

UNIVERSIDADE ESTADUAL DO OESTE DO PARANÁ
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA AGRÍCOLA

TÉCNICAS DE DIAGNÓSTICO EM MODELOS ESPACIAIS LINEARES
GAUSSIANOS

JOELMIR ANDRÉ BORSSOI

CASCADEL
DEZEMBRO DE 2007

JOELMIR ANDRÉ BORSSOI

**TÉCNICAS DE DIAGNÓSTICO EM MODELOS ESPACIAIS LINEARES
GAUSSIANOS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Agrícola em cumprimento parcial aos requisitos para obtenção do título de Mestre em Engenharia Agrícola, pela UNIOESTE/Campus de Cascavel, área de concentração Engenharia de Sistemas Agroindustriais.

Orientador: Professor Dr. Miguel Angel Uribe Opazo

Co-orientador: Professor Dr. Manuel Jesus Galea Rojas

CASCADEL

DEZEMBRO DE 2007

JOELMIR ANDRÉ BORSSOI

**TÉCNICAS DE DIAGNÓSTICOS EM MODELOS ESPACIAIS LINEARES
GAUSSIANOS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Agrícola em cumprimento parcial aos requisitos para obtenção do título de Mestre em Engenharia Agrícola, área de concentração Engenharia de Sistemas Agroindustriais, aprovada pela seguinte banca examinadora:

Orientador: Prof. Dr. Miguel Angel Uribe Opazo
Centro de Ciências Exatas e Tecnológicas, UNIOESTE

Prof^a. Dra. Lúcia Helena Pereira Nóbrega
Centro de Ciências Exatas e Tecnológicas, UNIOESTE

Prof. Dr. Eduardo Godoy de Souza
Centro de Ciências Exatas e Tecnológicas, UNIOESTE

Prof. Dr. Rubens Augusto Camargo Lamparelli
Faculdade de Engenharia Agrícola, UNICAMP

Cascavel, 04 de dezembro de 2007

À minha família, com muito carinho, dedico.

AGRADECIMENTOS

A Deus, pelo dom da vida e por estar comigo em todos os momentos.

Ao meu orientador, Prof. Dr. Miguel Angel Uribe Opazo, pela oportunidade de trabalharmos juntos há vários anos, pelos ensinamentos, amizade e exemplo de profissionalismo.

Ao meu co-orientador, Prof. Dr. Manuel Jesus Galea Rojas, pelo auxílio no desenvolvimento deste trabalho.

A todos os professores e colegas do curso de mestrado, pelos ensinamentos, amizade e companheirismo demonstrados durante este período.

Aos colegas do Grupo de Geoestatística Aplicada, em especial ao Prof. M.sc Jerry Adriani Johann, Prof. Dr. Erivelto Mercante, Prof. M.sc Edson Antonio Alves da Silva, Claudinei Lunkes, Gustavo Henrique Dalposso e Fabiano Giacomini, pelo companheirismo e apoio na coleta de dados em campo.

À UNIOESTE, por proporcionar meu aperfeiçoamento profissional.

À CONAB e ao CNPq, pelo apoio financeiro, concedendo uma bolsa de estudos.

Ao senhor Agassis Linhares Neto, pela parceria na implantação dos experimentos em sua propriedade.

Aos meus pais Adelino e Terezinha e a meus irmãos Adriana, Tatiani, Adilson e Marinho, pelo incentivo e por compartilharem comigo a realização de mais um sonho.

E a todos que não foram citados, mas que de alguma forma contribuíram com a realização deste trabalho.

A todos, muito obrigado!!!

SUMÁRIO

LISTA DE TABELAS	viii
LISTA DE FIGURAS	xi
RESUMO	xv
ABSTRACT	xvi
1 INTRODUÇÃO	1
2 REVISÃO BIBLIOGRÁFICA.....	3
2.1 A TEORIA DAS VARIÁVEIS REGIONALIZADAS.....	4
2.1.1 Características das Variáveis Regionalizadas	4
2.2 MODELOS ESPACIAIS LINEARES.....	7
2.2.1 Parâmetros do Semivariograma	9
2.2.2 Modelos Teóricos de Semivariograma, Funções de Covariância e de Correlação.....	10
2.3 ESTIMAÇÃO DE PARÂMETROS	18
2.3.1 Estimador de Máxima Verossimilhança – MV.....	18
2.3.2 Estimador de Máxima Verossimilhança Restrita – MVR.....	20
2.4 KRIGAGEM	21
2.4.1 Krigagem Ordinária.....	22
2.4.2 Krigagem com Tendência	24
2.5 MÉTODOS DE VALIDAÇÃO DE MODELOS.....	25
2.5.1 Critério de Akaike	26
2.5.2 Critério de Filliben	27
2.5.3 Validação Cruzada	29
2.5.4 Valor Máximo da Função Verossimilhança e Verossimilhança Restrita.....	31
2.6 TÉCNICAS DE DIAGNÓSTICO.....	32
2.6.1 Influência Local.....	33
2.6.1.1 Influência local no afastamento da verossimilhança (<i>Likelihood displacement</i>)	34
2.6.1.2 Influência local na matriz de covariância e no preditor linear	36

2.6.1.3	Influência local nos valores preditos (<i>Generalized leverage</i>).....	38
3	MATERIAL E MÉTODOS	40
3.1	ESTUDO COM DADOS SIMULADOS UTILIZANDO MONTE CARLO	40
3.2	ESTUDOS COM DADOS DA BIBLIOGRAFIA.....	44
3.3	ESTUDOS COM DADOS EXPERIMENTAIS	44
3.3.1	Localização e Caracterização das Áreas Experimentais	45
3.3.2	Amostragem	45
3.3.3	Variáveis Estudadas	47
3.3.3.1	Produtividade da cultura ($t\ ha^{-1}$).....	47
3.3.3.2	Densidade do solo ($Mg\ m^{-3}$).....	47
3.3.3.3	Resistência do solo à penetração (MPa).....	48
3.3.3.4	Fósforo ($mg\ dm^{-3}$) e manganês ($mg\ dm^{-3}$)	48
3.3.4	Análises de Dados Experimentais	49
4	RESULTADOS E DISCUSSÃO	50
4.1	ESTUDO COM DADOS SIMULADOS.....	50
4.2	ESTUDO COM DADOS DA BIBLIOGRAFIA	70
4.3	ESTUDOS COM DADOS EXPERIMENTAIS	73
4.3.1	Estudo do Teor Fósforo.....	74
4.3.1.1	Análise espacial do teor fósforo ($mg\ dm^{-3}$).....	74
4.3.1.2	Análises de diagnóstico	77
4.3.1.3	Análises de influência	82
4.3.1.3.1	Influência nas análises descritivas.....	82
4.3.1.3.2	Influência nas estimativas dos parâmetros	82
4.3.1.3.3	Influência na validação dos modelos	84
4.3.1.3.4	Influência na construção dos mapas temáticos	86
4.3.2	Estudo do Manganês	88
4.3.2.1	Análise espacial do manganês ($mg\ dm^{-3}$).....	88
4.3.2.2	Análises de diagnóstico	90
4.3.2.3	Análises de influência	95
4.3.2.3.1	Influência nas análises descritivas.....	95

4.3.2.3.2	Influência nas estimativas dos parâmetros	95
4.3.2.3.3	Influência na validação dos modelos	97
4.3.2.3.4	Influência na construção dos mapas temáticos	100
4.3.3	Estudo da Produtividade da Soja como Função de Covariáveis ..	101
4.3.3.1	Análises preliminares	102
4.3.3.2	Análises de diagnóstico	103
4.3.3.3	Análises de influência	108
4.3.3.3.1	Influência nas análises descritivas.....	108
4.3.3.3.2	Influência nas estimativas dos parâmetros	108
4.3.3.3.3	Influência na validação dos modelos	111
4.3.3.3.4	Influência na construção dos mapas temáticos	113
5	CONCLUSÕES.....	115
	REFERÊNCIAS	116
	APÊNDICES	121
	APÊNDICE A - CONJUNTOS DE DADOS SIMULADOS	122
	APÊNDICE B - ROTEIROS DE ANÁLISES NO PROGRAMA R E MÓDULO GEOESTATÍSTICO geoR	138
	ANEXO	150
	ANEXO A - TABELA DO COEFICIENTE DE CORRELAÇÃO r DE FILLIBEN	151

LISTA DE TABELAS

Tabela 1 -	Parâmetros estimados para a função exponencial com os estimadores de máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR).....	51
Tabela 2 -	Parâmetros estimados para a função gaussiana com os estimadores de máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR).....	56
Tabela 3 -	Parâmetros estimados para a função Matérn com os estimadores de máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR).....	61
Tabela 4 -	Resumo dos gráficos de diagnóstico que identificaram a observação perturbada como sendo influente	70
Tabela 5 -	Análise estatística descritiva para a variável $\ln(TCE)$	71
Tabela 6 -	Comparação entre os parâmetros de dependência espacial (φ_1 , φ_2 , $\varphi_1 + \varphi_2$ e φ_3) obtidos por PIEGORSCH e BAILER (2005), os calculados utilizando-se o programa <i>R</i> e retirando-se as observações influentes.....	72
Tabela 7 -	Estatística descritiva para a variável fósforo (P) [mg dm^{-3}].....	74
Tabela 8 -	Modelos espaciais ajustados e estimação dos parâmetros efeito pepita (φ_1), contribuição (φ_2) e função do alcance (φ_3), pelos métodos de máxima verossimilhança e máxima verossimilhança restrita, para a variável fósforo (mg dm^{-3}).....	76
Tabela 9 -	Análise estatística descritiva para a variável fósforo, com dados totais coletados (P) e retirando-se as observações 1, 26 e 45 (P-1-26-45).....	82
Tabela 10 -	Parâmetros estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR) para a variável fósforo, com dados	

	totais (P) e retirando-se as observações 1, 26 e 45 (P-1-26-45), utilizando os modelos teóricos exponencial, gaussiano e Matérn .	83
Tabela 11	- Resultados da validação pelos critérios de Akaike (\hat{A}), Filliben (r^*) e valor máximo da verossimilhança (VMV).....	84
Tabela 12	- Resultados da validação cruzada para a variável teor de fósforo com dados originais e sem dados influentes	86
Tabela 13	- Estatística descritiva para a variável teor de manganês (mg dm^{-3})	88
Tabela 14	- Modelos espaciais ajustados e estimação dos parâmetros efeito pepita (φ_1), contribuição (φ_2) e alcance (φ_3), pelos métodos de máxima verossimilhança e máxima verossimilhança restrita, para a variável teor de manganês (mg dm^{-3}).....	89
Tabela 15	- Análise estatística descritiva para a variável manganês, com dados totais coletados (Mn) e retirando-se as observações 5 e 6 (Mn-5-6)	95
Tabela 16	- Ajustes utilizando os modelos teóricos exponencial, gaussiano e Matérn, com parâmetros estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR) para a variável manganês, com dados totais e retirando-se as observações 5 e 6 (Sem influentes)	96
Tabela 17	- Resultados da validação dos modelos ajustados para a variável teor de manganês, pelos critérios de Akaike (\hat{A}), Filliben (r^*) e valor máximo do logaritmo da função verossimilhança (VMV).....	98
Tabela 18	- Resultados da validação cruzada para a variável teor de manganês com dados totais e sem dados influentes.....	99
Tabela 19	- Estatística descritiva para a variável produtividade (t ha^{-1})	102
Tabela 20	- Análise estatística descritiva para a variável produtividade, com dados totais coletados (Prod.) e retirando-se as observações 1, 26 e 45 (Prod-13-23-30).....	108
Tabela 21a	- Parâmetros β estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR) para a variável produtividade, com dados totais (Prod.) e retirando-se as	

observações 13, 23 e 30 (P-13-23-30), utilizando os modelos teóricos exponencial, gaussiano e Matérn.....	109
Tabela 21b - Parâmetros φ estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR) para a variável produtividade, com dados totais (Prod) e retirando-se as observações 13, 23 e 30 (P-13-23-30), utilizando os modelos teóricos exponencial, gaussiano e Matérn.....	110
Tabela 22 - Resultados dos critérios de Akaike (\hat{A}), Filliben (r^*) e valor máximo do logaritmo da função verossimilhança (VMV) para a produtividade com dados originais e sem dados influentes.....	111
Tabela 23 - Resultados da validação cruzada para a variável produtividade com dados originais e sem dados influentes	113

LISTA DE FIGURAS

Figura 1 -	Direções utilizadas na geoestatística em análises de dependência espacial.	6
Figura 2 -	Relação entre a covariância espacial $C(h)$ e a semivariância $\gamma(h)$	10
Figura 3 -	Área experimental de 57 ha e grade de amostragem para o ano agrícola de 2004/2005.	46
Figura 4 -	Área experimental de 71 ha e grade de amostragem para o ano agrícola de 2006/2007.	46
Figura 5 -	Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Exponencial 0-10-10, com os estimadores MV e MVR.	52
Figura 6 -	Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Exponencial 0-10-15, com os estimadores MV e MVR.	53
Figura 7 -	Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Exponencial 0-10-20, com os estimadores MV e MVR.	54
Figura 8 -	Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Exponencial 0-10-60, com os estimadores MV e MVR.	55
Figura 9 -	Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Gaussiano 0-10-10, com os estimadores MV e MVR.	57
Figura 10 -	Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Gaussiano 0-10-15, com os estimadores MV e MVR.	58

Figura 11 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Gaussiano 0-10-20, com os estimadores MV e MVR.....	59
Figura 12 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Gaussiano 0-10-60, com os estimadores MV e MVR.....	60
Figura 13 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-10 $k= 0,7$, com os estimadores MV e MVR.....	62
Figura 14 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-10 $k= 3,0$, com os estimadores MV e MVR.....	63
Figura 15 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-15 $k= 0,7$, com os estimadores MV e MVR.....	64
Figura 16 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-15 $k= 3,0$, com os estimadores MV e MVR.....	65
Figura 17 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-20 $k= 0,7$, com os estimadores MV e MVR.....	66
Figura 18 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-20 $k= 3,0$, com os estimadores MV e MVR.....	67
Figura 19 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-60 $k= 0,7$, com os estimadores MV e MVR.....	68
Figura 20 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-60 $k= 3,0$, com os estimadores MV e MVR.....	69
Figura 21 - Desenho esquemático <i>Box-plot</i> da variável $\ln(TCE)$	71

Figura 22 - Gráficos de diagnóstico $C_i \times i$, $ L_{max} \times i$, $lCd \times i$ e $GL \times i$, para a variável “ln (TCE) calculados”	72
Figura 23 - Mapas temáticos para as variáveis: “ln (TCE) calculados” a); “ln (TCE) sem influentes” b).	73
Figura 24 - Croqui da disposição amostral dos dados na área em estudo de 71 ha (a); Gráfico <i>Box-plot</i> da variável fósforo com os dados totais (b).....	75
Figura 25 - Gráficos de diagnóstico $C_i \times i$ e $ L_{max} \times i$, para a variável fósforo, com o estimador MV.....	78
Figura 26 - Gráficos de diagnóstico $lCd \times i$ e $GL \times i$, para a variável fósforo, com o estimador MV.....	79
Figura 27 - Gráficos de diagnóstico $C_i \times i$ e $ L_{max} \times i$, para a variável fósforo, com o estimador MVR.	80
Figura 28 - Gráficos de diagnóstico $lCd \times i$ e $GL \times i$, para a variável fósforo, com o estimador MVR.	81
Figura 29 - Mapas temáticos para a variável fósforo: a) com dados originais (P); b) retirando-se as observações 1, 26 e 45 (P-1-26-45).....	87
Figura 30 - Gráfico <i>Box-plot</i> da variável teor de manganês com os dados totais (a); Croqui da disposição amostral dos dados na área em estudo de 71 ha (b).....	89
Figura 31 - Gráficos de diagnóstico $C_i \times i$ e $ L_{max} \times i$, para a variável manganês, com o estimador MV.....	91
Figura 32 - Gráficos de diagnóstico $lCd \times i$ e $GL \times i$, para a variável manganês, com o estimador MV.....	92
Figura 33 - Gráficos de diagnóstico $C_i \times i$ e $ L_{max} \times i$, para a variável manganês, com o estimador MVR.	93
Figura 34 - Gráficos de diagnóstico $lCd \times i$ e $GL \times i$, para a variável manganês, com o estimador MVR.	94
Figura 35 - Mapas temáticos para a variável teor de manganês: a) com dados originais (Mn); b) retirando-se as observações 5 e 6 (Mn-5-6). ..	100

Figura 36 - Croqui da disposição amostral dos dados na área em estudo de 57 ha (a); Gráfico <i>Box-plot</i> da variável produtividade com os dados totais (b).....	102
Figura 37 - Gráficos de diagnóstico $C_i \times i$ e $ L_{max} \times i$, para a variável produtividade, com o estimador MV.	104
Figura 38 - Gráficos de diagnóstico $lCd \times i$ e $GL \times i$, para a variável produtividade, com o estimador MV.	105
Figura 39 - Gráficos de diagnóstico $C_i \times i$ e $ L_{max} \times i$, para a variável produtividade, com o estimador MVR.....	106
Figura 40 - Gráficos de diagnóstico $lCd \times i$ e $GL \times i$, para a variável produtividade, com o estimador MVR.....	107
Figura 41 - Mapas temáticos para a variável produtividade: a) com dados originais (Prod); b) retirando-se as observações 13, 23 e 30 (Prod-13-23-30).....	114

RESUMO

Conceitos de monitoramento e gerenciamento do processo de produção agrícola vêm sendo utilizados como ótima opção de estratégia gerencial na agricultura. Tais conceitos consideram a variabilidade espacial das variáveis em estudo. A modelagem da estrutura de dependência espacial pela abordagem da geoestatística é de fundamental importância para a definição de parâmetros que definem esta estrutura e que são utilizados na interpolação de valores em locais não amostrados, pela técnica de krigagem. Entretanto, a estimação de parâmetros pode ser muito afetada pela presença de observações atípicas nos dados amostrados. O desenvolvimento deste trabalho teve por objetivo utilizar técnicas de diagnóstico em modelos espaciais lineares gaussianos, utilizados em geoestatística, para avaliar a sensibilidade dos estimadores máxima verossimilhança e máxima verossimilhança restrita a pequenas perturbações nos dados. Realizaram-se estudos com dados simulados, com dados da bibliografia e também com dados experimentais, coletados em uma área agrícola comercial da região Oeste do Paraná. O estudo com dados simulados mostrou que as técnicas de diagnóstico utilizadas foram eficientes na identificação da perturbação nos dados. O estimador de máxima verossimilhança restrita produziu estimativas mais robustas para os parâmetros de dependência espacial. Pelos resultados obtidos com o estudo de dados reais, concluiu-se que a presença de valores atípicos entre os dados amostrados pode exercer forte influência nos mapas temáticos, alterando, assim, a dependência espacial. A aplicação de técnicas de diagnóstico deve fazer parte de toda análise geoestatística, garantindo que as informações contidas nos mapas temáticos tenham maior qualidade e possam ser utilizadas com maior segurança pelo agricultor.

PALAVRAS-CHAVE: Influência local, geoestatística, máxima verossimilhança, máxima verossimilhança restrita.

ABSTRACT

DIAGNOSTICS TECHNIQUES IN SPATIAL LINEAR GAUSSIANS MODELS

Tracking and management concepts of the process of agricultural production are being used as a great option of strategy management in agriculture. Such concepts consider the spatial variability of the variables at study. The modeling of the spatial dependence structure of the geostatistic approach is fundamental importance for the definition parameters that define this structure and are used in the interpolation of values in places not sampled, by kriging techniques. However, the estimation of parameters can be greatly affected by the presence atypical observations in the data sampled. The development of this work was aimed at using diagnostics techniques in spatial linear gaussians models, used in geostatistics, to evaluate the sensitivity of the maximum likelihood estimators and restrict maximum likelihood to small perturbations in the data. Studies were performed with simulated data, with literature data and with experimental data, collected in a commercial agricultural area in the region West of Paraná. The study with simulated data showed that the techniques used in diagnostics were efficient in identifying the perturbation data. The restrict maximum likelihood estimator produced more robust estimates for the parameters spatial dependence. Those results obtained from the study of real data, it was concluded that the presence atypical values between the sampled data can exert strong influence on thematic maps, changing, therefore, the spatial dependence. The application the diagnostic techniques should be part of any geostatistic analysis, ensuring that the information contained in thematic maps have better quality and can be used with greater security by the farmer.

KEYWORDS: Local influence, geostatistic, maximum likelihood, restrict maximum likelihood.

1 INTRODUÇÃO

Nas últimas décadas, conceitos de monitoramento e gerenciamento do processo de produção agrícola vêm sendo amplamente discutidos e aplicados, gerando grande quantidade de informações, obtidas pelo estudo dos fatores que interferem na produtividade.

Tais conceitos consideram a variabilidade espacial das variáveis georreferenciadas, principalmente aquelas relacionadas ao solo, como atributos físicos do solo (compactação e densidade do solo, teor de água, etc.), atributos químicos do solo (fósforo, potássio, magnésio, matéria orgânica, pH, etc.), e também atributos da planta (produtividade, altura, número de vagens, etc.).

Neste sentido, o estudo da variabilidade espacial é importante para o conhecimento da forma com que se distribuem espacialmente os atributos do solo e da planta na área em estudo. Segundo CRESSIE (1993), a não consideração da variabilidade espacial pode impedir que as diferenças reais sejam verificadas, o que impossibilitaria um tratamento diferenciado, de acordo com as exigências locais.

A geoestatística, que tem como fundamento a teoria das variáveis regionalizadas, surge como um dos métodos que considera a distribuição espacial das medidas, permitindo definir o raio de autocorrelação espacial entre amostras e, conseqüentemente, a distância máxima em que as amostras são dependentes.

A modelagem da estrutura de dependência espacial é de fundamental importância para a definição de parâmetros que a definem e que são utilizados na interpolação de valores não amostrados, pela técnica de krigagem, conduzindo a erros mínimos de predição e variância mínima na interpolação. Então, para que a interpolação produza previsões confiáveis o processo de modelagem deve ser feito com muita cautela, principalmente na presença de pontos discrepantes ou influentes.

Se o modelo ajustado aos dados não apresentar boa descrição do comportamento dos valores que foram observados, a utilização do modelo pode conduzir a inferências errôneas. Por esse motivo, a aplicação de técnicas de diagnóstico é muito valiosa, visto que consiste de métodos para avaliar o grau de sensibilidade das inferências a pequenas perturbações nos dados ou no modelo proposto.

Este trabalho justifica-se pela importância que a modelagem da variabilidade espacial representa, haja vista que este processo fornece parâmetros que serão utilizados no método de interpolação por krigagem.

A partir da interpolação por krigagem são gerados mapas temáticos que poderão ser utilizados para uma aplicação localizada de insumos ou manejo localizado do solo. E a qualidade dos mapas depende da qualidade das inferências dos modelos ajustados. Portanto, se o processo de ajuste for mal conduzido, resultará em mapas que não representam a real variabilidade local.

O objetivo deste trabalho foi utilizar técnicas de diagnóstico em modelos espaciais lineares gaussianos para avaliar a influência potencial de dados atípicos na estimativa dos parâmetros que definem a estrutura de dependência espacial e indicar os modelos mais robustos. Para tanto, foram realizados estudos sobre a influência local, utilizando-se os métodos de estimação de parâmetros por máxima verossimilhança e máxima verossimilhança restrita, para estudar a sensibilidade dos modelos na presença de valores influentes, por meio de técnicas de diagnóstico, tais como gráfico dos coeficientes de influências C_i , gráfico dos autovetores $|L_{max}|$, gráfico dos coeficientes de máxima inclinação local l_{Ch} e gráfico de alavanca generalizada GL .

2 REVISÃO BIBLIOGRÁFICA

Pesquisas mostram que uma área agrícola pode ter variações consideráveis em seus atributos físicos e químicos no solo. Com a utilização de grandes extensões de áreas destinadas à agricultura, as diferenças entre as necessidades de manejo do solo para a cultura e a taxa de aplicação de insumos empregada em função da média, tendem a ser maiores.

Neste processo, é importante o estudo da variabilidade espacial por meio dos métodos geoestatísticos, pois auxiliam no monitoramento e gerenciamento do processo de produção agrícola.

Dentre os trabalhos desenvolvidos, na região oeste do estado do Paraná, utilizando geoestatística, destacam-se SOUZA *et al.* (1999) que estudaram a variabilidade espacial de atributos químicos do solo em um Latossolo Vermelho Distroférrico da região Oeste do Paraná, em uma área de 1,33 ha; RIBEIRO JR (2002) que desenvolveu estudos estatísticos espaciais com modelos e métodos para análise de dados georreferenciados; SILVA *et al.* (2003) que confeccionaram mapas temáticos empregando o estimador de Cressie & Hawkins na construção de semivariogramas para as variáveis fósforo e potássio do solo e produtividade de soja, que apresentaram distribuição assimétrica; JOHANN *et al.* (2004) que utilizaram métodos geoestatísticos para confecção de mapas temáticos, pela interpolação por krigagem ordinária, da variabilidade espacial de atributos físicos do solo e da produtividade de soja, em uma área da região Oeste do Paraná; URIBE-OPAZO e KAVANAGH (2006) que estudaram ajustes de modelos teóricos a semivariogramas experimentais utilizados para modelar a estrutura de dependência espacial.

Todos os trabalhos supracitados mostram a importância de estudos acerca da variabilidade espacial de atributos físicos e químicos do solo e

atributos da planta, com o objetivo de melhorar a qualidade da confecção de mapas temáticos que representam tal variabilidade.

2.1 A TEORIA DAS VARIÁVEIS REGIONALIZADAS

A geoestatística tem como princípio que amostras de uma variável regionalizada não se dispõem independentemente até certa distância, elas têm influência sobre o ponto a ser predito. Assim, a geoestatística é utilizada quando é verificada a existência de dependência espacial entre as amostras e o uso da variância não tem sentido, se não levar em conta as distâncias entre as amostras.

O sucesso e as suposições das técnicas geoestatísticas estão fundamentadas na teoria das Variáveis Regionalizadas – VR, desenvolvida por Matheron, na década de 60. A variável regionalizada é uma variável distribuída no espaço (ou tempo), ou seja, são variáveis cujos valores são relacionados de algum modo com a posição espacial que ocupam (GUERRA, 1988).

A teoria das variáveis regionalizadas diz que alguma medida pode ser vista como realização de uma função aleatória, ou processo estocástico. Esta teoria consolida o alicerce da geoestatística.

2.1.1 Características das Variáveis Regionalizadas

As variáveis regionalizadas possuem características qualitativas estreitamente ligadas à estrutura do fenômeno natural que elas representam. Tais características são:

a) Localização

As variáveis regionalizadas somente tomam seus valores importantes no campo geométrico em que foram definidas, ou seja, no espaço geométrico onde as variáveis são susceptíveis de tomarem valores definidos e no interior do qual serão estudadas suas variações. A escolha do campo geométrico, onde será realizada a medição da variável, não deve ser feita de forma arbitrária, devendo respeitar certas condições de homogeneidade física (GUERRA, 1988).

b) Continuidade

A continuidade espacial apresentada pela VR entre amostras vizinhas reflete o grau de dependência ou independência espacial entre as amostras. Espera-se que amostras mais próximas apresentem maior dependência do que as que estão mais distantes (CRESSIE, 1993).

Segundo ISAAKS e SRIVASTAVA (1989), a teoria fundamental da geoestatística é a esperança de que, na média, as amostras próximas no tempo e espaço sejam mais similares entre si do que as que estiverem distantes.

Quando a continuidade espacial não é identificada, diz-se que há presença de “efeito de pepita puro”. Isto é, a distância entre as amostras pode ter sido muito grande a ponto de não identificar a variabilidade espacial da variável em estudo.

c) Isotropia

A isotropia é uma característica essencial da VR e é identificada quando a variabilidade espacial em uma área apresenta comportamento semelhante para distintas direções, ou seja, não há uma direção privilegiada, ao longo da qual os valores variem de forma significativa. Quando isso não ocorre diz-se que há anisotropia.

Normalmente, as direções consideradas são 0° , 45° , 90° e 135° , conforme a Figura 1.

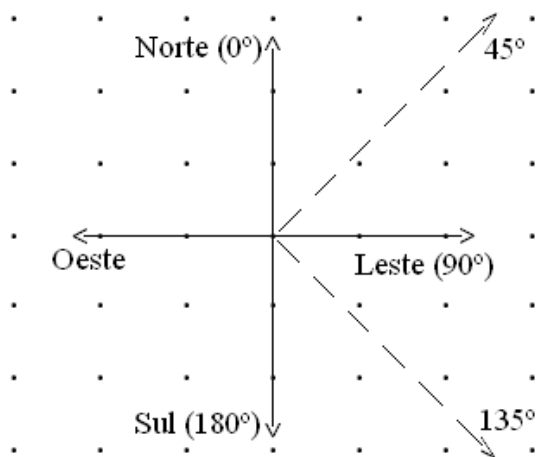


Figura 1 - Direções utilizadas na geoestatística em análises de dependência espacial.

A existência de direções que apresentam distintas estruturas de dependência, ou que não apresentam a mesma estrutura de continuidade espacial, caracteriza a anisotropia, que pode ser identificada pelo gráfico semivariograma, estimado para diferentes direções.

Segundo SOARES (2000), os modelos de anisotropia mais comuns são anisotropia geométrica, anisotropia zonal e combinada (geométrica e zonal).

A anisotropia geométrica caracteriza-se quando os semivariogramas apresentam o mesmo modelo com o mesmo patamar em todas as direções, mas com diferentes alcances, verificando-se os alcances máximos e mínimos em direções perpendiculares (exemplo: direções 0° e 90°). A anisotropia zonal caracteriza-se quando o semivariograma, construído para diferentes direções, apresenta diferentes patamares.

Estudo detalhado acerca da anisotropia na variabilidade espacial foi realizado por CARVALHO *et al.* (2007), com dados químicos do solo, em uma área experimental, localizada na região Oeste do Paraná.

2.2 MODELOS ESPACIAIS LINEARES

MARDIA e MARSHALL (1984) consideraram um processo estocástico gaussiano $\{Z(s_i), s_i \in S\}$, em que $S \subset \mathfrak{R}^d$, sendo \mathfrak{R}^d espaço euclidiano, d -dimensional ($d \geq 1$). Ou seja, o processo $Z = (Z(s_1), \dots, Z(s_n))^T$, em que s_i e s_u ($i, u = 1, \dots, n$) são localizações espaciais conhecidas, têm distribuição gaussiana n -variada com vetor de médias $\mu \mathbf{1}$ e matriz de covariância Σ , isto é, $Z \sim N_n(\mu \mathbf{1}, \Sigma)$, em que μ é uma constante; $\mathbf{1}$ é um vetor de uns de ordem $n \times 1$ e Σ é uma matriz definida positiva, de ordem n , dada por $\Sigma = [Cov(Z(s_i), Z(s_u))]$.

Supondo-se que os dados são descritos pelo modelo da Equação (1):

$$Z(s_i) = \mu(s_i) + \varepsilon(s_i) \quad (1)$$

em que os termos determinístico $\mu(s_i)$ e estocástico $\varepsilon(s_i)$ podem depender da localização espacial em que $Z(s_i)$ foi obtida. Assume-se que o erro estocástico ε tem média zero, $E[\varepsilon(s_i)] = 0$, e que a variação entre pontos no espaço é determinada por alguma função de covariância $C(s_i, s_u) = Cov\{\varepsilon(s_i), \varepsilon(s_u)\}$.

Assume-se também que para algumas funções conhecidas de s , como $w_1(s), w_2(s), \dots, w_p(s)$, a média do processo estocástico é dada por

$$\mu(s) = \sum_{u=1}^p w_u(s) \beta_u \quad (2)$$

em que: β_1, \dots, β_p são parâmetros desconhecidos e a serem estimados.

A função de covariância $C(s_i, s_u)$ é também especificada por um vetor q -dimensional $\varphi = (\varphi_1, \dots, \varphi_q)^T$. Por simplicidade, pode-se utilizar as seguintes notações: $Z(s_i) = z_i$, $Z = (z_1, \dots, z_n)^T$, $w_{iu} = w_u(s_i)$, $w_i^T = (w_{i1}, \dots, w_{ip})$, em que W é uma matriz $n \times p$ com suas linhas w_i^T , $\beta = (\beta_1, \dots, \beta_p)^T$, $\varepsilon_i = \varepsilon(s_i)$, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$, com $i = 1, \dots, n$ e $u = 1, \dots, p$.

Desta forma, $\mu(s_i) = w_i^T \beta$, e então,

$$z_i = w_i^T \beta + \varepsilon_i. \quad (3)$$

De forma equivalente, em notação matricial, tem-se:

$$Z = W \beta + \varepsilon \quad (4)$$

Então, $E(\varepsilon) = 0$ e a matriz de covariância de ε é $\Sigma = [(\sigma_{iu})]$, em que $\sigma_{iu} = C(s_i, s_u)$. Assume-se que Σ é não singular e que W tem colunas com posto Completo.

Considerando-se de maneira particular a forma paramétrica da matriz de covariância,

$$\Sigma = \varphi_1 I_n + \varphi_2 R \quad (5)$$

em que:

φ_1 : efeito pepita ou erro de variância (C_0);

φ_2 : contribuição ou variância de dispersão ($C1$);

R : é uma matriz que é função de φ_3 , $R=R(\varphi_3)=[(r_{iu})]$, matriz $n \times n$ simétrica com seus elementos da diagonal $r_{iu} = 1$, $i = 1, \dots, n$, em que φ_3 é função do alcance (a) do modelo;

I_n : matriz identidade de ordem n .

A forma paramétrica da matriz de covariância, dada na Equação 5, ocorre para vários processos isotrópicos, nos quais a covariância $C(s_i, s_u)$ é definida segundo a função de covariâncias $C(h_{iu}) = \varphi_2 r_{iu}$, em que $h_{iu} = ||s_i - s_u||$ é a distância euclidiana entre os pontos s_i e s_u .

Nas funções de covariâncias $C(h_{iu})$, a variância do processo estocástico Z é $C(0) = \varphi_1 + \varphi_2$.

Assim, o semivariograma pode ser definido como:

$$\gamma(h) = C(0) - C(h) \quad (6)$$

2.2.1 Parâmetros do Semivariograma

Considerando-se a Equação (6), os parâmetros que definem a estrutura de dependência espacial são:

- i) Efeito pepita ($\varphi_1 = C_0$): na teoria, quando a distância h é zero, o valor da semivariância é igual a zero, porém, na prática, isto não acontece e à medida que h aproxima-se de zero, o valor do semivariograma aproxima-se de um valor positivo chamado efeito pepita (φ_1). Este revela a descontinuidade do semivariograma para distâncias menores que a menor distância entre as amostras. Conforme relatou CRESSIE (1993), essa descontinuidade pode ser gerada por erros de análise ou de variabilidade de pequena escala não captada pela amostragem;
- ii) Contribuição ($\varphi_2 = C_1$): é denominada variância de dispersão e representa as diferenças espaciais entre os valores de uma variável tomada em dois pontos separados por distâncias cada vez maiores;
- iii) Alcance ($a = g(\varphi_3)$): é a distância dentro da qual as amostras apresentam autocorrelação espacial, ou seja, o alcance a marca a distância a partir da qual um ponto da variável em estudo não tem mais influência sobre o ponto vizinho; g representa uma função;
- iv) Patamar ($C = \varphi_1 + \varphi_2$): é o valor máximo do semivariograma correspondente ao seu alcance, ou seja, a variância total da variável em estudo. O patamar é o valor a partir do qual o semivariograma estabiliza-se.

Na Figura 2 é apresentado um exemplo gráfico da relação entre a covariância espacial $C(h)$ e a semivariância $\gamma(h)$.

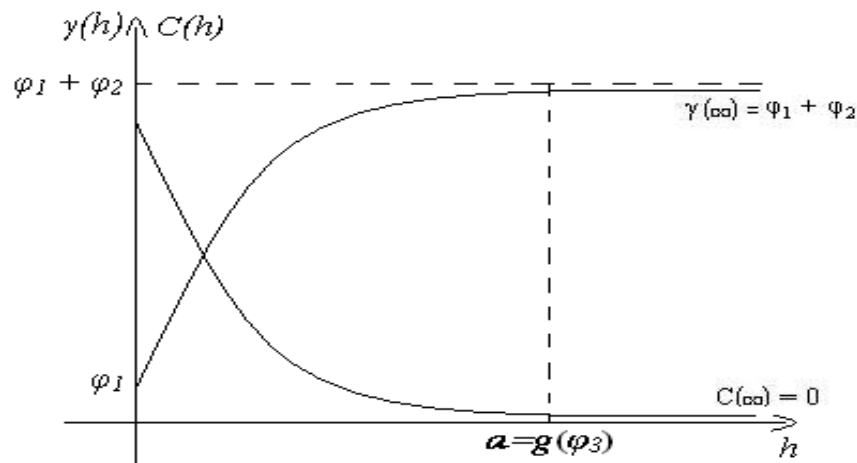


Figura 2 - Relação entre a covariância espacial $C(h)$ e a semivariância $\gamma(h)$.

2.2.2 Modelos Teóricos de Semivariograma, Funções de Covariância e de Correlação

Apresentados os parâmetros obtidos com o estudo da dependência espacial, a seguir são apresentados alguns dos modelos teóricos de semivariogramas, funções de covariâncias e funções de correlações espaciais mais utilizados na bibliografia.

Observa-se que a partir da função semivariância apresentada na Equação (6), tem-se a relação $C(h) = C(0) - \gamma(h)$, em que $C(0) = \varphi_1 + \varphi_2$ é a variância total (patamar).

Além disso, dividindo por $C(0)$, tem-se:

$$\frac{C(h)}{C(0)} = \frac{C(0)}{C(0)} - \frac{\gamma(h)}{C(0)}$$

$$\frac{C(h)}{C(0)} = 1 - \frac{\gamma(h)}{C(0)} \quad (7)$$

Logo, a função de correlação espacial é da forma:

$$\rho(h) = 1 - \frac{\gamma(h)}{\varphi_1 + \varphi_2}. \quad (8)$$

Os modelos ajustados para o semivariograma experimental, apresentados por ISAACS e SRIVASTAVA (1989) e CRESSIE (1993), são divididos em modelos transitivos, que possuem patamar, e modelos não transitivos, que não possuem patamar. Os principais modelos transitivos utilizados são: esférico, exponencial, gaussiano, circular, Matérn e exponencial potência. Segundo JOURNEL e HUIJBREGTS (1978), estes modelos consideram que a variável em estudo possui um mesmo padrão de continuidade espacial em todas as direções (isotropia), em consequência disso, o semivariograma deve ter uma mesma característica estrutural em todas as direções do espaço.

a) Modelo esférico

Este modelo apresenta crescimento rápido na origem e atinge o patamar a 2/3 do alcance. Segundo CRESSIE (1993), este modelo é válido em \mathfrak{R} , \mathfrak{R}^2 e \mathfrak{R}^3 e tem como expressão a Equação (9):

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \varphi_1 + \varphi_2 \left[\frac{3}{2} \left(\frac{h}{\varphi_3} \right) - \frac{1}{2} \left(\frac{h}{\varphi_3} \right)^3 \right], & 0 < h \leq \varphi_3 \\ \varphi_1 + \varphi_2, & h > \varphi_3 \end{cases} \quad (9)$$

A função de covariância é expressa por:

$$C(h) = \begin{cases} \varphi_1 + \varphi_2, & h = 0 \\ \varphi_2 \left[1 - \frac{3}{2} \left(\frac{h}{\varphi_3} \right) + \frac{1}{2} \left(\frac{h}{\varphi_3} \right)^3 \right], & 0 < h \leq \varphi_3 \\ 0, & h > \varphi_3 \end{cases} \quad (10)$$

A função de correlação espacial é definida como:

$$\rho(h) = \begin{cases} 1, & h = 0 \\ 1 - \frac{3}{2} \left(\frac{h}{\varphi_3} \right) + \frac{1}{2} \left(\frac{h}{\varphi_3} \right)^3, & 0 < h \leq \varphi_3 \end{cases} \quad (11)$$

É importante destacar que o modelo esférico não tem segunda derivada.

b) Modelo exponencial

Este modelo apresenta comportamento aproximadamente linear na origem e atinge o patamar assintoticamente com alcance prático definido como a distância na qual o valor do modelo é 95% de φ_2 , sendo o alcance prático dado por $a = 3\varphi_3$. Este modelo é válido em \mathfrak{R} , \mathfrak{R}^2 e \mathfrak{R}^3 (CRESSIE, 1993), e tem como expressão a Equação (12)

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \varphi_1 + \varphi_2 \left[1 - \exp\left(-\frac{h}{\varphi_3}\right) \right], & 0 < h \leq \varphi_3 \\ \varphi_1 + \varphi_2, & h > \varphi_3 \end{cases} \quad (12)$$

A função de covariância é expressa por:

$$C(h) = \begin{cases} \varphi_1 + \varphi_2, & h = 0 \\ \varphi_2 \left[\exp\left(-\frac{h}{\varphi_3}\right) \right], & 0 < h \leq \varphi_3 \\ 0, & h > \varphi_3 \end{cases} \quad (13)$$

A função de correlação espacial é definida como:

$$\rho(h) = \begin{cases} 1, & h = 0 \\ \exp\left(-\frac{h}{\varphi_3}\right), & 0 < h \leq \varphi_3 \end{cases} \quad (14)$$

c) Modelo gaussiano

Este modelo apresenta comportamento parabólico na origem e é utilizado para modelar um fenômeno extremamente contínuo. Também atinge o patamar apenas assintoticamente e o alcance prático é dado por $a = \sqrt{3}\varphi_3$. Este modelo é válido em \mathfrak{R} , \mathfrak{R}^2 e \mathfrak{R}^3 (CRESSIE, 1993) e tem como expressão a Equação (15):

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \varphi_1 + \varphi_2 \left\{ 1 - \exp \left[- \left(\frac{h}{\varphi_3} \right)^2 \right] \right\}, & 0 < h \leq \varphi_3 \\ \varphi_1 + \varphi_2, & h > \varphi_3 \end{cases} \quad (15)$$

A função de covariância é expressa por:

$$C(h) = \begin{cases} \varphi_1 + \varphi_2, & h = 0 \\ \varphi_2 \left\{ \exp \left[- \left(\frac{h}{\varphi_3} \right)^2 \right] \right\}, & 0 < h \leq \varphi_3 \\ 0, & h > \varphi_3 \end{cases} \quad (16)$$

A função de correlação espacial é definida como:

$$\rho(h) = \begin{cases} 1, & h = 0 \\ \exp \left[- \left(\frac{h}{\varphi_3} \right)^2 \right], & 0 < h \leq \varphi_3 \end{cases} \quad (17)$$

d) Modelo logarítmico ou de Wijs

Este é um modelo sem patamar, e foi um dos primeiros modelos aplicados à geoestatística (GUERRA, 1988), devido às suas propriedades analíticas. Entretanto, é inadequado para $h = 1$, e a função semivariância tem como expressão:

$$\gamma(h) = \varphi_1 + 3\alpha \ln(h), \quad (18)$$

em que α é uma constante conhecida como dispersão absoluta.

A função de covariância é expressa por:

$$C(h) = \varphi_2 - 3\alpha \ln(h), \quad (19)$$

A função de correlação espacial é definida como:

$$\rho(h) = \frac{\varphi_2 - 3\alpha \ln(h)}{\varphi_1 + \varphi_2}. \quad (20)$$

e) Modelo circular

O modelo teórico de semivariograma circular é válido para os espaços R e R^2 , mas não para R^3 , e tem como expressão a Equação (21):

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \varphi_1 + \varphi_2 \left[1 - \frac{2}{\pi} \cos^{-1}\left(\frac{h}{\varphi_3}\right) + \frac{2h}{\pi\varphi_3} \sqrt{1 - \frac{h^2}{\varphi_3^2}} \right], & 0 < h \leq \varphi_3 \\ \varphi_1 + \varphi_2, & h > \varphi_3 \end{cases} \quad (21)$$

A função de covariância é expressa pela Equação (22):

$$C(h) = \begin{cases} \varphi_1 + \varphi_2, & h = 0 \\ \varphi_2 \frac{2}{\pi} \left[\cos^{-1}\left(\frac{h}{\varphi_3}\right) - \frac{h}{\varphi_3} \sqrt{1 - \frac{h^2}{\varphi_3^2}} \right], & 0 < h \leq \varphi_3 \end{cases} \quad (22)$$

A função de correlação espacial é definida pela Equação (23):

$$\rho(h) = \begin{cases} 0, & h > \varphi_3 \\ \frac{2}{\pi} \left[\cos^{-1}\left(\frac{h}{\varphi_3}\right) - \frac{h}{\varphi_3} \sqrt{1 - \frac{h^2}{\varphi_3^2}} \right], & 0 \leq h \leq \varphi_3 \end{cases} \quad (23)$$

f) Modelo Matérn

MATÉRN (1986) apresentou uma função chamada família Matérn. Esta função é definida, em termos de modelo teórico de semivariograma, pela Equação (24):

$$\gamma(h) = \begin{cases} 0, & h = 0 \\ \varphi_1 + \varphi_2 \left[\left(1 - 2^{k-1} \Gamma(k) \right)^{-1} \left(\frac{h}{\varphi_3} \right)^k K_k \left(\frac{h}{\varphi_3} \right) \right], & h > 0 \end{cases} \quad (24)$$

em que:

$\varphi_1, \varphi_2, \varphi_3$ e k , são parâmetros;

K_k é a função de Bessel de terceiro tipo, de ordem k .

A família Matérn é válida para $\varphi_3, k > 0$ e corresponde à função de correlação exponencial quando $k = 0,5$; e à função de correlação gaussiana quando $k \rightarrow \infty$.

A função de covariância é expressa pela Equação (25):

$$C(h) = \begin{cases} 0, & h = 0 \\ \varphi_2 \left[\left(2^{k-1} \Gamma(k) \right)^{-1} \left(\frac{h}{\varphi_3} \right)^k K_k \left(\frac{h}{\varphi_3} \right) \right], & h > 0 \end{cases} \quad (25)$$

A função de correlação espacial é expressa pela Equação (26):

$$\rho(h) = \left\{ 2^{k-1} \Gamma(k) \right\}^{-1} \left(\frac{h}{\varphi_3} \right)^k K_k \left(\frac{h}{\varphi_3} \right), \quad h > 0, \quad (26)$$

g) Modelo exponencial-potência

O modelo exponencial-potência tem como expressão a Equação (27):

$$\gamma(h) = \begin{cases} \varphi_1 + \varphi_2, & h = 0 \\ \varphi_1 + \varphi_2 \left\{ 1 - \exp \left[- \left| \frac{h}{\varphi_3} \right|^\kappa \right] \right\}, & h > 0 \end{cases} \quad (27)$$

A função de covariância é expressa pela Equação (28):

$$C(h) = \begin{cases} \varphi_1 + \varphi_2, & h = 0 \\ \varphi_2 \exp \left[- \left| \frac{h}{\varphi_3} \right|^\kappa \right], & h > 0 \end{cases} \quad (28)$$

O modelo gaussiano é um caso especial do modelo exponencial-potência, quando $k = 2$. Já para $k = 1$, este modelo é equivalente ao exponencial.

O modelo exponencial-potência tem como função de correlação a expressão:

$$\rho(h) = \exp\left[-\left|\frac{h}{\varphi_3}\right|^\kappa\right], \quad 0 < \kappa \leq 2 \quad (29)$$

h) Modelos *wave (hole effect)* ou esquemas periódicos

Estas estruturas normalmente são ignoradas na maioria dos estudos, mas sua presença fornece informações valiosas a respeito da variabilidade espacial, pois, freqüentemente, indica um padrão de periodicidade, o que é uma característica espacial comum e legítima na geologia.

h.1) Efeito seno

Este modelo teórico de semivariograma tem como expressão a Equação (30):

$$\gamma(h) = \varphi_1 + \varphi_2 \left[1 - \frac{\text{sen}(h)}{h}\right], \quad h > 0 \quad (30)$$

A função de covariância é expressa pela Equação (31):

$$C(h) = \varphi_2 \left(\frac{\text{sen } h}{h}\right), \quad h > 0 \quad (31)$$

A função de correlação espacial é definida pela Equação (32):

$$\rho(h) = \left(\frac{\text{sen } h}{h}\right), \quad h > 0 \quad (32)$$

h.2) Efeito cosseno sem atenuação

Na Equação (33) é apresentado o modelo teórico de semivariograma para o efeito cosseno sem atenuação, em que p é o período.

$$\gamma(h) = \varphi_1 + \varphi_2 \left[1 - \cos\left(\frac{2\pi h}{p}\right) \right], \quad h > 0 \quad (33)$$

A função de covariância é expressa por:

$$C(h) = \varphi_2 \left(\cos\left(\frac{2\pi h}{p}\right) \right), \quad h > 0 \quad (34)$$

A função de correlação espacial é definida como:

$$\rho(h) = \left(\cos\left(\frac{2\pi h}{p}\right) \right), \quad h > 0 \quad (35)$$

h.3) Efeito cosseno com atenuação

Na Equação (36) é apresentado o modelo teórico de semivariograma para o efeito cosseno com atenuação.

$$\gamma(h) = \varphi_1 + \varphi_2 \left[1 - \cos\left(\frac{2\pi h}{p}\right) \exp\left(-\frac{h}{\lambda}\right) \right], \quad h > 0 \quad (36)$$

em que:

p : é o período;

λ : é o coeficiente de atenuação.

O período p é obtido no *eixo-h*, correspondendo ao valor que a senóide ou cossenóide se repete; o coeficiente λ de atenuação é obtido experimentalmente. Quanto maior o valor de λ , maior será a atenuação da cossenóide.

A função de covariância é expressa por:

$$C(h) = \varphi_2 \left[\cos\left(\frac{2\pi h}{p}\right) \exp\left(-\frac{h}{\lambda}\right) \right], \quad h > 0 \quad (37)$$

A função de correlação espacial é definida como:

$$\rho(h) = \cos\left(\frac{2\pi h}{p}\right) \exp\left(-\frac{h}{\lambda}\right), \quad h > 0 \quad (38)$$

2.3 ESTIMAÇÃO DE PARÂMETROS

2.3.1 Estimador de Máxima Verossimilhança – MV

O método de estimação de máxima verossimilhança (MV), considerado um estimador de parâmetros assintoticamente normal, não viciado e eficiente (COX; HINKLEY, 1974), consiste em maximizar a função de densidade de probabilidade conjunta do processo $Z(s)$, em relação aos efeitos fixos (processo estacionário) e aos componentes dos efeitos aleatórios. McBRATNEY e WEBSTER (1986) utilizaram esse método de ajuste de modelos e também afirmaram ser o mais eficiente. MARDIA e MARSHALL (1984) desenvolveram uma teoria para o caso em que o termo aleatório $\varepsilon(s)$ fosse um processo gaussiano.

O melhor modelo para um processo $Z(s)$ será aquele que apresentar o maior valor de maximização do logaritmo da função verossimilhança (log-verossimilhança).

A função de verossimilhança de n variáveis aleatórias Z_1, \dots, Z_n é definida como a densidade conjunta dada por $f_{Z_1}, \dots, f_{Z_n}(z_1, \dots, z_n; \theta)$, que deverá ser considerada como uma função do vetor de q -parâmetros desconhecidos $\theta = (\theta_1, \dots, \theta_q) \in \Theta$ (espaço paramétrico).

Seja (Z_1, \dots, Z_n) um vetor aleatório n -dimensional de uma população com função densidade de probabilidade conjunta $f_{Z_1}, \dots, f_{Z_n}(z_1, \dots, z_n; \theta)$, então a função de verossimilhança será:

$$L(\theta) = f_{Z_1}, \dots, f_{Z_n}(z_1, \dots, z_n; \theta). \quad (39)$$

A função de verossimilhança pode ser entendida como a intensidade de contribuições dos parâmetros na produção de uma dada amostra, ou seja, o quão fortemente os dados suportam os parâmetros desconhecidos.

Sob distribuição gaussiana n -variada dos erros, isto é, se $\varepsilon \sim N_n(0I, \Sigma)$, então, pela Equação (4) tem-se:

$$Z \sim N_n(W\beta, \Sigma),$$

os parâmetros desconhecidos do modelo $\hat{\theta} = (\beta_1, \dots, \beta_p, \varphi_1, \dots, \varphi_q)^T$ podem ser estimados maximizando-se a função de verossimilhança, como considerado por KATANIDIS (1983) e MARDIA e MARSHALL (1984).

Por motivos de simplicidade de cálculos, utiliza-se o logaritmo da função verossimilhança, chamada de função log-verossimilhança, definida a seguir:

$$l(\theta) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} (Z - W\beta)^T \Sigma^{-1} (Z - W\beta), \quad (40)$$

em que $\theta = (\beta^T, \varphi^T)^T$.

A função de verossimilhança é, basicamente, uma função do vetor de parâmetros $\theta \in \Theta$, em que Θ é o espaço paramétrico. Então a melhor estimativa

para o vetor θ de parâmetros será aquela que maximiza a função de verossimilhança ou de logaritmo da função verossimilhança, isto é:

$$l(\hat{\theta}) = \max l(\theta), \theta \in \Theta$$

em que $\hat{\theta}$ é o estimador de máxima verossimilhança.

2.3.2 Estimador de Máxima Verossimilhança Restrita – MVR

Outro método utilizado para estimar os parâmetros φ^T da função de covariância Σ é a máxima verossimilhança restrita – MVR (ZIMMERMAN, 1989; CRESSIE, 1993; CHRISTENSEN; JOHNSON; PEARSON, 1993). Neste caso, o logaritmo da função de verossimilhança restrita, também chamada de função log-verossimilhança restrita é dada pela Equação (41):

$$l_R(\theta) = -\frac{n-p}{2} \log(2\pi) + \frac{1}{2} \log(|W^T W|) - \frac{1}{2} \log|\Sigma| - \frac{1}{2} \log(|W^T \Sigma^{-1} W|) - \frac{1}{2} Z^T P Z, \quad (41)$$

em que $P = \Sigma^{-1}(I - A)$, com $A = W(W^T \Sigma^{-1} W)^{-1} W^T \Sigma^{-1}$.

Então, a melhor estimativa para o vetor de parâmetros $\theta \in \Theta$, será aquela que maximiza a função de verossimilhança restrita ou de log-verossimilhança restrita, isto é:

$$l_R(\hat{\theta}) = \max l_R(\theta), \theta \in \Theta.$$

em que $\hat{\theta}$ é o estimador de máxima verossimilhança restrita.

2.4 KRIGAGEM

A geoestatística tem um método de interpolação chamado krigagem, o qual usa a dependência espacial, expressa no semivariograma, para estimar valores em qualquer posição dentro de um espaço amostral, sem tendência e com variância mínima. Estas duas características fazem da krigagem um interpolador *Best Linear Unbiased Estimator* (BLUE) (BURGESS; WEBSTER, 1980) ou *Best Linear Unbiased Predictor* (BLUP), ou seja, seus estimadores não são tendenciosos, pois, em média, a diferença entre valores preditos e observados, para a mesma localização, deve ser nula; e possui variância mínima, pois este estimador possui a menor variância dentre todos os estimadores não tendenciosos.

O método de interpolação por krigagem possibilita a construção de mapas temáticos com alta precisão, uma vez que após a interpolação, a densidade espacial de dados será muito maior do que a obtida pela amostragem.

Conhecendo-se a estrutura de dependência espacial das variáveis regionalizadas em estudo, pode-se usar a interpolação por krigagem para prever valores em locais não amostrados e construir mapas temáticos.

A diferença entre krigagem e outros métodos de interpolação (inverso quadrado da distância, média móvel, linear simples, etc.) está na maneira como os pesos são atribuídos às diferentes amostras. No caso da interpolação linear simples, os pesos são todos iguais a $1/n$ (em que n é o número de amostras). Na interpolação baseada no inverso do quadrado das distâncias, os pesos são definidos como o inverso do quadrado da distância que separa o valor interpolado dos valores observados. Na krigagem, o procedimento é semelhante ao de interpolação por médias móveis ponderadas, exceto que neste caso os pesos são determinados a partir de uma análise espacial, baseado na estimação de parâmetros obtidos pela modelagem da dependência espacial.

Portanto, a krigagem é uma técnica especificamente desenvolvida para a geoestatística, pois considera os parâmetros obtidos pela modelagem da variabilidade espacial.

2.4.1 Krigagem Ordinária

Segundo JOURNEL (1989), a idéia básica de regressão linear é a estimativa de um valor desconhecido $\hat{Z}(s_0)$ por meio de uma combinação linear de n valores conhecidos: $Z(s_i)$, $i=1, 2, \dots, n$. Na krigagem, esses n valores conhecidos são de uma mesma característica ou propriedade em estudo. O estimador pode ser escrito como:

$$\hat{Z}(s_0) = \sum_{i=1}^n \lambda_i Z(s_i) \quad (42)$$

em que:

$\hat{Z}(s_0)$: valor predito no local s_0 , não amostrado;

n : número de valores $Z(s_i)$ medidos das variáveis em pontos amostrados;

λ_i : pesos associados a cada valor $Z(s_i)$ medido.

O estimador definido na equação (42) é tido como melhor estimador por produzir estimativas não viciadas e com variâncias mínimas. Tais qualidades são garantidas uma vez que os pesos λ_i são determinados impondo-se que a esperança do erro seja zero e a variância do erro seja mínima.

Sob a hipótese do processo $Z(\cdot)$ ser intrinsecamente estacionário, tem-se que o erro médio é:

$$E[Z(s_0) - \hat{Z}(s_0)] = \mu - \mu \sum_{i=1}^n \lambda_i = \mu \left(1 - \sum_{i=1}^n \lambda_i \right).$$

Esse erro será nulo se:

$$\sum_{i=1}^n \lambda_i = 1 \quad (43)$$

o que é necessário para garantir que o estimador será não tendencioso. Além disso, na krigagem, a variância da estimativa é mínima, ou seja,

$$\sigma^2 = \text{Var}[Z(s_0) - \hat{Z}(s_0)] = E[(Z(s_0) - \hat{Z}(s_0))^2] - \{E[(Z(s_0) - \hat{Z}(s_0))]\}^2 \quad (44)$$

Essas duas condições, (43) e (44), garantem que o estimador de krigagem seja BLUE (WEBSTER; OLIVER, 1990).

Substituindo (42) em (44), desenvolvendo-se a expressão resultante e minimizando-a, sujeita à restrição expressa em (43), chega-se a que os valores de λ_i , $i = 1, 2, \dots, n$ são soluções do sistema de equações

$$\sum_{j=1}^n \lambda_j \gamma(s_i, s_j) - \eta = \gamma(s_0, s_i), \quad i = 1, 2, \dots, n \quad (45)$$

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

em que a incógnita η é um multiplicador de Lagrange, introduzido ao minimizar a variância do erro.

O sistema expresso em (45), constituído de $n+1$ equações e $n+1$ incógnitas é conhecido como krigagem ordinária. Na forma matricial, o sistema pode ser escrito como $\gamma\lambda = A$

em que:

$$\lambda = [\lambda_1 \ \lambda_2 \ \lambda_3 \dots \ \lambda_n \ \eta]^T;$$

$$A = [\gamma(s_1, s_0) \ \gamma(s_2, s_0) \dots \ \gamma(s_n, s_0) \ 1]^T;$$

$$\gamma = \begin{bmatrix} \gamma(s_1, s_1) & \gamma(s_1, s_2) & \dots & \gamma(s_1, s_n) & 1 \\ \gamma(s_2, s_1) & \gamma(s_2, s_2) & \dots & \gamma(s_2, s_n) & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \gamma(s_n, s_1) & \gamma(s_n, s_2) & \dots & \gamma(s_n, s_n) & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix}$$

Os pesos λ_i podem, assim, ser obtidos mediante o produto da matriz inversa de γ pela matriz A , $\lambda = \gamma^{-1}A$.

Uma vez obtidos os pesos λ_i e o valor η , pode-se também calcular a variância do erro que é dada por (CRESSIE, 1993):

$$\sigma^2 = \sum_{i=1}^n \lambda_i \gamma(s_i, s_0) + \eta \quad (46)$$

2.4.2 Krigagem com Tendência

O interpolador linear denominado de krigagem é uma combinação linear dos valores amostrados pelo peso atribuído a cada vizinho. O preditor da krigagem pressupõe ausência de erros de medidas e semivariogramas corretos. É a média móvel ponderada dos valores observados na vizinhança em que os vizinhos mais próximos têm peso maior e os mais afastados, pesos cada vez menores, nulos ou até negativos (CRESSIE, 1993).

Para uma função qualquer de ordem k das coordenadas, genericamente denominadas por s , tem-se a definição de deriva

$$\mu(s) = \sum_{l=0}^k a_l f_l(s), \quad (47)$$

em que:

$f_l(s)$: são funções conhecidas de s ;

a_l : são coeficientes desconhecidos e constantes na vizinhança de s .

Assim, o estimador de um valor em s_0 é obtido por:

$$\hat{Z}(s_0) - \sum_{l=0}^k a_l f_l(s_0) = \sum_{\alpha=1}^n \lambda_{\alpha} \left[Z(s_{\alpha}) - \sum_{l=0}^k a_l f_l(s_{\alpha}) \right]$$

isto é,

$$\hat{Z}(s_0) = \sum_{\alpha=1}^n \lambda_{\alpha} Z(s_{\alpha}) + \sum_{l=0}^k a_l \left[f_l(s_0) - \sum_{\alpha=1}^n \lambda_{\alpha} f_l(s_{\alpha}) \right] \quad (48)$$

Para se calcular os coeficientes e derivar a_l por meio do estimador (48), impõem-se as seguintes restrições:

$$\sum_{\alpha=1}^n \lambda_{\alpha} f_l(s_{\alpha}) = f_l(s_0), l = 0, \dots, k.$$

O preditor de krigagem universal, que é interpretado como o valor interpolado de uma variável regionalizada Z , num local s_0 , pode ser determinado pela Equação (49):

$$\hat{Z}(s_0) = \sum_{\alpha=1}^n \lambda_{\alpha} Z(s_{\alpha}) \quad (49)$$

em que:

$\hat{Z}(s_0)$: valor predito no local s_0 ;

n : número de valores $Z(s_{\alpha})$, medidos das variáveis em pontos amostrados;

λ_{α} : pesos associados a cada valor $Z(s_{\alpha})$ medido.

O objetivo da krigagem é prever de forma mais precisa valores da variável nos locais não amostrados, a partir dos amostrados (ISAAKS; SRIVASTAVA, 1989).

2.5 MÉTODOS DE VALIDAÇÃO DE MODELOS

Para a tomada de decisão sobre o melhor modelo geoestatístico são usados os índices de desempenho como método de mínimos quadrados, ordinários e ponderados, máxima verossimilhança, máxima verossimilhança restrita e Bayesiano. Esses critérios são muito utilizados para a escolha do melhor modelo, porém, não ponderam sobre o número de parâmetros usados para o modelo estatístico ajustado.

A seguir serão apresentados alguns critérios que podem ser utilizados como métodos de validação de modelos por meio da análise de resíduos.

2.5.1 Critério de Akaike

O critério de informação de Akaike (*Akaike's Information Criterion - AIC*) (AKAIKE, 1973) busca demonstrar que se dois modelos representam dados igualmente satisfatórios, então com o modelo mais simples pode-se esperar um melhor desempenho para a predição de novos dados. Portanto, este critério busca a simplicidade, impondo uma penalidade para a complexidade.

O critério de informação de Akaike é definido como:

$$AIC = -2l(\theta) + 2r \quad (50)$$

em que:

$l(\theta)$: é a função de log-verossimilhança;

r : número de parâmetros do modelo ajustado.

Para o caso normal clássico, eliminadas as constantes arbitrárias, a definição de *AIC* é estimada por:

$$\hat{A} = n \ln \left(\frac{SQR}{n} \right) + 2r \quad (51)$$

em que:

n : número de pontos experimentais;

SQR : soma dos quadrados residuais $\left(\sum_{i=1}^n (Z(s_i) - \hat{Z}(s_i))^2 \right)$.

O critério de informação de Akaike é uma ligação da função de verossimilhança com uma medida de informação perdida quando um modelo representa, de modo aproximado, uma realidade. Esse critério penaliza a adição

de parâmetros, podendo ser aplicado quando se procura uma solução satisfatória entre o bom ajuste e o princípio da parcimônia.

A decisão para escolha entre os modelos utilizados no ajuste recai sobre aquele que apresentar o menor valor de \hat{A} .

Um dos problemas do critério AIC é que ele se baseia na log-verossimilhança e, portanto, o número n de observações não pode ser pequeno ($n/r < 40$). Nestes casos, é preferível utilizar o critério de Akaike de segunda ordem:

$$CAIC = -2l(\theta) + 2r + \frac{2r(r+1)}{n-r-1} \quad (52)$$

O valor de $CAIC$ tende para AIC quando n cresce.

Novamente, a decisão para escolha entre os modelos utilizados no ajuste recai sobre aquele que apresentar o menor valor de $CAIC$.

2.5.2 Critério de Filliben

O critério de Filliben, para resíduos ortonormais, pode ser de grande utilidade na determinação de qual o melhor modelo geoestatístico.

Considerando um vetor de dados observados de tamanho n , pode-se aleatorizar a ordem dos elementos do vetor e assim obter uma nova amostra.

Para a k -ésima posição ($k = p+1, p+2, \dots, n$) de um dado, pode-se estimar por krigagem ordinária o valor de $\hat{Z}(s_k)$, usando somente os $k - 1$ valores anteriores de dados, normalizados pelo desvio padrão do erro da krigagem, sendo o valor de p obtido da definição da função determinística $\mu(s_i)$, dada pela Equação (2).

Os erros normalizados (LEE, 1994) para $n - p$ variáveis são definidos como:

$$\varepsilon_k = \frac{Z(s_k) - \hat{Z}(s_k)}{\sigma_k} \quad (53)$$

para $k = p + 1, \dots, n$,

em que:

$Z(s_k)$: valores amostrados;

$\hat{Z}(s_k)$: valores estimados por krigagem ordinária;

σ_k : desvio padrão da estimativa da krigagem.

Os $n - p$ resíduos obtidos pela Equação (53) são chamados resíduos ortonormais, isto é, são não correlacionados, linearmente independentes e têm variância unitária.

Pelo critério de Filliben (FILLIBEN, 1975) calcula-se o coeficiente de correlação r^* (Equação 54), definido como o produto dos momentos entre as observações ordenadas $\varepsilon_{(i)}$ e a estatística das medianas ordenadas $M_{(i)}$, de distribuição normal $N(0, 1)$. Quanto mais próximo r^* for de 1 (um), mais normal é a distribuição dos dados. Se os valores tabelados de r para um determinado tamanho de amostra são maiores do que r^* , a hipótese de que os dados seguem uma distribuição normal é rejeitada, para um certo nível percentual de significância. O que se testa é se os resíduos ortonormais seguem uma distribuição normal, a um certo nível de significância.

Se ε_k representa o vetor de $n - p$ resíduos ortonormais e $\varepsilon_{(k)}$ uma amostra ordenada de ε_k , pode-se gerar estatísticas das medianas ordenadas m_i , a partir de uma população com distribuição uniforme $N(0, 1)$, por meio do seguinte algoritmo:

$$m_i = 1 - m_n, \text{ para } i = 1,$$

$$m_i = (i - 0,375)(n + 0,365), \text{ para } i = 2, 3, \dots, n - 1,$$

$$m_i = \sqrt[4]{0,5}, \text{ para } i = n.$$

A seguir obtém-se a estatística mediana ordenada M_i de uma população com distribuição normal usando:

$$m_i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{M_i} \exp\left(-\frac{t^2}{2}\right) dt$$

Assim, pode-se calcular o coeficiente de correlação r^* .

$$r^* = \frac{\sum_{i=1}^n (\varepsilon_{(i)} - \bar{\varepsilon})(M_{(i)} - \bar{M})}{\sqrt{\left[\sum_{i=1}^n (\varepsilon_{(i)} - \bar{\varepsilon})^2 \right] \left[\sum_{i=1}^n (M_{(i)} - \bar{M})^2 \right]}} \quad (54)$$

em que:

$\varepsilon_{(i)}$: resíduos normalizados ordenados;

$\bar{\varepsilon}$: média dos resíduos normalizados;

$M_{(i)}$: mediana da estatística de uma distribuição Normal $N(0, 1)$;

\bar{M} : média da estatística de uma distribuição Normal $N(0, 1)$.

2.5.3 Validação Cruzada

Segundo ISAACS e SRIVASTAVA (1989), a validação cruzada é uma técnica de avaliação de erros de estimativa que permite comparar valores estimados e amostrados tendo como informação a amostra de dados de que se dispõe. Tais comparações são úteis em ajudar a escolher entre os diferentes modelos de estimação.

Na aplicação da validação cruzada, o método de estimação é testado nos locais das amostras existentes. O valor da amostra, em uma certa localização, é temporariamente descartado do conjunto de dados da amostra. Um novo valor na mesma localização é, então, estimado por krigagem utilizando-se as amostras restantes. Uma vez que a estimativa é calculada, pode-se compará-la ao valor da amostra que foi inicialmente removida do conjunto de dados amostrais. Este procedimento é repetido para todas as amostras disponíveis e é designado com o método de “deixar-um-fora” (ISAACS; SRIVASTAVA, 1989).

Conhecendo-se os valores amostrados e os valores estimados, pode-se conhecer também a variância total da estimativa, sendo assim possível avaliar a qualidade ou precisão do processo.

Segundo GONÇALVES (1997), espera-se que os erros de estimação (Equação 55), tenham média nula, variância unitária e distribuição normal de probabilidade.

$$\varepsilon(s_i) = Z(s_i) - \hat{Z}(s_i) \quad (55)$$

em que $\hat{Z}(s_i)$ é o valor estimado sem a i -ésima observação, isto é, sem $Z(s_i)$.

As características do erro de estimação indicam não somente a eficácia do ajuste dos diferentes modelos teóricos ao semivariograma experimental e a modelagem do processo em questão, mas também a avaliação da estacionaridade e a influência da presença de dados atípicos.

A comparação entre os modelos, em cada método de ajuste, pode ser efetuada por meio do erro médio EM , erro médio reduzido (\overline{ER}), desvio padrão dos erros (DPe), desvio padrão dos erros reduzidos (S_{ER}) e do erro absoluto (EA).

O erro médio é definido como:

$$EM = \frac{1}{n} \sum_{i=1}^n (Z(s_i) - \hat{Z}(s_{(i)})) \quad (56)$$

em que:

n : número de dados;

$Z(s_i)$: valor observado no ponto s_i ;

$\hat{Z}(s_{(i)})$: valor predito por krigagem ordinária no ponto s_i , sem considerar a observação $Z(s_i)$.

Segundo McBRATNEY e WEBSTER (1986) e CRESSIE (1993), o erro médio reduzido é definido como

$$\overline{ER} = \frac{1}{n} \sum_{i=1}^n \frac{Z(s_i) - \hat{Z}(s_{(i)})}{\sigma(s_{(i)})} \quad (57)$$

em que:

$\sigma(s_{(i)})$: desvio padrão da krigagem no ponto s_i , sem considerar a observação $Z(s_i)$.

O desvio padrão dos erros reduzidos foi obtido a partir da seguinte expressão:

$$S_{ER} = \sqrt{\frac{1}{n} \sum_{i=1}^n \frac{|Z(s_i) - \hat{Z}(s_{(i)})|}{\sigma(s_{(i)})}} \quad (58)$$

Segundo McBRATNEY e WEBSTER (1986), CRESSIE (1993), VIEIRA (2000) e MELLO *et al.* (2005), aplicando-se a condição de não tendenciosidade, o valor populacional para o erro médio reduzido deve ser zero e do desvio padrão do erro reduzido deve ser igual a um. Portanto, o valor de EM e \overline{ER} mais próximo de zero, o valor DP menor e o valor de S_{ER} mais próximo de um, são os critérios para escolha do melhor modelo ajustado. O erro absoluto é uma medida da magnitude dos erros na unidade da variável. Conhecendo-se o conjunto de valores medidos e predito por krigagem ordinária $Z(s_i)$ e $\hat{Z}(s_{(i)})$, respectivamente, é possível definir o erro absoluto na unidade da variável estudada, como

$$EA = \sum_{i=1}^n |\hat{Z}(s_i) - Z(s_{(i)})| \quad (59)$$

2.5.4 Valor Máximo da Função Verossimilhança e Verossimilhança Restrita

Neste critério, o melhor modelo é escolhido observando o valor que maximiza a função log-verossimilhança. O modelo que apresentar o maior valor é dito o melhor. O mesmo critério vale para a função log-verossimilhança restrita.

2.6 TÉCNICAS DE DIAGNÓSTICO

Durante o processo de ajustes de modelos é importante verificar, entre outras características, a existência de observações influentes, que possam causar alguma interferência ou provoquem grandes desvios nos resultados do ajuste. Esta etapa de verificação é conhecida como análise de diagnóstico, a qual teve início com a análise de resíduos para detectar a presença de pontos extremos e avaliar a adequação da distribuição proposta para a variável resposta (PAULA, 2004).

Neste sentido, diversos trabalhos vêm sendo desenvolvidos, utilizando procedimentos que detectam a presença de observações que exercem grande influência em ajustes de modelos, por meio de técnicas de diagnósticos.

GALEA-ROJAS (1990) estudou técnicas de diagnóstico aplicadas em regressão linear, utilizando como estimador de parâmetros o método de mínimos quadrados; ORTEGA, BOLFARINE e PAULA (2002) utilizaram análises de diagnóstico em modelos de regressão *log-gamma* generalizado; OSORIO, PAULA e GALEA (2007) estudaram a curvatura de influência local sobre vários esquemas de perturbação para modelos lineares elípticos, com estrutura longitudinal; CHRISTENSEN, JOHNSON e PEARSON (1993) apresentaram estudos aplicando técnicas de diagnóstico para detectar observações que exercem influência na estimação da matriz de covariância, com modelos espaciais lineares, utilizados na krigagem universal.

Portanto, há grande interesse no estudo da influência que pontos discrepantes podem exercer sobre os resultados de quaisquer análises de dados, inclusive na modelagem da variabilidade espacial de um fenômeno, utilizada em análises geoestatísticas.

2.6.1 Influência Local

Em estudos de regressão, a prática da retirada de pontos é uma das técnicas mais conhecidas para avaliar o impacto de uma possível observação influente nas estimativas de parâmetros. Entretanto, um problema que pode ocorrer com a retirada individualizada de observações é o efeito de deixar de detectar pontos conjuntamente discrepantes. Existe também o procedimento de retirada múltipla de pontos, que é menos popular, e em virtude do custo computacional envolvido, existem vários métodos robustos para a detecção de pontos discrepantes e com baixo custo computacional (FUNG, 1993; PEÑA; YOHAI, 1999).

COOK (1986) apresentou uma proposta, que até hoje é considerada inovadora na área de diagnóstico, na qual propõe avaliar a influência conjunta das observações sob pequenas perturbações no modelo, em vez da avaliação pela retirada individual ou conjunta de pontos. A esta metodologia denomina-se “influência local”, e teve grande aceitação entre os pesquisadores da área.

Também na área de estatística espacial é de fundamental importância que se faça um estudo acerca da robustez dos resultados obtidos com o ajuste de um modelo, principalmente no que se refere à formulação e estimativas de seus parâmetros. Neste sentido, uma análise de diagnóstico consiste de métodos para avaliar a sensibilidade das inferências a pequenas perturbações nos dados ou no modelo postulado.

A idéia básica da metodologia de Cook consiste em estudar o comportamento de alguma medida particular de influência (como exemplo: nas estimativas dos parâmetros), segundo pequenas perturbações nos dados ou no modelo proposto. Ou seja, verificar a existência de pontos que sob pequenas modificações no modelo causam distorções nos resultados.

COOK (1986) propõe investigar o comportamento da função de afastamento da verossimilhança em uma vizinhança ω , que é o ponto no qual as duas funções de log-verossimilhança são iguais, ou seja, $l(\theta) = l(\theta/\omega)$, em que

$l(\theta)$ é definido como na Equação (40); o logaritmo da função verossimilhança perturbada $l(\theta/\omega_0)$, é definida pela Equação (60):

$$l(\theta/\omega) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma| - \frac{1}{2} (Z_\omega - W\beta)^T \Sigma^{-1} (Z_\omega - W\beta) \quad (60)$$

em que $Z_\omega = Z + \omega$, com $\omega = (\omega_1, \dots, \omega_n)^T$ vetor de perturbação das respostas, pertencente a um espaço de perturbações Ω ($\omega \in \Omega$).

Assume-se que existe um $\omega_0 \in \Omega$ tal que $l(\theta) = l(\theta/\omega_0)$, para todo θ e que $l(\theta/\omega)$ é duas vezes diferenciável em $(\theta^T, \omega^T)^T$.

A seguir, serão apresentadas algumas técnicas comuns em análises de diagnósticos, que são: influência local no afastamento da verossimilhança, influência local na matriz de covariância e no preditor linear e influência local nos valores preditos.

2.6.1.1 Influência local no afastamento da verossimilhança (*Likelihood displacement*)

A influência da perturbação ω no estimador de máxima verossimilhança do vetor de parâmetros θ pode ser avaliada pelo afastamento da verossimilhança, definido por:

$$LD(\omega) = 2(l(\hat{\theta}) - l(\hat{\theta}_\omega)) \quad (61)$$

em que:

$\hat{\theta}$: estimador de máxima verossimilhança de θ do modelo postulado;

$\hat{\theta}_\omega$: estimador de máxima verossimilhança de θ do modelo perturbado.

COOK (1986) propôs que se estudasse o comportamento local de $LD(\omega)$ em torno de ω_0 , utilizando a curva normal C_l de $LD(\omega)$ em ω_0 na direção de algum vetor unitário l , definido como:

$$C_l = 2|l^T \Delta^T L^{-1} \Delta l|, \quad (62)$$

com $\|Z\| = 1$,

em que:

L : é a matriz de informação observada, avaliada em $\theta = \hat{\theta}$;

Δ : é uma matriz de ordem $(p + q) \times n$ dada por $\Delta = (\Delta_\beta^T, \Delta_\varphi^T)^T$, avaliada em $\theta = \hat{\theta}$

e em $\omega = \omega_0$.

Para definir as matrizes Δ e L , considere-se a forma paramétrica da matriz de covariância Σ , definida na Equação (5).

Quando se utiliza a máxima verossimilhança (MV), tem-se:

$$\Delta_\beta = W^T \Sigma^{-1} \quad \text{e}$$

$$\Delta_\varphi = \frac{\partial^2 l(\theta/\omega)}{\partial \varphi \partial \omega^T}, \quad \text{com} \quad \frac{\partial^2 l(\theta/\omega)}{\partial \varphi \partial \omega^T} = (Z_\omega - W\beta)^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_i} \Sigma^{-1}, \quad i = 1, \dots, q.$$

A matriz de informação L é definida como:

$$L = \begin{pmatrix} L_{\beta\beta} & L_{\beta\varphi} \\ L_{\varphi\beta} & L_{\varphi\varphi} \end{pmatrix}$$

em que:

$$L_{\beta\beta} = - (W^T \Sigma^{-1} W);$$

$$L_{\beta\varphi} = \frac{\partial^2 l(\theta)}{\partial \beta \partial \varphi^T}, \quad \text{com} \quad \frac{\partial^2 l(\theta)}{\partial \beta \partial \varphi_j} = -W^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \varepsilon, \quad j = 1, \dots, q;$$

$$L_{\varphi\beta} = L_{\beta\varphi}^T;$$

$$L_{\varphi\varphi} = \frac{\partial^2 l(\theta)}{\partial \varphi \partial \varphi^T}, \quad \text{com elementos} \quad \frac{\partial^2 l(\theta)}{\partial \varphi_i \partial \varphi_j} = \frac{1}{2} \text{tr} \left\{ \Sigma^{-1} \left(\frac{\partial \Sigma}{\partial \varphi_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} - \frac{\partial^2 \Sigma}{\partial \varphi_i \partial \varphi_j} \right) \right\} +$$

$$+ \frac{1}{2} \varepsilon^T \Sigma^{-1} \left\{ \frac{\partial^2 \Sigma}{\partial \varphi_i \partial \varphi_j} - \frac{\partial \Sigma}{\partial \varphi_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} - \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_i} \right\} \Sigma^{-1} \varepsilon, \quad \text{sendo } i, j = 1, \dots, q.$$

Quando se utiliza a máxima verossimilhança restrita (MVR), $l_R(\varphi)$, a matriz Δ é substituída por:

$$\Delta_R = \frac{\partial^2 l_R(\theta/\omega)}{\partial \varphi \partial \omega^T}, \quad \text{com} \quad \frac{\partial^2 l_R(\theta/\omega)}{\partial \varphi_j \partial \omega^T} = -Z_\omega^T P \frac{\partial \Sigma}{\partial \varphi_j} P, \quad j = 1, \dots, q$$

e a matriz L é substituída por:

$$L_R = \frac{\partial^2 l_R(\theta)}{\partial \varphi \partial \varphi^T}, \text{ com elementos } \frac{\partial^2 l_R(\theta)}{\partial \varphi_i \partial \varphi_j} = \frac{1}{2} \text{tr} \left\{ P \left(\frac{\partial \Sigma}{\partial \varphi_i} P \frac{\partial \Sigma}{\partial \varphi_j} - \frac{\partial^2 \Sigma}{\partial \varphi_i \partial \varphi_j} \right) \right\} + \\ + \frac{1}{2} Z^T P \left\{ \frac{\partial^2 \Sigma}{\partial \varphi_i \partial \varphi_j} - \frac{\partial \Sigma}{\partial \varphi_i} P \frac{\partial \Sigma}{\partial \varphi_j} - \frac{\partial \Sigma}{\partial \varphi_j} P \frac{\partial \Sigma}{\partial \varphi_i} \right\} P Z, \text{ sendo } i, j = 1, \dots, q.$$

em que $P = \Sigma^l(I - A)$, com $A = W(W^T \Sigma^l W)^{-1} W^T \Sigma^l$.

Considerando-se a matriz

$$B = \Delta^T L^{-1} \Delta, \quad (63)$$

em que Δ e L são matrizes definidas em (62).

Seja L_{max} o autovetor, normalizado, associado ao maior autovalor, em módulo, da matriz B .

O gráfico dos elementos $|L_{max}|$ versus i (ordem dos dados) pode revelar qual o tipo de perturbação que tem a maior influência em $LD(\omega)$, na vizinhança de ω_0 (COOK, 1986).

Ainda, dada a matriz B , como apresentada em (63), e considerando-se $C_i = 2 * |b_{ii}|$, em que b_{ii} são os elementos da diagonal principal da matriz B . Pode-se também utilizar o gráfico de C_i versus i (ordem dos dados) como técnica de diagnóstico para avaliar a existência de observações influentes.

Sugere-se considerar a i -ésima observação como influente se seu valor for maior que duas vezes a média das medidas $|L_{max}|$ ou C_i .

2.6.1.2 Influência local na matriz de covariância e no preditor linear

A função de covariância $C(h)$ também depende do vetor de parâmetros θ e para a distância h fixa, denota-se $C_h(\theta)$. Este é um importante estudo da sensibilidade das estimativas de máxima verossimilhança, dados por $C_h(\hat{\theta})$. Para medir esta sensibilidade, segundo CADIGAN e FARREL (2002), pode-se utilizar a direção de máxima inclinação.

A influência local de primeira ordem da perturbação (neste caso $Z_\omega = Z + \omega$) pode ser medida usando a inclinação na direção l , denotada como $S(l)$, do gráfico de influência de $C_h(\hat{\theta}_\omega)$ versus ω . Neste caso tem-se,

$$S(l) = l^T \dot{C}_h \quad (64)$$

em que \dot{C}_h é um vetor $nx1$ dado por:

$$\dot{C}_h = \left\{ -\Delta^T L^{-1} \frac{\partial C_h(\theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}} \right\} \quad (65)$$

com $\frac{\partial C_h(\theta)}{\partial \theta^T} = \left(0^T, \frac{\partial C_h(\theta)}{\partial \varphi^T} \right)$, e $\frac{\partial C_h(\theta)}{\partial \varphi^T} = \left(\frac{\partial C_h(\theta)}{\partial \varphi_1}, \dots, \frac{\partial C_h(\theta)}{\partial \varphi_q} \right)$.

A direção de máxima inclinação local é dada pela Equação (66).

$$l_{C_h} = \frac{\dot{C}_h}{\|\dot{C}_h\|} \quad (66)$$

Outra direção de interesse é $l = e_{in}$, que corresponde ao i -ésimo vetor unitário de \mathfrak{R}^n . Neste caso, a inclinação é dada por

$$S(e_{in}) = S_i = \dot{C}_{ih}$$

em que \dot{C}_{ih} é o i -ésimo elemento de \dot{C}_h , com $i = 1, 2, \dots, n$.

Pode-se utilizar os gráficos de l_{C_h} e S_i para analisar a sensibilidade da função de covariância para pequenas perturbações no modelo estatístico.

Analogamente, para o preditor linear $p(s_0, \theta)$, tem-se que:

$$S(l) = l^T \dot{p}(s_0, \theta) \quad (67)$$

em que $\dot{p}(s_0, \theta)$ é um vetor $nx1$ dado por:

$$\dot{p}(s_0, \theta) = \left\{ -\Delta^T L^{-1} \frac{\partial p(s_0, \theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}} \right\} \quad (68)$$

Na Equação (68) tem-se que $\frac{\partial p(s_0, \theta)}{\partial \theta} = \left(\frac{\partial p(s_0, \theta)}{\partial \beta^T}, \frac{\partial p(s_0, \theta)}{\partial \varphi^T} \right)^T$

em que:

$$\frac{\partial p(s_0, \theta)}{\partial \beta} = w_0 - \Sigma^{-1} C_0;$$

$$\frac{\partial p(s_0, \theta)}{\partial \varphi} = \left[\left(\frac{\partial p(s_0, \theta)}{\partial \varphi_j} \right) \right];$$

$$\text{sendo } \frac{\partial p(s_0, \theta)}{\partial \varphi_j} = \left\{ \frac{\partial C_0^T}{\partial \varphi_j} - C_0^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \right\} \Sigma^{-1} (Z - W\beta) \text{ e}$$

$$\frac{\partial C_0^T}{\partial \varphi_j} = \left(\frac{\partial C(h_{10})}{\partial \varphi_j}, \dots, \frac{\partial C(h_{n0})}{\partial \varphi_j} \right) \text{ para } j = 1, 2, \dots, q.$$

A direção de máxima inclinação local é obtida pela Equação (69)

$$l_{p(s_0, \theta)} = \frac{\dot{p}(s_0, \theta)}{\|\dot{p}(s_0, \theta)\|} \quad (69)$$

Pode-se utilizar $l_{p(s_0, \theta)}$ para analisar a sensibilidade da função de covariância para pequenas perturbações no modelo estatístico.

2.6.1.3 Influência local nos valores preditos (*Generalized leverage*)

Outro importante conceito de diagnósticos de influência em estudos de modelos de regressão é a alavanca (*leverage*), segundo HOAGLIN e WELSH (1978). Esta idéia tem sido generalizada para modelos mais complexos. Por exemplo: ROSS (1987), ST. LAURENT e COOK (1992) e WEI, HU e FUNG (1998) estenderam o método de alavanca para modelos não-lineares.

A proposta do método de alavanca é medir a influência que os valores da variável resposta observada exercem sobre a predição de seus valores em locais não amostrados.

Seja $\mu = W\beta$ o valor esperado da variável Z. Então, segundo WEI, HU e FUNG (1998), a matriz de alavanca generalizada, $\frac{\partial \hat{Z}}{\partial Z^T}$, assume a forma

$$LG(\theta) = D_\theta (-L)^{-1} L_{\theta Z} \quad (70)$$

em que:

$$D_\theta = \frac{\partial \mu}{\partial \theta^T} = (W, 0);$$

$$L_{\theta Z} = \frac{\partial^2 l(\theta)}{\partial \theta \partial Z^T} = (L_{\beta Z}^T, L_{\phi Z}^T)^T, \text{ com } L_{\beta Z} = W^T \Sigma^{-1} \text{ e } L_{\phi Z} = \frac{\partial^2 l(\theta)}{\partial \phi \partial Z^T},$$

em que $\frac{\partial^2 l(\theta)}{\partial \phi_j \partial Z^T} = \varepsilon^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \phi_j} \Sigma^{-1}$, para $j = 1, 2, \dots, q$.

Após alguns cálculos algébricos tem-se que

$$LG(\theta) = W(L_{\beta\beta} - L_{\beta\phi} L_{\phi\phi}^{-1} L_{\phi\beta})^{-1} (L_{\phi\beta} L_{\phi\phi}^{-1} L_{\phi Z} - L_{\beta Z}) = LG_1 + LG_2 \quad (71)$$

em que:

$$LG_1 = W(W^T \Sigma^{-1} W + L_{\beta\phi} L_{\phi\phi}^{-1} L_{\phi\beta})^{-1} W^T \Sigma^{-1};$$

$$LG_2 = W(W^T \Sigma^{-1} W + L_{\beta\phi} L_{\phi\phi}^{-1} L_{\phi\beta})^{-1} (-L_{\beta\phi} L_{\phi\phi}^{-1} L_{\phi Z}).$$

3 MATERIAL E MÉTODOS

3.1 ESTUDO COM DADOS SIMULADOS UTILIZANDO MONTE CARLO

A primeira etapa do desenvolvimento do projeto foi a realização de experimentos Monte Carlo, em que foram simulados conjuntos de dados com estruturas de dependência espacial conhecidas e com distribuição normal de probabilidade, ou seja, processos espaciais gaussianos.

Segundo CRESSIE (1993), a simulação de processos espaciais estacionários de segunda ordem pode ser feita pelo método de decomposição de Cholesky. Esta é uma forma de garantir que a geração de seqüências aleatórias respeite uma matriz de correlações.

A decomposição de Cholesky é uma operação matricial que, aplicada ao vetor de números aleatórios sorteados, produz um outro vetor de números aleatórios que têm a característica de obedecerem a uma dada matriz de correlação entre eles.

O procedimento é descrito a seguir.

Considerando $Z = (z(s_1), \dots, z(s_n))^T$ o vetor $n \times 1$ dos dados simulados, os quais representam a realização de um processo estocástico ou função de variáveis aleatórias $\{Z(s); s \in S \subset \mathfrak{R}^2\}$, em n diferentes localizações: s_1, \dots, s_n .

Considere-se agora o vetor de médias do processo:

$$\mu = (E(Z(s_1)), \dots, E(Z(s_n)))^T, \quad (72)$$

e a matriz de covariância, definida em (5), como $\Sigma = [C(Z(s_i), Z(s_u))] = \varphi_1 I_n + \varphi_2 R$.

Para processos estocásticos que satisfazem a hipótese de estacionaridade de segunda ordem e isotropia, tem-se:

$$E[Z(s_i)] = \mu, \text{ para } i = 1, 2, \dots, n \quad (73)$$

e

$$C(Z(s_i), Z(s_u)) = C(h_{iu}), \text{ onde } h_{iu} = \|s_i - s_u\| \quad (74)$$

Nesse caso, cada elemento do vetor μ é igual a um valor constante μ e cada (i, u) - éximo elemento da matriz Σ , quadrada de ordem n , é igual a $C(h_{iu})$.

Assim, escolhendo-se o valor de μ e a função covariância $C(h)$, o vetor Z , satisfazendo (73) e (74), pôde ser simulado pela relação:

$$Z = \mu + L\varepsilon \quad (75)$$

em que: L é uma matriz triangular inferior $n \times n$, obtida mediante a decomposição de Σ no produto LL^T , chamada de decomposição de Cholesky, e $\varepsilon = (\varepsilon(s_1), \dots, \varepsilon(s_n))^T$ é um vetor de variáveis aleatórias não correlacionadas. Se $\varepsilon(s_1), \dots, \varepsilon(s_n)$ são identicamente distribuídas, com distribuição gaussiana de probabilidade, com média igual a zero e variância igual a 1, pode ser demonstrado que Z tem distribuição gaussiana multivariada, com vetor de médias μ e matriz de covariância Σ (JOHNSON; WICHERN, 1982).

Para que os dados simulados tivessem uma estrutura de dependência espacial conhecida foram fornecidos modelos teóricos de semivariogramas e vetores de parâmetros de dependência espacial do tipo $\varphi = (\varphi_1, \varphi_2, \varphi_3)^T$.

Foram feitas simulações de dados com uma estrutura de dependência espacial conhecida e com distribuição normal de probabilidade (Monte Carlo), pelo método de decomposição de Cholesky. Os modelos espaciais utilizados nas simulações foram: exponencial, gaussiano e Matérn com parâmetros κ iguais a 0,7 e 3,0. Em todos os casos considerou-se a média do processo $\mu = 9,45$. Em cada modelo, foram utilizados quatro vetores de parâmetros

$\varphi^T = (\varphi_1, \varphi_2, \varphi_3)$: 1º caso: $\varphi^T = (0, 10, 10)$; 2º caso: $\varphi^T = (0, 10, 15)$; 3º caso: $\varphi^T = (0, 10, 20)$; 4º caso: $\varphi^T = (0, 10, 60)$.

Os dados foram simulados em grade regular, com distância de 10 u.m entre pontos, totalizando 100 pontos.

O próximo passo foi utilizar um esquema de perturbação no vetor Z de dados, como definido a seguir.

Seja Z o vetor $nx1$ de dados simulados. O vetor de dados perturbados será Z_ω , em que $Z_\omega = Z + \omega$, com $\omega = (\omega_1, \dots, \omega_n)^T$ vetor de perturbação das respostas, pertencente a um espaço de perturbações Ω ($\omega \in \Omega$).

Neste trabalho foi utilizado o esquema de perturbação proposto por ORTEGA, BOLFARINE e PAULA (2002), como apresentado na Equação (76)

$$z_{max}^* = z_{max} + \sqrt{\|Z\|} \quad (76)$$

em que:

z_{max}^* : novo valor máximo do vetor Z ;

z_{max} : valor máximo do vetor Z ;

$$\|Z\| = \sqrt{z_1^2 + z_2^2 + \dots + z_n^2}.$$

Sendo assim, o vetor ω de perturbações tem a forma $\omega = (0, 0, \dots, \sqrt{\|Z\|}, \dots, 0)^T$.

A partir de então, foi realizada a modelagem da estrutura de dependência espacial para os conjuntos de dados simulados, mas somando-se o vetor ω de perturbações. Foram aplicadas técnicas de diagnóstico para avaliar a sensibilidade dos modelos ao esquema de perturbação.

Os modelos teóricos utilizados nas análises de diagnóstico foram o exponencial, gaussiano e família Matérn, com parâmetros estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MV).

Para o caso particular com $q = 3$ (parâmetros da estrutura de variabilidade espacial), tem-se que as derivadas de primeira e segunda ordem da

matriz de covariância Σ são, segundo URIBE-OPAZO, BORSSOI e GALEA-ROJAS (2007)¹, dadas por:

$$\frac{\partial \Sigma}{\partial \varphi_1} = I \text{ e } \frac{\partial^2 \Sigma}{\partial \varphi_1^2} = \frac{\partial^2 \Sigma}{\partial \varphi_1 \partial \varphi_2} = \frac{\partial^2 \Sigma}{\partial \varphi_1 \partial \varphi_3} = 0, \quad (77)$$

$$\frac{\partial \Sigma}{\partial \varphi_2} = R, \quad \frac{\partial^2 \Sigma}{\partial \varphi_2^2} = 0 \text{ e } \frac{\partial^2 \Sigma}{\partial \varphi_2 \partial \varphi_3} = \frac{\partial R}{\partial \varphi_3} = \left[\left(\frac{\partial r_{ij}}{\partial \varphi_3} \right) \right]. \quad (78)$$

Em particular, para a função de covariância exponencial tem-se:

$$\frac{\partial r_{ij}}{\partial \varphi_3} = \frac{r_{ij} h_{ij}}{\varphi_3^2}. \quad (79)$$

Para a função de covariância gaussiana tem-se:

$$\frac{\partial r_{ij}}{\partial \varphi_3} = 2 \frac{r_{ij} h_{ij}^2}{\varphi_3^3}. \quad (80)$$

E para a função de covariância Matérn tem-se:

$$\frac{\partial r_{ij}}{\partial \varphi_3} = - \left(\frac{1}{\varphi_3} \right) \left[k r_{ij} + \frac{1}{2^{k-1} \Gamma(k)} \left(\frac{h_{ij}}{\varphi_3} \right)^{k+1} K'_k \left(\frac{h_{ij}}{\varphi_3} \right) \right] \quad (81)$$

em que: $K'_k \left(\frac{h_{ij}}{\varphi_3} \right) = K'_k(u) = \frac{\partial K_k(u)}{\partial u} = \frac{1}{2} [K_{k-1}(u) + K_{k+1}(u)]$.

$$\frac{\partial \Sigma}{\partial \varphi_3} = \varphi_2 \frac{\partial R}{\partial \varphi_3} \text{ e } \frac{\partial^2 \Sigma}{\partial \varphi_3^2} = \varphi_2 \frac{\partial^2 R}{\partial \varphi_3^2}. \quad (82)$$

Em particular, para a função de covariância exponencial tem-se:

$$\frac{\partial^2 r_{ij}}{\partial \varphi_3^2} = \left(\frac{r_{ij} h_{ij}}{\varphi_3^3} \right) \left[\left(\frac{h_{ij}}{\varphi_3} \right) - 2 \right] \quad (83)$$

¹ Resultados não publicados.

Para a função de covariância gaussiana tem-se:

$$\frac{\partial^2 r_{ij}}{\partial \varphi_3^2} = \left(\frac{2r_{ij}h_{ij}^2}{\varphi_3^4} \right) \left[2 \left(\frac{h_{ij}}{\varphi_3} \right)^2 - 3 \right]. \quad (84)$$

E para a função de covariância Matérn tem-se:

$$\frac{\partial^2 r_{ij}}{\partial \varphi_3^2} = \left(\frac{k(k+1)}{\varphi_3^2} \right) + \left(\frac{1}{\varphi_3^2 2^{k-1} \Gamma(k)} \right) \left(\frac{h_{ij}}{\varphi_3} \right)^{k+1} \left[2(k+1)K_k' \left(\frac{h_{ij}}{\varphi_3} \right) + \left(\frac{h_{ij}}{\varphi_3} \right) K_k'' \left(\frac{h_{ij}}{\varphi_3} \right) \right] \quad (85)$$

em que $K_k'' \left(\frac{h_{ij}}{\varphi_3} \right) = K_k''(u) = \frac{\partial^2 K_k(u)}{\partial u^2} = \frac{1}{4} [K_{k-2}(u) + 2K_k(u) + K_{k+2}(u)]$.

3.2 ESTUDOS COM DADOS DA BIBLIOGRAFIA

Foram utilizados dados já estudados por PIEGORSCH e BAILER (2005) e publicados no livro *Analyzing enviromental data*. Os dados são relativos à seleção de concentração (em ppb) de *trichloroethylene* (TCE).

Estes dados foram estudados com o objetivo de comparar os resultados obtidos neste trabalho com os já analisados pelos autores, avaliando assim, se a metodologia aqui empregada é consistente.

3.3 ESTUDOS COM DADOS EXPERIMENTAIS

Na terceira etapa, os estudos foram realizados aplicando-se técnicas de diagnóstico e construindo-se os mapas temáticos com base nos resultados da

validação de modelos teóricos ajustados aos dados experimentais, obtidos em campo.

3.3.1 Localização e Caracterização das Áreas Experimentais

A coleta de dados foi realizada em dois anos agrícolas, na mesma área experimental. Os experimentos foram desenvolvidos na região Oeste do Paraná, em uma área comercial de produção de grãos, no município de Cascavel, cuja localização geográfica é, aproximadamente, 24,95° Sul de latitude, 53,57° Oeste de longitude e altitude média de 650 m. Os dados são referentes aos anos agrícolas de 2004/2005 e 2006/2007, em uma área de 57 ha e outra de 71 ha, respectivamente, com experimentos que foram conduzidos por pesquisadores do grupo de pesquisa Geoestatística Aplicada (GGEA) da Universidade Estadual do Oeste do Paraná – UNIOESTE, *campus* de Cascavel.

O solo é classificado como Latossolo Vermelho Distroférico. O clima da região é classificado como temperado mesotérmico e superúmido, tipo climático Cfa (Köppen) e a temperatura anual média é de 21°C.

3.3.2 Amostragem

No ano agrícola de 2004/2005, foi definida uma malha de amostragem, com distância regular entre pontos de 75 m, totalizando 125 pontos amostrais, obtendo-se uma densidade amostral de 2,2 amostras por hectare (Figura 3). Para o ano agrícola de 2006/2007 foi realizada uma amostragem sistemática centrada com pares de pontos próximos (*lattice plus close pairs*), com distância máxima de 141 m entre pontos e, em alguns locais pré-determinados, a amostragem foi realizada com distâncias menores, ou seja, 75 e 50 m entre pontos (Figura 4).

Todas as amostras foram georreferenciadas com o sistema de posicionamento global (GPS).

Nas Figuras 3 e 4 são ilustradas as áreas experimentais e as grades de amostragem, para os dois anos agrícolas em estudo. Cada número nas figuras representa o ponto de amostragem. A localização de cada ponto foi feita com o auxílio de um aparelho receptor de sinal de GPS (*Global Positioning System*).

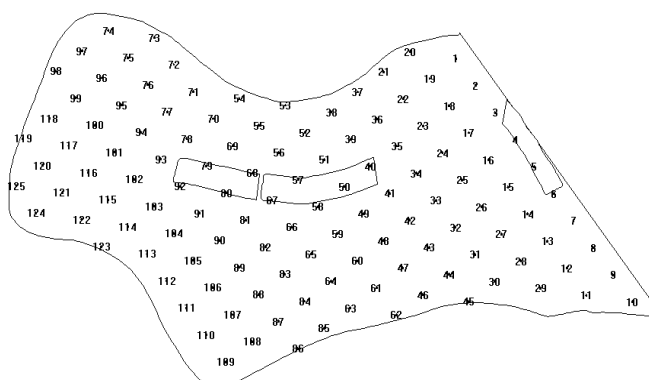


Figura 3 - Área experimental de 57 ha e grade de amostragem para o ano agrícola de 2004/2005.

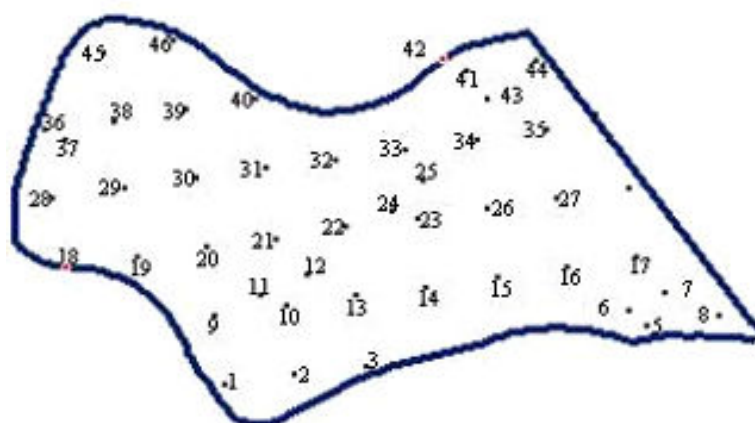


Figura 4 - Área experimental de 71 ha e grade de amostragem para o ano agrícola de 2006/2007.

3.3.3 Variáveis Estudadas

As variáveis avaliadas na área experimental no ano agrícola 2004/2005 (Figura 3) foram: produtividade da cultura da soja ($t\ ha^{-1}$), densidade do solo ($Mg\ m^{-3}$) e resistência do solo à penetração (MPa), ambas nas camadas do solo de 0,0 a 0,1 *m*; 0,1 a 0,2 *m* e 0,2 a 0,3 *m*. No ano agrícola 2006/2007 (Figura 4), as variáveis estudadas na área experimental foram: fósforo ($mg\ dm^{-3}$) e manganês ($mg\ dm^{-3}$).

3.3.3.1 Produtividade da cultura ($t\ ha^{-1}$)

A produtividade foi estimada considerando-se a quantidade de grãos de soja colhida de todas as plantas distribuídas em duas fileiras ao longo de um metro de comprimento, representando a parcela. Após a triagem, foi feita a pesagem de grãos para cada parcela e verificado o teor de água para posterior correção para 13%. O valor foi também convertido em $t\ ha^{-1}$.

3.3.3.2 Densidade do solo ($Mg\ m^{-3}$)

A variável densidade do solo foi coletada em três camadas: 0,0 a 0,1 *m*; 0,1 a 0,2 *m* e 0,2 a 0,3 *m* de profundidade, a partir da superfície, pelo método do anel volumétrico (EMBRAPA, 1997).

3.3.3.3 Resistência do solo à penetração (MPa)

As determinações da resistência do solo à penetração foram feitas com um penetrômetro. Este aparelho foi desenvolvido no Laboratório de Mecanização Agrícola da UNIOESTE (TIEPPO, 2004). O penetrômetro possui uma haste de 60 cm e fornece dados da resistência a cada 0,25 segundos, em Kgf, à medida que a haste é introduzida no solo.

Em cada parcela foram feitas as leituras em três camadas do solo: 0,0 a 0,1 m; 0,1 a 0,2 m e 0,2 a 0,3 m, com quatro repetições em cada uma. Os dados obtidos foram transformados em MPa e, posteriormente, determinou-se a resistência do solo à penetração, sendo utilizada a média das resistências em cada faixa para representar a parcela.

3.3.3.4 Fósforo (mg dm^{-3}) e manganês (mg dm^{-3})

A amostragem do solo para medir os níveis das variáveis químicas do solo fósforo e manganês foi realizada em cada ponto da Figura 4. Nas proximidades do ponto demarcado foram coletadas quatro subamostras de solo, de 0,0 a 0,2 m de profundidade. As subamostras foram misturadas e colocadas em sacos plásticos, com aproximadamente 500 g, compondo assim, a amostra representativa da parcela.

Após o trabalho de campo, as amostras foram encaminhadas ao laboratório, onde se procedeu a análise química, determinando os níveis de fósforo e manganês no solo.

3.3.4 Análises de Dados Experimentais

Inicialmente, foi realizada a análise estatística exploratória dos dados para avaliar o comportamento geral e identificar a presença de pontos discrepantes e suas possíveis causas.

O passo seguinte foi a análise espacial dos dados pelas técnicas geoestatísticas para identificar a estrutura de dependência espacial, por meio do ajuste dos modelos teóricos exponencial, gaussiano e Matérn, com parâmetros estimados por máxima verossimilhança e máxima verossimilhança restrita. Nesta etapa, foram aplicados os critérios de validação de modelos e os estudos das técnicas de diagnóstico para a posterior confecção dos mapas temáticos das variáveis em estudo.

Para os dados das variáveis produtividade da soja (Prod.), resistência do solo à penetração (RSP) e densidade do solo (DS), foi realizada uma análise diferenciada, em que as variáveis RSP e DS foram utilizadas como variáveis explicativas ou covariáveis da produtividade, pois estas foram coletadas no mesmo ano agrícola (2004/2005).

4 RESULTADOS E DISCUSSÃO

4.1 ESTUDO COM DADOS SIMULADOS

Nesta seção serão apresentados os resultados obtidos com os estudos de influência local aplicados a dados simulados. Foram simulados conjuntos de dados com distribuição gaussiana e com diferentes estruturas de dependência espacial, utilizando os modelos espaciais exponencial, gaussiano e Matérn, com parâmetros $kappa$ iguais a 0,7 e 3,0. Após a simulação, foi aplicado um esquema de perturbação no vetor de dados, com o objetivo de verificar se as técnicas de diagnóstico são capazes de identificar a observação perturbada.

Nas tabelas 1 a 3 são apresentados os parâmetros de dependência espacial φ_1 , φ_2 e φ_3 , estimados pelos métodos máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR), e a observação perturbada para variáveis simuladas com os modelos espaciais exponencial, gaussiano e Matérn, com parâmetros $kappa$ iguais a 0,7 e 3,0.

Nas figuras 5 a 20 são apresentados os gráficos de diagnóstico com o objetivo de identificar a observação à qual se aplicou o esquema de perturbação. Os gráficos " $C_i \times i$ " e " $|L_{max}| \times i$ ", utilizados no estudo da influência local no afastamento da verossimilhança: " $ICd \times i$ (equivalente à medida de diagnóstico l_{C_h})", empregado no estudo da influência local na matriz de covariância e " $GL \times i$ ", utilizado no estudo da influência local nos valores preditos.

Pelos resultados das análises de diagnóstico dos dados simulados pode-se observar que, para todas as variáveis, os gráficos " $C_i \times i$ " e " $|L_{max}| \times i$ " identificaram as observações às quais foi aplicado o esquema de perturbação,

tanto quando os parâmetros foram estimados por MV como para MVR (Figuras 5 a 20). Já os gráficos “ $ICd \times i$ ” e “ $GL \times i$ ”, quando as variáveis foram simuladas com os modelos exponencial e gaussiano, não se identificou a observação perturbada como sendo influente (Figuras 5 a 12). Entretanto, quando as variáveis foram simuladas com os modelos da família Matérn, em alguns casos, estes gráficos identificaram as observações perturbadas como sendo influentes (Figuras 13 a 20).

Tabela 1 - Parâmetros estimados para a função exponencial com os estimadores de máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR)

Variável simulada	Obs. perturbada	φ_1		φ_2		φ_3	
		MV	MVR	MV	MVR	MV	MVR
Exponencial 0-10-10	66	8,111	8,376	3,960	5,759	27,085	54,369
Exponencial 0-10-15	19	0,000	0,000	13,150	14,233	15,386	17,098
Exponencial 0-10-20	21	0,000	0,000	14,010	15,640	17,730	20,400
Exponencial 0-10-60	80	1,002	1,091	6,772	9,460	39,163	60,000

Analisando-se os resultados para a variável “Exponencial 0-10-10”, que foi simulada com modelo exponencial e parâmetros $\varphi_1 = 0$, $\varphi_2 = 10$ e $\varphi_3 = 10$ (Tabela 1), nota-se que os métodos MV e MVR superestimaram os parâmetros φ_1 e φ_3 e subestimaram o parâmetro φ_2 . Os parâmetros estimados por MV foram $\varphi_1 = 8,111$, $\varphi_2 = 3,960$ e $\varphi_3 = 27,085$ e os parâmetros estimados por MVR foram $\varphi_1 = 8,376$, $\varphi_2 = 5,759$ e $\varphi_3 = 54,369$. Nas estimativas de parâmetros para a variável simulada “Exponencial 0-10-60”, foi identificado um comportamento semelhante, exceto para o parâmetro φ_3 que teve valor subestimado, quando estimado por MV, e valor igual ao simulado para MVR.

Isso mostra que o valor perturbado das variáveis simuladas “Exponencial 0-10-10” e “Exponencial 0-10-60” foram mais influentes na estimação dos parâmetros do que os valores perturbados nas variáveis “Exponencial 0-10-15” e “Exponencial 0-10-20”, visto que nestas variáveis os parâmetros estimados foram próximos aos fornecidos para a simulação.

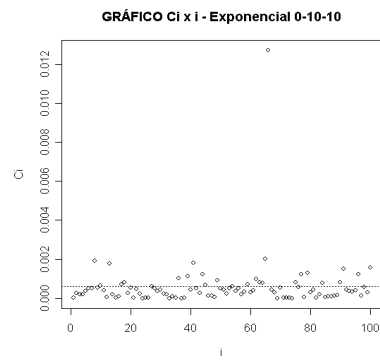
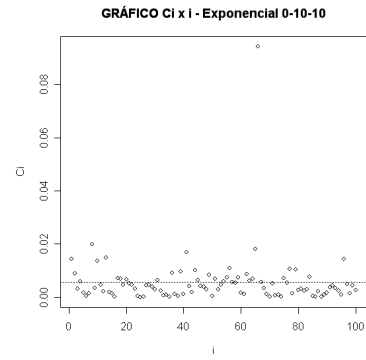
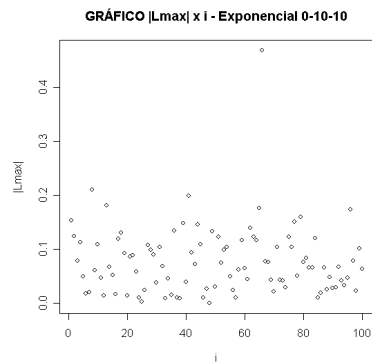
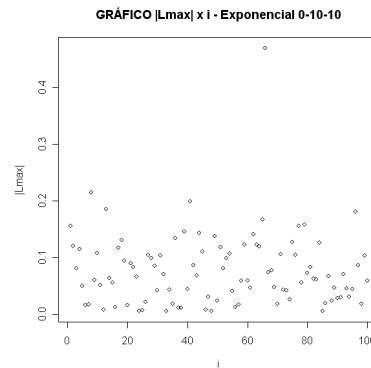
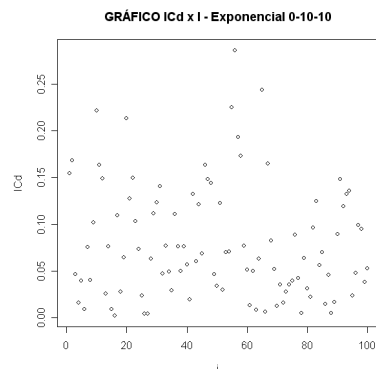
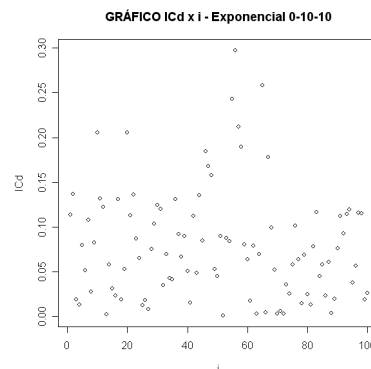
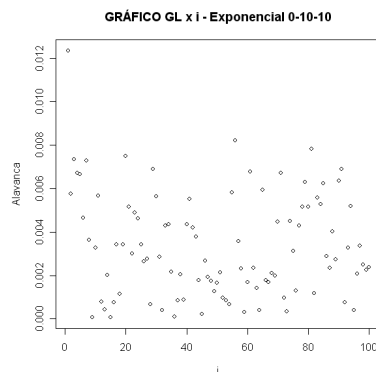
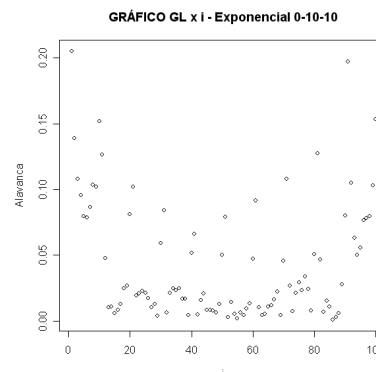
(a) C_i -MV(e) C_i -MVR(b) $|L_{max}|$ -MV(f) $|L_{max}|$ -MVR(c) ICd -MV(g) ICd -MVR(d) GL -MV(h) GL -MVR

Figura 5 - Gráficos de diagnóstico $C_i \times i$, $|L_{max}| \times i$, $ICd \times i$ e $GL \times i$, para a variável simulada Exponencial 0-10-10, com os estimadores MV e MVR.

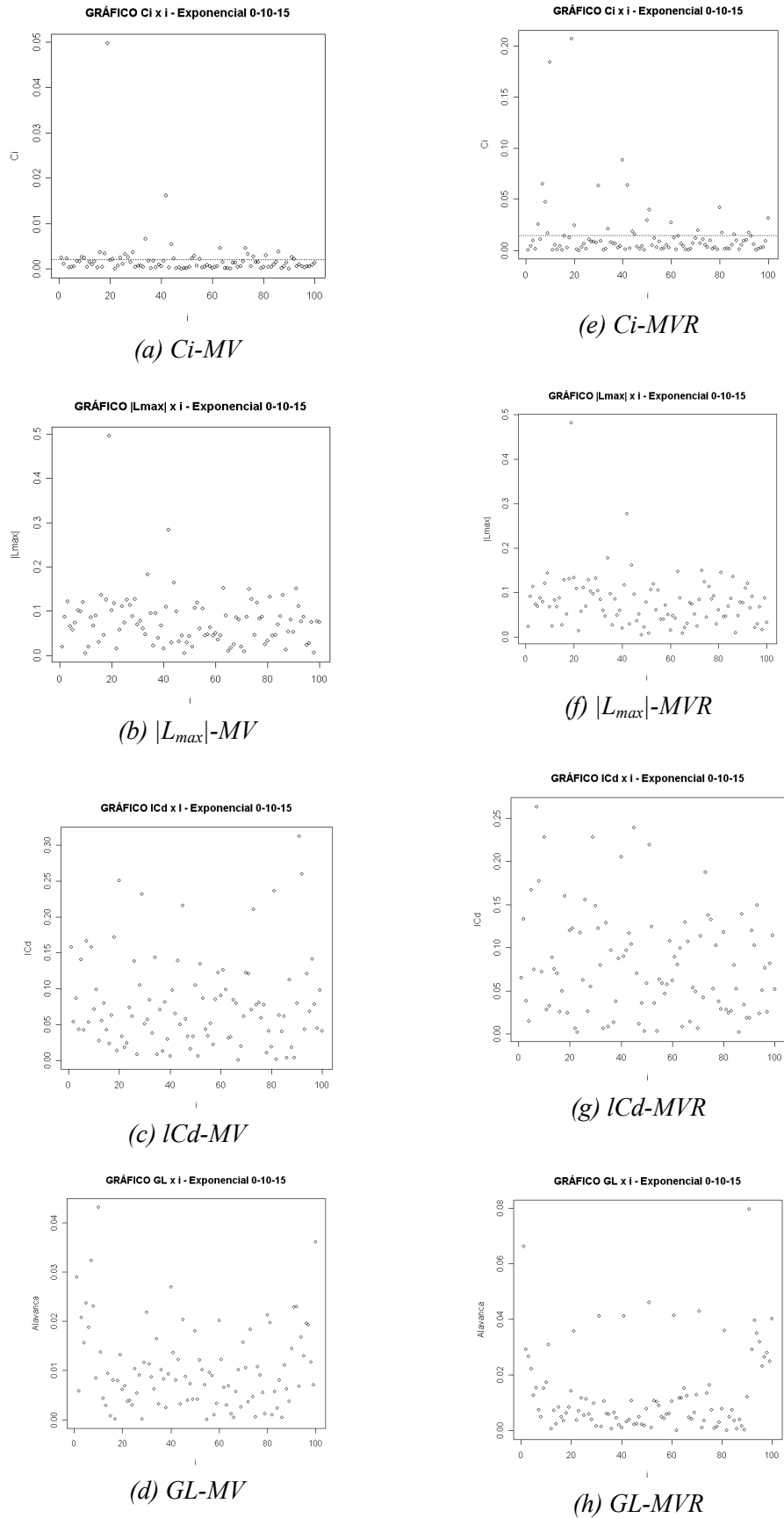


Figura 6 - Gráficos de diagnóstico C_i x i , $|L_{max}|$ x i , ICd x i e GL x i , para a variável simulada Exponencial 0-10-15, com os estimadores MV e MVR.

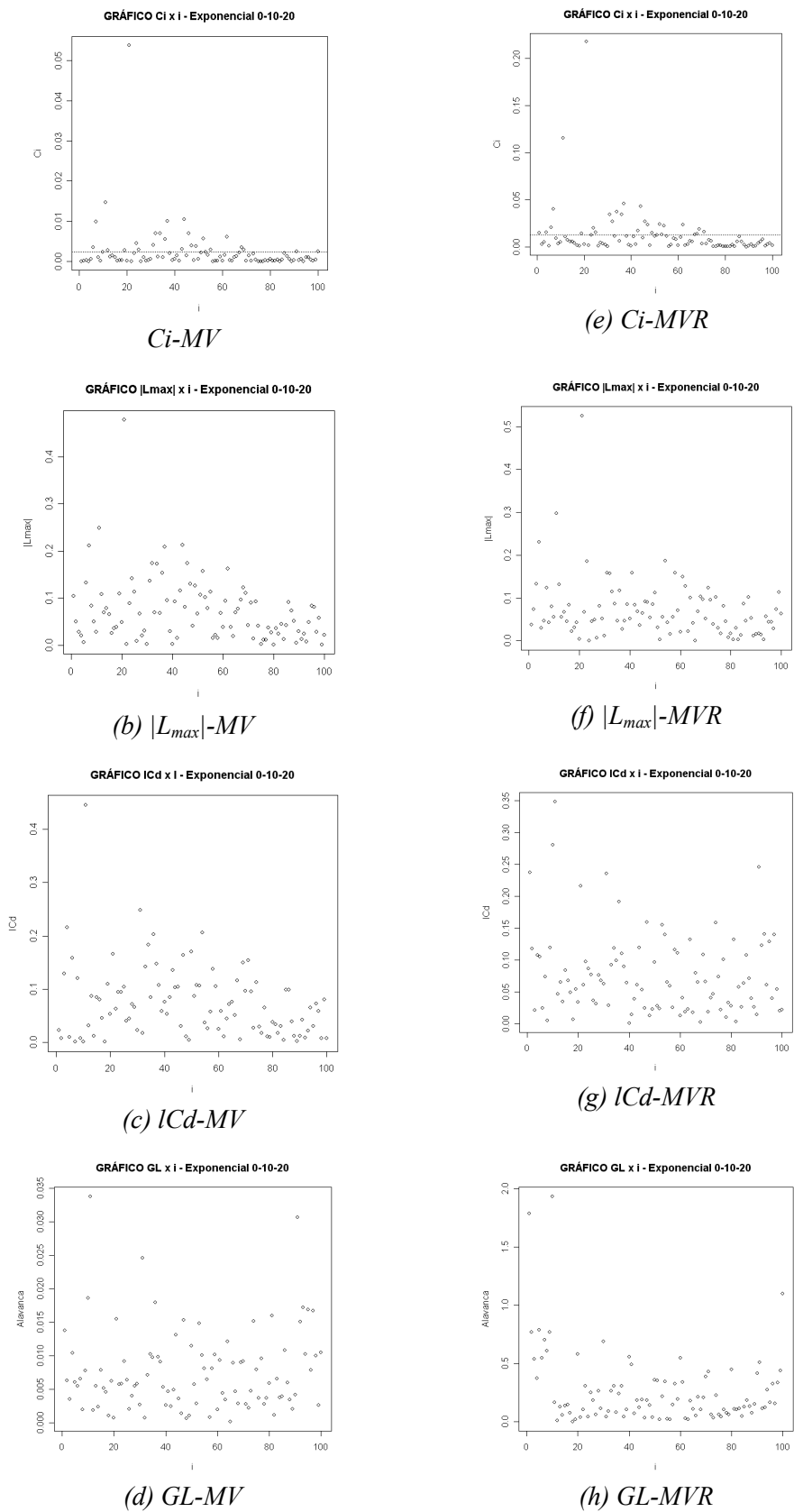


Figura 7 - Gráficos de diagnóstico C_i x i , $|L_{max}|$ x i , ICd x i e GL x i , para a variável simulada Exponencial 0-10-20, com os estimadores MV e MVR.

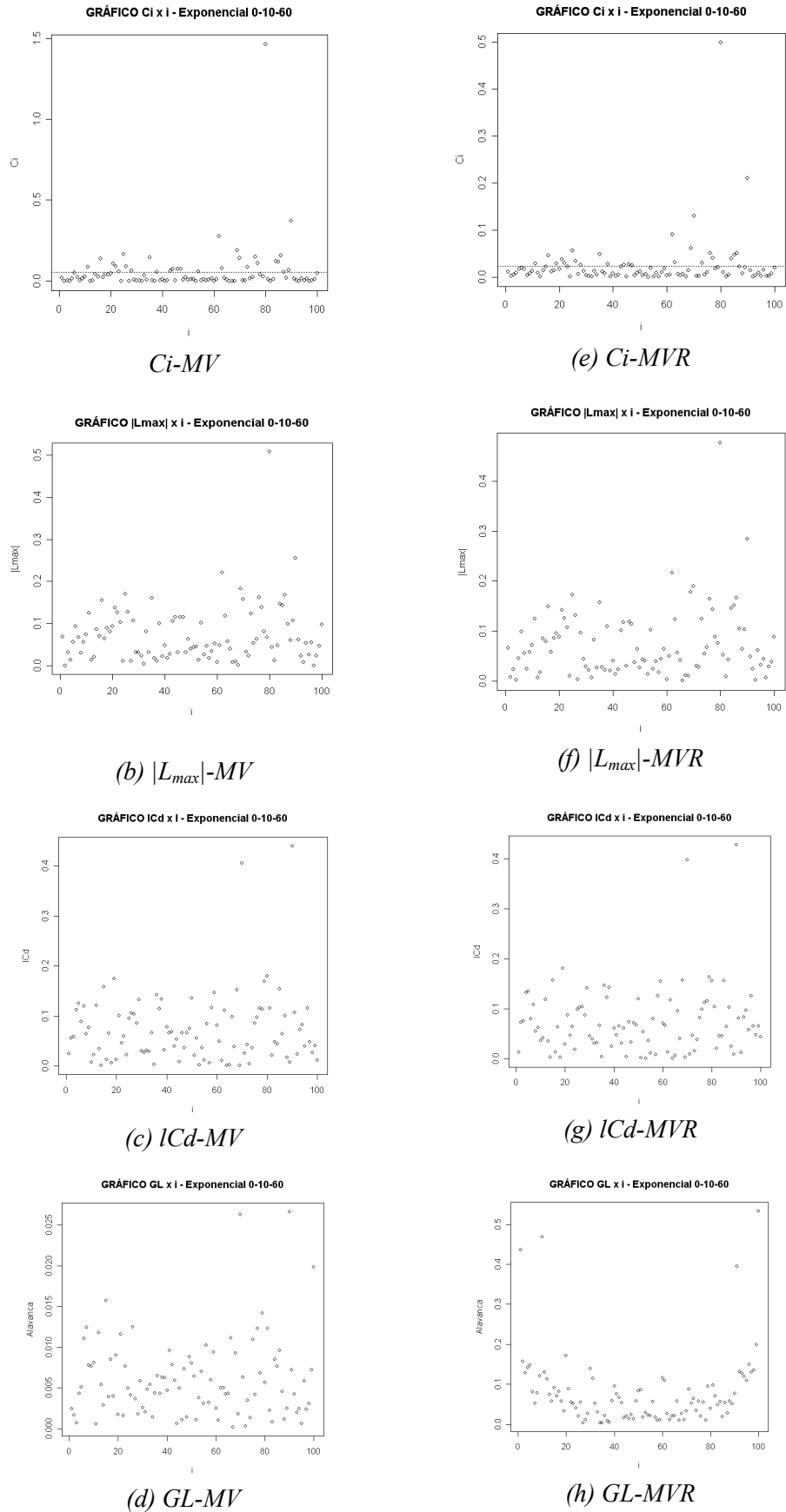


Figura 8 - Gráficos de diagnóstico $C_i x i$, $|L_{max}| x i$, $ICd x i$ e $GL x i$, para a variável simulada Exponencial 0-10-60, com os estimadores MV e MVR.

Na Tabela 2 são apresentados os resultados das análises espaciais realizadas para as variáveis simuladas com o modelo gaussiano e com parâmetros estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR).

Tabela 2 - Parâmetros estimados para a função gaussiana com os estimadores de máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR)

Variável Simulada	Obs. perturbada	φ_1		φ_2		φ_3	
		MV	MVR	MV	MVR	MV	MVR
Gaussiano 0-10-10	35	4,166	4,229	8,607	8,905	13,949	14,269
Gaussiano 0-10-15	44	2,487	2,564	9,014	9,426	16,920	17,431
Gaussiano 0-10-20	32	0,894	0,906	9,446	9,853	18,009	18,243
Gaussiano 0-10-60	84	1,027	1,030	3,462	4,258	45,314	48,087

Pelos resultados da Tabela 2, pode-se observar que para as variáveis simuladas “Gaussiano 0-10-10” e “Gaussiano 0-10-15”, os métodos MV e MVR superestimaram os parâmetros φ_1 e φ_3 e subestimaram o parâmetro φ_2 . Entretanto, para as variáveis “Gaussiano 0-10-20” e “Gaussiano 0-10-60”, os parâmetros estimados por MV e MVR tiveram valores mais próximos dos simulados. Isso indica que, quando as variáveis foram simuladas com o valor do parâmetro φ_3 , que é função do alcance, mais elevado, o valor perturbado teve menos influência na estimação dos parâmetros. Ainda, neste caso, observa-se que as estimativas por MVR obtiveram valores mais próximos dos simulados.

Portanto, os valores perturbados nas variáveis simuladas “Gaussiano 0-10-10” e “Gaussiano 0-10-15” foram mais influentes na estimação dos parâmetros do que os valores perturbados nas variáveis “Gaussiano 0-10-20” e “Gaussiano 0-10-60”, visto que nessas variáveis os parâmetros estimados são próximos aos fornecidos para a simulação.

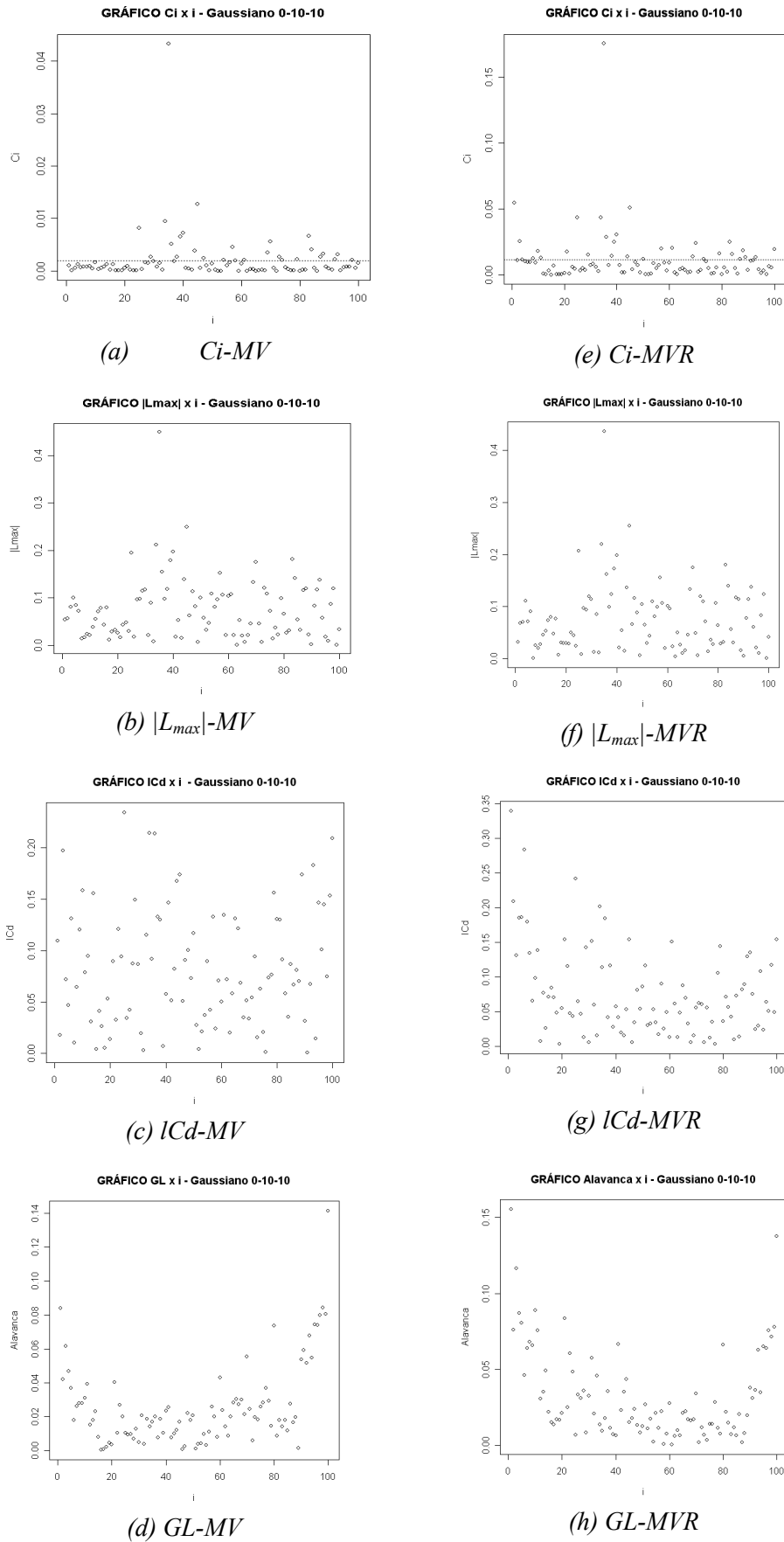


Figura 9 - Gráficos de diagnóstico $C_i \times i$, $|L_{max}| \times i$, $ICd \times i$ e $GL \times i$, para a variável simulada Gaussiano 0-10-10, com os estimadores MV e MVR.

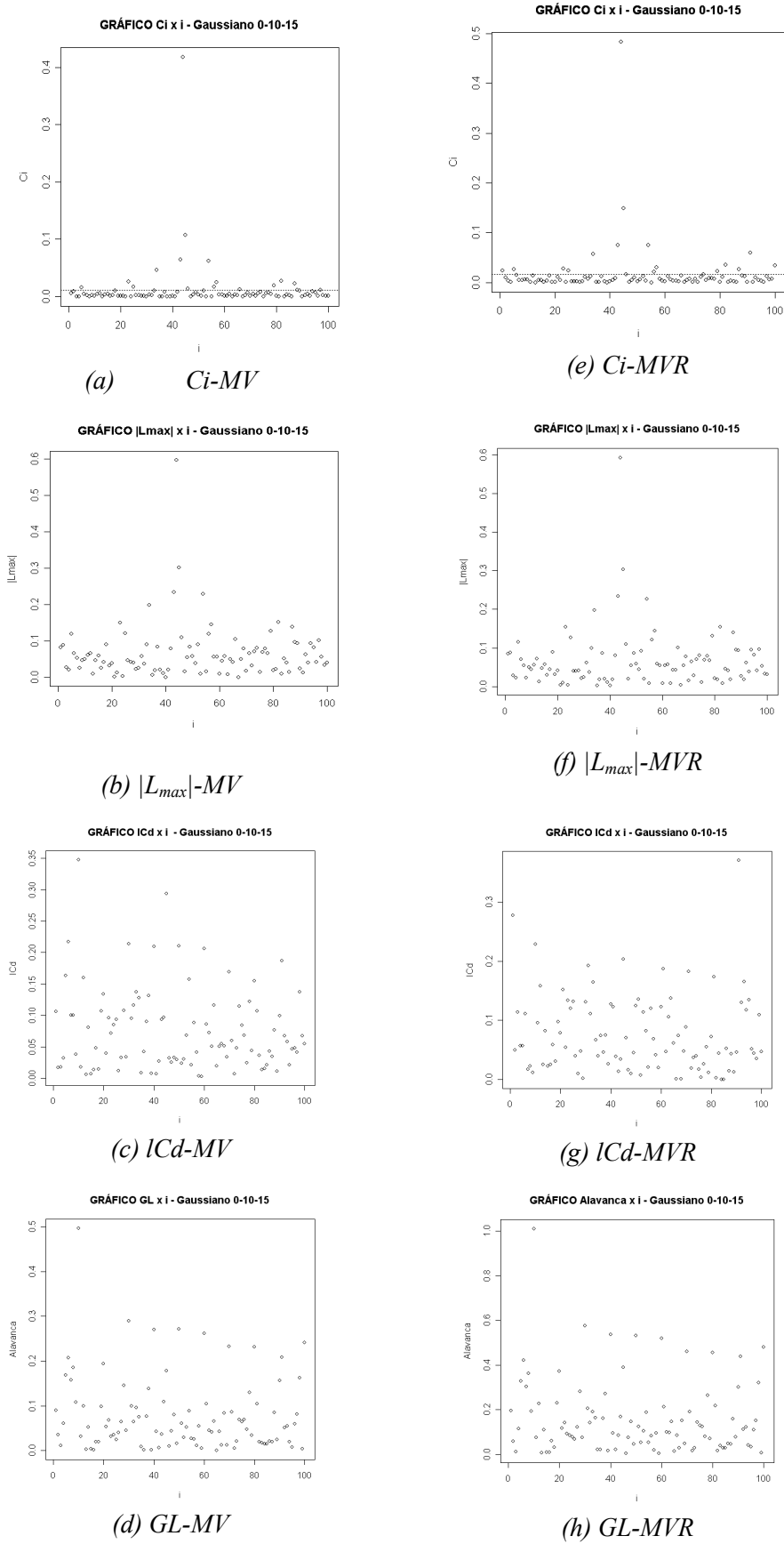


Figura 10 - Gráficos de diagnóstico $C_i \times i$, $|L_{max}| \times i$, $ICd \times i$ e $GL \times i$, para a variável simulada Gaussiano 0-10-15, com os estimadores MV e MVR.

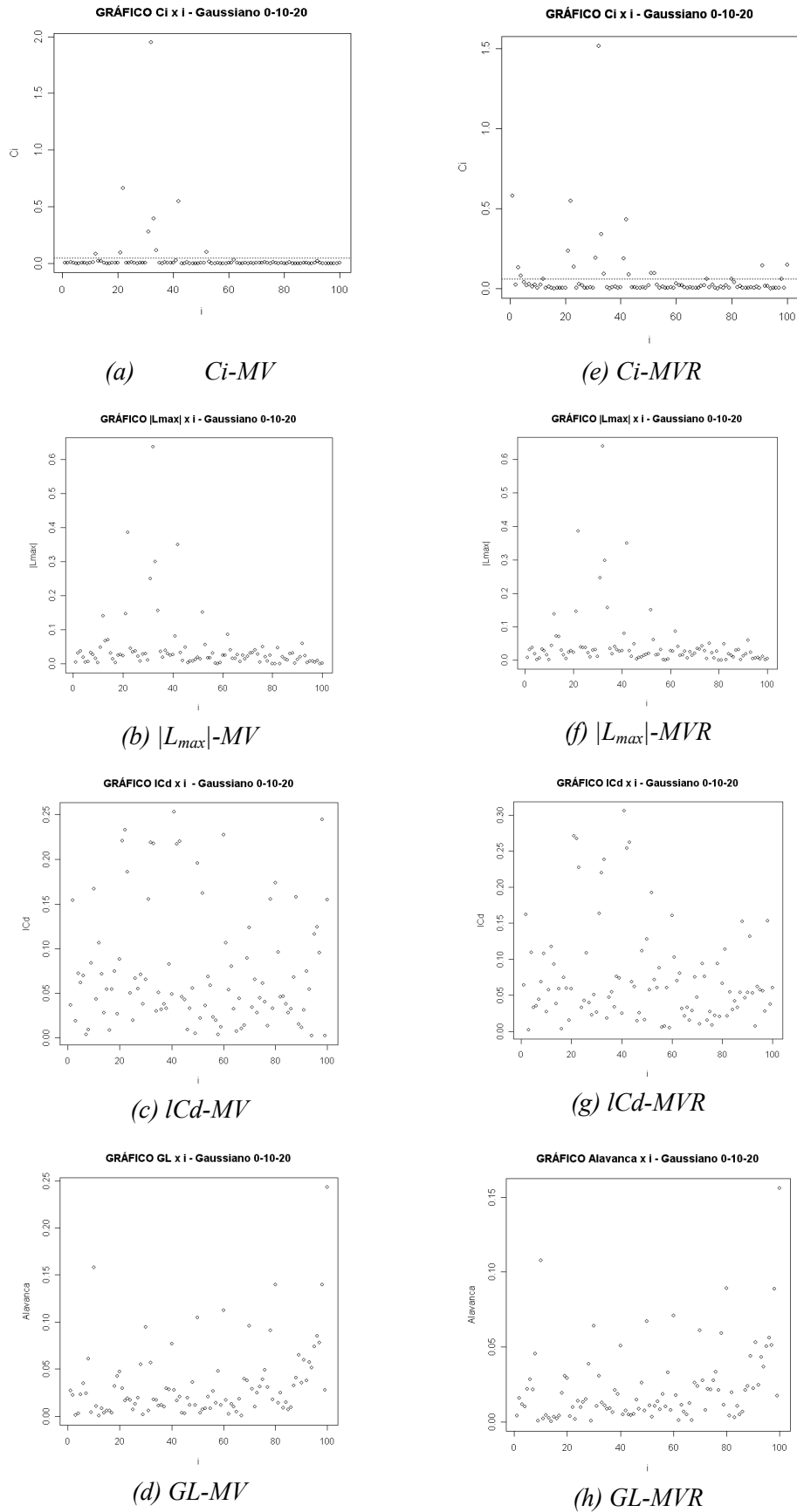


Figura 11 - Gráficos de diagnóstico C_i x i , $|L_{max}|$ x i , ICd x i e GL x i , para a variável simulada Gaussiano 0-10-20, com os estimadores MV e MVR.

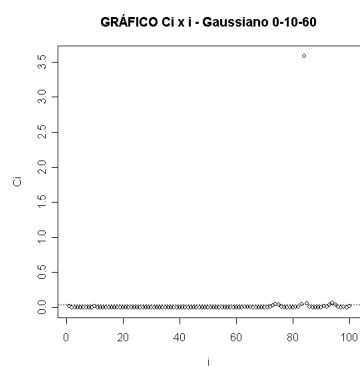
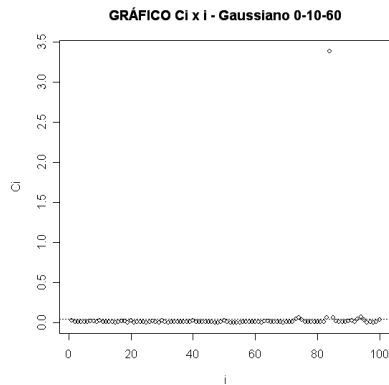
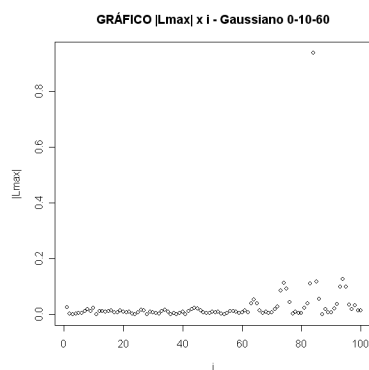
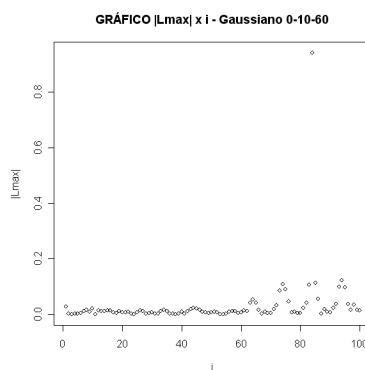
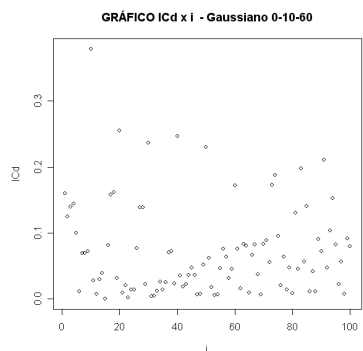
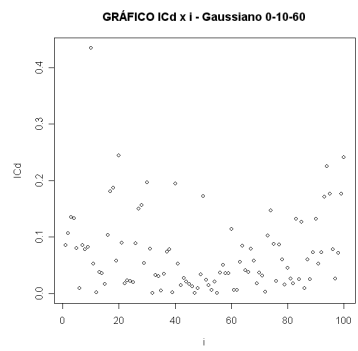
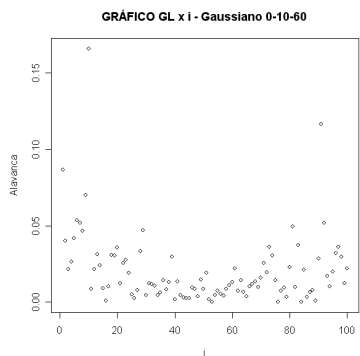
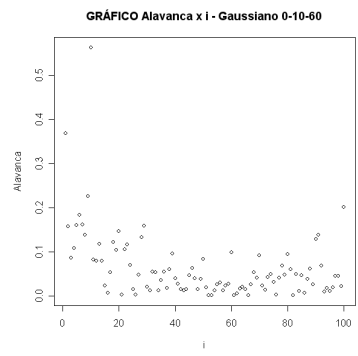
(a) C_i -MV(e) C_i -MVR(b) $|L_{max}|$ -MV(f) $|L_{max}|$ -MVR(c) ICd -MV(g) ICd -MVR(d) GL -MV(h) GL -MVR

Figura 12 - Gráficos de diagnóstico C_i x i , $|L_{max}|$ x i , ICd x i e GL x i , para a variável simulada Gaussiano 0-10-60, com os estimadores MV e MVR.

Na Tabela 3 são apresentados os resultados das análises espaciais realizadas para as variáveis simuladas com o modelo Matérn e parâmetros $kappa$ iguais a 0,7 e 3,0.

Tabela 3 - Parâmetros estimados para a função Matérn com os estimadores de máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR)

Variável Simulada	Obs. perturbada	φ_1		φ_2		φ_3	
		MV	MVR	MV	MVR	MV	MVR
Matérn 0-10-10 k= 0,7	42	1,603	2,140	8,332	8,457	9,828	11,669
Matérn 0-10-15 k= 0,7	91	0,000	0,000	14,409	16,170	13,732	15,409
Matérn 0-10-20 k= 0,7	06	0,223	0,552	9,225	9,912	13,473	15,556
Matérn 0-10-60 k= 0,7	31	0,009	0,093	6,825	8,374	23,241	28,945
Matérn 0-10-10 k= 3,0	11	0,232	0,218	93,082	75,689	16,084	15,000
Matérn 0-10-15 k= 3,0	03	0,766	0,781	14,791	21,770	20,278	23,238
Matérn 0-10-20 k= 3,0	99	0,698	0,708	15,346	21,845	21,354	21,020
Matérn 0-10-60 k= 3,0	50	0,814	0,822	3,098	5,410	28,524	35,666

Na maioria dos casos observa-se que, quando as variáveis foram simuladas com o valor do parâmetro φ_3 , que é função do alcance do modelo, mais elevado, os parâmetros estimados por MVR apresentam-se mais próximos dos simulados, quando comparados com os obtidos por MV.

De forma particular, para a variável “Matérn 0-10-15 k=0,7”, os valores dos parâmetros estimados φ_1 foram iguais aos fornecidos para simulação e os valores de φ_3 foram próximos aos simulados. Entretanto, os valores dos parâmetros φ_2 foram superestimados. E para a variável “Matérn 0-10-10 k=3,0”, observa-se que o parâmetro φ_2 teve seu valor superestimado de forma bem acentuada.

Portanto, os valores perturbados nas variáveis simuladas com o modelo Matérn com parâmetros $kappa$ iguais a 0,7 e 3,0 foram bastante influentes na estimação dos parâmetros, tanto por MV como por MVR.

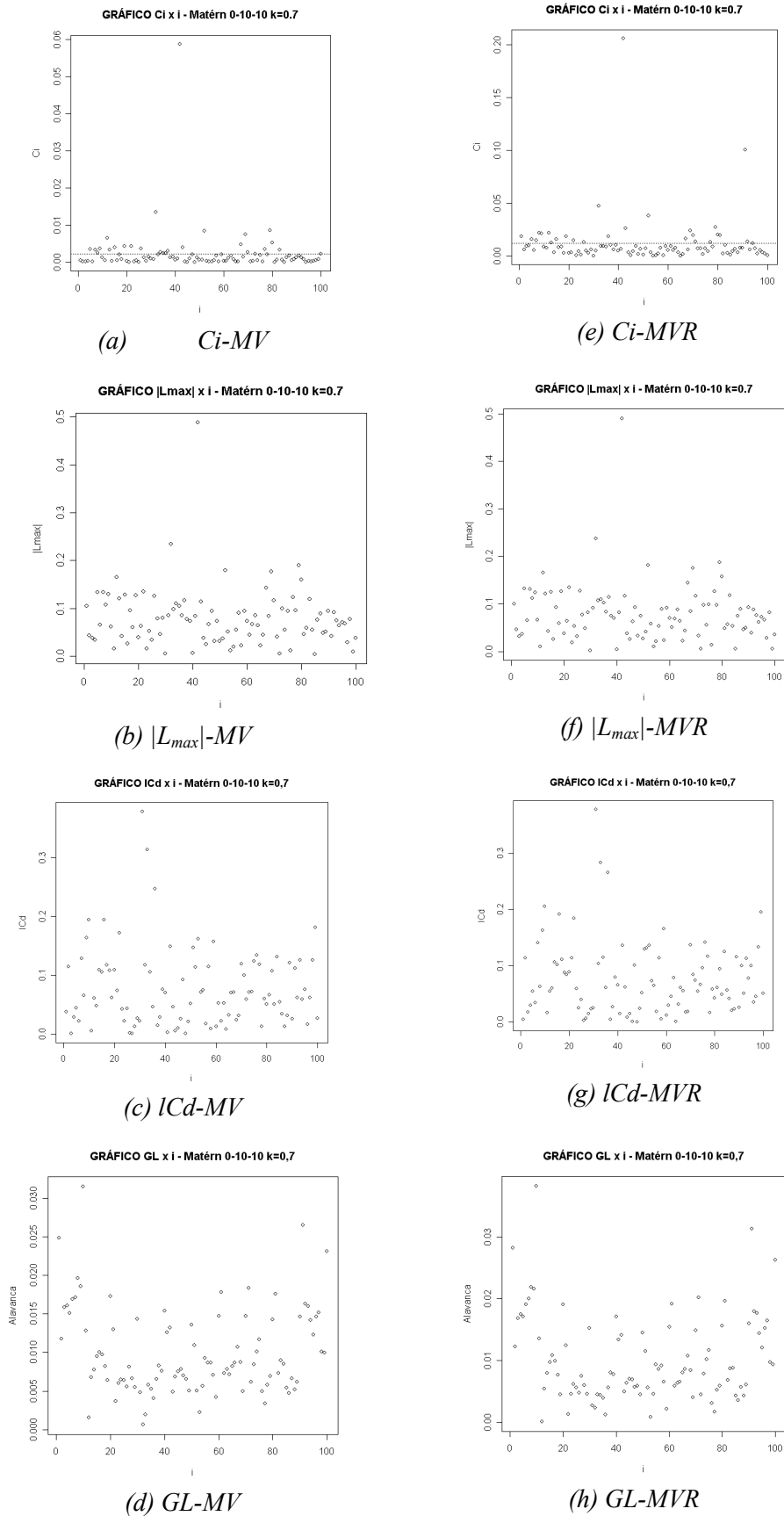


Figura 13 - Gráficos de diagnóstico C_i x i , $|L_{max}|$ x i , ICd x i e GL x i , para a variável simulada Matérn 0-10-10 $k=0,7$, com os estimadores MV e MVR.

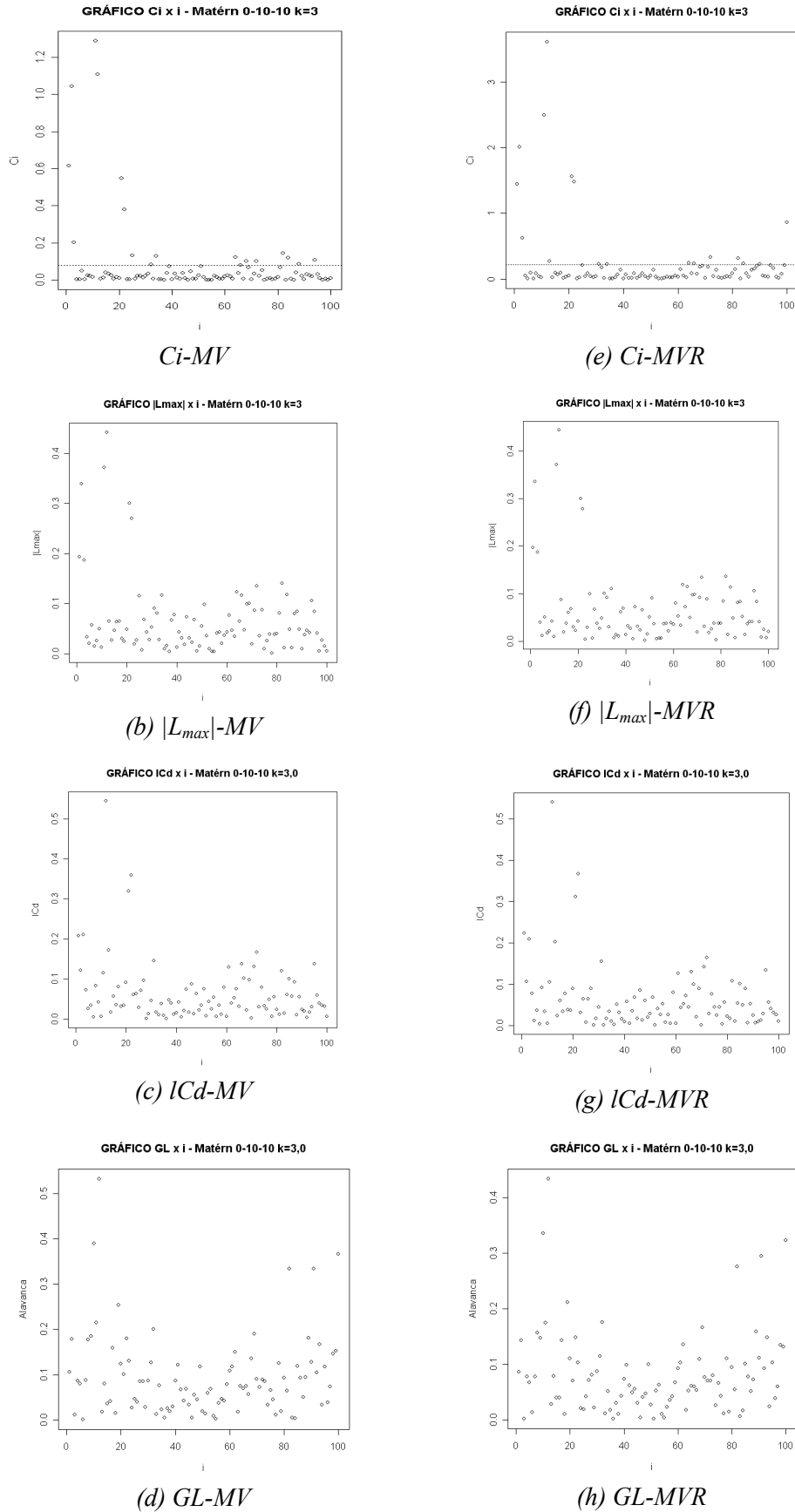


Figura 14 - Gráficos de diagnóstico $C_i \times i$, $|L_{max}| \times i$, $ICd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-10 $k=3,0$, com os estimadores MV e MVR.

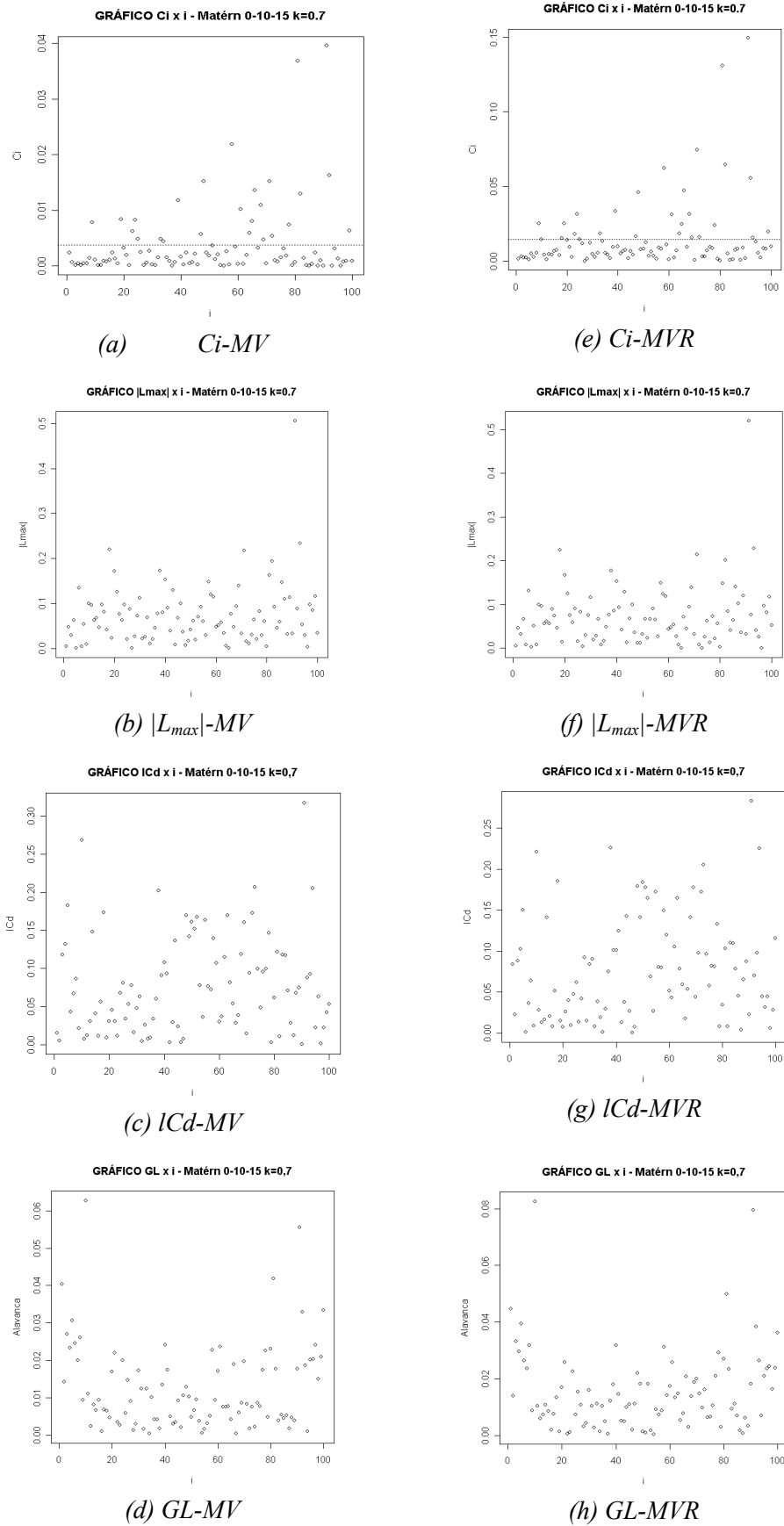


Figura 15 - Gráficos de diagnóstico $C_i \times i$, $|L_{max}| \times i$, $ICd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-15 $k=0,7$, com os estimadores MV e MVR.

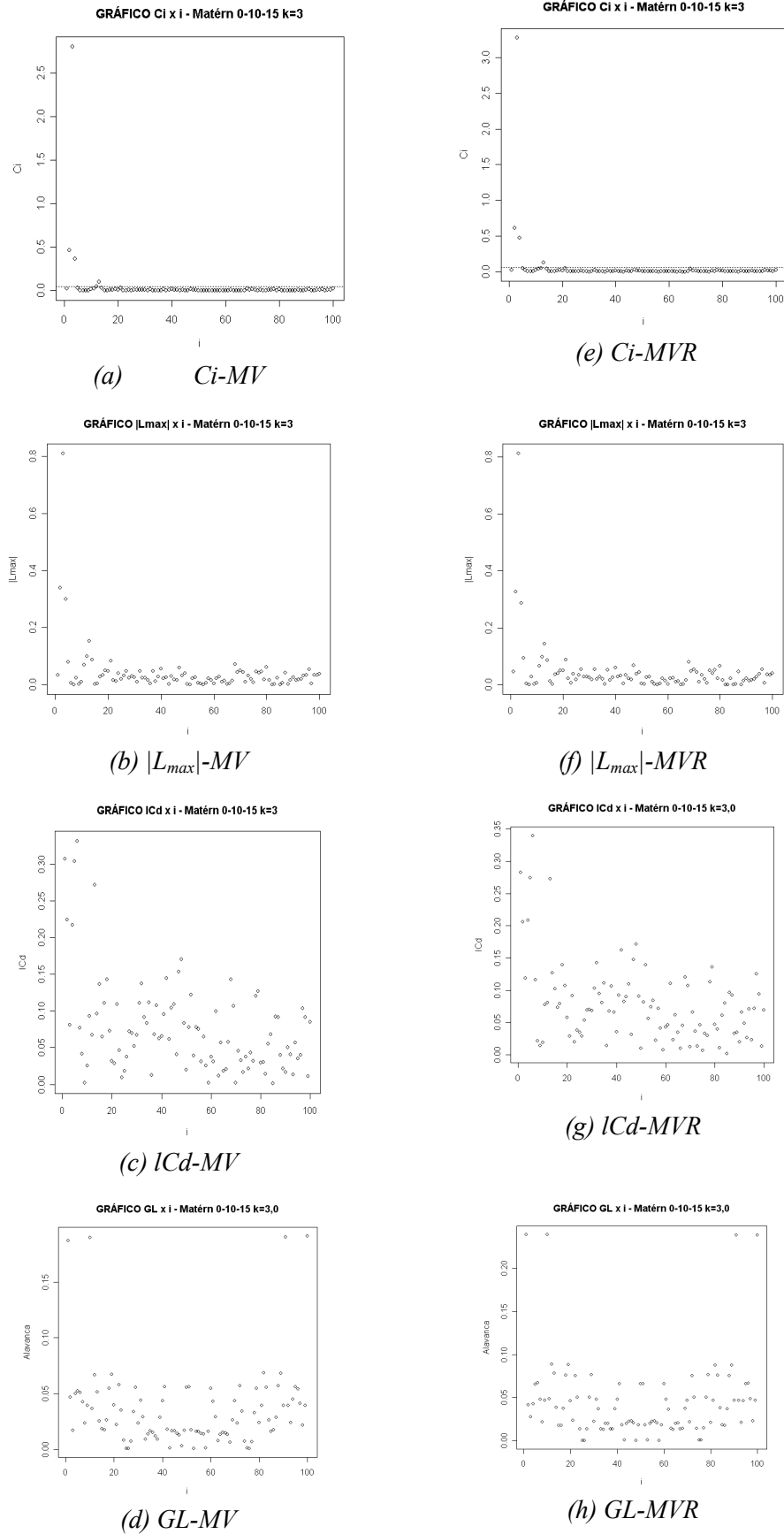


Figura 16 - Gráficos de diagnóstico C_i x i , $|L_{max}|$ x i , ICd x i e GL x i , para a variável simulada Matérn 0-10-15 $k=3,0$, com os estimadores MV e MVR.

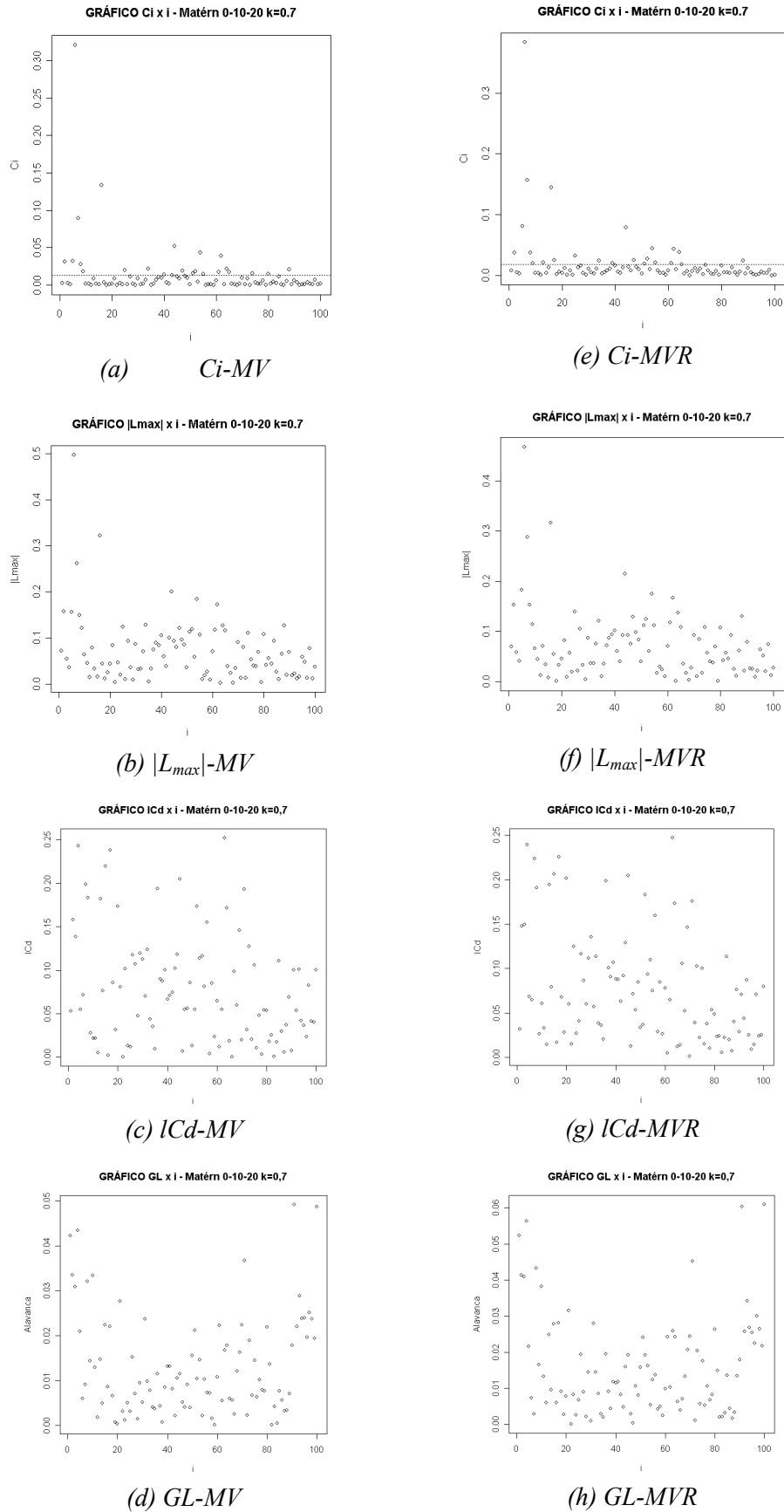


Figura 17 - Gráficos de diagnóstico $C_i \times i$, $|L_{max}| \times i$, $ICd \times i$ e $GL \times i$, para a variável simulada Matérn 0-10-20 $k=0.7$, com os estimadores MV e MVR.

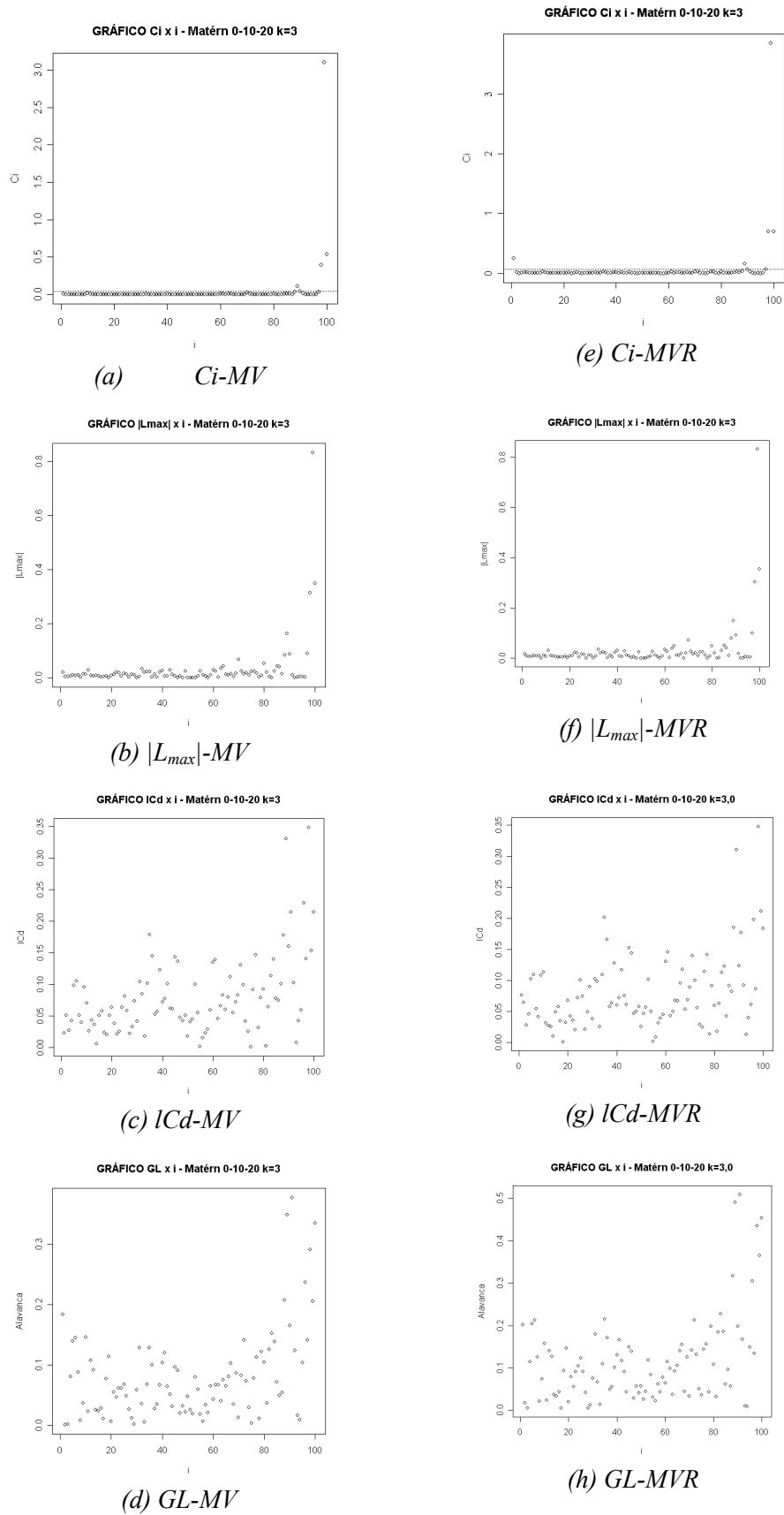


Figura 18 - Gráficos de diagnóstico C_i x i , $|L_{max}|$ x i , lCd x i e GL x i , para a variável simulada Matérn 0-10-20 $k=3,0$, com os estimadores MV e MVR.

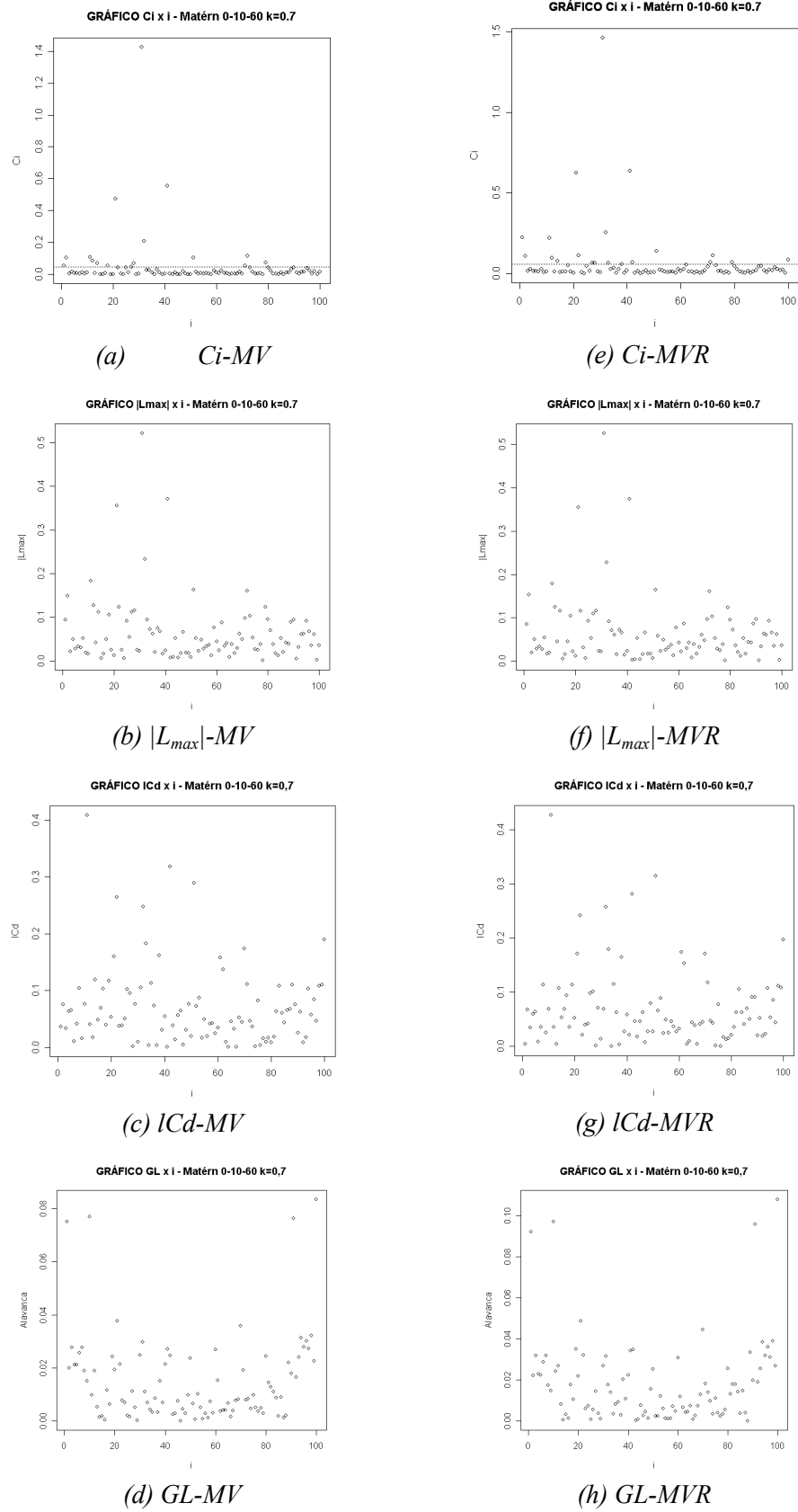


Figura 19 - Gráficos de diagnóstico C_i x i , $|L_{max}|$ x i , ICd x i e GL x i , para a variável simulada Matérn 0-10-60 $k=0,7$, com os estimadores MV e MVR.

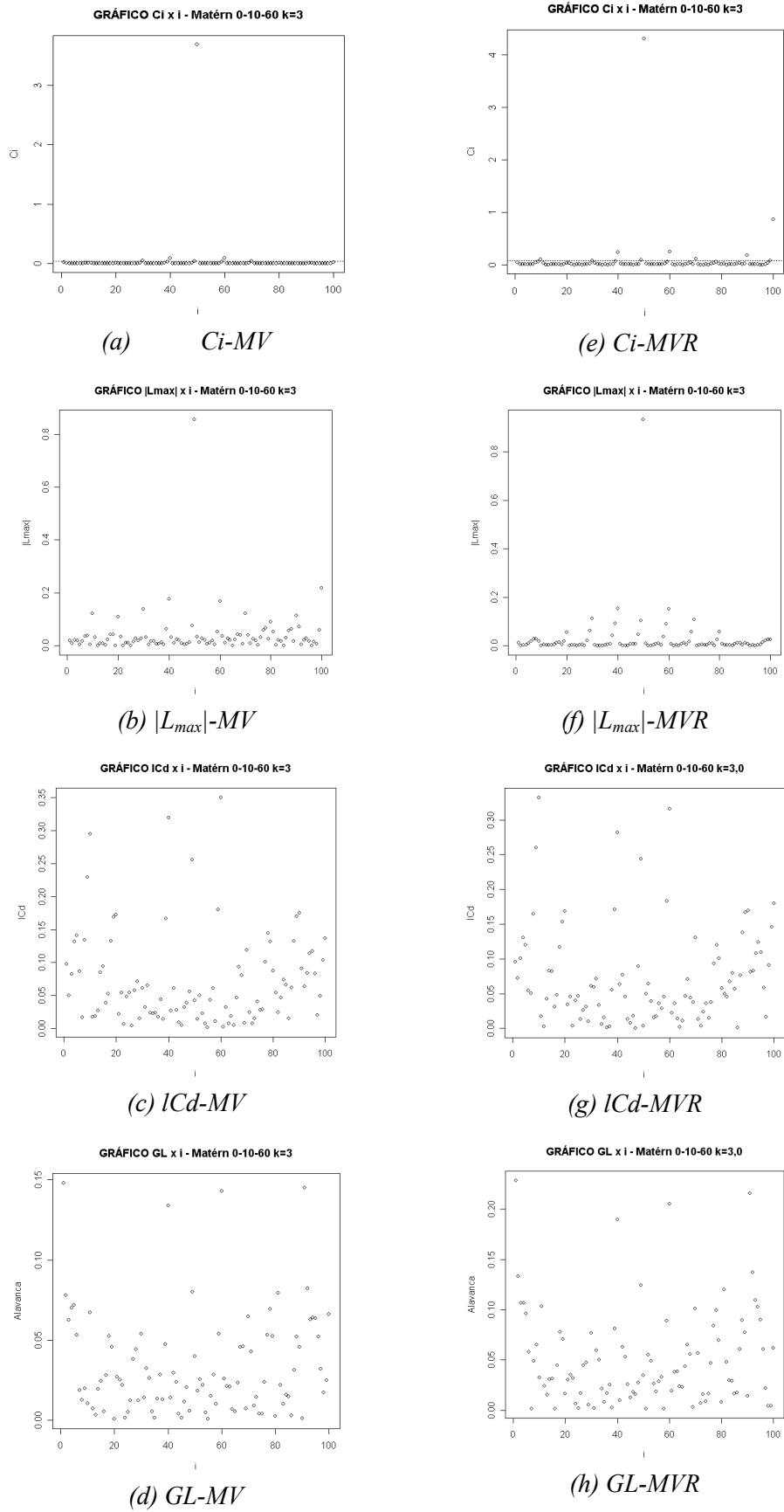


Figura 20 - Gráficos de diagnóstico C_i x i , $|L_{max}|$ x i , lCd x i e GL x i , para a variável simulada Matérn 0-10-60 $k=3,0$, com os estimadores MV e MVR.

Na Tabela 4 é apresentado um resumo com as variáveis simuladas, os respectivos pontos perturbados e os gráficos de diagnóstico utilizados.

Tabela 4 - Resumo dos gráficos de diagnóstico que identificaram a observação perturbada como sendo influente

Variável	Obs. perturbada	C_i		$ L_{max} $		l_{Cd}		GL	
		MV	MVR	MV	MVR	MV	MVR	MV	MVR
Exponencial 0-10-10	66	66	66	66	66	-	-	-	-
Exponencial 0-10-15	19	19	19	19	19	-	-	-	-
Exponencial 0-10-20	21	21	21	21	21	-	-	-	-
Exponencial 0-10-60	80	80	80	80	80	-	-	-	-
Gaussiano 0-10-10	35	35	35	35	35	-	-	-	-
Gaussiano 0-10-15	44	44	44	44	44	44	-	-	-
Gaussiano 0-10-20	32	32	32	32	32	-	-	-	-
Gaussiano 0-10-60	84	84	84	84	84	-	-	-	-
Matérn 0-10-10 k=0,7	42	42	42	42	42	-	-	-	-
Matérn 0-10-15 k=0,7	91	91	91	91	91	91	91	91	91
Matérn 0-10-20 k=0,7	06	06	06	06	06	06	06	06	06
Matérn 0-10-60 k=0,7	31	31	31	31	31	-	-	-	-
Matérn 0-10-10 k=3,0	11	11	11	11	11	11	11	11	11
Matérn 0-10-15 k=3,0	03	03	03	03	03	-	-	-	-
Matérn 0-10-20 k=3,0	99	99	99	99	99	99	99	99	99
Matérn 0-10-60 k=3,0	50	50	50	50	50	50	50	-	-

Pelos resultados da Tabela 4 observa-se que as medidas de diagnóstico C_i e $|L_{max}|$ foram eficientes na identificação da observação perturbada em todas as variáveis simuladas. Já as medidas l_{Cd} e GL não apresentaram o mesmo desempenho, pois apenas identificaram a observação perturbada para algumas variáveis simuladas com o modelo Matérn.

4.2 ESTUDO COM DADOS DA BIBLIOGRAFIA

Os dados em estudo são relativos à seleção de concentração (em ppb) de *trichloroethylene* (TCE), apresentadas por PIEGORSCH e BAILER (2005).

Para as análises destes dados, os autores transformaram os dados de TCE por logaritmo. Portanto, as análises realizadas foram desenvolvidas com dados transformados.

Na Tabela 5 é apresentada a análise estatística descritiva para a variável $\ln(TCE)$.

Tabela 5 - Análise estatística descritiva para a variável $\ln(TCE)$

Variável	Média	Mín.	Max.	Q1	Median.	Q3	CV(%)	DP	p-valor
$\ln(TCE)$	5,86	0,92	11,12	3,73	6,05	7,94	46,22	2,71	> 0,05

Nota: Mín.: valor mínimo; Máx.: valor máximo; Q1: primeiro quartil; Q3: terceiro quartil; CV: coeficiente de variação; DP: desvio-padrão; p-valor: nível descritivo do teste de normalidade de Shapiro & Wilks, ao nível de 5% de significância.

Na Figura 21 está representado o desenho esquemático *Box-plot* para o logaritmo de TCE. Pelo gráfico *Box-plot* não se observam pontos discrepantes neste conjunto de dados e, além disso, os dados apresentam-se simétricos.

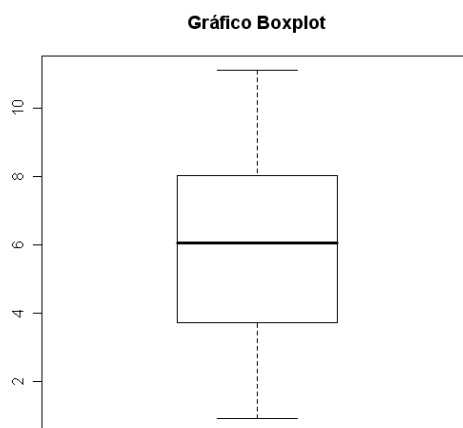


Figura 21 - Desenho esquemático *Box-plot* da variável $\ln(TCE)$.

Na Tabela 6 são apresentados os resultados obtidos por PIEGORSCH e BAILER (2005), utilizando o programa *SAS*, e o calculado utilizando o programa *R*, para o ajuste do modelo teórico com parâmetros efeito pepita (φ_1), contribuição (φ_2) e função do alcance (φ_3). O ajuste foi realizado utilizando-se o mesmo modelo e método de estimação de parâmetros dos autores, ou seja, modelo exponencial, com os parâmetros estimados por máxima verossimilhança restrita (MVR).

Como se pode observar, os resultados da estimação de parâmetros obtidos com o programa *R* foram muito semelhantes aos obtidos pelos autores citados acima.

Ao realizar o estudo de influência local, observou-se que os pontos 18, 25, 26, 27, 29 e 38 foram considerados influentes no afastamento da verossimilhança (técnicas C_i e $|L_{max}|$) e na função de covariância (técnica ICd) (Figura 22). Quando se retiraram as observações consideradas influentes no modelo ajustado, o efeito pepita (φ_1) permaneceu o mesmo, entretanto, os valores da contribuição (φ_2) e função do alcance (φ_3) elevaram-se, mostrando que as observações retiradas estavam exercendo influência na estimação dos parâmetros do modelo ajustado.

Tabela 6 - Comparação entre os parâmetros de dependência espacial (φ_1 , φ_2 , $\varphi_1 + \varphi_2$ e φ_3) obtidos por PIEGORSCH e BAILER (2005), os calculados utilizando-se o programa R e retirando-se as observações influentes

Variável	Método ¹	Modelo ²	φ_1	φ_2	$\varphi_1 + \varphi_2$	φ_3
ln (TCE) *	MVR	Exp.	0,0002	9,2040	9,2042	28,694
ln(TCE) calculados	MVR	Exp.	0,0000	9,2060	9,2060	28,698
ln(TCE) sem influentes	MVR	Exp.	0,0000	11,4530	11,4530	44,030

Nota: ¹MVR: estimador de máxima verossimilhança restrita; ²Exp.: modelo exponencial.

*Resultados obtidos por PIEGORSCH e BAILER (2005).

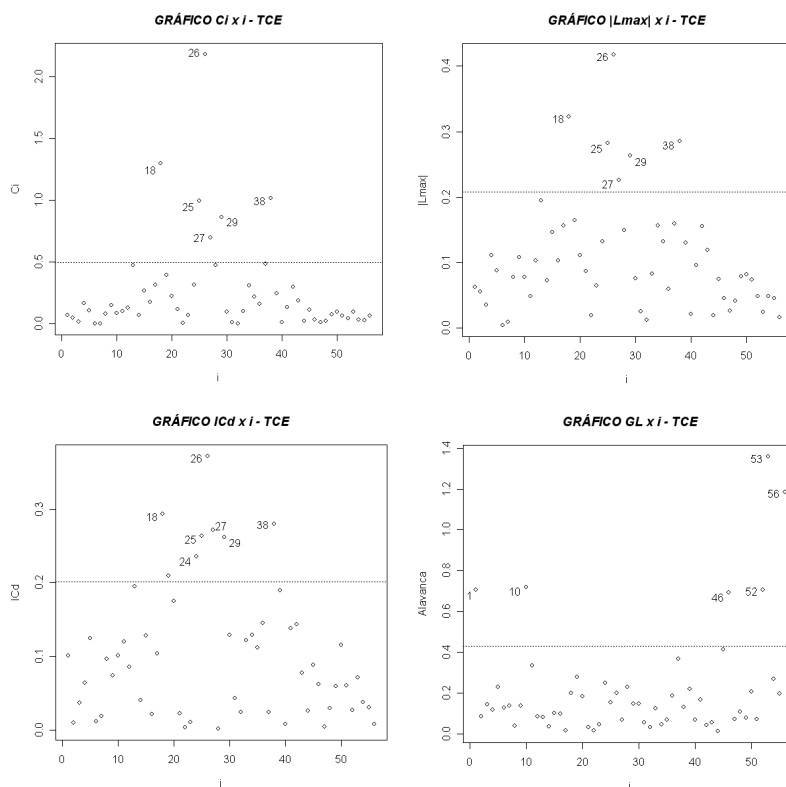


Figura 22 - Gráficos de diagnóstico $C_i \times i$, $|L_{max}| \times i$, $ICd \times i$ e $GL \times i$, para a variável “ln (TCE) calculados”.

Na Figura 23 são apresentados os mapas temáticos, construídos com a interpolação de valores por krigagem ordinária, para as variáveis “ln (TCE) calculados” e “ln (TCE) sem influentes”.

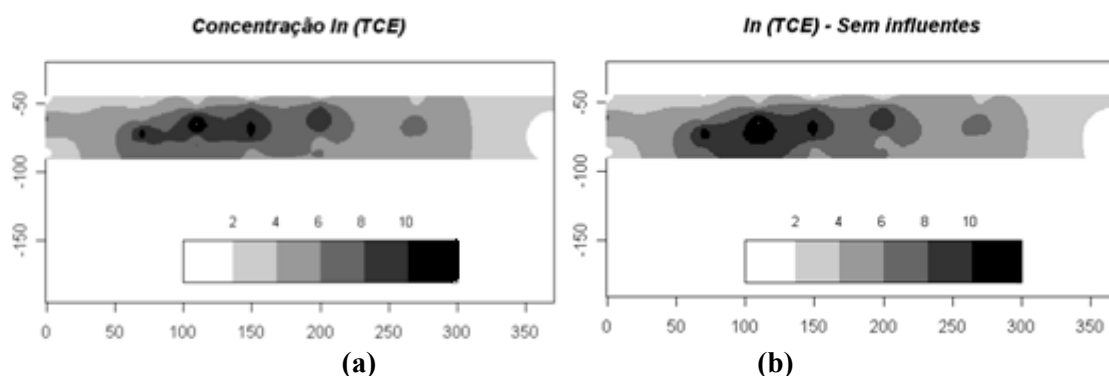


Figura 23 - Mapas temáticos para as variáveis: “ln (TCE) calculados” a); “ln (TCE) sem influentes” b).

Na Figura 23 pode-se observar quando se retirou as observações influentes (Figura 23b), houve alterações na distribuição espacial da concentração de *trichloroethylene* (TCE). Nota-se, principalmente um aumento da área com concentração entre 8 e 10 ppb.

4.3 ESTUDOS COM DADOS EXPERIMENTAIS

No ano agrícola de 2006/2007 foram realizadas análises químicas de amostras do solo de uma área de 71 ha, obtendo-se dados do teor de fósforo e manganês do solo, com um total de 46 observações, segundo a amostragem apresentada na Figura 4.

A seguir são apresentadas as análises estatísticas das variáveis teor de fósforo e manganês do solo, as quais foram georreferenciadas com o sistema GPS.

4.3.1 Estudo do Teor Fósforo

4.3.1.1 Análise espacial do teor fósforo (mg dm^{-3})

Após uma amostragem do solo, com elementos amostrais coletados em distâncias de $50 \times 50 \text{ m}$, $75 \times 75 \text{ m}$ e $141 \times 141 \text{ m}$, as amostras foram encaminhadas ao laboratório para análises químicas. Na primeira análise, o fósforo apresentou três valores discrepantes em relação aos demais, que são as observações das amostras números 1, 26 e 45, com valores $60,0$; $38,6$ e $60,0 \text{ mg dm}^{-3}$, respectivamente (Figura 24a e 24b). Foram repetidas as análises químicas do solo nessas amostras e constatou-se que não houve alteração dos valores. Portanto, verificou-se que as observações 1, 26 e 45 são valores reais e não erros de medição ou de análises de laboratório, não devendo ser descartados, pois, representam uma característica local, devido, provavelmente, ao excesso de adubação, segundo entrevistas com os técnicos que participaram da coleta dos dados.

Na Tabela 7 são apresentadas as estatísticas descritivas para a variável fósforo, com todos os dados coletados.

Tabela 7 - Estatística descritiva para a variável fósforo (P) [mg dm^{-3}]

Variável	N	Média	Mín.	Máx.	Q1	Med.	Q3	DP	CV (%)	p-valor
P	46	15,80	5,70	60,00	9,70	11,70	18,60	11,50	72,76	< 0,05*

Nota: N: número de dados; Mín.: valor mínimo; Máx.: valor máximo; Q1: primeiro quartil; Med.: mediana; Q3: terceiro quartil; DP: desvio-padrão; CV: coeficiente de variação; p-valor: nível descritivo do teste de normalidade de Shapiro & Wilks, ao nível de 5% de significância.

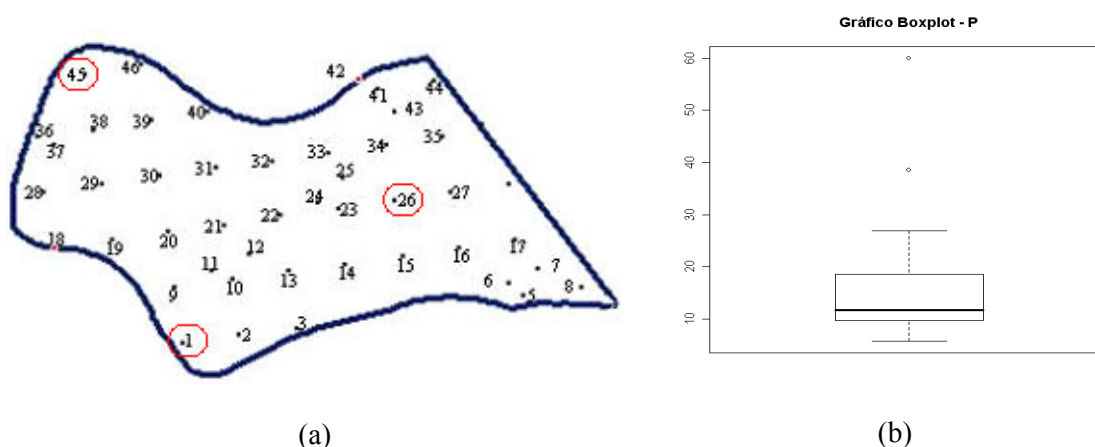


Figura 24 - Croqui da disposição amostral dos dados na área em estudo de 71 ha (a); Gráfico *Box-plot* da variável fósforo com os dados totais (b).

Pelos resultados da Tabela 7, verifica-se que a média do teor de fósforo ($15,80 \text{ mg dm}^{-3}$) está bem acima da faixa recomendada para a cultura da soja (EMBRAPA, 1979). Verifica-se que os dados são altamente heterogêneos, pois o coeficiente de variação foi muito elevado ($CV = 72,76\%$), devido aos valores discrepantes das observações 1, 26 e 45, identificados pelo gráfico *Box-plot* (Figura 24b).

Pelo p -valor do teste de normalidade de Shapiro & Wilks, a 5% de significância, apresentado na Tabela 7, rejeitou-se a hipótese de normalidade dos dados do fósforo. Como as técnicas estatísticas aplicadas neste trabalho supõem que os dados têm distribuição gaussiana, aplicou-se a transformação de Box-Cox com λ igual a $-0,7$ aos dados desta variável.

Na Tabela 8 são apresentadas as análises estatísticas espaciais para a variável fósforo, com todos os dados coletados (P), em que foram ajustados os modelos teóricos exponencial, gaussiano e Matérn, com parâmetros estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR).

Tabela 8 - Modelos espaciais ajustados e estimação dos parâmetros efeito pepita (ϕ_1), contribuição (ϕ_2) e função do alcance (ϕ_3), pelos métodos de máxima verossimilhança e máxima verossimilhança restrita, para a variável fósforo (mg dm^{-3})

Variável	Mod.	Mét.	Média β	Efeito pepita ϕ_1	Contribuição ϕ_2	Função do alcance ϕ_3	E (%)	
Fósforo	Exp.	MV	1,1822 (0,0155)	0,0043 (0,0023)	0,0020 (0,0024)	108,6275 (1,0833)	68,25	
		MVR	1,1824 (0,0234)	0,0046 (0,0016)	0,0024 (0,0021)	235,0387 (3,8662)	65,71	
		Gaus.	MV	1,1817 (0,0165)	0,0052 (0,0014)	0,0012 (0,0013)	214,8181 (0,6126)	81,25
			MVR	1,1817 (0,0217)	0,0052 (0,0013)	0,0018 (0,0015)	292,3174 (1,1816)	74,29
	Matérn k=0,7	MV	1,1822 (0,0156)	0,0046 (0,0052)	0,0017 (0,0034)	100,9344 (172,0850)	73,02	
		MVR	1,1823 (0,0230)	0,0048 (0,0036)	0,0022 (0,0022)	193,7198 (239,1845)	68,57	
	Matérn k=3,0	MV	1,1819 (0,0164)	0,0051 (0,0061)	0,0013 (0,0019)	60,0610 (72,1858)	79,69	
		MVR	1,1820 (0,0216)	0,0051 (0,0055)	0,0018 (0,0016)	86,5556 (74,2993)	73,91	

Nota: $E=(\phi_1/\phi_1+\phi_2)*100$: coeficiente de efeito pepita relativo; alcance a : modelo exponencial: $a = 3\phi_3$, modelo gaussiano: $a = \sqrt{3}\phi_3$; modelo Matérn: $a = \phi_3$.

Pelos resultados apresentados na Tabela 8, observa-se que não há grande diferença entre os valores dos parâmetros estimados efeito pepita (ϕ_1) e contribuição (ϕ_2). Já para o parâmetro ϕ_3 , que é função do alcance a , observa-se grande diferença entre os valores obtidos pelos diferentes modelos. Os menores valores foram obtidos pelo modelo Matérn, com parâmetro $kappa$ igual a 0,7 e 3,0, além de apresentar desvios padrão muito altos. Pelo coeficiente de efeito pepita relativo E , verifica-se que apenas os ajustes com os modelos gaussiano e Matérn ($kappa$ igual a 3,0), com parâmetros estimados por MV, apresentaram fraca dependência espacial ($E > 75\%$), os demais ajustes indicaram moderada dependência espacial ($25\% < E < 75\%$) (CAMBARDELLA *et al.* 1994).

4.3.1.2 Análises de diagnóstico

Foram aplicadas técnicas gráficas de diagnóstico com a finalidade de avaliar se os pontos discrepantes 1, 26 e 45 ou alguns outros, estariam exercendo algum tipo de influência no afastamento da verossimilhança, nas análises da estrutura de dependência espacial ($\varphi_1, \varphi_2, \varphi_3$), tais como na função de covariância e escolha do modelo a ser ajustado, e no preditor linear. As técnicas utilizadas foram o gráfico dos coeficientes de influências C_i versus i (ordem das observações coletadas) e gráfico dos autovetores $|L_{max}|$ versus i ; gráficos ICd versus i e gráficos de alavanca generalizada GL versus i .

A primeira análise de diagnóstico foi realizada para a variável fósforo com o total de 46 dados coletados (P), utilizando os modelos espaciais ajustados exponencial, gaussiano e Matérn, com parâmetro $kappa$ igual a 0,7 e 3,0.

Nas Figuras 25 e 26 são apresentados os gráficos de diagnóstico utilizando os modelos espaciais exponencial, gaussiano e Matérn ($kappa$ igual a 0,7 e 3,0) com o estimador de máxima verossimilhança (MV) e nas Figuras 28 e 29, os mesmos gráficos e modelos, mas utilizando o estimador de máxima verossimilhança restrita (MVR).

Pelos gráficos das figuras 25 a 28, pode-se verificar que as observações números 1, 26 e 45 foram consideradas altamente influentes no afastamento da verossimilhança, na matriz de covariância e nos valores preditos, pois apresentaram altos valores para as medidas de diagnóstico utilizadas. Então, estas observações foram retiradas do conjunto total de dados e novamente realizadas as análises.

Nos gráficos de diagnóstico, as linhas horizontais indicam que os pontos que estiverem acima são considerados valores influentes, de acordo com o valor de corte de cada uma das técnicas. Já as linhas verticais foram utilizadas para identificar quais foram as observações consideradas influentes em cada gráfico.

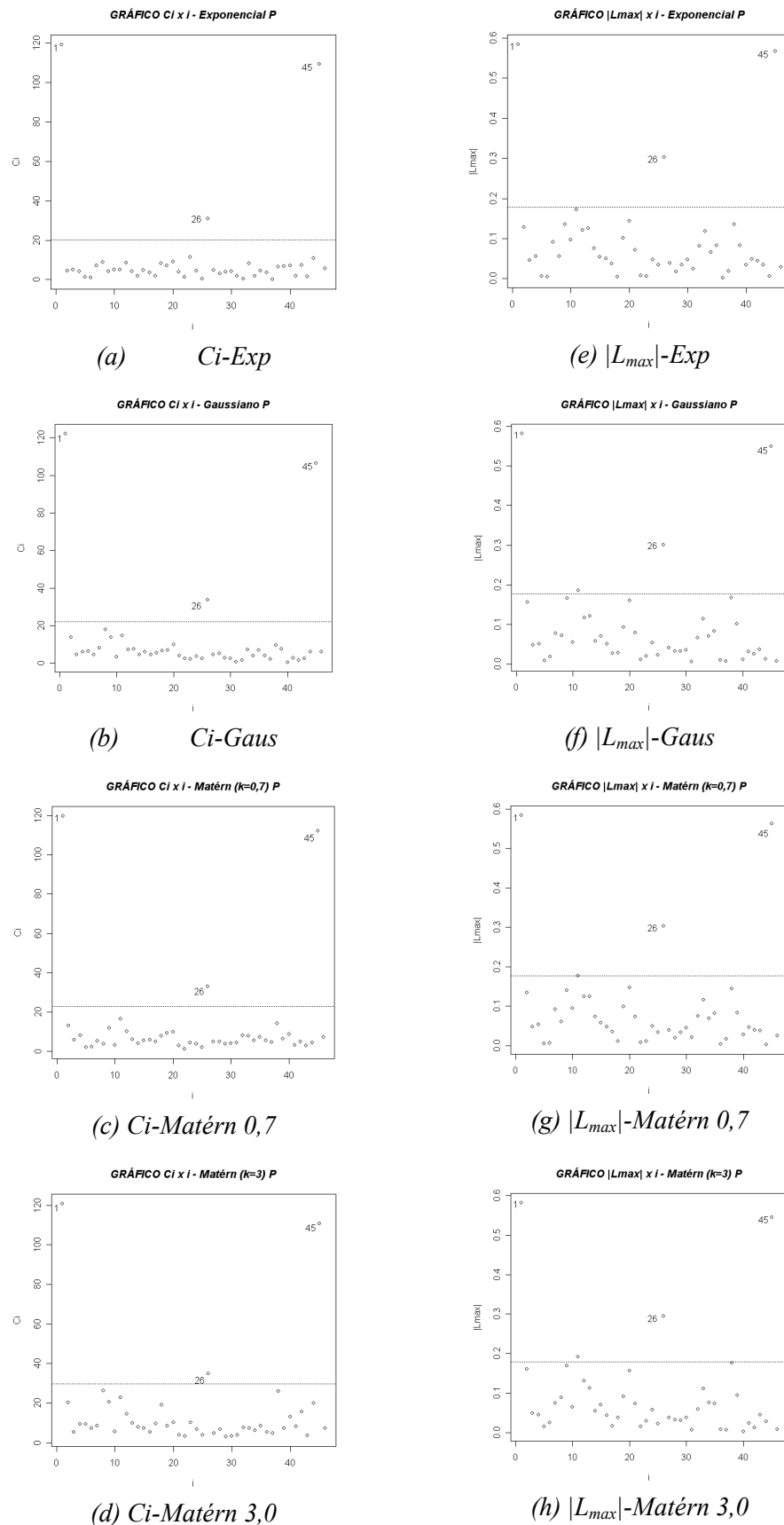


Figura 25 - Gráficos de diagnóstico $C_i \times i$ e $|L_{max}| \times i$, para a variável fósforo, com o estimador MV.

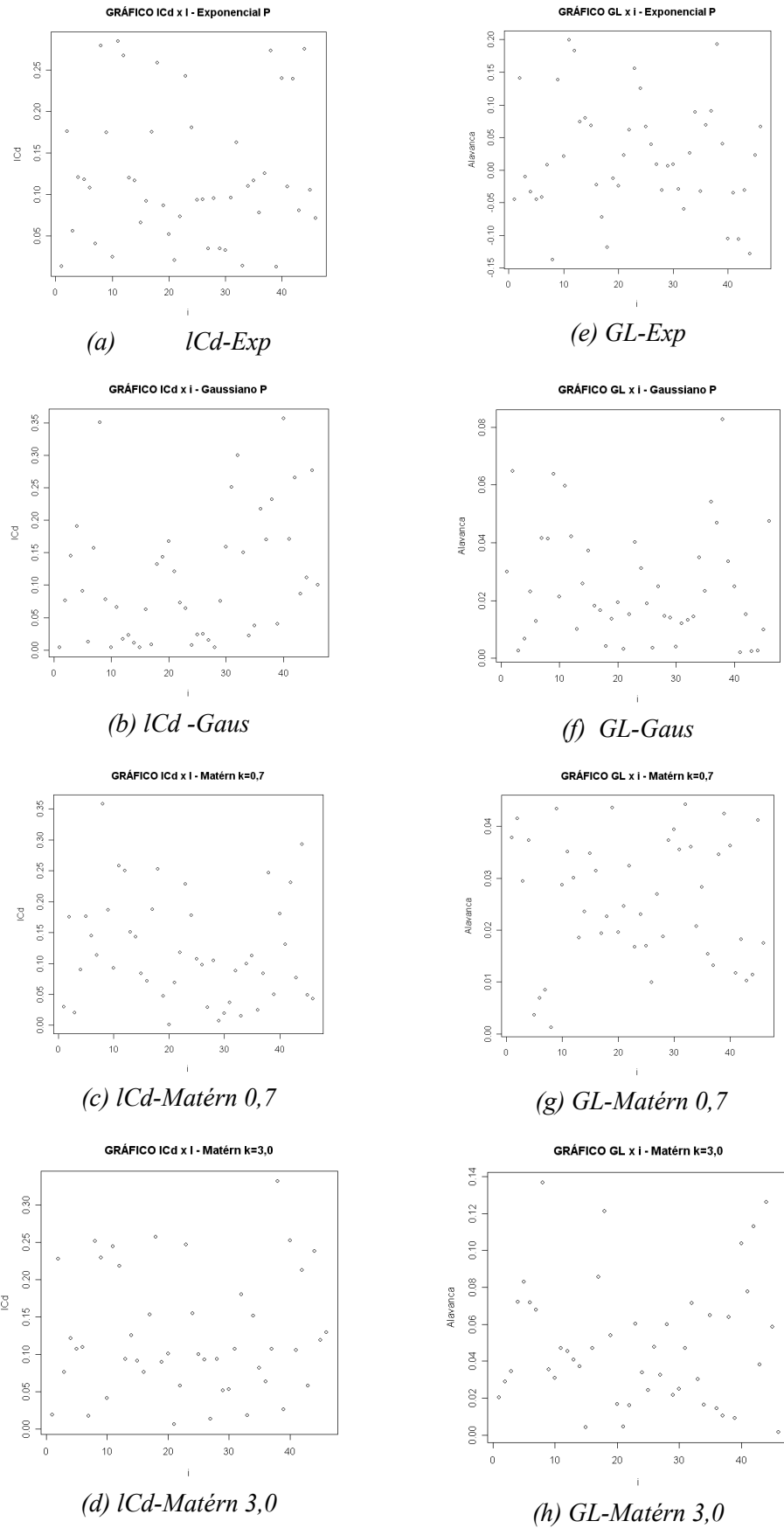


Figura 26 - Gráficos de diagnóstico *ICd x i* e *GL x i*, para a variável fósforo, com o estimador MV.

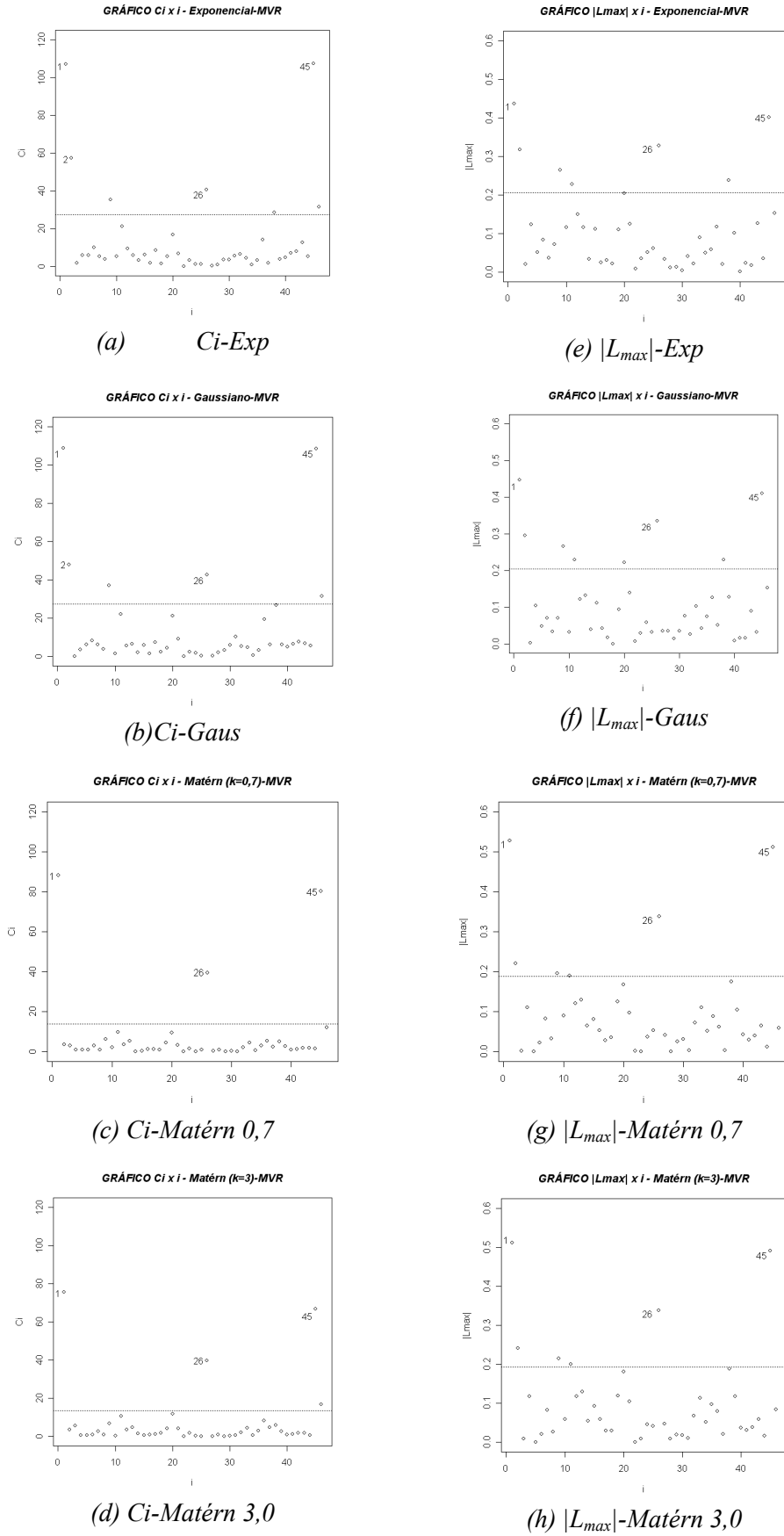


Figura 27 - Gráficos de diagnóstico $C_i \times i$ e $|L_{max}| \times i$, para a variável fósforo, com o estimador MVR.

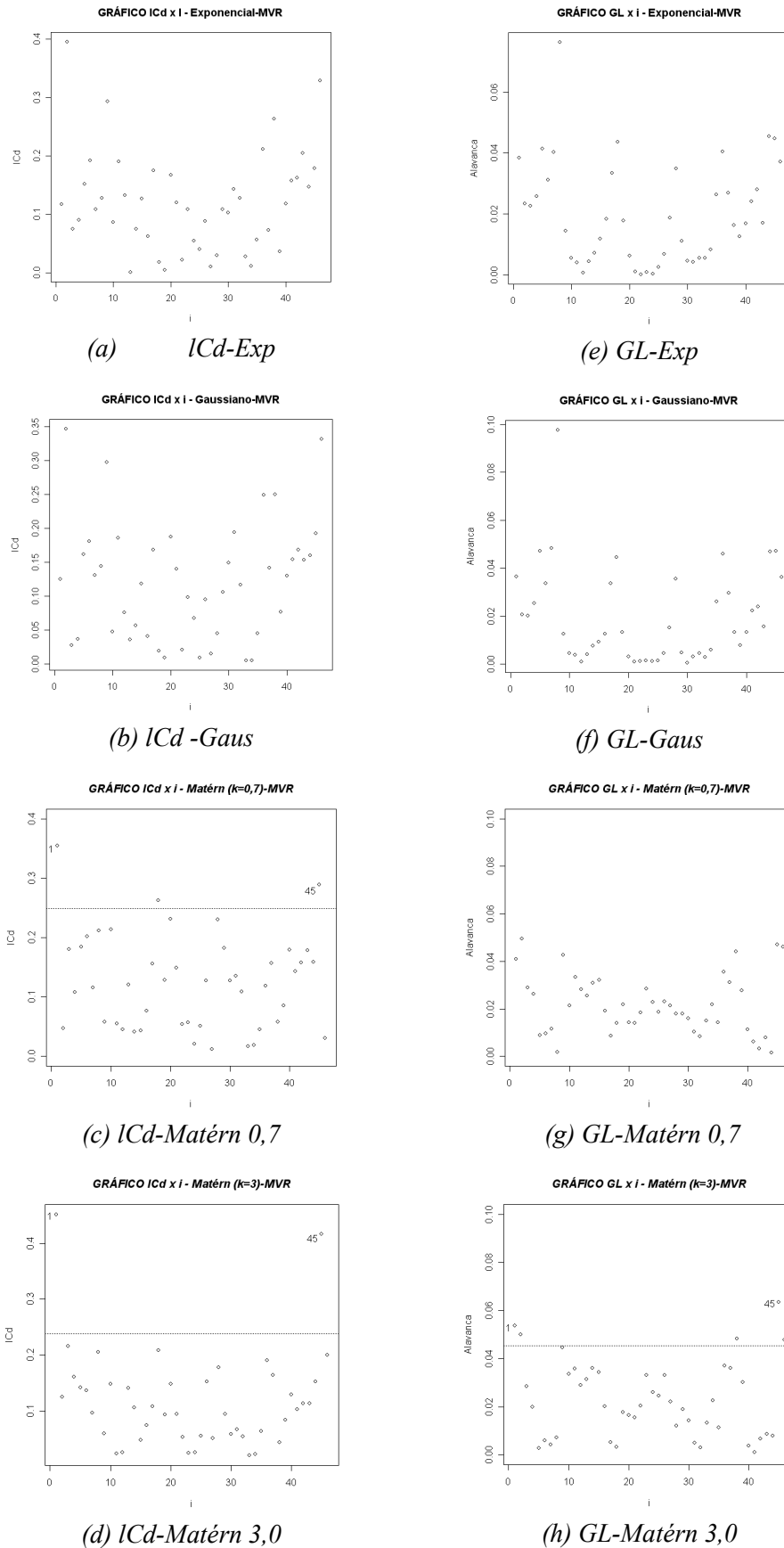


Figura 28 - Gráficos de diagnóstico ICd x i e GL x i , para a variável fósforo, com o estimador MVR.

4.3.1.3 Análises de influência

4.3.1.3.1 Influência nas análises descritivas

Na Tabela 9 são apresentadas as análises descritivas para o conjunto de dados com um total de 46 amostras do fósforo (P) e dados do fósforo, retirando-se as observações números 1, 26 e 45 (P-1-26-45) e verificando-se a influência na análise descritiva.

Tabela 9 - Análise estatística descritiva para a variável fósforo, com dados totais coletados (P) e retirando-se as observações 1, 26 e 45 (P-1-26-45)

Variável	N	Média	Mín.	Máx.	Q1	Med.	Q3	DP	CV (%)	p-valor
P	46	15,80	5,70	60,00	9,70	11,70	18,60	11,50	72,76	< 0,05*
P-1-26-45	43	13,21	5,70	26,90	9,70	11,70	16,55	5,42	40,99	< 0,05*

Nota: N: número de dados; Med.: mediana; DP: desvio-padrão; CV: coeficiente de variação; p-valor: nível descritivo do teste de normalidade de Shapiro & Wilks ao nível de 5% de significância.

Observa-se que há diferença entre as médias da variável fósforo com todos os dados, e retirando as observações influentes. Sem dados influentes, o valor da média baixou. O mesmo comportamento é observado no desvio padrão. Ao se analisar o coeficiente de variação (CV) do fósforo, verifica-se que, sem as observações 1, 26 e 45, o valor do CV também foi bem menor, quando se utilizaram os dados totais. Entretanto, como CV ficou acima de 30%, os dados ainda são considerados heterogêneos (GOMES, 2000).

4.3.1.3.2 Influência nas estimativas dos parâmetros

Na Tabela 10 são apresentados os resultados das análises de

variabilidade espacial, por meio da geoestatística, para os dados originais e retirando-se os dados considerados influentes. São apresentados os resultados da estimação dos parâmetros média (β), efeito pepita (φ_1), contribuição (φ_2) e alcance (φ_3), pelos modelos exponencial, gaussiano e Matérn, por máxima verossimilhança e máxima verossimilhança restrita. Apresentam-se também, entre parêntesis, os respectivos desvios padrão para os parâmetros estimados.

Tabela 10 - Parâmetros estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR) para a variável fósforo, com dados totais (P) e retirando-se as observações 1, 26 e 45 (P-1-26-45), utilizando os modelos teóricos exponencial, gaussiano e Matérn

Modelo	Método	Dados	Média β	Efeito pepita φ_1	Contribuição φ_2	Função do alcance φ_3
Exp.	MV	Dados	1,1822	0,0043	0,0020	108,6275
		originais	(0,0155)	(0,0023)	(0,0024)	(1,0833)
		Sem influentes	1,1698 (0,0137)	0,0038 (0,0017)	0,0011 (0,0017)	123,4113 (1,7771)
	MVR	Dados	1,1824	0,0046	0,0024	235,0387
		originais	(0,0234)	(0,0016)	(0,0021)	(3,8662)
		Sem influentes	1,1686 (0,1657)	0,0038 (0,3086)	0,0014 (0,1905)	221,8948 (398,2035)
Gaus.	MV	Dados	1,1817	0,0052	0,0012	214,8181
		originais	(0,0165)	(0,0014)	(0,0013)	(0,6126)
		Sem influentes	1,1690 (0,0151)	0,0042 (0,0011)	0,0008 (0,0009)	243,1303 (0,9982)
	MVR	Dados	1,1817	0,0052	0,0018	292,3174
		originais	(0,0217)	(0,0013)	(0,0015)	(1,1816)
		Sem influentes	1,1681 (0,0185)	0,0041 (0,0010)	0,0012 (0,0011)	292,5954 (1,3624)
Matérn k=0,7	MV	Dados	1,1822	0,0046	0,0017	100,9344
		originais	(0,0156)	(0,0052)	(0,0034)	(172,0850)
		Sem influentes	1,1700 (0,0141)	0,0040 (0,0037)	0,0010 (0,0024)	114,3420 (243,1572)
	MVR	Dados	1,1823	0,0048	0,0022	193,7198
		originais	(0,0230)	(0,0036)	(0,0022)	(239,1845)
		Sem influentes	1,1685 (0,0186)	0,0040 (0,0027)	0,0013 (0,0016)	189,6063 (287,1136)
Matérn k=3,0	MV	Dados	1,1819	0,0051	0,0013	60,0610
		originais	(0,0164)	(0,0061)	(0,0019)	(72,1858)
		Sem influentes	1,1693 (0,0147)	0,0042 (0,0045)	0,0008 (0,0013)	67,5349 (97,0722)
	MVR	Dados	1,1820	0,0051	0,0018	86,5556
		originais	(0,0216)	(0,0055)	(0,0016)	(74,2993)
		Sem influentes	1,1683 (0,0187)	0,0041 (0,0043)	0,0012 (0,0012)	88,5350 (87,4575)

De forma geral, como se pode observar na Tabela 10, os valores dos parâmetros β , φ_1 , φ_2 , estimados sem as observações influentes, foram inferiores aos estimados com dados originais. Já o parâmetro φ_3 , em todas as estimativas por MV, sem as observações influentes, teve seus valores elevados, com relação aos obtidos com dados originais. O mesmo não ocorreu para as estimativas do parâmetro φ_3 , quando estimado por MVR.

4.3.1.3.3 Influência na validação dos modelos

Na Tabela 11 são apresentados os resultados dos critérios de Akaike (\hat{A}), Filliben (r) e valor máximo do logaritmo da função verossimilhança (VMV), para validação de ajustes dos modelos geoestatísticos, para a variável fósforo com dados originais (P) e sem as observações influentes (P-1-26-45).

Tabela 11 - Resultados da validação pelos critérios de Akaike (\hat{A}), Filliben (r^*) e valor máximo da verossimilhança (VMV)

	Modelo	Mét.	Nº dados	\hat{A}	r^*	VMV
P Dados totais	Exp.	MV	46	-2671,605	0,811	-151,5
		MVR		-2670,189	0,815	-152,6
	Gaus.	MV	46	-2675,987	0,806	-151,3
		MVR		-2679,415	0,808	-152,5
	Matérn k=0,7	MV	46	-2653,826	0,811	-151,4
		MVR		-2682,174	0,814	-152,6
Matérn k=3,0	MV	46	-2712,174	0,808	-151,4	
	MVR		-2633,121	0,810	-152,5	
P Sem dados influentes	Exp.	MV	43	-2507,233	0,972	-129,4
		MVR		-2500,032	0,973	-130,8
	Gaus.	MV	43	-2489,941	0,973	-129,2
		MVR		-2503,918	0,974	-130,5
	Matérn k=0,7	MV	43	-2492,434	0,972	-129,4
		MVR		-2467,378	0,973	-130,8
	Matérn k=3,0	MV	43	-2503,918	0,973	-129,3
		MVR		-2492,434	0,974	-130,6

Nota: r^* : teste de Filliben, a 5% de significância, compara-se com $r = 0,9740$ (Dados originais) e $r = 0,9730$ (Sem dados influentes).

Segundo o critério de AKAIKE (1973), o menor valor de \hat{A} , para o conjunto original de dados, aponta como melhor ajuste o modelo Matérn, com

$kappa$ igual a 3,0, estimado pelo método MV. Entretanto, quando foram retiradas as observações influentes, o menor valor de \hat{A} indicou como melhor ajuste o modelo exponencial, com parâmetros estimados pelo método MV.

O critério de FILLIBEN (1975) testou se os resíduos ortonormais seguiam uma distribuição normal ao nível de 5 % de significância. Para os dados da variável fósforo com dados originais e sem dados influentes, têm-se 46 e 43 valores amostrados, respectivamente. Os valores tabelados de r , obtidos da tabela de FILLIBEN (1975), foram de 0,9740 e 0,9730, respectivamente. Comparando-se com os valores r^* calculados (Tabela 11), para dados originais do fósforo, verificou-se que os valores r^* são menores do que r tabelado, indicando que a hipótese de que os dados seguem distribuição normal foi rejeitada, para um nível de 5% de significância. Pelo mesmo critério de Filliben, para o fósforo sem dados influentes, apenas os modelos exponencial e Matérn ($k=0,7$), com parâmetros estimados por MV, obtiveram r^* menores do que r tabelado. Todos os outros modelos foram aceitos, pois os valores calculados r^* ficaram acima de 0,9730.

Finalmente, pelo valor máximo do logaritmo da função verossimilhança (VMV), tanto para fósforo com dados originais, quanto sem os valores influentes, o melhor modelo ajustado foi o gaussiano, com parâmetros estimados por MV.

Na Tabela 12, são apresentados os resultados da validação de modelos ajustados, utilizando o critério da validação cruzada para a variável fósforo com dados originais e sem as observações influentes, para os quais se ajustaram os modelos exponencial, gaussiano e Matérn, combinados com os estimadores máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR).

A comparação entre os modelos ajustados foi realizada por meio do erro médio (EM), erro médio reduzido (\overline{ER}), desvio padrão dos erros médios (DPe), desvio padrão dos erros reduzidos (S_{ER}) e do erro absoluto (EA).

Segundo McBRATNEY e WEBSTER (1986), CRESSIE (1993), VIEIRA (2000) e MELLO *et al.*, (2005), os valores de EM e \overline{ER} mais próximos de zero, o menor valor de DPe e o valor de S_{ER} mais próximo de um, são os

critérios para escolha do melhor modelo ajustado, assim como o menor valor do erro absoluto EA.

Tabela 12 - Resultados da validação cruzada para a variável teor de fósforo com dados originais e sem dados influentes

	Modelo	Mét.	EM	\overline{ER}	DPe	S_{ER}	EA	
Dados originais	Exp.	MV	1,2090	0,2343	11,54627	2,07292	316,3886	
		MVR	1,0980	0,2212	11,53704	2,02520	313,2067	
	Gaus.	MV	1,1880	0,2326	11,51368	2,08646	312,0864	
		MVR	1,0700	0,2216	11,50368	2,05184	308,1708	
	Matérn k=0,7	MV	1,2070	0,2339	11,54182	2,07344	315,9182	
		MVR	1,0950	0,2222	11,53642	2,03579	312,4415	
	Matérn k=3,0	MV	1,1970	0,2327	11,52192	2,07913	313,4758	
		MVR	1,0800	0,2222	11,51753	2,04884	309,8307	
	Sem dados influentes	Exp.	MV	0,0965	0,0248	5,34947	1,16735	184,3888
			MVR	0,0267	0,0146	5,30412	1,14151	183,0248
Gaus.		MV	0,0840	0,0195	5,25368	1,14969	181,2922	
		MVR	0,0093	0,0116	5,21748	1,13555	179,7801	
Matérn k=0,7		MV	0,0957	0,0241	5,33923	1,16464	184,1273	
		MVR	0,0252	0,0142	5,29354	1,13997	182,7251	
Matérn k=3,0		MV	0,0899	0,0211	5,28730	1,15375	182,5529	
		MVR	0,0163	0,0124	5,24524	1,13572	181,0536	

Os critérios da validação cruzada, aplicados aos modelos em estudo, para a variável fósforo com o conjunto original de dados, apresentados na Tabela 12, indicaram como melhor ajuste o modelo exponencial, com parâmetros estimados pelo método máxima verossimilhança restrita (MVR). Entretanto, quando foram retiradas as observações consideradas influentes pelas técnicas de diagnóstico, a validação cruzada indicou como melhor ajuste o modelo gaussiano, com parâmetros estimados por máxima verossimilhança restrita (MVR). Isso, indica que, pelos critérios da validação cruzada, as observações 1, 26 e 45 são influentes na escolha do melhor modelo espacial ajustado.

4.3.1.3.4 Influência na construção dos mapas temáticos

Na Figura 29 são apresentados os mapas temáticos da variável fósforo, com dados originais e sem dados influentes, confeccionados por meio da

interpolação por krigagem ordinária. Os mapas foram construídos utilizando-se os modelos indicados pelo critério da validação cruzada.

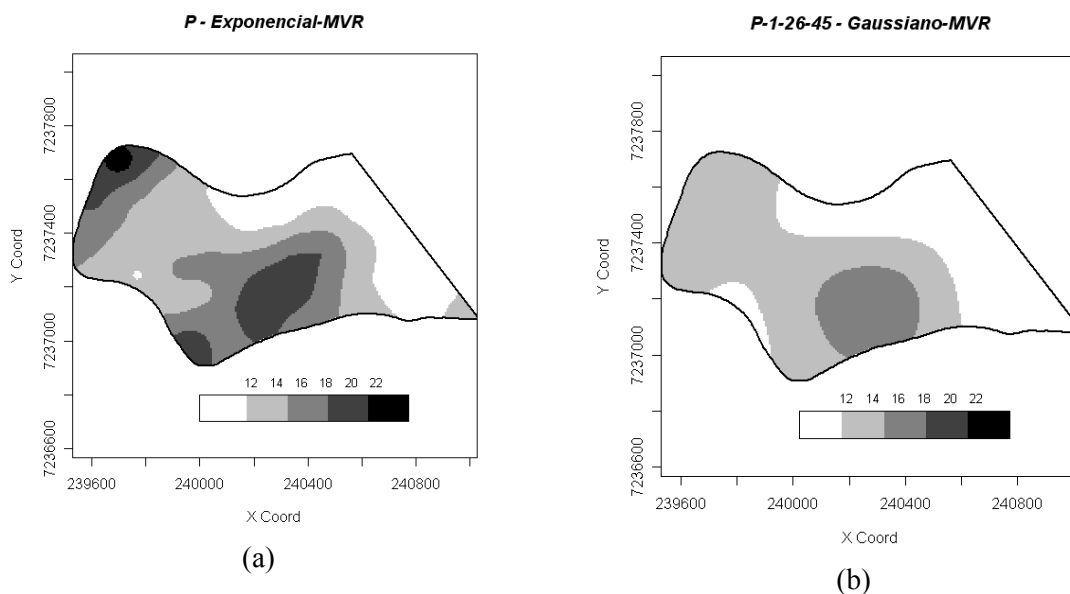


Figura 29 - Mapas temáticos para a variável fósforo: a) com dados originais (P); b) retirando-se as observações 1, 26 e 45 (P-1-26-45).

Pelas Figuras 29a e 29b, pode-se observar que há grande variação nas concentrações de fósforo entre os mapas. O mapa obtido para a variável fósforo com o conjunto original de dados (Figura 29a) mostra que a área possui regiões com concentração de fósforo acima de 18 mg dm^{-3} . Já o mapa confeccionado sem as observações consideradas influentes (Figura 29b) não apresenta região com valores acima de 18 mg dm^{-3} . Isso indica que as observações 1, 26 e 45, estão exercendo grande influência também na construção dos mapas temáticos, obtidos por técnicas geoestatísticas.

Assim, se a construção dos mapas temáticos não considerasse as análises de diagnóstico, que indicaram a presença dos valores influentes, o mapa da distribuição de fósforo que seria apresentado ao agricultor, estaria superestimando a concentração de fósforo na área em estudo. Em decorrência disso, os princípios da agricultura de precisão não estariam sendo atendidos, pois a correção do fósforo do solo deixaria de atender às exigências de forma localizada.

4.3.2 Estudo do Manganês

4.3.2.1 Análise espacial do manganês (mg dm^{-3})

As medidas do manganês, obtidas em análises químicas do solo, também foram coletadas em distâncias de $50 \times 50 \text{ m}$, $75 \times 75 \text{ m}$ e $141 \times 141 \text{ m}$. Na primeira análise laboratorial, o manganês apresentou um valor que não foi identificado como discrepante (Figura 30a), mas que, segundo técnicos, é um valor muito elevado. Esta observação é da parcela número 5 (Figura 30b). Foram repetidas as análises químicas do solo nesta parcela e constatou-se que não houve alteração dos valores. Portanto, verificou-se que a observação 5 é um valor real e não erro de medição ou de análise de laboratório, por isso, não foi descartado, pois representa uma característica local devido, provavelmente, a excesso de adubação, segundo entrevistas com os técnicos que participaram da coleta dos dados.

Na Tabela 13 são apresentadas as estatísticas descritivas para a variável teor de manganês, com todos os dados coletados.

Tabela 13 - Estatística descritiva para a variável teor de manganês (mg dm^{-3})

Variável	N	Média	Mín.	Máx.	Q ₁	Med.	Q ₃	DP	CV (%)	p-valor
Mn	46	75,43	51,00	114	64,25	74,50	84,25	13,50	17,89	> 0,05

Nota: N: número de dados; Mín.: valor mínimo; Máx.: valor máximo; Q₁: primeiro quartil; Med.: mediana; Q₃: terceiro quartil; DP: desvio-padrão; CV: coeficiente de variação; p-valor: nível descritivo do teste de normalidade de Shapiro & Wilks, ao nível de 5% de significância.

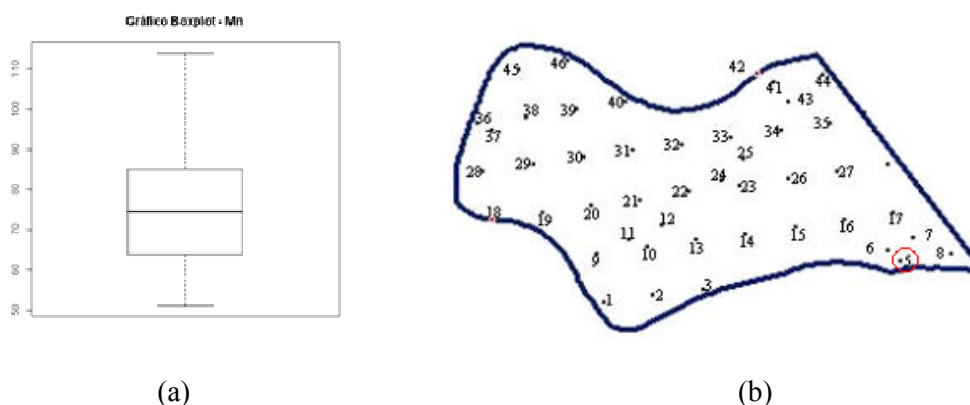


Figura 30 - Gráfico *Box-plot* da variável teor de manganês com os dados totais (a); Croqui da disposição amostral dos dados na área em estudo de 71 ha (b).

Pelos resultados da Tabela 13, verifica-se que os dados apresentam média homogeneidade ($CV = 17,89\%$), pois $10\% \leq CV \leq 20\%$ (GOMES, 2000).

Na Tabela 14 são apresentadas as análises estatísticas espaciais para a variável manganês, com todos os dados coletados (Mn), utilizando os modelos teóricos exponencial, gaussiano e Matérn com parâmetro $kappa$ igual 0,7 e 3,0, utilizando os métodos de estimação por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR).

Tabela 14 - Modelos espaciais ajustados e estimação dos parâmetros efeito pepita (φ_1), contribuição (φ_2) e alcance (φ_3), pelos métodos de máxima verossimilhança e máxima verossimilhança restrita, para a variável teor de manganês ($mg\ dm^{-3}$)

Variável	Mod..	Mét.	Média β	Efeito pepita φ_1	Contribuição φ_2	Função do alcance φ_3
Manganês	Exp.	MV	76,42 (3,321)	0,00 (59,050)	181,85 (87,763)	96,81 (0,500)
		MVR	76,8 (3,803)	0,0 (53,346)	197,6 (88,620)	113,3 (0,684)
	Gaus.	MV	76,35 (3,135)	25,87 (21,483)	168,06 (54,761)	126,92 (0,071)
		MVR	76,53 (3,336)	26,94 (20,632)	178,94 (57,473)	132,85 (0,081)
	Matérn k=0,7	MV	76,40 (3,324)	0,00 (141,676)	184,58 (92,776)	80,74 (43,383)
		MVR	76,73 (3,730)	0,00 (132,414)	199,98 (86,783)	91,17 (46,261)
	Matérn k=3,0	MV	76,32 (3,192)	15,32 (180,928)	176,01 (59,738)	37,11 (10,684)
		MVR	76,56 (3,463)	17,35 (179,360)	187,56 (61,349)	39,91 (10,908)

Nota: Alcance a : modelo exponencial: $a = 3\varphi_3$, modelo gaussiano: $a = \sqrt{3}\varphi_3$; modelo Matérn: $a = \varphi_3$.

Pelos resultados apresentados na Tabela 14, observa-se que não há grande diferença entre os valores dos parâmetros da média (β) e contribuição (φ_2) quando estimados pelos diferentes modelos. Entretanto, para o efeito pepita (φ_1), utilizando os modelos exponencial e Matérn ($k=0,7$) obteve-se valor igual a zero; Utilizando os modelos gaussiano e Matérn ($k=3,0$) os valores de φ_1 ficaram entre 15,32 e 26,94, apresentado grande alteração de valores estimados para este parâmetro. Para o parâmetro φ_3 , que é função do alcance a , também se observa grande diferença entre os valores obtidos pelos diferentes modelos. Os menores valores foram obtidos pelo modelo Matérn, com parâmetro $kappa$ igual a 0,7 e 3,0; e os maiores par o modelo gaussiano.

URIBE-OPAZO e KAVANAGH (2006) realizaram estudos acerca da modelagem da dependência espacial com dados simulados e constataram que, ao trabalhar com métodos automáticos de estimação de parâmetros, houve uma superestimação do parâmetro φ_1 , quando comparado com o valor verdadeiro. Com isso, os autores alertaram que para a modelagem da dependência espacial de dados reais, não se deve concluir a respeito de altos valores do efeito pepita φ_1 , tendo por base apenas os resultados obtidos com métodos automáticos de ajuste. É necessário, então, que o pesquisador, com sua experiência, decida se o valor obtido é razoável, pois um erro na estimativa de φ_1 tem implicações na krigagem e, conseqüentemente, na geração do mapa temático da variável em estudo.

4.3.2.2 Análises de diagnóstico

Pelos gráficos da Figura 31 e 32, em que os parâmetros foram estimados por máxima verossimilhança (MV), e Figuras 33 e 34, utilizando máxima verossimilhança restrita (MVR), pode-se observar que, em todos os modelos ajustados, tanto por MV como por MVR, as observações 5 e 6 apresentam influência local no afastamento da verossimilhança, identificadas pelos gráficos C_i e $|L_{max}|$. Quando se avaliou a influência local na matriz de covariância e nos valores preditos, pelos gráficos ICd e GL , apenas para os modelos exponencial e gaussiano os gráficos revelaram as observações 5 e 6 como influentes. Após as

análises, as observações 5 e 6 foram retiradas do conjunto total de dados para avaliar suas influências.

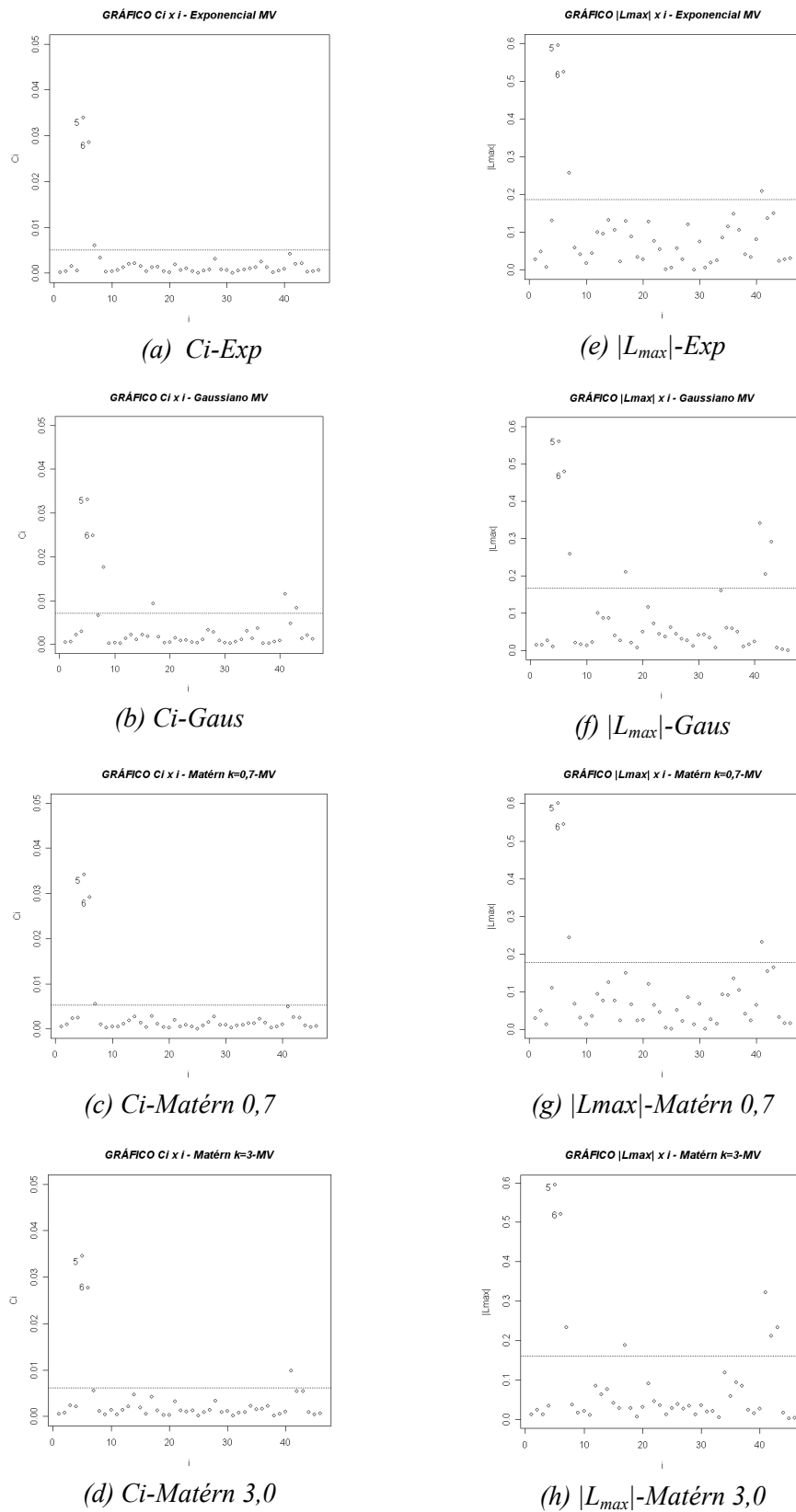
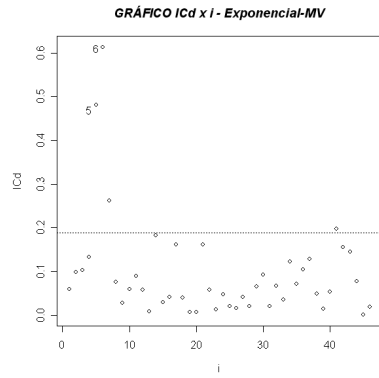
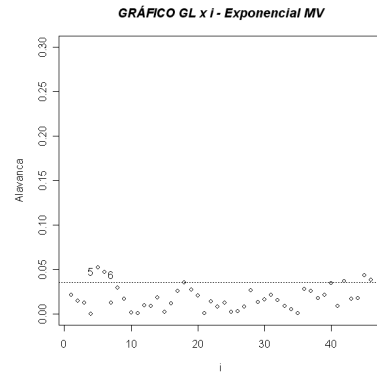


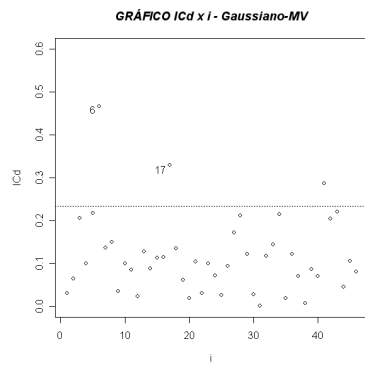
Figura 31 - Gráficos de diagnóstico $C_i \times i$ e $|L_{max}| \times i$, para a variável manganês, com o estimador MV.



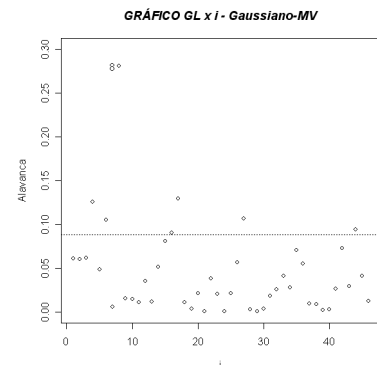
(a) *ICd-Exp*



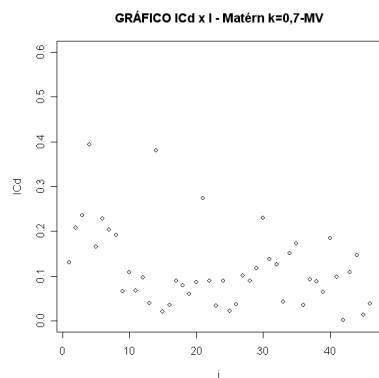
(e) *GL-Exp*



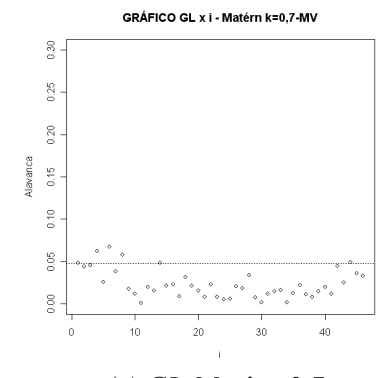
(b) *ICd-Gaus*



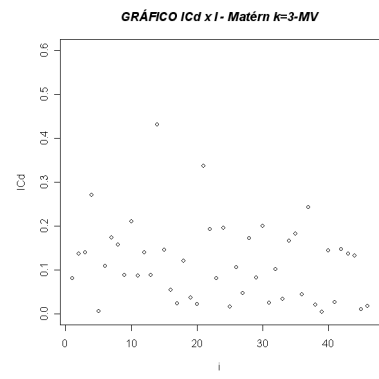
(f) *GL-Gaus*



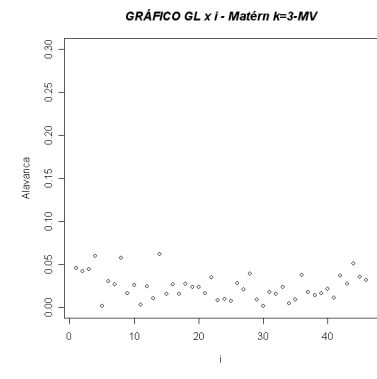
(c) *ICd-Matérn 0,7*



(g) *GL-Matérn 0,7*



(d) *ICd-Matérn 3,0*



(h) *GL-Matérn 3,0*

Figura 32 - Gráficos de diagnóstico $ICd \times i$ e $GL \times i$, para a variável manganês, com o estimador MV.

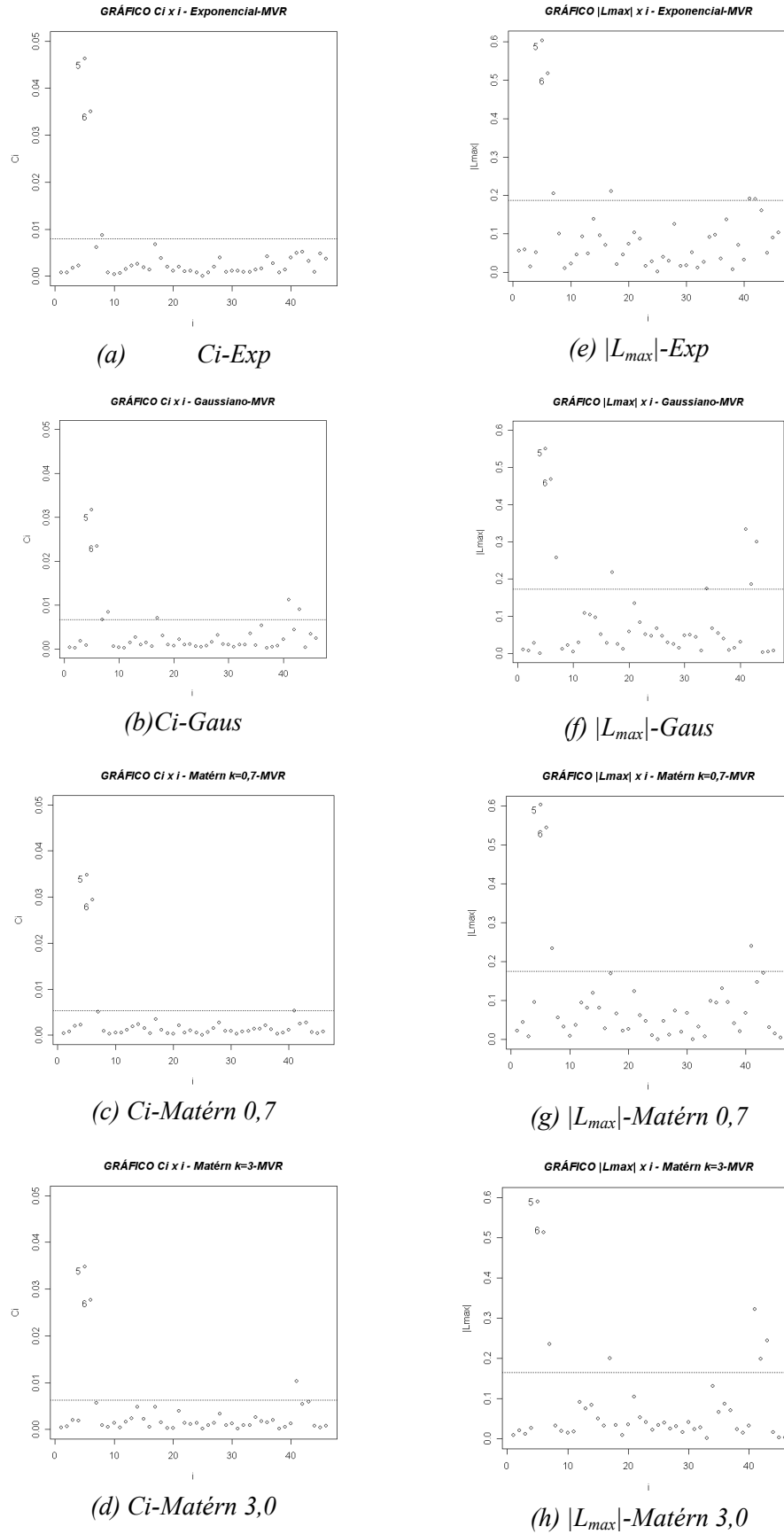


Figura 33 - Gráficos de diagnóstico $C_i \times i$ e $|L_{max}| \times i$, para a variável manganês, com o estimador MVR.

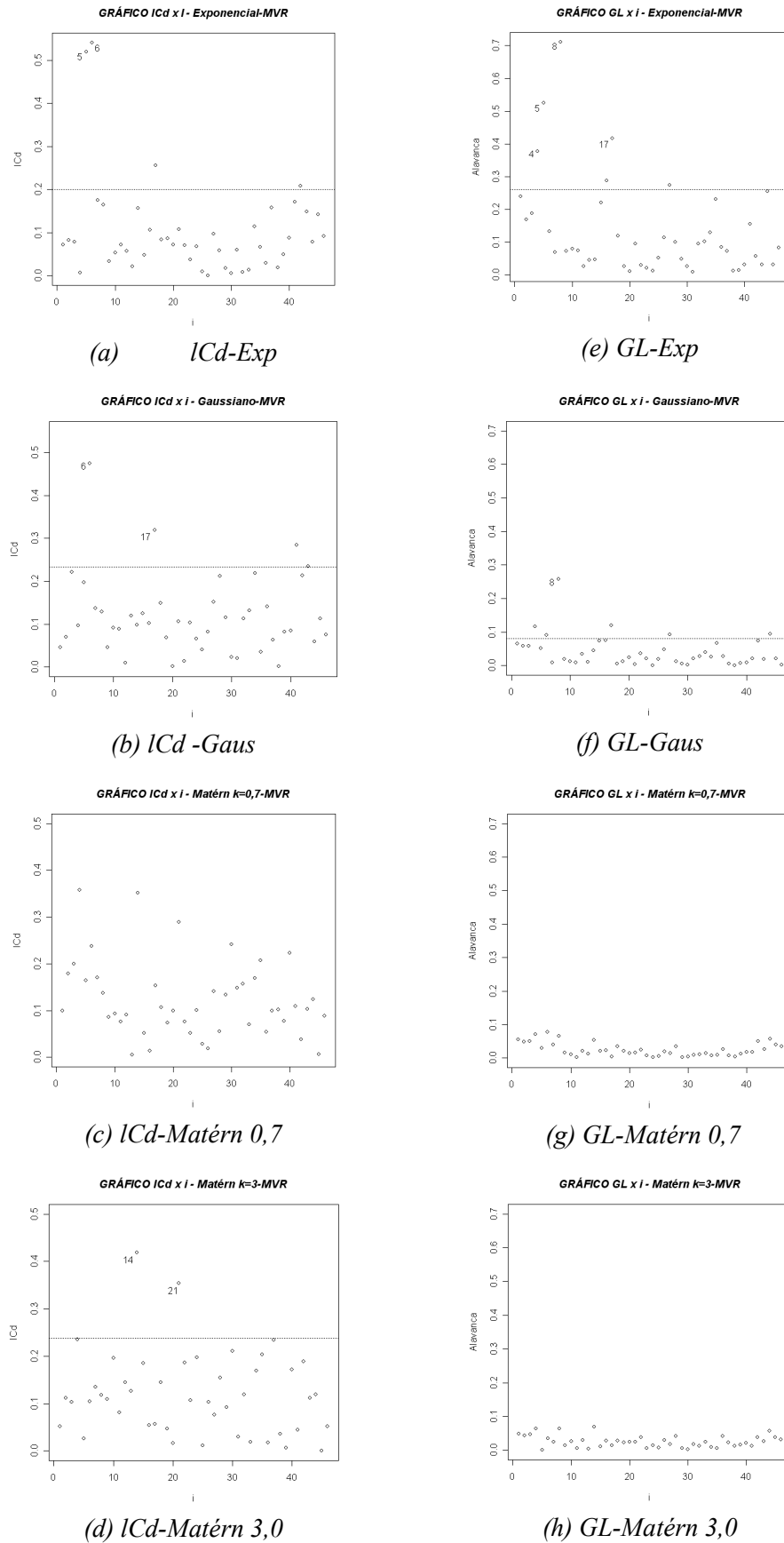


Figura 34 - Gráficos de diagnóstico $ICd \times i$ e $GL \times i$, para a variável manganês, com o estimador MVR.

4.3.2.3 Análises de influência

4.3.2.3.1 Influência nas análises descritivas

Na Tabela 15, são apresentadas as análises descritivas para o conjunto de dados com um total de 46 amostras do manganês (Mn) e dados do manganês, retirando-se as observações números 5 e 6 (Mn-5-6), verificando a mínima variação na análise descritiva.

Tabela 15 - Análise estatística descritiva para a variável manganês, com dados totais coletados (Mn) e retirando-se as observações 5 e 6 (Mn-5-6)

Variável	N	Média	Mín.	Máx.	Q1	Med.	Q3	DP	CV (%)	p-valor
Mn	46	75,4	51,0	114	64,3	74,5	84,3	13,5	17,9	>0,05*
Mn-5-6	44	74,2	51,0	103	63,8	73,0	82,0	12,2	16,4	>0,05*

Nota: N: número de dados; Med.: mediana; DP: desvio-padrão; CV: coeficiente de variação; p-valor: nível descritivo do teste de normalidade de Shapiro & Wilks, ao nível de 5% de significância.

Pela Tabela 15, observa-se que há pouca diferença nos valores médios, nos desvios padrão e nos CVs, entre os dados da variável manganês com todos os dados, retirando observações influentes.

4.3.2.3.2 Influência nas estimativas dos parâmetros

Na Tabela 16 são apresentados os resultados dos ajustes de modelos teóricos para modelagem da variabilidade espacial da variável manganês, por meio de métodos geoestatísticos, para os dados totais e retirando-se os dados considerados influentes. São apresentados os resultados da estimação dos parâmetros média (β), efeito pepita (ϕ_1), contribuição (ϕ_2) e alcance (ϕ_3), pelos modelos exponencial, gaussiano e Matérn ($k=0,7$ e $k=3,0$), por máxima

verossimilhança (MV) e máxima verossimilhança restrita (MVR). Apresentam-se também, entre parêntesis, os desvios padrão de cada parâmetro estimado.

Tabela 16 - Ajustes utilizando os modelos teóricos exponencial, gaussiano e Matérn, com parâmetros estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR) para a variável manganês, com dados totais e retirando-se as observações 5 e 6 (Sem influentes)

Modelo	Método	Dados	Média β	Efeito pepita φ_1	Contribuição φ_2	Função do alcance φ_3	
Exp.	MV	Dados totais	76,42 (3,321)	0,00 (59,050)	181,85 (87,763)	96,81 (0,500)	
		Sem influentes	75,48 (3,209)	0,00 (46,921)	150,86 (72,405)	105,81 (0,644)	
	MVR	Dados totais	76,80 (3,803)	0,00 (53,345)	197,60 (88,619)	113,30 (0,684)	
		Sem influentes	75,79 (3,646)	0,00 (43,109)	163,93 (73,965)	122,26 (0,860)	
	Gaus.	MV	Dados totais	76,35 (3,135)	25,87 (21,483)	168,06 (54,761)	126,92 (0,071)
			Sem influentes	75,26 (2,851)	3,29 (10,243)	153,78 (41,241)	122,93 (0,068)
MVR		Dados totais	76,53 (3,336)	26,94 (20,632)	178,94 (57,473)	132,85 (0,081)	
		Sem influentes	75,35 (2,970)	3,99 (10,111)	160,58 (42,993)	126,12 (0,073)	
Matérn k=0,7		MV	Dados totais	76,40 (3,324)	0,00 (141,676)	184,58 (92,776)	80,74 (43,383)
			Sem influentes	75,57 (3,280)	0,00 (108,017)	154,04 (70,730)	90,67 (47,409)
	MVR	Dados totais	76,73 (3,730)	0,00 (132,414)	199,98 (86,783)	91,17 (46,261)	
		Sem influentes	75,83 (3,658)	0,00 (103,647)	167,19 (68,756)	101,23 (50,788)	
	Matérn k=3,0	MV	Dados totais	76,32 (3,192)	15,32 (180,928)	176,01 (59,738)	37,11 (10,684)
			Sem influentes	75,50 (3,158)	0,00 (126,204)	161,11 (45,633)	39,92 (9,255)
MVR		Dados totais	76,56 (3,463)	17,35 (179,360)	187,56 (61,349)	39,91 (10,908)	
		Sem influentes	75,60 (3,332)	0,00 (127,772)	170,95 (48,203)	41,36 (9,254)	

Analisando os resultados do ajuste de modelos para os dados da variável manganês, na Tabela 16, verifica-se que as observações 5 e 6 foram pouco influentes na estimação dos parâmetros que definem a estrutura de dependência espacial. Verifica-se que o parâmetro que estima a média (β) teve seu valor reduzido, quando se comparam ajustes para dados totais e sem os dados influentes. O mesmo comportamento foi observado para o parâmetro contribuição (φ_2), que mede a variância de dispersão dos dados.

Quanto à influência no parâmetro φ_3 , que é função do alcance (a) do modelo, em todos os modelos ajustados houve variação nos valores estimados com dados totais e sem os dados influentes. Entretanto, para os modelos exponencial, Matérn ($k = 0,7$) e Matérn ($k = 3,0$), com parâmetros estimados por MV ou MVR, houve aumento nos valores de φ_3 , quando retiradas as observações influentes. Apenas nos ajustes com o modelo gaussiano verificou-se comportamento oposto.

Finalmente, as observações 5 e 6 também foram influentes na estimação do efeito pepita (φ_1), que é o parâmetro que mais contribui na construção de mapas temáticos (URIBE-OPAZO; KAVANAGH, 2006), mas isso somente ocorreu nos ajustes com os modelos gaussiano e Matérn ($k = 3,0$). Nos ajustes com os modelos exponencial e Matérn ($k = 0,7$), com parâmetros também estimados por MV e MVR, os valores do efeito pepita não tiveram alteração ($\varphi_1 = 0,00$).

4.3.2.3.3 Influência na validação dos modelos

Na Tabela 17 são apresentados os resultados dos critérios de Akaike (\hat{A}), Filliben (r^*) e valor máximo do logaritmo da função verossimilhança (VMV), para validação de ajustes dos modelos geoestatísticos, para a variável manganês com dados totais (Mn) e sem os dados influentes (Mn-5-6).

Tabela 17 - Resultados da validação dos modelos ajustados para a variável teor de manganês, pelos critérios de Akaike (\hat{A}), Filliben (r^*) e valor máximo do logaritmo da função verossimilhança (VMV)

	Modelo	Mét.	Nº dados	\hat{A}	r^*	VMV
Dados totais	Exp.	MV	46	-2742,028	0,981	-181,2
		MVR		-2711,072	0,981	-177,1
	Gaus.	MV	46	-2779,331	0,977	-180,1
		MVR		-2770,562	0,977	-176,1
	Matérn k=0,7	MV	46	-2799,860	0,978	-180,9
		MVR		-2748,375	0,978	-176,8
Matérn k=3,0	MV	46	-2863,629	0,976	-180,3	
	MVR		-2826,327	0,977	-176,3	
Sem dados influentes	Exp.	MV	44	-2689,617	0,994	-168,8
		MVR		-2703,182	0,993	-164,8
	Gaus.	MV	44	-2649,842	0,994	-166,1
		MVR		-2764,179	0,994	-162,2
	Matérn k=0,7	MV	44	-2635,141	0,993	-168,2
		MVR		-2719,226	0,993	-164,1
Matérn k=3,0	MV	44	-2642,185	0,992	-166,3	
	MVR		-2658,229	0,992	-162,3	

Nota: r^* : teste de Filliben, a 5% de significância, compara-se com $r = 0,9740$ (Dados originais) e $r = 0,9730$ (Sem dados influentes).

Pelo critério de Akaike, o menor valor de \hat{A} (-2863.629), para o conjunto original de dados, aponta como melhor ajuste o modelo Matérn, $kappa$ igual a 3,0, com parâmetros estimados poro MV. Entretanto, quando foram retiradas as observações influentes, o menor valor de \hat{A} (-2764,179) indicou como melhor ajuste o modelo gaussiano, com parâmetros estimados pelo método MVR.

Para a aplicação do critério de Filliben, ao nível de 5 % de significância, sabe-se que os dados da variável manganês com dados originais e sem dados influentes, têm-se 46 e 44 valores amostrados, respectivamente. Os valores tabelados de r , obtidos da tabela de FILLIBEN (1975), foram de 0,9740 e 0,9730, respectivamente. Comparando-se com os valores r^* calculados (Tabela 17), para dados originais do manganês e sem dados influentes, verificou-se que, em todos os modelos ajustados, os valores r^* são maiores do que r tabelado. Isto indica que a hipótese de que os dados seguem distribuição normal foi aceita, para um nível de 5% de significância.

Pelo valor máximo do logaritmo da função verossimilhança (VMV), tanto para manganês com dados originais, quanto sem os valores influentes, o melhor modelo ajustado foi o gaussiano, com parâmetros estimados por MVR.

Na Tabela 18, são apresentados os resultados da validação de modelos ajustados, utilizando o critério da validação cruzada para a variável manganês com dados originais e sem as observações influentes.

Tabela 18 - Resultados da validação cruzada para a variável teor de manganês com dados totais e sem dados influentes

	Modelo	Mét.	EM	\overline{ER}	DPe	S_{ER}	EA	
Dados totais	Exp.	MV	-0,1809	-0,0085	11,3828	0,9661	420,0082	
		MVR	-0,1818	-0,0086	11,3161	0,9618	416,1526	
	Gaus.	MV	-0,1637	-0,0097	10,6322	0,9804	390,9073	
		MVR	-0,1316	-0,0080	10,5965	0,9784	393,2419	
	Matérn k=0,7	MV	-0,2284	-0,0112	11,1444	0,9652	408,9312	
		MVR	-0,2276	-0,0114	11,0799	0,9616	404,6158	
	Matérn k=3,0	MV	-0,2542	-0,0143	10,7444	0,9761	389,3410	
		MVR	-0,2253	-0,0129	10,6968	0,9734	390,8827	
	Sem dados influentes	Exp.	MV	0,0184	0,0003	10,492	0,9684	371,1455
			MVR	0,0251	0,0005	10,4531	0,9630	366,8021
Gaus.		MV	0,0864	-0,0005	9,6907	1,0031	345,7955	
		MVR	0,0949	0,0001	9,6717	0,9999	346,1121	
Matérn k=0,7		MV	0,0039	-0,0009	10,2609	0,9651	358,4471	
		MVR	0,0074	-0,0008	10,2326	0,9605	356,0652	
Matérn k=3,0		MV	0,0247	-0,0034	9,6964	0,9814	342,9460	
		MVR	0,0318	-0,0032	9,6869	0,9805	343,1075	

Pelos resultados da Tabela 18, para a variável manganês com o conjunto original de dados e sem dados influentes, os critérios da validação cruzada indicaram como melhor ajuste o modelo gaussiano, com parâmetros estimados por MVR.

Sendo assim, pelo critério de Akaike, as observações 5 e 6 foram influentes na escolha do melhor modelo espacial ajustado, mas essa influência não foi registrada para o critério da validação cruzada, nem para o valor máximo do logaritmo da função verossimilhança (VMV). Portanto, a validação cruzada e o critério VMV foram mais robustos na escolha do melhor modelo geoestatístico ajustado, porquanto não sofreram a influência dos pontos 5 e 6.

4.3.2.3.4 Influência na construção dos mapas temáticos

Na Figura 35 são apresentados os mapas temáticos da variável teor de manganês, com dados originais e sem dados considerados influentes, confeccionados por meio da interpolação por krigagem ordinária. Os mapas foram construídos utilizando-se os modelos indicados pelos critérios da validação cruzada e *VMV*, ou seja, modelo gaussiano, com parâmetros estimados por MVR.

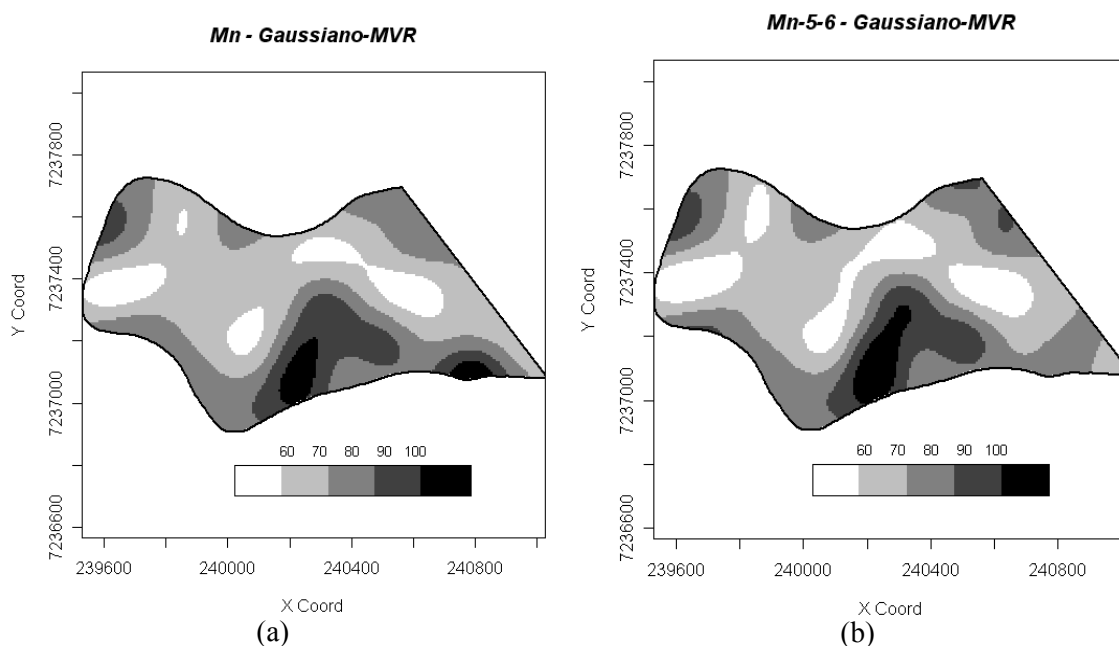


Figura 35 - Mapas temáticos para a variável teor de manganês: a) com dados originais (Mn); b) retirando-se as observações 5 e 6 (Mn-5-6).

Ao compararmos os mapas temáticos da Figura 35, observa-se que a distribuição espacial da variável manganês, com e sem as observações 5 e 6, figuras 35a e 35b, respectivamente, é semelhante. Entretanto, em uma análise mais cuidadosa, verificam-se algumas regiões nas quais os teores de manganês no solo (mg dm^{-3}) são diferentes.

Isso indica que as observações 5 e 6 exercem influência na construção dos mapas temáticos, obtidos por técnicas geoestatísticas, o que prejudicaria a correção do manganês do solo de forma localizada.

4.3.3 Estudo da Produtividade da Soja como Função de Covariáveis

No ano agrícola de 2004/2005 foram coletados dados da resistência do solo à penetração (RSP) [MPa], densidade do solo (DS) [Mg m^{-3}] e produtividade da soja (t ha^{-1}), em uma área agrícola comercial de 57 ha, segundo a amostragem apresentada na Figura 36a.

O principal interesse foi estudar a variabilidade espacial da produtividade da soja na área em estudo, mas buscando aproveitar que as variáveis densidade do solo e resistência do solo à penetração, nas camadas 0,0 a 0,1 m, 0,1 a 0,2 m e 0,2 a 0,3 m de profundidade, foram coletadas nos mesmos pontos onde foi medida a produtividade, e foram utilizadas como covariáveis, ou seja, como variáveis explicativas da produtividade.

JOHANN *et al.* (2004) concluíram que, dentre os atributos físicos dos solos estudados, a resistência do solo à penetração, na camada 0,0 a 0,1 m, foi a variável que melhor correlacionou-se com a produtividade.

Como há diferença entre o número de amostras coletadas com produtividade e as covariáveis, foi necessário realizar uma filtragem para compatibilizar aquelas que possuíam a mesma posição, fazendo com que as coordenadas e os dados da variável produtividade coincidissem com as covariáveis.

Assim, definiu-se:

$$Z(s_i) = \mu(s_i) + \varepsilon(s_i)$$

com $E[\varepsilon(s_i)] = 0$ e $C(s_i, s_u) = \text{Cov}\{\varepsilon(s_i), \varepsilon(s_u)\}$.

Assim, a média do processo estocástico será dada por:

$$\mu(s) = \beta_0 + \beta_1 \text{RSP}_{0,0-0,1} + \beta_2 \text{RSP}_{0,1-0,2} + \beta_3 \text{RSP}_{0,2-0,3} + \beta_4 \text{DS}_{0,0-0,1} + \beta_5 \text{DS}_{0,1-0,2} + \beta_6 \text{DS}_{0,2-0,3}$$

em que β_0, \dots, β_6 são parâmetros desconhecidos e a serem estimados.

4.3.3.1 Análises preliminares

Na Tabela 19 são apresentadas as estatísticas descritivas para a variável produtividade.

Tabela 19 - Estatística descritiva para a variável produtividade ($t\ ha^{-1}$)

Variável	N	Média	Mín.	Máx.	Q1	Mediana	Q3	DP	CV (%)	p-valor
Prod	47	3,23	2,09	4,09	3,01	3,19	3,49	0,38	11,79	> 0,05*

Nota: N: número de dados; Mín.: valor mínimo; Máx.: valor máximo; Q1: primeiro quartil; Q3: terceiro quartil; DP: desvio-padrão; CV: coeficiente de variação; p-valor: nível descritivo do teste de normalidade de Shapiro & Wilks, ao nível de 5% de significância.

Pelos resultados da Tabela 19, verifica-se que a produção média da soja foi de $3,23\ t\ ha^{-1}$, com desvio padrão de $0,38\ t\ ha^{-1}$. Pelo valor do coeficiente de variação ($CV = 11,79\%$), os dados foram classificados como média homogeneidade (GOMES, 2000).

Pelo gráfico *Box-plot* (Figura 36b), observa-se a presença de um valor discrepante, com valor igual a $2,09\ t\ ha^{-1}$, sendo o 13º valor da série de dados.

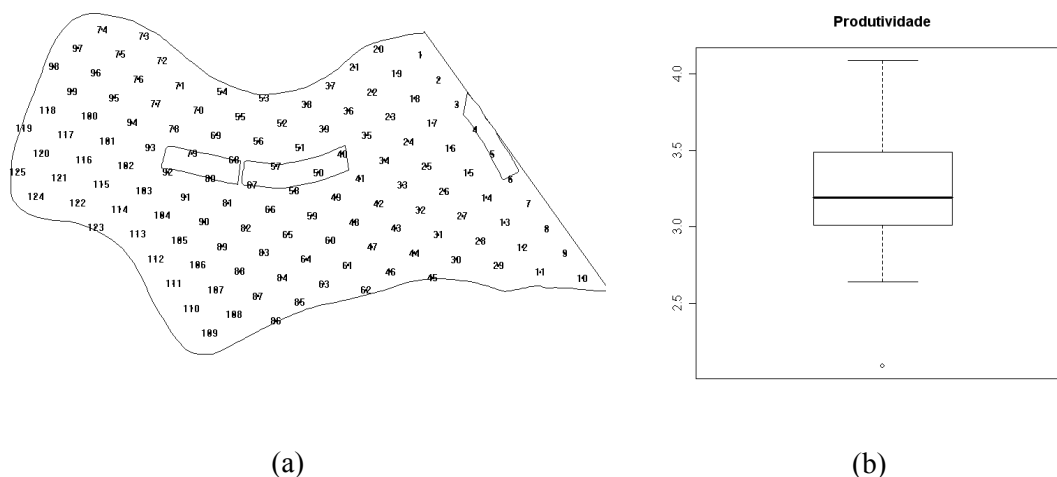


Figura 36 - Croqui da disposição amostral dos dados na área em estudo de 57 ha (a); Gráfico *Box-plot* da variável produtividade com os dados totais (b).

4.3.3.2 Análises de diagnóstico

Foram aplicadas técnicas gráficas de diagnóstico com a finalidade de avaliar se o ponto discrepante (13º valor) ou alguns outros estariam exercendo algum tipo de influência no afastamento da verossimilhança, nas análises da estrutura de dependência espacial, tais como na função de covariância e escolha do modelo a ser ajustado, e no preditor linear.

Por simplicidade, doravante, quando foi citada a variável produtividade, subentende-se que as análises incluem as covariáveis resistência do solo à penetração e densidade do solo.

A primeira análise de diagnóstico foi realizada para a variável produtividade com o total de 47 dados coletados (Prod.), utilizando os modelos espaciais ajustados exponencial, gaussiano e Matérn, com parâmetro $kappa$ igual a 0,7 e 3,0.

Nas figuras 37 e 38 são apresentados os gráficos de diagnóstico, utilizando os modelos espaciais exponencial, gaussiano e Matérn ($kappa$ igual a 0,7 e 3,0) com o estimador de máxima verossimilhança (MV), e nas figuras 39 e 40, os mesmos gráficos e modelos, mas utilizando o estimador de máxima verossimilhança restrita (MVR).

Pelos gráficos, pode-se verificar que as observações números 13, 23 e 30 foram as que mais se destacaram como influentes, pois apresentaram altos valores para a maioria das medidas de diagnóstico da influência local no afastamento da verossimilhança, identificadas pelos gráficos C_i e $|L_{max}|$, na matriz de covariância e nos valores preditos, pelos gráficos ICd e GL , respectivamente. Então, estas observações foram retiradas do conjunto total de dados e novamente realizadas as análises para avaliar suas influências.

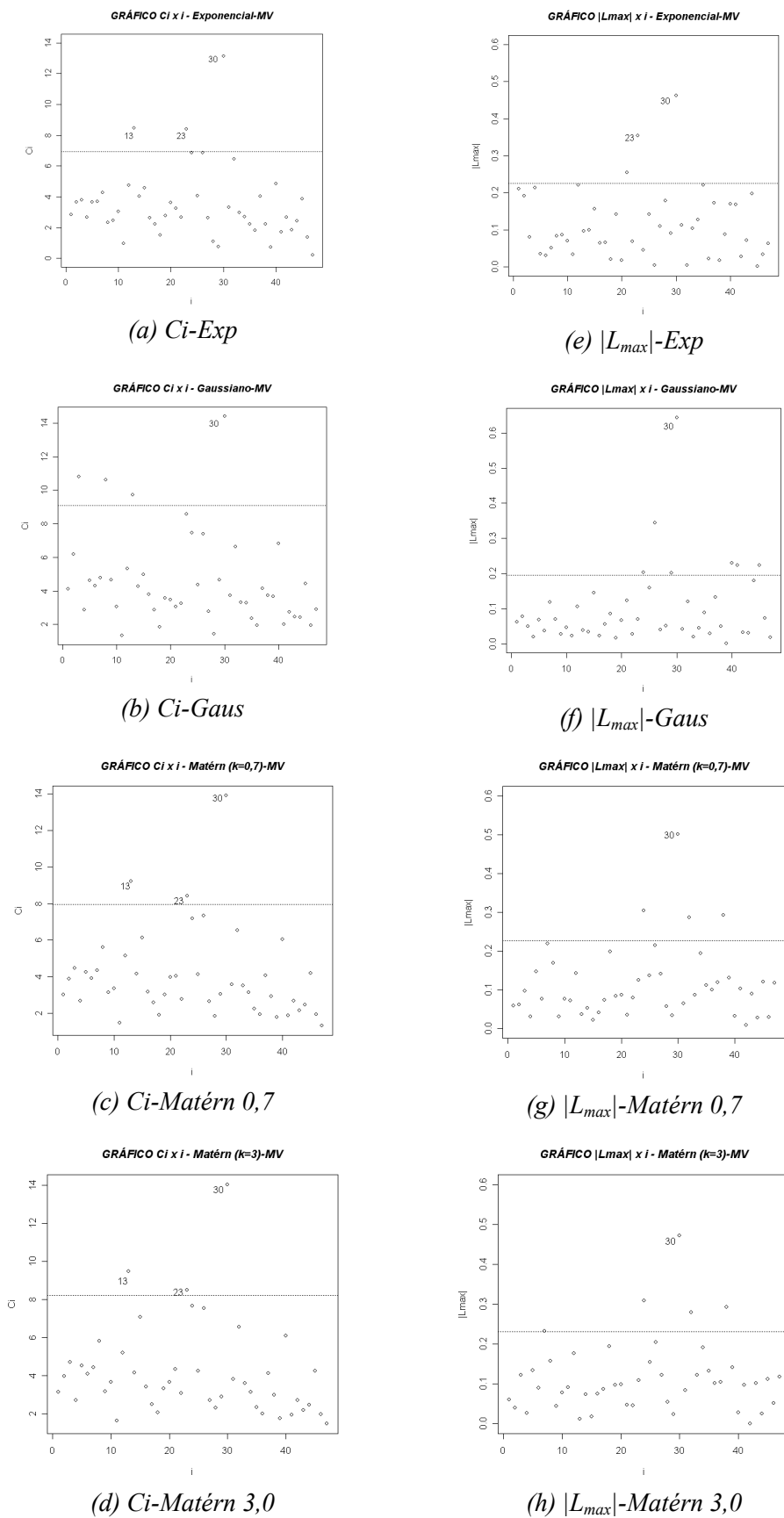


Figura 37 - Gráficos de diagnóstico C_i x i e $|L_{max}|$ x i , para a variável produtividade, com o estimador MV.

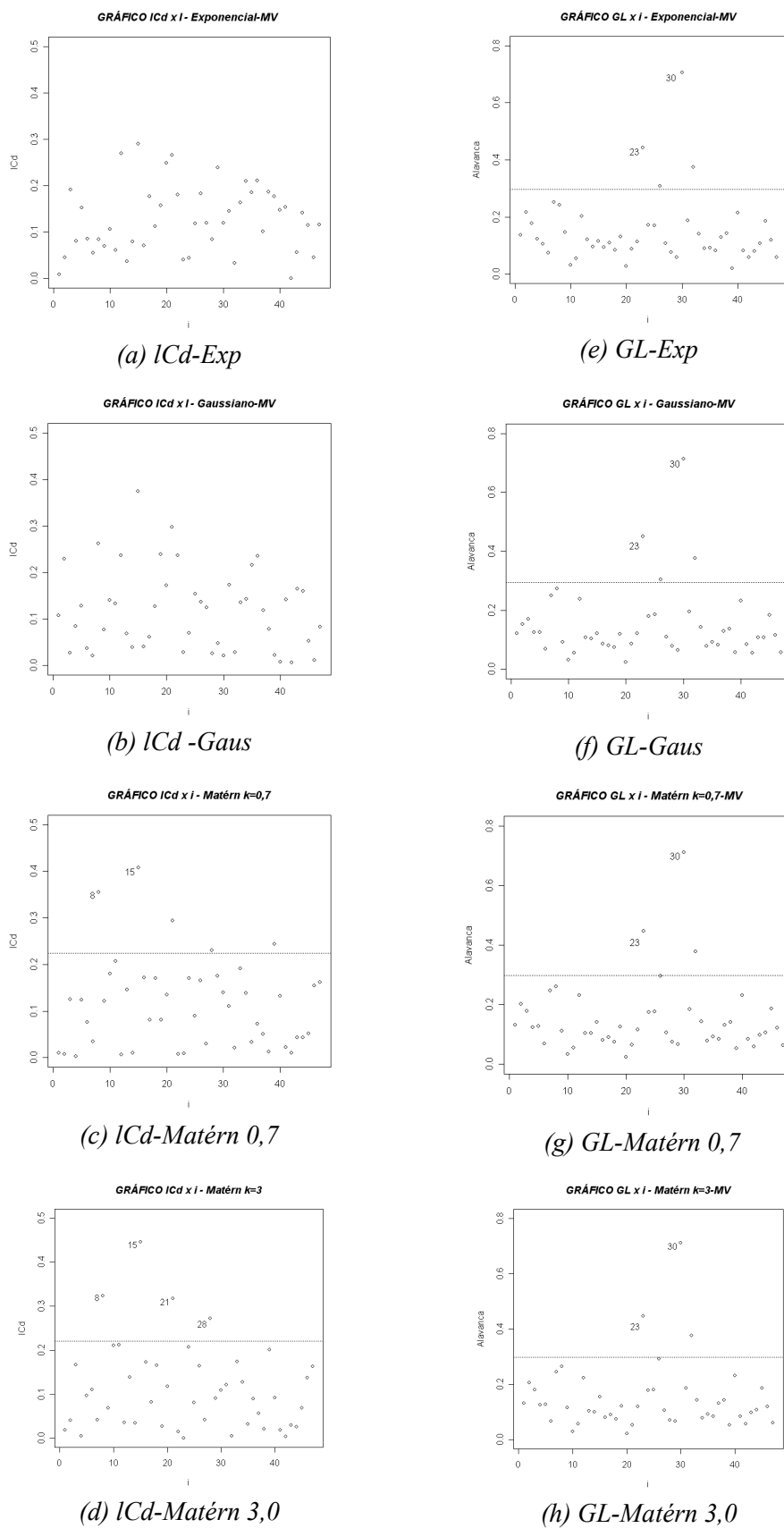


Figura 38 - Gráficos de diagnóstico ICd x i e GL x i , para a variável produtividade, com o estimador MV.

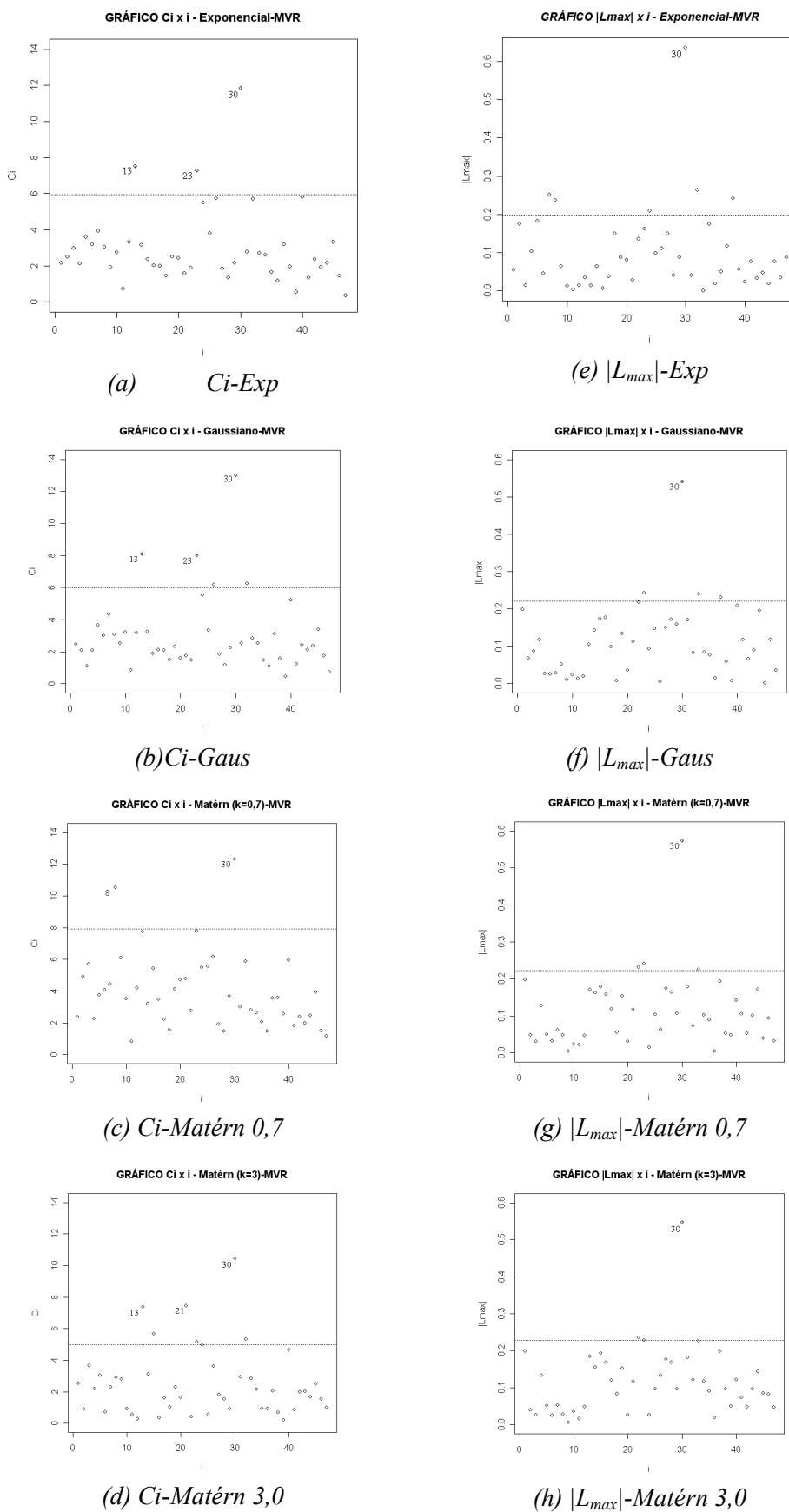


Figura 39 - Gráficos de diagnóstico C_i x i e $|L_{max}|$ x i , para a variável produtividade, com o estimador MVR.

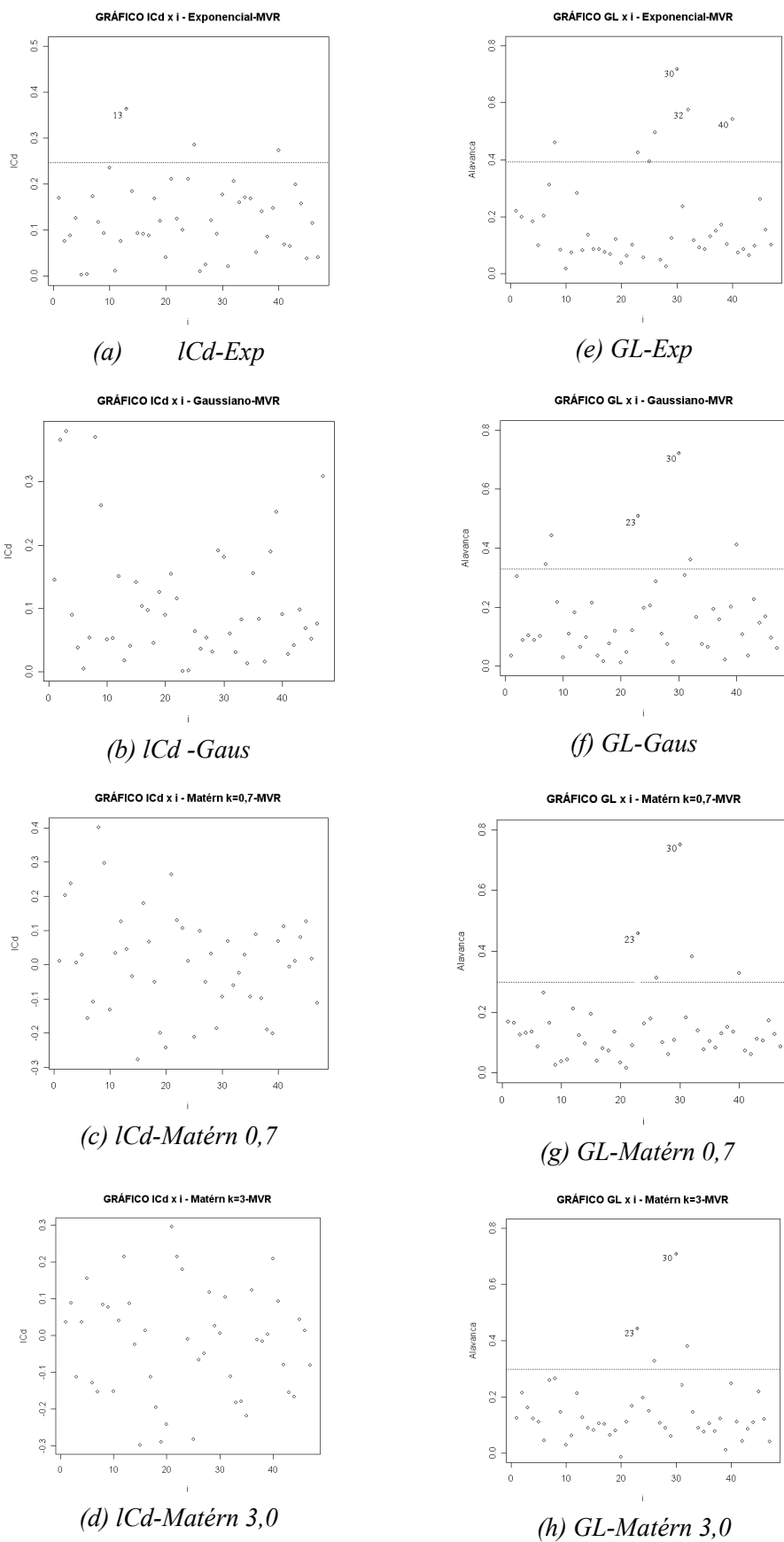


Figura 40 - Gráficos de diagnóstico $ICd \times i$ e $GL \times i$, para a variável produtividade, com o estimador MVR.

4.3.3.3 Análises de influência

4.3.3.3.1 Influência nas análises descritivas

Na Tabela 20, são apresentadas as análises descritivas para o conjunto de dados com um total de 47 amostras da produtividade (Prod.) e dados da produtividade, retirando-se as observações números 13, 23 e 30 (Prod-13-23-30).

Tabela 20 - Análise estatística descritiva para a variável produtividade, com dados totais coletados (Prod.) e retirando-se as observações 1, 26 e 45 (Prod-13-23-30)

Variável	N	Média	Mín.	Máx.	Q1	Med.	Q 3	DP	CV (%)	p-valor
Prod.	47	3,23	2,09	4,09	3,01	3,19	3,49	0,38	11,79	> 0,05*
Prod.-13-23-30	44	3,27	2,64	4,09	3,02	3,23	3,50	0,12	10,58	> 0,05*

Nota: N: número de dados; Med.: mediana; DP: desvio-padrão; CV: coeficiente de variação; p-valor: nível descritivo do teste de normalidade de Shapiro & Wilks ao nível de 5% de significância.

Pela Tabela 20, observa-se que há pouca diferença entre as médias da variável produtividade com todos os dados, e retirando as observações influentes. Sem dados influentes, o valor da média elevou-se. No desvio padrão, sem as observações influentes, este valor diminuiu. Ao analisar o coeficiente de variação (CV) da produtividade, verifica-se que, sem as observações 13, 23 e 30, o valor do CV também foi menor para os dados totais, considerados homogêneos (GOMES, 2000).

4.3.3.3.2 Influência nas estimativas dos parâmetros

Nas tabelas 21a e 21b são apresentados os resultados das análises de variabilidade espacial da produtividade da soja, por meio da geoestatística, para

os dados originais e retirando-se os dados considerados influentes. São apresentados os resultados da estimação dos parâmetros $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$, efeito pepita (φ_1), contribuição (φ_2) e alcance (φ_3), pelos modelos exponencial, gaussiano e Matérn, por máxima verossimilhança e máxima verossimilhança restrita. Apresentam-se também, entre parêntesis, os desvios padrão para cada parâmetro estimado.

Tabela 21a - Parâmetros β estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR) para a variável produtividade, com dados totais (Prod.) e retirando-se as observações 13, 23 e 30 (P-13-23-30), utilizando os modelos teóricos exponencial, gaussiano e Matérn

Modelo	Mét.	Dad.*	β_0	β_1	β_2	β_3	β_4	β_5	β_6	
Exp.	MV	DO	4,5641 (1,0738)	-0,1027 (0,1153)	0,0633 (0,1231)	-0,0150 (0,0919)	-0,2885 (0,6392)	0,0611 (0,5224)	-0,8662 (0,7995)	
		SDI	3,4192 (1,0446)	0,0893 (0,1129)	-0,1449 (0,1263)	0,2129 (0,1551)	-0,1401 (0,5612)	-0,1401 (0,6057)	-1,0294 (0,716)	
	MVR	DO	4,7484 (1,1682)	-0,1133 (0,1240)	0,0519 (0,1335)	-0,0175 (0,0994)	-0,2945 (0,689)	-0,0016 (0,5591)	-0,9168 (0,8590)	
		SDI	3,4047 (1,1583)	0,0884 (0,1228)	-0,1512 (0,1373)	0,2065 (0,1709)	-0,1014 (0,6085)	0,6739 (0,6564)	-1,0371 (0,7716)	
	Gaus.	MV	DO	4,4487 (1,0660)	-0,0965 (0,1137)	0,0612 (0,1224)	-0,0176 (0,0904)	-0,2633 (0,6361)	0,1143 (0,5195)	-0,8467 (0,7884)
			SDI	3,3777 (1,0345)	0,0916 (0,1111)	-0,1557 (0,1252)	0,2094 (0,1529)	-0,1107 (0,5548)	0,6922 (0,5934)	-1,0048 (0,7023)
MVR		DO	4,5134 (1,1538)	-0,1009 (0,1229)	0,0538 (0,1330)	-0,0184 (0,0981)	-0,2558 (0,6890)	0,0960 (0,5613)	-0,8678 (0,8533)	
		SDI	3,4134 (1,1349)	0,0877 (0,1212)	-0,1616 (0,1362)	0,2058 (0,1683)	-0,0860 (0,6016)	0,6678 (0,6424)	-1,0169 (0,7620)	
Matérn k=0,7	MV	DO	4,5615 (1,0728)	-0,1026 (0,1152)	0,0622 (0,1230)	-0,0152 (0,0918)	-0,2864 (0,6389)	0,0633 (0,5220)	-0,8660 (0,7984)	
		SDI	3,4141 (1,0440)	0,0895 (0,1127)	-0,1463 (0,1261)	0,2120 (0,1549)	-0,1349 (0,5603)	0,6837 (0,6040)	-1,0257 (0,7142)	
	MVR	DO	4,7920 (1,1685)	-0,1165 (0,1239)	0,0508 (0,1336)	-0,0178 (0,0995)	-0,2982 (0,6895)	-0,0161 (0,5575)	-0,9283 (0,8586)	
		SDI	3,4044 (1,1556)	0,0882 (0,1226)	-0,1526 (0,1372)	0,2062 (0,1705)	-0,0967 (0,6075)	0,6711 (0,6541)	-1,0338 (0,7696)	
Matérn k=3,0	MV	DO	4,5184 (1,0679)	-0,1003 (0,1144)	0,0597 (0,1227)	-0,0163 (0,0910)	-0,2750 (0,6371)	0,0855 (0,5200)	-0,8592 (0,7927)	
		SDI	3,3931 (1,0386)	0,0903 (0,1116)	-0,1524 (0,1254)	0,2096 (0,1537)	-0,1163 (0,5564)	0,6838 (0,5967)	-1,0108 (0,7058)	
	MVR	DO	4,7316 (1,1540)	-0,1137 (0,1236)	0,0460 (0,1338)	-0,0171 (0,0991)	-0,2791 (0,6892)	0,0135 (0,5572)	-0,9156 (0,8581)	
		SDI	3,4076 (1,1428)	0,0876 (0,1217)	-0,1584 (0,1366)	0,2056 (0,1691)	-0,0855 (0,6039)	0,6649 (0,6461)	-1,0200 (0,7640)	

Nota: *Dad.: dados, em que: DO: dados originais; SDI: sem dados influentes.

Tabela 221b - Parâmetros φ estimados por máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR) para a variável produtividade, com dados totais (Prod) e retirando-se as observações 13, 23 e 30 (P-13-23-30), utilizando os modelos teóricos exponencial, gaussiano e Matérn

Modelo	Mét.	Dad.*	φ_1	φ_2	φ_3	
Exp	MV	DO	0,0791 (0,0324)	0,0455 (0,0404)	142,2595 (1,7333)	
		SDI	0,0611 (0,0246)	0,0368 (0,0323)	168,2458 (2,4582)	
	MVR	DO	0,1061 (0,0297)	0,0648 (0,0568)	475,0023 (16,6251)	
		SDI	0,0768 (0,0249)	0,0538 (0,0443)	342,8133 (9,0962)	
	Gaus	MV	DO	0,0909 (0,0236)	0,0333 (0,0269)	192,3540 (0,5085)
			SDI	0,0701 (0,0180)	0,0280 (0,0226)	239,4322 (0,9292)
MVR		DO	0,1084 (0,0270)	0,0401 (0,0322)	220,3495 (0,7236)	
		SDI	0,0842 (0,0208)	0,0376 (0,0298)	293,4564 (1,5802)	
Matérn k=0,7		MV	DO	0,0844 (0,0708)	0,0403 (0,0435)	126,7328 (156,4998)
			SDI	0,0650 (0,0526)	0,0331 (0,0324)	151,0009 (180,8037)
	MVR	DO	0,1108 (0,0727)	0,0653 (0,0581)	430,0033 (519,4543)	
		SDI	0,0801 (0,0573)	0,0495 (0,0398)	272,4014 (303,1239)	
	Matérn k=3,0	MV	DO	0,0913 (0,1006)	0,0332 (0,0300)	61,3444 (51,0390)
			SDI	0,0700 (0,0785)	0,0284 (0,0240)	75,4043 (59,1810)
MVR		DO	0,1130 (0,1129)	0,0422 (0,0359)	100,0084 (78,7116)	
		SDI	0,0841 (0,0983)	0,0405 (0,0324)	99,2289 (71,6130)	

Pelos resultados da Tabela 21a, pode-se observar que para todos os modelos ajustados, o parâmetro β_0 teve seu valor reduzido quando retiradas as observações influentes, ficando mais próximo da média amostral dos dados. Os parâmetros de β_1 a β_6 tiveram seus valores invertidos, ora sendo maiores, ora sendo menores, quando retiradas as observações influentes. Observando a Tabela 21b, verifica-se que os parâmetros φ_1 e φ_2 , para todos os modelos ajustados, quando estimados sem as observações influentes, tiveram seus valores reduzidos e o parâmetro φ_3 , quando retiradas as observações consideradas influentes e estimado por máxima verossimilhança, teve seu valor elevado. O contrário ocorreu quando as estimativas foram realizadas por máxima

verossimilhança restrita, em que os valores do parâmetro φ_3 diminuíram, exceto para o ajuste com o modelo gaussiano.

4.3.3.3.3 Influência na validação dos modelos

Na Tabela 22 são apresentados os resultados dos critérios de Akaike (\hat{A}), Filliben (r) e valor máximo da verossimilhança (VMV), para validação de ajustes dos modelos geoestatísticos, para a variável produtividade com dados originais (Prod) e sem as observações influentes (Prod-13-23-30).

Tabela 232 - Resultados dos critérios de Akaike (\hat{A}), Filliben (r^*) e valor máximo do logaritmo da função verossimilhança (VMV) para a produtividade com dados originais e sem dados influentes

	Modelo	Mét.	N	\hat{A}	r^*	VMV	
Dados originais	Exp.	MV	47	-79,04015	0,9913485	-16,18	
		MVR	47	-66,56789	0,9912173	-16,42	
	Gaus.	MV	47	-86,19630	0,9909290	-15,86	
		MVR	47	-80,73545	0,9904414	-16,36	
	Matérn k=0,7	MV	47	-79,11687	0,9912577	-16,14	
		MVR	47	-64,07900	0,9911578	-16,39	
	Matérn k=3,0	MV	47	-81,66521	0,9910980	-15,98	
		MVR	47	-67,21813	0,9906132	-16,38	
	Sem dados influentes	Exp.	MV	44	-18,15724	0,9898262	-9,49
			MVR	44	-16,83605	0,9895991	-10,55
Gaus.		MV	44	-18,15000	0,9911868	-9,01	
		MVR	44	-16,66785	0,9905340	-10,26	
Matérn k=0,7		MV	44	-18,18989	0,9900218	-9,43	
		MVR	44	-16,84991	0,9896835	-10,50	
Matérn k=3,0		MV	44	-18,27464	0,9906348	-9,16	
		MVR	44	-16,92014	0,9900728	-10,32	

Nota: r^* : teste de Filliben, compara-se com $r = 0,9740$ (Dados originais) e $r = 0,9730$ (Sem dados influentes).

Segundo o critério de AKAIKE (1973), para o conjunto original de dados, o menor valor de \hat{A} aponta como melhor ajuste o modelo gaussiano, estimado pelo método MV. Entretanto, quando foram retiradas as observações influentes, o menor valor de \hat{A} indicou como melhor ajuste o modelo Matérn, com $kappa$ igual a 3,0, com parâmetros também estimados pelo método MV.

O critério de FILLIBEN (1975) testou se os resíduos ortonormais seguiam uma distribuição normal, ao nível de 5 % de significância. Para o caso dos dados da variável produtividade com dados originais e sem dados influentes têm-se 47 e 44 valores amostrados, respectivamente. Os valores tabelados de r , obtidos da tabela de FILLIBEN (1975), foram de 0,9740 e 0,9730, respectivamente. Comparando-se com os valores r^* calculados (Tabela 22), para dados originais da produtividade e sem as observações influentes, verificou-se que, para todos os modelos ajustados, os valores de r^* são maiores do que r tabelado, indicando que a hipótese de que os dados seguem distribuição normal foi aceita, para um nível de 5% de significância, portanto, todos os ajustes foram aceitos.

Por último, pelo critério do valor máximo do logaritmo da função verossimilhança (VMV), tanto para produtividade com dados originais, quanto sem os valores influentes, o melhor modelo ajustado foi o gaussiano, com parâmetros estimados por MV.

Na Tabela 23, são apresentados os resultados da validação de modelos ajustados, utilizando o critério da validação cruzada para a variável produtividade com dados originais e sem as observações influentes, para os quais se ajustaram os modelos exponencial, gaussiano e Matérn, combinados com os estimadores máxima verossimilhança (MV) e máxima verossimilhança restrita (MVR).

O erro médio (EM), erro médio reduzido (\overline{ER}), desvio padrão dos erros (DPe), desvio padrão dos erros reduzidos (S_{ER}) e do erro absoluto (EA) foram utilizados como critérios para a comparação entre os modelos ajustados.

Segundo McBRATNEY e WEBSTER (1986), CRESSIE (1993) e MELLO *et al.*, (2005), o valor do erro médio reduzido deve ser zero e do desvio padrão do erro reduzido deve ser igual a um. Portanto, os valores de EM e \overline{ER} mais próximos de zero, o menor valor de DP e o valor de S_{ER} mais próximo de um, são os critérios para escolha do melhor modelo ajustado. O erro absoluto EA é uma medida da magnitude dos erros na unidade da variável.

Tabela 243 - Resultados da validação cruzada para a variável produtividade com dados originais e sem dados influentes

	Modelo	Mét.	EM	\overline{ER}	DPe	S_{ER}	EA	
Dados originais	Exp.	MV	-0,00546	-0,00388	0,39311	1,08127	14,29101	
		MVR	-0,00415	-0,00252	0,38967	1,00024	14,28861	
	Gaus.	MV	-0,00531	-0,00384	0,38918	1,08186	14,05426	
		MVR	-0,00497	-0,00325	0,38861	0,99977	14,13366	
	Matérn k=0,7	MV	-0,00541	-0,00385	0,39253	1,08123	14,26218	
		MVR	-0,00394	-0,00237	0,38915	1,00038	14,29727	
	Matérn k=3,0	MV	-0,00531	-0,00382	0,39039	1,08142	14,13835	
		MVR	-0,00384	-0,00243	0,38895	1,00094	14,32584	
	Sem dados influentes	Exp.	MV	0,00111	0,00211	0,34253	1,08298	11,88536
			MVR	0,00087	0,00154	0,34145	0,99756	11,75648
Gaus.		MV	0,00115	0,00220	0,33656	1,08216	11,65584	
		MVR	0,00099	0,00173	0,33669	0,99613	11,58336	
Matérn k=0,7		MV	0,00108	0,00206	0,34166	1,08283	11,85134	
		MVR	0,00087	0,00154	0,34073	0,99746	11,71984	
Matérn k=3,0		MV	0,00105	0,00201	0,33822	1,08239	11,71264	
		MVR	0,00092	0,00161	0,33796	0,99673	11,61951	

Os critérios da validação cruzada, aplicados aos modelos em estudo, para a variável produtividade com o conjunto original de dados, apresentados na Tabela 23, indicaram como melhor ajuste o modelo Matérn, com $kappa$ igual a 3,0, parâmetros estimados pelo método máxima verossimilhança restrita (MVR). Entretanto, quando foram retiradas as observações consideradas influentes pelas técnicas de diagnóstico, a validação cruzada indicou como melhor ajuste o modelo exponencial, com parâmetros estimados por máxima verossimilhança restrita (MVR). Assim sendo, pelos critérios da validação cruzada, as observações 13, 23 e 30 são influentes também na escolha do melhor modelo espacial ajustado.

4.3.3.3.4 Influência na construção dos mapas temáticos

A Figura 41 apresenta os mapas temáticos da variável produtividade com dados originais (Figura 41a) e sem dados influentes (Figura 41b), confeccionados por meio da interpolação por krigagem com tendência, visto que neste estudo foram introduzidas covariáveis ou variáveis explicativas. Os mapas foram construídos utilizando-se os modelos indicados pelo critério da validação cruzada.

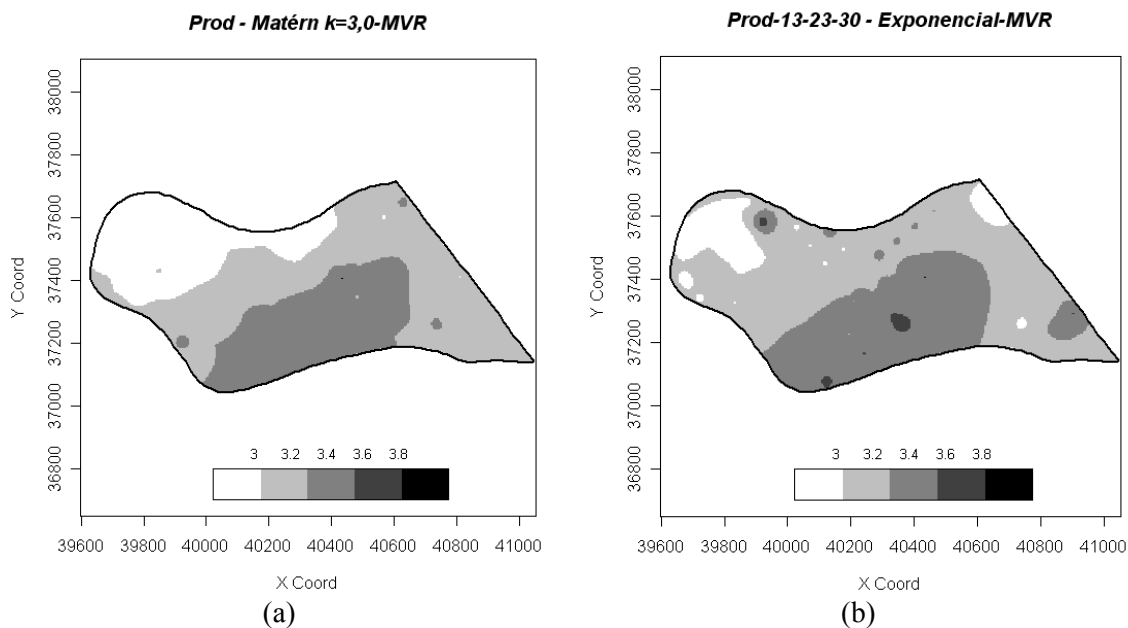


Figura 41 - Mapas temáticos para a variável produtividade: a) com dados originais (Prod); b) retirando-se as observações 13, 23 e 30 (Prod-13-23-30).

Pelas figuras 41a e 41b, pode-se observar que há grande variação entre as áreas dos mapas. O mapa obtido para a variável produtividade com o conjunto original de dados (Figura 41a) mostra que a área possui uma grande região com produtividade abaixo de $3,0 \text{ t ha}^{-1}$. No mapa confeccionado sem as observações consideradas influentes (Figura 41b) a região abaixo de $3,0 \text{ t ha}^{-1}$ é bem menor, e observa-se a presença de outras regiões nesta mesma faixa. Além disso, na Figura 41b nota-se a presença de subáreas com produção acima de $3,8 \text{ t ha}^{-1}$, o que não ocorre no mapa da Figura 41a.

Estas constatações indicam que as observações 13, 23 e 30, estão exercendo grande influência também na construção dos mapas temáticos, obtidos pela interpolação por krigagem com tendência.

Portanto, a aplicação das técnicas de diagnóstico mostrou-se importante na identificação de dados influentes, os quais exercem grande influência na construção dos mapas temáticos. Assim, o mapa da distribuição espacial da produtividade que seria apresentado ao agricultor, não estaria representando de forma correta as tendências da produtividade na área em estudo. Isso acarretaria interpretações errôneas quanto às futuras intervenções no trato do solo, de acordo com os princípios da agricultura de precisão de atender às exigências locais.

5 CONCLUSÕES

Pelos estudos com dados simulados, verificou-se que as técnicas de diagnóstico aplicadas foram eficientes na identificação dos valores perturbados. Quando se aplicaram as mesmas técnicas para dados experimentais, elas identificaram a presença de valores atípicos, os quais foram altamente influentes na determinação da estrutura de dependência espacial e na escolha do melhor modelo ajustado. Por meio destes resultados observou-se que a presença de valores atípicos pode exercer grande influência nos mapas temáticos.

Nos ajustes de modelos espaciais geoestatísticos, tanto para estudos com dados simulados como para estudos com dados experimentais, o método de máxima verossimilhança restrita para estimação de parâmetros mostrou-se mais robusto à presença de observações influentes do que o método de máxima verossimilhança.

Os critérios de validação aplicados nos estudos não convergiram para o mesmo modelo na escolha do melhor ajuste, sendo o critério de Filliben o menos conclusivo.

Finalmente a aplicação de técnicas de diagnóstico, para avaliar a influência local de observações atípicas nos resultados das análises, deve fazer parte de toda análise geoestatística. Desta forma garante-se que as informações contidas nos mapas temáticos tenham maior qualidade e possam ser utilizadas com maior segurança pelo usuário final, que é o agricultor.

REFERÊNCIAS

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. **Journal of the Royal Statistical Society**, London, v. 51, p. 469-483, 1973.

BURGESS, T. M.; WEBSTER, R. Optimal interpolation and isarithmic mapping of soil properties: I. The semi-variogram and punctual kriging. **Journal of Soil Science**, Oxford, v. 31, n. 2, p. 315-331, 1980.

CADIGAN, N. G; FARRELL, P. J. Generalized local influence with applications to fish stock cohort analysis. **Journal of the Royal Statistical Society**, London, v. 51, p. 469-483, 2002.

CAMBARDELLA, C. A.; MOORMAN, T. B.; NOVAK, J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA, A. E. Field-scale variability of soil properties in Central Iowa soils. **Journal of Soil Science**, Madison, v. 58, n. 8, p. 1501-1511, 1994.

CARVALHO, L. P.; URIBE-OPAZO, M. A.; JOHANN, J. A.; SOUZA, E. G. Estudo da anisotropia na variabilidade espacial em propriedades químicas do solo. **Revista Brasileira de Ciência do Solo**, Viçosa - MG, 2007. Artigo aceito para ser publicado em 2008.

CHRISTENSEN, R.; JOHNSON, W.; PEARSON, L. Covariance function diagnostics for spatial linear models. **International Association for Mathematical geology**, New York, v. 25, n.2, p. 145-160, 1993.

COOK, R. D. Assessment of local influence (with discussion). **Journal of the Royal Statistical Society**, Series B, London, v. 48, p. 133-169, 1986.

COX, D. R.; HINKLEY, D. U. **Theoretical Statistics**, London: Chapman and Hall, 1974. 532 p.

CRESSIE, N. A. C. **Statistic for spatial data**. Edição revisada. New York: John Wiley & Sons, 1993. 900p.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA. Serviço Nacional de Levantamento e Conservação de Solos. **Manual de métodos de análise de solo**. Rio de Janeiro: Ministério da Agricultura, 1979. 247p.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA. Serviço Nacional de Levantamento e Conservação de Solos. **Manual de métodos de análise do solo**. Rio de Janeiro: Centro Nacional de Pesquisas de Solos, 2 ed, 1997. 212p.

FILLIBEN, J. J. The probability plot correlation coefficient test for normality. **Technometrics**, Alexandria, v. 17, n. 1, p. 111-117, 1975.

FUNG, W. K. Unmasking outliers and leverage points: a confirmation. **Journal of the American Statistical Association**, Alexandria, v. 88, p. 515-519, 1993.

GALEA-ROJAS, M. G. Técnicas de diagnóstico en regresión lineal. **Revista de la Sociedad Chilena de Estadística**, Santiago-Chile, v. 42, n. 1-2, p. 23-43, 1990.

GOMES, P. **Curso de estatística experimental**. 14 ed. Piracicaba: Degaspari, 2000. 477 p.

GONÇALVES, A. C. A. **A variabilidade especial de propriedades físicas do solo para fins de manejo da irrigação**. 1997. 118 f. Tese (Doutorado em Agronomia) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1997.

GUERRA, P. A. G. **Geoestatística operacional**. Brasília: Ministério das Minas e Energia, Departamento Nacional da Produção Mineral, 1988, 145p.

HOAGLIN, D.; WELSH, R. The hat matrix in regression and ANOVA. **The American Statistician Journal**, Alexandria, v. 32, p. 17-22, 1978.

ISAAKS, E. H.; SRIVASTAVA, R. M. **An introduction to applied geostatistics**. New York: Oxford University Press, 1989. 561p.

JOHANN, J. A.; URIBE-OPAZO, M. A.; SOUZA, E. G.; ROCHA, J. V. Variabilidade espacial dos atributos físicos do solo e da produtividade em uma área experimental de agricultura de precisão com e sem manejo localizado.

Revista Brasileira de Engenharia Agrícola e Ambiental, Campina Grande-PB, v. 8, n. 2, p. 1-8, 2004.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. Madison: Prentice Hall International, 1982. 607p.

JOURNEL, A. G. **Fundamentals of geostatistics in five lessons**. Washington: American Geophysical Union, 1989. 40p.

JOURNEL, A. G.; HUIJBREGTS, C. J. **Mining geostatistics**. London: Academic Press, 1978. 600p.

KITANIDIS, P. Statistical estimation of polynomial generalized covariance functions and hydrologic applications. **Water Resources Research**, Novo México, v. 19, p. 909-921, 1983.

LEE, S. Validation of geostatistical models using the Filliben test for orthonormal residual. **Journal of Hydrology**, Amsterdam, v. 158, p. 319-332, 1994.

MARDIA, K; MARSHALL, R. Maximum likelihood estimation of models for residual covariance in spatial regression. **Biometrika**, Alemanha, v. 71, p. 135-146, 1984.

MATÉRN, B. Spatial variation. **Lecture notes in statistics**. 2 ed. New York: Springer, 1986. 151p.

McBRATNEY, A.; WEBSTER, R. Choosing functions for semi-variograms of soil properties and fitting them to sample estimates. **Journal of Soil Science**, Oxford, v. 37, p. 617-639, 1986.

MELLO, J.M.; BATISTA, J.L.F.; RIBEIRO JR, P.J.; OLIVEIRA, M.S. Ajuste e seleção de modelos espaciais de semivariograma visando à estimativa volumétrica de *Eucalyptus grandis*. **Scientia Forestalis**, Piracicaba, n. 69, p.25-37, dez. 2005.

ORTEGA, E; BOLFARINE, H; PAULA G. Influence diagnostics in generalized log-gamma regression models. **Computacional Statistics & Data Analysis Journal**, Holanda, v. 42, p. 165-186, 2002.

OSORIO, F.; PAULA, G. A.; GALEA, M. Assessment of local influence in elliptical linear models with longitudinal structure. **Computacional Statistics & Data Analysis Journal**, Holanda, v. 51, p. 4354-4368, 2007.

PAULA, G. A. **Modelos de regressão com apoio computacional**. São Paulo - SP: Instituto de Matemática e Estatística – IME; USP, 2004. 233p.

PEÑA, D; YOHAI, V. A fast procedure for outlier diagnostics in large regression problems. **Journal of the American Statistical Association**, Alexandria, v. 94, p. 434-445, 1999.

PIEGORSCH, W. W.; BAILER, A. J. **Analyzing environmental data**. New York: John Wiley & Sons, 2005. 512p.

RIBEIRO JR, P. J. Estatística no espaço: modelos e métodos para análise de dados georreferenciados. REUNIÃO REGIONAL DA ASSOCIAÇÃO BRASILEIRA DE ESTATÍSTICA - ABE, 33, 2002, Curitiba - PR **Anais...** Curitiba – PR: Associação Brasileira de Estatística – ABE, 2002.

ROSS, W. The geometry of case deletion and the assessment of influence in nonlinear regression. **The Canadian Journal of Statistics**, Canadá, v. 15, p. 91-103, 1987.

SILVA, E. A. A.; URIBE-OPAZO, M. A.; SOUZA, E. G.; ROCHA, J. V. Um estimador robusto e o semivariograma cruzado na análise de variabilidade espacial de atributos de solo e planta, **Revista Acta Scientiarum**, Maringá - PR, v. 25, n. 2, p. 365-371, 2003.

SOARES, A. **Geoestatística para as ciências da terra e do ambiente**. Portugal: IST Press, 2000. 186p.

SOUZA, E. G.; JOHANN, J. A.; ROCHA, J. V.; RIBEIRO, S. R. A.; SILVA, M. S.; URIBE-OPAZO, M. A.; MOLIN, J. P.; OLIVEIRA, E. F.; NÓBREGA, L. H. P. Variabilidade espacial dos atributos químicos do solo em um latossolo roxo distrófico da região de Cascavel - PR. **Revista Engenharia Agrícola**, Jaboticabal - SP, v. 18, n. 3, p. 80-92, 1999.

ST. LAURENT, R.; COOK, D. Leverage and superleverage in nonlinear regression. **Journal of the American Statistical Association**, Alexandria, v. 87, p. 985-990, 1992.

TIEPPO, R. C. **Desenvolvimento de um penetrógrafo eletrônico**. 2004. 85 f. Dissertação (Mestrado em Engenharia Agrícola). Centro de Ciências Exatas e Tecnológicas, Universidade Estadual do Oeste do Paraná, Cascavel – PR, 2004.

URIBE-OPAZO, M. A.; KAVANAGH, C. R. Estimação de Parâmetros no ajuste de modelos teóricos a semivariogramas experimentais. **Revista Brasileira de Estatística**, Rio de Janeiro - RJ, v. 67, p. 97-122, 2006.

VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial do solo. **In: UFV – UNIVERSIDADE FEDERAL DE VIÇOSA. Tópicos em Ciências do Solo**. Viçosa: SBCS, 2000.

WEBSTER, R.; OLIVER, M.A. **Statistical methods in soil and land resource survey**. Oxford: Oxford University Press, 1990. 316p.

WEI, B.; HU, Y.; FUNG, W. Generalized leverage and its applications. **Scandinavian Journal of Statistics**, Oxford, v. 25, p. 25-37, 1998.

ZIMMERMAN, D. Computationally efficient restricted maximum likelihood estimation of generalized covariance functions. **International Association for Mathematical geology**, New York, v. 21, n. 7, p. 655-672, 1989.

APÊNDICES

APÊNDICE A - CONJUNTOS DE DADOS SIMULADOS

Variável simulada “Exponencial 0-10-10”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	0.1781637	51	0.000	55.556	9.4540788
2	11.111	0.000	8.4918966	52	11.111	55.556	4.6077179
3	22.222	0.000	8.7316483	53	22.222	55.556	5.5679896
4	33.333	0.000	10.4833479	54	33.333	55.556	12.6324887
5	44.444	0.000	6.4753887	55	44.444	55.556	10.4096755
6	55.556	0.000	8.4548315	56	55.556	55.556	12.3146135
7	66.667	0.000	8.0670423	57	66.667	55.556	12.3967550
8	77.778	0.000	13.6489685	58	77.778	55.556	9.5871930
9	88.889	0.000	4.2927205	59	88.889	55.556	13.5421767
10	100.000	0.000	1.3870834	60	100.000	55.556	7.2107200
11	0.000	11.111	3.0796910	61	0.000	66.667	8.4401405
12	11.111	11.111	5.0619201	62	11.111	66.667	3.6523591
13	22.222	11.111	0.7675412	63	22.222	66.667	11.5723676
14	33.333	11.111	8.3447319	64	33.333	66.667	12.8377451
15	44.444	11.111	6.1024138	65	44.444	66.667	6.5542934
16	55.556	11.111	8.4361741	66	55.556	66.667	24.067972
17	66.667	11.111	5.5998004	67	66.667	66.667	9.1247557
18	77.778	11.111	11.4892639	68	77.778	66.667	7.9500828
19	88.889	11.111	3.8178726	69	88.889	66.667	10.8424736
20	100.000	11.111	4.4029792	70	100.000	66.667	7.7189716
21	0.000	22.222	7.6398684	71	0.000	77.778	10.0113450
22	11.111	22.222	7.3716991	72	11.111	77.778	8.4163624
23	22.222	22.222	3.7551426	73	22.222	77.778	6.6016707
24	33.333	22.222	6.7937489	74	33.333	77.778	8.6537197
25	44.444	22.222	7.5609244	75	44.444	77.778	5.3530999
26	55.556	22.222	9.1410422	76	55.556	77.778	6.8797932
27	66.667	22.222	11.7144749	77	66.667	77.778	5.0049648
28	77.778	22.222	5.9035447	78	77.778	77.778	7.4140407
29	88.889	22.222	9.3133367	79	88.889	77.778	12.7624546
30	100.000	22.222	7.3837141	80	100.000	77.778	5.6267839
31	0.000	33.333	8.2102021	81	0.000	88.889	4.3072914
32	11.111	33.333	3.9648606	82	11.111	88.889	7.9997767
33	22.222	33.333	6.4841110	83	22.222	88.889	7.9011192
34	33.333	33.333	8.6362208	84	33.333	88.889	3.3374915
35	44.444	33.333	8.3089457	85	44.444	88.889	7.7453529
36	55.556	33.333	6.0739297	86	55.556	88.889	9.3729284
37	66.667	33.333	10.2742662	87	66.667	88.889	7.2485117
38	77.778	33.333	9.1177446	88	77.778	88.889	9.3274251
39	88.889	33.333	4.7245625	89	88.889	88.889	7.0380352
40	100.000	33.333	8.6591957	90	100.000	88.889	6.2813627
41	0.000	44.444	0.5023872	91	0.000	100.000	6.4358087
42	11.111	44.444	8.3496461	92	11.111	100.000	3.6409581
43	22.222	44.444	9.0535840	93	22.222	100.000	4.4165782
44	33.333	44.444	5.1757959	94	33.333	100.000	7.0366346
45	44.444	44.444	13.1915219	95	44.444	100.000	6.7825492
46	55.556	44.444	11.0775076	96	55.556	100.000	14.2312639
47	66.667	44.444	12.1566424	97	66.667	100.000	12.0758968
48	77.778	44.444	11.0131218	98	77.778	100.000	8.8056652
49	88.889	44.444	13.5317499	99	88.889	100.000	10.9389412
50	100.000	44.444	7.9217520	100	100.000	100.000	5.5936598

Variável simulada “Exponencial 0-10-15”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	10.445333	51	0.000	55.556	10.631982
2	11.111	0.000	9.698180	52	11.111	55.556	8.202423
3	22.222	0.000	11.666401	53	22.222	55.556	9.180271
4	33.333	0.000	6.074574	54	33.333	55.556	9.244385
5	44.444	0.000	3.578422	55	44.444	55.556	6.806847
6	55.556	0.000	4.854117	56	55.556	55.556	9.538845
7	66.667	0.000	7.775321	57	66.667	55.556	11.088872
8	77.778	0.000	6.448764	58	77.778	55.556	11.744773
9	88.889	0.000	11.349199	59	88.889	55.556	14.044394
10	100.000	0.000	11.391167	60	100.000	55.556	14.527957
11	0.000	11.111	10.338289	61	0.000	66.667	8.876921
12	11.111	11.111	11.723110	62	11.111	66.667	4.184170
13	22.222	11.111	8.761854	63	22.222	66.667	4.495870
14	33.333	11.111	9.563003	64	33.333	66.667	11.467714
15	44.444	11.111	9.991040	65	44.444	66.667	12.660780
16	55.556	11.111	12.369209	66	55.556	66.667	11.377121
17	66.667	11.111	10.673154	67	66.667	66.667	9.239994
18	77.778	11.111	12.790319	68	77.778	66.667	12.805276
19	88.889	11.111	26.942078	69	88.889	66.667	15.579114
20	100.000	11.111	14.320975	70	100.000	66.667	13.121989
21	0.000	22.222	8.211002	71	0.000	77.778	7.292277
22	11.111	22.222	9.642327	72	11.111	77.778	2.928135
23	22.222	22.222	7.151575	73	22.222	77.778	7.656131
24	33.333	22.222	6.998902	74	33.333	77.778	8.021888
25	44.444	22.222	9.594213	75	44.444	77.778	13.422358
26	55.556	22.222	9.321937	76	55.556	77.778	11.567964
27	66.667	22.222	15.018382	77	66.667	77.778	5.087127
28	77.778	22.222	13.480019	78	77.778	77.778	7.273180
29	88.889	22.222	15.318747	79	88.889	77.778	11.368191
30	100.000	22.222	10.935527	80	100.000	77.778	10.287541
31	0.000	33.333	12.873012	81	0.000	88.889	7.540465
32	11.111	33.333	9.036641	82	11.111	88.889	4.828307
33	22.222	33.333	8.496260	83	22.222	88.889	7.611070
34	33.333	33.333	9.593764	84	33.333	88.889	8.069447
35	44.444	33.333	4.834159	85	44.444	88.889	10.455285
36	55.556	33.333	7.070171	86	55.556	88.889	4.572729
37	66.667	33.333	13.820063	87	66.667	88.889	6.216603
38	77.778	33.333	13.272141	88	77.778	88.889	10.562714
39	88.889	33.333	14.188313	89	88.889	88.889	13.147534
40	100.000	33.333	10.963231	90	100.000	88.889	10.852956
41	0.000	44.444	12.864159	91	0.000	100.000	1.310266
42	11.111	44.444	2.331437	92	11.111	100.000	6.306861
43	22.222	44.444	5.699805	93	22.222	100.000	6.680226
44	33.333	44.444	3.820574	94	33.333	100.000	10.202205
45	44.444	44.444	5.770801	95	44.444	100.000	7.800194
46	55.556	44.444	6.318104	96	55.556	100.000	5.598716
47	66.667	44.444	9.784132	97	66.667	100.000	9.711335
48	77.778	44.444	12.496285	98	77.778	100.000	11.349207
49	88.889	44.444	14.003191	99	88.889	100.000	11.523076
50	100.000	44.444	11.705595	100	100.000	100.000	13.779627

Variável simulada “Exponencial 0-10-20”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	15.265641	51	0.000	55.556	7.546353
2	11.111	0.000	11.867687	52	11.111	55.556	8.013658
3	22.222	0.000	15.453535	53	22.222	55.556	4.743170
4	33.333	0.000	16.624306	54	33.333	55.556	5.767564
5	44.444	0.000	12.821168	55	44.444	55.556	5.414627
6	55.556	0.000	10.675468	56	55.556	55.556	8.913054
7	66.667	0.000	16.683169	57	66.667	55.556	8.844146
8	77.778	0.000	10.046720	58	77.778	55.556	7.885507
9	88.889	0.000	8.164027	59	88.889	55.556	10.560461
10	100.000	0.000	9.280200	60	100.000	55.556	12.975417
11	0.000	11.111	15.404131	61	0.000	66.667	2.175004
12	11.111	11.111	10.810421	62	11.111	66.667	1.150504
13	22.222	11.111	12.635940	63	22.222	66.667	7.006452
14	33.333	11.111	14.229663	64	33.333	66.667	9.201726
15	44.444	11.111	11.306934	65	44.444	66.667	9.212844
16	55.556	11.111	11.023935	66	55.556	66.667	7.666657
17	66.667	11.111	11.047283	67	66.667	66.667	9.551473
18	77.778	11.111	9.898386	68	77.778	66.667	5.467743
19	88.889	11.111	11.650668	69	88.889	66.667	10.443011
20	100.000	11.111	9.394029	70	100.000	66.667	10.816752
21	0.000	22.222	29.559996	71	0.000	77.778	6.062394
22	11.111	22.222	13.471082	72	11.111	77.778	5.165011
23	22.222	22.222	6.061512	73	22.222	77.778	9.526844
24	33.333	22.222	7.085911	74	33.333	77.778	9.394339
25	44.444	22.222	11.576264	75	44.444	77.778	9.575448
26	55.556	22.222	10.372751	76	55.556	77.778	9.718799
27	66.667	22.222	9.822609	77	66.667	77.778	10.813866
28	77.778	22.222	10.626838	78	77.778	77.778	8.271456
29	88.889	22.222	11.611537	79	88.889	77.778	8.999018
30	100.000	22.222	11.406223	80	100.000	77.778	10.503259
31	0.000	33.333	12.033652	81	0.000	88.889	7.206511
32	11.111	33.333	4.958493	82	11.111	88.889	6.980116
33	22.222	33.333	6.803060	83	22.222	88.889	7.730240
34	33.333	33.333	9.194261	84	33.333	88.889	8.582433
35	44.444	33.333	8.522932	85	44.444	88.889	9.007482
36	55.556	33.333	9.484113	86	55.556	88.889	9.018988
37	66.667	33.333	14.337992	87	66.667	88.889	12.731838
38	77.778	33.333	10.665144	88	77.778	88.889	11.219545
39	88.889	33.333	14.098005	89	88.889	88.889	9.590693
40	100.000	33.333	13.422046	90	100.000	88.889	10.315212
41	0.000	44.444	5.257622	91	0.000	100.000	9.775900
42	11.111	44.444	5.867958	92	11.111	100.000	9.749563
43	22.222	44.444	6.859470	93	22.222	100.000	9.088687
44	33.333	44.444	2.387884	94	33.333	100.000	9.248501
45	44.444	44.444	8.771964	95	44.444	100.000	7.386072
46	55.556	44.444	13.842652	96	55.556	100.000	10.718546
47	66.667	44.444	9.461751	97	66.667	100.000	10.023195
48	77.778	44.444	9.813206	98	77.778	100.000	7.850544
49	88.889	44.444	15.432538	99	88.889	100.000	7.784654
50	100.000	44.444	13.193756	100	100.000	100.000	8.179645

Variável simulada “Exponencial 0-10-60”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	6.923456	51	0.000	55.556	8.502575
2	11.111	0.000	7.313487	52	11.111	55.556	6.280729
3	22.222	0.000	6.446856	53	22.222	55.556	6.671088
4	33.333	0.000	6.079397	54	33.333	55.556	5.281108
5	44.444	0.000	5.490540	55	44.444	55.556	6.778884
6	55.556	0.000	6.645976	56	55.556	55.556	7.765337
7	66.667	0.000	9.357596	57	66.667	55.556	6.333030
8	77.778	0.000	8.175770	58	77.778	55.556	7.112282
9	88.889	0.000	6.736885	59	88.889	55.556	8.955974
10	100.000	0.000	8.088163	60	100.000	55.556	11.435297
11	0.000	11.111	9.959942	61	0.000	66.667	7.764108
12	11.111	11.111	9.142959	62	11.111	66.667	3.616486
13	22.222	11.111	8.253712	63	22.222	66.667	7.456221
14	33.333	11.111	8.617238	64	33.333	66.667	5.439929
15	44.444	11.111	7.156841	65	44.444	66.667	5.893051
16	55.556	11.111	9.567370	66	55.556	66.667	7.521045
17	66.667	11.111	8.289673	67	66.667	66.667	7.042878
18	77.778	11.111	5.151512	68	77.778	66.667	7.164128
19	88.889	11.111	6.574913	69	88.889	66.667	7.628032
20	100.000	11.111	5.391407	70	100.000	66.667	12.410076
21	0.000	22.222	6.976046	71	0.000	77.778	8.053990
22	11.111	22.222	9.876576	72	11.111	77.778	6.862924
23	22.222	22.222	6.665498	73	22.222	77.778	8.035379
24	33.333	22.222	7.517551	74	33.333	77.778	6.583838
25	44.444	22.222	8.658902	75	44.444	77.778	5.753674
26	55.556	22.222	5.084327	76	55.556	77.778	8.387058
27	66.667	22.222	5.064478	77	66.667	77.778	6.009843
28	77.778	22.222	3.652408	78	77.778	77.778	9.280316
29	88.889	22.222	5.953299	79	88.889	77.778	11.225019
30	100.000	22.222	6.543982	80	100.000	77.778	21.595369
31	0.000	33.333	7.803239	81	0.000	88.889	7.711835
32	11.111	33.333	8.319171	82	11.111	88.889	6.571939
33	22.222	33.333	6.615711	83	22.222	88.889	5.074881
34	33.333	33.333	5.969433	84	33.333	88.889	3.808384
35	44.444	33.333	3.771207	85	44.444	88.889	6.872130
36	55.556	33.333	5.055523	86	55.556	88.889	4.101184
37	66.667	33.333	4.934007	87	66.667	88.889	8.009051
38	77.778	33.333	6.559702	88	77.778	88.889	9.528038
39	88.889	33.333	6.984605	89	88.889	88.889	9.316096
40	100.000	33.333	9.490018	90	100.000	88.889	11.348135
41	0.000	44.444	8.832620	91	0.000	100.000	9.177502
42	11.111	44.444	8.765904	92	11.111	100.000	6.725878
43	22.222	44.444	9.187825	93	22.222	100.000	5.486503
44	33.333	44.444	8.100933	94	33.333	100.000	5.908560
45	44.444	44.444	5.832492	95	44.444	100.000	5.694468
46	55.556	44.444	7.268671	96	55.556	100.000	4.404689
47	66.667	44.444	4.487751	97	66.667	100.000	6.577328
48	77.778	44.444	6.674493	98	77.778	100.000	8.442294
49	88.889	44.444	7.678757	99	88.889	100.000	8.837658
50	100.000	44.444	10.995226	100	100.000	100.000	12.204616

Variável simulada “Gaussiano 0-10-10”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	12.6022264	51	0.000	55.556	10.2282161
2	11.111	0.000	9.5445917	52	11.111	55.556	10.7335312
3	22.222	0.000	8.4890031	53	22.222	55.556	9.9334656
4	33.333	0.000	6.8512850	54	33.333	55.556	7.7419994
5	44.444	0.000	14.3855936	55	44.444	55.556	10.8159233
6	55.556	0.000	15.5216061	56	55.556	55.556	12.1687672
7	66.667	0.000	13.5047856	57	66.667	55.556	9.4570332
8	77.778	0.000	9.2169081	58	77.778	55.556	11.8361266
9	88.889	0.000	7.2799414	59	88.889	55.556	6.6837498
10	100.000	0.000	5.4782760	60	100.000	55.556	6.0678114
11	0.000	11.111	8.6871654	61	0.000	66.667	17.2111130
12	11.111	11.111	8.9607728	62	11.111	66.667	11.5736782
13	22.222	11.111	6.1759702	63	22.222	66.667	9.1687596
14	33.333	11.111	8.3510185	64	33.333	66.667	12.4136149
15	44.444	11.111	11.8077398	65	44.444	66.667	12.8874557
16	55.556	11.111	16.4918764	66	55.556	66.667	13.4043817
17	66.667	11.111	12.6730666	67	66.667	66.667	12.4094443
18	77.778	11.111	9.3634134	68	77.778	66.667	9.9804418
19	88.889	11.111	8.9867625	69	88.889	66.667	5.0127042
20	100.000	11.111	7.8493956	70	100.000	66.667	12.4340203
21	0.000	22.222	9.1075044	71	0.000	77.778	9.7167780
22	11.111	22.222	9.4718701	72	11.111	77.778	7.1590257
23	22.222	22.222	8.2624548	73	22.222	77.778	5.7000905
24	33.333	22.222	10.7480194	74	33.333	77.778	9.8334279
25	44.444	22.222	14.9528274	75	44.444	77.778	7.6981059
26	55.556	22.222	13.7438475	76	55.556	77.778	9.8666700
27	66.667	22.222	7.6946618	77	66.667	77.778	8.6198263
28	77.778	22.222	11.4599422	78	77.778	77.778	4.1837938
29	88.889	22.222	11.1532787	79	88.889	77.778	3.9223172
30	100.000	22.222	11.0098279	80	100.000	77.778	6.8457015
31	0.000	33.333	10.6451446	81	0.000	88.889	6.0978208
32	11.111	33.333	11.4459998	82	11.111	88.889	9.8783394
33	22.222	33.333	7.0664460	83	22.222	88.889	12.2149312
34	33.333	33.333	11.6250959	84	33.333	88.889	5.3308717
35	44.444	33.333	28.4762467	85	44.444	88.889	8.0133845
36	55.556	33.333	13.8711394	86	55.556	88.889	9.1886702
37	66.667	33.333	10.3033121	87	66.667	88.889	8.6354443
38	77.778	33.333	10.7877100	88	77.778	88.889	0.5319427
39	88.889	33.333	15.9223730	89	88.889	88.889	0.7519928
40	100.000	33.333	5.7198462	90	100.000	88.889	5.0759893
41	0.000	44.444	6.4711566	91	0.000	100.000	5.0461823
42	11.111	44.444	6.3621089	92	11.111	100.000	12.9910914
43	22.222	44.444	6.5370736	93	22.222	100.000	9.2272262
44	33.333	44.444	10.6308520	94	33.333	100.000	9.2256147
45	44.444	44.444	9.7507055	95	44.444	100.000	11.4018967
46	55.556	44.444	8.8869970	96	55.556	100.000	10.6386014
47	66.667	44.444	11.1134925	97	66.667	100.000	8.2220914
48	77.778	44.444	10.7050654	98	77.778	100.000	10.6954647
49	88.889	44.444	11.5973967	99	88.889	100.000	9.5326405
50	100.000	44.444	8.3038225	100	100.000	100.000	10.7333997

Variável simulada “Gaussiano 0-10-15”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	8.79719045	51	0.000	55.556	12.71808338
2	11.111	0.000	12.25319116	52	11.111	55.556	14.25716116
3	22.222	0.000	8.76168139	53	22.222	55.556	12.80101463
4	33.333	0.000	8.70557365	54	33.333	55.556	14.01779504
5	44.444	0.000	13.83774478	55	44.444	55.556	13.89153111
6	55.556	0.000	11.66210877	56	55.556	55.556	8.49796082
7	66.667	0.000	8.82708943	57	66.667	55.556	2.07652505
8	77.778	0.000	4.41521831	58	77.778	55.556	5.47475054
9	88.889	0.000	5.09657777	59	88.889	55.556	6.67705115
10	100.000	0.000	5.48906674	60	100.000	55.556	9.86566885
11	0.000	11.111	13.42166918	61	0.000	66.667	11.67059256
12	11.111	11.111	13.26455671	62	11.111	66.667	12.76169554
13	22.222	11.111	9.26538468	63	22.222	66.667	12.44836176
14	33.333	11.111	4.94748091	64	33.333	66.667	12.00745734
15	44.444	11.111	6.33632480	65	44.444	66.667	11.94877087
16	55.556	11.111	7.34604366	66	55.556	66.667	7.64170757
17	66.667	11.111	6.76403281	67	66.667	66.667	5.67807807
18	77.778	11.111	3.42167244	68	77.778	66.667	7.68943232
19	88.889	11.111	5.28850307	69	88.889	66.667	8.84563787
20	100.000	11.111	7.48090325	70	100.000	66.667	8.65830820
21	0.000	22.222	11.63246358	71	0.000	77.778	12.32851755
22	11.111	22.222	11.43769726	72	11.111	77.778	12.22329372
23	22.222	22.222	9.27050468	73	22.222	77.778	14.74319768
24	33.333	22.222	4.98221776	74	33.333	77.778	13.55457337
25	44.444	22.222	6.21338524	75	44.444	77.778	13.65986374
26	55.556	22.222	5.42336871	76	55.556	77.778	12.34794754
27	66.667	22.222	7.86915690	77	66.667	77.778	10.83221549
28	77.778	22.222	8.64801710	78	77.778	77.778	9.43699737
29	88.889	22.222	7.69776898	79	88.889	77.778	7.43904620
30	100.000	22.222	8.25648402	80	100.000	77.778	8.09264062
31	0.000	33.333	8.73195806	81	0.000	88.889	8.71488043
32	11.111	33.333	7.63858402	82	11.111	88.889	6.18568850
33	22.222	33.333	6.37279970	83	22.222	88.889	11.13860797
34	33.333	33.333	10.09897332	84	33.333	88.889	12.85816316
35	44.444	33.333	10.57027821	85	44.444	88.889	10.94926541
36	55.556	33.333	7.86858470	86	55.556	88.889	9.55650951
37	66.667	33.333	8.00365723	87	66.667	88.889	8.82706232
38	77.778	33.333	10.21832877	88	77.778	88.889	11.91158730
39	88.889	33.333	9.24083803	89	88.889	88.889	8.97846001
40	100.000	33.333	9.95287895	90	100.000	88.889	5.57587347
41	0.000	44.444	12.33216260	91	0.000	100.000	7.92006196
42	11.111	44.444	11.41035695	92	11.111	100.000	8.65585354
43	22.222	44.444	10.99845960	93	22.222	100.000	9.68802885
44	33.333	44.444	26.56310576	94	33.333	100.000	8.31101240
45	44.444	44.444	14.39702145	95	44.444	100.000	9.97296974
46	55.556	44.444	10.44885916	96	55.556	100.000	9.10478855
47	66.667	44.444	7.08944779	97	66.667	100.000	11.20366526
48	77.778	44.444	7.48995295	98	77.778	100.000	8.40164424
49	88.889	44.444	6.62292086	99	88.889	100.000	3.04350378
50	100.000	44.444	10.93093491	100	100.000	100.000	0.05725343

Variável simulada “Gaussiano 0-10-20”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	13.011297	51	0.000	55.556	9.149808
2	11.111	0.000	13.810465	52	11.111	55.556	8.707389
3	22.222	0.000	11.232900	53	22.222	55.556	8.270870
4	33.333	0.000	9.449275	54	33.333	55.556	7.835110
5	44.444	0.000	10.782871	55	44.444	55.556	6.834085
6	55.556	0.000	11.602592	56	55.556	55.556	7.327579
7	66.667	0.000	9.805157	57	66.667	55.556	6.188600
8	77.778	0.000	7.881445	58	77.778	55.556	5.000731
9	88.889	0.000	6.163393	59	88.889	55.556	8.653675
10	100.000	0.000	7.525524	60	100.000	55.556	12.651943
11	0.000	11.111	10.354010	61	0.000	66.667	10.936938
12	11.111	11.111	12.212640	62	11.111	66.667	8.479795
13	22.222	11.111	10.504971	63	22.222	66.667	9.165791
14	33.333	11.111	8.388579	64	33.333	66.667	10.028423
15	44.444	11.111	8.853560	65	44.444	66.667	9.134910
16	55.556	11.111	10.183837	66	55.556	66.667	8.584331
17	66.667	11.111	11.240511	67	66.667	66.667	7.861445
18	77.778	11.111	9.887020	68	77.778	66.667	7.936241
19	88.889	11.111	7.455257	69	88.889	66.667	8.947509
20	100.000	11.111	8.002624	70	100.000	66.667	11.034655
21	0.000	22.222	13.622938	71	0.000	77.778	9.051743
22	11.111	22.222	14.969654	72	11.111	77.778	6.797065
23	22.222	22.222	13.693238	73	22.222	77.778	7.639541
24	33.333	22.222	11.270891	74	33.333	77.778	9.331406
25	44.444	22.222	7.953924	75	44.444	77.778	10.529588
26	55.556	22.222	7.500435	76	55.556	77.778	10.223488
27	66.667	22.222	9.656367	77	66.667	77.778	10.183051
28	77.778	22.222	7.902248	78	77.778	77.778	12.351342
29	88.889	22.222	6.325029	79	88.889	77.778	10.886412
30	100.000	22.222	6.099589	80	100.000	77.778	8.904258
31	0.000	33.333	15.009049	81	0.000	88.889	4.999437
32	11.111	33.333	26.906222	82	11.111	88.889	4.317765
33	22.222	33.333	16.026371	83	22.222	88.889	5.721424
34	33.333	33.333	12.888323	84	33.333	88.889	8.206472
35	44.444	33.333	8.108273	85	44.444	88.889	10.438635
36	55.556	33.333	6.977628	86	55.556	88.889	10.642186
37	66.667	33.333	7.365112	87	66.667	88.889	12.129737
38	77.778	33.333	4.412568	88	77.778	88.889	15.441665
39	88.889	33.333	5.149792	89	88.889	88.889	11.508461
40	100.000	33.333	6.719740	90	100.000	88.889	5.894752
41	0.000	44.444	10.833507	91	0.000	100.000	4.923677
42	11.111	44.444	12.316297	92	11.111	100.000	6.819669
43	22.222	44.444	10.499071	93	22.222	100.000	6.616374
44	33.333	44.444	7.975104	94	33.333	100.000	7.642787
45	44.444	44.444	5.965850	95	44.444	100.000	10.460463
46	55.556	44.444	6.501294	96	55.556	100.000	11.588993
47	66.667	44.444	5.770542	97	66.667	100.000	12.744463
48	77.778	44.444	3.600121	98	77.778	100.000	13.502262
49	88.889	44.444	7.303693	99	88.889	100.000	8.389416
50	100.000	44.444	11.397352	100	100.000	100.000	2.818048

Variável simulada “Gaussiano 0-10-60”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	8.505012	51	0.000	55.556	9.482800
2	11.111	0.000	8.065897	52	11.111	55.556	10.135022
3	22.222	0.000	8.234467	53	22.222	55.556	10.658701
4	33.333	0.000	8.925842	54	33.333	55.556	10.849884
5	44.444	0.000	9.869060	55	44.444	55.556	10.583858
6	55.556	0.000	10.681939	56	55.556	55.556	9.895056
7	66.667	0.000	10.985209	57	66.667	55.556	8.955663
8	77.778	0.000	10.549049	58	77.778	55.556	7.974798
9	88.889	0.000	9.408617	59	88.889	55.556	7.107323
10	100.000	0.000	7.858336	60	100.000	55.556	6.430205
11	0.000	11.111	8.131811	61	0.000	66.667	9.980391
12	11.111	11.111	7.851408	62	11.111	66.667	10.801898
13	22.222	11.111	8.101126	63	22.222	66.667	11.375050
14	33.333	11.111	8.777198	64	33.333	66.667	11.495196
15	44.444	11.111	9.604812	65	44.444	66.667	11.104763
16	55.556	11.111	10.231563	66	55.556	66.667	10.334255
17	66.667	11.111	10.340847	67	66.667	66.667	9.428844
18	77.778	11.111	9.768233	68	77.778	66.667	8.613206
19	88.889	11.111	8.578702	69	88.889	66.667	8.000591
20	100.000	11.111	7.046727	70	100.000	66.667	7.599215
21	0.000	22.222	8.161671	71	0.000	77.778	10.302392
22	11.111	22.222	8.090075	72	11.111	77.778	11.243127
23	22.222	22.222	8.415065	73	22.222	77.778	11.881903
24	33.333	22.222	9.003795	74	33.333	77.778	12.018742
25	44.444	22.222	9.590135	75	44.444	77.778	11.647687
26	55.556	22.222	9.884410	76	55.556	77.778	10.956253
27	66.667	22.222	9.674738	77	66.667	77.778	10.210843
28	77.778	22.222	8.897408	78	77.778	77.778	9.610356
29	88.889	22.222	7.663655	79	88.889	77.778	9.215742
30	100.000	22.222	6.222700	80	100.000	77.778	8.988282
31	0.000	33.333	8.471942	81	0.000	88.889	10.303647
32	11.111	33.333	8.651252	82	11.111	88.889	11.304744
33	22.222	33.333	9.054147	83	22.222	88.889	12.023221
34	33.333	33.333	9.515033	84	33.333	88.889	22.135444
35	44.444	33.333	9.794942	85	44.444	88.889	12.045159
36	55.556	33.333	9.697756	86	55.556	88.889	11.564454
37	66.667	33.333	9.144089	87	66.667	88.889	11.060073
38	77.778	33.333	8.183843	88	77.778	88.889	10.675257
39	88.889	33.333	6.970379	89	88.889	88.889	10.421093
40	100.000	33.333	5.713090	90	100.000	88.889	10.241988
41	0.000	44.444	8.944892	91	0.000	100.000	9.901271
42	11.111	44.444	9.378123	92	11.111	100.000	10.880231
43	22.222	44.444	9.849493	93	22.222	100.000	11.660420
44	33.333	44.444	10.167562	94	33.333	100.000	12.050817
45	44.444	44.444	10.142314	95	44.444	100.000	12.066314
46	55.556	44.444	9.691837	96	55.556	100.000	11.872596
47	66.667	44.444	8.873356	97	66.667	100.000	11.643856
48	77.778	44.444	7.834956	98	77.778	100.000	11.451869
49	88.889	44.444	6.747251	99	88.889	100.000	11.266223
50	100.000	44.444	5.759974	100	100.000	100.000	11.039574

Variável simulada “Matérn 0-10-10 k=0,7”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	14.552123	51	0.000	55.556	9.322305
2	11.111	0.000	6.880724	52	11.111	55.556	10.968351
3	22.222	0.000	5.198228	53	22.222	55.556	10.671751
4	33.333	0.000	6.466961	54	33.333	55.556	10.241822
5	44.444	0.000	5.411746	55	44.444	55.556	9.720611
6	55.556	0.000	9.002170	56	55.556	55.556	9.307367
7	66.667	0.000	4.981783	57	66.667	55.556	8.279903
8	77.778	0.000	6.062378	58	77.778	55.556	9.282773
9	88.889	0.000	1.540289	59	88.889	55.556	9.475185
10	100.000	0.000	5.651552	60	100.000	55.556	7.000266
11	0.000	11.111	13.337455	61	0.000	66.667	7.270698
12	11.111	11.111	6.953687	62	11.111	66.667	8.094756
13	22.222	11.111	10.524539	63	22.222	66.667	8.649930
14	33.333	11.111	9.750571	64	33.333	66.667	6.843612
15	44.444	11.111	13.506764	65	44.444	66.667	7.694101
16	55.556	11.111	12.440775	66	55.556	66.667	7.650594
17	66.667	11.111	11.690683	67	66.667	66.667	11.892656
18	77.778	11.111	6.744138	68	77.778	66.667	11.257247
19	88.889	11.111	7.279348	69	88.889	66.667	5.818883
20	100.000	11.111	5.706372	70	100.000	66.667	8.320925
21	0.000	22.222	13.059700	71	0.000	77.778	4.072515
22	11.111	22.222	12.759347	72	11.111	77.778	4.271742
23	22.222	22.222	11.332066	73	22.222	77.778	4.921151
24	33.333	22.222	10.735827	74	33.333	77.778	8.670091
25	44.444	22.222	10.855148	75	44.444	77.778	9.544443
26	55.556	22.222	7.474522	76	55.556	77.778	6.324406
27	66.667	22.222	10.882388	77	66.667	77.778	6.327486
28	77.778	22.222	9.212831	78	77.778	77.778	9.273645
29	88.889	22.222	8.616836	79	88.889	77.778	13.119440
30	100.000	22.222	9.424677	80	100.000	77.778	6.342344
31	0.000	33.333	9.005107	81	0.000	88.889	6.705985
32	11.111	33.333	9.670892	82	11.111	88.889	6.432525
33	22.222	33.333	10.318661	83	22.222	88.889	10.387516
34	33.333	33.333	13.045189	84	33.333	88.889	8.312447
35	44.444	33.333	6.984194	85	44.444	88.889	7.577912
36	55.556	33.333	6.882577	86	55.556	88.889	7.354494
37	66.667	33.333	6.497718	87	66.667	88.889	12.310783
38	77.778	33.333	12.521900	88	77.778	88.889	12.839936
39	88.889	33.333	14.225369	89	88.889	88.889	11.667613
40	100.000	33.333	12.725537	90	100.000	88.889	11.616955
41	0.000	44.444	12.103737	91	0.000	100.000	8.195173
42	11.111	44.444	25.454660	92	11.111	100.000	11.660870
43	22.222	44.444	13.983241	93	22.222	100.000	10.025610
44	33.333	44.444	12.953585	94	33.333	100.000	10.481552
45	44.444	44.444	9.495639	95	44.444	100.000	7.829557
46	55.556	44.444	6.189247	96	55.556	100.000	12.208870
47	66.667	44.444	9.154378	97	66.667	100.000	13.776952
48	77.778	44.444	9.901204	98	77.778	100.000	10.443428
49	88.889	44.444	10.378978	99	88.889	100.000	10.207068
50	100.000	44.444	11.258408	100	100.000	100.000	9.296322

Variável simulada “Matérn 0-10-15 k=0,7”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	7.971957	51	0.000	55.556	8.659870
2	11.111	0.000	10.551311	52	11.111	55.556	5.802922
3	22.222	0.000	8.841363	53	22.222	55.556	2.025967
4	33.333	0.000	9.628423	54	33.333	55.556	3.825745
5	44.444	0.000	10.708146	55	44.444	55.556	6.696611
6	55.556	0.000	12.914911	56	55.556	55.556	7.208454
7	66.667	0.000	10.201467	57	66.667	55.556	4.835604
8	77.778	0.000	6.604080	58	77.778	55.556	1.861533
9	88.889	0.000	11.742936	59	88.889	55.556	3.950397
10	100.000	0.000	11.290989	60	100.000	55.556	5.828027
11	0.000	11.111	6.829466	61	0.000	66.667	2.258808
12	11.111	11.111	10.375722	62	11.111	66.667	3.446275
13	22.222	11.111	10.855805	63	22.222	66.667	4.029910
14	33.333	11.111	11.269064	64	33.333	66.667	3.204131
15	44.444	11.111	12.376914	65	44.444	66.667	9.222944
16	55.556	11.111	13.598495	66	55.556	66.667	7.184513
17	66.667	11.111	7.826462	67	66.667	66.667	11.106134
18	77.778	11.111	4.878982	68	77.778	66.667	12.674401
19	88.889	11.111	7.247722	69	88.889	66.667	12.142389
20	100.000	11.111	14.367294	70	100.000	66.667	8.744893
21	0.000	22.222	4.719538	71	0.000	77.778	6.069243
22	11.111	22.222	9.631575	72	11.111	77.778	1.872016
23	22.222	22.222	13.143305	73	22.222	77.778	3.809574
24	33.333	22.222	10.293827	74	33.333	77.778	5.126402
25	44.444	22.222	13.059376	75	44.444	77.778	7.739803
26	55.556	22.222	10.559770	76	55.556	77.778	12.217133
27	66.667	22.222	9.355841	77	66.667	77.778	11.552683
28	77.778	22.222	8.817499	78	77.778	77.778	9.149046
29	88.889	22.222	8.298445	79	88.889	77.778	10.134655
30	100.000	22.222	10.888143	80	100.000	77.778	9.056979
31	0.000	33.333	7.596613	81	0.000	88.889	7.694012
32	11.111	33.333	8.262117	82	11.111	88.889	6.405583
33	22.222	33.333	7.641054	83	22.222	88.889	1.400635
34	33.333	33.333	10.454937	84	33.333	88.889	3.584284
35	44.444	33.333	7.963878	85	44.444	88.889	6.812888
36	55.556	33.333	8.482178	86	55.556	88.889	12.197090
37	66.667	33.333	9.466837	87	66.667	88.889	13.808601
38	77.778	33.333	10.390108	88	77.778	88.889	9.220278
39	88.889	33.333	3.325011	89	88.889	88.889	7.337744
40	100.000	33.333	3.801986	90	100.000	88.889	8.230286
41	0.000	44.444	7.996102	91	0.000	100.000	23.88505
42	11.111	44.444	4.667720	92	11.111	100.000	7.505123
43	22.222	44.444	4.366723	93	22.222	100.000	1.990869
44	33.333	44.444	6.453355	94	33.333	100.000	5.350951
45	44.444	44.444	4.461246	95	44.444	100.000	5.818042
46	55.556	44.444	5.192502	96	55.556	100.000	10.305892
47	66.667	44.444	4.248641	97	66.667	100.000	10.459535
48	77.778	44.444	9.175738	98	77.778	100.000	8.066003
49	88.889	44.444	6.473994	99	88.889	100.000	3.979684
50	100.000	44.444	6.233749	100	100.000	100.000	9.568859

Variável simulada “Matérn 0-10-20 k=0,7”

n	X	Y	Z	n	X	Y	Z
1	0.00000	0.00000	10.782706	51	0.00000	55.55556	10.495540
2	11.11111	0.00000	6.914007	52	11.11111	55.55556	8.076895
3	22.22222	0.00000	11.045447	53	22.22222	55.55556	6.240443
4	33.33333	0.00000	10.943174	54	33.33333	55.55556	2.289221
5	44.44444	0.00000	13.477198	55	44.44444	55.55556	6.783017
6	55.55556	0.00000	23.123135	56	55.55556	55.55556	10.026318
7	66.66667	0.00000	10.305522	57	66.66667	55.55556	10.536272
8	77.77778	0.00000	8.551040	58	77.77778	55.55556	10.416211
9	88.88889	0.00000	4.041771	59	88.88889	55.55556	8.426939
10	100.00000	0.00000	6.993877	60	100.00000	55.55556	9.368782
11	0.00000	11.11111	10.579752	61	0.00000	66.66667	8.043951
12	11.11111	11.11111	10.718752	62	11.11111	66.66667	12.867653
13	22.22222	11.11111	13.288084	63	22.22222	66.66667	9.814592
14	33.33333	11.11111	10.971097	64	33.33333	66.66667	6.546047
15	44.44444	11.11111	7.287037	65	44.44444	66.66667	4.997121
16	55.55556	11.11111	3.683465	66	55.55556	66.66667	9.938954
17	66.66667	11.11111	5.450229	67	66.66667	66.66667	11.676855
18	77.77778	11.11111	4.563016	68	77.77778	66.66667	9.378713
19	88.88889	11.11111	4.370852	69	88.88889	66.66667	7.083376
20	100.00000	11.11111	6.085179	70	100.00000	66.66667	7.549741
21	0.00000	22.22222	7.100629	71	0.00000	77.77778	8.647218
22	11.11111	22.22222	9.391317	72	11.11111	77.77778	9.716371
23	22.22222	22.22222	9.015851	73	22.22222	77.77778	8.898076
24	33.33333	22.22222	6.681404	74	33.33333	77.77778	6.849633
25	44.44444	22.22222	6.386925	75	44.44444	77.77778	8.196643
26	55.55556	22.22222	3.963328	76	55.55556	77.77778	8.590875
27	66.66667	22.22222	6.109805	77	66.66667	77.77778	9.897076
28	77.77778	22.22222	3.522065	78	77.77778	77.77778	8.433018
29	88.88889	22.22222	4.406654	79	88.88889	77.77778	8.894252
30	100.00000	22.22222	7.054145	80	100.00000	77.77778	11.400704
31	0.00000	33.33333	7.917204	81	0.00000	88.88889	11.540651
32	11.11111	33.33333	8.731858	82	11.11111	88.88889	11.813980
33	22.22222	33.33333	6.569623	83	22.22222	88.88889	10.387730
34	33.33333	33.33333	2.499022	84	33.33333	88.88889	11.961522
35	44.44444	33.33333	4.460105	85	44.44444	88.88889	11.037574
36	55.55556	33.33333	6.466262	86	55.55556	88.88889	8.862800
37	66.66667	33.33333	7.751294	87	66.66667	88.88889	7.814475
38	77.77778	33.33333	4.419972	88	77.77778	88.88889	10.587120
39	88.88889	33.33333	4.990238	89	88.88889	88.88889	10.468313
40	100.00000	33.33333	4.097350	90	100.00000	88.88889	10.815601
41	0.00000	44.44444	7.946597	91	0.00000	100.00000	10.717743
42	11.11111	44.44444	7.267850	92	11.11111	100.00000	11.077742
43	22.22222	44.44444	3.823424	93	22.22222	100.00000	11.312757
44	33.33333	44.44444	5.195105	94	33.33333	100.00000	11.047026
45	44.44444	44.44444	4.994968	95	44.44444	100.00000	9.693447
46	55.55556	44.44444	8.990069	96	55.55556	100.00000	9.646519
47	66.66667	44.44444	7.474635	97	66.66667	100.00000	7.819674
48	77.77778	44.44444	8.590480	98	77.77778	100.00000	7.403397
49	88.88889	44.44444	5.865414	99	88.88889	100.00000	10.317161
50	100.00000	44.44444	7.076761	100	100.00000	100.00000	11.993860

Variável simulada “Matérn 0-10-60 k=0,7”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	5.647086	51	0.000	55.556	10.253399
2	11.111	0.000	7.320879	52	11.111	55.556	10.267192
3	22.222	0.000	6.162752	53	22.222	55.556	9.630374
4	33.333	0.000	6.211360	54	33.333	55.556	7.172201
5	44.444	0.000	7.435150	55	44.444	55.556	5.723834
6	55.556	0.000	9.056123	56	55.556	55.556	4.420197
7	66.667	0.000	8.150092	57	66.667	55.556	5.276954
8	77.778	0.000	6.687131	58	77.778	55.556	6.338116
9	88.889	0.000	5.646663	59	88.889	55.556	5.900444
10	100.000	0.000	6.091747	60	100.000	55.556	6.748743
11	0.000	11.111	7.134102	61	0.000	66.667	7.343499
12	11.111	11.111	4.461426	62	11.111	66.667	9.231730
13	22.222	11.111	5.452729	63	22.222	66.667	9.950673
14	33.333	11.111	8.136387	64	33.333	66.667	8.221480
15	44.444	11.111	7.984287	65	44.444	66.667	6.758200
16	55.556	11.111	8.741812	66	55.556	66.667	4.954506
17	66.667	11.111	8.461146	67	66.667	66.667	4.148604
18	77.778	11.111	6.418025	68	77.778	66.667	5.310122
19	88.889	11.111	5.975718	69	88.889	66.667	5.841450
20	100.000	11.111	5.629207	70	100.000	66.667	6.363850
21	0.000	22.222	9.700694	71	0.000	77.778	5.703736
22	11.111	22.222	7.445244	72	11.111	77.778	9.112649
23	22.222	22.222	6.136761	73	22.222	77.778	7.954217
24	33.333	22.222	5.885145	74	33.333	77.778	8.506529
25	44.444	22.222	5.318157	75	44.444	77.778	7.256863
26	55.556	22.222	7.720025	76	55.556	77.778	4.510597
27	66.667	22.222	8.596789	77	66.667	77.778	3.350012
28	77.778	22.222	9.764343	78	77.778	77.778	3.711628
29	88.889	22.222	7.399985	79	88.889	77.778	3.150846
30	100.000	22.222	5.541860	80	100.000	77.778	5.067174
31	0.000	33.333	19.34440	81	0.000	88.889	7.085153
32	11.111	33.333	8.919952	82	11.111	88.889	6.930147
33	22.222	33.333	6.148232	83	22.222	88.889	6.285262
34	33.333	33.333	4.282855	84	33.333	88.889	6.494532
35	44.444	33.333	5.021922	85	44.444	88.889	5.519016
36	55.556	33.333	6.766531	86	55.556	88.889	4.005362
37	66.667	33.333	9.099508	87	66.667	88.889	3.932526
38	77.778	33.333	9.129877	88	77.778	88.889	3.940752
39	88.889	33.333	7.358315	89	88.889	88.889	3.409895
40	100.000	33.333	4.836894	90	100.000	88.889	2.387401
41	0.000	44.444	10.299011	91	0.000	100.000	6.490439
42	11.111	44.444	8.227956	92	11.111	100.000	5.746882
43	22.222	44.444	6.989143	93	22.222	100.000	5.900542
44	33.333	44.444	6.076871	94	33.333	100.000	5.142252
45	44.444	44.444	5.182202	95	44.444	100.000	5.187145
46	55.556	44.444	5.284758	96	55.556	100.000	2.600428
47	66.667	44.444	6.215158	97	66.667	100.000	2.332607
48	77.778	44.444	7.457402	98	77.778	100.000	1.701824
49	88.889	44.444	6.537895	99	88.889	100.000	2.360041
50	100.000	44.444	5.413173	100	100.000	100.000	2.993435

Variável simulada “Matérn 0-10-10 k=3,0”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	26.039123	51	0.000	55.556	4.810298
2	11.111	0.000	14.252518	52	11.111	55.556	4.879249
3	22.222	0.000	11.150729	53	22.222	55.556	5.454559
4	33.333	0.000	9.430681	54	33.333	55.556	5.984499
5	44.444	0.000	8.913611	55	44.444	55.556	6.905096
6	55.556	0.000	8.205039	56	55.556	55.556	7.267323
7	66.667	0.000	4.718410	57	66.667	55.556	7.814090
8	77.778	0.000	2.647974	58	77.778	55.556	9.410038
9	88.889	0.000	5.086127	59	88.889	55.556	9.877994
10	100.000	0.000	8.061216	60	100.000	55.556	8.302283
11	0.000	11.111	14.938919	61	0.000	66.667	2.754431
12	11.111	11.111	13.682376	62	11.111	66.667	3.753371
13	22.222	11.111	11.536846	63	22.222	66.667	4.454219
14	33.333	11.111	10.044582	64	33.333	66.667	4.778488
15	44.444	11.111	9.596564	65	44.444	66.667	6.353526
16	55.556	11.111	9.495584	66	55.556	66.667	6.188555
17	66.667	11.111	8.160465	67	66.667	66.667	6.771948
18	77.778	11.111	6.831814	68	77.778	66.667	7.861561
19	88.889	11.111	8.143729	69	88.889	66.667	9.546375
20	100.000	11.111	9.932489	70	100.000	66.667	8.270890
21	0.000	22.222	13.956384	71	0.000	77.778	1.752388
22	11.111	22.222	13.446817	72	11.111	77.778	3.537470
23	22.222	22.222	13.280250	73	22.222	77.778	5.702563
24	33.333	22.222	11.725873	74	33.333	77.778	6.165187
25	44.444	22.222	12.021502	75	44.444	77.778	6.088013
26	55.556	22.222	11.593231	76	55.556	77.778	6.234268
27	66.667	22.222	10.442628	77	66.667	77.778	7.464611
28	77.778	22.222	9.875335	78	77.778	77.778	9.436425
29	88.889	22.222	8.992227	79	88.889	77.778	10.044885
30	100.000	22.222	9.853276	80	100.000	77.778	8.978688
31	0.000	33.333	11.501185	81	0.000	88.889	1.822855
32	11.111	33.333	13.787259	82	11.111	88.889	4.979940
33	22.222	33.333	13.061594	83	22.222	88.889	6.846568
34	33.333	33.333	10.622242	84	33.333	88.889	7.238458
35	44.444	33.333	10.401557	85	44.444	88.889	7.236107
36	55.556	33.333	9.939982	86	55.556	88.889	7.078710
37	66.667	33.333	9.990958	87	66.667	88.889	8.734030
38	77.778	33.333	10.569236	88	77.778	88.889	11.559361
39	88.889	33.333	9.060287	89	88.889	88.889	10.952826
40	100.000	33.333	8.830287	90	100.000	88.889	9.154657
41	0.000	44.444	7.328472	91	0.000	100.000	4.913601
42	11.111	44.444	9.017446	92	11.111	100.000	6.306202
43	22.222	44.444	9.576192	93	22.222	100.000	8.887064
44	33.333	44.444	8.712455	94	33.333	100.000	11.256138
45	44.444	44.444	7.911973	95	44.444	100.000	10.404639
46	55.556	44.444	7.723288	96	55.556	100.000	9.531066
47	66.667	44.444	8.410472	97	66.667	100.000	10.007845
48	77.778	44.444	10.451645	98	77.778	100.000	11.356127
49	88.889	44.444	10.273364	99	88.889	100.000	11.081422
50	100.000	44.444	9.032310	100	100.000	100.000	9.636414

Variável simulada “Matérn 0-10-15 k=3,0”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	12.445266	51	0.000	55.556	5.072230
2	11.111	0.000	12.982159	52	11.111	55.556	6.087490
3	22.222	0.000	22.774341	53	22.222	55.556	7.496484
4	33.333	0.000	13.670358	54	33.333	55.556	8.656440
5	44.444	0.000	12.881461	55	44.444	55.556	8.375288
6	55.556	0.000	11.441081	56	55.556	55.556	7.347131
7	66.667	0.000	9.890114	57	66.667	55.556	6.200189
8	77.778	0.000	8.732555	58	77.778	55.556	5.469429
9	88.889	0.000	7.209931	59	88.889	55.556	4.635970
10	100.000	0.000	6.136582	60	100.000	55.556	4.458152
11	0.000	11.111	11.139539	61	0.000	66.667	5.642576
12	11.111	11.111	12.096538	62	11.111	66.667	6.839047
13	22.222	11.111	13.117909	63	22.222	66.667	7.761531
14	33.333	11.111	13.208980	64	33.333	66.667	8.322415
15	44.444	11.111	12.242025	65	44.444	66.667	8.120444
16	55.556	11.111	10.663597	66	55.556	66.667	7.246532
17	66.667	11.111	9.424938	67	66.667	66.667	6.517623
18	77.778	11.111	8.096888	68	77.778	66.667	6.436818
19	88.889	11.111	6.138130	69	88.889	66.667	5.757696
20	100.000	11.111	5.161449	70	100.000	66.667	4.592435
21	0.000	22.222	9.138293	71	0.000	77.778	6.218293
22	11.111	22.222	10.053688	72	11.111	77.778	7.701307
23	22.222	22.222	11.681625	73	22.222	77.778	8.739762
24	33.333	22.222	12.250898	74	33.333	77.778	8.822990
25	44.444	22.222	11.282246	75	44.444	77.778	8.110857
26	55.556	22.222	9.925153	76	55.556	77.778	6.693741
27	66.667	22.222	8.504813	77	66.667	77.778	5.986315
28	77.778	22.222	6.929228	78	77.778	77.778	6.334900
29	88.889	22.222	5.434750	79	88.889	77.778	5.894696
30	100.000	22.222	5.191337	80	100.000	77.778	5.018854
31	0.000	33.333	6.798079	81	0.000	88.889	7.857266
32	11.111	33.333	7.724022	82	11.111	88.889	8.854561
33	22.222	33.333	9.783790	83	22.222	88.889	9.322740
34	33.333	33.333	10.641753	84	33.333	88.889	8.921497
35	44.444	33.333	10.160305	85	44.444	88.889	8.141287
36	55.556	33.333	8.558232	86	55.556	88.889	6.895623
37	66.667	33.333	6.527350	87	66.667	88.889	5.537979
38	77.778	33.333	5.464576	88	77.778	88.889	5.492205
39	88.889	33.333	4.997813	89	88.889	88.889	5.647986
40	100.000	33.333	4.926967	90	100.000	88.889	5.866841
41	0.000	44.444	5.657519	91	0.000	100.000	9.482279
42	11.111	44.444	6.528948	92	11.111	100.000	10.102589
43	22.222	44.444	8.270970	93	22.222	100.000	9.928283
44	33.333	44.444	9.538452	94	33.333	100.000	9.275998
45	44.444	44.444	9.202053	95	44.444	100.000	8.567777
46	55.556	44.444	7.598026	96	55.556	100.000	7.172160
47	66.667	44.444	5.675442	97	66.667	100.000	5.443190
48	77.778	44.444	4.732095	98	77.778	100.000	4.668924
49	88.889	44.444	4.091274	99	88.889	100.000	4.840952
50	100.000	44.444	4.253436	100	100.000	100.000	5.883578

Variável simulada “Matérn 0-10-20 k=3,0”

n	X	Y	Z	n	X	Y	Z
1	0.00000	0.00000	9.053156	51	0.00000	55.55556	8.402052
2	11.11111	0.00000	8.788767	52	11.11111	55.55556	7.383788
3	22.22222	0.00000	8.160736	53	22.22222	55.55556	6.522752
4	33.33333	0.00000	7.444191	54	33.33333	55.55556	5.957403
5	44.44444	0.00000	6.612182	55	44.44444	55.55556	5.828281
6	55.55556	0.00000	6.189693	56	55.55556	55.55556	5.384293
7	66.66667	0.00000	6.291945	57	66.66667	55.55556	5.075744
8	77.77778	0.00000	6.504116	58	77.77778	55.55556	5.390020
9	88.88889	0.00000	6.955111	59	88.88889	55.55556	5.970739
10	100.00000	0.00000	7.723753	60	100.00000	55.55556	6.965150
11	0.00000	11.11111	8.858674	61	0.00000	66.66667	9.553413
12	11.11111	11.11111	8.330836	62	11.11111	66.66667	8.214384
13	22.22222	11.11111	7.804410	63	22.22222	66.66667	6.779417
14	33.33333	11.11111	7.445601	64	33.33333	66.66667	5.897153
15	44.44444	11.11111	6.849630	65	44.44444	66.66667	5.656457
16	55.55556	11.11111	6.204501	66	55.55556	66.66667	5.493306
17	66.66667	11.11111	5.877048	67	66.66667	66.66667	5.755678
18	77.77778	11.11111	5.766350	68	77.77778	66.66667	6.789305
19	88.88889	11.11111	5.896489	69	88.88889	66.66667	8.020169
20	100.00000	11.11111	6.384790	70	100.00000	66.66667	9.073022
21	0.00000	22.22222	8.024319	71	0.00000	77.77778	10.230611
22	11.11111	22.22222	7.435034	72	11.11111	77.77778	9.193602
23	22.22222	22.22222	7.297732	73	22.22222	77.77778	7.918312
24	33.33333	22.22222	7.334209	74	33.33333	77.77778	7.219821
25	44.44444	22.22222	6.763175	75	44.44444	77.77778	7.035822
26	55.55556	22.22222	6.135175	76	55.55556	77.77778	7.052797
27	66.66667	22.22222	5.535829	77	66.66667	77.77778	7.701085
28	77.77778	22.22222	5.061314	78	77.77778	77.77778	9.009715
29	88.88889	22.22222	4.888268	79	88.88889	77.77778	10.390046
30	100.00000	22.22222	5.132070	80	100.00000	77.77778	11.060113
31	0.00000	33.33333	7.444836	81	0.00000	88.88889	10.066248
32	11.11111	33.33333	6.814620	82	11.11111	88.88889	9.693844
33	22.22222	33.33333	7.121809	83	22.22222	88.88889	9.213826
34	33.33333	33.33333	7.139499	84	33.33333	88.88889	9.034113
35	44.44444	33.33333	6.542567	85	44.44444	88.88889	9.030562
36	55.55556	33.33333	6.030628	86	55.55556	88.88889	9.322539
37	66.66667	33.33333	5.324213	87	66.66667	88.88889	10.322134
38	77.77778	33.33333	4.530186	88	77.77778	88.88889	11.574057
39	88.88889	33.33333	4.182009	89	88.88889	88.88889	12.458953
40	100.00000	33.33333	4.471608	90	100.00000	88.88889	12.692470
41	0.00000	44.44444	7.599202	91	0.00000	100.00000	9.908982
42	11.11111	44.44444	6.907587	92	11.11111	100.00000	9.999910
43	22.22222	44.44444	6.805385	93	22.22222	100.00000	10.087041
44	33.33333	44.44444	6.533099	94	33.33333	100.00000	10.327667
45	44.44444	44.44444	6.226250	95	44.44444	100.00000	10.677416
46	55.55556	44.44444	5.661992	96	55.55556	100.00000	11.496787
47	66.66667	44.44444	5.083970	97	66.66667	100.00000	12.601396
48	77.77778	44.44444	4.656872	98	77.77778	100.00000	13.478428
49	88.88889	44.44444	4.594427	99	88.88889	100.00000	22.827946
50	100.00000	44.44444	5.274761	100	100.00000	100.00000	13.782820

Variável simulada “Matérn 0-10-60 k=3,0”

n	X	Y	Z	n	X	Y	Z
1	0.000	0.000	6.175885	51	0.000	55.556	6.368920
2	11.111	0.000	6.897401	52	11.111	55.556	6.881635
3	22.222	0.000	7.557186	53	22.222	55.556	7.449957
4	33.333	0.000	8.122890	54	33.333	55.556	8.027841
5	44.444	0.000	8.591793	55	44.444	55.556	8.557445
6	55.556	0.000	8.964236	56	55.556	55.556	9.002985
7	66.667	0.000	9.230371	57	66.667	55.556	9.350709
8	77.778	0.000	9.406318	58	77.778	55.556	9.619924
9	88.889	0.000	9.533750	59	88.889	55.556	9.833247
10	100.000	0.000	9.623763	60	100.000	55.556	9.996125
11	0.000	11.111	6.139878	61	0.000	66.667	6.535615
12	11.111	11.111	6.863212	62	11.111	66.667	7.008986
13	22.222	11.111	7.543659	63	22.222	66.667	7.503977
14	33.333	11.111	8.144354	64	33.333	66.667	8.003785
15	44.444	11.111	8.651423	65	44.444	66.667	8.465130
16	55.556	11.111	9.062894	66	55.556	66.667	8.851099
17	66.667	11.111	9.371112	67	66.667	66.667	9.153647
18	77.778	11.111	9.583830	68	77.778	66.667	9.396059
19	88.889	11.111	9.727226	69	88.889	66.667	9.585426
20	100.000	11.111	9.820959	70	100.000	66.667	9.748348
21	0.000	22.222	6.117592	71	0.000	77.778	6.707046
22	11.111	22.222	6.811273	72	11.111	77.778	7.161279
23	22.222	22.222	7.498481	73	22.222	77.778	7.595069
24	33.333	22.222	8.116341	74	33.333	77.778	8.000947
25	44.444	22.222	8.653117	75	44.444	77.778	8.356333
26	55.556	22.222	9.107460	76	55.556	77.778	8.649660
27	66.667	22.222	9.454143	77	66.667	77.778	8.884823
28	77.778	22.222	9.702064	78	77.778	77.778	9.083590
29	88.889	22.222	9.879656	79	88.889	77.778	9.251132
30	100.000	22.222	9.994111	80	100.000	77.778	9.421650
31	0.000	33.333	6.129966	81	0.000	88.889	6.901618
32	11.111	33.333	6.775789	82	11.111	88.889	7.340448
33	22.222	33.333	7.446246	83	22.222	88.889	7.719958
34	33.333	33.333	8.078638	84	33.333	88.889	8.028543
35	44.444	33.333	8.645543	85	44.444	88.889	8.271999
36	55.556	33.333	9.128025	86	55.556	88.889	8.452057
37	66.667	33.333	9.495569	87	66.667	88.889	8.597068
38	77.778	33.333	9.761928	88	77.778	88.889	8.737307
39	88.889	33.333	9.973096	89	88.889	88.889	8.892163
40	100.000	33.333	10.118111	90	100.000	88.889	9.071727
41	0.000	44.444	6.217077	91	0.000	100.000	7.130703
42	11.111	44.444	6.793525	92	11.111	100.000	7.545080
43	22.222	44.444	7.422516	93	22.222	100.000	7.869268
44	33.333	44.444	8.049046	94	33.333	100.000	8.084951
45	44.444	44.444	8.620070	95	44.444	100.000	8.212925
46	55.556	44.444	9.100836	96	55.556	100.000	8.283275
47	66.667	44.444	9.466969	97	66.667	100.000	8.348495
48	77.778	44.444	9.741087	98	77.778	100.000	8.437221
49	88.889	44.444	9.965744	99	88.889	100.000	8.569275
50	100.000	44.444	19.33767	100	100.000	100.000	8.752368

APÊNDICE B - ROTEIROS DE ANÁLISES NO PROGRAMA R E MÓDULO GEOESTATÍSTICO **geoR**

Roteiro para variável produtividade com covariáveis

```

require(geoR) #carregando o módulo geoR
d=read.table("D:/JOELMIR/Análise Covariáveis/PCovariáveis.txt") #Arquivo com
coordenadas, variável principal e covariáveis
d=as.geodata(d,coords.col=2:3,data.col=4,covar.col=5:10,head=T) #Transforma o
arquivo "d" em um objeto no formato "geodata", data.col=4: coluna da variável
principal; covar.col=5:10: colunas de 5 a 10 com as covariáveis

#ANÁLISE DESCRITIVA
summary(d$data) # média, mediana, Q1, Q3, min, máx
var(d$data) # Variância
sd(d$data) # Desvio padrão
CV<- sd(d$data)*100/mean(d$data) # Coeficiente de variação

plot(d) #Mostra uma tela com gráficos de tendências e gráfico histograma.
boxplot(d$data,main='Produtividade') #
library(stats) #Carregando pacote "stats" que contém testes estatísticos.
shapiro.test(d$data) #Teste de normalidade de Shapiro & Wilk.

require(MASS) # Carregando o pacote MASS para a função Boxcox
boxcox(d$data~1, lambda=seq(-2,2)) # Gráfico Boxcox para verificar a necessidade de
transformação dos dados: a linha central indica o lambda ótimo, estimado por Log-
verossimilhança.

#Ajuste de curva

d.mle=likfit(d,ini=c(0.4,150),method="ML",lambda=1,trend=trend.spatial(~V5+V6+V7
+V8+V9+V10),cov.model='exp') #
d.mle

d.mlg=likfit(d,ini=c(0.4,150),lambda=1,trend=trend.spatial(~V5+V6+V7+V8+V9+V10
),method="ML",cov.model='gaus') #V5: busca a covariável V5, que está no objeto "d",
e assim por diante
d.mlg
lines(d.mle)
lines(d.mlg,lty=2)
legend(200,0.05,legend=c('Exponencial','Gaussiano'),lty=c(1
,2),lwd=c(1,2),cex=0.9)

###
##KRIGAGEM

```

```

####
bor <- read.table("D:/Joelmir/Análise Covariáveis/Bor.txt", head=F)
plot(bor)
polygon(bor)
points(d, pt.div= "quint", cex.max=1, cex.min=1)
apply(bor,2,range) #mostra mínimo e máximo das coordenadas, para determinar o grid
de interpolação
gr <- expand.grid(x=seq(39600, 41050, by=5), y=seq(37040,37715, by=5))
#monta o grid. By = espaçamento entre as interpolações. Seq(início, fim, by=distância
entre as interpolações) quanto mais próximas forem as interpolações, mas fina a grade,
melhor a qualidade do mapa.
plot(gr)
require(splancs)
gi <- polygrid(gr,bor=bor)
points(gi, pch="+", col=2)

#####
## Valores das covariáveis nos locais de predição

c1=read.geodata("D:/Joelmir/Análise Covariáveis/Dados/PENETR 75 00
10.txt",head=F,coords.col=2:3,data.col=4)
c1.mle=likfit(c1,ini=c(0.4,150),lambda=1,cov.model='exp')
kc1=krige.control(obj=c1.mle,lam=1)
c1.k=krige.conv(c1,loc=gr,krige=kc1)
#####
c2=read.geodata("D:/Joelmir/Análise Covariáveis/Dados/PENETR 75 10
20.txt",head=F,coords.col=2:3,data.col=4)
c2.mle=likfit(c2,ini=c(0.4,150),lambda=1,cov.model='exp')
kc2=krige.control(obj=c2.mle,lam=1)
c2.k=krige.conv(c2,loc=gr,krige=kc2)
#####
c3=read.geodata("D:/Joelmir/Análise Covariáveis/Dados/PENETR 75 20
30.txt",head=F,coords.col=2:3,data.col=4)
c3.mle=likfit(c3,ini=c(0.4,150),lambda=1,cov.model='exp')
kc3=krige.control(obj=c3.mle,lam=1)
c3.k=krige.conv(c3,loc=gr,krige=kc3)
#####
c4=read.geodata("D:/Joelmir/Análise Covariáveis/Dados/DENS 75 00
10.txt",head=F,coords.col=2:3,data.col=4)
c4.mle=likfit(c4,ini=c(0.4,150),lambda=1,cov.model='exp')
kc4=krige.control(obj=c4.mle,lam=1)
c4.k=krige.conv(c4,loc=gr,krige=kc4)
#####
c5=read.geodata("D:/Joelmir/Análise Covariáveis/Dados/DENS 75 10
20.txt",head=F,coords.col=2:3,data.col=4)
c5.mle=likfit(c5,ini=c(0.4,150),lambda=1,cov.model='exp')
kc5=krige.control(obj=c5.mle,lam=1)
c5.k=krige.conv(c5,loc=gr,krige=kc5)
#####
c6=read.geodata("D:/Joelmir/Análise Covariáveis/Dados/DENS 75 20

```

```

30.txt",head=F,coords.col=2:3,data.col=4)
c6.mle=likfit(c6,ini=c(0.003,50),lambda=1,cov.model='exp')
kc6=krige.control(obj=c6.mle,lam=1)
c6.k=krige.conv(c6,loc=gr,krige=kc6)
#####

KC=krige.control(obj=d.mle,type.krige="ok",trend.d=~V5+V6+V7+V8+V9+V10,trend
.l=~c1.k$predict+c2.k$predict+c3.k$predict+c4.k$predict+c5.k$predict+c6.k$predict,la
m=1)
# trend.d: recebe o valor das covariáveis nas localizações
dos dados; trend.l: recebe o valor das covariáveis nas localizações a serem preditas
d.k=krige.conv(d,loc=gr,krige=KC)

image(d.k, loc=gr, border=bor,col=gray(seq(1,0,l=3)), x.leg=c(40000,40800),
y.leg=c(36700,36800),zlim=range(d.k$predict))
2.827985 3.899321
title(main = "Variável Produtividade", font.main = 4)

####Para criar classes com tamanhos irregulares na legenda
image(d.k,loc=gr,border=bor,col=gray(seq(1,0,l=3)),x.leg=c(40000,40800),y.leg=c(36
700,36800),breaks=c(2.8,3,3.5,4),zlim=c(2.8,4))
#"Breaks" determina intervalos de [2.8;3), [3;3.5),
[3.5;4]: 3 classes
#"col" determina as cores das classes, em escala de
cinza:"l=3"cores

#ANÁLISE DE DIAGNÓSTICO
f1=0.0791 #Ef. Pepita
f2= 0.0455 #Contribuição
f3=142.2664 #Alcance
n=47 #Número de dados
#Parâmetros beta estimados
B0= 4.5641
B1= -0.1027
B2= 0.0633
B3= -0.0150
B4= -0.2885
B5= 0.0611
B6= -0.8662

## MATRIZ DE COVARIÂNCIA
C<-
varcov.spatial(coords=d$coords,dists.lowertri=NULL,cov.model="exp",nugget=f1,cov.
pars=c(f2,f3),det=FALSE,scaled=FALSE,sqrt.inv=FALSE,try.another.decomposition=
FALSE)

C=C$varcov #Retorna uma a matriz de covariância
IC=solve(C)#Inversa da matriz de covariância
X= read.table("D:/Joelmir/Análise Covariáveis/Matriz X.txt", head=T)

```

```

#Matriz de 1s e covariáveis
X=c(X)
X=cbind(X$C,X$R1,X$R2,X$R3,X$D1,X$D2,X$D3)#Matrix X com as
covariáveis

B=rbind(B0,B1,B2,B3,B4,B5,B6)
#Matriz Beta de parâmetros desconhecidos, estimados por MV
XB= X%*%B

E=d$data-XB #Erro
I=matrix(0,nr=n,nc=n)
I[row(I)==col(I)]=1 #Matriz identidade I (Derivada da
matriz de covariância com relação ao efeito pepita f1)
R=1/f2*(C - f1*I) #Derivada da matriz de covariância com relação à contribuição f2
H=dist(d$coords, method = "euclidean", diag = TRUE, upper = TRUE)
H=as.matrix(H) #Matriz nxn com as distâncias entre amostras
F=f2*((R%*%H)/f3^2) #Derivada da matriz de covariância com relação ao alcance f3
(Mod. Exponencial)
#Para gerar a matriz Delta beta:
Db=t(X)%*%IC #Matriz Delta beta (1xn)
#Para gerar as matrizes Delta fi:
Df1=(t(E))%*%IC%*%I%*%IC #
Df2=(t(E))%*%IC%*%R%*%IC
Df3=(t(E))%*%IC%*%F%*%IC
Df=rbind(Df1,Df2,Df3) #Junta as matrizes Df1,Df2,Df3 numa
matriz (3xn)
#Para gerar a matriz Delta, juntando a matriz Delta beta
com Delta fi:
D=rbind(Db,Df) # Gera uma matriz [(p+q) x n]= [(1+3) x 100]
#GERANDO AS MATRIZES L:
Lbb=-((t(X))%*%IC%*%X)
Lbf1=-t(X)%*%IC%*%I%*%IC%*%E
Lbf2=-t(X)%*%IC%*%R%*%IC%*%E
Lbf3=-t(X)%*%IC%*%F%*%IC%*%E
Lbf=cbind(Lbf1,Lbf2,Lbf3)
Lfb=t(Lbf)
Lf1f1=1/2*(sum(diag(IC%*%(I%*%IC%*%I))))+1/2*t(E)%*%IC%*%(-
I%*%IC%*%I - I%*%IC%*%I )%*%IC%*%E
Lf1f2=1/2*(sum(diag(IC%*%(I%*%IC%*%R))))+1/2*t(E)%*%IC%*%(-
I%*%IC%*%R - R%*%IC%*%I )%*%IC%*%E
Lf1f3=1/2*(sum(diag(IC%*%(I%*%IC%*%F))))+1/2*t(E)%*%IC%*%(-
I%*%IC%*%F - F%*%IC%*%I )%*%IC%*%E
Lf2f2=1/2*(sum(diag(IC%*%(R%*%IC%*%R))))+1/2*t(E)%*%IC%*%(-
R%*%IC%*%R - R%*%IC%*%R )%*%IC%*%E
Lf2f3=1/2*(sum(diag(IC%*%(R%*%IC%*%F-
(R%*%H)/f3^2))))+1/2*t(E)%*%IC%*%((R%*%H)/f3^2-R%*%IC%*%F-
F%*%IC%*%R )%*%IC%*%E
Lf3f3=1/2*(sum(diag(IC%*%(F%*%IC%*%F-
f2*(((R%*%H)/f3^3)%*%((H/f3)-
2)))))+1/2*t(E)%*%IC%*%(f2*(((R%*%H)/f3^3)%*%((H/f3))-2)-
F%*%IC%*%F-F%*%IC%*%F)%*%IC%*%E

```

```

Lff=cbind(c(Lf1f1, Lf1f2, Lf1f3),c(Lf1f2, Lf2f2, Lf2f3),c(Lf1f3,
Lf2f3, Lf3f3))
L1=cbind(Lbb,Lbf)
L2= cbind(Lfb,Lff)
L=rbind(L1,L2)
B=t(D)%*%solve(L)%*%D
#####
#Erro- padrão para o estimador de máxima verossimilhança
Kb=(t(X))%*%IC%*%X
Kf1f1=1/2*(sum(diag(IC%*%I%*%IC%*%I)))
Kf1f2=1/2*(sum(diag(IC%*%I%*%IC%*%R)))
Kf1f3=1/2*(sum(diag(IC%*%I%*%IC%*%F)))
Kf2f2=1/2*(sum(diag(IC%*%R%*%IC%*%R)))
Kf2f3=1/2*(sum(diag(IC%*%R%*%IC%*%F)))
Kf3f3=1/2*(sum(diag(IC%*%F%*%IC%*%F)))
Kf=cbind(c(Kf1f1,Kf1f2,Kf1f3),c(Kf1f2,Kf2f2,Kf2f3),c(Kf1f3,Kf2f3,Kf3f3))
K1=cbind(Kb,matrix(0,nr=7,nc=3))
K2= cbind(matrix(0,nr=3,nc=7),Kf)
Kt=rbind(K1,K2) #Matriz de informação esperada
IKt=solve(Kt) # Inversa da matriz de informação esperada, que dá os erros padrão
DIKt=diag(IKt)# Diagonal principal da matriz IKt
DP=sqrt(DIKt)# Desvio padrão dos parâmetros estimados por MV:  $\beta_0, \varphi_1, \varphi_2, \varphi_3$ 

#Gráfico para Diagnósticos:  $C_i$  vs  $i$ :
Ci=2*abs(diag(B))
plot(Ci,xlab="i",ylim=c(0,14),main='GRÁFICO  $C_i$  x  $i$  – Exponencial-MV')
abline(h=2*mean(Ci), col = "black", lty = 3) #Desenha uma linha horizontal no valor
VC; lty=3: para tipo de linha
abline(v=c(13,23,30), lty = 3)

#Gráfico para Diagnósticos:  $L_{max}$  vs  $i$ :
BA=eigen(B, symmetric=TRUE, only.values = FALSE)
Lmax=BA$vectors[44] #Retorna o autovetor, associado ao maior autovalor da matriz B
Lmax=abs(Lmax/sqrt(sum(Lmax^2))) #Autovetor Lmax normalizado e em módulo.
plot(Lmax, ylim=c(0,0.6),xlab="i",ylab="|Lmax|",main='GRÁFICO |Lmax| x  $i$  –
Exponencial-MV')
abline(h=2*mean(Lmax), col = "black", lty = 3)
abline(v=c(23,30), lty = 3)

#Gráfico de Diagnóstico LG vs  $i$ :
Gbz=t(X) %*%IC
Gf1z=t(E)%*%IC%*%I%*%IC
Gf2z=t(E)%*%IC%*%R%*%IC
Gf3z=t(E) %*%IC%*%F%*%IC
Gfz=rbind(Gf1z,Gf2z,Gf3z)
GL=X%*%(solve(t(X)%*%IC%*%X+Lbf%*%solve(Lff)%*%Lfb))%*%t(X)%*%IC+
X%*%(solve(t(X)%*%IC%*%X+Lbf%*%solve(Lff)%*%Lfb))%*%(-
Lbf%*%solve(Lff)%*%Gfz)
G=abs(diag(GL))
plot(G,xlab="i", ylim=c(0,0.8),ylab="Alavanca", main='GRÁFICO GL x  $i$  –

```

```

Exponencial-MV')
abline(h=2*mean(G), col = "black", lty = 3)
abline(v=c(23,30), col = "black", lty = 3)
LG=X%%(solve(Lbb- Lbf%%solve(Lff)%Lfb))%%(Lfb%%solve(Lff)
%%Gfz-Gbz)

#Gráfico de Diagnóstico lCd: Exponencial
h=100
dCf1=0 #Derivadas da matriz de covariância, com relação a  $\varphi_1$ 
dCf2=exp(-(h/f3))
dCf3=(h*f2*exp(-(h/f3)))/(f3)^2
dC=t(c(0,0,0,0,0,0,dCf1,dCf2,dCf3))
dC=matrix(dC) # Junção das derivadas da matriz de covariância, com relação aos 3
parâmetros ( $\varphi_1, \varphi_2, \varphi_3$ )

Cd=-t(D)%solve(L)%dC # Matriz  $\dot{C}_\delta$ , nx1, leia-se: "C delta (ponto)"
lCd= Cd/sqrt(sum(Cd^2)) #Medida de diagnóstico, com Cd normalizado.
plot(abs(lCd),xlab="i",ylab="lCd",main='GRÁFICO lCd x I – Exponencial-MV ')
abline(v=26, lty = 3)

```

```
#####
```

Roteiro para a variável fósforo e manganês

```

require(geoR) #carregando o módulo geoR
d<-read.geodata("D:/JOELMIR/Área Mapas/Análise P/Dados P-Mn.txt
",head=F,coords.col=2:3,data.col=4) #Importando os dados no formato geodata.
O argumento "head=F" é o argumento da função read.geodata, que determina que
o arquivo de dados não conterá cabeçalho.

```

```

plot(d) #Mostra uma tela com gráficos de tendências e gráfico histograma.
boxplot(d$data,main='Gráfico Boxplot') #
library(stats) #Carregando pacote "stats" que contém testes estatísticos.
shapiro.test(d$data) #Teste de normalidade de Shapiro & Wilk.

```

```

require(MASS) # Carregando o pacote MASS para a função Boxcox
boxcox(d$data~1, lambda=seq(-5,8)) # Gráfico Boxcox para verificar a
necessidade de transformação dos dados: a linha central indica o lambda ótimo,
estimado por Log-verossimilhança.

```



```
points(d,l=0,pt.div="equal",col=gray(seq(1,0,l=11)),main="Post-Plot") #Post-plot
```

```
#Ajuste de curva
```

```
d.ml=likfit(d,ini=c(.002,200),lambda=-.7,cov.model='matern',kappa=0.7)
```

```
d.ml
```

```
mean(d$data)
```

```
#ANÁLISE DE DIAGNÓSTICO
```

```
f1= 0.0046#Ef. Pepita
```

```
f2= 0.0017#Contribuição
```

```
f3= 100.9346#Alcance
```

```
m= 1.1822#Média
```

```
k=0.7 #kappa
```

```
n=46 #Número de dados
```

```
## MATRIZ DE COVARIÂNCIA
```

```
C<-
```

```
varcov.spatial(coords=d$coords,dists.lowertri=NULL,cov.model="matern",kappa=k,nugget=f1,cov.pars=c(f2,f3),det=FALSE,scaled=FALSE,sqrt.inv=FALSE,try.another.decomposition=FALSE)
```

```
C=C$varcov #Retorna uma a matriz de covariância
```

```
IC=solve(C)#Inversa da matriz de covariância
```

```
X=matrix(1,nr=n,nc=1) #Matriz de 1s (Isotrópico)
```

```
XB=matrix(m,nr=n,nc=1) #Matriz X*parâmetros B (nesse caso a média geral da variável resposta)
```

```
E=d$data-XB #Erro
```

```
I=matrix(0,nr=n,nc=n)
```

```
I[row(I)==col(I)]=1 #Matriz identidade I (Derivada da matriz de covariância com relação ao efeito pepita f1)
```

```

R=1/f2*(C - f1*I) #Derivada da matriz de covariância com relação à contribuição
f2
H=dist(d$coords, method = "euclidean", diag = TRUE, upper = TRUE)
H=as.matrix(H) #Matriz nxn com as distâncias entre amostras
#Derivada da função de Bessel de terceiro tipo e ordem k
dK=function (H, f3, k)
{
  if (is.vector(H))
    names(H) <- NULL
  if (is.matrix(H))
    dimnames(H) <- list(NULL, NULL)
  uphi <- H/f3
  uphi <- ifelse(H > 0, -1/2*(besselK(x = uphi, nu = k-1)+ besselK(x = uphi, nu
= k+1)), 1)
  uphi[H > 600 * f3] <- 0
  return(uphi)
}
F=f2*(-(1/f3)*(k*R+(1/(2^(k-1)*gamma(k)))*(H/f3)^(k+1)*dK(H,f3,k)))
#Derivada da matriz de covariância com relação ao alcance f3 (Mod. Matérn)
#Para gerar a matriz Delta beta:
Db=t(X)%*%IC #Matriz Delta beta (1xn)
#Para gerar as matrizes Delta fi:
Df1=(t(E))%*%IC%*%I%*%IC #
Df2=(t(E))%*%IC%*%R%*%IC
Df3=(t(E))%*%IC%*%F%*%IC
Df=rbind(Df1,Df2,Df3) #Junta as matrizes Df1,Df2,Df3
numa matriz (3xn)
#Para gerar a matriz Delta, juntando a matriz Delta
beta com Delta fi:
D=rbind(Db,Df) # Gera uma matriz [(p+q) x n]= [(1+3) x
66]

```

```
#GERANDO AS MATRIZES L:
```

```
Lbb=-((t(X))%*%IC%*%X)
```

```
Lbf1=-t(X)%*%IC%*%I%*%IC%*%E
```

```
Lbf2=-t(X)%*%IC%*%R%*%IC%*%E
```

```
Lbf3=-t(X)%*%IC%*%F%*%IC%*%E
```

```
Lbf=cbind(Lbf1, Lbf2, Lbf3)
```

```
Lfb=t(Lbf)
```

```
Lf1f1=1/2*(sum(diag(IC%*%(I%*%IC%*%I))))+1/2*t(E)%*%IC%*%(-I%*%IC%*%I-I%*%IC%*%I)%*%IC%*%E
```

```
Lf1f2=1/2*(sum(diag(IC%*%(I%*%IC%*%R))))+1/2*t(E)%*%IC%*%(-I%*%IC%*%R-R%*%IC%*%I)%*%IC%*%E
```

```
Lf1f3=1/2*(sum(diag(IC%*%(I%*%IC%*%F))))+1/2*t(E)%*%IC%*%(-I%*%IC%*%F-F%*%IC%*%I)%*%IC%*%E
```

```
Lf2f2=1/2*(sum(diag(IC%*%(R%*%IC%*%R))))+1/2*t(E)%*%IC%*%(-R%*%IC%*%R-R%*%IC%*%R)%*%IC%*%E
```

```
dR.f3=-((1/f3)*(k*R+(1/(2^(k-1))*gamma(k)))*(H/f3)^(k+1)*dK(H,f3,k))
```

```
#Derivada da matriz R com relação a f3 (mod. Matérn)
```

```
Lf2f3=1/2*(sum(diag(IC%*%(R%*%IC%*%F-(dR.f3)))))+1/2*t(E)%*%IC%*%((dR.f3)-R%*%IC%*%F-F%*%IC%*%R)%*%IC%*%E
```

```
#Derivada segunda da função de Bessel de terceiro tipo e ordem k
```

```
d2K=function(H, f3, k)
```

```
{
```

```
  if (is.vector(H))
```

```
    names(H) <- NULL
```

```
  if (is.matrix(H))
```

```
    dimnames(H) <- list(NULL, NULL)
```

```
  uphi <- H/f3
```

```
  uphi <- ifelse(H > 0, 1/4*(besselK(x = uphi, nu = k-2)+ 2*besselK(x = uphi, nu = k)+ besselK(x = uphi, nu = k+2)), 1)
```

```

uphi[H > 600 * f3] <- 0
return(uphi)
}
d2R.f3=((k*(k+1)*R)/f3^2+(1/(f3^2*2^(k-
1)*gamma(k)))*(H/f3)^(k+1)*(2*(k+1)*dK(H,f3,k)+(H/f3)*d2K(H,f3,k)))
#Derivada segunda da matriz R com relação a f3f3 (mod. Matérn)
Lf3f3=1/2*(sum(diag(IC%%(F%%IC%%F-
f2*(d2R.f3)))))+1/2*t(E)%%IC%%(f2*(d2R.f3)-F%%IC%%F-
F%%IC%%F)%%IC%%E
Lff=cbind(c(Lf1f1, Lf1f2, Lf1f3),c(Lf1f2, Lf2f2, Lf2f3),c(Lf1f3,
Lf2f3, Lf3f3))
L1=cbind(Lbb,Lbf)
L2= cbind(Lfb,Lff)
L=rbind(L1,L2)
B=t(D)%%solve(L)%%D
#Erro- padrão para o estimador de máxima verossimilhança
Kb=(t(X))%%IC%%X
Kf1f1=1/2*(sum(diag(IC%%I%%IC%%I)))
Kf1f2=1/2*(sum(diag(IC%%I%%IC%%R)))
Kf1f3=1/2*(sum(diag(IC%%I%%IC%%F)))
Kf2f2=1/2*(sum(diag(IC%%R%%IC%%R)))
Kf2f3=1/2*(sum(diag(IC%%R%%IC%%F)))
Kf3f3=1/2*(sum(diag(IC%%F%%IC%%F)))
Kf=cbind(c(Kf1f1,Kf1f2,Kf1f3),c(Kf1f2,Kf2f2,Kf2f3),c(Kf1f3,Kf2f3,Kf3f3))
K1=cbind(Kb,0,0,0)
K2= cbind(c(0,0,0),Kf)
Kt=rbind(K1,K2) #Matriz de informação esperada
IKt=solve(Kt) # Inversa da matriz de informação esperada, que dá os erros
padrão
DIKt=diag(IKt)# Diagonal principal da matriz IKt
DP=sqrt(DIKt)# Desvio padrão dos parâmetros estimados por MV:  $\beta_0, \varphi_1, \varphi_2, \varphi_3$ 

```

#Gráfico de Diagnóstico C_i vs i :

```
Ci=2*abs(diag(B))
```

```
plot(Ci,xlab="i",main='GRÁFICO Ci x i - Matérn k=0,7')
```

```
abline(h=2*mean(Ci), lty = 3)#Desenha uma linha horizontal; lty=3: para tipo de linha
```

```
abline(v=c(1,26,45), lty = 3)
```

#Gráfico de Diagnóstico L_{max} vs i :

```
BA=eigen(B, symmetric=TRUE, only.values = FALSE)
```

```
Lmax=BA$vectors[,46] #Retorna o autovetor, associado ao maior autovalor da matriz B
```

```
Lmax=abs(Lmax/sqrt(sum(Lmax^2))) #Autovetor Lmax normalizado e em módulo.
```

```
plot(Lmax,xlab="i",ylab="|Lmax|",main='GRÁFICO |Lmax| x i - Matérn k=0,7')
```

```
abline(h=2*mean(Lmax), lty = 3)
```

```
abline(v=c(1,26,45), lty = 3)
```

#Gráfico de Diagnóstico LG vs i :

```
Gf1z=t(E)%*%IC%*%I%*%IC
```

```
Gf2z=t(E)%*%IC%*%R%*%IC
```

```
Gf3z=t (E) %*%IC%*%F%*%IC
```

```
Gfz=rbind(Gf1z, Gf2z, Gf3z)
```

```
GL=X%*(solve(t(X)%*%IC%*%X+Lbf%*%solve(Lff)%*%Lfb))%*%t(X)%*%IC+X%*(solve(t(X)%*%IC%*%X+Lbf%*%solve(Lff)%*%Lfb))%*%(-Lbf%*%solve(Lff)%*%Gfz)
```

```
G=abs(diag(GL))
```

```
G=abs(G/sqrt(sum(G^2)))#Teste
```

```
plot(G,xlab="i",ylab="Alavanca",main='GRÁFICO GL x i - Matérn k=0,7')
```

```
abline(v=0.04, lty = 3)
```

```
#####
```

```
Gbz=t(X) %*%IC
```

```

LG=X%*(solve(Lbb-      Lbf%*solve(Lff)%*Lfb))%*(t(Lfb)%*solve(Lff)
%*Gfz-Gbz)
GL=diag(GL)
plot(GL,xlab="i",ylab="Alavanca",main='GRÁFICO Alavanca x i')

#Gráfico de Diagnóstico ICd: Matérn (k=)

h=100
dCf1=0 #Derivadas da função de covariância, com relação a  $\varphi_1$ 
Km=function (h, f3, k)
{
  if (is.vector(h))
    names(h) <- NULL
  if (is.matrix(h))
    dimnames(h) <- list(NULL, NULL)
  uphi <- h/f3
  uphi <- ifelse(h > 0, besselK(x = uphi, nu = k), 1)
  uphi[h > 600 * f3] <- 0
  return(uphi)
}

dCf2=(2^(k-1)* gamma(k))^(-1)*(h/f3)^k* Km(h, f3, k) #K: função de Bessel de
terceiro tipo e ordem k

dKm=function (h, f3, k)
{
  if (is.vector(h))
    names(h) <- NULL
  if (is.matrix(h))
    dimnames(h) <- list(NULL, NULL)
  uphi <- h/f3
  uphi <- ifelse(h > 0, -1/2*(besselK(x = uphi, nu = k-1)+ besselK(x = uphi, nu =
k+1)), 1)
  uphi[h > 600 * f3] <- 0
  return(uphi)
}

dCf3=f2*(2^(k-1)*gamma(k))^(-1)*h^k*(((k)/f3^(k+1))*Km(h,f3,k)+(h/f3)^k*dKm(h,
f3, k))

dC=t(c(0,dCf1,dCf2,dCf3))
dC=matrix(dC) # Junção das derivadas da função de covariância, com relação aos 3
parâmetros ( $\varphi_1, \varphi_2, \varphi_3$ )

Cd=-t(D)%*solve(L)%*dC # Matriz  $\hat{C}_\delta$ , nx1, leia-se: "C delta (ponto)"
ICd= Cd/sqrt(sum(Cd^2)) #Medida de diagnóstico, com Cd normalizado.
plot(abs(ICd),xlab="i",ylab="ICd",main='GRÁFICO ICd x I – Matérn k=0,7')
abline(v=10, lty = 3)

```

ANEXO

**ANEXO A - TABELA DO COEFICIENTE DE CORRELAÇÃO r DE
FILLIBEN**

$n \backslash \alpha$,000	,005	,010	,025	,050	,750	,900	,950	,975	,990	,995
3	,866	,867	,869	,872	,879	,991	,999	1,000	1,000	1,000	1,000
4	,784	,813	,822	,845	,868	,979	,992	,996	,998	,999	1,000
5	,726	,803	,822	,855	,879	,977	,988	,992	,995	,997	,998
6	,683	,818	,835	,868	,890	,977	,986	,990	,993	,996	,997
7	,648	,828	,847	,876	,899	,978	,986	,990	,992	,995	,996
8	,619	,841	,859	,886	,905	,979	,986	,990	,992	,995	,996
9	,595	,851	,868	,893	,912	,980	,987	,990	,992	,994	,995
10	,574	,860	,876	,900	,917	,981	,987	,990	,992	,994	,995
11	,556	,868	,883	,906	,922	,982	,988	,990	,992	,994	,995
12	,539	,875	,889	,912	,926	,982	,988	,990	,992	,994	,995
13	,525	,882	,895	,917	,931	,983	,988	,991	,993	,994	,995
14	,512	,888	,901	,921	,934	,984	,989	,991	,993	,994	,995
15	,500	,894	,907	,925	,937	,984	,989	,991	,993	,994	,995
16	,489	,899	,912	,928	,940	,985	,989	,991	,993	,994	,995
17	,478	,903	,916	,931	,942	,986	,990	,992	,993	,994	,995
18	,469	,907	,919	,934	,945	,986	,990	,992	,993	,995	,995
19	,460	,909	,923	,937	,947	,987	,990	,992	,993	,995	,995
20	,452	,912	,925	,939	,950	,987	,991	,992	,994	,995	,995
21	,445	,914	,928	,942	,952	,987	,991	,993	,994	,995	,996
22	,437	,918	,930	,944	,954	,988	,991	,993	,994	,995	,996
23	,431	,922	,933	,947	,955	,988	,991	,993	,994	,995	,996
24	,424	,926	,936	,949	,957	,988	,992	,993	,994	,995	,996
25	,418	,928	,937	,95	,958	,989	,992	,993	,994	,995	,996
26	,412	,930	,939	,952	,959	,989	,992	,993	,994	,995	,996
27	,407	,932	,941	,953	,960	,989	,992	,994	,995	,995	,996
28	,402	,934	,943	,955	,962	,990	,992	,994	,995	,995	,996
29	,397	,937	,945	,956	,962	,990	,992	,994	,995	,995	,996
30	,392	,938	,947	,957	,964	,990	,993	,994	,995	,996	,996
31	,388	,939	,948	,958	,965	,990	,993	,994	,995	,996	,996
32	,383	,939	,949	,959	,966	,990	,993	,994	,995	,996	,996
33	,379	,940	,950	,960	,967	,991	,993	,994	,995	,996	,996
34	,375	,941	,951	,960	,967	,991	,993	,994	,995	,996	,996
35	,371	,943	,952	,961	,968	,991	,993	,995	,995	,996	,997
36	,367	,945	,953	,962	,968	,991	,994	,995	,996	,996	,997
37	,364	,947	,955	,962	,969	,991	,994	,995	,996	,996	,997
38	,360	,948	,956	,964	,970	,992	,994	,995	,996	,996	,997
39	,357	,949	,957	,965	,971	,992	,994	,995	,996	,996	,997
40	,354	,949	,958	,966	,972	,992	,994	,995	,996	,996	,997
41	,351	,950	,958	,967	,972	,992	,994	,995	,996	,996	,997
42	,348	,951	,959	,967	,973	,992	,994	,995	,996	,997	,997
43	,345	,953	,959	,967	,973	,992	,994	,995	,996	,997	,997
44	,342	,954	,960	,968	,973	,992	,994	,995	,996	,997	,997
45	,339	,955	,961	,969	,974	,993	,994	,995	,996	,997	,997
46	,336	,956	,962	,969	,974	,993	,995	,995	,996	,997	,997
47	,334	,956	,963	,970	,974	,993	,995	,995	,996	,997	,997
48	,331	,957	,963	,970	,975	,993	,995	,996	,996	,997	,997
49	,329	,957	,964	,971	,975	,993	,995	,996	,996	,997	,997
50	,326	,959	,965	,972	,977	,993	,995	,996	,996	,997	,997
55	,315	,962	,967	,974	,978	,994	,995	,996	,997	,997	,997
60	,305	,965	,970	,976	,980	,994	,995	,996	,997	,997	,988
65	,296	,967	,972	,977	,981	,994	,996	,996	,997	,997	,988
70	,288	,969	,974	,978	,982	,995	,996	,997	,997	,998	,988
75	,281	,971	,975	,979	,983	,995	,996	,997	,997	,998	,988
80	,274	,973	,976	,980	,984	,995	,996	,997	,997	,998	,988
85	,268	,974	,977	,981	,985	,995	,997	,997	,997	,998	,988
90	,263	,976	,978	,982	,985	,996	,997	,997	,998	,998	,988
95	,257	,977	,979	,983	,986	,996	,997	,997	,998	,998	,988
100	,252	,979	,981	,984	,987	,996	,997	,998	,998	,998	,988