1,nu2,phi)
## Fazendo a maximização
source("monta.V.rho.R")
source("log.vero.rho.R")
estima <- mec(dados.comp=dados.comp)
## ic.inf.delta = limite inferior para rho pelo metodo delta
## ic.sup.delta = limite superior para rho pelo metodo delta

```

```

ic_i.delta.rho <- estima[[1]]$LI.Delta[8]
ic_s.delta.rho <- estima[[1]]$LS.Delta[8]
rho <- seq(ic_i.delta.rho*min.delta,ic_i.delta.rho*max.delta,l=tamanho)
rho[which(rho > 1*0.9995)] <- 1*0.9995
lista.res <- list()
for(i in 1:length(rho)){
  teste <- optim(theta1, log.vero.rho, rho=rho[i],
                dados.comp=dados.comp, method="L-BFGS-B",
                lower=c(1e-32,1e-32,1e-32,1e-32),
                upper=c(20,20,20,200), hessian=TRUE)
  ## Todos os resultados para cada valor de rho
  lista.res[i][[1]] <- teste
}
perf.rho <- c()
for(i in 1:length(rho)){
  perf.rho[i] <- lista.res[[i]][2]$value
}
perf.rho
## Gráfico da perfilhada: rho versus log-verossimilhança
#postscript('perf.rho.eps')
#plot(rho,perf.rho,ylab="Log-verossimilhanca Perfilhada")
#dev.off()
maximo <- -estima[[2]][2]
deviance.rho <- 2*(perf.rho-maximo)
#postscript('deviance.rho.eps')
#plot(rho,deviance.rho,ylab='Deviance',type="l")
#dev.off()
#abline(h=qchisq(0.95,df=1))
val.rho.perf.dev <- data.frame(rho,perf.rho,deviance.rho)
#intervalo.rho <- range(val.rho.perf.dev[which(val.rho.perf.dev$deviance<
      qchisq(0.95,df=1)),]$rho)
intervalo.rho <- val.rho.perf.dev[which(val.rho.perf.dev$deviance.rho<
      qchisq(alpha,df=1)),]$rho
retorna <- list()
retorna[[1]] <- val.rho.perf.dev
retorna[[2]] <- intervalo.rho
return(retorna)
}

```

Arquivo `script.perfilhada.rho.R`

```

rm(list=ls(all=T))
#par.ori <- par(no.readonly=TRUE)
#require(compositions)
###require(geoComp)

## Carregando o arquivo de dados do geoComp
source('pivo.R')
###data(pivo)
pivo

## Selecionando os componentes e as coordenadas
dados <- pivo[,c(6,7,8,1,2)]
#head(dados)

## Convertendo o arquivo para o formato geoComp
## transformacao alr ( denominador=argila )
## dados[[1]]= Y1 Y2 ; dados[[2]]= Y ; dados[[3]]= Coord.X Coord.Y
source('as.geoComp.R')
dados <- as.geoComp(dados)

source('monta.V.R')
source('log.vero.R')
source('mec.R')
dados.comp <- dados
estima <- mec(dados)

## Fazendo a perfilhada para rho com alpha=095
source('perf.rho.R')
perfilhada.rho <- perf.rho(dados,tamanho=150,min.delta=0.85,max.delta=1.10,
                          alpha=0.95)
perfilhada.rho.1 <- perfilhada.rho[[1]]
write.table(perfilhada.rho.1,'dataframePerfrho095p150.txt')
perfilhada.rho.2 <- perfilhada.rho[[2]]
write.table(perfilhada.rho.2 , 'IC.Perf.rho.095p150.txt')
## Lendo os resultados
dad.per.rho.095 <- read.table('dataframePerfrho095p150.txt')

```

```

dad.ic.rho.095 <- read.table('IC.Perf.rho.095p150.txt')
## Organizando os intervalos
dad.ic.rho.095 <- data.frame(dad.ic.rho.095[1,],dad.ic.rho.095[90,])
names(dad.ic.rho.095) <- c('Li.rho.095','Ls.rho.095')
## Gráficos das perfilhadas: rho versus log-verossimilhanca
postscript('perfrho095.eps')
par(mfrow=c(1,1),mar=c(3,3,1.5,1.5),mgp=c(1.8,0.8,0))
plot(dad.per.rho.095$rho,-dad.per.rho.095$perf.rho,xlab=expression(rho),
      ylab="Log-verossimilhanca Perfilhada",type='l')
dev.off()
## Gráficos das Deviances: rho versus Deviance
postscript('deviancerho095.eps')
par(mfrow=c(1,1),mar=c(3,3,1.5,1.5),mgp=c(1.8,0.8,0))
plot(dad.per.rho.095$rho,dad.per.rho.095$deviance.rho,xlab=expression(rho),
      ylab='Deviance',type="l")
segments(x0=dad.ic.rho.095$Li.rho.095,x1=dad.ic.rho.095$Ls.rho.095,
          y0=qchisq(0.95,df=1),y1=qchisq(0.95,df=1))
arrows(x0=dad.ic.rho.095$Li.rho.095,y0=qchisq(0.95,df=1),
        x1=dad.ic.rho.095$Li.rho.095,y1=0)
text(x=dad.ic.rho.095$Li.rho.095,y=-0.2,
      label=as.character(round(dad.ic.rho.095$Li.rho.095,2)),cex=0.8)
dev.off()

```

Script para fazer inferência Bayesiana:

```

#.libPaths("/home/users/est/paulojus/Rlibs")
## carregando pacotes
require(geoR) # necessário para usar rinvchisq
require(geoComp)
require(MCMCpack)
require(MASS)
source('posteriori.R')
source('FunçãoInfBayes.R')
source('resultado.mh.R')
source('mycokri.bayes.3.R')
## Lendo dados
dados <- pivo[,c(6,7,8,1,2)]
## Transformando os dados para a classe geoComp
dados <- as.geoComp(dados)

```

```

estima.bayes <- infBayes(dados,tune=c(0.5,0.5,0.5,0.5,0.5),
                        queima=1000,salto=10,nsim=12000,mediahiper=c())
write.table(estima.bayes[[1]],'estima.bayes1.txt')
write.table(estima.bayes[[2]],'estima.bayes2.txt')
## Opção para o usuário entrar com os parâmetros das prioris:
#estima.bayes <- infBayes(dados,tune=c(0.5,0.5,0.5,0.5,0.5),
                        queima=1000,salto=10,nsim=12000,
                        mediahiper=c(1,1,1,1,1,1,1,1,1,1))
summary(estima.bayes[[2]])
res.mh <- resul.mh(resultado=estima.bayes)
res.mh.ic <- data.frame(res.mh[[2]])
names(res.mh.ic) <- c('2.5%', 'media', '97.5%')
write.table(res.mh.ic, 'res.mh.ic.txt')
metro.bayes <- data.frame(res.mh[[1]])
write.table(metro.bayes, 'metro.bayes.txt')
bor <- cbind(c(0,seq(0,200,l=100),0),
            c(0,sqrt(200^2-seq(0,200,l=100)^2),0))
gr <- pred_grid(bor, by=4)
cokri.bayes <- cokri.bayes(estim.par=metro.bayes,locations=gr,
                        dados.comp=dados)
write.table(cokri.bayes[[1]], 'cokri1.txt')
write.table(cokri.bayes[[2]], 'cokri2.txt')
# Fazendo o mapa
tes1 <- read.table('cokri1_by4ns1200.txt')#composição
tes2 <- read.table('cokri2_by4ns1200.txt')#coordenadas
tes1 <- as.matrix(tes1)
tes2 <-as.matrix(tes2)
postscript('mapabayesby4ns1200.eps',horizontal=TRUE)
#pdf('mapabayesby4ns1200.pdf',width=8,height=2)
par(mfrow=c(1,3), mar=c(3,3,.5,.5), mgp=c(1.7,0.7, 0),cex=0.9)
#par(mfrow=c(1,3), mar=c(3.5, 3.5, 0.5, 0.5), mgp=c(1.8,0.8,0))
xleg <- c(85, 198); yleg <- c(185, 200)
image(structure(list(predict=tes1[,1]), class="kriging"),loc=tes2,bor,
      col=terrain.colors(21),x.leg=xleg, y.leg=yleg,ylim=c(0,220))
text(x=165,y=170,"(A)Areia/IB")
image(structure(list(predict=tes1[,2]), class="kriging"),loc=tes2,bor,
      col=terrain.colors(21),
      x.leg=xleg, y.leg=yleg,ylim=c(0,220))

```

```

text(x=165,y=170,"(B)Silte/IB")
image(structure(list(predict=tes1[,3]), class="kriging"),loc=tes2,bor,
      col=terrain.colors(21),
      x.leg=xleg, y.leg=yleg,ylim=c(0,220))
text(x=165,y=170,"(C)Argila/IB")
dev.off()
### Fazendo diagnóstico
require(coda)
metro.emv <- read.table('metro.bayes_by4ns1200.txt')
metro.emv <- as.mcmc(metro.emv)
metro.emv.GR <- read.table('metro.bayes.GR.txt')
metro.emv.GR <- as.mcmc(metro.emv.GR)
fac <- autocorr.diag(metro.emv)
crosscor <-crosscorr(metro.emv)
hpd <-HPDinterval(metro.emv)
plot(density(metro.emv[,1]), main='post.mu1')
rug(metro.emv[,1])
lines(seq(unique(hpd[,1][1]),unique(hpd[,2][1]),l=10),rep(0,10),
      type="l",lty=1)
teste <- geweke.diag(metro.emv,frac1=0.90,frac2=0.5)
## Rodando uma segunda cadeia com
## theta1.inicial <-c(1,1,1,40,0.8) na FunçãoInfBayes.R
## para fazer o teste Gelman eRubin
metro.emv <- read.table('mbmh1.txt')
metro.emv.GR<- read.table('mbmh2.txt')
metro.emv<- as.mcmc(metro.emv)
metro.emv.GR<- as.mcmc(metro.emv.GR)
dadoPgelman <- list()
dadoPgelman[[1]] <- metro.emv
dadoPgelman[[2]] <- metro.emv.GR
res.gelman <- gelman.diag(dadoPgelman,confidence=0.95,transform=FALSE,
                          autoburnin=FALSE)
gelmanplot <- gelman.plot(dadoPgelman,bin.width=10,max.bins=50,
                          confidence=0.95, transform=FALSE,
                          auto.layout=TRUE)

```

Scripts referentes aos funcionais:

Arquivo **lcpadFuncionais.R:**

```

## Script para obter 1000 simulações de areia, silte e argila em cada
## localização
rm(list=ls(all=TRUE))
require(MASS)
require(statmod)
require(geoR)
require(geoComp)
data(pivo)
dados <- pivo[,c(6,7,8,1,2)]
dados <- as.geoComp(dados)
bor <- cbind(c(0,seq(0, 200, l=100), 0), c(0,sqrt(200^2-seq(0, 200, l=100)^2), 0))
estima <- mec(dados)
gr <- pred_grid(bor, by=4)
md.cov.ck <- cokrigagem(estima[[1]]$Estimativas, loc=gr, dados.comp=dados)
#preditos.gh <- volta.quad(md.cov.ck,n.pontos=7,Variancia=FALSE)
#preditos.gh <- data.frame(preditos.gh)
#names(preditos.gh) <- c("Areia","Silte","Argila")
#write.table(preditos.gh,"predGH_by4k7ns1000.txt")
preditos.simu <- volta.cokri(md.cov.ck,num.simu=1000,retorna.tudo=TRUE,int.conf=0.95)
preditos.simu.areia <- data.frame(preditos.simu[[1]])
preditos.simu.silte <- data.frame(preditos.simu[[2]])
preditos.simu.argila <- data.frame(preditos.simu[[3]])
write.table(preditos.simu.areia,"predsimu.are_by4k7ns1000.txt")
write.table(preditos.simu.silte,"predsimu.sil_by4k7ns1000.txt")
write.table(preditos.simu.argila,"predsimu.arg_by4k7ns1000.txt")
### Se retorna.tudo=FALSE use as 4 linhas seguintes
##preditos.simu <- data.frame(preditos.simu[[1]])
##preditos.simu.ic <- data.frame(preditos.simu[[2]])
##write.table(preditos.simu,"predsimu_by4k7ns1000.txt")
##write.table(preditos.simu.ic,"predsimuic_by4k7ns1000.txt")

```

Arquivo **funcionalEst.R**:

```

## Função que calcula funcionais dos valores preditos simulados
funcionalEst <- function(dados,funcional,prob.quartil){
  if(funcional=="Maximo"){
    media <- apply(dados,1,max)
    return(media)
  }
}

```



```
if(funcional=="Minimo"){
  media <- apply(dados,1,min)
  return(media)
}
## Outros possíveis funcionais
if(funcional=="Media"){
  media <- apply(dados,1,mean)
  return(media)
}
if(funcional=="Variância"){
  require(stats)
  var <- apply(dados,1,var)
  desv.pad <- sqrt(var)
  resultado <- data.frame(var,desv.pad)
  names(resultado) <- c("Variância","Desvio padrao")
  return(resultado)
}
if(funcional=="Quartil"){
  quartil <- data.frame(t(apply(dados,1,quantile,prob=c(prob.quartil))))
  return(quartil)
}
}
```