



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

ROSANGELA APARECIDA BOTINHA ASSUMPÇÃO

**INFLUÊNCIA LOCAL EM MODELOS GEOESTATÍSTICOS T-STUDENT
COM APLICAÇÕES A DADOS AGRÍCOLAS**

CASCABEL
2010

ROSANGELA APARECIDA BOTINHA ASSUMPÇÃO

**INFLUÊNCIA LOCAL EM MODELOS GEOESTATÍSTICOS T-STUDENT
COM APLICAÇÕES A DADOS AGRÍCOLAS**

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

Orientador: Prof. Dr. Miguel Angel Uribe Opazo

Co-Orientador: Prof. Dr. Manuel Galea

**CASCABEL
2010**

ROSANGELA APARECIDA BOTINHA ASSUMPCÃO

INFLUÊNCIA LOCAL EM MODELOS GEOESTATÍSTICOS T-STUDENT COM APLICAÇÕES A DADOS AGRÍCOLAS

Tese apresentada ao Programa de Pós-Graduação *"Stricto Sensu"* em Engenharia Agrícola em cumprimento parcial aos requisitos para a obtenção do título de doutor 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. Dr. Paulo Justiniano Ribeiro Junior
Departamento de Estatística - UFPR

Prof. Dr. Victor Leiva Sanchez
Departamento de Estatística - Universidade de
Valparaíso, UV - Chile

Prof. Dr. Ronny Obed Vallejos Arriagada
Departamento de Matemática - Universidade
Técnica Federico Santa Maria - Chile.

Prof. Dr. Erivelto Mercante
Centro de Ciências Exatas e Tecnológicas
UNIOESTE

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

Cascavel, 16 dezembro 2010.

BIOGRAFIA

Nome: Rosangela Aparecida Botinha Assumpção

Ano de nascimento: 1971

Naturalidade: Perdões - MG

Curso de graduação: Licenciatura em Matemática - Unioeste, 1999

Mestrado: Engenharia Agrícola - Unioeste, 2004

Vínculos empregatícios anteriores: FAG - Faculdade Assis Gurgacz - Cascavel Pr, de agosto de 2001 a dezembro de 2008.

Atualmente: Bolsista Capes

Aos meus filhos, Renan e Roxane, e ao meu esposo Rono, sempre incentivador e companheiro.

AGRADECIMENTOS

A Deus, por estar comigo em todos os momentos me iluminando;

Aos meus pais, por me ensinarem a ser persistente e me mostrarem que, se acreditarmos e trabalharmos, então os caminhos se abrem;

Ao meu orientador, Prof. Dr. Miguel Angel Uribe Opazo, pelos ensinamentos, amizade e paciência, e ainda pela oportunidade de trabalharmos juntos todos esses anos;

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

Aos colegas do Grupo de Geoestatística Aplicada, em especial ao Joelmir, com o auxílio ao software R, e a Naymara, com o L^AT_EX;

À minha colega Luciana, pelo apoio, incentivo e auxílio;

A todos os professores, mestres sensíveis às dificuldades e limitações de seus alunos; orientadores capazes na facilitação do processo ensino-aprendizagem.

À UNIOESTE, por proporcionar meu aperfeiçoamento profissional;

À CAPES, pelo apoio financeiro, concedendo-me a bolsa de estudos;

E a todos que de alguma forma me auxiliaram na realização desse trabalho e que não foram citados:

OBRIGADA!!!

Sumário

Lista de Figuras	viii
Lista de Tabelas	x
RESUMO	xiii
ABSTRACT	xiv
1 INTRODUÇÃO	1
2 REVISÃO BIBLIOGRÁFICA	3
2.1 Distribuição <i>t</i> -student	3
2.1.1 Definição	4
2.1.2 Propriedades	4
2.2 Modelos espaciais lineares	8
2.2.1 A teoria das variáveis regionalizadas	8
2.2.2 Características da variável regionalizada	8
2.2.3 Modelo espacial linear	9
2.2.4 Parâmetros do semivariograma	11
2.2.5 Modelos teóricos de semivariograma, funções de covariância e de correlação	12
2.3 Análise de Diagnóstico	15
2.3.1 Influência local	15
2.4 Krigagem	16
2.4.1 Krigagem ordinária	17
2.4.2 Krigagem com tendência	18
2.5 Métodos de validação do modelo	19
2.5.1 Critério de Akaike	20
2.5.2 Validação cruzada	20
2.5.3 Valor máximo da função verossimilhança	21
2.6 Méridas para a comparação de mapas	21
3 MATERIAL E MÉTODOS	24

3.1	Estimação de parâmetros por máxima verossimilhança para o processo t -student utilizando o algoritmo EM	24
3.2	Estimação de parâmetros por máxima verossimilhança para o processo t -student utilizando o algoritmo de Newton-Raphson (NR)	26
3.3	Estimação do erro padrão das estimativa dos parâmetros $\hat{\theta}$	27
3.4	Influência local	27
3.4.1	Influência local no afastamento da verossimilhança usual (<i>likelihood displacement</i>)	27
3.4.2	Influência local na matriz de covariância usual	30
3.4.3	Influência local no preditor linear usual	31
3.4.4	Influência local nos valores preditos usual (<i>generalized leverage</i>)	32
3.5	Influência local segundo a proposta do <i>Q</i> -afastamento	33
3.5.1	Curvatura normal conformal	35
3.5.2	Influência local no afastamento da verossimilhança utilizando o <i>Q</i> -afastamento da função verossimilhança completa (<i>Q-displacement</i>)	35
3.5.3	Influência local na matriz de covariância utilizando o <i>Q</i> -afastamento da verossimilhança completa (<i>Q-displacement</i>)	36
3.5.4	Influência local no preditor linear utilizando o <i>Q</i> -afastamento da verossimilhança completa (<i>Q-displacement</i>)	37
3.5.5	Influência local nos valores preditos utilizando o <i>Q</i> -afastamento da verossimilhança completa (<i>Q-displacement</i>)	38
3.6	Simulação de processos espaciais	38
3.7	Perturbação dos dados simulados	39
3.8	Estudo com dados experimentais	40
3.9	Recursos computacionais	41
4	RESULTADOS E DISCUSSÃO	42
4.1	Estudos de influência local aplicados a dados simulados	42
4.2	Estudo do nutriente potássio	57
4.2.1	Análise espacial da influência local utilizando a técnica do <i>Q</i> -afastamento	57
4.2.2	Análise espacial da influência local utilizando o afastamento da verossimilhança usual	63
4.3	Estudo da produtividade de soja com covariáveis	69
4.3.1	Análise espacial	71
4.3.1.1	Análise espacial da influência local pelo <i>Q</i> -afastamento	72
4.3.1.2	Análise espacial utilizando o afastamento da verossimilhança usual	78

5 CONCLUSÕES	86
6 CONSIDERAÇÕES FINAIS	87
BIBLIOGRAFIA	88
Apêndice A - Demonstrações da obtenção das matrizes de informação.....	96
A.1 Obtenção da matriz de informação usual do modelo t -student n -variado	96
A.1.1 Função escore	96
A.1.2 Matriz de informação observada usual (L)	97
A.1.3 Matriz delta usual	98
A.2 Obtenção da matriz de informação \ddot{Q} do modelo t -student n -variado	99
A.2.1 Matriz de informação \ddot{Q}	99
A.2.2 Matriz $\Delta\omega$ obtida a partir de $Q(\theta, \omega \hat{\theta})$	100
A.3 Matriz de informação esperada de Fisher	100
Apêndice B - Roteiros para análise no programa R e módulo geoestatístico geoR	107
Anexo A - Dados.....	129

Lista de Figuras

Figura 1	Direções utilizadas no estudo da anisotropia da variável regionalizada.	9
Figura 2	Relação entre a covariância espacial $C(h)$ e a semivariância $\gamma(h)$	12
Figura 3	Representação gráfica do enfoque de influência local (Verbeke e Molenberghs, 2000).	28
Figura 4	<i>Boxplot</i> dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 3$	42
Figura 5	<i>Boxplot</i> dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 5$	42
Figura 6	<i>Boxplot</i> dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 7$	43
Figura 7	<i>Boxplot</i> dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 10$	43
Figura 8	<i>Boxplot</i> dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 100$	43
Figura 9	<i>Boxplot</i> dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 1000$	44
Figura 10	Gráficos de diagnóstico C_i e C_{Qi} para dados simulados com grau de liberdade $v = 3$, $v = 5$ e $v = 7$	47
Figura 11	Gráficos de diagnóstico C_i e C_{Qi} para dados simulados com grau de liberdade $v = 10$, $v = 100$ e $v = 1000$	48
Figura 12	Gráficos de diagnóstico $ l_{max} $ e $ l_{Qmax} $ para dados simulados com grau de liberdade $v = 3$, $v = 5$ e $v = 7$	49
Figura 13	Gráficos de diagnóstico $ l_{max} $ e $ l_{Qmax} $ para dados simulados com grau de liberdade $v = 10$, $v = 100$ e $v = 1000$	50
Figura 14	Gráficos de diagnóstico l_{Ch} e l_{QCh} para dados simulados com grau de liberdade $v = 3$, $v = 5$ e $v = 7$	51
Figura 15	Gráficos de diagnóstico l_{Ch} e l_{QCh} para dados simulados com grau de liberdade $v = 10$, $v = 100$ e $v = 1000$	52
Figura 16	Gráficos de diagnóstico l_p e l_{Qp} para dados simulados com grau de liberdade $v = 3$, $v = 5$ e $v = 7$	53
Figura 17	Gráficos de diagnóstico l_p e l_{Qp} para dados simulados com grau de liberdade $v = 10$, $v = 100$ e $v = 1000$	54
Figura 18	Gráficos de diagnóstico LG e LG_Q para dados simulados com grau de liberdade $v = 3$, $v = 5$ e $v = 7$	55

Figura 19	Gráficos de diagnóstico LG e LG_Q para dados simulados com grau de liberdade $v = 10$, $v = 100$ e $v = 1000$.	56
Figura 20	<i>Boxplot</i> e croqui da área experimental.	57
Figura 21	Gráficos de diagnóstico de influência local C_{Qi} e $ l_{Qmax} $ construídos com os dados da variável potássio.	59
Figura 22	Gráficos de diagnóstico de influência local l_{QCh} e l_{Qp} construídos com os dados da variável potássio.	59
Figura 23	Gráficos de diagnóstico de influência local LG_Q construídos com os dados da variável potássio.	60
Figura 24	Mapas de contorno para a variável potássio com e sem as observações influentes utilizando o algoritmo EM na estimação dos parâmetros.	62
Figura 25	Gráficos de diagnóstico de influência local C_i e $ l_{max} $ construídos com os dados da variável potássio.	65
Figura 26	Gráficos de diagnóstico de influência local l_{Ch} e l_p construídos com os dados da variável potássio.	65
Figura 27	Gráfico de diagnóstico de influência local LG construídos com os dados da variável potássio.	66
Figura 28	Mapas de contorno para a variável potássio com e sem as observações influentes, utilizando o algoritmo de NR na estimação dos parâmetros.	68
Figura 29	<i>Boxplot</i> da produtividade e croqui da área experimental.	71
Figura 30	Scaterplot da produtividade.	71
Figura 31	Gráficos de diagnóstico de influência local C_{Qi} e $ l_{Qmax} $ para a produtividade de soja	73
Figura 32	Gráficos de diagnóstico de influência local l_{QCh} e l_{Qp} para a produtividade de soja	74
Figura 33	Gráfico de diagnóstico de influência local LG_Q para a produtividade de soja	74
Figura 34	Mapa de produtividade em $t \text{ ha}^{-1}$ com e sem as observações influentes, utilizando o algoritmo EM na estimação dos parâmetros.	77
Figura 35	Gráficos de diagnóstico de influência local C_i e $ l_{max} $ para a produtividade de soja	80
Figura 36	Gráficos de diagnóstico de influência local l_{Ch} e l_p para a produtividade de soja	80
Figura 37	Gráfico de diagnóstico de influência local LG para a produtividade de soja	81
Figura 38	Mapa de produtividade em $t \text{ ha}^{-1}$ com e sem as observações influentes, utilizando o algoritmo de NR na estimação dos parâmetros.	84

Lista de Tabelas

Tabela 1	Identificação dos pontos discrepantes nos dados simulados	44
Tabela 2	Parâmetros espaciais estimados por meio do algoritmo EM assumindo-se distribuição <i>t</i> -student para os dados simulados sem e com perturbação	45
Tabela 3	Parâmetros espaciais estimados por meio do algoritmo de NR assumindo-se distribuição <i>t</i> -student para os dados simulados sem e com perturbação	45
Tabela 4	Parâmetros espaciais estimados por meio do algoritmo EM assumindo-se distribuição <i>t</i> -student com grau de liberdade igual a $v=5$ para a variável potássio	58
Tabela 5	Resultados da validação cruzada para a variável potássio	58
Tabela 6	Análise estatística descritiva para a variável potássio [$cmol\ dm^{-3}$], com dados totais coletados (K) e retirando-se a observação 6 (K-6)	60
Tabela 7	Parâmetros espaciais estimados por meio do algoritmo EM antes e após a eliminação do valor influente da variável potássio	61
Tabela 8	Resultados da validação cruzada para a variável potássio após a eliminação do valor influente (K-6)	61
Tabela 9	Matriz de erros entre os mapas temáticos K e K-6 da variável potássio [$cmol\ dm^{-3}$] em número de <i>pixels</i>	63
Tabela 10	Medidas obtidas a partir da matriz de erros entre os mapas temáticos K e K-6 da variável potássio [$cmol\ dm^{-3}$]	63
Tabela 11	Parâmetros espaciais estimados por meio do algoritmo NR assumindo-se distribuição <i>t</i> -student com grau de liberdade $v=5$ para a variável potássio	64
Tabela 12	Resultados da validação cruzada para a variável potássio	64
Tabela 13	Análise estatística descritiva para a variável potássio [$cmol\ dm^{-3}$], com dados totais coletados (K) e retirando-se as observações 5 e 6 (K-5-6)	66
Tabela 14	Parâmetros espaciais estimados por meio do algoritmo EM antes e após a eliminação do valor influente da variável potássio	67
Tabela 15	Resultados da validação cruzada para a variável potássio após a eliminação do valor influente (K-5-6)	67
Tabela 16	Matriz de erros entre os mapas temáticos K e K-5-6 da variável potássio [$cmol\ dm^{-3}$] em número de <i>pixels</i>	68
Tabela 17	Medidas obtidas a partir da matriz de erros entre os mapas temáticos K e K-5-6 da variável potássio [$cmol\ dm^{-3}$]	69
Tabela 18	Estatísticas descritivas para a variável Produtividade da soja (Prod)[t]	

ha ^{-1]} , Resistência do solo a penetração [Mpa] e Densidade do solo [mg m ⁻³]	70
Tabela 19 Parâmetros espaciais estimados por meio do algoritmo EM assumindo-se distribuição t-student com grau de liberdade 3 para a variável produtividade	72
Tabela 20 Resultados da validação cruzada para os parâmetros estimados para a variável produtividade	73
Tabela 21 Estatísticas descritivas para as variáveis Prod e Prod-13-19	75
Tabela 22 Parâmetros $\hat{\beta}$ estimados pelo algoritmo EM para as variáveis Prod e Prod-13-19	75
Tabela 23 Parâmetros espaciais estimados pelo algoritmo EM para as variáveis Prod e Prod-13-19	76
Tabela 24 Resultados da validação cruzada para a variável Prod-13-19	76
Tabela 25 Porcentagem que cada classe da legenda representa nos mapas da variável	77
Tabela 26 Matriz de erros entre as classes da variável Prod e Prod-13-19 [$t\ ha^{-1}$] em número de <i>pixels</i>	78
Tabela 27 Medidas obtidas a partir da matriz de erros entre as classes da variável Prod e Prod-13-19 $t\ ha^{-1}$	78
Tabela 28 Parâmetros espaciais estimados por meio do algoritmo de NR assumindo-se distribuição t-student com grau de liberdade $v=3$ para a variável produtividade	79
Tabela 29 Resultados da validação cruzada para os parâmetros estimados pelo algoritmo de NR para a variável produtividade	79
Tabela 30 Estatísticas descritivas para as variáveis Prod e Prod-20-24	81
Tabela 31 Parâmetros $\hat{\beta}$ estimados pelo algoritmo NR para as variáveis Prod e Prod-20-24	82
Tabela 32 Parâmetros espaciais estimados pelo algoritmo EM para as variáveis Prod e Prod-20-24	82
Tabela 33 Resultados da validação cruzada para a variável Prod-20-24	83
Tabela 34 Porcentagem que cada classe da legenda representa nos mapas da variável	84
Tabela 35 Matriz de erros entre as classes da variável Prod e Prod-20-24 [$t\ ha^{-1}$] em número de <i>pixels</i>	84
Tabela 36 Medidas obtidas a partir da matriz de erros entre as classes da variável Prod e Prod-20-24 $t\ ha^{-1}$	85
Tabela 37 Dados experimentais do nutriente potássio [$cmol\ dm^{-3}$]	129
Tabela 38 Dados experimentais da produtividade [$t\ ha^{-1}$]	130
Tabela 39 Dados experimentais das covariáveis Resistência do solo a penetração	

[Mpa]	131
Tabela 40 Dados experimentais das covariáveis Densidade [$mg\ m^{-3}$]	132

RESUMO

INFLUÊNCIA LOCAL EM MODELOS GEOESTATÍSTICOS T-STUDENT COM APLICAÇÕES A DADOS AGRÍCOLAS

A presença de observações discrepantes torna imprópria a análise do processo gaussiano, sendo assim, como é encontrado na literatura, esse processo deve ser substituído por modelos da classe das distribuições simétricas, tal como a distribuição *t*-student, que incorpora parâmetros adicionais para reduzir a influência dos pontos discrepantes. Neste trabalho, assumiu-se que o processo apresenta distribuição *t*-student *n*-variada. Essa distribuição tem como parâmetro adicional o grau de liberdade v , que aqui considerou-se fixo. Dessa forma, desenvolveu-se o algoritmo EM e o algoritmo de NR para a estimação dos parâmetros da estrutura de dependência espacial e do modelo espacial linear. Após a estimativa dos parâmetros, utilizou-se duas técnicas de diagnósticos de influência local, ambas com o intuito de avaliar a qualidade do ajuste do modelo pelas suposições feitas e pela robustez dos resultados das estimativas quando há perturbações no modelo ou nos dados. A primeira técnica, denominada "usual", já utilizada por diversos autores, avalia o afastamento da verossimilhança pela função do logaritmo da verossimilhança e a segunda técnica que aqui apresentamos propõe a análise de influência local pelo Q-afastamento da função de verossimilhança para dados completos. Essas técnicas permitiram verificar a influência no afastamento da verossimilhança, na matriz de covariância, no preditor linear e nos valores preditos por meio da análise gráfica. Para ilustrar a aplicação da técnica usual e da nossa proposta, realizou-se a análise de dados simulados e dados reais provenientes de experimentos agrícolas.

PALAVRAS-CHAVE: Geoestatística, Algoritmo EM, Máxima verossimilhança.

ABSTRACT

LOCAL INFLUENCE IN GEOESTATISTIC T-STUDENT MODELS APPLIED TO AGRICULTURAL DATA

The presence of inconsistent observations make it improper to consider the gaussian process, as it is found in the literature. This process should be replaced by models of the symmetric distribution classes, such as the t -student distribution, which incorporates additional parameters to reduce the influence of inconsistent points. This work has developed the EM algorithm for estimating the structure of the spatial dependence of the parameters and of the spatial linear model, assuming that the process shows t -student n -varied distribution. This distribution has the degree of freedom v as the additional parameter, which has been considered to be fixed in this research. Techniques to diagnose influence are used after the estimation of parameters, in order to assess the quality of the adjustment of the model by the assumptions made and for the robustness of the results of the estimates when there are disturbances in the model or data. In the present work, diagnostic techniques for the assessment of local influence in linear spatial models have been developed, considering the process with t -student n -varied distribution. The usual diagnostic technique evaluates the withdrawing of the likelihood rate by the function of the likelihood logarithm. In this proposal, in addition to considering the usual technique, we use the withdrawing of the likelihood by Q-displacement of the complete likelihood. The application of the usual technique and of the one proposed here are illustrated through the analyses of both simulated and real data, provenient of agricultural experiments.

KEYWORDS: Geostatistics, EM Algorithm, Maximum Likelihood.

1 INTRODUÇÃO

O Brasil apresenta grande potencial de crescimento para sua produção agrícola, pois conta com clima favorável, que possibilita duas ou mais safras por ano, grandes extensões de áreas agricultáveis ainda não aproveitadas, disponibilidade de água, demanda mundial por alimentos em crescimento, grande potencial de aumento no consumo interno e, acima de tudo, da elevada tecnologia utilizada no campo.

O avanço dessa tecnologia se deu a partir do momento em que satélites, computadores e sensores passaram a auxiliá-la. Surgiu, então, um novo sistema de produção que, há alguns anos, já é utilizada pelos agricultores de países de tecnologia avançada, chamado de *Precision Agriculture* ou *Precision Farming*, sendo que no Brasil é chamado de Agricultura de Precisão. Este sistema resgata a capacidade de conhecer cada metro quadrado da lavoura, que foi perdida à medida em que as áreas cultivadas foram crescendo.

A geoestatística tem sido relatada como ferramenta eficiente de suporte para o uso efetivo da tecnologia de agricultura de precisão, garantindo a confiabilidade dessa técnica, pois possibilita a construção de mapas que caracterizam a variabilidade espacial e temporal dos atributos do solo, em suas características químicas e físicas correlacionadas à produtividade, permite detalhar os fatores de produção, definir a aplicação localizada de insumos e o manejo diferenciado das culturas.

Segundo Landim (2006), duas são as ferramentas fundamentais dos métodos geostatísticos: o semivariograma, por possibilitar a análise do grau de dependência espacial entre amostras dentro de um campo experimental, e a krigagem, por ser um processo de interpolação dos valores de variáveis distribuídas no espaço e/ou no tempo, que usa informações a partir do semivariograma para encontrar os pesos ótimos a serem associados às amostras que irão estimar um ponto, uma área ou um bloco.

O uso do semivariograma para a estimativa por krigagem não exige que os dados tenham distribuição normal, mas a presença de distribuição assimétrica, com muitos valores anômalos, deve ser levada em consideração. No entanto, é possível encontrar trabalhos que têm usado transformações de variáveis para alcançar a normalidade e, em muitas situações, os resultados são satisfatórios. Porém, como apontam Azzaline e Capitanio (1999), existem problemas, tais como a interpretação das variáveis transformadas na análise multivariada, especialmente quando cada variável é transformada utilizando uma função diferente.

Quando a suposição de normalidade não é satisfeita, Cysneiros *et al.* (2005) sugerem como alternativa modelos cujas estimativas sejam menos sensíveis a valores discrepantes, tais como os modelos da classe das distribuições simétricas que permitem reduzir a influência dos

outliers, incorporando parâmetros adicionais que ajustam a curtose da distribuição dos dados. A distribuição *t*-student é um desses modelos e apresenta como parâmetro adicional o grau de liberdade (v).

Apesar da distribuição *t*-student ser uma distribuição com cauda mais pesada que a normal e acomodar melhor os *outliers*, ainda é possível que essa distribuição sofra o efeito de observações influentes. Sendo assim, é importante que se realize estudos de sensibilidade sobre ela, por meio de análise de diagnóstico de influência local.

O objetivo desse trabalho foi desenvolver estudos da inferência geoestatística utilizando processos estocásticos com distribuição *t*-student n -variada considerando o grau de liberdade fixo, desenvolver o algoritmo EM e o algoritmo de NR para a estimativa dos parâmetros da estrutura de dependência espacial e do modelo espacial linear; desenvolver técnicas de diagnósticos de influência local para modelos espaciais lineares, considerando processos estocásticos com distribuição *t*-student n -variada, utilizando o afastamento da verossimilhança usual e o Q-afastamento da verossimilhança completa; e aplicar a metodologia desenvolvida na estimativa de parâmetros de dados simulados e na estimativa dos parâmetros de dados reais de atributos físico/químicos do solo e de produtividade de soja de um experimento de agricultura de precisão obtidos nos anos agrícolas de 2005/2006 e 2006/2007, em uma área comercial de 71 ha.

2 REVISÃO BIBLIOGRÁFICA

2.1 Distribuição t -student

Cornish (1954) e Dunnett e Sobel (1954) introduziram, separadamente na literatura, a primeira versão da t -student n -variada, em que consideraram a distribuição do vetor aleatório ($n \times 1$), definido por $\mathbf{X} = \mathbf{Z}/S$, em que $\mathbf{Z} \sim N_n(0, \Sigma)$ e $vS^2/\sigma^2 \sim X_v^2$ são independentes. Para tais autores, a matriz de covariância $\Sigma = \sigma^2\Gamma$, onde $\Gamma = (\rho_{ij})$ é a matriz das correlações entre os componentes do vetor \mathbf{Z} , não-singular e conhecida. Consideram ainda que σ^2 é a variância dos componentes de \mathbf{Z} e é supostamente desconhecida. Assim, S^2 representa um estimador baseado em v graus de liberdade do parâmetro σ^2 , obtendo uma versão da distribuição t -student n -variada, não-singular e centrada na origem. Kshirsagar (1960) apresenta essas definições para o caso não-centrado.

Bennett (1961) apresentou formalmente a distribuição t -student n -variada não singular com vetor de parâmetros de locação $\boldsymbol{\mu} \in \Re^n$, matriz escala $\Sigma > 0$ e v graus de liberdade como $t_n(\boldsymbol{\mu}, \Sigma, v)$ em termos de densidade e, ainda, algumas propriedades, descrevendo como estimar $\boldsymbol{\mu}$ e Σ por meio de máxima verossimilhança, sendo v conhecido. Ainda hoje é bastante utilizada em análise multivariada para v pequeno devido às suas caudas pesadas, que tornam os estimadores de máxima verossimilhança (EMV) de $\boldsymbol{\mu}$ e Σ resistentes a *outliers*.

Dickey (1967) e Dawid (1977, 1981) apresentaram a distribuição t -student n -variada na forma matricial. Dempster *et al.* (1977) mostraram que o algoritmo EM pode ser usado para obter estimadores de máxima verossimilhança univariada com dados completos e v graus de liberdade fixo. Dempster *et al.* (1980) estenderam esse resultado para casos de regressão. Little (1988) ampliou ainda mais os resultados que mostraram como o algoritmo EM pode tratar os casos com dados perdidos.

Considerando o fato de que as distribuições da família $\{t_n(\boldsymbol{\mu}, \Sigma, v); v > 0\}$ têm caudas mais pesadas que a distribuição normal $N_n(\boldsymbol{\mu}, \Sigma)$ e ainda que, quando $v \rightarrow \infty$, a distribuição $t_n(\boldsymbol{\mu}, \Sigma, v)$ aproxima-se da normal, Lange *et al.* (1989) consideraram uma situação mais geral com v graus de liberdade desconhecido, na qual utilizaram o algoritmo EM para encontrar todos os parâmetros, e também forneceram aplicações do modelo geral.

Fang *et al.* (1990), motivados pelo fato da distribuição t -student ser uma alternativa robusta à distribuição normal, apresentam um desenvolvimento teórico em torno dela. Lang e Sinsheimer (1993) apresentam a família da distribuição normal/independente, entre elas a distribuição t -student n -variada, e utilizaram o algoritmo EM na estimação dos parâmetros por máxima verossimilhança para a regressão robusta.

Arellano-Valle (1994) discute diferentes formas de definir a distribuição t -student n -variada, tomando por base que esta distribuição pertence à classe das distribuições normais compostas. Arellano e Bolfarine (1995) discutem três caracterizações para a distribuição t -student n -variada como classe das distribuições elípticas, mostrando que esta distribuição pode ser caracterizada em termos marginal, não-condicional e condicional e de forma quadrática, assim como também a apresentam como uma subclasse da distribuição normal composta (ou mistura da distribuição normal).

Liu e Rubin (1994, 1995) descreveram os algoritmos EM, ECM e ECME, mostrando sua eficiência computacional na estimativação por máxima verossimilhança dos parâmetros da distribuição t -student n -variada, com grau de liberdade conhecido e também desconhecido, com ou sem dados pedidos e com ou sem covariáveis. Mostram ainda que todos os três algoritmos têm convergência monótona estável, sendo que a taxa de convergência do algoritmo ECME pode ser significativamente mais rápida. Nadarajah e Koltz (2005) apresentam propriedades matemáticas da distribuição t -student n -variada.

2.1.1 Definição

Um vetor aleatório n -dimensional $\mathbf{Y} = (Y_1, \dots, Y_n)^T$ tem distribuição t -student n -variada, com vetor de locação $\boldsymbol{\mu} \in \Re^n$, matriz de escala $\boldsymbol{\Sigma}$ ($n \times n$) definida positiva e v graus de liberdade, se sua função densidade de probabilidade é dada por:

$$f_Y(y) = K(v, n) \left[1 + \frac{1}{v}(y - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(y - \boldsymbol{\mu}) \right]^{-\frac{(v+n)}{2}}, \quad y \in \Re^n \quad (2.1)$$

onde

$$K(v, n) = \frac{\Gamma(\frac{v+n}{2})}{(\pi v)^{n/2} \Gamma(\frac{v}{2}) |\boldsymbol{\Sigma}|^{1/2}}, \quad v > 0.$$

Se um vetor aleatório \mathbf{Y} tem uma distribuição com função densidade dada por (2.1) denota-se por $\mathbf{Y} \sim t_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v)$.

A distribuição t -student pertence a classe das distribuições elípticas, de modo que as suas propriedades podem ser obtidas a partir dessa classe (Fang et al., 1990).

2.1.2 Propriedades

Propriedade 1 Seja $\mathbf{Y} = \boldsymbol{\mu} + \mathbf{Z}/\sqrt{U}$, onde $\mathbf{Z} \sim N_n(0, \boldsymbol{\Sigma})$, independente de $\mathbf{U} \sim \chi^2(v)/v$. Então, tem-se que $\mathbf{Y} \sim t_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v)$.

Note que dado $\mathbf{U} = u$, tem-se que

$$\mathbf{Y}|(\mathbf{U} = u) \sim \text{N}_n(\boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma}).$$

A densidade de Y , dada pela Equação (2.1), pode ser escrita como

$$f_Y(y) = \int_0^\infty |\boldsymbol{\Sigma}|^{-1/2} (2\pi u^{-1})^{-n/2} e^{(-\frac{u}{2}(y-u)^T \boldsymbol{\Sigma}^{-1}(y-u))} f_U(u) du,$$

em que $f_U(u)$ é a densidade de \mathbf{U} . Isto mostra que a distribuição t -student n -variada pode ser obtida como uma mistura gaussiana, ou seja, mistura de uma distribuição normal multivariada e uma distribuição χ^2 .

Propriedade 2 Seja $\mathbf{Y}|(\mathbf{U} = u) \sim \text{N}_n(\boldsymbol{\mu}, u^{-1}\boldsymbol{\Sigma})$, onde $\mathbf{U} \sim \chi^2(v)/v$, $v > 0$. Então, tem-se que:

- (i) $\mathbf{Y} \sim t_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v)$;
- (ii) $\mathbf{U}|(\mathbf{Y} = y) \sim \chi^2(v + n)/(v + \delta^2)$, onde $\delta^2 = (\mathbf{Y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \boldsymbol{\mu})$.

Propriedade 3 Seja $\mathbf{Y} \sim t_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v)$. Então,

$$\text{E}(\mathbf{Y}) = \boldsymbol{\mu}, \text{ se } v > 1 \text{ e}$$

$$\text{Var}(\mathbf{Y}) = \frac{v}{v-2} \boldsymbol{\Sigma}, \text{ se } v > 2.$$

Usando a Propriedade 1, tem-se que:

$$\text{E}(\mathbf{Y}) = \boldsymbol{\mu} + \text{E}\left(\frac{1}{\sqrt{\mathbf{U}}}\right) \text{E}(\mathbf{Z}) \text{ e } \text{Var}(\mathbf{Y}) = \text{E}(\mathbf{U}^{-1})\text{E}(\mathbf{Z}\mathbf{Z}^T)$$

e o resultado segue,

$$\text{E}\left(\frac{1}{\sqrt{\mathbf{U}}}\right) = \frac{\left(\frac{v}{2}\right)^{1/2} \Gamma\left(\frac{v-1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \text{ se } v > 1 \text{ e}$$

$$\text{E}(\mathbf{U}^{-1}) = \frac{v}{(v-2)} \text{ se } v > 2.$$

em que $\Gamma(\cdot)$ é a função gama.

Propriedade 4 Seja $\mathbf{Y} \sim t_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v)$ e $a \in \mathbb{R}^q$, $q \leq n$, \mathbf{A} uma matriz $q \times n$ de posto q e

$\mathbf{X} = a + \mathbf{A}\mathbf{Y}$. Então, tem-se que

$$\mathbf{X} \sim t_q(a + \mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T, v).$$

Considere-se \mathbf{Y} , $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$ como

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \end{pmatrix}; \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}; \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$$

onde \mathbf{Y}_1 e $\boldsymbol{\mu}_1$ são vetores $q \times 1$ sendo $q \leq n$; $\boldsymbol{\Sigma}_{11}$ é uma matriz $q \times q$; $\boldsymbol{\Sigma}_{22}$ é uma matriz $(n - q) \times (n - q)$; $\boldsymbol{\Sigma}_{12}$ é uma matriz $(n - q) \times q$ e $\boldsymbol{\Sigma}_{21}$ é uma matriz $q \times (n - q)$.

Propriedade 5 Seja $\mathbf{Y} \sim t_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v)$. Então, tem-se que

$$\mathbf{Y}_1 \sim t_n(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}, v).$$

Propriedade 6 Seja $\mathbf{Y} \sim t_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v)$. Então, tem-se que

$$\mathbf{Y}_1 | (\mathbf{Y}_2 = y_2) \sim t_n(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(y_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{1.2}, v + n - q),$$

onde $\boldsymbol{\Sigma}_{1.2} = \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$, (Arellano-Valle, 1994).

Daqui segue que

$$\mathbb{E}(\mathbf{Y}_1 | (\mathbf{Y}_2 = y_2)) = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(y_2 - \boldsymbol{\mu}_2), \text{ se } v + n - q > 1$$

e

$$\text{Var}(\mathbf{Y}_1 | (\mathbf{Y}_2 = y_2)) = \left[\frac{v + (y_2 - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}_{22}^{-1} (y_2 - \boldsymbol{\mu}_2)}{v + n - q - 2} \right] \boldsymbol{\Sigma}_{1.2},$$

se $v + n - q > 2$.

Propriedade 7 Seja $\mathbf{Y} \sim t_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, v)$. Então, o coeficiente de curtose n -variado é dado por

$$\beta_{2,n} = \frac{n(n+2)(v-2)}{(v-4)}, \text{ para } v > 4. \quad (2.2)$$

O coeficiente de curtose n -variado, $\beta_{2,n}$, é dado por:

$$\beta_{2,n} = \mathbb{E} \left\{ (\mathbf{Y} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \boldsymbol{\mu}) \right\}^2.$$

Considerando a Equação (2.2) e a Propriedade 1, tem-se que:

$$\beta_{2,n} = \left(\frac{v-2}{v} \right)^2 \mathbb{E}(\mathbf{U}^{-2}) \mathbb{E}[(\mathbf{Z}^T \boldsymbol{\Sigma}^{-1} \mathbf{Z})^2]$$

e o resultado segue, já que $E(\mathbf{U}^{-2}) = \frac{v^2}{(v-2)(v-4)}$, $v > 4$, e como $\mathbf{Z} \sim N_n(0, \Sigma)$,

$$E[(\mathbf{Z}^T \Sigma^{-1} \mathbf{Z})^2] = n(n+2) \text{ (Galea, 1995).}$$

Propriedade 8 Seja $\mathbf{Y} \sim t_n(\boldsymbol{\mu}, \Sigma, v)$ com densidade $f_Y(y, v)$ dada pela Propriedade 1 e seja $f_N(y)$ a função de densidade de um vetor aleatório com distribuição $N_n(\boldsymbol{\mu}, \Sigma)$. Então,

$$\lim_{v \rightarrow \infty} f_Y(y, v) = f_N(y), \quad \forall y \in \Re^n.$$

Esta propriedade segue diretamente do fato que (Tong, 1990):

$$\lim_{v \rightarrow \infty} \left(1 + \frac{1}{v} (\mathbf{y} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \right)^{-\frac{1}{2}(v+n)} = \exp \left\{ -\frac{1}{2} (\mathbf{y} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \right\},$$

e que

$$\lim_{v \rightarrow \infty} K(v, n) = \frac{1}{(2\pi)^{n/2}} |\Sigma|^{1/2}.$$

Combinando a Propriedade 8 com o Teorema de convergência de Scheffé, tem-se a seguinte propriedade.

Propriedade 9 Seja $\mathbf{Y} \sim t_n(\boldsymbol{\mu}, \Sigma, v)$ e $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \Sigma)$. Então, tem-se que

$$\lim_{v \rightarrow \infty} P\{\mathbf{Y} \in B\} = P\{\mathbf{X} \in B\},$$

para todo subconjunto B Borel-mensurável de \Re^n . Como caso especial, temos que $\mathbf{Y} \xrightarrow{D} \mathbf{X}$, quando $v \rightarrow \infty$.

Por outro lado, na prática, é comum supor a existência de segundo momento (Sutradhar, 1993) e trabalhar com a distribuição $t_n(\boldsymbol{\mu}, \Sigma, v)$ com matriz de covariância dada por $\Sigma^* = \text{Var}(\mathbf{Y})$. Neste caso, $\Sigma = ((v-2)/v) \Sigma^*$ e, substituindo Σ por Σ^* em (2.1), têm-se que

$$f_Y(y) = C(v, n) \left\{ (v-2) + (\mathbf{y} - \boldsymbol{\mu})^T \Sigma^{*-1} (\mathbf{y} - \boldsymbol{\mu}) \right\}^{-\left(\frac{v+n}{2}\right)}, \quad (2.3)$$

onde

$$C(v, n) = \frac{\left[(v-2)^{-\frac{1}{2}} \Gamma \left(\frac{v+n}{2} \right) \right]}{\pi^{\frac{n}{2}} \Gamma \left(\frac{v}{2} \right)}, \quad v > 2.$$

Se um vetor \mathbf{Y} tem uma distribuição com função de densidade dada por (2.3), escreve-se $\mathbf{Y} \sim t_n^*(\boldsymbol{\mu}, \Sigma^*, v)$. Ou seja, com esta versão reparametrizada da distribuição t -student n -variada, tem-se que $E(\mathbf{Y}) = \boldsymbol{\mu}$ e $\text{Var}(\mathbf{Y}) = \Sigma^*$, que também permite uma comparação direta com a distribuição $N_n(\boldsymbol{\mu}, \Sigma)$.

2.2 Modelos espaciais lineares

A agricultura de precisão baseia-se no gerenciamento agrícola de informação sobre a variabilidade espacial e temporal dos fatores de produção e da produtividade, que é apresentada por mapas temáticos de superfície ou isolinhas. A construção desses mapas temáticos requer a utilização de um conjunto de técnicas de estatística espacial denominado Geoestatística, a qual baseia-se na teoria das variáveis regionalizadas, formalizada por Matheron (1963), que considera as características estruturais e aleatórias de uma variável espacialmente distribuída, de forma a descrevê-la adequadamente.

Diversos trabalhos na região têm sido desenvolvidos utilizando geoestatística, entre eles destacam-se Souza *et al.* (1999), que estudaram a variabilidade espacial de atributos químicos do solo em um latossolo vermelho distroférrego da região Oeste do Paraná. Também, Ribeiro Jr (2002) desenvolveu estudos estatísticos espaciais com modelos e métodos para análise de dados georreferenciados; Mercante *et al.* (2003) e Johann *et al.* (2004) estudaram a variabilidade espacial de atributos do solo utilizando métodos estatísticos com semivariogramas escalonados e análise multivariada de componentes principais; e Guedes *et al* (2008) realizaram a análise da anisotropia no estudo da variabilidade espacial de algumas propriedades químicas do solo espacialmente referenciadas.

2.2.1 A teoria das variáveis regionalizadas

A variável regionalizada é uma variável distribuída no espaço, isto é, uma variável aleatória cujos valores são relacionados de algum modo com a posição espacial que ocupam. 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.

2.2.2 Características da variável regionalizada

A variável regionalizada possui características qualitativas estreitamente ligadas à estrutura do fenômeno natural que elas representam. Tais características são:

a) Localização

A variável regionalizada somente toma seus valores importantes no campo geométrico em que foram definidas, isto é, no espaço geométrico onde a variável é suscetível de tomar 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.

b) Continuidade

A continuidade espacial apresentada pela VR entre elementos amostrais vizinhos reflete o grau de dependência entre eles. Espera-se que elementos amostrais mais próximos apresentem maior dependência do que aqueles que estão mais distantes (Cressie, 1993).

Segundo Isaaks e Srivastava (1989), a teoria fundamental da geoestatística é a de que, em média, os elementos amostrais próximos no espaço sejam mais similares entre si do que aqueles que estiverem mais distantes. Quando a continuidade espacial não é identificada, diz-se que há presença de efeito de pepita puro. Isto é, a distância entre os elementos amostrais pode ter sido muito grande, a ponto de não identificar a variabilidade espacial da variável em estudo.

c) Isotropia

A isotropia é uma característica importante da VR e é identificada quando a variabilidade espacial em uma área apresenta comportamento semelhante para distintas direções, isto é, não há uma direção privilegiada, ao longo da qual os valores variem de forma acentuada. Quando isso não ocorre, diz-se que há anisotropia. As direções consideradas no estudo da anisotropia são 0° , 45° , 90° e 135° (Guedes et al., 2008), conforme a Figura 1.

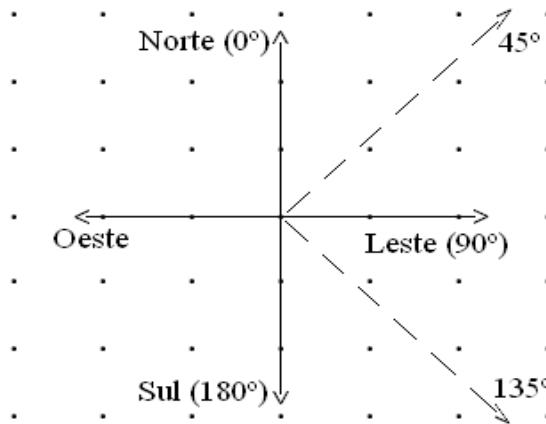


Figura 1: Direções utilizadas no estudo da anisotropia da variável regionalizada.

A existência de direções que apresentem distintas estruturas de dependência ou que não apresentem a mesma estrutura de continuidade espacial, caracteriza a anisotropia, que pode ser identificada pelo gráfico do semivariograma, para diferentes direções (Soares, 2000).

2.2.3 Modelo espacial linear

Considere um processo estocástico¹ $\{Y(s_i), s_i \in S\}$, em que $S \subset \mathbb{R}^d$, sendo \mathbb{R}^d espaço euclidiano, d -dimensional ($d \geq 1$).

¹Processo estocástico é um conjunto de variáveis aleatórias definidas em um espaço paramétrico.

Cada $\mathbf{Y} = (Y(s_1), \dots, Y(s_n))^T$ corresponde a localizações espaciais conhecidas em s_i e s_j para $i \neq j = 1, 2, \dots, n$.

Pode-se escrever $Y(s_i)$ como:

$$Y(s_i) = \mu(s_i) + \epsilon(s_i), \quad i = 1, \dots, n, \quad (2.4)$$

em que $\mu(s_i)$ é um termo determinístico de locação e $\epsilon(s_i)$ é um termo estocástico, ambos dependentes do espaço paramétrico onde $Y(s_i)$ atua.

Assume-se que o erro estocástico ϵ tem $E(\epsilon(s_i)) = 0$, para $i = 1, \dots, n$ e que a variação entre os pontos no espaço é determinada por alguma função de covariância $C(\cdot)$ de s_i e s_j . Essa função de covariância é também especificada por um vetor de parâmetros $\Phi = (\varphi_1, \varphi_2, \varphi_3)^T$, que são os parâmetros que definem a estrutura de dependência espacial.

Considera-se o modelo espacial linear do termo determinístico $\mu(\cdot)$ como

$$\mu(s_i) = \beta_1 X_1(s_i) + \beta_2 X_2(s_i) + \beta_3 X_3(s_i) + \dots + \beta_p X_p(s_i), \quad (2.5)$$

$$\mu(s_i) = \sum_{j=1}^p \beta_j \mathbf{X}_j(s_i), \quad i = 1, \dots, n,$$

em que:

\mathbf{X}_1 é um vetor de um's;

X_2, \dots, X_p são covariáveis ou variáveis explicativas tomadas na posição (s_i) e β_1, \dots, β_p são parâmetros desconhecidos a serem estimados.

A Equação (2.5) pode ser escrita em forma matricial, tal como:

$$\mu = \mathbf{X}\boldsymbol{\beta} \quad (2.6)$$

em que:

$\boldsymbol{\mu} = (\mu(s_1), \dots, \mu(s_n))^T$ é um vetor $n \times 1$;

\mathbf{X} é uma matriz $n \times p$ composta pelo vetor de um's e pelas $(p - 1)$ covariáveis;

$\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ é o vetor de parâmetros desconhecidos a serem estimados, os quais definem a relação linear na Equação (2.6).

Assim, pode-se escrever a Equação (2.4) em forma matricial

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

sendo $\mathbf{Y} = (Y(s_1), \dots, Y(s_n))^T$ e $\boldsymbol{\epsilon} = (\epsilon(s_1), \dots, \epsilon(s_n))^T$ vetores $n \times 1$.

A esperança do vetor de erros aleatórios $\boldsymbol{\epsilon}$, $E(\boldsymbol{\epsilon}) = \mathbf{0}$ (vetor nulo) e a sua matriz escala é da forma $\boldsymbol{\Sigma} = [(\sigma_{ij})]$, em que $\sigma_{ij} = C(s_i, s_j)$.

Assumindo que $\boldsymbol{\Sigma}$ é não-singular e que a matriz \mathbf{X} é de posto completo ($\text{rank}(\mathbf{X}) = p$),

a matriz escala Σ tem estrutura espacial da forma

$$\Sigma = [(\sigma_{ij})] = \varphi_1 \mathbf{I}_n + \varphi_2 R, \quad (2.7)$$

em que:

\mathbf{I}_n é a matriz identidade $n \times n$;

φ_1 é o parâmetro que define o efeito pepita puro;

φ_2 é o parâmetro que define o *sill* ou a contribuição;

$R = R(\varphi_3)$ é uma matriz $n \times n$, simétrica, sendo que o alcance (a) é uma função de φ_3 ($a = g(\varphi_3)$).

Assim, os elementos C_{ij} da matriz Σ são $C(s_i, s_j) = C(h_{ij}) = \varphi_2 r_{ij}$ em que $h_{ij} = \|s_i - s_j\|$, sendo os elementos r_{ij} de R da forma

$$r_{ij} = \frac{1}{\varphi_2} C(h_{ij}), \text{ para } i \neq j, r_{ii} = 1, i \text{ e } j = 1, \dots, n.$$

Nas funções de covariâncias $C(h_{iu})$, a variância do processo estocástico \mathbf{Y} , no caso da distribuição normal, $\mathbf{Y} \sim N_n(\mathbf{X}\boldsymbol{\beta}, \Sigma)$, é dada por

$$C(0) = \varphi_1 + \varphi_2.$$

E no caso da distribuição *t*-Student, $\mathbf{Y} \sim t_n(\mathbf{X}\boldsymbol{\beta}, \Sigma, v)$, é dada por:

$$C(0) = \frac{v}{v-2} (\varphi_1 + \varphi_2). \quad (2.8)$$

Assim, a semivariância pode ser definido como:

$$\gamma(h) = C(0) - C(h) \quad (2.9)$$

nos dois casos.

2.2.4 Parâmetros do semivariograma

Considerando-se a Equação (2.7), os parâmetros que definem a estrutura de dependência espacial são:

- i) Efeito pepita (φ_1): 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 em 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 captados pela amostragem;
- ii) Contribuição (φ_2): 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 dependência 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, sendo que o alcance é uma função do parâmetro φ_3 ;

iv) Patamar ($C = \varphi_1 + \varphi_2$): é o valor máximo do semivariograma correspondente ao seu alcance no qual estabiliza-se.

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

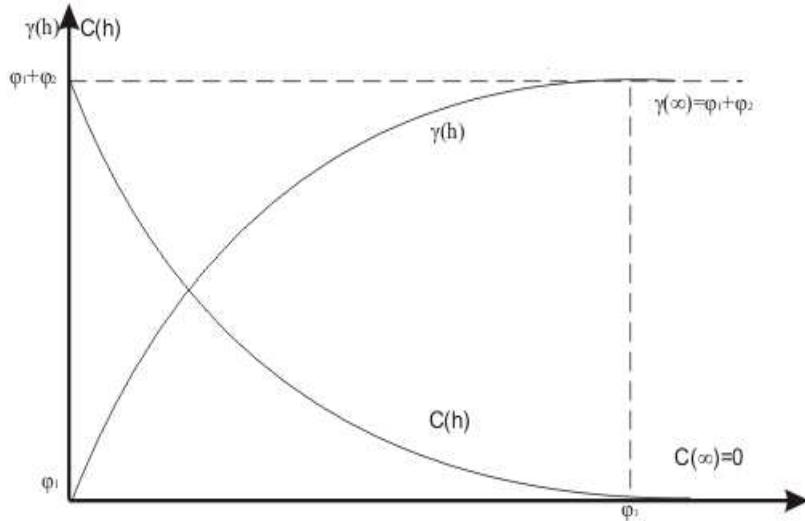


Figura 2: Relação entre a covariância espacial $C(h)$ e a semivariância $\gamma(h)$.

2.2.5 Modelos teóricos de semivariograma, funções de covariância e de correlação

Apresentados os parâmetros obtidos com o estudo da dependência espacial, segue-se com a apresentação de alguns dos modelos teóricos de semivariogramas, funções de covariâncias e funções de correlações espaciais mais utilizados na literatura.

A partir da função semivariância apresentada na Equação (2.9), tem-se a relação $C(h) = C(0) - \gamma(h)$. Além disso, dividindo por $C(0)$, tem-se:

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

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

A função de correlação espacial é da forma:

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

Os modelos espaciais utilizados para a estimativa da estrutura de dependência espacial do semivariograma experimental ou da matriz escala Σ , apresentados por Isaaks 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 e família Matérn. 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 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 de semivariograma é válido em \mathbb{R} , \mathbb{R}^2 e \mathbb{R}^3 (Cressie, 1993), sendo expressa por

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

A função de covariância é expressa por

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

A função de correlação espacial é definida como

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

b) Modelo gaussiano

Este modelo apresenta comportamento parabólico na origem e é utilizado para descrever um fenômeno extremamente contínuo. Também atinge o patamar apenas assintoticamente e o alcance prático é dado por $a = \sqrt{3}\varphi_3$. Este modelo de semivariograma é válido em \mathbb{R} , \mathbb{R}^2 e \mathbb{R}^3

(Cressie, 1993) e é dado por

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

A função de covariância é expressa por

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

A função de correlação espacial é definida como

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

c) Modelo da família Matérn

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

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

em que $\varphi_1, \varphi_2, \varphi_3$ e k são parâmetros e K_k é a função 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 ainda à função de correlação gaussiana, quando $k \rightarrow \infty$.

A função de covariância é expressa por

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

A função de correlação espacial é definida como

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

2.3 Análise de Diagnóstico

Diagnóstico de influência é uma técnica utilizada para avaliar a qualidade do ajuste do modelo, pela avaliação das suposições feitas a ele e também pela avaliação da robustez de suas estimativas quando são introduzidas perturbações no modelo e/ou nos dados (Ferreira, 2008).

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. Ortega *et al.* (2002) utilizaram análises de diagnóstico em modelos de regressão log-gamma generalizado; Osório *et al.* (2007) estudaram a curvatura de influência local sobre vários esquemas de perturbação para modelos lineares elípticos, com estrutura longitudinal; Christensen *et al.* (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.

São duas as linhas de diagnóstico de influência: *Influência global*, também denominada como eliminação de pontos, que caracteriza-se pela identificação de observações influentes mediante eliminação de casos, e *Influência local*, que consiste em estudar o efeito de introduzir pequenas perturbações no modelo (ou nos dados), usando uma medida de influência apropriada.

2.3.1 Influência local

A metodologia de influência local, inicialmente proposta por Cook (1986), tem se tornado uma ferramenta popular de diagnóstico para identificação conjunta de observações influentes em regressão linear e não-linear, com a grande vantagem de poder ser aplicada a qualquer modelo paramétrico.

Schall e Dunne (1992) demonstraram a relação próxima entre o conceito da colinearidade dos parâmetros e a influência local. Ainda nesse trabalho, definiram a curvatura normal proposta por Cook no espaço \mathbb{R}^d , quando $d = 1$ ou $d > 1$.

Fung e Kwan (1997) mostraram que a influência local baseada na curvatura normal não é suficientemente fundamentada para produzir resultados úteis quando aplicada à estatística em que a primeira derivada não é zero. Assim, distinguem a aplicação da influência local baseada na probabilidade de deslocamento e a abordagem baseada em estatísticas como a soma dos quadrados dos resíduos.

Poon e Poon (1999) apresentaram uma reparametrização da curvatura normal de Cook, denominando-a por curvatura normal conformal. Estudaram várias aproximações usando essa curvatura e a relação entre essas aproximações.

Zhu e Lee (2001) propuseram um método para avaliar a influência local em dados incompletos utilizando o algoritmo EM.

Lee e Xu (2003) estudaram a influência local em modelos lineares generalizados mistos e, ainda, Lee e Xu (2004) estudam a influência local em modelos com efeitos misto não-lineares

sob normalidade.

Galea *et al.* (2002) estudaram a influência local nos modelos com erros nas variáveis sob a distribuição *t*-student, enquanto que Galea *et al.* (2003) investigaram a influência das observações nas estimativas dos parâmetros usando o enfoque de influência local na classe dos modelos simétricos lineares.

Militino *et al.* (2004) apresentaram um estudo de detecção de outliers em modelos espaciais lineares multivariados.

Zhu e Zang (2004) propuseram um procedimento de diagnóstico baseado na influência local para avaliar a má especificação do modelo e detectar pontos influentes simultaneamente.

Xie *et al.* (2007) apresentaram uma técnica para análise de influência local baseada na mistura de distribuições e algoritmo EM, considerando a distribuição *t*-student multivariada como um caso particular da mistura gaussiana.

Considerando-se a análise de diagnóstico em modelos espaciais lineares, têm-se Christensen *et al.* (1992a, 1993), que apresentaram estudos para a predição de parâmetros em modelos espaciais lineares aplicando técnicas de diagnóstico para detectar observações que exercem influência na estimativa da matriz de covariância, utilizados na krigagem universal e, Borssoi *et al.* (2009), que apresentaram estudos de diagnóstico em modelos espaciais lineares gaussianos, desenvolvendo diversas medidas de influência local.

2.4 Krigagem

A geoestatística tem um método de interpolação chamado krigagem. Este método usa a dependência espacial, expressa no semivariograma ou função de covariância, para predizer valores em qualquer posição dentro de um espaço amostral, sem tendência e com variância mínima. Estas duas características fazem da krigagem um interpolador BLUP - *Best Linear Unbiased Predictor* (Ribeiro Jr., 1995).

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 a estrutura de dependência espacial das variáveis regionalizadas em estudo, pode-se usar a interpolação por krigagem para predizer 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) é a 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 elementos amostrais). 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. Já na krigagem, o procedimento é semelhante

ao de interpolação por médias móveis ponderada, exceto que aqui os pesos são determinados a partir de uma análise espacial, baseado na estimação de parâmetros obtidos pela modelagem da dependência espacial. Além disso, como mencionado anteriormente, a krigagem oferece estimadores BLUE (*Best Linear Unbiased Estimator*) ou BLUP(*Best Linear Unbiased Predictor*), isto é, seus estimadores não são tendenciosos, pois, em média, a diferença entre valores preditos e observados, para a mesma localização, deve ser nula. Além disso, a krigagem possui variância mínima, pois este estimador possui a menor variância dentre todos os estimadores não tendenciosos. Portanto, a krigagem é uma técnica especificamente desenvolvida para a geoestatística, pois leva em consideração 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 por meio de uma combinação linear de n valores conhecidos, $Y(s_i)$, $i = 1, 2, \dots, n$. Na krigagem, esses n valores conhecidos são de uma mesma característica ou propriedade em estudo. O preditor pode ser escrito como

$$\hat{Y}(s_0) = \sum_{i=1}^n \lambda_i Y(s_i), \quad (2.10)$$

em que

$\hat{Y}(s_0)$ é o valor predito no local s_0 não amostrado;

n é o número de valores $Y(s_i)$ medidos das variáveis em pontos amostrados;

λ_i são os pesos associados com cada valor $Y(s_i)$ medido.

O estimador definido em (2.10) é tido como melhor estimador não-viciado e com variância mínima. 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 $Y(\cdot)$ ser intrinsecamente estacionário, tem-se que o erro médio é

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

Esse erro será nulo se

$$\sum_{i=1}^n \lambda_i = 1, \quad (2.11)$$

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

$$\sigma^2 = \text{Var}[Y(s_0) - \hat{Y}(s_0)] = E[(Y(s_0) - \hat{Y}(s_0))^2] - [E(Y(s_0) - \hat{Y}(s_0))]^2. \quad (2.12)$$

As duas condições citadas em (2.11) e (2.12) garantem que o estimador de krigagem seja BLUE (Webster e Oliver, 2007).

Substituindo (2.10) em (2.12), desenvolvendo-se a expressão resultante e minimizando-a, sujeita à restrição expressa em (2.11), chega-se aos valores de λ_i , para $i = 1, 2, \dots, n$, sendo estes soluções do sistema de equações

$$\sum_{j=1}^n \lambda_j \gamma(s_{ij}) - \eta = \gamma(s_{i0}), \quad i = 1, 2, \dots, n, \quad (2.13)$$

em que a incógnita η é um multiplicador de Lagrange introduzido ao minimizar a variância do erro, $\gamma(s_{ij})$ refere-se à função de semivariância correspondente a um vetor h , com origem em s_i e extremidade em s_j , $\gamma(s_{i0})$ refere-se a função de semivariância correspondente a um vetor h , com origem em s_i e extremidade em s_0 .

O sistema expresso em (2.13), constituído de $n+1$ equações e $n+1$ incógnitas, é conhecido como krigagem ordinária. Na forma matricial, o sistema pode ser escrito como $\gamma\lambda = A$, sendo que

$$\boldsymbol{\lambda} = [\lambda_1 \ \lambda_2 \ \lambda_3 \ \cdots \ \lambda_n \ \eta]^T,$$

$$\mathbf{A} = [\gamma(s_{10}) \ \gamma(s_{20}) \ \cdots \ \gamma(s_{n0}) \ 1]^T \text{ e}$$

$$\boldsymbol{\gamma} = \begin{bmatrix} \gamma(s_{11}) & \gamma(s_{12}) & \cdots & \gamma(s_{1n}) & 1 \\ \gamma(s_{21}) & \gamma(s_{22}) & \cdots & \gamma(s_{2n}) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma(s_{n1}) & \gamma(s_{n2}) & \cdots & \gamma(s_{nn}) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix}.$$

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

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):

$$\sigma^2 = \sum_{i=1}^n \lambda_i \gamma(s_{i0}) + \eta.$$

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

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

em que $f_l(s)$ são funções conhecidas de s e a_l são coeficientes desconhecidos e constantes na vizinhança de s .

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

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

isto é,

$$\hat{Y}(s_0) = \sum_{\alpha=1}^n \lambda_\alpha Y(s_\alpha) + \sum_{l=0}^k a_l \left[f_l(s_0) - \sum_{\alpha=1}^n \lambda_\alpha f_l(s_\alpha) \right]. \quad (2.14)$$

Para se calcular os coeficientes e derivar a_l por meio do estimador (2.14), impõe-se as seguintes restrições:

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

Segundo Cressie (1993), a krigagem universal é um método de interpolação sem tendência e com variância mínima que é aplicado quando há dependência espacial entre os elementos amostrais, obtido pelo semivariograma da variável regionalizada. Tal interpolador visa obter valores para a variável regionalizada em pontos não-amostrados, para a construção de mapas temáticos.

O preditor de krigagem universal, que é interpretado como o valor interpolado de uma variável regionalizada Y , num local s_0 , pode ser determinado por:

$$\hat{Y}(s_0) = \sum_{\alpha=1}^n \lambda_\alpha Y(s_\alpha),$$

em que $\hat{Y}(s_0)$ é o valor predito no local s_0 ; n é o número de valores $Y(s_\alpha)$ medidos das variáveis em pontos amostrados e λ_α são os pesos associados com cada valor $Y(s_\alpha)$ medido.

2.5 Métodos de validação do modelo

O melhor modelo de semivariância poderá ser selecionado com base na análise dos erros, na comparação entre o modelo teórico geoestatístico escolhido e os valores empíricos obtidos na amostragem. Os métodos de validação do modelo fazem a análise desse erro. Entre os principais critérios de validação encontra-se o critério de Akaike, método da validação cruzada e valor máximo da função de verossimilhança.

2.5.1 Critério de Akaike

O critério de informação de Akaike (AIC, do inglês *Akaike's information criterion*) (Akaike, 2002) 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.

AIC é definido como:

$$\text{AIC} = -2l(\theta) + 2r,$$

em que $l(\theta)$ é o logaritmo da função de verossimilhança e r é o número de parâmetros do modelo ajustado.

AIC 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 procura-se uma solução satisfatória entre o bom ajuste e o princípio da parcimônia.

2.5.2 Validação cruzada

Segundo Isaaks 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, predito por krigagem, usando 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’ (Isaaks e Srivastava, 1989).

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

Segundo Gonçalves (1997), espera-se que os erros de predição dados pela Equação (2.15) tenham média nula e variância mínima. Estes erros são dados por

$$\epsilon(s_i) = Y(s_i) - \hat{Y}(s_{(i)}) \quad (2.15)$$

em que, $\hat{Y}(s_{(i)})$ é o valor predito sem a i -ésima observação.

2.5.3 Valor máximo da função verossimilhança

Neste critério, o melhor modelo é escolhido observando o valor máximo do logaritmo da função de verossimilhança. O modelo que apresentar o maior valor é dito o melhor.

2.6 Méridas para a comparação de mapas

Para uma mesma variável é possível que sejam construídos diversos tipos de mapas temáticos. Para que esses mapas possam ser comparados, é necessário quantificar as áreas e determinar a concordância entre os mesmos. Estas informações podem ser obtidas a partir da matriz de erros dos mapas dadas no Quadro 1. Esta matriz tem como unidade de medida o *pixel*² e o número total de *pixels* equivale à área total do mapa.

Considera-se como mapa de referência ou mapa real aquele que possui todos os dados e mapa modelo aquele obtido após a eliminação das observações. A área do mapa real é quantificada nas colunas, enquanto que a área do mapa modelo é quantificada nas linhas. Cada elemento da matriz representa a área pertencente à classe i -ésima do mapa modelo e à classe j -ésima do mapa de referência, sendo N a área total. A diagonal principal (quando $i = j$) representa casos em que a área apresentou a mesma classificação nos dois mapas, enquanto que os elementos fora da diagonal principal representam as classificações errôneas. Observa-se que, se os dois mapas forem iguais, todos os elementos fora da diagonal principal serão nulos.

A avaliação da acurácia é útil para avaliar a qualidade da “predição da superfície”. No caso, avaliar a qualidade do mapa modelo, ou seja, do mapa estimado por um modelo em relação ao real ou de referência é útil também para melhorar a qualidade do mapa identificando e corrigindo fontes de erro, facilitar a comparação de vários algoritmos, técnicas, modelos criadores (desenvolvedores) e interpretadores (tradutores) e determinar a relevância dos dados produzidos no processo de tomada de decisão (Congalton e Green, 1999).

²*pixel* é o menor elemento num dispositivo de exibição ao qual é possível atribuir uma cor. De uma forma mais simples, um *pixel* é o menor ponto que forma uma imagem digital, sendo que o conjunto de milhares de *pixels* formam a imagem inteira.

Quadro 1: Matriz de erros com r classes

	Mapa referência				
Mapa modelo	1	2	...	r	Total
1	x_{11}	x_{12}	...	x_{1r}	$X_{1+} = \sum_{j=1}^r x_{1j}$
2	x_{21}	x_{22}	...	x_{2r}	$X_{2+} = \sum_{j=1}^r x_{2j}$
⋮	⋮	⋮	⋮	⋮	⋮
r	x_{r1}	x_{r2}	...	x_{rr}	$X_{r+} = \sum_{j=1}^r x_{rj}$
Total	$X_{+1} = \sum_{i=1}^r x_{i1}$	$X_{+2} = \sum_{i=1}^r x_{i2}$...	$X_{+r} = \sum_{i=1}^r x_{ir}$	N

Uma medida utilizada para avaliar a acurácia dos mapas construídos é a exatidão global (EG), que tem por objetivo determinar o percentual de acerto da classificação realizada corretamente e pode ser extraída da matriz dos erros por:

$$\text{EG} = \frac{\sum_{i=1}^r x_{ii}}{N}.$$

De acordo com Anderson *et al.* (1976), o nível mínimo de precisão aceitável da exatidão global é de 0,85.

EG é uma estatística de avaliação utilizada para mensurar a similaridade entre algo de referência e um modelo ajustado. Porém, é conveniente apresentar esta medida associada com outras métricas que levem em consideração os demais valores contidos na matriz dos erros e não somente os valores da diagonal principal.

Segundo Gong e Howarth (1990), o índice kappa (Cohen, 1960) é uma estatística utilizada para mensurar a exatidão das classificações temáticas e é recomendada como uma medida apropriada para esta medição, por utilizar todos os elementos da matriz dos erros em vez de apenas utilizar aqueles que se situam na diagonal principal da mesma. Esse índice fornece uma medida de concordância entre os valores estimados pelo modelo e os valores observados de referência, conforme a seguinte equação:

$$K = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} x_{+i})},$$

em que

K: índice kappa de concordância;

N : total de área (equivalente ao número total de *pixels*);

r : número de linhas e colunas da matriz;

x_{ii} : área da i -ésima linha e a i -ésima coluna;

x_{i+} : área da i -ésima linha, sendo $x_{i+} = \sum_{j=1}^r x_{ij}$;

x_{+i} : área da i -ésima coluna, sendo $x_{+i} = \sum_{u=1}^r x_{ui}$.

Kripendorff (1980) classifica o índice kappa como de baixa exatidão se $K < 0,67$, como de média exatidão de $0,67 \leq K < 0,80$ e como de alta exatidão se $K \geq 0,80$. Forbes (1995) observou que o índice kappa pode ser sensível ao tamanho da amostra e é falho quando o tamanho de uma classe é muito superior ao tamanho das outras.

Além dos índices de exatidão global e kappa, que são muito utilizados, Ma e Redmond (1995) apresentaram o índice tau (T) que fornece uma medida quantitativa relativamente precisa e intuitiva sobre a acurácia da classificação. O índice tau é similar ao índice kappa e é calculado conforme a Equação (2.16), em que p indica a probabilidade *a priori* para cada i -ésima classe, para $i = 1, 2, \dots, r$.

$$T = \frac{\sum_{i=1}^r x_{ii}/N - p_i}{1 - p_i}. \quad (2.16)$$

Quando as probabilidades *a priori* para as classes forem iguais, temos $p_i = 1/r$, onde r representa o número de classes da matriz dos erros. O índice tau pode seguir a mesma classificação do índice kappa, segundo Kripendorff (1980).

3 MATERIAL E MÉTODOS

3.1 Estimação de parâmetros por máxima verossimilhança para o processo t -student utilizando o algoritmo EM

Considere-se $\{Y(s), s \in S\}$ um processo estocástico estacionário, em que $S \subset \mathbb{R}^d$ e \mathbb{R}^d é um espaço euclíadiano d -dimensional ($d \geq 1$) e que \mathbf{Y} tem uma distribuição t -student n -variada com vetor de parâmetros de média $\mathbf{X}\boldsymbol{\beta}$, matriz escala Σ e v graus de liberdade (fixo), sendo $v > 0$, isto é, $\mathbf{Y} \sim t_n(\mathbf{X}\boldsymbol{\beta}, \Sigma, v)$.

Seja $\mathbf{Y}_c = (\mathbf{Y}, \mathbf{U})$ o vetor de dados completos, tal que \mathbf{Y} é o vetor de dados observados e \mathbf{U} o vetor de dados não observados. A distribuição condicional de \mathbf{Y} dado \mathbf{U} é normal n -variada com vetor de locação $\boldsymbol{\mu}$ e matriz de covariância $u^{-1}\Sigma$, ou seja, $(\mathbf{Y}|\mathbf{U} = u) \sim N_n(\mathbf{X}\boldsymbol{\beta}, u^{-1}\Sigma)$ e $u \sim h(u)$ é uma variável aleatória.

A função densidade conjunta dos dados completos $f_{Y,U}(y, u)$ é dada por

$$f_{Y,U}(y, u) = f_{Y|U}(y|u) \cdot h(u),$$

e o logaritmo da verossimilhança completa por

$$\log(f_{Y,U}(y, u)) = \log(f_{Y|U}(y|u)) + \log(h(u)).$$

Tal que, $\log(f_{Y,U}(y, u)) = l_c(\boldsymbol{\theta}|\mathbf{Y}_c)$ com $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \boldsymbol{\Phi}^T)^T$, sendo $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_p)^T$ e $\boldsymbol{\Phi} = (\varphi_1, \varphi_2, \varphi_3)^T$.

A função densidade condicional

$$f_{Y|U}(y|u) = \frac{1}{(2\pi)^{n/2} u^{-n/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}u\delta},$$

sendo $\delta = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$.

Tem-se que

$$\log(f_{Y|U}(y|u)) = -\frac{n}{2}\log(2\pi) + \frac{n}{2}\log(u) - \frac{1}{2}\log|\Sigma| - \frac{u}{2}\delta.$$

Então

$$E \left\{ \log(f_{Y|U}(y|u)|Y, \hat{\boldsymbol{\theta}}) \right\} = -\frac{n}{2}\log(2\pi) + \frac{n}{2}E \left\{ \log(U|Y, \hat{\boldsymbol{\theta}}) \right\} - \frac{1}{2}\log|\Sigma| - \frac{\delta}{2}E \left\{ U|Y, \hat{\boldsymbol{\theta}} \right\}.$$

Sendo

$$h(u) = \frac{\left(\frac{v}{2}\right)^{v/2} u^{v/2-1}}{\Gamma\left(\frac{v}{2}\right)} e^{(-uv/2)},$$

em que \mathbf{U} tem distribuição Gamma, $G(v/2, v/2)$, assim obtém-se a distribuição $t_n(\boldsymbol{\mu}, \Sigma, v)$ com densidade

$$f_Y(y) = \frac{\Gamma\left(\frac{v+n}{2}\right) |\Sigma|^{-1/2}}{\Gamma\left(\frac{v}{2}\right) (v\pi)^{n/2}} \left(1 + \frac{\delta}{v}\right)^{-1/2(v+n)},$$

Neste caso, $(U|Y = u) \sim \frac{\chi^2(v+n)}{v+\delta}$.

Logo,

$$\mathbb{E}\{U|Y = y\} = \left(\frac{v+n}{v+\delta}\right) \text{ e}$$

$$\mathbb{E}\{\log(U|Y = y)\} = \psi\left(\frac{v+n}{2}\right) - \log\left(\frac{\delta+v}{2}\right),$$

onde ψ é a função digama.

O logaritmo da função verossimilhança para os dados completos é dado por

$$l_c(\boldsymbol{\theta}|\mathbf{Y}_c) = -\frac{n}{2}\log(2\pi) + \frac{n}{2}\log(u) - \frac{1}{2}\log(|\Sigma|) - \frac{u}{2}\delta + \log(h(u)),$$

isto é,

$$\begin{aligned} l_c(\boldsymbol{\theta}|\mathbf{Y}_c) = & -\frac{n}{2}\log(2\pi) + \frac{n}{2}\log(u) - \frac{1}{2}\log(|\Sigma|) - \frac{u}{2}\delta + \\ & + \frac{v}{2}\log\left(\frac{v}{2}\right) + \left(\frac{v}{2} - 1\right)\log(u) - \log\left(\Gamma\left(\frac{v}{2}\right)\right) - \left(u\frac{v}{2}\right). \end{aligned}$$

Desse modo, têm-se a esperança condicional do logaritmo da função de verossimilhança para dados completos

$$Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) = \mathbb{E}\left\{l_c(\boldsymbol{\theta}|\mathbf{Y}_c)|\mathbf{U}, \hat{\boldsymbol{\theta}}\right\},$$

em que

$$Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}}) = -\frac{n}{2}\log(2\pi) + \frac{n}{2}\mathbb{E}\left\{\log\left(\mathbf{U}|\mathbf{Y}, \hat{\boldsymbol{\theta}}\right)\right\} - \frac{1}{2}\log(|\Sigma|) - \frac{\delta}{2}\mathbb{E}\left\{\mathbf{U}|\mathbf{Y}, \hat{\boldsymbol{\theta}}\right\} +$$

$$+ \frac{v}{2}\log\left(\frac{v}{2}\right) + \left(\frac{v}{2} - 1\right)\mathbb{E}\left\{\log\left(\mathbf{U}|\mathbf{Y}, \hat{\boldsymbol{\theta}}\right)\right\} - \log\left(\Gamma\left(\frac{v}{2}\right)\right) - \frac{v}{2}\mathbb{E}\left\{\mathbf{U}|\mathbf{Y}, \hat{\boldsymbol{\theta}}\right\}. \quad (3.1)$$

Considere-se:

$$\vartheta = E\left\{\mathbf{U}|\mathbf{Y}, \hat{\boldsymbol{\theta}}\right\} = \frac{v+n}{v+\delta}$$

e

$$c = E \left\{ \log(\mathbf{U} | \mathbf{Y}, \hat{\boldsymbol{\theta}}) \right\} = \psi \left(\frac{v+n}{2} \right) - \log \left(\frac{\hat{\delta} + v}{2} \right).$$

Então, tem-se:

$$\begin{aligned} Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}) &= -\frac{n}{2} \log(2\pi) + \frac{n}{2} c - \frac{1}{2} \log(|\Sigma|) - \frac{\delta}{2} \vartheta + \\ &+ \frac{v}{2} \log \left(\frac{v}{2} \right) + \left(\frac{v}{2} - 1 \right) c - \log \left(\Gamma \left(\frac{v}{2} \right) \right) - \frac{v}{2} \vartheta. \end{aligned} \quad (3.2)$$

Denota-se por $\hat{\boldsymbol{\theta}}^{(k)} = (\hat{\boldsymbol{\beta}}^{(k)^T}, \hat{\Phi}^{(k)^T})^T$ a estimativa de $\boldsymbol{\theta}$ para a k -ésima iteração do algoritmo EM, que é composto pelas seguintes etapas:

Etapa (E): dado $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^{(k)}$, calcula-se $Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}^{(k)})$

Etapa (M): Atualiza-se $\hat{\boldsymbol{\theta}}^{(k+1)}$ maximizando $Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}}^{(k)})$ em $\boldsymbol{\theta}$, que leva as soluções apresentadas nas Equações (3.3) e (3.4).

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1} \mathbf{Y} \quad (3.3)$$

$$\text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \right) = \vartheta \mathbf{r}^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \mathbf{r} \quad (3.4)$$

A sequência obtida a partir das iterações do algoritmo EM converge para a estimativa de máxima verossimilhança de $\hat{\boldsymbol{\theta}}$.

3.2 Estimação de parâmetros por máxima verossimilhança para o processo t -student utilizando o algoritmo de Newton-Raphson (NR)

Seja $\hat{\boldsymbol{\theta}}^{(0)}$ a estimativa incial de $\boldsymbol{\theta}$ e $\hat{\boldsymbol{\theta}}^{(k)}$ a estimativa na k -ésima iteração. Segundo o algoritmo de Newton-Raphson, a estimativa de $\boldsymbol{\theta}$ na iteração $k + 1$ é dada por (Little e Rubin, 1987) como

$$\hat{\boldsymbol{\theta}}^{(k+1)} = \hat{\boldsymbol{\theta}}^{(k)} - \left(\frac{\partial^2 l(\hat{\boldsymbol{\theta}}^{(k)})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right)^{-1} U^{(k)}$$

onde $U^{(k)} = U(\hat{\boldsymbol{\theta}})^{(k)} = (U_{\boldsymbol{\beta}}^{(k)}, U_{\Phi}^{(k)})^T$ é o vetor score e $\partial^2 l(\hat{\boldsymbol{\theta}}^{(k)}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T$ é a matriz de informação observada na k -ésima iteração, apresentados nos anexos A.1.1 e A.1.2, respectivamente.

A partir de uma estimativa incial de $\hat{\boldsymbol{\theta}}^{(0)}$, o algoritmo vai atualizando a estimativa de $\hat{\boldsymbol{\theta}}$

até obter a convergência, o que não é garantido (ref?).

3.3 Estimação do erro padrão das estimativas dos parâmetros $\hat{\theta}$

O erro padrão das estimativas de máxima verossimilhança pode ser obtido mediante a determinação da matriz de informação esperada de Fisher para o modelo t -student, que assume a forma bloco diagonal (Lange et al., 1989).

$$\mathbf{F} = \mathbf{F}(\boldsymbol{\theta}) = \begin{pmatrix} \mathbf{F}_{\beta\beta} & 0 \\ 0 & \mathbf{F}_{\varphi\varphi} \end{pmatrix},$$

com elementos

$$\mathbf{F}_{\beta\beta} = \frac{4a_g}{n} \mathbf{X}^T \Sigma^{-1} \mathbf{X} \text{ e}$$

$$\mathbf{F}_{\varphi\varphi} = \frac{b_g}{4} \left(\frac{4f_g}{n(n+2)-1} \right) + \frac{2f_g}{n(n+2)} \text{tr} \left\{ \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \right\} \text{ para } j = 1, 2, 3,$$

em que

$$a_g = \frac{n}{4} \left(\frac{v+n}{v+n+2} \right),$$

$$b_g = \text{tr} \left\{ \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \right\} \text{tr} \left\{ \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \right\} \text{ para } j = 1, 2, 3 \text{ e}$$

$$f_g = \frac{n(n+2)}{4} \left(\frac{v+n}{v+n+2} \right).$$

A obtenção das expressões $\mathbf{F}_{\beta\beta}$, $\mathbf{F}_{\varphi\varphi}$ encontram-se no apêndice A.3.

3.4 Influência local

Considerou-se como influência local usual o estudo do comportamento do afastamento do logaritmo da função de verossimilhança de Cook e influência local pelo Q-afastamento do logaritmo da função de verossimilhança a esperança do logaritmo da verossimilhança completa, ambos sobre o modelo espacial linear t -student.

3.4.1 Influência local no afastamento da verossimilhança usual (*likelihood displacement*)

Para definir a influência local, considera-se um vetor de perturbação $\omega = (\omega_1, \dots, \omega_n)^T$ de dimensão $n \times 1$, sendo que ω pode refletir qualquer esquema de perturbação.

Sejam $\hat{\theta}$ e $\hat{\theta}_\omega$ os estimadores de MV de θ sob o modelo postulado e perturbado, respectivamente. Cook (1986) sugere estudar o comportamento do afastamento do logaritmo da função de verossimilhança definido na Equação 3.5, comparando $\hat{\theta}$ e $\hat{\theta}_\omega$ quando ω varia em Ω :

$$\text{LD}(\omega) = 2[l(\hat{\theta}) - l(\hat{\theta}_\omega)], \quad \omega \in \Omega. \quad (3.5)$$

Segundo Cook (1987), o sentido da distância entre $\hat{\theta}$ e $\hat{\theta}_\omega$, baseado em $\text{LD}(\omega)$, pode depender da concavidade do logaritmo da função de verossimilhança. Se $l(\theta)$ é suficientemente achatada, pode-se dizer que $\hat{\theta}$ e $\hat{\theta}_\omega$ estão próximos entre si, enquanto que se $l(\theta)$ for suficientemente concentrada em torno de $\hat{\theta}$, estas estimativas podem estar distantes entre si.

O gráfico de $\text{LD}(\omega)$ versus ω fornece informações essenciais sobre a influência do esquema de perturbação quando a superfície $\alpha(\omega) = (\omega^T, \text{LD}(\omega))^T$ desvia-se de seu plano tangente em ω_0 .

Essa análise pode ser realizada com o estudo das curvaturas normais, as quais correspondem às curvaturas das seções normais do plano tangente em ω_0 , conforme ilustra a Figura 3 (Verbeke e Molenberghs, 2000).

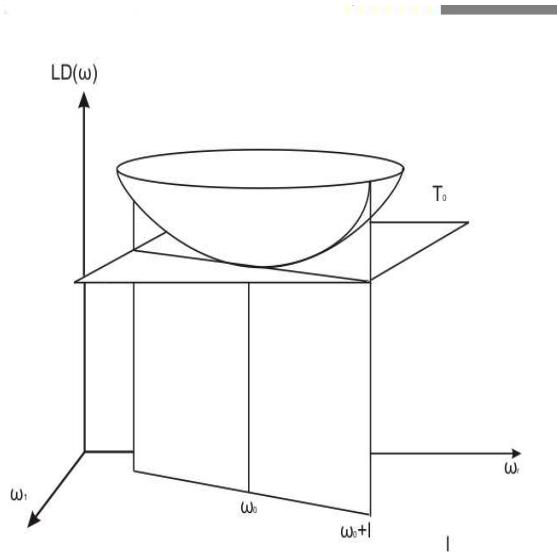


Figura 3: Representação gráfica do enfoque de influência local (Verbeke e Molenberghs, 2000).

A ideia principal do método de Cook (1986), baseado na superfície $\alpha(\omega)$ em torno de ω_0 , consiste em considerar o plano (T_0) tangente à superfície $\alpha(\omega)$ em ω_0 . Como $\text{LD}(\omega)$ atinge o mínimo em ω_0 , temos que T_0 é paralelo a Ω . Cada vetor l em ω , de norma um, determina um plano que contém l e que é ortogonal a T_0 .

Cook (1986) ainda apresenta a curvatura normal na direção de um vetor l (com $\|l\| = 1$),

como a seguir:

$$C_l = 2 [\mathbf{l}^T \Delta_{\boldsymbol{\omega}}^T \mathbf{L}^{-1} \Delta_{\boldsymbol{\omega}} \mathbf{l}],$$

em que

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

$\Delta_{\boldsymbol{\omega}}$ é a matriz delta usual de ordem $(p+3) \times n$ dada por $\Delta_{\boldsymbol{\omega}} = (\Delta_{\beta(\boldsymbol{\omega})}^T, \Delta_{\varphi(\boldsymbol{\omega})}^T)^T$, avaliada em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e em $\boldsymbol{\omega} = \boldsymbol{\omega}_0$.

Para definir as matrizes \mathbf{L} e $\Delta_{\boldsymbol{\omega}}$, considere-se a forma paramétrica da matriz escala Σ definida na Equação (2.7) e o logaritmo da função de verossimilhança para dados completos da distribuição $t_n(\boldsymbol{\mu}, \Sigma, v)$, definido na Equação (3.1).

A matriz de informação observada usual \mathbf{L} , no caso t -student n -variado com v graus de liberdade, é definida como

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

em que

$$\mathbf{L}_{\beta\beta} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \beta^T} = 2W_{g(\delta)} \mathbf{X}^T \Sigma^{-1} \mathbf{X} + 4W'_{g(\delta)} [\mathbf{X}^T \Sigma^{-1} \mathbf{r}] [\mathbf{r}^T \Sigma^{-1} \mathbf{X}];$$

$$\mathbf{L}_{\beta\varphi} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta \partial \varphi_j^T} = 2W_{g(\delta)} \mathbf{X}^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \mathbf{r} + 2W'_{g(\delta)} [\mathbf{X}^T \Sigma^{-1} \mathbf{r}] \left[\mathbf{r}^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \mathbf{r} \right],$$

para $j = 1, 2, 3$;

$$\mathbf{L}_{\beta\varphi} = \mathbf{L}_{\varphi\beta}^T \text{ e}$$

$$\mathbf{L}_{\varphi\varphi} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \varphi_j \partial \varphi_i^T} = -\frac{1}{2} \left\{ \text{tr} \left[\Sigma^{-1} \left(\frac{\partial^2 \Sigma}{\partial \varphi_j \partial \varphi_i^T} - \frac{\partial \Sigma}{\partial \varphi_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \right) \right] \right\} +$$

$$+ W_{g(\delta)} \left\{ \mathbf{r}^T \Sigma^{-1} \left[\frac{\partial^2 \Sigma}{\partial \varphi_j \partial \varphi_i^T} - \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_i} - \frac{\partial \Sigma}{\partial \varphi_i} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \right] \Sigma^{-1} \mathbf{r} \right\} +$$

$$+ W'_{g(\delta)} \left\{ \left[\mathbf{r}^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \mathbf{r} \right] \left[\mathbf{r}^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_i} \Sigma^{-1} \mathbf{r} \right] \right\}, \text{ para } i, j = 1, 2, 3;$$

sendo que $W_{g(\delta)} = -\frac{1}{2} \left(\frac{v+n}{v+\delta} \right)$, $W'_{g(\delta)} = - \left(\frac{v+n}{(v+\delta)^2} \right)$, $\delta = \mathbf{r}^T \Sigma^{-1} \mathbf{r}$ e $\mathbf{r} = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$.

Para determinar a matriz delta usual, considera-se o esquema de perturbação linear na variável resposta, da forma $\tilde{Y} = Y + \epsilon$, o qual permite verificar o mau condicionamento de Y . Dessa forma a matriz delta usual é dada por

$$\epsilon_{(1)} = 2W_{g((1))}X^{-1} + 2W_{g((1))}^0X^{-1}r_{(1)}r_{(1)}^T e$$

$$\epsilon'_{(1)} = 2r_{(1)}^T - \frac{\partial}{\partial x_j} W_{g((1))} + W_{g((1))}^0 r_{(1)}^T r_{(1)}^{-1} r_{(1)}^T e$$

onde

$$W_{g((1))} = \frac{1}{2} \frac{v+n}{v+(1)}, \quad W_{g((1))}^0 = \frac{v+n}{(v+(1))^2}, \quad (1) = r_{(1)}^T r_{(1)} e$$

$$r_{(1)} = (Y \quad X).$$

Seja a matriz

$$B = \tilde{Y} L^{-1} \epsilon;$$

se $C_i = 2 \sum_j b_{ij}$, em que b_{ij} são os elementos da diagonal principal da matriz L , pode-se utilizar o gráfico C_i versus i (ordem dos dados), como técnica de diagnóstico para compreensão da existência de dados inuentes.

Ainda sobre a matriz B , considera-se e_{max} como sendo o autovetor normalizado associado ao maior autovalor, em módulo, dessa matriz. Então, os elementos de e_{max} versus i (ordem dos dados), temos um gráfico que pode revelar qual parte da perturbação que tem a maior influência em $LD(\epsilon)$ em ϵ_0 (Cook, 1986).

Segundo Cysneiros et al. (2005), o gráfico e_{max} versus j pode mostrar como se deve perturbar, por exemplo, o parâmetro de escala para obter maiores mudanças nas estimativas de

3.4.2 Influência local na matriz de covariância usual

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

A influência local de primeira ordem da perturbação (nesta $\tilde{Y} = Y + \epsilon$) pode

ser medida usando a inclinação na direção denotada com $S(I)$, do gráfico de influência de $C_h(\hat{I})$ versus I . Neste caso tem-se,

$$S(I) = I^T C_h;$$

em que C_h é um vetor de 1 dado por

$$C_h = \left[L^{-1} \frac{\partial C(\cdot)}{\partial I_j} \right]_{j=1}^n;$$

sendo L é a matriz de observação usual,

$$\begin{aligned} \frac{\partial C(\cdot)}{\partial I^T} &= 0^T; \frac{\partial C(\cdot)}{\partial I^T} \quad \text{e} \\ \frac{\partial C(\cdot)}{\partial I^T} &= \frac{\partial C(\cdot)}{\partial I_1}, \frac{\partial C(\cdot)}{\partial I_2}, \frac{\partial C(\cdot)}{\partial I_3}. \end{aligned}$$

A direção de máxima inclinação local é dada por:

$$I_{C_h} = \frac{C_h}{\|C_h\|};$$

Utiliza-se o gráfico C_h versus I (ordem dos dados) para avaliar a sensibilidade da função de covariância.

3.4.3 Influência local no preditor linear usual

No estudo de dados espaciais tem-se como medida de preditor da krigagem universal, que visa obter valores para a variável regionalizada em pontos amostrados. Seja $y_0 = Y(s_0)$, o preditor da krigagem universal na localização D . A média de y_0 é dada por x_0^T , onde $x_0^T = (x_{01}; \dots; x_{0p})$ e $x_{0j} = x_j(s_0)$ para $j = 1, \dots, p$.

O preditor do menor erro quadrático médio é dado por

$$p_Q(s_0; \cdot) = x_0^T + C_0^T I^{-1}(Y - X \cdot);$$

sendo $C_0^T = (C(h_{10}); \dots; C(h_{n0}))$, com $h_{i0} = ks_i - s_0k$ para $i = 1, \dots, n$.

Então, um estimador pontual para é

$$\hat{y}_0 = p_Q(s_0; \hat{I}) = x_0^T \hat{I} + C_0^T \hat{I}^{-1}(Y - X \hat{I});$$

Tem-se que

$$S(l) = l^T p(s_0; \cdot);$$

em que $p(s_0; \cdot)$ é um vetor $n \times 1$ dado por

$$p(s_0; \cdot) = l^T L^{-1} \frac{\partial p(s_0; \cdot)}{\partial l} \quad : \quad (3.6)$$

Na Equação (3.6), tem-se que $\frac{\partial p(s_0; \cdot)}{\partial l} = \left[\frac{\partial p(s_0; \cdot)}{\partial l^T}, \frac{\partial p(s_0; \cdot)}{\partial l^T} \right]^T$,
em que:

$$\frac{\partial p(s_0; \cdot)}{\partial l^T} = x_0^T X^{-1} C_0;$$

$$\frac{\partial p(s_0; \cdot)}{\partial l^T} = \left[\frac{\partial p(s_0; \cdot)}{\partial l_j^T} \right]^T;$$

sendo que

$$\frac{\partial p(s_0; \cdot)}{\partial l_j^T} = \frac{\partial \tilde{C}}{\partial l_j^T} C_0^T l^{-1} \frac{\partial}{\partial l_j^T} l^T (Y - X \beta) \text{ e}$$

$$\frac{\partial \tilde{C}}{\partial l_j^T} = \left[\frac{\partial \tilde{C} h_{10}}{\partial l_j^T}, \dots, \frac{\partial \tilde{C} h_{n0}}{\partial l_j^T} \right]^T, \text{ para } j = 1, 2, \dots, q.$$

A direção de máxima inclinação local é obtida por:

$$l_p = \frac{p(s_0; \cdot)}{kp(s_0; \cdot)k};$$

Utiliza-se o gráfico l_p versus i (ordem dos dados) para avaliar a influência no preditor linear.

3.4.4 Influência local nos valores preditos usual (generalized leverage)

O método da alavancagem (*generalized leverage*) tem como objetivo principal medir a influência da resposta observada no seu próprio valor ajustado, seja, avaliar a influência local nos valores preditos (Wei et al., 1998). St. Laurent e Cook (1992) generalizaram esse método para modelos mais complexos, como modelos não-lineares. Chen et al., (1992b) sugeriram uma medida de alavancagem para modelos lineares mistos. Paula (1999) considerou o método de alavancagem em modelos de regressão, usando o vetor de parâmetros

é restrito por desigualdades lineares e, mais recentemente, Nobile e Singer (2006) propuseram incorporar a informação dos efeitos aleatórios ajustados. Consideraram a matriz de alavancas generalizada em modelos lineares com efeitos mistos.

A proposta do método de alavancas é medir a influência que ~~o resultado~~ variável resposta observada exercem sobre a predição de seus valores ~~esassim~~ amostrados.

Seja $\mu = X$ o valor esperado da variável, de modo em (2.4), $\hat{Y} = X \wedge$ o seu valor estimado utilizando máxima verossimilhança e $(\beta; \sigma^2)^T$ os parâmetros do modelo, então, segundo Weit al. (1998), a matriz de alavancas generalizada dada por

$$LG(\mu) = \frac{\partial \hat{Y}}{\partial \beta^T} = \frac{\partial \hat{Y}}{\partial \beta} \frac{\partial \beta}{\partial \beta^T};$$

assume a forma:

$$LG(\mu) = D(L^{-1}L_Y);$$

em que:

$$D = \frac{\partial \hat{Y}}{\partial \beta^T} = (X; 0) \text{ e}$$

$$L_Y = \frac{\partial I(\mu)}{\partial \beta^T} = (L_Y^T; L^T Y)^T; \text{ com } L^T Y = X^T - 1 \text{ e}$$

$$L^T Y = \frac{\partial I(\mu)}{\partial \beta_j^T}, \text{ onde } \frac{\partial I(\mu)}{\partial \beta_j^T} = r^T - 1 \frac{\partial \mu}{\partial \beta_j} - 1, \text{ para } j = 1, 2, \dots, q.$$

Logo

$$LG(\mu) = X(L + L^T L^{-1} L^T)^{-1}(L^T L^{-1} L^T Y - L^T Y)$$

ou ainda

$$LG(\mu) = LG_1 + LG_2;$$

sendo

$$LG_1 = X(L + L^T L^{-1} L^T)^{-1} L^T Y \text{ e}$$

$$LG_2 = X(L + L^T L^{-1} L^T)^{-1}(-L^T L^{-1} L^T Y).$$

3.5 Influência local segundo a proposta do Q-afastamento

Zhu e Lee (2001) apresentaram uma alternativa ao procedimento desenvolvido na seção (3.4.1), a qual propõe uma função de deslocamento $f_Q(\mu)$ para substituir $LD(\mu)$. Essa função é dada por

$$f_Q(\mu) = 2 \sum_{j=1}^h Q(\mu_j) - Q(\mu_{j+1}) ; \quad (3.7)$$

em que $Q(\hat{j})$ é a esperança condicional do logaritmo da função de verossimilhança dos dados completos de nida na Equação (3.2). $\hat{\theta}$ é a estimativa de que maximiza $Q(\cdot; \hat{j})$.

O gráfico de inuência local de $Q(\cdot)$ é de nido como $(\cdot) = (\cdot^T; f_Q(\cdot))^T$. Nesse caso a curva normal $C_{f_Q; l}$ de (\cdot) em $\hat{\theta}_0$ na direção de algum vetor unitário pode ser usada para resumir o comportamento local de $f_Q(\cdot)$.

A curvatura normal $C_{f_Q; l}$ para (\cdot) em $\hat{\theta}_0$ é de nido como

$$C_{f_Q; l} = 2 \cdot l^T B_Q \cdot l ; \quad (3.8)$$

em que

$$B_Q = \begin{smallmatrix} & n & 0 & \dots & 0 \\ T & Q & & & \vdots \\ & & & & 1 \end{smallmatrix} ;$$

sendo

$$Q = \frac{\partial Q(\hat{j})}{\partial \theta^T} j = \hat{j}^T e \cdot \hat{\theta} = \frac{\partial Q(\cdot; \hat{j})}{\partial \theta^T} j = \hat{j}^T (\cdot) ;$$

ambas avaliadas em $\hat{j}^T e \cdot \hat{\theta} = \hat{\theta}_0$.

Para definir as matrizes Q e $\hat{\theta}$, também considera-se a forma paramétrica da matriz escala $\hat{\theta}$, de nida na Equação (2.7).

A matriz hessiana Q para a distribuição studentn-variada com graus de liberdade é obtida sobre $Q(\hat{j})$ da forma:

$$Q = \begin{smallmatrix} 0 & & & 1 \\ @ & Q & Q & A \\ & Q & Q & \end{smallmatrix} ;$$

sendo:

$$Q = \frac{\partial Q(\hat{j})}{\partial \theta^T} = \#X^T X^{-1} X ;$$

$$Q = \frac{\partial Q(\hat{j})}{\partial \theta_j^T} = \#X^T \frac{\partial}{\partial \theta_j^T} r \text{ para } j = 1, 2, 3 ;$$

$$Q = Q^T \text{ e}$$

$$Q = \frac{\partial Q(\hat{j})}{\partial \theta_i^T \partial \theta_i^T} = \frac{1}{2} \text{tr} \left(\frac{\partial}{\partial \theta_i^T} \frac{\partial}{\partial \theta_i^T} \right) +$$

$$+ \frac{1}{2} \#r^T \left(\frac{\partial}{\partial \theta_i^T} \frac{\partial}{\partial \theta_i^T} \right)^{-1} \frac{\partial}{\partial \theta_i^T} \frac{\partial}{\partial \theta_i^T} \frac{\partial}{\partial \theta_i^T} \frac{\partial}{\partial \theta_i^T} \frac{\partial}{\partial \theta_i^T} \text{ para } i = 1, 2, 3 ,$$

onde

$$\vartheta = E \{ \mathbf{U} | \mathbf{Y}, \boldsymbol{\theta} \} = \left(\frac{v+n}{v+\delta} \right); \delta = \mathbf{r}^T \boldsymbol{\Sigma}^{-1} \mathbf{r} \text{ e } \mathbf{r} = (\mathbf{Y} - \mathbf{X} \boldsymbol{\beta}).$$

A matriz $\Delta_{\boldsymbol{\omega}}$, obtida a partir de $Q(\boldsymbol{\theta} | \hat{\boldsymbol{\theta}})$, considerando-se o mesmo esquema de perturbação linear, onde $\mathbf{Y}_{\boldsymbol{\omega}} = \mathbf{Y} + \boldsymbol{\omega}$, é dada por:

$$\Delta_{\boldsymbol{\beta}(\boldsymbol{\omega})} = \vartheta_{\boldsymbol{\omega}} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} - 2\vartheta'_{\boldsymbol{\omega}} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{r}_{(\boldsymbol{\omega})} \mathbf{r}_{(\boldsymbol{\omega})}^T \boldsymbol{\Sigma}^{-1} \text{ e}$$

$$\Delta_{\Phi(\boldsymbol{\omega})} = \mathbf{r}_{(\boldsymbol{\omega})}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{-1} \left\{ \vartheta_{\boldsymbol{\omega}} + \vartheta'_{\boldsymbol{\omega}} \boldsymbol{\Sigma}^{-1} \mathbf{r}_{(\boldsymbol{\omega})} \mathbf{r}_{(\boldsymbol{\omega})}^T \right\},$$

sendo que

$$\vartheta_{(\boldsymbol{\omega})} = \frac{v+n}{v+\delta_{\boldsymbol{\omega}}(\hat{\boldsymbol{\theta}})}, \delta_{(\boldsymbol{\omega})} = \mathbf{r}_{(\boldsymbol{\omega})}^T \boldsymbol{\Sigma}^{-1} \mathbf{r}_{(\boldsymbol{\omega})} \text{ e } \mathbf{r}_{(\boldsymbol{\omega})} = (\mathbf{Y}_{\boldsymbol{\omega}} - \mathbf{X} \boldsymbol{\beta}).$$

3.5.1 Curvatura normal conformal

Com o objetivo de obter uma curvatura invariante na troca da escala, Poon e Poon (1999) introduziram a curvatura normal conformal $B_{f_Q,l}$ para $\boldsymbol{\omega}_0$ na direção de algum vetor unitário \mathbf{l} e a definem como

$$B_{f_Q,l} = \frac{C_{f_Q,l}}{\|\mathbf{A}\|_F}$$

onde $\|\cdot\|_F$ é definida como: $\|\mathbf{A}\|_F = [\text{tr}(\mathbf{A} \mathbf{A}^T)]^{1/2}$.

Neste caso, $\mathbf{A} = 2 \left\{ \boldsymbol{\Delta}_{\boldsymbol{\omega}}^T \left[-\ddot{\mathbf{Q}} \right]^{-1} \boldsymbol{\Delta}_{\boldsymbol{\omega}} \right\}$, avaliadas em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$.

Uma propriedade interessante da curvatura normal conformal é que, para qualquer direção unitária \mathbf{l} , se tem que $0 \leq B_{f_Q,l} \leq 1$. Isso permite, por exemplo, a comparação da curvatura entre diferentes modelos.

3.5.2 Influência local no afastamento da verossimilhança utilizando o Q-afastamento da função verossimilhança completa (*Q-displacement*)

Dada a matriz $B_{f_Q,l}$ e considerando-se

$$C_{Qi} = 2 * |b_{ii}|, \quad (3.9)$$

em que b_{ii} são os elementos da diagonal principal da matriz $B_{f_Q,l}$, obtém-se o gráfico C_{Qi} versus i (ordem dos dados), que, assim como o gráfico C_i , tem por objetivo avaliar a existência de observações influentes, assumindo-se distribuição *t*-student n -variada aos dados.

Segundo Zhu e Lee (2001), é usual considerar o i -ésimo ponto como inente se

$$b_i > B_{f_Q;i} + 2S_{B_{f_Q};i}; \text{ para } i = 1, \dots, n; \quad (3.10)$$

em que

$$B_{f_Q;i} = \frac{1}{n} \sum_{i=1}^n b_i \text{ e } S_{B_{f_Q};i}^2 = \frac{1}{n} \sum_{i=1}^n (b_i - B_{f_Q;i})^2.$$

Considere-se $\mathbf{e}_{Q_{\max}}$ o autovetor, normalizado, associado ao maior autovalor mádulo, da matriz $B_{f_Q;i}$, sendo esse autovetor utilizado na construção do gráco de sentos $\mathbf{l}_{Q_{\max}}$ versus i (ordem dos dados).

3.5.3 Inencia local na matriz de covariância utilizando o Q-afastamento da verossimilhana completa (Q-displacement)

Como foi denotado em 3.4. $\mathbf{C}_h(\cdot)$ é a covariância para uma distância fixa, dependente do vetor de parâmetros. Considere-se a inencia local de primeira ordem da perturbação medida usando a inclinação na direção \mathbf{l} , notada com $\mathbf{s}_Q(\mathbf{l})$, do gráco de inencia de $\mathbf{C}_h(\cdot)$ versus \mathbf{l} . Assim, tem-se,

$$\mathbf{s}_Q(\mathbf{l}) = \mathbf{l}^T \mathbf{Q}_{Ch};$$

em que \mathbf{Q}_{Ch} é um vetor $n \times 1$ dado por

$$\mathbf{Q}_{Ch} = \mathbf{l}^T \mathbf{Q}^{-1} \frac{\partial \mathbf{C}(\cdot)}{\partial \mathbf{l}} \Big|_{\mathbf{l}=\mathbf{0}};$$

sendo que as matrizes \mathbf{l} e \mathbf{Q} são definidas na seção 4.1.4,

$$\begin{aligned} \frac{\partial \mathbf{C}(\cdot)}{\partial \mathbf{l}^T} &= \mathbf{0}^T; \frac{\partial \mathbf{C}(\cdot)}{\partial \mathbf{l}^T} \text{ e} \\ \frac{\partial \mathbf{C}(\cdot)}{\partial \mathbf{l}^T} &= \frac{\partial \mathbf{C}(\cdot)}{\partial l_1}, \frac{\partial \mathbf{C}(\cdot)}{\partial l_2}, \frac{\partial \mathbf{C}(\cdot)}{\partial l_3}. \end{aligned}$$

A direção de máxima inclinação local é dada por

$$\mathbf{l}_{Q_{Ch}} = \frac{\mathbf{Q}_{Ch}}{\|\mathbf{Q}_{Ch}\|};$$

3.5.4 Influência local no preditor linear utilizando o Q-afastamento da verosimilhança completa (*Q-displacement*)

Considere-se $Y_0 = Y(s_0)$ o preditor da krigagem universal na localização $s_0 \in D$. A média de Y_0 é dada por $x_0^T \boldsymbol{\beta}$, onde $x_0^T = (x_{01}, \dots, x_{0p})$ e $x_{0j} = x_j(s_0)$ para $j = 1, \dots, p$.

O preditor do menor erro quadrático médio, como foi denotado em 3.4.3, é dado por

$$p_Q(s_0, \boldsymbol{\theta}) = x_0^T \boldsymbol{\beta} + C_0^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X} \boldsymbol{\beta}),$$

onde $C_0^T = (C(h_{10}), \dots, C(h_{n0}))$, com $h_{i0} = \|s_i - s_0\|$ para $i = 1, \dots, n$.

Então, um estimador pontual para Y_0 é

$$\hat{Y}_0 = p_Q(s_0, \hat{\boldsymbol{\theta}}) = x_0^T \hat{\boldsymbol{\beta}} + \hat{C}_0^T \hat{\boldsymbol{\Sigma}}^{-1} (\mathbf{Y} - \mathbf{X} \hat{\boldsymbol{\beta}}).$$

Tem-se que

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

em que $\dot{p}(s_0, \boldsymbol{\theta})$ é um vetor $n \times 1$ dado por

$$\dot{p}_Q(s_0, \boldsymbol{\theta}) = \left\{ -\Delta_{\boldsymbol{\omega}}^T \ddot{\mathbf{Q}}^{-1} \frac{\partial p_Q(s_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \right\}. \quad (3.11)$$

Na Equação (3.11), tem-se que

$$\frac{\partial p_Q(s_0, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left(\frac{\partial p_Q(s_0, \boldsymbol{\theta})}{\partial \boldsymbol{\beta}^T}, \frac{\partial p_Q(s_0, \boldsymbol{\theta})}{\partial \Phi^T} \right)^T,$$

em que

$$\frac{\partial p_Q(s_0, \boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = x_0 - \mathbf{X}^T \boldsymbol{\Sigma}^{-1} C_0 \text{ e}$$

$$\frac{\partial p_Q(s_0, \boldsymbol{\theta})}{\partial \Phi} = \left[\left(\frac{\partial p_Q(s_0, \boldsymbol{\theta})}{\partial \varphi_j} \right) \right],$$

sendo que

$$\frac{\partial p_Q(s_0, \boldsymbol{\theta})}{\partial \varphi_j} = \left\{ \frac{\partial C_0^T}{\partial \varphi_j} - C_0^T \boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \right\} \boldsymbol{\Sigma}^{-1} \mathbf{r} \text{ e}$$

$$\frac{\partial C_0^T}{\partial \varphi_j} = \left(\frac{\partial C(h_{10})}{\partial \varphi_j}, \dots, \frac{\partial C(h_{n0})}{\partial \varphi_j} \right), \text{ para } j = 1, 2, 3.$$

Nesse caso, a direção de máxima inclinação local é obtida por:

$$l_{Qp} = \frac{\dot{p}_Q(s_0, \boldsymbol{\theta})}{\|\dot{p}_Q(s_0, \boldsymbol{\theta})\|}.$$

O gráfico $l_{Q_{Ch}}$ versus i (ordem dos dados) e o gráfico l_{Q_p} versus i (ordem dos dados) são utilizados para avaliar a influência na matriz de covariância e no preditor linear, respectivamente.

3.5.5 Influência local nos valores preditos utilizando o Q-afastamento da verossimilhança completa (Q -displacement)

Baseado na proposta de Wei *et al.* (1998), assim como nos trabalhos de Zhu e Lee (2001) e Zhu *et al.* (2001), Salgado (2006) propõe “a matriz de alavancas generalizadas para modelos com dados incompletos”.

Considere-se $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ o valor estimado de \mathbf{Y} pela máxima verossimilhança completa, sendo $\boldsymbol{\theta} = (\boldsymbol{\beta}^T, \boldsymbol{\Phi}^T)^T$ os parâmetros do modelo e $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$ é o valor esperado da variável \mathbf{Y} .

A matriz de alavancas generalizadas para dados incompletos é dada por

$$\text{LG}_Q(\boldsymbol{\theta}) = \frac{\partial \hat{\mathbf{Y}}}{\partial \hat{\mathbf{Y}}^T} = \frac{\partial \hat{\mathbf{Y}}}{\partial \boldsymbol{\theta}} \frac{\partial \boldsymbol{\theta}}{\partial \mathbf{Y}^T},$$

Logo,

$$\text{LG}_Q(\boldsymbol{\theta}) = D_{\boldsymbol{\theta}} \left[-\ddot{Q}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}}) \right]^{-1} Q_{\boldsymbol{\theta}, \mathbf{Y}}(\hat{\boldsymbol{\theta}}),$$

onde $Q_{\boldsymbol{\theta}, \mathbf{Y}}(\hat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \mathbf{Y}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$ e $Q(\hat{\boldsymbol{\theta}}) = \frac{\partial^2 Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$, sendo que a função $Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})$ é proveniente da esperança condicional do logaritmo da função verossimilhança para dados completos definido na Equação (3.2).

3.6 Simulação de processos espaciais

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 sequência aleatória respeite uma matriz de correlações.

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

Seja $\mathbf{Y} = (Y(s_1), \dots, Y(s_n))^T$ o vetor $n \times 1$ dos dados simulados, os quais representam a realização de um processo estocástico ou função de variáveis aleatórias $Y(s_i)$, $s \in S$, em que $s \subset \mathbb{R}^2$ e \mathbb{R}^2 é um espaço eucladiano bi-dimensional em diferentes localizações s_1, \dots, s_n .

Considere-se agora o vetor de médias do processo, $\boldsymbol{\mu} = (\mathbb{E}(Y(s_1)), \dots, \mathbb{E}(Y(s_n)))^T$, e a matriz de covariância para a distribuição t -student n -variada, definida na Equação (2.8). Para processos estocásticos que satisfazem a hipótese de estacionaridade de segunda ordem e

isotropia, tem-se:

$$\mathbb{E}[Y(s_i)] = \mu, \quad (3.12)$$

para $i = 1, 2, \dots, n$ e

$$C(Y(s_i), Y(s_u)) = C(h_{iu}), \text{ onde } h_{iu} = \|s_i - s_u\|. \quad (3.13)$$

Nesse caso, cada elemento do vetor é igual a um valor constante μ e cada (i, u) -ésimo elemento da matriz $\Sigma, n \times n$, é igual a $C(h_{iu})$.

Assim, escolhendo-se o valor de μ e a função covariância $C(h)$, o vetor \mathbf{Y} , satisfazendo as Equações (3.12) e (3.13), pode ser simulado pela relação:

$$\mathbf{Y} = \boldsymbol{\mu} + \mathbf{L}\boldsymbol{\epsilon} \quad (3.14)$$

em que \mathbf{L} é uma matriz triângular inferior $n \times n$, obtida mediante a decomposição de Σ no produto $\mathbf{L}\mathbf{L}^T$, chamada de decomposição de Cholesky, e $\boldsymbol{\epsilon} = (\epsilon(s_1), \dots, \epsilon(s_n))^T$ é um vetor de variáveis aleatórias não correlacionadas.

Seja $\epsilon(s_1), \dots, \epsilon(s_n)$ identicamente distribuídos, com distribuição t -student de probabilidade, média igual a zero, variância igual a 1 e v grau de liberdade fixo.

Para que os dados simulados tivessem uma estrutura de dependência espacial conhecida, foi fornecido o modelo teórico exponencial, o vetor de parâmetros de dependência espacial do tipo $\Phi = (\varphi_1, \varphi_2, \varphi_3)^T$ e os graus de liberdade iguais a $v = 3, v = 5, v = 7, v = 10, v = 100$ e $v = 1000$. Considerou-se a média do processo $\mu = 50$ e utilizou-se o vetor de parâmetros espaciais igual $\Phi = (0, 1, 2.5)$. Os dados foram simulados em grade regular, com distância de 1 u.m. entre pontos, totalizando 144 pontos.

3.7 Perturbação dos dados simulados

Seja \mathbf{Y} o vetor $n \times 1$ de dados simulados. O vetor de dados perturbados será \mathbf{Y}_ω , em que $\mathbf{Y}_\omega = \mathbf{Y} + \omega$, com $\omega = (\omega_1, \dots, \omega_n)^T$ o vetor de perturbação das respostas, pertencente a um espaço de perturbações Ω ($\omega \in \Omega$). O esquema de perturbação utilizado foi:

$$\mathbf{Y}_{max^*} = \mathbf{Y}_{max} + \sqrt{\|\mathbf{Y}\|}, \quad (3.15)$$

proposto por Ortega *et al.* (2002), em que:

\mathbf{Y}_{max^*} : novo valor máximo do vetor \mathbf{Y} ;

\mathbf{Y}_{max} : valor máximo do vetor \mathbf{Y} , e

$$\mathbf{p} \frac{\partial}{\partial Y_k} = \mathbf{p} \frac{\partial}{\partial y_1^2 + \dots + y_n^2}, \text{ a perturbação.}$$

Sendo assim, o vetor de perturbações tem a forma $(0; \dots; \mathbf{p} \frac{\partial}{\partial Y_k}; \dots; 0)^T$.

3.8 Estudo com dados experimentais

Para ilustrar as metodologias apresentadas nas seções anteriores, realizou-se a análise de dois estudos com conjuntos de dados experimentais reais. A análise de dados reais é univariada, realizada para o nutriente potássio do ano agrícola 2005/2006, e a segunda análise é multivariada, realizada com a produtividade de soja do ano agrícola 2006/2007, considerando como suas covariáveis a resistência do solo à penetração (RSP), densidade do solo (Des), ambas em três diferentes camadas (0; 10m, 0; 10 - 0; 20m e 0; 20 - 0; 30m de profundidade).

Realizou-se experimentos em uma área de 71 ha localizada na região Oeste do Paraná, aproximadamente 26,95°S de latitude, 53,57°Oeste de longitude e altitude média de 650 m. O solo desta área é classificado como Latossolo Vermelho Distroférico e o clima da região é temperado mesotérmico e superúmido climático Cfa (Köeppen). Todas as amostras foram georreferenciadas com auxílio do receptor de dados GPS (Global Positioning System) GEOEXPLORE 3, num sistema espacial de coordenadas (UTM).

No ano agrícola de 2005/2006, realizou-se análises laboratoriais a partir da coleta de amostras de solo dessa área, obtendo-se os dados do potássio [mol dm⁻³], num total de 46 observações, segundo o croqui da área apresentado na Figura 20(b).

A amostragem do solo para medir o nível da variável química foi realizada nas proximidades do ponto demarcado, sendo coletadas quatro amostras de solo, de 0,0 a 0,2 m de profundidade. As subamostras foram misturadas e colocadas em sacos plásticos, com aproximadamente 500 gramas, compondo, assim, a amostra representativa da parcela. Após o trabalho de campo, as amostras foram encaminhadas ao laboratório e procedeu-se a análise química, determinando os níveis de potássio.

Segundo a Coamo/Coodetec (1998), o potássio é classificado como baixo quando apresenta valores até 0,10 mol dm⁻³, médio acima de 0,10 mol dm⁻³ e abaixo de 0,20 mol dm⁻³, alto acima de 0,20 mol dm⁻³ e abaixo de 0,30 mol dm⁻³ e muito alto se acima de 0,30 mol dm⁻³.

No ano agrícola de 2006/2007 foram coletados dados da Produtividade da soja (Prod) [t ha⁻¹], Densidade do solo (Des) [mg m⁻³] e Resistência mecânica do solo à penetração (RSP) [MPa], totalizando 47 observações para cada uma das variáveis.

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 comprimento, representando a parcela. Após a trilha, foi feita a pesagem de grãos da parcela e verificado o teor de água para posterior correção para 13%. O valor foi também convertido em t ha⁻¹.

A variável densidade do solo foi coletada em três camadas: nas camadas de 0,00 a 0,10 m; 0,10 a 0,20 m e 0,20 a 0,30 m de profundidade, a partir da superfície, pelo método do anel volumétrico (Embrapa, 1997).

As determinações da resistência do solo à penetração foram realizadas com um penetrômetro. Este aparelho, desenvolvido no laboratório de mecanização agrícola da UNIOESTE (Tieppo, 2004), possui uma haste de 60 cm e fornece dados da resistência em Kgf a cada 0,25 segundos, à medida em 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. Considerou-se a classificação de Canarache (1990), em que valores da RPS no intervalo de [2,6 – 5,0] Mpa limitam o crescimento da raiz e no intervalo de [1,1 – 2,59] Mpa, apresentam pouca limitação da raiz.

3.9 Recursos computacionais

O pacote estatístico R na versão 2.9.2 (R Development Core Team, 2009), o módulo geoestatístico geoR (Ribeiro Jr; Diggle, 2001) e o pacote Mass (Ihaka & Gentleman, 1996) foram utilizados para a simulação, análises estatística e geoestatística dos dados.

4 RESULTADOS E DISCUSSÃO

4.1 Estudos de influência local aplicados a dados simulados

As Figuras 4 a 9 apresentam os gráficos *Boxplot* construídos a partir dos dados simulados sem perturbação e com perturbação, com graus de liberdade iguais a $v = 3, 5, 7, 10, 100$ e 1000 . Os dados simulados com graus de liberdade $v = 3, 5$ e 7 possuem pontos discrepantes que tornam as caudas mais longas. No entanto, os dados simulados com graus de liberdade $v = 10, 100$ e 1000 apresentam distribuição mais simétrica, sem pontos discrepantes, conforme espera-se, já que, quando $v \rightarrow \infty$, obtém-se a distribuição normal.

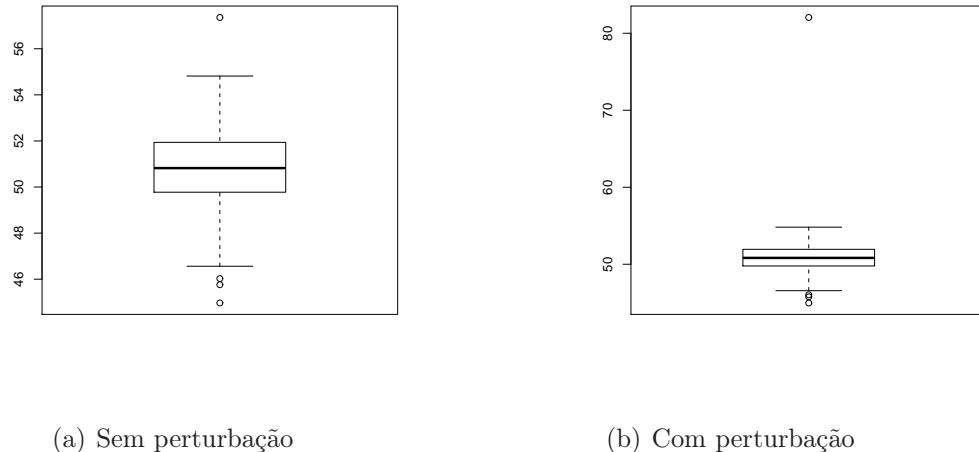


Figura 4: *Boxplot* dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 3$.

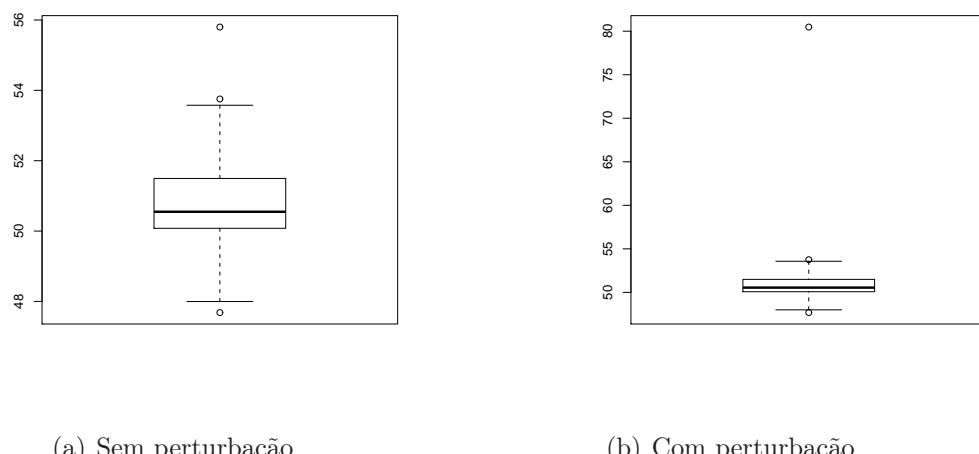


Figura 5: *Boxplot* dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 5$.

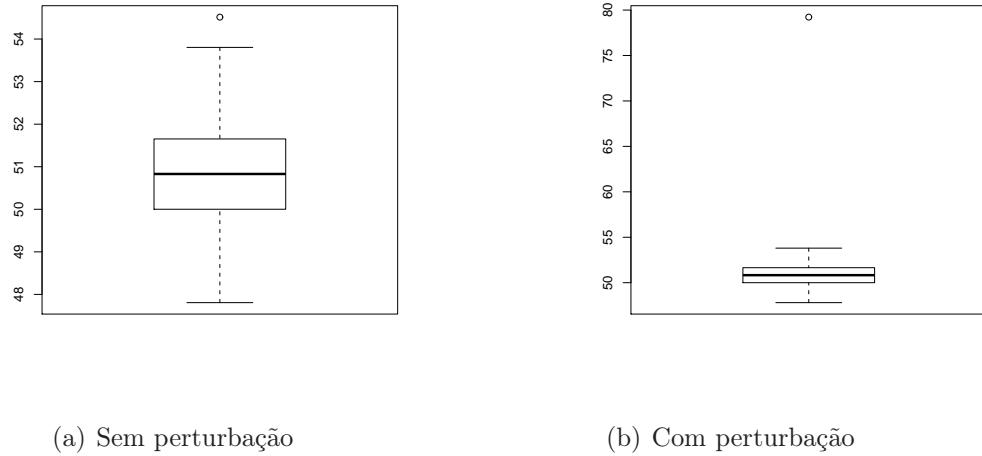


Figura 6: *Boxplot* dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 7$.

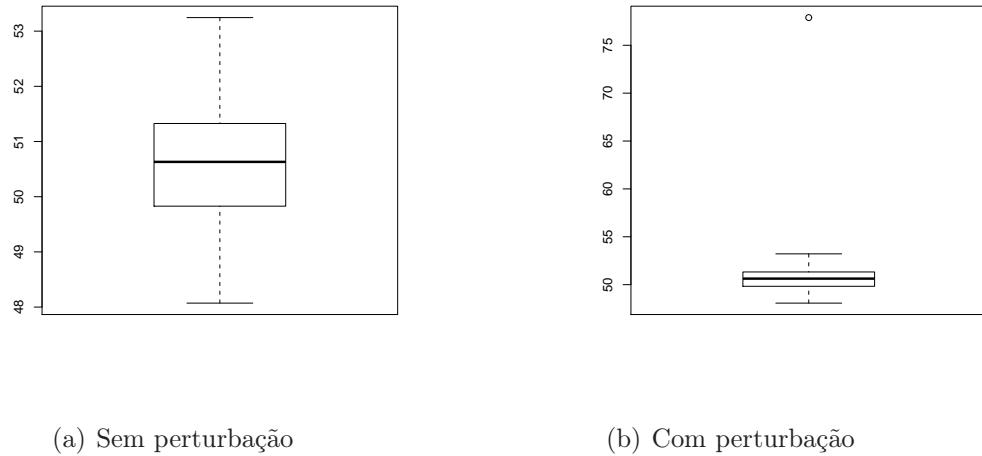


Figura 7: *Boxplot* dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 10$.

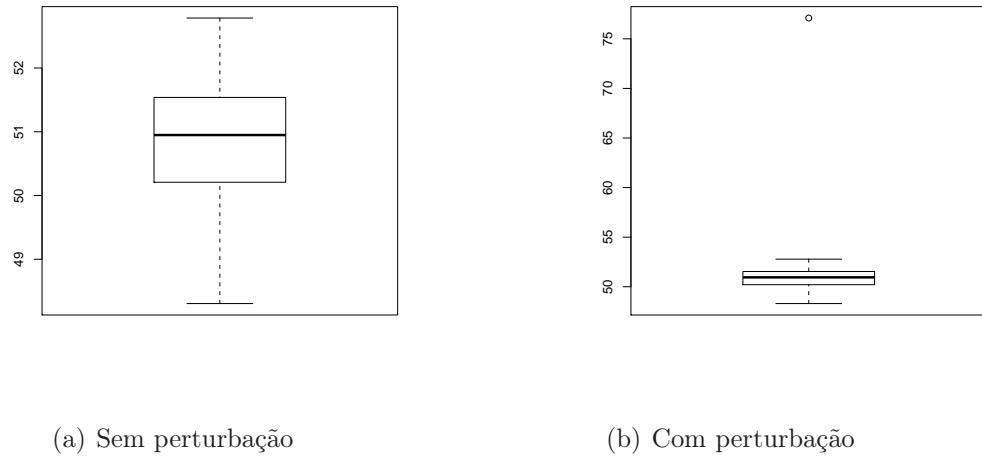


Figura 8: *Boxplot* dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 100$.

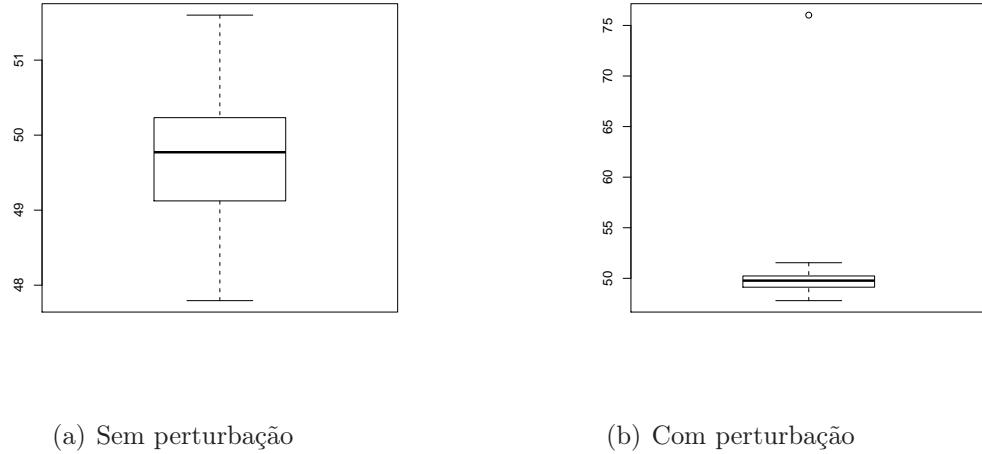


Figura 9: *Boxplot* dos dados simulados sem perturbação e com perturbação, com grau de liberdade igual a $v = 1000$.

Os pontos discrepantes que aparecem nos gráficos *Boxplot* são identificados na Tabela 1. Os dados simulados com $v = 3$, Figura 4(a), apresentam quatro pontos discrepantes, sendo que, destes, três correspondem a valores mínimos e um corresponde ao valor de máximo. Para os dados simulados com $v = 5$, Figura 5(a), têm-se três pontos discrepantes, sendo um de mínimo e dois de máximo e, ainda, os dados simulados com $v = 7$, Figura 6(a), apresentam um ponto discrepante de máximo.

Tabela 1: Identificação dos pontos discrepantes nos dados simulados

	$v = 3$				$v = 5$			$v = 7$
Valor	44,97	45,76	46,02	57,36	47,68	53,75	55,80	54,52
Ordem	(25)	(14)	(26)	(48)	(136)	(14)	(46)	(122)

Os valores de máximo de cada um dos conjuntos de dados simulados foram substituídos pelos valores perturbados, conforme a Equação (3.15).

As Tabelas 2 e 3 apresentam o elemento perturbado e as estimativas dos parâmetros sem e com a perturbação pelos algoritmos EM e NR. Como é possível observar, as estimativas obtidas para $\hat{\beta}$ pelo algoritmo EM apresentaram aumento com a perturbação, exceto nos casos em que $v = 7$ e $v = 10$, em que houve pequena redução. As estimativas obtidas para $\hat{\varphi}_1$ com a perturbação foram superiores em todos os casos, enquanto que para $\hat{\varphi}_2$ apresentaram redução na maioria dos casos e para $\hat{\varphi}_3$ não apresentaram grandes variações. Pelo algoritmo de NR as estimativas obtidas para $\hat{\beta}$ apresentaram aumento com a perturbação nos casos em que $v = 3$, $v = 5$ e $v = 1000$, e nas demais houve pequena redução. As estimativas obtidas para $\hat{\varphi}_1$ com a perturbação foram superiores em todos os casos, enquanto que para $\hat{\varphi}_2$ apresentaram redução nos casos em que $v = 3$, $v = 7$ e $v = 10$ e para $\hat{\varphi}_3$ na maioria dos casos apresentaram redução.

Tabela 2: Parâmetros espaciais estimados por meio do algoritmo EM assumindo-se distribuição t -student para os dados simulados sem e com perturbação

Grau de liberdade	Sem perturbação				Elemento Perturbado	Com perturbação			
	$\hat{\beta}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$		$\hat{\beta}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$
$v = 3$	50,62	0,38	0,61	2,49	48	51,06	0,54	0,44	2,50
$v = 5$	50,80	0,10	0,82	2,48	46	50,86	0,50	0,07	2,50
$v = 7$	50,81	0,01	1,01	2,49	122	50,63	0,43	0,06	2,50
$v = 10$	50,55	0,02	1,07	2,50	21	50,52	0,39	0,06	2,50
$v = 100$	50,44	0,10	0,74	2,49	94	50,51	0,54	0,07	2,50
$v = 1000$	49,80	0,09	0,67	2,49	133	51,67	0,53	0,45	2,50

Tabela 3: Parâmetros espaciais estimados por meio do algoritmo de NR assumindo-se distribuição t -student para os dados simulados sem e com perturbação

Grau de liberdade	Sem perturbação				Elemento Perturbado	Com perturbação			
	$\hat{\beta}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$		$\hat{\beta}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$
$v = 3$	50,61	0,06	0,98	2,49	48	51,08	0,14	0,97	2,49
$v = 5$	50,80	0,00	0,97	2,50	46	50,88	0,08	0,98	2,49
$v = 7$	50,80	0,00	0,99	2,50	122	50,58	0,29	0,83	2,44
$v = 10$	50,56	0,00	0,99	2,50	21	50,44	0,30	0,82	2,43
$v = 100$	50,81	0,05	0,43	2,58	94	50,37	0,31	0,54	2,43
$v = 1000$	49,80	0,00	0,97	2,50	133	50,12	0,09	1,01	2,54

As Figuras 10 a 19 apresentam a análise gráfica da influência local nos dados simulados com $v = 3, 5, 7, 10, 100$ e 1000 .

Nas Figuras 10 e 11 avaliou-se a influência no afastamento da verossimilhança pelos gráficos C_i e C_{Qi} , segundo as duas técnicas apresentadas neste trabalho. Os gráficos C_i identificaram os valores perturbados, enquanto que os gráficos C_{Qi} não identificou o valor perturbado apenas para $v = 5$.

Ainda avaliando a influência no afastamento da verossimilhança, tem-se os gráficos $|l_{max}|$ e $|l_{Qmax}|$, apresentados nas Figuras 12 e 13. Nos gráficos $|l_{max}|$ os valores perturbados foram identificados nas Figuras 12(a), 12(c) e 13(e), porém nos gráficos $|l_{Qmax}|$ a perturbação foi identificada somente nas Figuras 12(b) e 13(f).

A influência da perturbação na matriz de covariância foi avaliada pelos gráficos l_{Ch} e l_{QCh} apresentados nas Figuras 14 e 15. Mediante a análise desses gráficos é possível observar que tanto os gráficos l_{Ch} quanto os gráficos l_{QCh} identificaram poucos valores perturbados. A influência no preditor linear, Figuras 16 e 17, pela técnica usual identificou os valores perturbados para $v = 3$, $v = 10$ e $v = 100$, enquanto que para a técnica do Q-afastamento foram

identificadas as perturbações apenas para os dados simulados com grau de liberdade igual a $v = 3$ e $v = 1000$. A influência nos valores preditos, Figuras 18 e 19, nenhum elemento perturbado foi identificado pela técnica usual, enquanto que pela técnica do Q-afastamento foi identificada a perturbação apenas para os dados simulados com grau de liberdade igual a 1000 ($v = 1000$).

Nos casos em que os valores perturbados foram identificados, observa-se que nem sempre esse valor é o mais influente. Verifica-se a presença de outros valores e em nenhum dos casos estes outros valores correspondem aos pontos discrepantes identificados na Tabela 1.

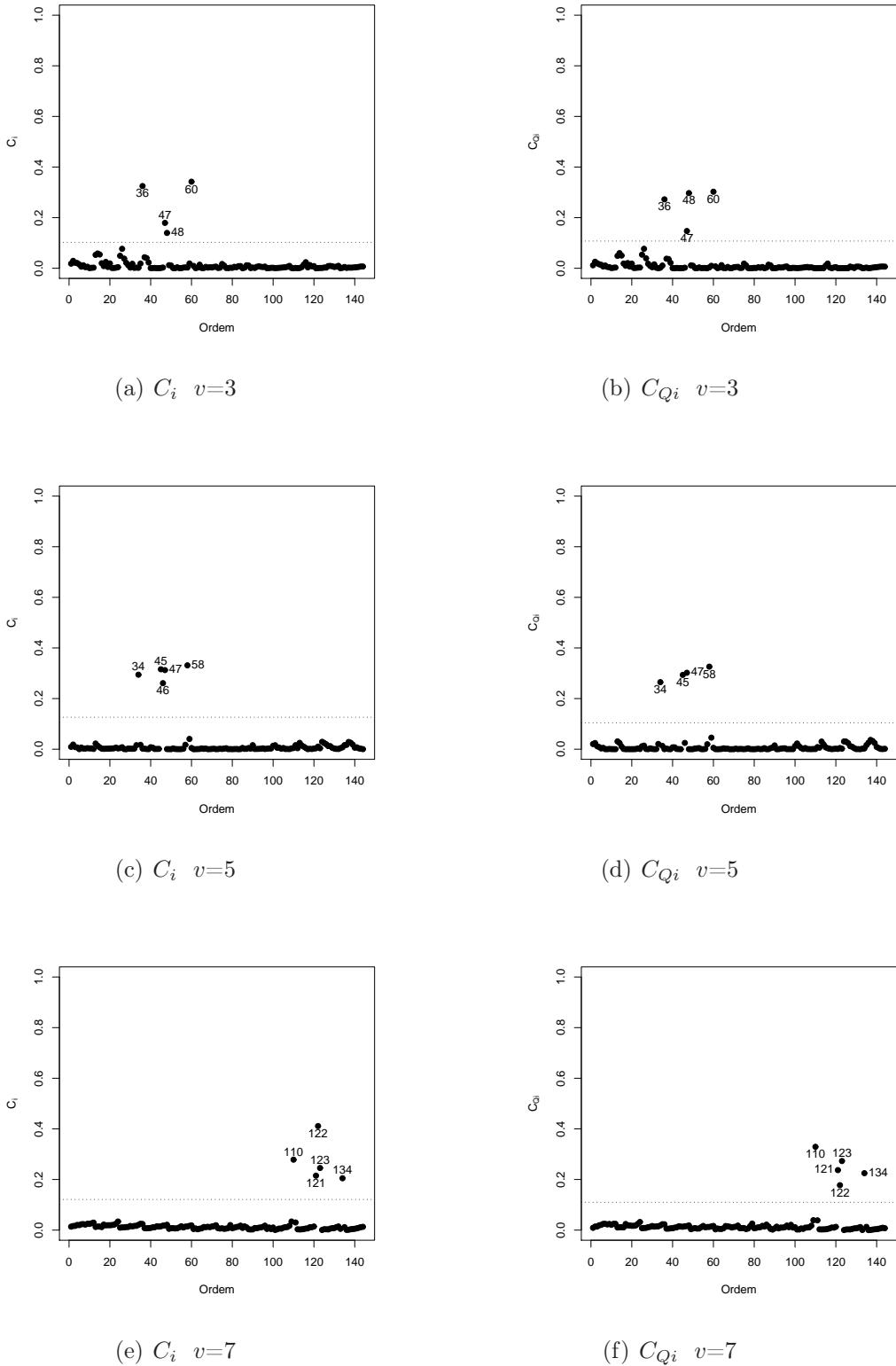


Figura 10: Gráficos de diagnóstico C_i e C_{Qi} para dados simulados com grau de liberdade $v = 3$, $v = 5$ e $v = 7$.

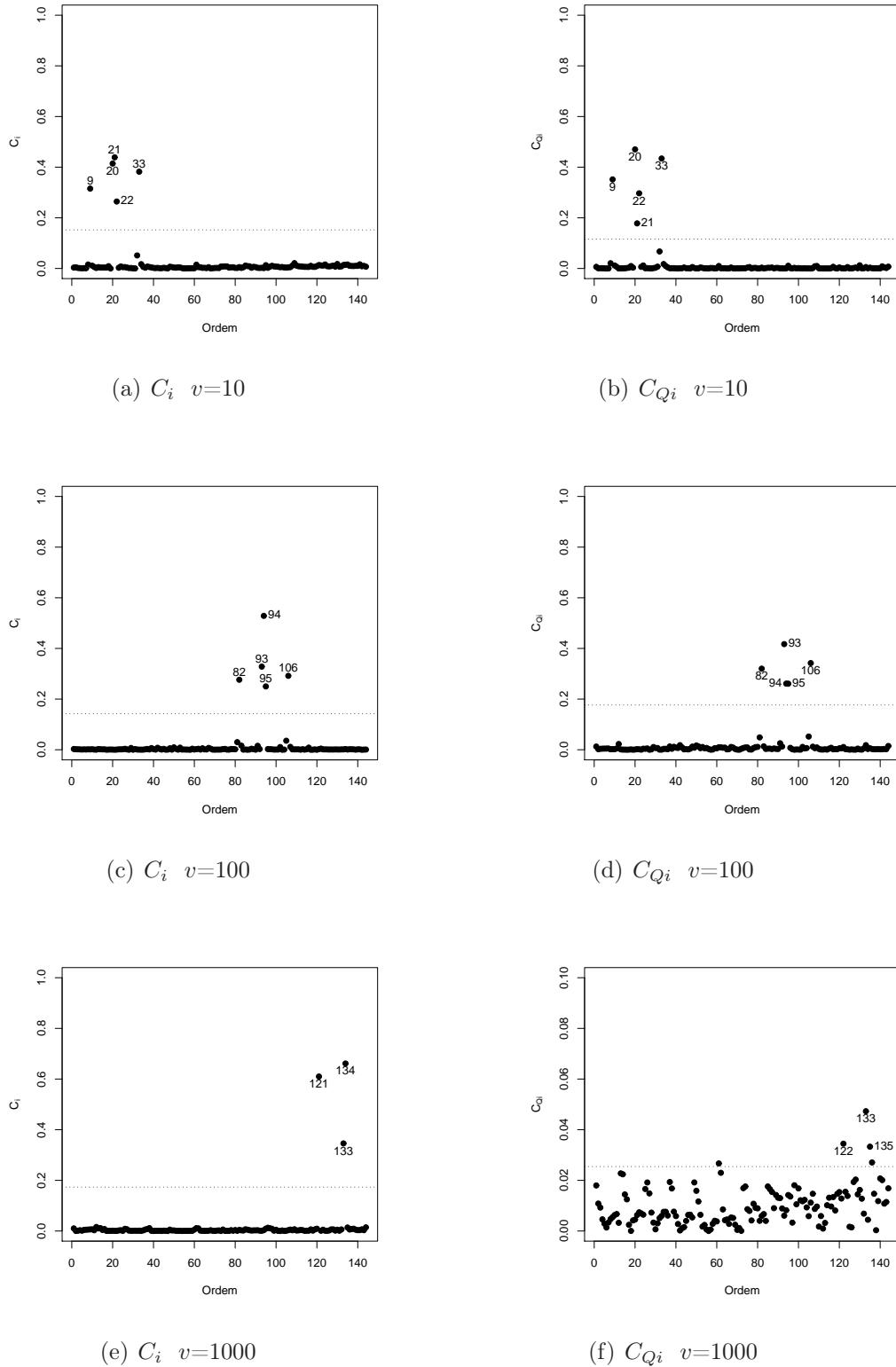


Figura 11: Gráficos de diagnóstico C_i e C_{Qi} para dados simulados com grau de liberdade $v = 10$, $v = 100$ e $v = 1000$.

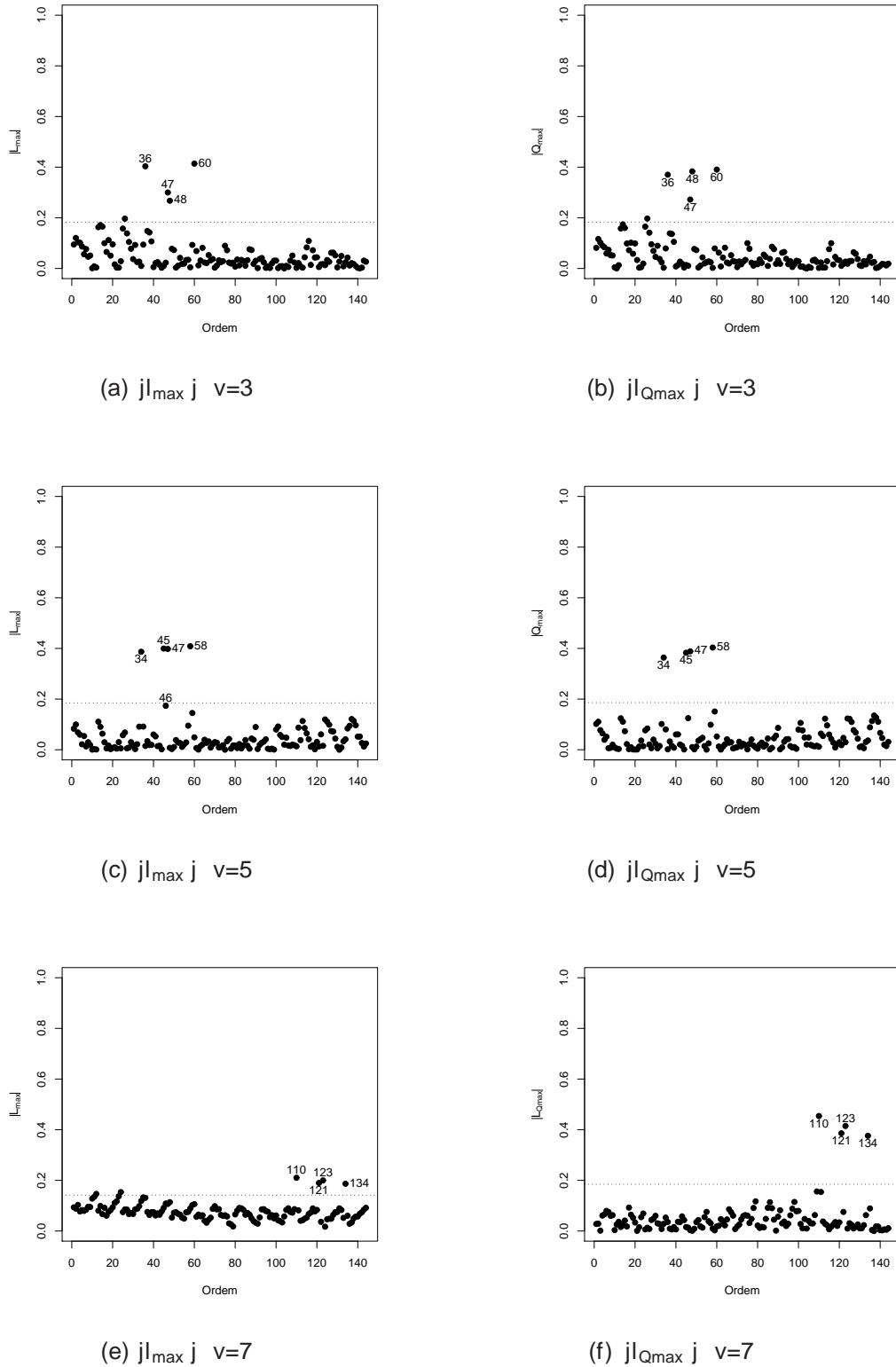


Figura 12: Gráficos de diagnóstico $\|l_{\text{mod}}\|$ e $|Q_{\text{mod}}|$ para dados simulados com grau de liberdade $v = 3$, $v = 5$ e $v = 7$.

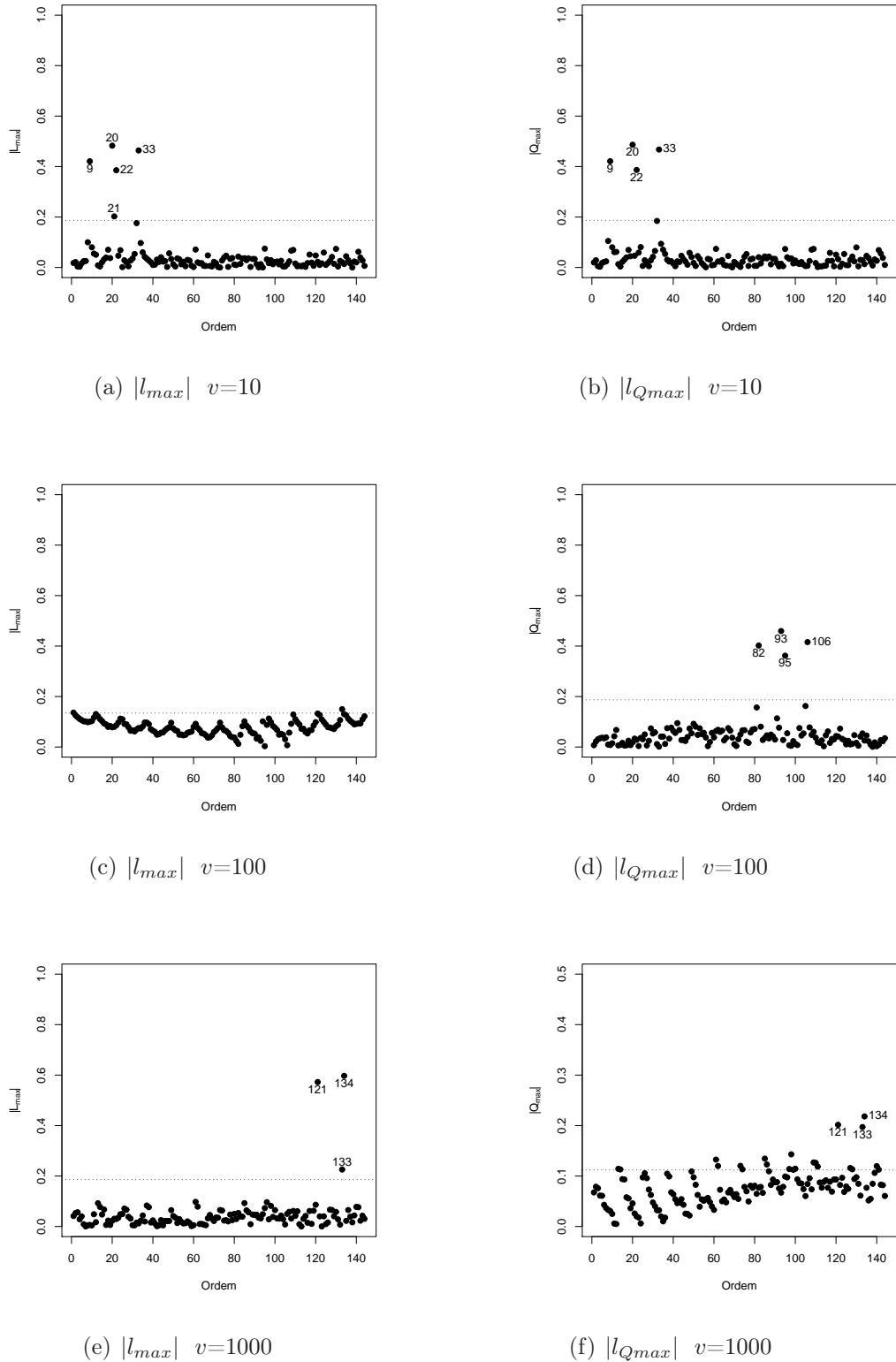


Figura 13: Gráficos de diagnóstico $|l_{max}|$ e $|l_{Qmax}|$ para dados simulados com grau de liberdade $v = 10$, $v = 100$ e $v = 1000$.

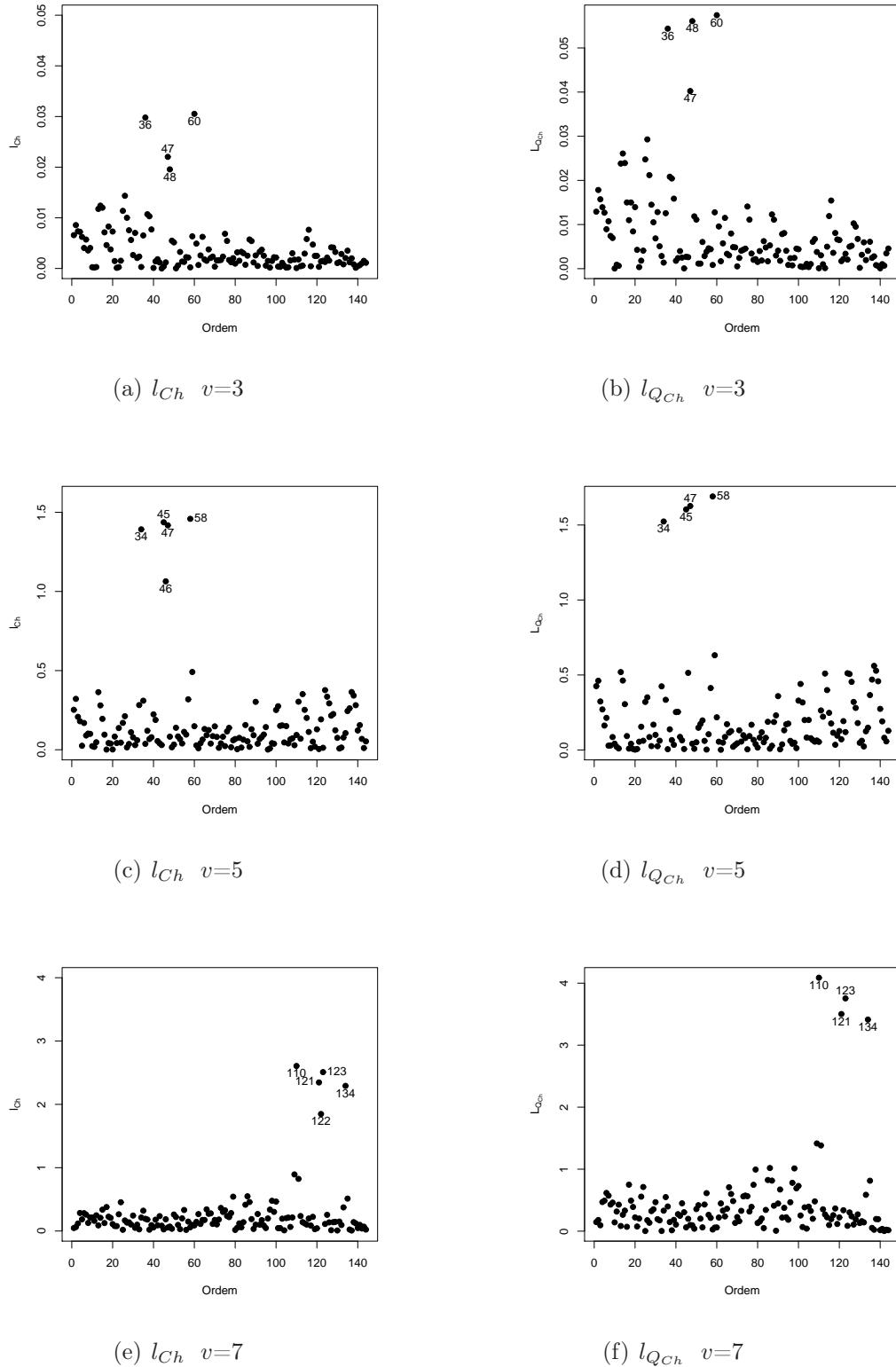


Figura 14: Gráficos de diagnóstico l_{Ch} e l_{QCh} para dados simulados com grau de liberdade $v = 3$, $v = 5$ e $v = 7$.

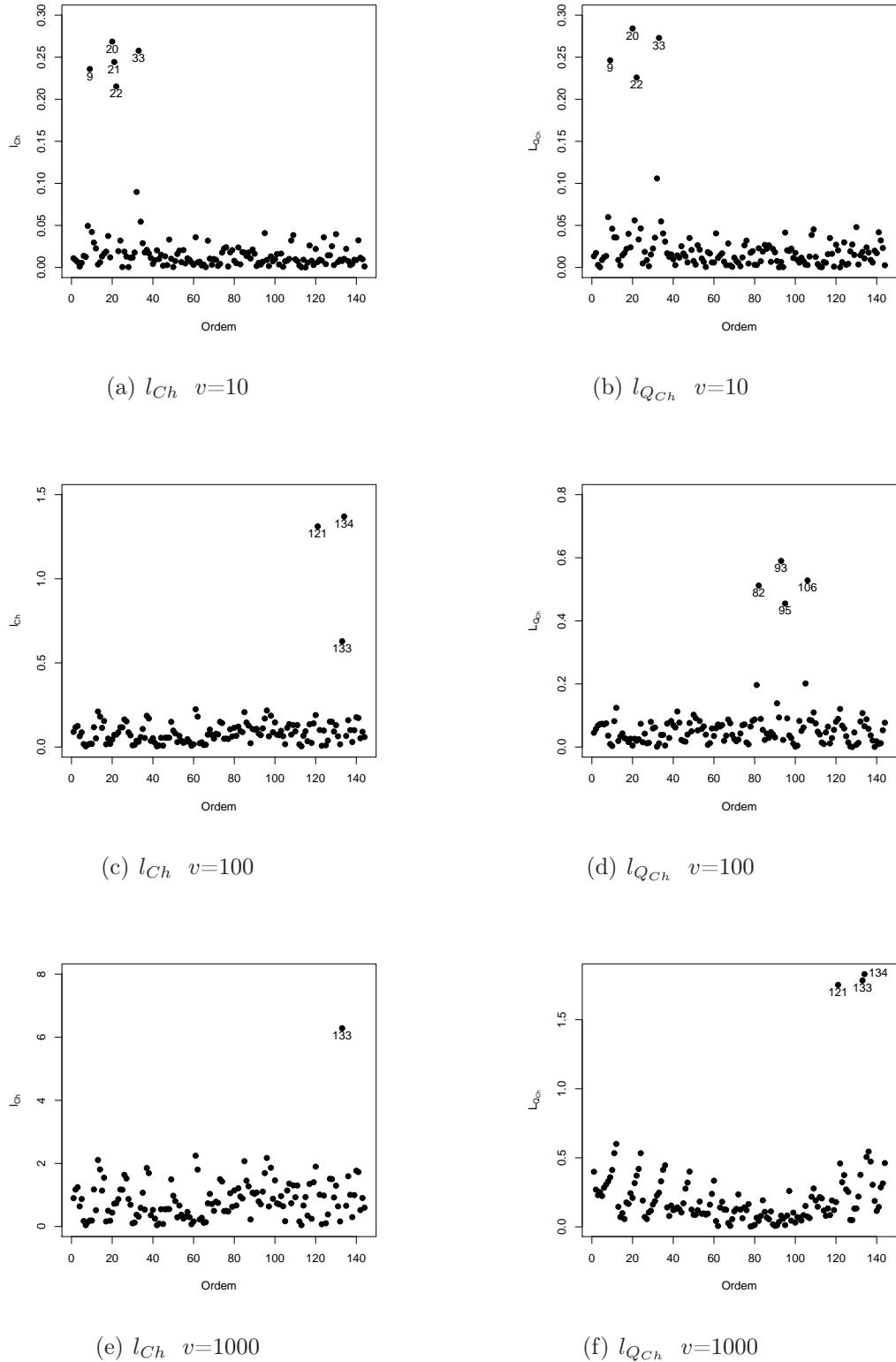


Figura 15: Gráficos de diagnóstico l_{Ch} e l_{QCh} para dados simulados com grau de liberdade $v = 10$, $v = 100$ e $v = 1000$.

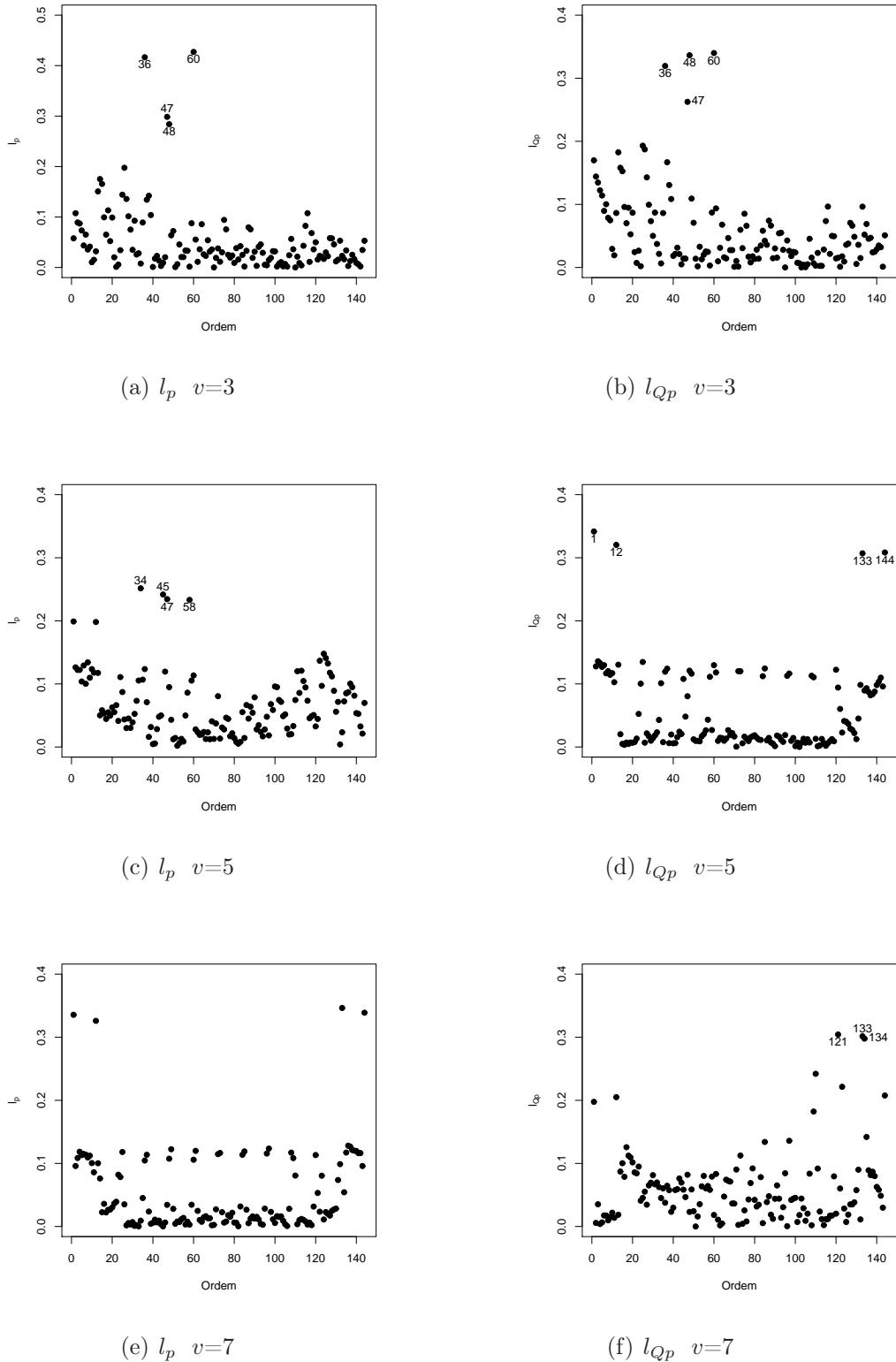


Figura 16: Gráficos de diagnóstico l_p e l_{Qp} para dados simulados com grau de liberdade $v = 3$ $v = 5$ e $v = 7$.

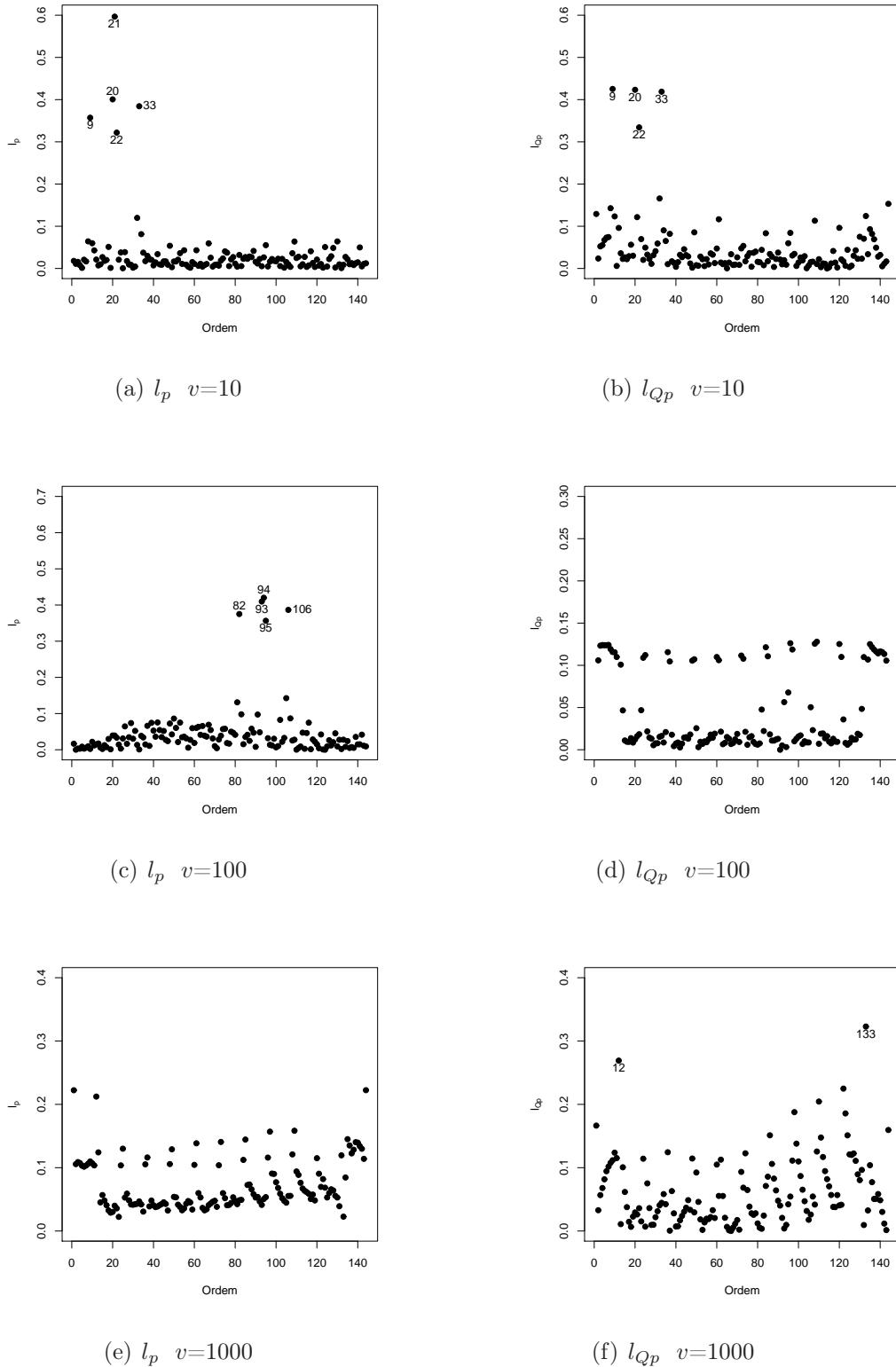


Figura 17: Gráficos de diagnóstico l_p e l_{Qp} para dados simulados com grau de liberdade $v = 10$, $v = 100$ e $v = 1000$.

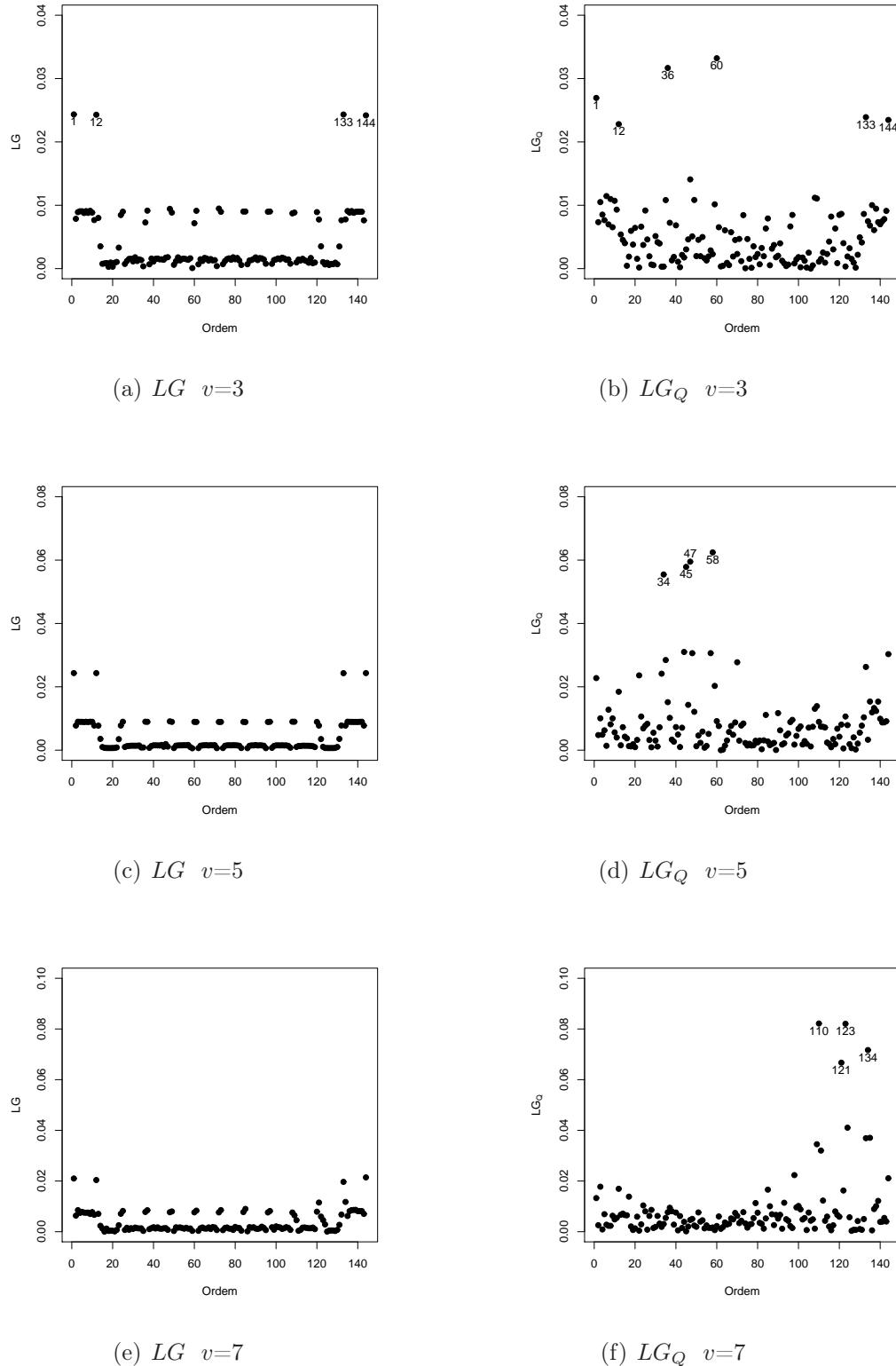


Figura 18: Gráficos de diagnóstico LG e LG_Q para dados simulados com grau de liberdade $v = 3$, $v = 5$ e $v = 7$.

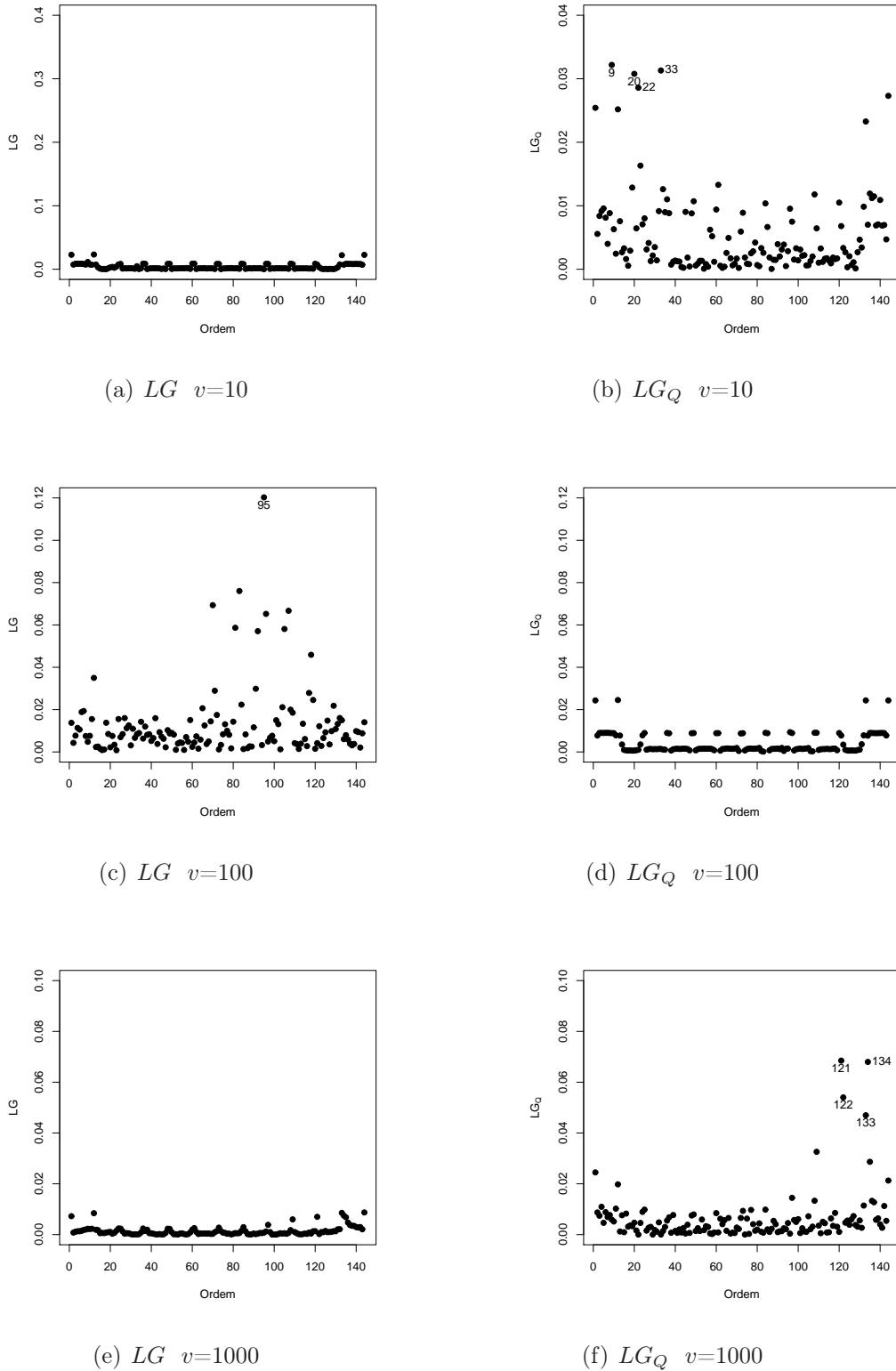


Figura 19: Gráficos de diagnóstico LG e LG_Q para dados simulados com grau de liberdade $v = 10$, $v = 100$ e $v = 1000$.

4.2 Estudo do nutriente potássio

Por meio da estatística descritiva dos dados da variável potássio, verificou-se que os valores de média ($0,27 \text{ cmol dm}^{-3}$) e mediana ($0,26 \text{ cmol dm}^{-3}$) são próximos, com desvio padrão de $0,08 \text{ cmol dm}^{-3}$ e coeficiente de variação igual a 29%, o que os torna heterogêneos os dados. De acordo com a classificação da Coamo/Coodetec (1998), 75% da área em estudo apresenta-se com médio ou alto teor de potássio e os 25% restantes com muito alto teor de potássio.

No gráfico *Boxplot* apresentado na Figura 20(a), observa-se que existem dois pontos discrepantes, sendo eles os elementos de ordem 6 ($0,51 \text{ cmol dm}^{-3}$) e 30 ($0,48 \text{ cmol dm}^{-3}$) localizados conforme apresenta o croqui da área experimental na Figura 20(b). De acordo com a classificação já citada, estes dois elementos apresentam muito alto teor de potássio.

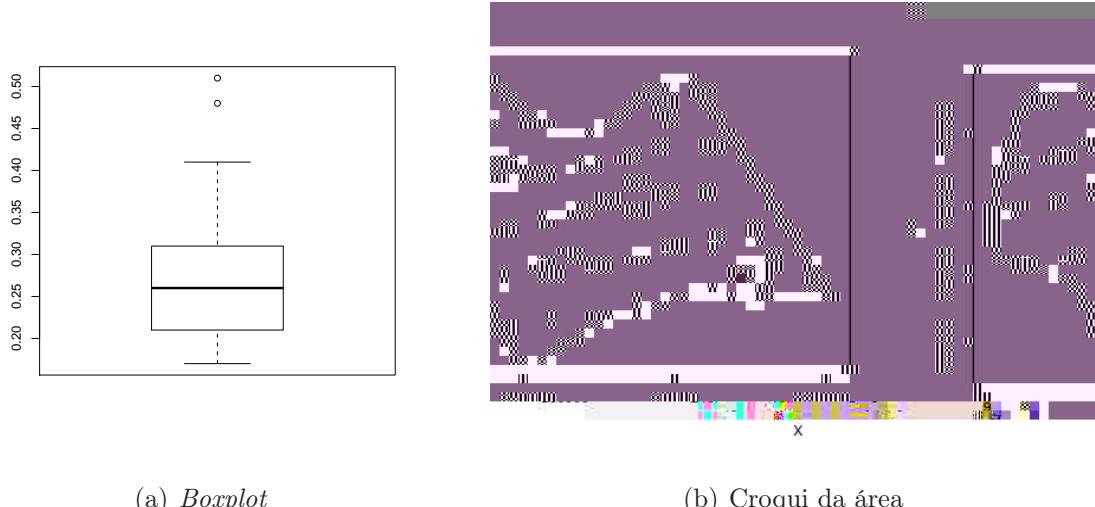


Figura 20: *Boxplot* e croqui da área experimental.

4.2.1 Análise espacial da influência local utilizando a técnica do *Q*-afastamento

A Tabela 4 apresenta as estimativas dos parâmetros espaciais (considerando $\mu(s) = 1\beta$). Para a estimativa desses parâmetros, por meio do algoritmo EM com a distribuição *t*-student, utilizou-se grau de liberdade igual a 5 ($v = 5$), o qual foi definido pelo teste de Liliford.

Com respeito ao coeficiente linear estimado(β), o modelo exponencial foi superior ao resultado apresentado pelo modelo gaussiano, enquanto que para $\hat{\varphi}_2$, a maior estimativa foi para o modelo Matérn (com $k=1,5$).

Entre parênteses encontram-se os erros padrões das estimativas. O modelo Matérn com $k=1,5$ obteve os menores erros padrões para os parâmetros β e $\hat{\varphi}_1$. Para o parâmetro $\hat{\varphi}_2$ o menor erro padrão ficou com o modelo exponencial e para o parâmetro $\hat{\varphi}_3$ o modelo gaussiano.

Tabela 4: Parâmetros espaciais estimados por meio do algoritmo EM assumindo-se distribuição *t*-student com grau de liberdade igual a $v=5$ para a variável potássio

Modelo	$\hat{\beta}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$
Exponencial	0,274 (0,020)	0,0023 (0,009)	0,0007 (0,001)	121,61 (1,050)
Gaussiano	0,282 (0,014)	0,0020 (0,002)	0,0019 (0,002)	200,00 (0,022)
Matérn ($k=1,5$)	0,286 (0,013)	0,0023 (0,001)	0,0069 (0,007)	119,84 (0,102)

Entre parênteses são apresentados os erros padrões de cada parâmetro estimado.

Os resultados dos critérios de seleção dos modelos por validação cruzada (Faraco et al., 2008), valor máximo do logaritmo da função verossimilhança (MVL) e Critério de Akaike -AIC (Akaike, 2002) para o nutriente potássio, utilizando os modelos teóricos: exponencial (Exp), gaussiano (Gaus) e Matérn com parâmetro kappa igual a 1,5, apresentados na Tabela 5. Considerando esses critérios verifica-se que o melhor modelo ajustado é o exponencial.

Tabela 5: Resultados da validação cruzada para a variável potássio

Modelo	<i>EM</i>	<i>EMR</i>	<i>S</i>	<i>S_{EMR}</i>	<i>EA</i>	<i>MVL</i>	<i>AIC</i>
Exponencial	-0,0011	-0,0406	0,0825	1,57	0,062	-172,07	350,13
Gaussiano	-0,0015	-0,0324	0,1282	1,74	0,095	-284,62	575,24
Matérn	-0,0038	-0,0557	0,0970	1,55	0,073	-265,28	536,57

EM: erro médio; *EMR*: erro médio reduzido; *S*: desvio padrão dos erros; *S_{EMR}*: desvio padrão do erro médio; *EA*: erro absoluto; *MVL*: máximo valor do logaritmo da função verossimilhança e *AIC*.

Os gráficos de diagnóstico de influência local foram construídos com o intuito de avaliar se os pontos discrepantes 6 e 30, ou algum(ns) outro(s), 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.

Esses gráficos foram construídos pela técnica do Q-afastamento da verossimilhança completa e denominados por: gráfico dos coeficientes de influências C_{Qi} versus i (ordem das observações coletadas) e gráfico dos auto-vetores $|l_{Qmax}|$ versus i; gráficos l_{QCh} versus i; gráficos l_{Qp} versus i e gráfico de alavanca generalizada LG_Q versus i.

Nos gráficos de diagnósticos C_{Qi} e $|l_{Qmax}|$ da Figura 21, seguiu-se a proposta da curvatura normal conformal de Poon e Poon (1999), apresentada na seção 3.5.1. A linha segmentada corresponde à proposta de Zhu e Lee (2001), apresentada na Equação (3.10). Pode-se observar a partir desses gráficos que a observação 6 aparece como influente no afastamento da verossimilhança.

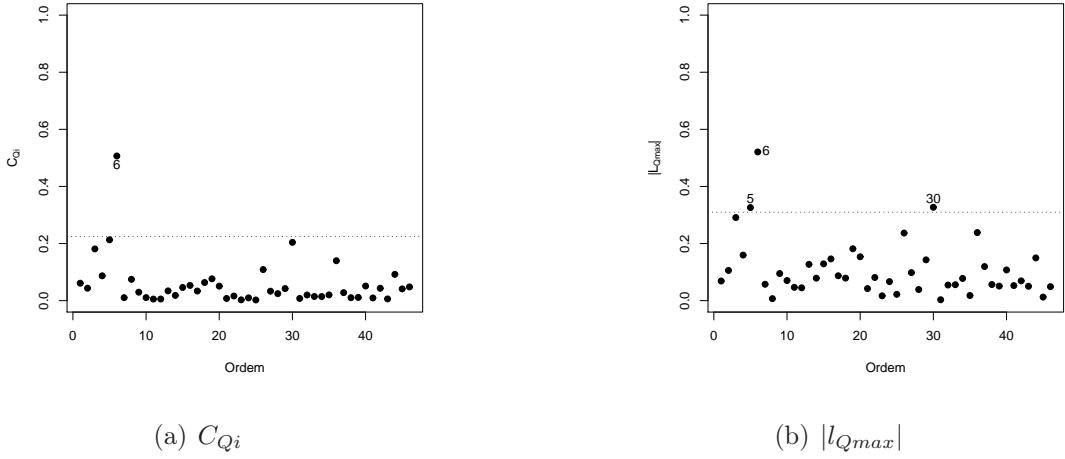


Figura 21: Gráficos de diagnóstico de influência local C_{Qi} e $|l_{Qmax}|$ construídos com os dados da variável potássio.

A avaliação da influência na matriz de covariância, gráfico $l_{Q_{Ch}}$, e no preditor linear, gráfico l_{Qp} , são apresentados na Figura 22. O elemento 6 apresenta-se como influente na matriz de covariância. Porém, o gráfico que avalia a influência local no preditor linear identificou o elemento 8 como influente.

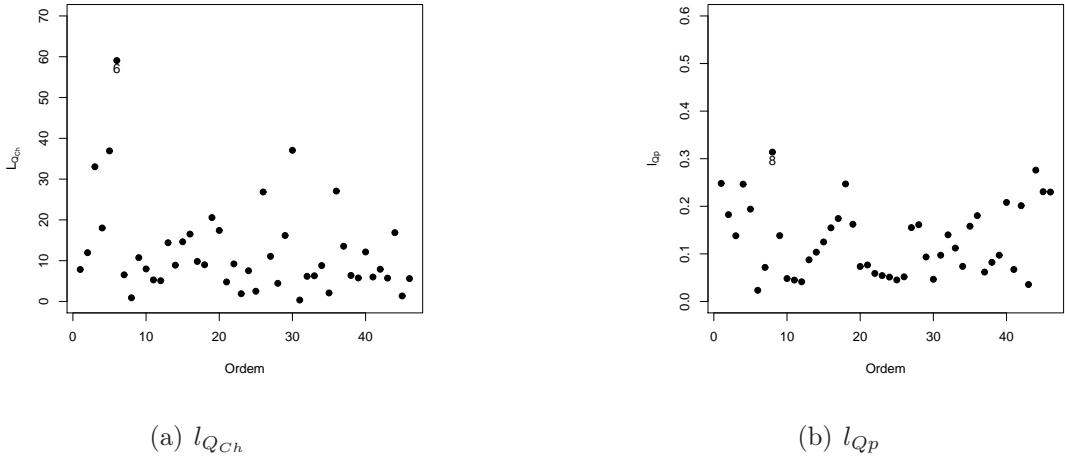


Figura 22: Gráficos de diagnóstico de influência local $l_{Q_{Ch}}$ e l_{Qp} construídos com os dados da variável potássio.

O gráfico LG_Q apresentou as observações 5, 8 e 17 como influentes nos valores preditos.

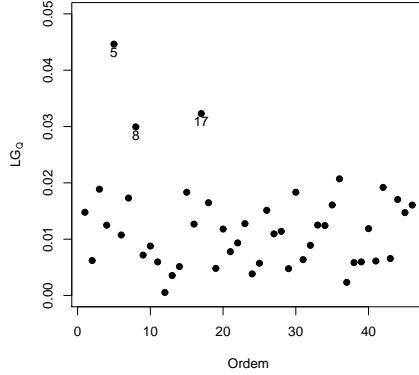
(a) LG_Q

Figura 23: Gráficos de diagnóstico de influência local LG_Q construídos com os dados da variável potássio.

Como a maior parte dos gráficos de diagnóstico identificaram o elemento 6 ($0,51 \text{ cmol dm}^{-3}$) como influente, decidiu-se pela eliminação deste e denominou-se o novo conjunto de dados por K-6.

A eliminação deste elemento (6) provocou pequenas alterações na análise descritiva, como é possível observar na Tabela 6.

Tabela 6: Análise estatística descritiva para a variável potássio [cmol dm^{-3}], com dados totais coletados (K) e retirando-se a observação 6 (K-6)

Variável	Média	Mínimo	Máximo	Q_1	Mediana	Q_3	DP	CV
Potássio (K)	0,27	0,17	0,51	0,21	0,26	0,31	0,08	29%
Potássio (K-6)	0,27	0,17	0,48	0,21	0,26	0,30	0,07	27%

Q_1 : primeiro quartil; Q_3 : terceiro quartil; DP: desvio padrão; CV: coeficiente de variação.

A Tabela 7 apresenta as estimativas do vetor de parâmetros $\hat{\theta}$ antes e após a eliminação do valor influente. Para a estimativa dos parâmetros espaciais dos modelos exponencial, gaussiano e Matérn ($\kappa=1,5$), utilizou-se novamente o grau de liberdade igual a 5 ($v = 5$).

A estimativa do coeficiente linear $\hat{\beta}$ obtida pelo modelo exponencial não apresentou diferença com relação à estimativa anterior e sofreu pequenas alterações com os outros dois modelos. As estimativas do parâmetro $\hat{\varphi}_1$ apresentaram aumento para os três modelos. O parâmetro $\hat{\varphi}_2$ apresentou aumento com os modelos exponencial e gaussiano e redução com o modelo Matérn ($\kappa=1,5$). Para $\hat{\varphi}_3$, houve redução com os modelos exponencial e gaussiano e aumento com o modelo Matérn ($\kappa=1,5$).

Tabela 7: Parâmetros espaciais estimados por meio do algoritmo EM antes e após a eliminação do valor influente da variável potássio

Variável	Modelo	$\hat{\beta}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$
K	Exponencial	0,274	0,0023	0,0007	121,61
		(0,020)	(0,009)	(0,001)	(1,050)
	Gaussiano	0,282	0,0020	0,0019	200,00
		(0,014)	(0,002)	(0,002)	(0,022)
K-6	Matérn (k=1,5)	0,286	0,0023	0,0069	119,84
		(0,013)	(0,001)	(0,007)	(0,102)
	Exponencial	0,274	0,0040	0,0017	119,61
		(0,028)	(0,005)	(0,002)	(0,940)
K-6	Gaussiano	0,272	0,0040	0,0020	120,00
		(0,022)	(0,005)	(0,002)	(1,009)
	Matérn (k=1,5)	0,288	0,0042	0,0021	123,85
		(0,009)	(0,001)	(0,004)	(0,937)

Entre parênteses são apresentados os erros padrões de cada parâmetro estimado.

Considerando os mesmos critérios utilizados anteriormente para a avaliação da validação cruzada, apresentada na Tabela 8, o modelo exponencial foi novamente o que melhor se ajustou aos dados.

Tabela 8: Resultados da validação cruzada para a variável potássio após a eliminação do valor influente (K-6)

Modelo	<i>EM</i>	<i>EMR</i>	<i>S</i>	<i>S_{EMR}</i>	<i>EA</i>	<i>MVL</i>	<i>AIC</i>
Exponencial	-0,0003	-0,0218	0,0737	1,55	0,056	-166,38	338,76
Gaussiano	-0,0007	-0,0222	0,0699	1,55	0,054	-169,87	345,32
Matérn (k=1,5)	-0,0021	-0,0279	0,0801	1,90	0,060	-213,17	432,35

EM: erro médio; *EMR*: erro médio reduzido; *S*: desvio padrão dos erros; *S_{EMR}*: desvio padrão do erro médio e *EA*: erro absoluto; *MVL*: máximo valor do logaritmo da função verossimilhança; *AIC*.

Os gráficos de contorno apresentados na Figura 24, foram confeccionados por meio da interpolação por krigagem ordinária pelo modelo exponencial, antes e após a eliminação do valor influente, ambos considerando o grau de liberdade igual a 5 ($v = 5$).

Pelas Figuras 24(a) e 24(b), pode-se observar que após a eliminação do valor influente houve aumento da área da 1^a classe 0,22-0,24 $cmol\ dm^{-3}$ e redução da área da 4^a classe 0,30-0,33 $cmol\ dm^{-3}$. Isso se justifica pelo fato que o elemento influente (6) é o valor de máximo do conjunto de dados. Consequentemente, estaria influenciando no aumento dos valores preditos obtidos pela krigagem, aumentando a área com muito alto teor de potássio.

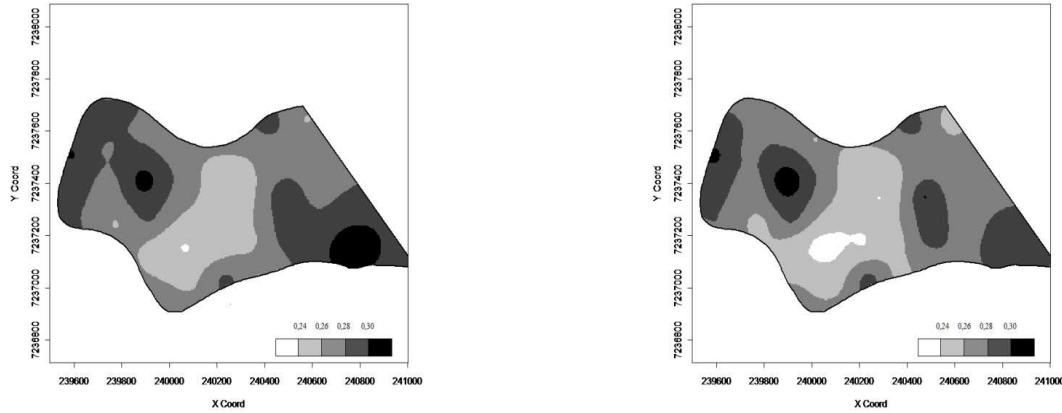


Figura 24: Mapas de contorno para a variável potássio com e sem as observações influentes utilizando o algoritmo EM na estimativa dos parâmetros.

Na Tabela 9 apresenta-se um estudo de comparação dos mapas temáticos por meio da matriz de erros entre suas classes conforme apresenta-se na seção 3.7. Considerou-se o mapa K como referência por possuir todos os dados e o mapa K-6 como modelo. Analisando essa tabela, é possível verificar que houve pequena concordância no número de áreas para a classe 0,22|–0,24 $cmol\ dm^{-3}$. A classe 0,26|–0,28 $cmol\ dm^{-3}$ apresentou a maior concordância, sendo esta de 9514 *pixels*, o que corresponde a 33,4% da área total, ou seja, aproximadamente 27 ha.

Considerando-se a classificação da Coamo/Coodetec (1998) e a Tabela 9, as quatro primeiras classes, juntas, possuem alto teor de potássio (0,22| – – 0,30) e, a última classe (0,30| – – 0,33), muito alto teor de potássio. Somando-se o número de *pixels* dessas quatro primeiras classes na vertical, verifica-se que o mapa K apresenta 27132 *pixels* com alto teor de potássio, o que corresponde a 68 ha da área total e 1357 *pixels* com muito alto teor de potássio correspondendo a 3 ha. Somando-se as quatro primeiras classes na horizontal observa-se que o mapa K-6 apresenta 28049 *pixels* com alto teor e 449 com muito alto teor, o que corresponde a 70 ha e 1 ha, respectivamente. Portanto, houve redução da área com muito alto teor de potássio.

Tabela 9: Matriz de erros entre os mapas temáticos K e K-6 da variável potássio [$cmol\ dm^{-3}$] em número de *pixels*

		Mapa K					
Classes		0,22 −0,24	0,24 −0,26	0,26 −0,28	0,28 −0,30	0,30 −0,33	Total
Mapa K-6	0,22 −0,24	32	627	0	0	0	659
	0,24 −0,26	627	5171	23187	0	0	28358
	0,26 −0,28	0	0	9514	3406	142	13062
	0,28 −0,30	0	0	60	5779	991	6830
	0,30 −0,33	0	0	0	225	224	449
Total		32	5798	32761	9410	1357	49358

Na Tabela 10 são apresentadas as medidas para a comparação dos mapas. Segundo Anderson *et al.* (1976), o índice de exatidão global obtido ($EG = 0,42$) não é aceitável. O índice *Kappa*, de acordo com a classificação de Kripendorff (1980), é de baixa exatidão e por último, o índice *Tau*, que é classificado como o índice *kappa*, também é de baixa exatidão, isto significa que os mapas não são semelhantes.

Tabela 10: Medidas obtidas a partir da matriz de erros entre os mapas temáticos K e K-6 da variável potássio [$cmol\ dm^{-3}$]

Medidas	Potássio
Exatidão Global - <i>EG</i>	0,42
<i>Kappa</i> - <i>K</i>	0,21
<i>Tau</i> - <i>T</i>	0,27

4.2.2 Análise espacial da influência local utilizando o afastamento da verosimilhança usual

A Tabela 11 apresenta as estimativas dos parâmetros espaciais ($\beta, \varphi_1, \varphi_2, \varphi_3$). Para a estimativa desses parâmetros, por meio do algoritmo NR com a distribuição *t*-student, também utilizou-se grau de liberdade igual a 5 ($v = 5$), considerando $\mu(s) = 1\beta$.

O modelo exponencial apresentou a estimativa de β igual a média. Para $\hat{\varphi}_1$, o modelo exponencial apresentou a menor estimativa e, o modelo Mátern ($k=1,5$), a maior. Para $\hat{\varphi}_2$, a maior estimativa foi para o modelo exponencial, seguido por Matérn.

Analizando os erros padrões das estimativas, observa-se que o erro padrão para β foi o mesmo nos três modelos, $\hat{\varphi}_1$ apresentou o menor erro padrão com o modelo exponencial e, para os parâmetros $\hat{\varphi}_2$ e $\hat{\varphi}_3$, o modelo gaussiano apresentou os menores erros padrões.

Tabela 11: Parâmetros espaciais estimados por meio do algoritmo NR assumindo-se distribuição t -student com grau de liberdade $v=5$ para a variável potássio

Modelo	$\hat{\beta}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$
Exponencial	0,270 (0,030)	0,0007 (0,002)	0,0031 (0,008)	100,60 (0,562)
Gaussiano	0,273 (0,030)	0,0038 (0,010)	0,0022 (0,006)	120,00 (0,067)
Matérn (k=1,5)	0,276 (0,030)	0,0015 (0,005)	0,0026 (0,007)	100,00 (67,85)

Entre parênteses são apresentados os erros padrões de cada parâmetro estimado.

Considerando os critérios de seleção dos modelos verifica-se que o melhor modelo ajustado é o gaussiano.

Tabela 12: Resultados da validação cruzada para a variável potássio

Modelo	EM	EMR	S	S_{EMR}	EA	MVL	AIC
Exponencial	-0,0010	-0,0661	0,0812	0,83	0,061	-192,12	390,82
Gaussiano	<u>-0,0009</u>	<u>-0,0134</u>	0,0864	1,64	0,066	<u>-184,21</u>	374,42
Matérn (k=1,5)	-0,0013	-0,0179	0,0939	1,36	0,071	-233,30	472,60

EM : erro médio; EMR : erro médio reduzido; S : desvio padrão dos erros; S_{EMR} : desvio padrão do erro médio; EA : erro absoluto; MVL : máximo valor do logaritmo da função verossimilhança e AIC .

Esse gráficos foram gerados pela técnica usual, utilizando o critério de Cook (1986), e as estimativas obtidas por NR. Com a técnica usual, os gráficos dos coeficientes de influências foram denominados por C_i versus i (ordem das observações coletadas) e gráfico dos auto-vetores $|l_{max}|$ versus i; gráficos l_{Ch} versus i; gráficos l_p versus i e gráfico de alavanca generalizada LG versus i.

Pode-se observar a partir dos gráficos de diagnósticos de influência local C_i e $|l_{max}|$ da Figura 25 que as observações 5 e 6 aparecem influentes no afastamento da verossimilhança.

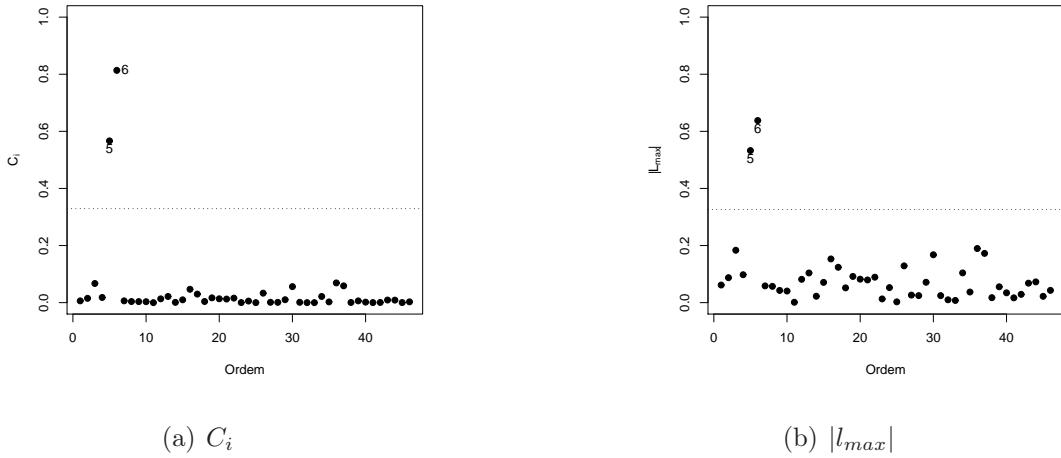


Figura 25: Gráficos de diagnóstico de influência local C_i e $|l_{max}|$ construídos com os dados da variável potássio.

O elemento 6 apresentou-se como influentes na matriz de covariância, Figura 26(a). A avaliação da influência no preditor linear identificou o elemento 6 como influente.

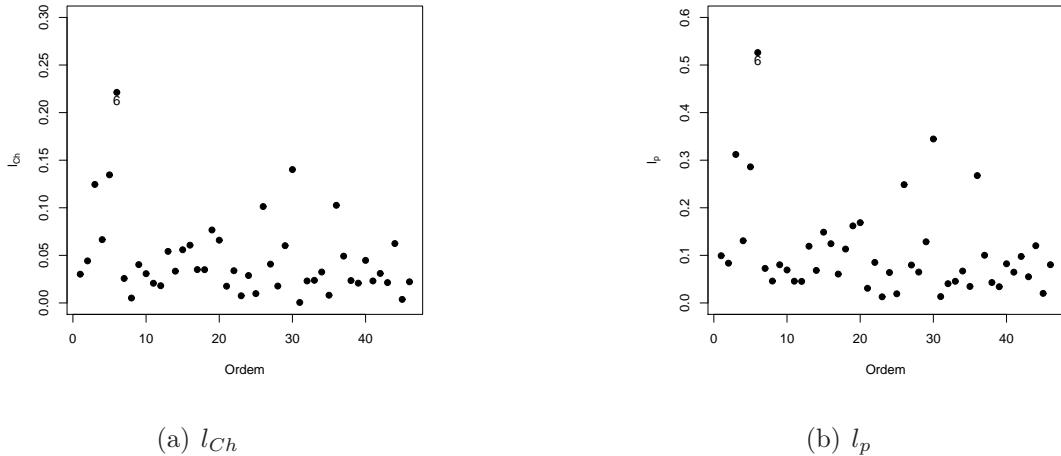


Figura 26: Gráficos de diagnóstico de influência local l_{Ch} e l_p construídos com os dados da variável potássio.

A influência nos valores preditos avaliada pela técnica usual não identificou valores influentes.

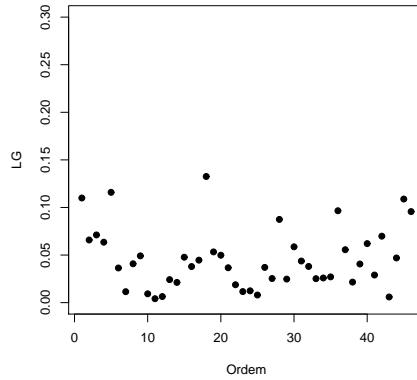
(a) LG

Figura 27: Gráfico de diagnóstico de influência local LG construídos com os dados da variável potássio.

Como a maior parte dos gráficos de diagnóstico identificaram os elementos 5 ($0,29 \text{ cmol dm}^{-3}$) e 6 ($0,51 \text{ cmol dm}^{-3}$) como influentes, decidiu-se pela eliminação destes e denominou-se o novo conjunto de dados por K-5-6.

A eliminação destes elementos provocou pequenas alterações na análise descritiva, como é possível observar na Tabela 13.

Tabela 13: Análise estatística descritiva para a variável potássio [cmol dm^{-3}], com dados totais coletados (K) e retirando-se as observações 5 e 6 (K-5-6)

Variável	Média	Mínimo	Máximo	Q_1	Mediana	Q_3	DP	CV
Potássio (K)	0,27	0,17	0,51	0,21	0,26	0,31	0,08	29%
Potássio (K-5-6)	0,27	0,17	0,48	0,21	0,26	0,30	0,07	27%

Q_1 : primeiro quartil; Q_3 : terceiro quartil; DP: desvio padrão; CV: coeficiente de variação.

A Tabela 14 apresenta as estimativas do vetor de parâmetros $\hat{\theta}$ antes e após a eliminação dos valores influentes. Utilizou-se novamente o grau de liberdade igual a 5 ($v = 5$).

As estimativas do coeficiente linear $\hat{\beta}$ apresentaram pequenas diferenças com relação à estimativa anterior. $\hat{\phi}_1$ apresentou redução para o modelo exponencial e aumento para os outros dois modelos. O parâmetro $\hat{\phi}_2$ apresentou redução para os três modelos.

Tabela 14: Parâmetros espaciais estimados por meio do algoritmo EM antes e após a eliminação do valor influente da variável potássio

Variável	Modelo	$\hat{\beta}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$
K	Exponencial	0,270	0,0007	0,0031	100,60
		(0,030)	(0,002)	(0,008)	(0,562)
	Gaussiano	0,273	0,0038	0,0022	120,00
		(0,030)	(0,010)	(0,006)	(0,067)
K-5-6	Matérn (k=1,5)	0,276	0,0015	0,0026	100,00
		(0,030)	(0,005)	(0,007)	(67,85)
	Exponencial	0,267	0,0001	0,0030	100,29
		(0,021)	(0,002)	(0,004)	(0,628)
K-5-6	Gaussiano	0,272	0,0040	0,0020	120,01
		(0,021)	(0,002)	(0,004)	(0,039)
	Matérn (k=1,5)	0,275	0,0020	0,0020	100,00
		(0,029)	(0,005)	(0,005)	(84,83)

Entre parênteses são apresentados os erros padrões de cada parâmetro estimado.

Considerando os mesmos critérios utilizados anteriormente para a avaliação da validação cruzada apresentada na Tabela 15, o modelo gaussiano foi novamente o que melhor se ajustou aos dados.

Tabela 15: Resultados da validação cruzada para a variável potássio após a eliminação do valor influente (K-5-6)

Modelo	<i>EM</i>	<i>EMR</i>	<i>S</i>	<i>S_{EMR}</i>	<i>EA</i>	<i>MVL</i>	<i>AIC</i>
Exponencial	-0,00010	-0,0110	0,0735	1,97	0,056	-175,228	370,43
Gaussiano	-0,00002	-0,0051	0,0704	1,35	0,054	-163,26	346,52
Matérn (k=1,5)	-0,00090	-0,0201	0,0787	1,18	0,060	-211,48	442,95

EM: erro médio; *EMR*: erro médio reduzido; *S*: desvio padrão dos erros; *S_{EMR}*: desvio padrão do erro médio e *EA*: erro absoluto; *MVL*: máximo valor do logaritmo da função verossimilhança; *AIC*.

Os gráficos de contorno, apresentados na Figura 28, foram confeccionados por meio da interpolação por krigagem ordinária pelo modelo gaussiano antes e após a eliminação dos valores influentes, ambos considerando o grau de liberdade igual a 5 ($v = 5$).

Pelas Figuras 28(a) e 28(b), pode-se observar que, após a eliminação dos valores influentes, houve aumento da área da 1^a classe 0,22-0,24 $cmol\ dm^{-3}$ e redução da área das demais classes.

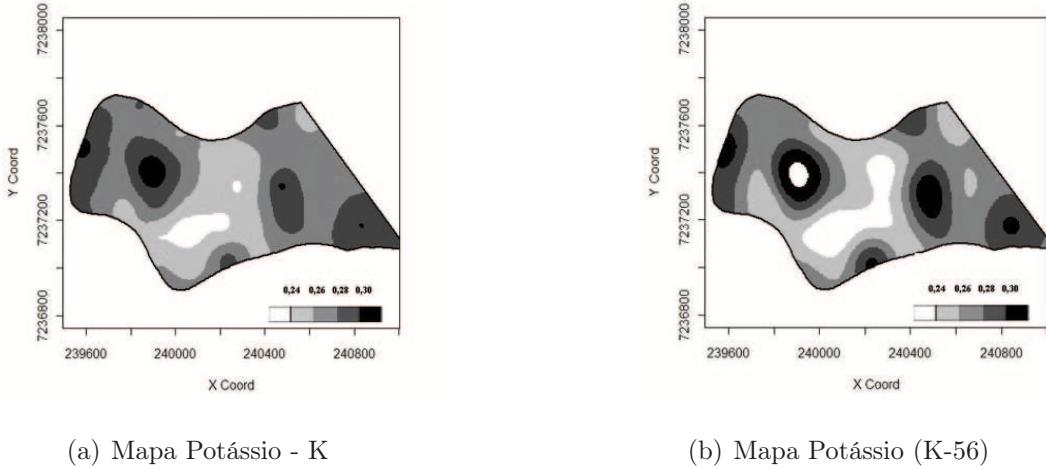


Figura 28: Mapas de contorno para a variável potássio com e sem as observações influentes, utilizando o algoritmo de NR na estimação dos parâmetros.

Analizando a Tabela 16, que apresenta um estudo de comparação dos mapas temáticos por meio da matriz de erros entre suas classes, onde o mapa K é referência (ou real) e o mapa K-5-6 é o modelo, é possível verificar que houve grande concordância no número de áreas em todas as classes.

Somando-se as quatro primeiras classes na vertical, observa-se que o mapa K apresenta 25996 *pixels* com alto teor e 2493 com muito alto teor, o que corresponde a 65 ha e 6 ha, respectivamente. Porém, somando-se as quatro primeiras classes na horizontal, observa-se que o mapa K-5-6 apresenta 25555 *pixels* com alto teor e 2934 com muito alto teor, o que corresponde a 64 ha e 7 ha, respectivamente. Portanto, com a retirada dos elementos 5 e 6 houve um pequeno aumento da área com muito alto teor de potássio.

Tabela 16: Matriz de erros entre os mapas temáticos K e K-5-6 da variável potássio [$cmol\ dm^{-3}$] em número de *pixels*

		Mapa K					Total
Classes	0,22 −0,24	0,24 −0,26	0,26 −0,28	0,28 −0,30	0,30 −0,33		
Mapa K-5-6	0,22 −0,24	2694	414	0	0	0	3108
	0,24 −0,26	0	6770	395	0	0	7165
	0,26 −0,28	0	0	9543	0	0	9543
	0,28 −0,30	0	0	294	5445	0	5739
	0,30 −0,33	0	0	0	441	2493	2934
Total	2694	7184	10232	5886	2493	28489	

Segundo Anderson *et al.* (1976), o índice de exatidão global obtido ($EG = 0,95$) é aceitável. O índice *Kappa*, de acordo com a classificação de Kripendorff (1980), apresenta alta

exatidão e o índice Tau também é de alta exatidão, indicando semelhança nos mapas temáticos.

Tabela 17: Medidas obtidas a partir da matriz de erros entre os mapas temáticos K e K-5-6 da variável potássio [$cmol\ dm^{-3}$]

Medidas	Potássio
Exatidão Global -EG	0,95
Kappa - K	0,93
Tau - T	0,94

4.3 Estudo da produtividade de soja com covariáveis

Na Tabela 18 são apresentadas as estatísticas descritivas para as variáveis: produtividade da soja; resistência do solo à penetração nas camadas de 0 a 0,10; 0,1 a 0,20 e 0,20 a 0,3 m de profundidade ($RSP_{0,0-0,1}$, $RSP_{0,1-0,2}$ e $RSP_{0,2-0,3}$, respectivamente); e densidade do solo nas camadas de 0 a 0,10; 0,10 a 0,20 e 0,20 a 0,30 m de profundidade ($Des_{0-0,1}$, $Des_{0,1-0,2}$ e $Des_{0,2-0,3}$, respectivamente).

A média da produtividade é um pouco superior quando comparada com a mediana, indicando uma pequena assimetria à esquerda, como pode ser verificado pelo coeficiente de assimetria ($As = -0,13$). Apresenta, ainda, curtose positiva igual 0,76. Quanto ao desvio padrão ($DP=0,38\ t\ ha^{-1}$), verifica-se que é pequeno, comparado com o valor da média, que foi de $3,23\ t\ ha^{-1}$; logo, existe, homogeneidade nos dados da produtividade ($CV = 11,79\%$).

É sabido que quando aumenta a resistência à penetração do solo, o sistema radicular apresenta desenvolvimento reduzido, podendo comprometer a produtividade da área. No entanto, na área em estudo, a resistência do solo à penetração apresenta redução da média, do 1º quartil (Q_1), da mediana e do 3º quartil (Q_3), e aumento da assimetria e da curtose conforme aumenta a profundidade. Em todas as camadas observam-se valores superiores a 2,60 Mpa. Na camada de 0 a 0,10 m de profundidade, 50% dos dados encontram-se acima de 2,64 Mpa. Segundo Canarache (1990), valores da RPS no intervalo de $[2,6 - 5,0]$ Mpa limitam o crescimento da raiz, enquanto que no intervalo de $[1,1 - 2,59]$ Mpa apresentam pouca limitação da raiz.

Observa-se que desvio padrão não varia muito entre as três camadas, ficando entre 0,49 Mpa e 0,56 Mpa. A camada de 0,20 a 0,30 m de profundidade apresenta aumento tanto da assimetria ($As = 3,0$) quanto da curtose ($K = 14,64$), quando comparada às camadas anteriores. As três camadas apresentam homogeneidade ($CV < 30\%$), sendo a camada de 0,20 a 0,30 m de profundida menos homogênea que as demais.

Quanto à densidade do solo, todas as medidas estatísticas apresentaram aumento da camada de 0-0,10 para a camada de 0,10-0,20 m de profundidade e redução da camada de 0,10-0,20 para 0,20-0,30 m de profundidade. A média, o desvio padrão e o coeficiente de variação

não variam muito entre as três camadas, tanto que todas apresentam homogeneidade dos dados.

Para Letey (1985), há influência da densidade do solo no desenvolvimento radicular das plantas, pois com o aumento da densidade ocorre diminuição do volume de poros, contribuindo para baixas produtividades da área cultivada. Portanto, a resistência do solo à penetração (RSP), juntamente com a densidade do solo (Des), são atributos físicos que influenciam diretamente o crescimento das raízes e, consequentemente, a parte aérea das plantas (Dexter, 2004). Logo, influenciam a produtividade, justificando o motivo pelo qual foram consideradas como cova-riáveis para essa análise.

Tabela 18: Estatísticas descritivas para a variável Produtividade da soja (Prod)[t ha⁻¹], Resistência do solo a penetração [Mpa] e Densidade do solo [mg m⁻³]

Variável	Prod	RSP (0,0–0,1)	RSP (0,1–0,2)	RSP (0,2–0,3)	Des (0,0–0,1)	Des (0,1–0,2)	Des (0,2–0,3)
n	47	47	47	47	47	47	47
Média	3,23	2,74	2,44	2,08	1,09	1,13	1,09
Mín.	2,09	1,40	1,36	1,34	0,77	0,98	0,95
Máx	4,09	3,69	3,71	4,96	1,22	1,55	1,30
Q_1	3,01	2,39	2,14	1,73	1,05	1,06	1,04
Mediana	3,19	2,64	2,36	2,05	1,10	1,12	1,08
Q_3	3,49	3,14	2,75	2,26	1,16	1,17	1,13
DP	0,38	0,53	0,49	0,56	0,08	0,10	0,07
CV(%)	11,79	19,49	19,97	26,88	7,61	8,84	6,59
As	-0,13	0,05	0,67	3,00	-1,32	1,79	0,62
K	0,76	-0,30	0,62	14,64	3,50	6,04	0,48

n: número de dados; Mín.: valor mínimo; Máx.: valor máximo; Q1: primeiro quartil; Q3: terceiro quartil; DP: desvio-padrão; CV: coeficiente de variação; As: Assimetria e K: curtose.

Analizando o gráfico *Boxplot* da produtividade na Figura 29(a), observa-se a presença de um ponto discrepante, que corresponde ao valor de mínimo do conjunto de dados, sendo igual a 2,09 t ha⁻¹. Esse é o 13º valor da série de dados e está localizado conforme mostra o croqui da área experimental na Figura 29(b).

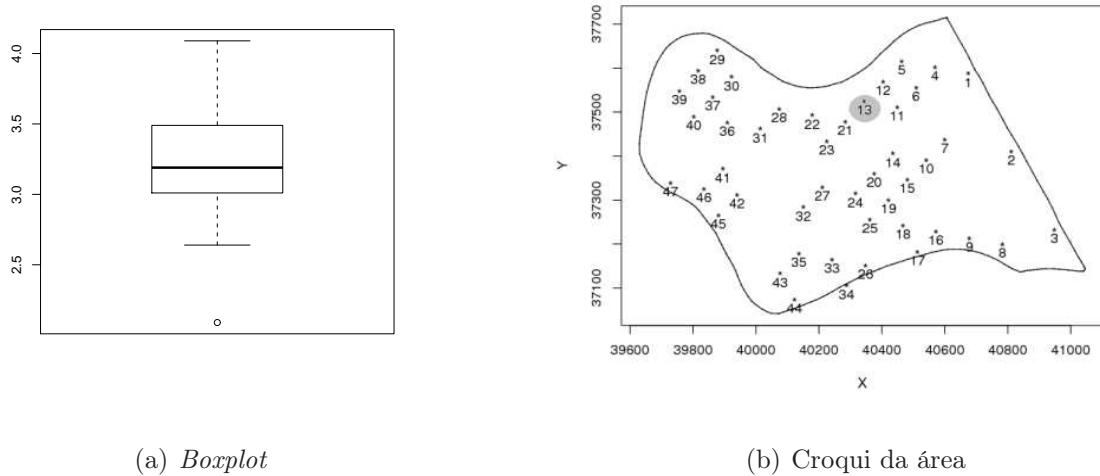


Figura 29: *Boxplot* da produtividade e croqui da área experimental.

Na Figura 30 tem-se o gráfico scaterplot da produtividade. Verifica-se que a observação 13 ($2,09 \text{ t ha}^{-1}$), identificada no *Boxplot* como discrepante, encontra-se em uma região onde os vizinhos mais próximos possuem valores superiores a $3,0 \text{ t ha}^{-1}$.

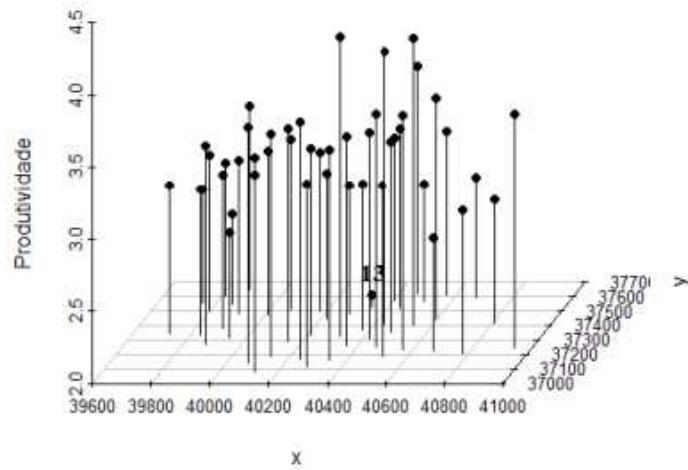


Figura 30: Scaterplot da produtividade.

4.3.1 Análise espacial

Para a estimativa dos parâmetros da modelagem espacial da produtividade da soja como função de covariáveis, considerou-se a definição de $Y(s_i)$ (Equação 2.4), a qual apresenta o termo determinístico $\mu(s_i)$ (Equação 2.6) como sendo a média do processo estocástico obtido o modelo espacial linear da produtividade em função das covariáveis *RSP* e *Des* nas três camadas,

como apresentado na equação abaixo:

$$\begin{aligned}\mu(s_i) = & \beta_1 + \beta_2 \text{RSP}_{0,0-0,1} + \beta_3 \text{RSP}_{0,1-0,2} \\ & + \beta_4 \text{RSP}_{0,2-0,3} + \beta_5 \text{Des}_{0,0-0,1} + \beta_6 \text{Des}_{0,1-0,2} + \beta_7 \text{Des}_{0,2-0,3}\end{aligned}\quad (4.1)$$

em que, β_1, \dots, β_7 são os parâmetros desconhecidos e a serem estimados.

4.3.1.1 Análise espacial da influência local pelo Q -afastamento

Na Tabela 19 são apresentados os resultados das estimativas dos parâmetros do modelo espacial linear, da Equação (4.1), e dos parâmetros da estrutura de variabilidade espacial ($\Sigma = \varphi_1 \mathbf{I}_n + \varphi_2 R(\varphi_3)$) para a produtividade da soja por meio do algoritmo EM. Foram ajustados os modelos teóricos exponencial (Exp), gaussiano (Gaus) e Matérn com $\kappa=0,7$ (Mat), assumindo-se distribuição t-student com grau de liberdade $v=3$.

Para confirmar a hipótese de que a distribuição da população da qual se retirou a determinada amostra segue uma distribuição de probabilidade t-student com grau de liberdade $v = 3$, utilizou-se o teste de Liliford.

Verifica-se na Tabela 19 que os parâmetros espaciais $\hat{\varphi}_1$ (efeito pepita), $\hat{\varphi}_2$ (contribuição) e $\hat{\varphi}_3$ (função do alcance) apresentam pequenas variações entre as estimativas dos três modelos. Entre parênteses são apresentados os desvios padrões das estimativas.

Tabela 19: Parâmetros espaciais estimados por meio do algoritmo EM assumindo-se distribuição t-student com grau de liberdade 3 para a variável produtividade

Modelo	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$
Exp	4,56 (1,07)	-0,10 (0,12)	0,06 (0,12)	-0,01 (0,09)	-0,29 (0,64)	0,06 (0,52)	-0,86 (0,80)	0,08 (0,03)	0,05 (0,04)	138,4 (1,64)
Gaus	4,35 (1,07)	-0,09 (0,12)	0,07 (0,12)	-0,01 (0,09)	-0,27 (0,64)	0,13 (0,53)	-0,81 (0,80)	0,09 (0,03)	0,04 (0,03)	147,1 (0,25)
Mat	4,71 k=0,7 (1,07)	-0,11 (0,11)	0,05 (0,12)	-0,02 (0,09)	-0,29 (0,64)	0,02 (0,52)	-0,90 (0,79)	0,08 (0,08)	0,05 (0,05)	192,9 (220)

Entre parênteses são apresentados os erros padrões de cada parâmetro estimado.

A Tabela 20 apresenta os resultados dos critérios de validação dos modelos por validação cruzada (Faraco et al., 2008), critério de informação de Akaike (Akaike, 2002) e máximo valor do logaritmo da função verossimilhança (MVL) para a variável produtividade, utilizando os modelos teóricos exponencial (Exp), gaussiano (Gaus) e família Matérn com $\kappa=0,7$ (Matérn), pelo algoritmo EM. Considerando esses resultados, verifica-se que o melhor modelo ajustado é o gaussiano.

Tabela 20: Resultados da validação cruzada para os parâmetros estimados para a variável produtividade

Modelo	EM	EMR	S	$S_{E\bar{M}R}$	EA	IAC	MVL
Exp	-0,0004	-0,0005	0,3484	0,6989	12,04	-94,11	45,95
Gaus	-0,0004	-0,0005	0,3451	0,6686	11,98	-95,01	44,26
Matérn k=0,7	-0,0005	-0,0008	0,3503	0,7291	12,01	-93,60	46,70

EM: erro médio; EMR: erro médio reduzido; S : desvio padrão dos erros; $S_{E\bar{M}R}$: desvio padrão do erro médio e EA: erro absoluto; IAC: critério de transformação de Akaike; MVL: máximo valor do logaritmo da função verossimilhança.

Nos gráficos C_{Qi} , Figura 31(a), e $|l_{Q_{max}}|$, Figura 31(b), considerando o limite do gráfico definida na Equação (3.10), observa-se que ambos identificaram o elemento 19 como influente.

A influência na função de covariância foi avaliada pelo gráfico $l_{Q_{Ch}}$, que identificou os pontos 12, 13 e 21 como influentes.

O gráfico l_{Qp} da Figura 32(b) avalia a influência no preditor linear. Segundo esse gráfico os elementos 12, 13, 21 e 23 foram identificados como influentes.

Optou-se pela retirada de duas observações 13 ($2,09 t ha^{-1}$) e 19 ($3,44 t ha^{-1}$) para dar continuidade nas análises. Para diferenciar os conjuntos de dados, considerou-se Prod: dados totais, Prod-13-19: dados com exclusão dos elementos 13 e 19.

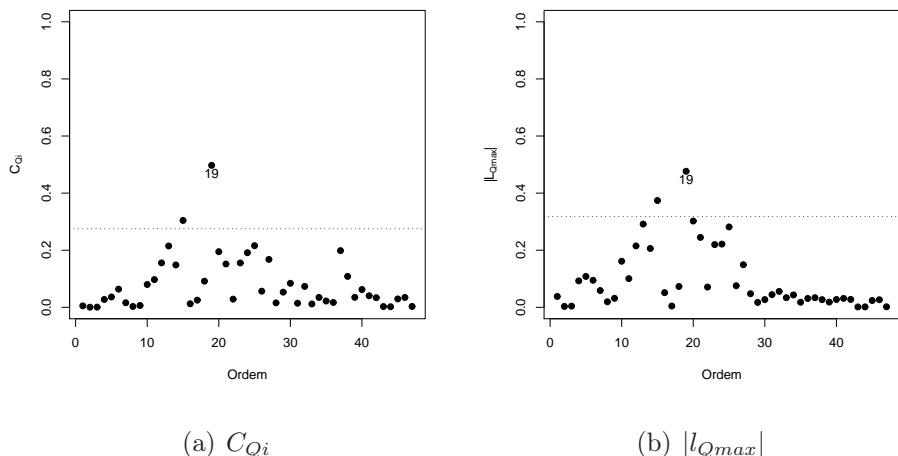


Figura 31: Gráficos de diagnóstico de influência local C_{Qi} e $|l_{Q_{max}}|$ para a produtividade de soja

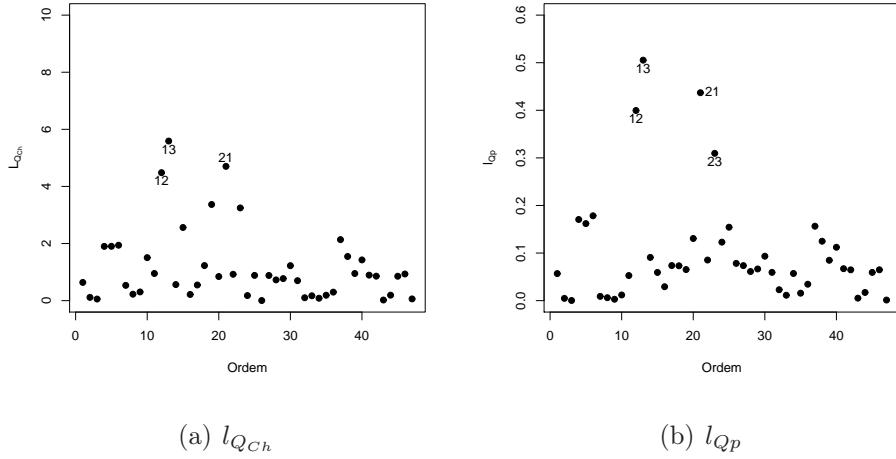


Figura 32: Gráficos de diagnóstico de influência local l_{QCh} e l_{Qp} para a produtividade de soja

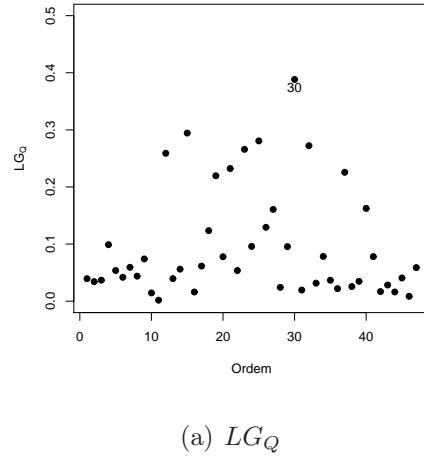


Figura 33: Gráfico de diagnóstico de influência local LG_Q para a produtividade de soja

Com respeito à influência nas medidas descritivas, observa-se na Tabela 30 que, a eliminação dos elementos 13 e 19 alterou a média, o desvio padrão e o coeficiente de variação, que passou de 11,79% para 10,59%. O novo conjunto de dados permanece homogêneo. A retirada desses elementos também alterou a assimetria e a curtose dos dados, havendo aumento da assimetria (passando para positiva) e redução da curtose (passando para negativa).

Tabela 21: Estatísticas descritivas para as variáveis Prod e Prod-13-19

Variável	Prod	Prod-13-19
n	47	45
Média	3,23	3,25
Mín.	2,09	2,64
Máx	4,09	4,09
Q_1	3,01	3,02
Mediana	3,19	3,19
Q_3	3,49	3,49
DP	0,38	0,34
CV(%)	11,8	10,7
As	-0,13	0,43
K	0,76	-0,34

N: número de dados; Mín.: valor mínimo; Máx.: valor máximo; Q_1 : primeiro quartil; Q_3 : terceiro quartil; DP: desvio-padrão; CV: coeficiente de variação; As: assimetria e K: curtose.

As estimativas obtidas pelos três modelos consideraram a distribuição t -student com grau de liberdade igual a 10 ($v = 10$).

Tabela 22: Parâmetros $\hat{\beta}$ estimados pelo algoritmo EM para as variáveis Prod e Prod-13-19

Variável	Modelo	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$
Prod	Exp	4,56 (1,07)	-0,10 (0,12)	0,06 (0,12)	-0,01 (0,09)	-0,29 (0,64)	0,06 (0,52)	-0,86 (0,80)
	Gaus	4,35 (1,07)	-0,09 (0,12)	0,07 (0,12)	-0,01 (0,09)	-0,27 (0,64)	0,13 (0,53)	-0,81 (0,80)
	Matérn	4,71 (1,07)	-0,11 (0,11)	0,05 (0,12)	-0,02 (0,09)	-0,29 (0,64)	0,02 (0,52)	-0,90 (0,79)
	k=0,7	3,96 (1,01)	0,05 (0,11)	-0,07 (0,11)	-0,02 (0,08)	0,03 (0,57)	0,26 (0,46)	-0,89 (0,71)
Prod-13-19	Exp	4,83 (1,78)	0,15 (0,19)	0,23 (0,20)	-0,11 (0,14)	0,39 (1,00)	1,38 (0,82)	-3,94 (1,26)
	Gaus	3,62 (1,01)	0,06 (0,10)	-0,09 (0,11)	-0,03 (0,07)	0,24 (0,54)	0,38 (0,43)	-0,90 (0,67)
	Matérn	3,62 (1,01)	0,06 (0,10)	-0,09 (0,11)	-0,03 (0,07)	0,24 (0,54)	0,38 (0,43)	-0,90 (0,67)
	k=0,7	3,62 (1,01)	0,06 (0,10)	-0,09 (0,11)	-0,03 (0,07)	0,24 (0,54)	0,38 (0,43)	-0,90 (0,67)

Entre parênteses são apresentados os erros padrões de cada parâmetro estimado.

Observando a Tabela 32, verifica-se as estimativas dos parâmetros $\hat{\phi}_1$, $\hat{\phi}_2$ e $\hat{\phi}_3$ apresen-

taram variações diversas após a exclusão dos valores influentes.

Tabela 23: Parâmetros espaciais estimados pelo algoritmo EM para as variáveis Prod e Prod-13-19

Variável	Modelo	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$
Prod	Exp	0,08	0,05	138,4
		(0,03)	(0,04)	(1,64)
	Gaus	0,09	0,04	147,1
		(0,03)	(0,03)	(0,25)
	Mat	0,08	0,05	192,9
		(0,03)	(0,05)	(220)
Prod-13-19	Exp	0,09	0,001	105,3
		(0,02)	(0,001)	(0,96)
	Gaus	0,03	0,02	169,9
		(0,01)	(0,01)	(0,28)
	Mat	0,04	0,01	177,8
		(0,01)	(0,01)	(576,3)

Entre parênteses são apresentados os erros padrões de cada parâmetro estimado.

Entre os três modelos, considerando os critérios apresentados na Tabela 24, optou-se novamente pelo modelo gaussiano para a construção dos mapas temáticos.

Tabela 24: Resultados da validação cruzada para a variável Prod-13-19

Modelo	EM	EMR	S	S_{EMR}	EA	IAC	MVL
Exp	-0,0007	0,2677	0,3399	3,8617	0,27	272,1	-126,07
Gaus	-0,0093	-0,0357	0,8265	3,5767	0,64	564,5	-272,2
Matérn k=0,7	-0,0024	0,0255	0,3684	1,7433	0,29	388,2	-184,2

EM: erro médio; EMR: erro médio reduzido; S: desvio padrão dos erros; S_{EMR} : desvio padrão do erro médio e EA: erro absoluto; IAC: critério de transformação de Akaike; MVL: máximo valor do logaritmo da função verossimilhança.

A Figura 34 apresenta os mapas temáticos dos conjuntos de dados Prod e Prod-13-19, confeccionados por meio da interpolação por krigagem universal, a qual utiliza as covariáveis para a estimação dos parâmetros ($\theta = (\beta^T, \Phi^T)^T$). Os mapas foram construídos utilizando-se o modelo gaussiano assumindo distribuição t-student com grau de liberdade igual a 10 ($v = 10$). A Tabela 25 apresenta a porcentagem que cada classe da legenda representa nos mapas da variável.

Analizando os mapas da Figura 34 e a Tabela 25, observa-se que houve redução das

áreas com produtividade nos intervalos da 1^a classe ($2,70$ a $3,09$ t ha $^{-1}$), o qual passou de 33% para 19% da área, e também da 4^a classe ($3,50$ a $3,90$ t ha $^{-1}$), passando de 11% para 6%. Consequentemente, as áreas das demais classes apresentaram aumento, sendo mais acentuado na classe de $3,10$ a $3,29$ t ha $^{-1}$, que passou de 32% para 46%.

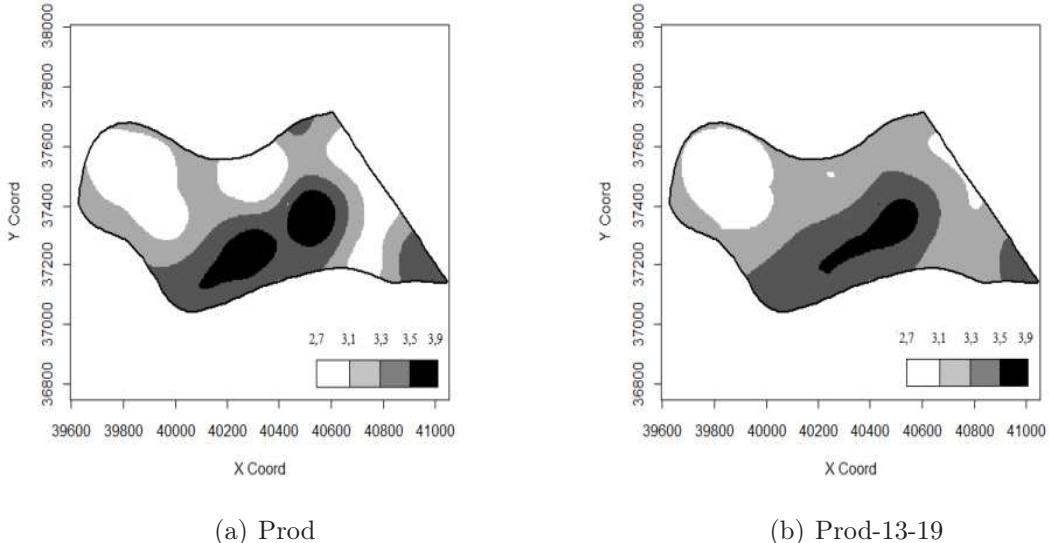


Figura 34: Mapa de produtividade em t ha $^{-1}$ com e sem as observações influentes, utilizando o algoritmo EM na estimação dos parâmetros.

Tabela 25: Porcentagem que cada classe da legenda representa nos mapas da variável

Variável	1 ^a classe ($2,70 - 3,09$)	2 ^a classe ($3,10 - 3,29$)	3 ^a classe ($3,30 - 3,49$)	4 ^a classe ($3,50 - 3,90$)
Prod	33%	32%	24%	11%
Prod-13-19	19%	46%	29%	6%

A Tabela 26 apresenta um estudo de comparação dos mapas temáticos por meio da matriz de erros entre as classes dos mapas Prod e Prod-13-19. O mapa Prod foi considerado como referência ou real, por possuir todos os dados e o mapa Prod-13-19 foi considerado como modelo. Observando-a, é possível verificar que há concordância no número de áreas para todas as classes, sendo mais acentuada na classe $3,10$ a $3,29$ t ha $^{-1}$.

Tabela 26: Matriz de erros entre as classes da variável Prod e Prod-13-19 [$t ha^{-1}$] em número de *pixels*

		Mapa Prod				Total
Classes		2,70–3,09	3,10–3,29	3,30–3,49	3,50–3,90	
Mapa	2,70–3,09	3857	387	0	0	4244
	3,10–3,29	3693	6333	456	0	10482
	3,30–3,49	0	478	4835	1334	6647
	3,50–3,90	0	0	226	1155	1381
Total		7550	7198	5517	2489	22754

Na Tabela 36 são apresentadas as medidas para a comparação dos mapas. Segundo Anderson *et al.* (1976), o índice de exatidão global obtido (0,71) é baixo. Esse índice foi utilizado para verificar a similaridade entre os mapas; logo, comprova-se que há diferenças entre eles. O índice *Kappa* obtido foi de 0,60 e, de acordo com a classificação de Kripendorff (1980), é baixo. Para finalizar, foi determinado o índice *Tau*, que foi classificado como baixo, já que o resultado encontrado foi de 0,61, isto significa que os mapas com e sem os pontos influentes não são semelhantes.

Tabela 27: Medidas obtidas a partir da matriz de erros entre as classes da variável Prod e Prod-13-19 $t ha^{-1}$

Medidas	Produtividade $t ha^{-1}$
Exatidão Global -EG	0,71
Kappa - K	0,60
Tau - T	0,61

Portanto, conclui-se que os elementos identificados na análise de influência local realmente são influentes na construção dos mapas temáticos da produtividade com as covariáveis RSP e Des.

4.3.1.2 Análise espacial utilizando o afastamento da verossimilhança usual

A Tabela 28 apresenta os resultados das estimativas dos parâmetros do modelo espacial linear da Equação (4.1) e dos parâmetros da estrutura de variabilidade espacial da Equação (2.7) para a produtividade da soja, por meio do algoritmo de NR. Foram ajustados os modelos teóricos exponencial (Exp), gaussiano (Gaus) e Matérn com parâmetro kappa=0,7 (Mat), assumindo-se distribuição t-student com grau de liberdade igual a 3.

A hipótese de que a distribuição da população da qual se retirou a determinada amostra segue uma distribuição de probabilidade t-student com grau de liberdade $v = 3$ foi confirmada pelo teste de Liliford.

Tabela 28: Parâmetros espaciais estimados por meio do algoritmo de NR assumindo-se distribuição t-student com grau de liberdade $v=3$ para a variável produtividade

Modelo	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$
Exp	4,38 (1,81)	-0,10 (0,18)	0,11 (0,19)	-0,004 (0,15)	-0,27 (1,04)	0,12 (0,86)	-0,87 (1,33)	0,09 (0,12)	0,06 (0,11)	60,01 (0,33)
Gaus	4,05 (1,68)	-0,10 (0,17)	0,09 (0,18)	0,00 (0,14)	-0,03 (0,96)	0,13 (0,80)	-0,80 (1,23)	0,03 (0,10)	0,07 (0,08)	70,02 (0,02)
Mat	3,78 k=0,7 (1,93)	-0,06 (0,19)	0,07 (0,21)	-0,01 (0,16)	0,02 (1,09)	0,18 (0,91)	-0,69 (1,36)	0,03 (0,45)	0,11 (0,16)	70,00 (206,99)

Entre parênteses são apresentados os erros padrões de cada parâmetro estimado.

A Tabela 29 apresenta os resultados dos critérios de validação dos modelos para a variável produtividade, utilizando os modelos teóricos exponencial (Exp), gaussiano (Gaus) e família Matérn com $\kappa=0,7$ (Matérn), pelo algoritmo de NR. Considerando esses resultados, verifica-se que o melhor modelo ajustado é o gaussiano.

Tabela 29: Resultados da validação cruzada para os parâmetros estimados pelo algoritmo de NR para a variável produtividade

Modelo	EM	EMR	S	$S_{E\bar{M}R}$	EA	IAC	MVL
Exp	-0,0005	-0,0079	0,3556	1,3367	0,2696	169,45	-74,72
Gaus	-0,0004	-0,0005	0,3451	0,6686	11,98	-95,01	-58,25
Matérn k=0,7	-0,0013	0,0078	0,3653	1,1528	0,282	193,66	-86,76

EM: erro médio; EMR: erro médio reduzido; S: desvio padrão dos erros; $S_{E\bar{M}R}$: desvio padrão do erro médio e EA: erro absoluto; IAC: critério de transformação de Akaike; MVL: máximo valor do logaritmo da função verossimilhança.

Nos gráficos C_i e $|l_{max}|$ apresentados nas Figuras 35(a) e Figuras 35(b), considerando o limite do gráfico definido na Equação (3.10), observa-se que ambos identificaram os elementos 20 e 24 como influentes no afastamento da verossimilhança.

A influência na função de covariância avaliada pelo gráfico l_{Ch} e a influência no preditor linear avaliada pelo gráfico l_p , ambos construídos a partir da técnica usual, também identificaram os elementos 20 e 24 como influentes.

Seguindo com a análise do gráfico LG , Figura 37(a), observa-se que o elemento 13 foi identificado como influente.

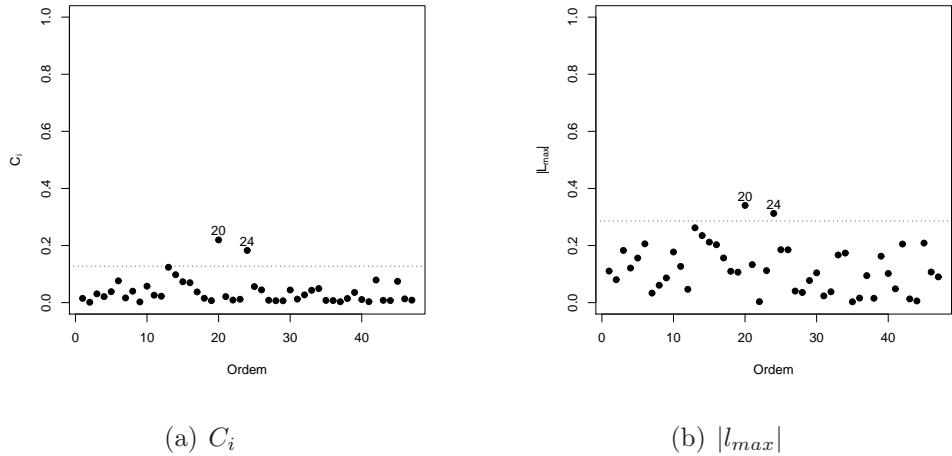


Figura 35: Gráficos de diagnóstico de influência local C_i e $|l_{max}|$ para a produtividade de soja

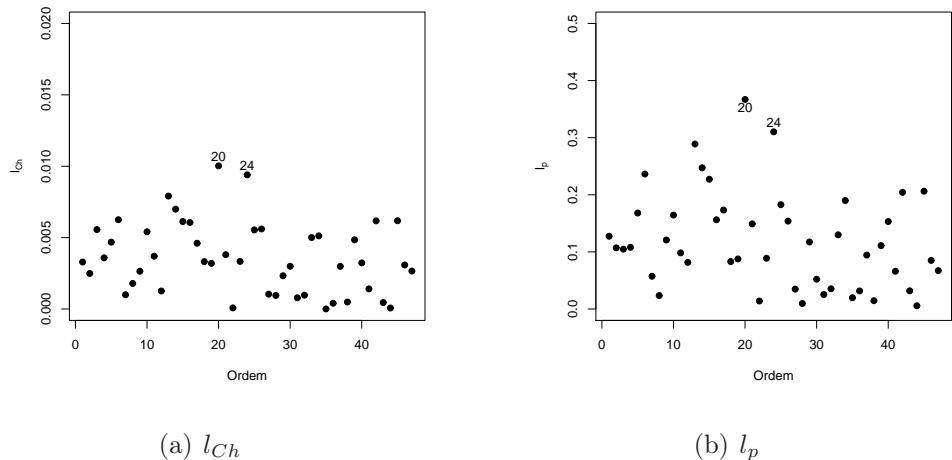
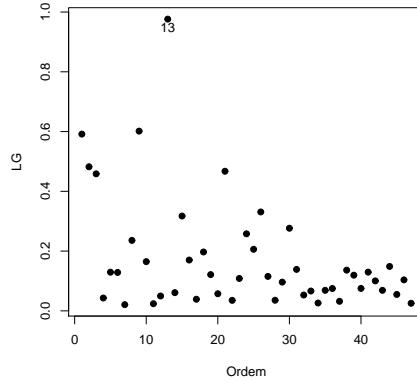


Figura 36: Gráficos de diagnóstico de influência local l_{Ch} e l_p para a produtividade de soja

(a) LG Figura 37: Gráfico de diagnóstico de influência local LG para a produtividade de soja

Optou-se pela retirada das observações 20 ($3,02 \text{ t ha}^{-1}$) e 24 ($4,09 \text{ t ha}^{-1}$) para dar continuidade nas análises. Os conjuntos de dados foram denominados por Prod: dados totais; Prod-20-24: dados com exclusão dos elementos 20 e 24.

Observa-se na Tabela 30 que a eliminação dos elementos 20 e 24 alterou a média, máximo, o desvio padrão e o coeficiente de variação, que passou de 11,79% para 11,4%. A retirada desses elementos também alterou a assimetria e a curtose dos dados.

Tabela 30: Estatísticas descritivas para as variáveis Prod e Prod-20-24

Variável	Prod	Prod-20-24
n	47	45
Média	3,23	3,21
Mín.	2,09	2,09
Máx	4,09	4,00
Q_1	3,01	3,00
Mediana	3,19	3,19
Q_3	3,49	3,49
DP	0,38	0,36
CV(%)	11,8	11,4
As	-0,13	-0,34
K	0,76	0,82

N: número de dados; Mín.: valor mínimo; Máx.: valor máximo; Q_1 : primeiro quartil; Q_3 : terceiro quartil; DP: desvio-padrão; CV: coeficiente de variação; As: assimetria e K: curtose.

As estimativas obtidas pelos três modelos consideraram a distribuição t -student com grau de liberdade igual a 10 ($v = 10$).

As estimativas do vetor de parâmetros $\hat{\beta}$, obtidas a partir do conjunto de dados Prod-

20-24, apresentaram algumas variações comparadas com as estimativas obtidas com os dados totais (Prod).

Tabela 31: Parâmetros $\hat{\beta}$ estimados pelo algoritmo NR para as variáveis Prod e Prod-20-24

Variável	Modelo	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$
Prod	Exp	4,38 (1,81)	-0,10 (0,18)	0,11 (0,19)	-0,004 (0,15)	-0,27 (1,04)	0,12 (0,86)	-0,87 (1,33)
	Gaus	4,05 (1,68)	-0,10 (0,17)	0,09 (0,18)	0,00 (0,14)	-0,03 (0,96)	0,13 (0,80)	-0,80 (1,23)
	Mat	3,78 k=0,7	-0,06 (1,93)	0,07 (0,19)	-0,01 (0,21)	0,02 (0,16)	0,18 (1,09)	-0,69 (0,91)
	Exp	3,97 (1,61)	-0,08 (0,61)	0,03 (0,17)	-0,003 (0,13)	-0,52 (0,92)	0,19 (0,75)	-0,26 1.17
	Gaus	4,18 (2,26)	-0,09 (0,24)	0,09 (0,25)	0,02 (0,19)	-0,71 (1,34)	0,13 (1,09)	-0,31 (1,76)
	Matérn	3,76 k=0,7	-0,06 (1,71)	0,05 (0,17)	0,003 (0,18)	-0,39 (0,14)	0,14 (0,98)	-0,22 (0,79)
Prod-20-24	Gaus	4,18 (2,26)	-0,09 (0,24)	0,09 (0,25)	0,02 (0,19)	-0,71 (1,34)	0,13 (1,09)	-0,31 (1,76)
	Matérn	3,76 k=0,7	-0,06 (1,71)	0,05 (0,17)	0,003 (0,18)	-0,39 (0,14)	0,14 (0,98)	-0,22 (0,79)
	Exp	3,97 (1,61)	-0,08 (0,61)	0,03 (0,17)	-0,003 (0,13)	-0,52 (0,92)	0,19 (0,75)	-0,26 1.17
	Gaus	4,18 (2,26)	-0,09 (0,24)	0,09 (0,25)	0,02 (0,19)	-0,71 (1,34)	0,13 (1,09)	-0,31 (1,76)
	Matérn	3,76 k=0,7	-0,06 (1,71)	0,05 (0,17)	0,003 (0,18)	-0,39 (0,14)	0,14 (0,98)	-0,22 (0,79)
	Exp	3,97 (1,61)	-0,08 (0,61)	0,03 (0,17)	-0,003 (0,13)	-0,52 (0,92)	0,19 (0,75)	-0,26 1.17

Entre parênteses são apresentados os erros padrões de cada parâmetro estimado.

Observando a Tabela 32, verifica-se que as estimativas dos parâmetros $\hat{\varphi}_1$, $\hat{\varphi}_2$ e $\hat{\varphi}_3$ apresentaram variações diversas após a exclusão dos valores influentes.

Tabela 32: Parâmetros espaciais estimados pelo algoritmo EM para as variáveis Prod e Prod-20-24

Variável	Modelo	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_3$
Prod	Exp	0,09 (0,12)	0,06 (0,11)	60,01 (0,33)
	Gaus	0,03 (0,10)	0,07 (0,08)	70,02 (0,02)
	Mat	0,03 k=0,7	0,11 (0,16)	70,00 (206,9)
	Exp	0,04 (0,09)	0,06 (0,07)	100,17 (0,68)
	Gaus	0,10 (0,28)	0,06 (0,07)	60,01 (0,01)
	Mat	0,05 k=0,7	0,07 (0,13)	69,98 (254,9)
Prod-20-24	Matérn	0,03 k=0,7	0,11 (0,16)	70,00 (206,9)
	Exp	0,04 (0,09)	0,06 (0,07)	100,17 (0,68)
	Gaus	0,10 (0,28)	0,06 (0,07)	60,01 (0,01)
	Mat	0,05 k=0,7	0,07 (0,13)	69,98 (254,9)
	Exp	0,04 (0,09)	0,06 (0,07)	100,17 (0,68)
	Gaus	0,10 (0,28)	0,06 (0,07)	60,01 (0,01)

Entre parênteses são apresentados os erros padrões de cada parâmetro estimado.

Entre os três modelos, considerando os critérios apresentados na Tabela 33, optou-se novamente pelo modelo gaussiano para a construção dos mapas temáticos.

Tabela 33: Resultados da validação cruzada para a variável Prod-20-24

Modelo	EM	EMR	<i>S</i>	<i>S_{EMR}</i>	EA	IAC	MVL
Exp	0,0023	0,0106	0,3311	<u>1,8322</u>	<u>0,25</u>	276,03	<u>-128,02</u>
Gaus	<u>0,0019</u>	-0,0182	<u>0,3382</u>	2,9912	0,26	<u>218,86</u>	-99,13
Matérn	0,0027	<u>0,0069</u>	0,3311	1,8948	<u>0,25</u>	265,53	-122,76
k=0,7							

EM: erro médio; EMR: erro médio reduzido; *S*: desvio padrão dos erros; *S_{EMR}*: desvio padrão do erro médio e EA: erro absoluto; IAC: critério de transformação de Akaike; MVL: máximo valor do logaritmo da função verossimilhança; kappa=0,7.

A Figura 34 apresenta os mapas temáticos dos conjuntos de dados Prod e Prod-20-24, confeccionados por meio da interpolação por krigagem universal, a qual utiliza as covariáveis para a estimação dos parâmetros ($\theta = (\beta^T, \Phi^T)^T$). Os mapas foram construídos utilizando-se o modelo gaussiano assumindo distribuição t-student com grau de liberdade igual a 10 ($v = 10$). A Tabela 34 apresenta a porcentagem que cada classe da legenda representa nos mapas da variável.

Analizando os mapas da Figura 38 e a Tabela 34, observa-se que houve aumento das áreas com produtividade no intervalo da 1^a classe (2,7 a 3,1 t ha⁻¹), o qual passou de 16,5% para 56,1% da área, e redução das demais classes.

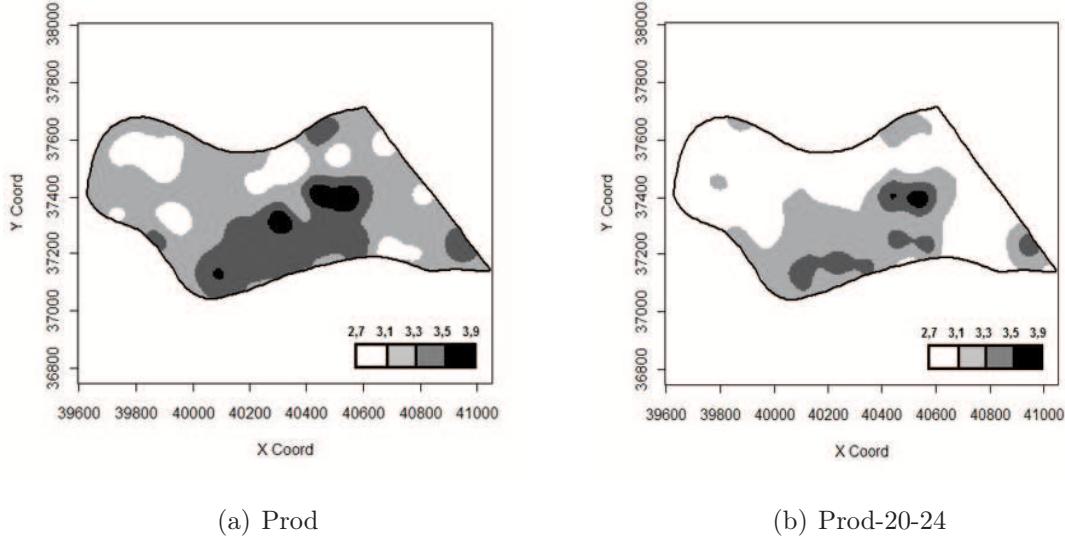


Figura 38: Mapa de produtividade em $t \text{ ha}^{-1}$ com e sem as observações influentes, utilizando o algoritmo de NR na estimativa dos parâmetros.

Tabela 34: Porcentagem que cada classe da legenda representa nos mapas da variável

Variável	1 ^a classe (2,70 – 3,09)	2 ^a classe (3,10 – 3,29)	3 ^a classe (3,30 – 3,49)	4 ^a classe (3,50 – 3,90)
Prod	16,5%	53,8%	26,1%	3,6%
Prod-20-24	56,1%	34,4%	8,7%	0,8%

A Tabela 35 apresenta um estudo de comparação entre os mapas temáticos por meio da matriz de erros considerando o mapa da Prod como referência ou real, por possuir todos os dados, e o mapa Prod-20-24 como modelo. Observando-a, é possível verificar que há concordância no número de áreas para todas as classes, sendo mais acentuada na classe 2,70 a 3,09 $t \text{ ha}^{-1}$.

Tabela 35: Matriz de erros entre as classes da variável Prod e Prod-20-24 [$t \text{ ha}^{-1}$] em número de *pixels*

		Mapa Prod				Total
Classes		2,70–3,09	3,10–3,29	3,30–3,49	3,50–3,90	
Mapa	2,70–3,09	3765	8944	62	0	12771
	3,10–3,29	0	3299	4284	244	7827
Prod 20-24	3,30–3,49	0	8	1581	402	1991
	3,50–3,90	0	0	0	165	165
Total		3765	12251	5927	811	22754

O índice de exatidão global obtido (0,38) é baixo, indicando pequena similaridade entre os mapas. O índice *Kappa* obtido foi baixo (0,12) e o índice *Tau* também classificado como baixo, já que o resultado encontrado foi de 0,18. Isto implica na influência dos pontos 20 e 24 na construção do mapa temático.

Tabela 36: Medidas obtidas a partir da matriz de erros entre as classes da variável Prod e Prod-20-24 $t\ ha^{-1}$

Medidas	Produtividade $t\ ha^{-1}$
Exatidão Global - <i>EG</i>	0,38
Kappa - <i>K</i>	0,12
Tau - <i>T</i>	0,18

5 CONCLUSÕES

Neste trabalho, desenvolveu-se um estudo de inferência e análise de diagnósticos pela influência local em modelos espaciais lineares com distribuição t -student n -variada. Foram realizadas implementações computacionais e apresentadas duas técnicas para a análise de influência local.

O algoritmo de NR foi eficiente na estimativa dos parâmetros em modelos espaciais lineares com distribuição t -student n -variada. No entanto, a proposta da construção do algoritmo EM para a estimativa desses mesmos parâmetros mostrou-se satisfatória, já que a presença de assimetria nos dados tornou o algoritmo EM computacionalmente instável.

A análise de diagnóstico foi baseada em técnicas gráficas de influência local, construídas a partir da técnica usual apresentada por Cook (1986) e da proposta de Zhu e Lee (2001) sobre o Q-afastamento da verossimilhança para dados completos. Em muitos dos gráficos de diagnóstico construídos, identificou-se pela técnica do Q-afastamento da verossimilhança para dados completos os mesmos valores que pela técnica usual como possíveis valores influentes. Esses gráficos, confeccionados por meio das duas técnicas, permitiram a análise de influência local no afastamento da verossimilhança, na matriz de covariância, no preditor linear e ainda nos valores preditos. As duas técnicas apresentaram, na maioria dos casos, ponderações similares.

Dessa forma, através deste trabalho, confirmou-se a viabilidade e eficiência da técnica que foi proposta.

6 CONSIDERAÇÕES FINAIS

O construção dos algoritmos NR e EM, assim como as análises de influência local pelo afastamento da verossimilhança usual e pelo Q-afastamento da verossimilhança para os dados completos, podem ser extendidos para outros modelos da classe das distribuições simétricas.

A simulação foi realizada com o modelo exponencial, pois, inserida a perturbação nos dados simulados, causou-se assimetria nos mesmos, e os modelos gaussiano e Matérn apresentaram maior sensibilidade a essa assimetria. Dessa forma que não foi possível estimar os parâmetros pelo algoritmo EM. Sendo assim, sugere-se que seja realizada outra forma de perturbação nos dados que não cause assimetria. Ou, então, que o algoritmo EM seja desenvolvido para estimar, além dos parâmetros $\theta = (\beta^T, \Phi^T)^T$, o grau de liberdade v e a assimetria.

No contexto de diagnósticos, pode-se considerar outros esquemas de perturbação para realizar a análise de influência local em análise espacial de dados.

Este trabalho espera ter contribuído com o estudo da variabilidade espacial ao introduzir os métodos de influência local pelo afastamento da verossimilhança usual e pelo Q-afastamento da verossimilhança para os dados completos.

BIBLIOGRAFIA

- AKAIKE, H. Information theory and an extension of the maximum likelihood principle. *Journal of the Royal Statistical Society, London*, v. 51, p. 469-483, 2002.
- ANDERSON, J. F.; HARDY, E. E.; ROACH, J. T.; WUTMER, R. E. A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey Professional Paper 964, U.S. Geological Survey, Washington, DC, p.28, 1976.
- ARELLANO-VALLE, R.B. Distribuições Elípticas: Propriedades, Inferência e Aplicações a Modelos de Regressão. Tese, IME-USP, 1994.
- ARELLANO-VALLE, R.B.; BOLFARINE, H. On some characterizations of the t-distribution, *Statist. Probab. Letters*, v.25, n.1, p.79-85, 1995.
- AZALLINI, A.; CAPITANIO, A. Statistical applications of the multivariate skew normal distribution. *J. R. Stat. Soc., B*, v.61, parte 3, p.579-602, 1999.
- BENNETT, B.M. On a certain multivariate non-normal distribution. *Proceedings of the Cambridge Philosophical Society*, v.57, p.434-436, 1961.
- BORSSOI, J. A.; URIBE-OPAZO, M. A.; GALEA-ROJAS, M. J. Diagnostics techniques applied in geoestatistics for agricultural data analysis. *Revista Brasileira de Ciência do Solo*, v. 6, n. 6, p.1-16, 2009.
- CADIGAN, N. G.; FARREL, P. J. Generalized Local Influence with Applications to Fish Stock Cohort Analysis, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, v. 51, n. 4, p.469-483, 2002.
- CANARACHE, A. P. A generalized semi-empirical model estimating soil resistance to penetration. *Soil Till Res*, Amsterdam, v.16, n.1, p.51-70, 1990.
- CHRISTENSEN,R.; PEARSON, L. M.; JHONSON, W. Prediction diagnostics for spatial linear models. *Biometrics*, v.79, n.3, p.583-591, 1992a.

CHRISTENSEN,R.; PEARSON, L. M.; JHONSON, W. Case-deletion diagnostics for mixed models. *Technometrics*, v.34, p.38-45, 1992b.

CHRISTENSEN, R.; JOHNSON, W.; PEARSON, L. M. Covariance function Diagnostics for Spatial Linear Models. International Association for Mathematical Geology, New York, v. 25, n.2, p. 145-160, 1993.

COAMO/COODETEC. Fertilidade do solo e nutrição de plantas. Campo Mourão/Cascavel, 1998.

CONGALTON, R. G.; GREEN, K. Assesing the accuracy of remotely sensed data: principles and practices. New York: Lewis Publisher, p.130, 1999.

COHEN, J.A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, v. XX, n.1, 1960.

COOK, R.D. Influence assessment.*Journal of Applied Statistics*, v.14,p.117-131, 1987.

COOK, R. D. Assessment of Local Influence (with discussion). *Journal of the Royal Statistical Society, Series B*, London, v. 48, n.2, p. 133-169, 1986.

CORNISH, E. A. The multivariate t-distribution associated with a set of normal sample deviates. *Austral. J. Physics*, v.7, p.531-542, 1954.

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

CYSNEIROS, F.J; PAULA, G. A; GALEA, M. Modelos Simétricos Aplicados. 9^a Escola de Modelos de Regressão, Águas de São Pedro, SP, 2005, 89p.

DAWID, A. P. Spherical matrix distributions and a multivariate model. *J. R. Statist. Soc. B*, v.39, n. 2, p.254-61, 1977.

DAWID, A.P. Some matrix-variate distribution theory: notational considerations and a Bayesian application, *Biometrika*, v.68, n.1, p.265-274, 1981.

DE BASTIANI, F.; URIBE-OPAZO, M. A.; DALPOSSO, G. H.; BORSSOI, J. A.; ASSUMPÇÃO, R. A. B. Uso de Covariáveis na Variabilidade Espacial da Resistência do Solo ‘a Penetração. Avances em Ingenieria Rural 2007-2009. 1 ed. Rosario: CADIR 2009, v. 1, p. 28-35, 2009.

DEMPSSTER, A. P., LAIRD, N. M. RUBIN, D. B. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, B*, v.39, n.1, p.1-38, 1977.

DEMPSSTER, A. P.; LAIRD, N. M., RUBIN, D. B. Iteratively reweighted least squares for linear regression when errors are normal independent distributed. In *Multivariate Analysis*. North-Holland, Amsterdam, Ed. , p. 35-57, 1980.

DEXTER, A. R. Soil physical quality. Part I. Theory, effects of soil texture, density, and organic matter, and effects on root growth. *Geoderma*, v.120, p.201-214, 2004.

DICKEY, J. M. Matricvariate Generalizations of the Multivariate t Distribution and the Inverted Multivariate t Distribution. *Annals of Mathematical Statistics*, v.38, n. 2, p.511-518, 1967.

DUNNETT, C. W., SOBEL, M. A Bivariate Generalization of Student’s t-Distribution, With Tables for Certain Special Cases. *Biometrika*, v.41, n.1/2, p.153-169, 1954.

EMBRAPA. Serviço Nacional de Levantamento e Conservação de Solos. Manual de Métodos de Análise de Solo. Rio de Janeiro. Centro Nacional de Pesquisas de Solos, 2 ed., 1997, 212p.

ESCOBAR, L. A.; MEEKER, W. Q. Assessing influence in regression analysis with censored data. *Biometrics*, v.48, p.507-528, 1992.

FANG, K.T.; KOTZ, S.; Ng, K.W. *Symmetric Multivariate and Related Distributions*, Chapman & Hall, London, 1990, 319p.

FARACO M. A.; URIBE-OPAZO, M. A.; SILVA, E. A.; JOHANN J. A.; BORSSOI, J. A. Seleção de modelos de variabilidade espacial para elaboração de mapas temáticos de atributos físicos do solo e produtividade da soja. *RBCS*, v. 32, n. 2, p. 463-476. 2008.

FERREIRA, C. S. Inferência e diagnóstico em modelos assimétricos. Tese (Doutorado), IME-USP, 2008.

- FORBES, A. D. Classification algorithm evaluation: five performance measures based on confision matrices. *Journal of Clinical Monitoring* 11:, p.189-206, 1995.
- FUNG, W.K; KWAN,C.W. A Note on Local Influence Based on Normal Curvature. *Journal of the Royal Statistical Society. Series B*, v.59, n.4, p.839-843, 1997.
- GALEA, M. Calibração Comparativa Estrutural e Funcional. Tese de doutorado. IME-USP, 1995.
- GALEA, M.; BOLFARINE, H.; VILCA-LABRA, M. Local influence in comparative calibration models. *Biometrical Journal*,v.44, p.59-81, 2002.
- GALEA, M.; PAULA, G.A.; URIBE-OPAZO, M.A. On influence diagnostic in univariate elliptical linear regression models. *Statistical Papers*, v.44, p.23-45, 2003.
- GUEDES, L. P. C.; URIBE-OPAZO, M. A.; JOHANN, J. A.; SOUZA, E. G. Anisotropia no estudo da variabilidade espacial de algumas variáveis químicas do solo. *Revista Brasileira de Ciência do Solo*, v.32, n.6, p.2217-2226, 2008.
- GONG, P.; HOWARTH, P.J. An assessment of some factors influencing multispectral land-cover classification. *Photogrammetric Engineering and Remote Sensing*, v.56, n.5, p.597-603, 1990.
- GONÇALVES, A. C. A. A variabilidade especial de propriedades físicas do solo para fins de manejo da irrigação. Piracicaba, 1997. 118 f. Tese (doutorado) - Escola Superior de Agricultura 'Luiz de Queiroz', Universidade de São Paulo, 1997.
- IHAKA, R.; GENTLEMAN, R. A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, Salt Like City, v.5, n.3, p.229-314. 1996. Software disponível em: <<http://www.R-project.org>>. Acesso: 31 de ago. 2003.
- ISAAKS, E. H.; SRIVASTAVA, R. M. *An Introduction to Applied Geoestatistics*. New York: Oxford University Press, 1989, 560p.
- 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.

JOURNEL, A. G. Fundamentals of geoestatistics in five lessons short course in Geology. American Geophysical Union, v.8, Washington, 1989, 40p.

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

KSHIRSAGAR, A. M. Some extensions of the multivariate t-distribution and the multivariate generalization of the distribution of the regression coefficient. Proc. Cambridge Phil. Soc., v.57, p.80-85, 1960.

KRIPPENDORFF, K. Content Analysis: an Introduction to its Methodology. Beverly Hills: Sage Publications. p.189, 1980.

LANDIM, P.M.B. Sobre geoestatística e mapas. Terræ Didática, v. 2, p. 19-33, 2006.

LANGE, K.; LITTLE, R. J. A.; TAYLOR, J. Robust Statistical Using the T Distribution, Journal of the American Statistical Association, v.84, n.408, p.881-896, 1989.

LANGE, K., SINSHEIMER, J. Normal/Independent Distributions and Their Applications in Robust Regression. Journal of Computational and Graphical Statistics,v.2, p.175-198, 1993.

LEE, S.-Y; XU, L. Case-Deletion Diagnostics for Factor Analysis Models With Continuous and Ordinal Categorical Data. Sociological Methods & Research, v.31 , n.3, p.389-419, 2003.

LEE, S.-Y; XU, L. Influence analyses of nonlinear mixed-effects models. Computational Statistics and Data Analysis, v.45, p.321-341, 2004.

LETEY, J. Relationship between soil physical properties and crop reduction. Adv. Soil Science, v.1, p. 277-294, 1985.

LITTLE, R. J. A. A Test of Missing Completely at Random for Multivariate Data With Missing Values, Journal of the American Statistical Association, v. 83, n.404, p.1198-1202, 1988.

LITTLE, R. J. A.; RUBIN, D. B. Statistical Analysis with Missing Data. New York: Wiley,

1987.

LIU, C. H.; RUBIN, D. B. The ECME Algorithm: A Simple Extension of EM and ECM with Faster Monotone Convergence. *Biometrika*, v.81,n.4, p.633-648, 1994.

LIU, C.; RUBIN, D. B. ML Estimativa of the Distribution Using EM and its Extensions, ECM and ECME. *Sinica Statistica*, v.5, n.5, p.19-39, 1995.

MA, Z. ; REDMOND, R. L. Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineering and Remote Sensing*. Bethesda, v.61(4), p.453-439. 1995.

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.

MATERON, G. Principles of geoestatistics. *Economic Geology*, v.58, n.1, p.1246-1266, 1963.

MERCANTE, E.; URIBE-OPAZO, M. A.; SOUZA, E. G. de . Variabilidade Espacial e Temporal da Resistência Mecânica do Solo à Penetração em Áreas com e sem Manejo Químico Localizado. *Revista Brasileira de Ciência do Solo*, v. 27, n. 6, p. 1149-1159, 2003.

MILITINO, A. F.; PALACIUS, M. B.; UGARTE, M. D. Outliers Detection in Multivariate Spatial Linear Models. *Journal statistical planning and inference*, v.136: 125-146, 2004.

NADARAJAH S.; KOTZ, S. Mathematical Properties of the Multivariate t Distribution. *Acta Ap. Math.*, v.89 , p.53-84, 2005.

NOBRE, J. S.; SINGER, J. M. Fixed and random effects leverage for influence analysis in linear mixed models. Departamento de Estatística IME-USP, 2006.

ORTEGA, E; BOLFARINE, H; PAULA G. Influence diagnostics in generalized log-gamma regression models. *Computacional Statistics & Data Analisys 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 Analisys Journal*, Holanda, v. 51, p. 4354-4368, 2007.

PAULA, G. Leverage in Inequality - Constrained Regression Models. *Journal of the Royal Statistical Society. Série D*, v.48, n.4, p.529-538, 1999.

POON, W.-Y.; POON, Y. S. Conformal Normal Curvature and Assessment of Local Influence. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, v.61, n.1, p.51-61, 1999.

R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Disponível em: <<http://www.R-project.org>>. Acesso em 3 jun. 2009.

RIBEIRO JR, P. J. Métodos geoestatísticos no estudo da variabilidade espacial de parâmetros do solo. Piracicaba, 1995. 99f. Dissertação (Mestrado) - Escola Superior de Agricultura ‘Luiz de Queiroz’, Universidade de São Paulo.

RIBEIRO JR, P. J.; DIGGLE P. J. geoR: A package for geostatistical analysis. *R-NEWS*, 01, <<http://cran.r-project.org/doc/Rnews>>. 2001.

RIBEIRO JR, P. J. Estatística no espaço: modelos e métodos para análise de dados georreferenciados. 33^a Reunião Regional da Associação Brasileira de Estatística-ABE, Curitiba-PR, 2002. Anais 33^a Reunião Regional da Associação Brasileira de Estatística-ABE. 2002.

SALGADO, F. A. O. Diagnóstico de Influência em Modelos Elípticos com Efeitos Mistos. Tese (Doutorado). IME-USP, 2006.

SCHALL, R.; DUNNE, T. T. A Note on the Relationship Between Parameter Collinearity and Local Influence. *Biometrika*, v.79, n.2, p.399-404, 1992.

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.; URIBEOPAZO, 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. 03, p. 80-92, 1999.

St. LAURENT, R.; COOK, D. Leverage, local influence and curvature in nonlinear regression. Journal of the American Statistical Association, Alexandria, v. 87, p. 985-990, 1992.

SUTRADHAR, B. C. Score tests for the covariance matrix of the Elliptical t-distribution. Journal of Multivariate Analysis, p.46, p.1-12, 1993.

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

TONG, Y. L. The multivariate Normal Distribution. Springer-Verlag, Nova York, 1990.

VERBEKE G.; MOLENBERGHS, G. Linear mixed models for longitudinal data. New York: Springer. 306p, 2000.

WEBSTER, W.; OLIVER, M. Geoestatistics for Environmental Scientists, ed.2, 315p, 2007.

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

XIE, F.C.; WEI, B. C.; LIN, J. G. Case-deletion Influence Measures for the Data from Multivariate t Distributions. Journal of Applied Statistics, Vol. 34, No. 8, p. 907-921, 2007.

ZHU, H. T.; LEE, S. Y. Local influence for incomplete-data models. J. R. Statist. Soc. B, v.63, Part.1, p.111-126, 2001.

ZHU, H.; ZANG, H. A diagnostic procedure based on local influence. Biometrika, v.91, n.3, p.579-589, 2004.

APÊNDICE A - Demonstrações da obtenção das matrizes de informação

A.1 Obtenção da matriz de informação usual do modelo t -student n -variado

Seja o logaritmo da função de máxima verossimilhança dado por:

$$l(\boldsymbol{\theta}) = K_n - \frac{1}{2} \log |\boldsymbol{\Sigma}| + \log(g(\delta)) \quad (\text{A.1})$$

em que $K_n = \log \left[\frac{\Gamma(\frac{v+n}{2})}{(\pi v)^{n/2} \Gamma(v/2)} \right]$ e $g(\delta) = \left(1 + \frac{\delta}{v}\right)^{-\left(\frac{v+n}{2}\right)}$.

A.1.1 Função escore

A função escore é dada por $U(\boldsymbol{\theta}) = (U(\boldsymbol{\beta})^T, U(\boldsymbol{\Phi})^T)^T$, sendo $\boldsymbol{\Phi} = (\varphi_1, \varphi_2, \varphi_3)^T$ e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$.

Seja $U(\boldsymbol{\beta}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} \{\log g(\delta)\}$, ou seja

$$U(\boldsymbol{\beta}) = W_{g(\boldsymbol{\delta})} \frac{\partial \boldsymbol{\delta}}{\partial \boldsymbol{\beta}},$$

$$\text{onde } W_{g(\boldsymbol{\delta})} = \frac{g'(\boldsymbol{\delta})}{g(\boldsymbol{\delta})} = -\frac{1}{2} \left(\frac{v+n}{v+\boldsymbol{\delta}} \right).$$

Porém, $\delta = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ e $\frac{\partial \boldsymbol{\delta}}{\partial \boldsymbol{\beta}} = -2\mathbf{X}^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$, então:

$$U(\boldsymbol{\beta}) = -2W_{g(\boldsymbol{\delta})} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$$

$$U(\boldsymbol{\Phi}) = \frac{\partial l(\boldsymbol{\theta})}{\partial \varphi_j} = \frac{\partial}{\partial \varphi_j} \left\{ -\frac{1}{2} \log |\boldsymbol{\Sigma}| + \log g(\delta) \right\} \text{ para } j = 1, 2, 3.$$

Como

$$\frac{\partial}{\partial \varphi_j} \{\log |\boldsymbol{\Sigma}|\} = \text{tr} \left(\boldsymbol{\Sigma}^{-1} \frac{\partial \boldsymbol{\Sigma}}{\partial \varphi_j} \right) \text{ e}$$

$$\frac{\partial \delta}{\partial \varphi_j} = -(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}),$$

então

$$U(\Phi) = -\frac{1}{2} \text{tr} \left(\Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \right) - W_{g(\delta)} \left[(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \right], \text{ para } j = 1, 2, 3.$$

A.1.2 Matriz de informação observada usual (L)

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

Os elementos da matriz observada usual são obtidos da seguinte forma:

$$\begin{aligned} L_{\boldsymbol{\beta}\boldsymbol{\beta}} &= \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}; \\ L_{\boldsymbol{\beta}\boldsymbol{\beta}} &= W_{g(\delta)} \frac{\partial^2 \delta}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} + W'_{g(\delta)} \frac{\partial \delta}{\partial \boldsymbol{\beta}} \cdot \frac{\partial \delta}{\partial \boldsymbol{\beta}^T}, \\ \text{onde } W'_{g(\delta)} &= \frac{\partial W_{g(\delta)}}{\partial \delta}. \end{aligned}$$

Assim:

$$L_{\boldsymbol{\beta}\boldsymbol{\beta}} = 2W_{g(\delta)} \mathbf{X}^T \Sigma^{-1} \mathbf{X} + 4W'_{g(\delta)} [\mathbf{X}^T \Sigma^{-1} \mathbf{r}] [\mathbf{r}^T \Sigma^{-1} \mathbf{X}];$$

$$L_{\boldsymbol{\beta}\boldsymbol{\varphi}} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\varphi}^T}, \text{ com elementos}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\varphi}_j^T} = W_{g(\delta)} \frac{\partial^2 \delta}{\partial \boldsymbol{\beta} \partial \boldsymbol{\varphi}_j^T} + W'_{g(\delta)} \frac{\partial \delta}{\partial \boldsymbol{\beta}} \cdot \frac{\partial \delta}{\partial \boldsymbol{\varphi}_j^T}, \text{ para } j = 1, 2, 3;$$

Então,

$$\begin{aligned} L_{\boldsymbol{\beta}\boldsymbol{\varphi}} &= 2W_{g(\delta)} \mathbf{X}^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) + \\ &\quad + 2W'_{g(\delta)} \left[\left(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta} \right)^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \right] [\mathbf{X}^T \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})] \end{aligned}$$

para $j = 1, 2, 3$.

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

$$L_{\varphi\varphi} = \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \varphi_i \partial \varphi_j^T}, \text{ com elementos}$$

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \varphi_i \partial \varphi_j^T} = -\frac{1}{2} \frac{\partial}{\partial \varphi_i^T} \left\{ \frac{\partial}{\partial \varphi_j} \log |\Sigma| \right\} + W_{g(\delta)} \frac{\partial^2 \delta}{\partial \varphi_j \partial \varphi_i^T} + W'_{g(\delta)} \frac{\partial \delta}{\partial \varphi_j} \frac{\partial \delta}{\partial \varphi_i^T}, \text{ para } i, j = 1, 2, 3.$$

Segue que

$$\begin{aligned} L_{\varphi\varphi} &= -\frac{1}{2} \left\{ \text{tr} \left[\Sigma^{-1} \left(\frac{\partial^2 \Sigma}{\partial \varphi_j \partial \varphi_j} - \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \right) \right] \right\} + \\ &\quad + W_{g(\delta)} \left\{ (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1} \left[\frac{\partial^2 \Sigma}{\partial \varphi_j \partial \varphi_j} - 2 \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \right] \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \right\} - \\ &\quad - W'_{g(\delta)} \left\{ \left[(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \right] \left[(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \right] \right\}. \end{aligned}$$

A.1.3 Matriz delta usual

$$\Delta_{\boldsymbol{\omega}} = (\Delta_{\boldsymbol{\beta}_{\boldsymbol{\omega}}}^T, \Delta_{\Phi_{\boldsymbol{\omega}}}^T)^T$$

Seja $\mathbf{Y}_{\boldsymbol{\omega}} = \mathbf{Y} + \boldsymbol{\omega}$, então

$$\delta_{\boldsymbol{\omega}} = (\mathbf{Y}_{\boldsymbol{\omega}} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1} (\mathbf{Y}_{\boldsymbol{\omega}} - \mathbf{X}\boldsymbol{\beta})^T \text{ e } \frac{\partial \delta_{\boldsymbol{\omega}}}{\partial \boldsymbol{\omega}} = 2\Sigma^{-1}(\mathbf{Y}_{\boldsymbol{\omega}} - \mathbf{X}\boldsymbol{\beta}).$$

Logo

$$\Delta_{\boldsymbol{\beta}_{\boldsymbol{\omega}}} = \frac{\partial U(\boldsymbol{\beta})}{\partial \boldsymbol{\omega}}$$

$$\Delta_{\boldsymbol{\beta}_{\boldsymbol{\omega}}} = -2W_{g(\delta_{\boldsymbol{\omega}})} \mathbf{X}^T \Sigma^{-1} - 4W'_{g(\delta_{\boldsymbol{\omega}})} \mathbf{X}^T \Sigma^{-1} (\mathbf{Y}_{\boldsymbol{\omega}} - \mathbf{X}\boldsymbol{\beta}) (\mathbf{Y}_{\boldsymbol{\omega}} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1};$$

e

$$\Delta_{\boldsymbol{\varphi}_{\boldsymbol{\omega}}} = \frac{\partial U(\boldsymbol{\varphi})}{\partial \boldsymbol{\omega}}$$

$$\Delta_{\boldsymbol{\varphi}_{\boldsymbol{\omega}}} = -2(\mathbf{Y}_{\boldsymbol{\omega}} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} \left[W_{g(\delta_{\boldsymbol{\omega}})} + W'_{g(\delta_{\boldsymbol{\omega}})} (\mathbf{Y}_{\boldsymbol{\omega}} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1} (\mathbf{Y}_{\boldsymbol{\omega}} - \mathbf{X}\boldsymbol{\beta}) \right].$$

A.2 Obtenção da matriz de informação \ddot{Q} do modelo t -student n -variado

Considere-se a esperança do logaritmo da função de verossimilhança completa, dado por:

$$Q(\theta|\hat{\theta}) = -\frac{n}{2}\log 2\pi + \frac{n}{2}c - \frac{1}{2}\log|\Sigma| - \frac{\delta\vartheta}{2} + \frac{v}{2} + \left(\frac{v}{2} - 1\right)c - \frac{v\vartheta}{2} - \log\Gamma\left(\frac{v}{2}\right)$$

com

$$c = \psi\left(\frac{v-n}{2}\right) - \log\left(\frac{\delta(\hat{\theta})+v}{2}\right);$$

$$\vartheta = \frac{v+n}{v+\delta(\hat{\theta})};$$

$$\delta = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^T \Sigma^{-1} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}).$$

A.2.1 Matriz de informação \ddot{Q}

A matriz de informação observada \ddot{Q} da distribuição t -student n -variada

$$\ddot{Q} = \begin{pmatrix} Q_{\boldsymbol{\beta}\boldsymbol{\beta}} & Q_{\boldsymbol{\beta}\Phi} \\ Q_{\Phi\boldsymbol{\beta}} & Q_{\Phi\Phi} \end{pmatrix},$$

com v graus de liberdade, é dada por:

$$Q_{\boldsymbol{\beta}\boldsymbol{\beta}} = \frac{\partial^2 Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}$$

$$Q_{\boldsymbol{\beta}\boldsymbol{\beta}} = -\vartheta \mathbf{X}^T \Sigma^{-1} \mathbf{X}.$$

Temos que $Q_{\boldsymbol{\beta}\Phi} = Q_{\Phi\boldsymbol{\beta}}^T$,

$$\text{logo } Q_{\boldsymbol{\beta}\Phi} = \frac{\partial^2 Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta} \partial \varphi_j^T}, \text{ com elementos}$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}|\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\beta} \partial \varphi_j^T} = \frac{\partial}{\partial \varphi_j} (\vartheta \mathbf{X}^T \Sigma^{-1} \mathbf{r}).$$

Assim temos

$$Q_{\boldsymbol{\beta}\Phi} = -\vartheta \mathbf{X}^T \frac{\partial \Sigma^{-1}}{\partial \Phi} \mathbf{r}.$$

$$Q_{\Phi\Phi} = -\frac{1}{2}\text{tr} \left[\Sigma^{-1} \left(\frac{\partial \Sigma}{\partial \Phi} \Sigma^{-1} \frac{\partial \Sigma}{\partial \Phi} - \frac{\partial^2 \Sigma}{\partial \Phi \partial \Phi^T} \right) \right] + \\ + \frac{1}{2} \left\{ \vartheta \mathbf{r}^T \left[\Sigma^{-1} \left(\frac{\partial^2 \Sigma}{\partial \Phi \partial \Phi^T} - \frac{\partial \Sigma}{\partial \Phi} \Sigma^{-1} \frac{\partial \Sigma}{\partial \Phi} - \frac{\partial \Sigma}{\partial \Phi} \Sigma^{-1} \frac{\partial \Sigma}{\partial \Phi} \right) \Sigma^{-1} \right] \mathbf{r} \right\}.$$

A.2.2 Matriz Δ_ω obtida a partir de $Q(\theta, \omega | \hat{\theta})$

Como tem-se

$$\Delta_\omega = \frac{\partial^2 Q(\theta, \omega | \hat{\theta})}{\partial \theta \partial \omega^T} = (\Delta_{\beta(\omega)}, \Delta_{\Phi(\omega)})^T$$

com

$$\Delta_{\beta(\omega)} = \frac{\partial^2 Q(\theta, \omega | \hat{\theta})}{\partial \beta \partial \omega^T} \text{ e} \\ \Delta_{\beta(\Phi)} = \frac{\partial^2 Q(\theta, \omega | \hat{\theta})}{\partial \Phi \partial \omega^T}$$

então

$$\Delta_{\beta(\omega)} = \vartheta \omega \mathbf{X}^T \Sigma^{-1} - 2\vartheta' \omega X^T \Sigma^{-1} (Y_\omega - X\beta) (Y_\omega - X\beta)^T \Sigma^{-1} \text{ e}$$

$$\Delta_{\beta(\Phi)} = \vartheta(\omega) (\mathbf{Y}_\omega - \mathbf{X}\beta)^T \Sigma^{-1} - \frac{\vartheta'(\omega)}{2} (\mathbf{Y}_\omega - \mathbf{X}\beta)^T \Sigma^{-1} \frac{\partial \Sigma}{\partial \varphi_j} \Sigma^{-1} (\mathbf{Y}_\omega - \mathbf{X}\beta) (\mathbf{Y}_\omega - \mathbf{X}\beta)^T \Sigma^{-1}$$

$$\text{sendo } \vartheta'(\omega) = \frac{v+n}{(v+\delta(\omega))^2}.$$

A.3 Matriz de informação esperada de Fisher

Pode-se escrever:

$$\mathbf{F} = \mathbf{F}(\theta) = \text{E}(UU^T), \text{ onde}$$

$$\mathbf{F} = \begin{pmatrix} F_{\beta\beta} & F_{\beta\varphi} \\ F_{\varphi\beta} & F_{\varphi\varphi} \end{pmatrix}.$$

Para determinar $F_{\beta\beta}$ tem-se que:

$$U_{\beta} = U(\beta) = W_g(\delta) \frac{\partial \delta}{\partial \beta} \text{ onde } W_g(\delta) = \frac{g'(\delta)}{g(\delta)}.$$

Porém,

$$\frac{\partial \delta}{\partial \beta} = -2\mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{r};$$

$$U_{\beta} = -2W_g(\delta) \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{r} \text{ ou}$$

$$U_{\beta} = -2W_g(\delta) \mathbf{X}^T \boldsymbol{\Sigma}^{-1/2} \mathbf{Z}, \text{ onde } \mathbf{Z} = \boldsymbol{\Sigma}^{-1/2} \mathbf{r} \text{ e } \mathbf{r} = (\mathbf{Y} - \mathbf{X}\beta).$$

Considere-se que $\mathbf{Z} \sim t_n(\mathbf{0}, \mathbf{I}_n, v)$, onde \mathbf{I}_n é a matriz indentidade e v é o grau de liberdade.

Sendo assim, pode-se escrever:

$$U_{\beta} U_{\beta}^T = 4W_g^2(\delta) \mathbf{X}^T \boldsymbol{\Sigma}^{-1/2} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1/2} \mathbf{X}$$

Porém,

$$\mathbb{E} \{ W_g^2(\delta) \mathbf{Z} \mathbf{Z}^T \} = \frac{a_g}{n} \mathbf{I}_n;$$

então:

$$F_{\beta\beta} = \frac{4a_g}{n} \mathbf{X}^T \boldsymbol{\Sigma}^{-1} \mathbf{X}$$

Para determinar $F_{\beta\varphi}$, tem-se:

$$U_{\varphi} = U(\varphi) = -\frac{1}{2} \frac{\partial}{\partial \varphi} (\log |\boldsymbol{\Sigma}|) + W_g(\delta) \frac{\partial \delta}{\partial \varphi},$$

como:

$$\frac{\partial}{\partial \varphi} (\log |\boldsymbol{\Sigma}|) = \frac{\partial \text{vec}^T(\boldsymbol{\Sigma})}{\partial \varphi} \text{vec}(\boldsymbol{\Sigma}^{-1}) \text{ e}$$

$$\frac{\partial \delta}{\partial \varphi} = \frac{\partial \text{vec}^T(\boldsymbol{\Sigma})}{\partial \varphi} \text{vec}(\boldsymbol{\Sigma}^{-1} \mathbf{r} \mathbf{r}^T \boldsymbol{\Sigma}^{-1}),$$

então:

$$U_{\boldsymbol{\varphi}} = U(\boldsymbol{\varphi}) = -\frac{1}{2} \frac{\partial \text{vec}^T(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}} \text{vec}(\boldsymbol{\Sigma}^{-1}) - W_g(\delta) \frac{\partial \text{vec}^T(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}} \text{vec}(\boldsymbol{\Sigma}^{-1} \mathbf{r} \mathbf{r}^T \boldsymbol{\Sigma}^{-1}).$$

Logo,

$$U_{\boldsymbol{\beta}} U_{\boldsymbol{\varphi}}^T = -2W_g(\delta) \mathbf{X}^T \boldsymbol{\Sigma}^{-1/2} \mathbf{Z}$$

$$\left\{ -\frac{1}{2} \text{vec}^T(\boldsymbol{\Sigma}^{-1}) \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}^T} - W_g(\delta) \text{vec}^T(\boldsymbol{\Sigma}^{-1/2} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1/2}) \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}^T} \right\},$$

ou ainda

$$\begin{aligned} U_{\boldsymbol{\beta}} U_{\boldsymbol{\varphi}}^T &= W_g(\delta) \mathbf{X}^T \boldsymbol{\Sigma}^{-1/2} \mathbf{Z} \text{vec}^T(\boldsymbol{\Sigma}^{-1}) \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}^T} + \\ &\quad + 2W_g^2(\delta) \mathbf{X}^T \boldsymbol{\Sigma}^{-1/2} \mathbf{Z} \text{vec}^T(\boldsymbol{\Sigma}^{-1/2} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1/2}) \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}^T}. \end{aligned}$$

Porém é possível mostrar que:

- i) $E\{W_g(\delta) \mathbf{Z}\} = 0$;
- ii) $E\left\{W_g^2(\delta) \boldsymbol{\Sigma}^{-1/2} \mathbf{Z} \text{vec}^T(\boldsymbol{\Sigma}^{-1/2} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1/2})\right\} = 0$.

Então obtêm-se que

$$F_{\boldsymbol{\beta} \boldsymbol{\varphi}} = E\left\{U_{\boldsymbol{\beta}} U_{\boldsymbol{\varphi}}^T\right\} = 0.$$

Agora para $F_{\boldsymbol{\varphi} \boldsymbol{\varphi}}$, tem-se:

$$\begin{aligned} U_{\boldsymbol{\varphi}} U_{\boldsymbol{\varphi}}^T &= \left[-\frac{1}{2} \frac{\partial \text{vec}^T(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}} \text{vec}(\boldsymbol{\Sigma}^{-1}) - W_g(\delta) \frac{\partial \text{vec}^T(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}} \text{vec}(\boldsymbol{\Sigma}^{-1/2} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1/2}) \right] \times \\ &\quad \times \left[-\frac{1}{2} \text{vec}^T(\boldsymbol{\Sigma}^{-1}) \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}^T} - W_g(\delta) \text{vec}(\boldsymbol{\Sigma}^{-1}) \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}^T} \right]; \\ U_{\boldsymbol{\varphi}} U_{\boldsymbol{\varphi}}^T &= \frac{1}{4} \frac{\partial \text{vec}^T(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}} \text{vec}(\boldsymbol{\Sigma}^{-1}) \text{vec}(\boldsymbol{\Sigma}^{-1}) \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}^T} + \\ &\quad + \frac{1}{2} W_g(\delta) \frac{\partial \text{vec}^T(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}} \text{vec}(\boldsymbol{\Sigma}^{-1}) \text{vec}^T(\boldsymbol{\Sigma}^{-1/2} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1/2}) \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}^T} + \\ &\quad + \frac{1}{2} W_g(\delta) \frac{\partial \text{vec}^T(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}} \text{vec}(\boldsymbol{\Sigma}^{-1/2} \mathbf{Z} \mathbf{Z}^T \boldsymbol{\Sigma}^{-1/2}) \text{vec}^T(\boldsymbol{\Sigma}^{-1}) \frac{\partial \text{vec}(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}^T} + \end{aligned}$$

$$+W_g^2(\delta) \frac{\partial \text{vec}^T(\Sigma)}{\partial \varphi} \text{vec}(\Sigma^{-1/2} \mathbf{Z} \mathbf{Z}^T \Sigma^{-1}) \text{vec}^T(\Sigma^{-1/2} \mathbf{Z} \mathbf{Z}^T \Sigma^{-1/2}) \frac{\partial \text{vec}(\Sigma)}{\partial \varphi^T}.$$

Denomine-se $U_\varphi U_\varphi^T$ por:

$U_\varphi U_\varphi^T = T_1 + T_2 + T_3 + T_4$, então:

$$\mathbb{E}(U_\varphi U_\varphi^T) = \mathbb{E}(T_1) + \mathbb{E}(T_2) + \mathbb{E}(T_3) + \mathbb{E}(T_4).$$

Porém,

$$\mathbb{E}\{W_g(\delta) \mathbf{Z} \mathbf{Z}^T\} = \frac{b_g}{n} \mathbf{I}_n, \text{ de forma que } b_g = -\frac{n}{2},$$

$$\text{então } \mathbb{E}\{W_g(\delta) \mathbf{Z} \mathbf{Z}^T\} = -\frac{1}{2} \mathbf{I}_n.$$

$$\mathbb{E}(T_2) = -\frac{1}{4} \frac{\partial \text{vec}^T(\Sigma)}{\partial \varphi} \text{vec}(\Sigma^{-1}) \text{vec}^T(\Sigma^{-1}) \frac{\partial \text{vec}(\Sigma)}{\partial \varphi^T};$$

$$\mathbb{E}(T_3) = -\frac{1}{4} \frac{\partial \text{vec}^T(\Sigma)}{\partial \varphi} \text{vec}(\Sigma^{-1}) \text{vec}^T(\Sigma^{-1}) \frac{\partial \text{vec}(\Sigma)}{\partial \varphi^T}.$$

No entanto,

$$\text{vec}(\Sigma^{-1/2} \mathbf{Z} \mathbf{Z}^T \Sigma^{-1/2}) = (\mathbf{A} \otimes \mathbf{A}) \text{vec}(\mathbf{Z} \mathbf{Z}^T) = (\mathbf{A} \otimes \mathbf{A})(\mathbf{Z} \otimes \mathbf{Z}),$$

$$\text{de onde } \mathbf{A} = \Sigma^{-1/2}.$$

Logo,

$$\text{vec}(\mathbf{A} \mathbf{Z} \mathbf{Z}^T \mathbf{A}) \text{vec}^T(\mathbf{A} \mathbf{Z} \mathbf{Z}^T \mathbf{A}) = (\mathbf{A} \otimes \mathbf{A})(\mathbf{Z} \otimes \mathbf{Z})(\mathbf{Z}^T \otimes \mathbf{Z}^T)(\mathbf{A} \otimes \mathbf{A});$$

$$\text{vec}(\mathbf{A} \mathbf{Z} \mathbf{Z}^T \mathbf{A}) \text{vec}^T(\mathbf{A} \mathbf{Z} \mathbf{Z}^T \mathbf{A}) = (\mathbf{A} \otimes \mathbf{A}) [\mathbf{Z} \mathbf{Z}^T \otimes \mathbf{Z} \mathbf{Z}^T] (\mathbf{A} \otimes \mathbf{A}).$$

Porém,

$$\mathbb{E}(W_g^2(\delta) [\mathbf{Z} \mathbf{Z}^T \otimes \mathbf{Z} \mathbf{Z}^T]) = \frac{C_g}{n(n+2)} [\mathbf{I}_{n^2} + K_n + \text{vec}(\mathbf{I}_n) \text{vec}^T(\mathbf{I}_n)] \text{ e}$$

$$C_g = \frac{n(n+2)(v+n)}{4(v+n+2)}$$

$$\text{E}(T_4) = \frac{\partial \text{vec}^T(\Sigma)}{\partial \varphi} \left\{ (\mathbf{A} \otimes \mathbf{A}) [\mathbf{I}_{n^2} + K_n + \text{vec}(\mathbf{I}_n) \text{vec}^T(\mathbf{I}_n)] (\mathbf{A} \otimes \mathbf{A}) \frac{\partial \text{vec}(\Sigma)}{\partial \varphi^T} \frac{C_g}{n(n+2)} \right\};$$

$$\begin{aligned} \text{E}(T_4) &= \frac{C_g}{n(n+2)} \left\{ \frac{\partial \text{vec}^T(\Sigma)}{\partial \varphi} [(\Sigma^{-1} \otimes \Sigma^{-1}) + \right. \\ &\quad \left. (\mathbf{A} \otimes \mathbf{A}) K_n (\mathbf{A} \otimes \mathbf{A}) + (\mathbf{A} \otimes \mathbf{A}) \text{vec}(\mathbf{I}_n) \text{vec}^T(\mathbf{I}_n) (\mathbf{A} \otimes \mathbf{A})] \frac{\partial \text{vec}(\Sigma)}{\partial \varphi^T} \right\}; \end{aligned}$$

tal que, K_n é uma matriz de comutacão, de forma que $K_n \text{vec}(\mathbf{A}) = \text{vec}(\mathbf{A}^T)$.

Contudo, como $K_n(\mathbf{A} \otimes \mathbf{A}) = (\mathbf{A} \otimes \mathbf{A})K_n$, então:

$$\begin{aligned} (\Sigma^{-1} \otimes \Sigma^{-1}) + (\mathbf{A} \otimes \mathbf{A}) K_n (\mathbf{A} \otimes \mathbf{A}) &= \\ &= (\Sigma^{-1} \otimes \Sigma^{-1}) + (\mathbf{A} \otimes \mathbf{A})(\mathbf{A} \otimes \mathbf{A}) K_n = (\Sigma^{-1} \otimes \Sigma^{-1}) + (\Sigma^{-1} \otimes \Sigma^{-1}) K_n; \\ &= (\Sigma^{-1} \otimes \Sigma^{-1})(\mathbf{I}_{n^2} + K_n). \end{aligned}$$

Sendo assim,

$$\text{E}(T_4) = \frac{C_g}{n(n+2)} \frac{\partial \text{vec}^T(\Sigma)}{\partial \varphi} \{ (\Sigma^{-1} \otimes \Sigma^{-1})(\mathbf{I}_n + K_n) + \text{vec}(\Sigma^{-1}) \text{vec}^T(\Sigma^{-1}) \} \frac{\partial \text{vec}(\Sigma)}{\partial \varphi^T}.$$

Consequentemente

$$\begin{aligned} \text{E}(U_\varphi U_\varphi^T) &= -\frac{1}{4} \frac{\partial \text{vec}^T(\Sigma)}{\partial \varphi} \text{vec}(\Sigma^{-1}) \text{vec}^T(\Sigma^{-1}) \frac{\partial \text{vec}(\Sigma)}{\partial \varphi} \\ &\quad + \frac{C_g}{n(n+2)} \frac{\partial \text{vec}(\Sigma)}{\partial \varphi} \{ (\Sigma^{-1} \otimes \Sigma^{-1})(\mathbf{I}_n + K_n) + \text{vec}(\Sigma^{-1}) \text{vec}^T(\Sigma^{-1}) \} \frac{\partial \text{vec}(\Sigma)}{\partial \varphi^T}. \\ E(U_\varphi U_\varphi^T) &= \left(-\frac{1}{4} + \frac{C_g}{n(n+2)} \right) \frac{\partial \text{vec}^T(\Sigma)}{\partial \varphi} \{ \text{vec}(\Sigma^{-1}) \text{vec}^T(\Sigma^{-1}) \} \frac{\partial \text{vec}(\Sigma)}{\partial \varphi} + \\ &\quad + \frac{C_g}{n(n+2)} \frac{\partial \text{vec}^T(\Sigma)}{\partial \varphi} \{ (\Sigma^{-1} \otimes \Sigma^{-1})(\mathbf{I}_n + K_n) \} \frac{\partial \text{vec}(\Sigma)}{\partial \varphi^T}. \end{aligned}$$

Seja $\Sigma = \varphi_1 \mathbf{I}_n + \varphi_2 \text{vec}(R)$, tem-se que:

$$(a) \text{vec}(\Sigma) = \varphi_1 \text{vec}(\mathbf{I}_n) + \varphi_2 \text{vec}(R);$$

$$(b) \frac{\partial \text{vec}(\Sigma)}{\partial \varphi^T} = \frac{\partial(\varphi_1 \text{vec}(\mathbf{I}_n))}{\partial \varphi^T} + \frac{\partial(\varphi_2 \text{vec}(R))}{\partial \varphi^T};$$

$$(c) \frac{\partial(\varphi_1 \text{vec}(\mathbf{I}_n))}{\partial \varphi^T} = \left(\frac{\partial(\varphi_1 \text{vec}(\mathbf{I}_n))}{\partial \varphi_1}, \frac{\partial(\varphi_1 \text{vec}(\mathbf{I}_n))}{\partial \varphi_2}, \frac{\partial(\varphi_1 \text{vec}(\mathbf{I}_n))}{\partial \varphi_3} \right)$$

$$= (\text{vec}(\mathbf{I}_n), 0, 0), n^2 \times 3;$$

$$(d) \frac{\partial(\varphi_2 \text{vec}(R))}{\partial \varphi^T} = \left(\frac{\partial(\varphi_2 \text{vec}(R))}{\partial \varphi_1}, \frac{\partial(\varphi_2 \text{vec}(R))}{\partial \varphi_2}, \frac{\partial(\varphi_2 \text{vec}(R))}{\partial \varphi_3} \right)$$

$$= \left(0, \text{vec}(R), \frac{\partial(\varphi_2 \text{vec}(R))}{\partial \varphi_3} \right);$$

$$= \left(0, \text{vec}(R), \varphi_2 \text{vec} \left(\frac{\partial R}{\partial \varphi_3} \right) \right), n^2 \times 3;$$

$$= (0, \text{vec}(R), \varphi_2 \text{vec}(R')), R' = \frac{\partial R}{\partial \varphi_3}.$$

Logo:

$$\frac{\partial \text{vec}(\Sigma)}{\partial \varphi^T} = ((\text{vec}(\mathbf{I}_n), 0, 0) + (0, \text{vec}(R), \varphi_2 \text{vec}(R')))$$

$$(\text{vec}(\mathbf{I}_n), \text{vec}(R), \varphi_2 \text{vec}(R')), n^2 \times 3.$$

Tem-se ainda que:

$$(a) \text{vec}^T(\varphi_1(\mathbf{I}_n) + \varphi_2 \text{vec}(R)) = \varphi_1 \text{vec}^T(\mathbf{I}_n) + \varphi_2 \text{vec}^T(R), 1 \times n^2;$$

$$(b) \frac{\partial \text{vec}^T(\Sigma)}{\partial \varphi} = \frac{\partial \varphi_1 \text{vec}^T(\mathbf{I}_n)}{\partial \varphi} + \frac{\partial \varphi_2 \text{vec}^T(R)}{\partial \varphi};$$

$$(c) \frac{\partial \varphi_1 \text{vec}^T(\mathbf{I}_n)}{\partial \varphi} = \begin{bmatrix} \frac{\partial \varphi_1 \text{vec}^T(\mathbf{I}_n)}{\partial \varphi_1} \\ \frac{\partial \varphi_1 \text{vec}^T(\mathbf{I}_n)}{\partial \varphi_2} \\ \frac{\partial \varphi_1 \text{vec}^T(\mathbf{I}_n)}{\partial \varphi_3} \end{bmatrix} = \begin{bmatrix} \text{vec}^T(\mathbf{I}_n) \\ 0 \\ 0 \end{bmatrix}, 3 \times n^2;$$

$$(d) \frac{\partial \varphi_2 \text{vec}^T(R)}{\partial \varphi} = \begin{bmatrix} \frac{\partial \varphi_2 \text{vec}^T(R)}{\partial \varphi_1} \\ \frac{\partial \varphi_2 \text{vec}^T(R)}{\partial \varphi_2} \\ \frac{\partial \varphi_2 \text{vec}^T(R)}{\partial \varphi_3} \end{bmatrix} = \begin{bmatrix} 0 \\ \text{vec}^T(R') \\ \varphi_2 \text{vec}^T(R') \end{bmatrix};$$

$$(e) \frac{\partial vec^T(\boldsymbol{\Sigma})}{\partial \boldsymbol{\varphi}} = \begin{bmatrix} vec^T(\boldsymbol{I}_n) \\ vec^T(R) \\ \varphi_2 vec^T(R) \end{bmatrix}, 3 \times n^2.$$

APÊNDICE B – Roteiros para análise no programa R e módulo geoestatístico geoR

```
#####
##          Algoritmo EM          ##
##          Distribuição t-student modelo exponencial      ##
##          Análise sem covariáveis      ##
#####

require(geoR)
P<-read.geodata("C:/Desktop/potássio.txt",h=FALSE,coords.col=2:3,data.col=4)
Y<-P$data
summary(Y)
sd(Y)
CV<- sd(Y)*100/mean(Y)
CV
hist(Y)
boxplot(Y)
g1<3
ks.test(Y,"pt",g1)
## Ajuste do modelo pelo algoritmo EM
n<-length(P$data)
Y<-matrix(P$data,n,1)
X<-as.matrix(rep(1,n))
coords<-P$coords
D<-as.matrix(dist(coords,method="euclidean",diag=T,upper=T))
## Chute inicial para os parâmetros beta e phi
## modelo exponencial
beta<-solve(t(X)%*%X)%*%t(X)%*%Z
fi1<0
fi2<-1
fi3<-2.5
teta<-cbind(beta,fi1,fi2,fi3)

t=0
g<-function(num) {
nr<-matrix(0,num,5)
while( t < num ) {
t = t+1
E<-cov.spatial(D,cov.model="exp", cov.par=c(fi2,fi3))
IE<-solve(E)
beta.<-solve(t(X)%*%IE%*%X)%*%t(X)%*%IE%*%Y
r<-(Y-X)%*%beta..
u<-(t(r)%*%IE%*%r)
v<-as.numeric((g1+n)/(g1+u))
vi<-as.numeric((g1+n)/(g1+u)^2)
I<-(diag(1,nr=n))
R<-1/fi2*(E-fi1*I)
F<-fi2*R%*%D/(fi3^2)
H<-fi2*R%*%(D/fi3^3)%*%((D/fi3)-2)
S1<-v*(t(r)%*%IE%*%I%*%IE%*%r)
S2<-v*(t(r)%*%IE%*%R%*%IE%*%r)
S<-cbind(S1,S2)
a11<-sum(diag(IE%*%I%*%IE))
a12<-sum(diag(IE%*%I%*%IE%*%R))
a21<-sum(diag(IE%*%R%*%IE))
a22<-sum(diag(IE%*%R%*%IE%*%R))
A<-matrix(c(a11,a21,a12,a22),2,2)
Fi<-S%*%solve(A)
fi11<-Fi[1,1]
fi22<-Fi[1,2]
Qfi3<-1/2*(sum(diag(IE%*%F)))+1/2*v*(t(r)%*%IE%*%F%*%IE%*%r)
```

```

Qffi3<-1/2*(sum(diag(IE%*%(H-F%*%IE%*%F))))+1/2*v*(t(r)%*%IE%*%(H-2*F%*%IE%*%F)%*%IE%*%r)
fi33<-as.numeric(fi3-(solve(Qffi3)%*%Qffi3))
tetax<-cbind(beta,fii,fi2,fi3)
tetai<-cbind(beta.,fi11,fi22,fi33)
l<-(sum((tetax-teta1)^2)^0.5))/((sum(tetax^2))^0.5)
nr[t,]=c(teta1,1)
beta<-beta.
fi1<-fi11
fi2<-fi22
fi3<-fi33
}
return(nr)
}
nr<-g(100)
nr
min(nr[,5])
EM<-matrix(nr[1,])
beta<-EM[1,]
fi1<-EM[2,]
fi2<-EM[3,]
fi3<-EM[4,]

cov<-cov.spatial(D,cov.model="exp", cov.par=c(fi2,fi3))
plot(D,cov,col='blue',pch="*")

## Validação Cruzada
C<-cov

kc<-function(n) {
kcx<-matrix(0,n,3)
i=0
while( i < n ) {
i = i+1
kcx[i,1]<- Z[i,1]

K<-function (i){
Kx<-matrix(0,(n-1),(n-1))
if (i==1){
Kx<-as.matrix(C[(i+1):n,(i+1):n])
return(Kx)
}
else if (i==n){
Kx<-as.matrix(C[(1:(n-1)),(1:(n-1))])
return(Kx)
}
else if (i==2){
K1<-as.matrix(C[1,1])
K2<-as.matrix(C[(i+1):n,1])
K3<-as.matrix(t(C[1,(i+1):n]))
K4<-as.matrix(C[(i+1):(i+1),n])
Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
return(Kx)
}
else if (i==(n-1)){
K1<-as.matrix(C[1:(i-1),1:(i-1)])
K2<-as.matrix(t(C[n,1:(i-1)]))
K3<-as.matrix(C[1:(i-1),n])
K4<-as.matrix(C[n,n])
Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
return(Kx)
}
else {
K1<-as.matrix(C[(1:(i-1)),(1:(i-1))])
K2<-as.matrix(C[(i+1):n,1:(i-1)])
K3<-as.matrix(C[1:(i-1),(i+1):n])
K4<-as.matrix(C[(i+1):n,(i+1):n])
Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
return(Kx)
}
}
K<-K(i)

L<-function (i){
Lx<-matrix(0,n,1)
if (i==1){
L<-as.matrix(C[i,(1+i):n])
L<-rbind(L,1)
}
}
K<-K(i)

```

```

Lx<-L
return(Lx)
}
else if (i==n){
L<-as.matrix(C[n,1:(n-1)])
L<-rbind(L,L)
Lx<-L
return(Lx)
}
else {
L1<-as.matrix(C[i,1:(i-1)])
L2<-as.matrix(C[i,(i+1):n])
L<-rbind(L1,L2,1)
Lx<-L
return(Lx)
}
}
L.<-L(i)

Tk<-as.matrix(rep(1,n-1),n-1,1)
K<-cbind(K,T)
T<-t(rbind(T,0))
K<-rbind(K,T)
Lamb<-solve(K)%%L.
Lambda<-Lamb[1:(n-1),]
Lambda1<-as.matrix(Lambda,(n-1),1)

Z.<-function (i){
Z.x<-matrix(0,n,1)
if (i==1){
Z.x<-as.matrix(Z[(i+1):n,1])
return(Z.x)
}
else if (i==n){
Z.x<-as.matrix(Z[1:(i-1),1])
return(Z.x)
}
else {
Z.1<-as.matrix(Z[1:(i-1),1])
Z.2<-as.matrix(Z[(i+1):n,1])
Z.x<-rbind(Z.1,Z.2)
return(Z.x)
}
}
Z.<-Z.(i)
Z1<-t(Lambda1)%%Z.
kcx[i,2]<- Z1
sd<-(sum(t(Lambda1))%%L.[(1:(n-1)),1])+Lamb[n,])^0.5
kcx[i,3]<- sd
}
return(kcx)
}
KC<-kc(n)
KC
DF<-(KC[,1]-KC[,2])
s<-summary(DF)
s<-t(as.matrix(s,1,6))
Sdf=as.matrix(sd(DF))
erro<-cbind(s,Sdf)
DF.<-(KC[,1]-KC[,2])/(KC[,3])
s.<-summary(DF.)
s.<-t(as.matrix(s.,1,6))
Sdf.=as.matrix(sd(DF.))
erro_std<-cbind(s.,Sdf.)
summary<-rbind(erro,erro_std)
summary
EA<-sum(abs(DF))/n
EA
## AKAIKE
r=3
AIC=-2*Lv+2*r
AIC
#####
##          Análise de diagnóstico usual      ##
##      l(teta)=K(v)-1/2 log|Sigma|-1/2(v+n) log(1+delta/v)  ##
##          Distribuição t-student usual        ##
#####

```

```

tetabeta=cbind(beta,fi1,fi2,fi3)

##Calculando a matriz de covariância com as estimativas dos
##parâmetros phi(f1, f2, f3)
Es=cov.spatial(D,cov.model="exp", cov.par=c(1,2.5))
Covariância=g1/(g1-2)*Es
IC<-solve(Covariância)

##Derivadas
D1<-diag(1,nr=n)
D2<-1/fi2*(Covariância-fi1*D1)
D3<-fi2*((D2%*%D)/(fi3^2))
D11<-matrix(0,n,n)
D22<-matrix(0,n,n)
D33<-fi2*((D2%*%D)/(fi3^3))%*%((D/fi3)-2)
D23<-(D2%*%D)/(fi3^2)

##Matriz delta beta e delta phi
r<-(Y-X%*%beta)
delta<-(t(r)%*%IC%*%(r))
v<-as.numeric((gl+n)/(gl+delta))
vi<-as.numeric((gl+n)/(gl+delta))^2
Db<-v*t(X)%*%IC -2*vi*t(X)%*%IC%*%r%*%t(r)%*%IC
Df11<-v*t(r)%*%IC%*%D1-vi*t(r)%*%IC%*%D1%*%IC%*%r%*%t(r)%*%IC
Df12<-v*t(r)%*%IC%*%D2-vi*t(r)%*%IC%*%D2%*%IC%*%r%*%t(r)%*%IC
Df13<-v*t(r)%*%IC%*%D3-vi*t(r)%*%IC%*%D3%*%IC%*%r%*%t(r)%*%IC

##Matriz delta
Df<-rbind(Df11,Df12,Df13)
Delta<-rbind(Db,Df)

##Matriz de informação observada (L)
##L com respeito a beta/beta
Lbb<-v*t(X)%*%X-2*vi*t(X)%*%IC%*%r%*%t(X)%*%IC%*%r

##L com respeito a beta/phi
Lbf11<-v*t(X)%*%IC%*%D1%*%IC%*%r+vi*t(X)%*%IC%*%r%*%t(r)
    %*%IC%*%D1%*%IC%*%r
Lbf12<-v*t(X)%*%IC%*%D2%*%IC%*%r+vi*t(X)%*%IC%*%r%*%t(r)
    %*%IC%*%D2%*%IC%*%r
Lbf13<-v*t(X)%*%IC%*%D3%*%IC%*%r+vi*t(X)%*%IC%*%r%*%t(r)
    %*%IC%*%D3%*%IC%*%r
Lbf<-cbind(Lbf11,Lbf12,Lbf13)
Lfb<-t(Lbf)

##L com respeito a phii/phij
Lf1f1<-1/2*(sum(diag(IC%*%(D11-D1%*%IC%*%D1))))+
1/2*v*(t(r)%*%IC%*%(D11-D1%*%IC%*%D1-D1%*%IC%*%D1)%*%IC%*%r)+ 
1/2*vi*(t(r)%*%IC%*%D1%*%IC%*%r)*(t(r)%*%IC%*%D1%*%IC%*%r)

Lf2f2<-1/2*(sum(diag(IC%*%(D11-D2%*%IC%*%D2))))+
1/2*v*(t(r)%*%IC%*%(D22-D2%*%IC%*%D2-D2%*%IC%*%D2)%*%IC%*%r)+ 
1/2*vi*(t(r)%*%IC%*%D2%*%IC%*%r)*(t(r)%*%IC%*%D2%*%IC%*%r)

Lf3f3<-1/2*(sum(diag(IC%*%(D11-D3%*%IC%*%D3))))+
1/2*v*(t(r)%*%IC%*%(D33-D3%*%IC%*%D3-D3%*%IC%*%D3)%*%IC%*%r)+ 
1/2*vi*(t(r)%*%IC%*%D3%*%IC%*%r)*(t(r)%*%IC%*%D3%*%IC%*%r)

Lf1f2<-1/2*(sum(diag(IC%*%(D11-D2%*%IC%*%D1))))+
1/2*v*(t(r)%*%IC%*%(D11-D1%*%IC%*%D1-D2%*%IC%*%D1)%*%IC%*%r)+ 
1/2*vi*(t(r)%*%IC%*%D1%*%IC%*%r)*(t(r)%*%IC%*%D2%*%IC%*%r)

Lf1f3<-1/2*(sum(diag(IC%*%(D11-D3%*%IC%*%D1))))+
1/2*v*(t(r)%*%IC%*%(D11-D1%*%IC%*%D3-D3%*%IC%*%D1)%*%IC%*%r)+ 
1/2*vi*(t(r)%*%IC%*%D1%*%IC%*%r)*(t(r)%*%IC%*%D3%*%IC%*%r)

Lf2f3<-1/2*(sum(diag(IC%*%(D23-D3%*%IC%*%D2))))+
1/2*v*(t(r)%*%IC%*%(D23-D2%*%IC%*%D3-D3%*%IC%*%D2)%*%IC%*%r)+ 
1/2*vi*(t(r)%*%IC%*%D2%*%IC%*%r)*(t(r)%*%IC%*%D3%*%IC%*%r)

Lff<-rbind(cbind(Lf1f1,Lf1f2,Lf1f3), cbind(Lf1f2,Lf2f2,Lf2f3),
            cbind(Lf1f3,Lf2f3,Lf3f3))

L1<-cbind(Lbb,Lbf)
L2<-cbind(Lfb,Lff)
L<-rbind(L1,L2)

```

```

##Erro padrão do estimador de máxima verossimilhança L

n=length(Y)
ag=n/4*((v+n)/(v+n+2))
bg11=sum(diag(IC%*%D1))*sum(diag(IC%*%D1))
bg12=sum(diag(IC%*%D1))*sum(diag(IC%*%D2))
bg13=sum(diag(IC%*%D1))*sum(diag(IC%*%D3))
bg22=sum(diag(IC%*%D2))*sum(diag(IC%*%D2))
bg23=sum(diag(IC%*%D2))*sum(diag(IC%*%D3))
bg33=sum(diag(IC%*%D3))*sum(diag(IC%*%D3))
fg=((n*(n+2))/4)*((v+n)/(v+n+2))
Kb<-4*ag/n*t(X)%*%IC%*%X
Kf1f1<-bg11/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*(
  (sum(diag(IC%*%D1%*%IC%*%D1))))
Kf1f2<-bg12/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*(
  (sum(diag(IC%*%D1%*%IC%*%D2))))
Kf1f3<-bg13/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*(
  (sum(diag(IC%*%D1%*%IC%*%D3))))
Kf2f2<-bg22/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*(
  (sum(diag(IC%*%D2%*%IC%*%D2))))
Kf2f3<-bg23/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*(
  (sum(diag(IC%*%D2%*%IC%*%D3))))
Kf3f3<-bg33/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*(
  (sum(diag(IC%*%D3%*%IC%*%D3))))
Kf<-rbind(cbind(Kf1f1,Kf1f2,Kf1f3),cbind(Kf1f2,Kf2f2,Kf2f3),
           cbind(Kf1f3,Kf2f3,Kf3f3))

K1<-cbind(Kb,matrix(0,nr=1,nc=3))
K2<-cbind(matrix(0,nr=3,nc=1),Kf)
Kt<-rbind(K1,K2)
IKt<-solve(Kt)
DIKt<-diag(IKt)
DP<-sqrt(DIKt)
DP<-rbind(DP)
T<-rbind(teta,DP)
T

#####
##### INFLUÊNCIA LOCAL USUAL #####
#### Influência no afastamento da verossimilhança
B<-t(Delta)%*%solve(L)%*%Delta
Af<-(sum(diag(t(B)%*%B)))^0.5
Bd<-B/Af
Bd

##Gráfico de coeficientes Ci
Ci<-2*abs(diag(Bd))
plot(Ci, xlab="Ordem", ylab=expression(C[i]), ylim=c(0,1), pch=19)
abline(h=mean(Ci)+2*sd(Ci), col = "black", lty = 3)
identify(Ci,offset=0.3)

##Gráfico do autovalor |Lmax|
BA<-eigen(Bd,symmetric=TRUE, only.value=FALSE)
abs(BA$values)
Lmax<-BA$vector[,144]
Lmax<-abs(Lmax/sqrt(sum(Lmax^2)))
plot(Lmax, xlab="Ordem", ylab=expression(abs(L[max])),
      ylim=c(0,1), pch=19)
abline(h=mean(Lmax)+2*sd(Lmax), col = "black", lty = 3)
identify(Lmax,offset=0.3)

## Influência na matriz de covariância e no preditor linear
## Gráfico de diagnóstico 1[Ch]
h=10
dCf1=0

```

```

dCf2=gl/(gl-2)*exp(-(h/fi3))
dCf3=gl/(gl-2)*(h*fi2*exp(-(h/fi3)))/(fi3)^2
DCh=(cbind(0,dCf1,dCf2,dCf3))
DCh=matrix(DCh)
Ch=-t(Delta) %*% solve(L) %*% DCh
lCh= Ch/sum(diag(t(Ch)) %*% Ch)
plot(abs(lCh),xlab="Ordem",ylim=c(0,800),
     ylab=expression(l[Ch]),pch=19)
identify(lCh,offset=0.3)

## Gráfico de diagnóstico l[p]
bor=read.table("C:/Desktop/Bordas.txt")
plot(bor)
coords=P$coords
n=length(P$data)
s0<-coords[20,]
s0<-matrix(s0,1,2)
points(s0)
x0=1
dist_s0=apply(1:n,function(p)
  sqrt(((matrix(coords[,p])[,1])-s0[,1])^2+
    ((matrix(coords[,p])[,2])-s0[,2])^2))

C0=cov.spatial(dist_s0,cov.model="exp",cov.par=c(fi1+fi2,fi3))
cov_s0=gl/(gl-2)*C0
dCf1=0
dCf2=gl/(gl-2)*exp(-(dist_s0/fi3))
dCf3=gl/(gl-2)*(dist_s0*fi2*exp(-(dist_s0/fi3)))/(fi3)^2
pb=x0-t(X) %*% IC %*% r
pf1=(dCf1-t(CO)) %*% IC %*% D1) %*% IC %*% r
pf2=(dCf2-t(CO)) %*% IC %*% D2) %*% IC %*% r
pf3=(dCf3-t(CO)) %*% IC %*% D3) %*% IC %*% r
pf=cbind(pf1,pf2,pf3)
pf=matrix(pf)
pt=rbind(pb,pf)
p.=-t(Delta) %*% solve(L) %*% pt
Lp=abs(p./sqrt(sum(p.^2)))
plot(Lp,xlab="Ordem", ylab=expression(l[p]),
     ylim=c(0,1),pch=19)
identify(Lp,offset=0.3)

##Influência nos valores preditos
LbY=-1/2*v*t(X) %*% IC + v*i*t(X) %*% IC %*% r %*% t(r) %*% IC
Lf1Y=-v*t(r) %*% IC %*% D1 %*% IC + v*i*t(r) %*% IC %*% D1 %*% IC %*% (r) %*% t(r) %*% IC
Lf2Y=-v*t(r) %*% IC %*% D2 %*% IC + v*i*t(r) %*% IC %*% D2 %*% IC %*% (r) %*% t(r) %*% IC
Lf3Y=-v*t(r) %*% IC %*% D3 %*% IC + v*i*t(r) %*% IC %*% D3 %*% IC %*% (r) %*% t(r) %*% IC
LfY=rbind(Lf1Y,Lf2Y,Lf3Y)
LG1=X %*% solve(Lbb+Lbf) %*% solve(Lff) %*% Lfb
LG2=X %*% solve(Lbb+Lbf) %*% solve(Lff) %*% Lfb %*% (-Lbf) %*% solve(Lff) %*% LfY
LG=LG1+LG2
G=abs(diag(LG))
plot(G,xlab="Ordem",ylab=expression(LG),font.main=4, pch=19)
identify(G,offset=0.3)

#####
## Análise de diagnóstico Q-deslocamento
## Distribuição t-student modelo exponencial
#####
teta=cbind(beta,fi1,fi2,fi3)

##Calculando a matriz de covariância com as estimativas dos
##parâmetros phi(f1, f2, f3)
IC<-solve(Covariância)

##Primeira derivadas
D1<-diag(1,nr=n)
D2<-1/fi2*(Covariância-fi1*D1)
D3<-fi2*((D1%D)/(fi3^2))
D11<-matrix(0,n,n)
D22<-matrix(0,n,n)
D33<-fi2*((D1%D)/(fi3^3)) %*% ((D/fi3)-2)
D23<-(D1%D)/(fi3^2)

##Matriz delta beta e delta phi
r<-Y-X %*% beta
delta<-(t(r)) %*% IC %*% r
W<-as.numeric((gl+n)/(gl+delta))

```

```

Wl<-as.numeric((gl+n)/(gl+delta)^2)
Db<-W*t(X)%%IC -2*Wl*t(X)%%IC%%r%*%t(r)%%IC
Df11<-W*t(r)%%IC%%D1-Wl/2*t(r)%%IC%%D1%%r%*%t(r)%%IC
Df12<-W*t(r)%%IC%%D2-Wl/2*t(r)%%IC%%D2%%r%*%t(r)%%IC
Df13<-W*t(r)%%IC%%D3-Wl/2*t(r)%%IC%%D3%%r%*%t(r)%%IC

##Matriz delta
Df<-rbind(Df11,Df12,Df13)
Delta<-rbind(Db,Df)

##Matriz de informação observada (Q)
##Q com respeito a beta
Qbb<-W*t(X)%%IC%%X

##L com respeito a beta e phi
Qbfi1<-W*t(X)%%IC%%D1%%IC%%(Z-X%%beta)
Qbfi2<-W*t(X)%%IC%%D2%%IC%%(Z-X%%beta)
Qbfi3<-W*t(X)%%IC%%D3%%IC%%(Z-X%%beta)
Qbf<-cbind(Qbfi1,Qbfi2,Qbfi3)
Qfb<-t(Qbf)

##Q com respeito a phii e phij
Qf1f1<-1/2*(sum(diag(IC%%(D11-D1%%IC%%D1))))+W/2*(t(Z-X%%beta)
  %%IC%%(D11-D1%%IC%%D1-D1%%IC%%D1))%%IC%%(Z-X%%beta)

Qf2f2<-1/2*(sum(diag(IC%%(D22-D2%%IC%%D2))))+W/2*(t(Z-X%%beta)
  %%IC%%(D22-D2%%IC%%D2-D2%%IC%%D2))%%IC%%(Z-X%%beta)

Qf3f3<-1/2*(sum(diag(IC%%(D33-D3%%IC%%D3))))+W/2*(t(Z-X%%beta)
  %%IC%%(D33-D3%%IC%%D3-D3%%IC%%D3))%%IC%%(Z-X%%beta)

Qf1f2<-1/2*(sum(diag(IC%%(D11-D2%%IC%%D1))))+W/2*(t(Z-X%%beta)
  %%IC%%(D11-D2%%IC%%D1-D1%%IC%%D2))%%IC%%(Z-X%%beta)

Qf1f3<-1/2*(sum(diag(IC%%(D11-D3%%IC%%D1))))+W/2*(t(Z-X%%beta)
  %%IC%%(D11-D3%%IC%%D1-D1%%IC%%D3))%%IC%%(Z-X%%beta)

Qf2f3<-1/2*(sum(diag(IC%%(D23-D3%%IC%%D2))))+W/2*(t(Z-X%%beta)
  %%IC%%(D23-D3%%IC%%D2-D2%%IC%%D3))%%IC%%(Z-X%%beta)

Qff<-rbind(cbind(Qf1f1,Qf1f2,Qf1f3), cbind(Qf1f2,Qf2f2,Qf2f3),
            cbind(Qf1f3,Qf2f3,Qf3f3))

Q1<-cbind(Qbb,Qbf)
Q2<-cbind(Qfb,Qff)
Q<-rbind(Q1,Q2)

#####
##### INFLUÊNCIA LOCAL #####
#####

Bq<-t(Delta)%%solve(Q)%%Delta
Afq<-(sum(diag(t(Bq)%%Bq)))^0.5
Bdq<-Bq/Afq
Bdq

##Gráfico de coeficientes CQi
Cqi<-2*abs(diag(Bdq))
plot(Cqi, xlab="Ordem", ylab=expression(C[Qi]),
      ylim=c(0,1), pch=19)

abline(h=mean(Ci)+2*sd(Ci), col = "black", lty = 3)
identify(Cqi,offset=0.3)

##Gráfico do autovalor |Qmax|
BAq<-eigen(Bdq,symmetric=TRUE, only.value=FALSE)
abs(BAq$values)
Qmax<-BAq$vector[,144]
Qmax<-abs(Qmax/sqrt(sum(Qmax^2)))
plot(Qmax, xlab="Ordem", ylab=expression(abs(Q[max])),
      ylim=c(0,1), pch=19)

abline(h=mean(Lmax)+2*sd(Lmax), col = "black", lty = 3)
identify(Qmax,offset=0.3)

##Influência na matriz de covariância e no preditor linear
## Gráfico de diagnóstico na matriz de covariância 1C[Qh]
h=10
dCf1=0

```

```

dCf2=gl/(gl-2)*exp(-(h/fi3))
dCf3=gl/(gl-2)*(h*fi2*exp(-(h/fi3)))/(fi3)^2
DChq=t(cbind(0,dCf1,dCf2,dCf3))
DCqh=matrix(DChq)
Chq=-t(Delta) %*% solve(Q) %*% DChq
lChq=Chq/sum(diag(t(Chq)) %*% Chq)
plot(abs(lChq),xlab="Ordem",ylim=c(0,1),ylab=expression(Q[Ch]),pch=19)
identify(lChq,offset=0.3)

## Gráfico de diagnóstico nno preditor linear lp[Q]
plot(bor)
coords=P$coords
n=length(P$data)
s0<-coords[20,]
s0<-matrix(s0,1,2)
points(s0)
x0=1
dist_s0=sapply(1:n,function(p)
sqrt(((matrix(coords[p,]) [1,])-s0[,1])^2+
((matrix(coords[p,]) [2,])-s0[,2])^2))

dist_s0
C0=cov.spatial(dist_s0,cov.model="exp", cov.par=c(fi2,fi3))
cov_s0=gl/(gl-2)*C0
dCf1=0
dCf2=gl/(gl-2)*exp(-(dist_s0/fi3))
dCf3=gl/(gl-2)*(dist_s0*fi2*exp(-(dist_s0/fi3)))/(fi3)^2
pbq=x0-t(X) %*% IC %*% CO
pf1q=(dCf1-t(CO)) %*% IC %*% D1) %*% IC %*% (r)
pf2q=(dCf2-t(CO)) %*% IC %*% D2) %*% IC %*% (r)
pf3q=(dCf3-t(CO)) %*% IC %*% D3) %*% IC %*% (r)
pfq=cbind(pf1q,pf2q,pf3q)
pfq=matrix(pfq)
ptq=rbind(pbq,pfq)
pq.=-t(Delta) %*% solve(Q) %*% ptq
Qp=abs(pq./sqrt(sum(pq.^2)))
plot(Qp,xlab="Ordem", ylab=expression(l[pQ]),ylim=c(0,1),pch=19)
identify(Lp,offset=0.3)

##Influência nos valores preditos - LG[Q]
QbY=-1/2*W*t(X) %*% IC+W1*t(X) %*% IC %*% (r) %*% t(r) %*% IC
Qf1Y=-W*t(r) %*% IC %*% D1 %*% IC+W1*t(r) %*% IC %*% D1 %*% IC %*% (r) %*% t(r) %*% IC
Qf2Y=-W*t(r) %*% IC %*% D2 %*% IC+W1*t(r) %*% IC %*% D2 %*% IC %*% (r) %*% t(r) %*% IC
Qf3Y=-W*t(r) %*% IC %*% D3 %*% IC+W1*t(r) %*% IC %*% D3 %*% IC %*% (r) %*% t(r) %*% IC
QfY=rbind(Lf1Y,Lf2Y,Lf3Y)
QG1=X %*% solve(Qbb+Qbf) %*% solve(Qff) %*% Qfb) %*% QbY
QG2=X %*% solve(Qbb+Qbf) %*% solve(Qff) %*% Qfb) %*% (-Qbf) %*% solve(Qff) %*% QfY)
LGQ=QG1+QG2
LGQ=abs(diag(LGQ))
plot(LGQ,xlab="Ordem", ylab=expression(LG[Q]),font.main=4, pch=19)
identify(LGQ,offset=0.3)

#####
## Algoritmo EM
## Distribuição t-student modelo exponencial
## Análise com covariáveis
#####

require(geoR)
P=read.table("C:/Desktop/Análise Covariáveis/PCovariáveis.txt")
P$as.geodata(P,coords.col=2:3,data.col=4,covar.col=5:10,head=F)
Z<-P$data
summary(Z)
sd(Z)
CV<- sd(Z)*100/mean(Z)
CV
hist(Z)
boxplot(Z)
gl=3
ks.test(Z,pt,gl)

## Ajuste do modelo pelo algoritmo EM
n=length(P$data)
Z<-matrix(P$data,n,1)

## Chute inicial para os parâmetros beta e phi
## modelo exponencial

```

```

X=read.table("C:/Análise Covariáveis/Matriz X.txt", head=T)
#Matriz de ls e covariáveis
X=as.matrix(X)

beta0=mean(X[,1])
beta1=mean(X[,2])
beta2=mean(X[,3])
beta3=mean(X[,4])
beta4=mean(X[,5])
beta5=mean(X[,6])
beta6=mean(X[,7])
beta=<rbind(beta0,beta1,beta2,beta3,beta4,beta5,beta6)
fi1=0.1
fi2=0.05
fi3=140
coords<-matrix(P$coords,n,2)
D<-as.matrix(dist(coords,method="euclidean",diag=T,upper=T))

## Algoritmo EM
t=0
g<-function(num) {
nr<-matrix(0,num,11)
while( t < num ) {
t = t+1
A<-cov.spatial(D,cov.model="exp", cov.par=c(fi1+fi2,fi3))
E<-(gl/(gl-2))*A
IE<-solve(E)
beta.<-solve(t(X)%*%IE%*%X)%*%t(X)%*%IE%*%Z
beta.
r<-(Z-X)%*%beta.)
u<-(t(r)%*%IE%*%r)
v<-as.numeric((gl+n)/(gl+u))
vi<-as.numeric((gl+n)/(gl+u)^2)
I<-(diag(1, nr=n))
R<-1/fi2*(E-fi1*I)
F<-fi2*((R%*%D)/fi3^2)
H<-fi2*((R%*%D)/fi3^3)%*%((D/fi3)-2)
S1<-v*(t(r)%*%IE%*%R%*%IE%*%r)
S2<-v*(t(r)%*%IE%*%R%*%IE%*%r)
S<-cbind(S1,S2)
a11<-sum(diag(IE%*%I%*%IE))
a12<-sum(diag(IE%*%I%*%IE%*%R))
a21<-sum(diag(IE%*%R%*%IE))
a22<-sum(diag(IE%*%R%*%IE%*%R))
A<-matrix(c(a11,a21,a12,a22),2,2)
Fi<-S%*%solve(A)
fi11<-Fi[1,1]
fi22<-Fi[1,2]
Qff13<-1/2*(sum(diag(IE%*%F)))+1/2*v*(t(r)%*%IE%*%F%*%IE%*%r)
Qff13<-1/2*(sum(diag(IE%*%(H-F)%*%IE%*%F)))+
1/2*v*(t(r)%*%IE%*%(H-2*F%*%IE%*%F)%*%IE%*%r)

fi33<-as.numeric(fi3-(solve(Qff13)%*%Qff13))
teta<-cbind(beta0,beta1,beta2,beta3,beta4,beta5,
beta6,fi1,fi2,fi3)

teta1<-cbind(beta.[1],beta.[2],beta.[3],beta.[4],
beta.[5],beta.[6],beta.[7],fi11,fi22,fi33)

l<-((sum((teta-teta1)^2)^0.5))/((sum(teta1^2))^0.5)
nr[t,]=c(teta1,l)
beta<-beta.
fi1<-fi11
fi2<-fi22
fi3<-fi33
}
return(nr)
}
nr<-g(100)
nr
min(nr[,11])
EM<-matrix(nr[1,])
beta<-EM[1:7,1]
fi1<-EM[8,]
fi2<-EM[9,]
fi3<-EM[10,]

```

```

E=cov.spatial(D,cov.model="exp", cov.par=c(fi1+fi2,fi3))
covariancia=gl/(gl-2)*E
plot(D,covariancia,xlab="Distância (h)",ylab="Covariância C(h)"
      ,col='blue',pch="*")

## Validação Cruzada
C<-covariancia
kc<-function(n) {
  kcx<-matrix(0,n,3)
  i=0
  while( i < n ) {
    i = i+1
    kcx[i,1]<- Z[i,1]

    K<-function (i){
      Kx<-matrix(0,(n-1),(n-1))
      if (i==1){
        Kx<-as.matrix(C[(i+1):n,(i+1):n])
        return(Kx)
      }
      else if (i==n){
        Kx<-as.matrix(C[(1:(n-1)),(1:(n-1))])
        return(Kx)
      }
      else if (i==2){
        K1<-as.matrix(C[1,1])
        K2<-as.matrix(C[(i+1):n,1])
        K3<-as.matrix(t(C[1,(i+1):n]))
        K4<-as.matrix(C[(i+1):n,(i+1):n])
        Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
        return(Kx)
      }
      else if (i==(n-1)){
        K1<-as.matrix(C[1:(i-1),1:(i-1)])
        K2<-as.matrix(t(C[n,1:(i-1)]))
        K3<-as.matrix(C[1:(i-1),n])
        K4<-as.matrix(C[n,n])
        Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
        return(Kx)
      }
      else {
        K1<-as.matrix(C[(1:(i-1)),(1:(i-1))])
        K2<-as.matrix(C[(i+1):n,1:(i-1)])
        K3<-as.matrix(C[1:(i-1),(i+1):n])
        K4<-as.matrix(C[(i+1):n,(i+1):n])
        Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
        return(Kx)
      }
    }
    K<-K(i)
    L<-function (i){
      Lx<-matrix(0,n,1)
      if (i==1){
        L<-as.matrix(C[i,(1+i):n])
        L<-rbind(L,1)
        Lx<-L
        return(Lx)
      }
      else if (i==n){
        L<-as.matrix(C[n,1:(n-1)])
        L<-rbind(L,1)
        Lx<-L
        return(Lx)
      }
      else {
        L1<-as.matrix(C[i,1:(i-1)])
        L2<-as.matrix(C[i,(i+1):n])
        L<-rbind(L1,L2,1)
        Lx<-L
        return(Lx)
      }
    }
    L.<-L(i)
    T<-as.matrix(rep(1,n-1),n-1,1)
    K<-cbind(K,T)
    T<-t(rbind(T,0))
    K<-rbind(K,T)
  }
}

```

```

Lamb<-solve(K)/*%L.
Lambda<-Lamb[1:(n-1),]
Lambda<-as.matrix(Lambda,(n-1),1)
Z.<-function (i){
Z.x<-matrix(0,n,1)
if (i==1){
Z.x<-as.matrix(Z[(i+1):n,1])
return(Z.x)
}
else if (i==n){
Z.x<-as.matrix(Z[1:(i-1),1])
return(Z.x)
}
else {
Z.1<-as.matrix(Z[1:(i-1),1])
Z.2<-as.matrix(Z[(i+1):n,1])
Z.x<-rbind(Z.1,Z.2)
return(Z.x)
}
}
Z.<-Z.(i)
Z1<-t(Lamb1)/*%Z.
kcx[i,2]<- Z1
sd<-(sum(t(Lamb1)/*%L.[(1:(n-1)),1])+Lamb[n,])^0.5
kcx[i,3]<- sd
}
return(kcx)
}
KC<-kc(n)
KC
DF<-(KC[,1]-KC[,2])
s<-summary(DF)
s<-t(as.matrix(s,1,6))
Sdf=as.matrix(sd(DF))
erro<as.matrix(cbind(s,Sdf),1,7)
DF.<-(KC[,1]-KC[,2])/(KC[,3])
s.<-summary(DF.)
s.<-t(as.matrix(s.,1,6))
Sdf.=as.matrix(sd(DF.))
erro_std<as.matrix(cbind(s.,Sdf.),1,7)
summary<-rbind(erro,erro_std)
summary
summary
EA<sum(abs(DF))/n
EA

## AKAIKE
r=10
AIC=-2*Lv+2*r
AIC
CAIC=-2*Lv+2*r+(2*r*(r+1))/(n-r-1)
CAIC

#####
##DIAGNÓSTICO COM COVARIÁVEIS PARA AS ESTIMATIVAS PELA T-STUDENT##
##          USUAL "1"          ##
#####
beta=c(4.35,-0.09,0.07,-0.01,-0.27,0.13,-0.81)
fi1=0.0941
fi2=0.0367
fi3=147.14
Y=Z
beta<-matrix(beta,7,1)
teta=cbind(t(beta),fi1,fi2,fi3)
E=cov.spatial(D,cov.model="gaus", cov.par=c(fi1+fi2,fi3))
covariancia=gl/(gl-2)*E

##Calculando a matriz de covariância com as estimativas dos
##parâmetros phi(f1, f2, f3)
IC<-solve(covariancia)

##Derivadas
D1<-diag(1,nr=n)
D2<-1/fi2*(covariancia-fi1*D1)
D3<-fi2*((D2/*%D)/(fi3^2))
D11<-matrix(0,n,n)
D22<-matrix(0,n,n)
D33<-fi2*((D2/*%D)/(fi3^3))/*%((D/fi3)-2)

```

```

D23<-(D2%*%D)/(fi3^2)

##Matriz delta beta e delta phi
r<-(Y-X%*%beta)
delta<-(t(r)*%IC%*%(r))
v<-as.numeric((g1+n)/(g1+delta))
vi<-as.numeric((g1+n)/(g1+delta)^2)
Db<-v*t(X)%*%IC -2*v*i*t(X)%*%IC%*%r%*%t(r)%*%IC
Dfi1<-v*t(r)*%IC%*%D1-vi*t(r)*%IC%*%D1%*%IC%*%r%*%t(r)%*%IC
Dfi2<-v*t(r)*%IC%*%D2-vi*t(r)*%IC%*%D2%*%IC%*%r%*%t(r)%*%IC
Dfi3<-v*t(r)*%IC%*%D3-vi*t(r)*%IC%*%D3%*%IC%*%r%*%t(r)%*%IC

##Matriz delta
Df<-rbind(Dfi1,Dfi2,Dfi3)
Delta<-rbind(Db,Df)

##Matriz de informação observada (L)
##L com respeito a beta/beta
Lbb<-v*t(X)%*%IC%*%X-2*v*i*t(X)%*%IC%*%r%*%t(r)%*%IC%*%X

##L com respeito a beta/phi
Lbfi1<-v*t(X)%*%IC%*%D1%*%IC%*%r+v*i*t(X)%*%IC%*%r%*%
t(r)%*%IC%*%D1%*%IC%*%r
Lbfi2<-v*t(X)%*%IC%*%D2%*%IC%*%r+v*i*t(X)%*%IC%*%r%*%
t(r)%*%IC%*%D2%*%IC%*%r
Lbfi3<-v*t(X)%*%IC%*%D3%*%IC%*%r+v*i*t(X)%*%IC%*%r%*%
t(r)%*%IC%*%D3%*%IC%*%r
Lbf<-cbind(Lbfi1,Lbfi2,Lbfi3)
Lbf<-t(Lbf)

##L com respeito a phii/phij
Lf1f1<-1/2*(sum(diag(IC%*%(D11-D1%*%IC%*%D1))))+
1/2*v*(t(r)*%IC%*%(D11-D1%*%IC%*%D1-D1%*%IC%*%D1)*%*%IC%*%r) +
1/2*vi*(t(r)*%IC%*%D1%*%IC%*%r)*(t(r)*%IC%*%D1%*%IC%*%r)

Lf2f2<-1/2*(sum(diag(IC%*%(D11-D2%*%IC%*%D2))))+
1/2*v*(t(r)*%IC%*%(D22-D2%*%IC%*%D2-D2%*%IC%*%D2)*%*%IC%*%r) +
1/2*vi*(t(r)*%IC%*%D2%*%IC%*%r)*(t(r)*%IC%*%D2%*%IC%*%r)

Lf3f3<-1/2*(sum(diag(IC%*%(D11-D3%*%IC%*%D3))))+
1/2*v*(t(r)*%IC%*%(D33-D3%*%IC%*%D3-D3%*%IC%*%D3)*%*%IC%*%r) +
1/2*vi*(t(r)*%IC%*%D3%*%IC%*%r)*(t(r)*%IC%*%D3%*%IC%*%r)

Lf1f2<-1/2*(sum(diag(IC%*%(D11-D2%*%IC%*%D1))))+
1/2*v*(t(r)*%IC%*%(D11-D1%*%IC%*%D2-D2%*%IC%*%D1)*%*%IC%*%r) +
1/2*vi*(t(r)*%IC%*%D1%*%IC%*%r)*(t(r)*%IC%*%D2%*%IC%*%r)

Lf1f3<-1/2*(sum(diag(IC%*%(D11-D3%*%IC%*%D1))))+
1/2*v*(t(r)*%IC%*%(D11-D1%*%IC%*%D3-D3%*%IC%*%D1)*%*%IC%*%r) +
1/2*vi*(t(r)*%IC%*%D1%*%IC%*%r)*(t(r)*%IC%*%D3%*%IC%*%r)

Lf2f3<-1/2*(sum(diag(IC%*%(D23-D3%*%IC%*%D2))))+
1/2*v*(t(r)*%IC%*%(D23-D2%*%IC%*%D3-D3%*%IC%*%D2)*%*%IC%*%r) +
1/2*vi*(t(r)*%IC%*%D2%*%IC%*%r)*(t(r)*%IC%*%D3%*%IC%*%r)

Lff<-rbind(cbind(Lf1f1,Lf1f2,Lf1f3),cbind(Lf1f2,Lf2f2,Lf2f3),
            cbind(Lf1f3,Lf2f3,Lf3f3))

L1<-cbind(Lbb,Lbf)
L2<-cbind(Lbf,Lff)
L<-rbind(L1,L2)

##Erro padrão do estimador de máxima verossimilhança
n=length(Y)
ag=n/4*((v+n)/(v+n+2))
bg11=sum(diag(IC%*%D1))*sum(diag(IC%*%D1))
bg12=sum(diag(IC%*%D1))*sum(diag(IC%*%D2))
bg13=sum(diag(IC%*%D1))*sum(diag(IC%*%D3))
bg22=sum(diag(IC%*%D2))*sum(diag(IC%*%D2))
bg23=sum(diag(IC%*%D2))*sum(diag(IC%*%D3))
bg33=sum(diag(IC%*%D3))*sum(diag(IC%*%D3))
fg<((n*(n+2))/4)*((v+n)/(v+n+2))
Kb<-4*ag/n*t(X)%*%IC%*%X
Kf1f1<-bg11/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*
        (sum(diag(IC%*%D1%*%IC%*%D1)))
Kf1f2<-bg12/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*
        (sum(diag(IC%*%D1%*%IC%*%D1)))

```

```

(sum(diag(IC%D1%*%IC%D2)))

Kf1f3<-bg13/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*  

  (sum(diag(IC%D1%*%IC%D3)))

Kf2f2<-bg22/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*  

  (sum(diag(IC%D2%*%IC%D2)))

Kf2f3<-bg23/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*  

  (sum(diag(IC%D2%*%IC%D3)))

Kf3f3<-bg33/4*((4*fg)/(n*(n+2))-1)+2*fg/(n*(n+2))*  

  (sum(diag(IC%D3%*%IC%D3)))

Kf<-rbind(cbind(Kf1f1,Kf1f2,Kf1f3),cbind(Kf1f2,Kf2f2,Kf2f3),  

  cbind(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)  

IKt<-solve(Kt)  

DIKt<-diag(IKt)  

DP<-sqrt(DIKt)  

DP<-rbind(DP)  

T<-rbind(teta,DP)  

T

####INFLUÊNCIA LOCAL NO AFASTAMENTO DA VEROSSIMILHANÇA  

B<-t(Delta)%*%solve(L)%*%Delta  

Af<-(sum(diag(t(B)%*%B)))^0.5  

Bd<--B/Af  

Bd

##Gráfico de coeficientes Ci  

Ci<-abs(diag(Bd))  

plot(Ci, xlab="Ordem", ylab=expression(C[i]),ylim=c(0,1),pch=19)  

abline(h=mean(Ci)+2*sd(Ci), col = "black", lty = 3)  

identify(Ci,offset=0.3)

##Gráfico do autovalor |Lmax|
BA<-eigen(Bd,symmetric=TRUE, only.value=FALSE)
abs(BA$values)
Lmax<-BA$vector[,1]
Lmax<-abs(Lmax/sqrt(sum(Lmax^2)))
plot(Lmax, xlab="Ordem", ylab=expression(abs(L[max])),  

  ylim=c(0,1), pch=19)
abline(h=mean(Lmax)+2*sd(Lmax), col = "black", lty = 3)
identify(Lmax,offset=0.3)

#####
##          INFLUÊNCIA LOCAL          ##
#####

## Influência no afastamento da verossimilhança

## Gráfico de diagnóstico l[Ch]
h=50
dCf1=0
dCf2=gl/(gl-2)*exp(-(h/fi3))
dCf3=gl/(gl-2)*(h*fi2*exp(-(h/fi3)))/(fi3)^2
DCh=t(cbind(0,0,0,0,0,0,dCf1,dCf2,dCf3))
DCh=matrix(DCh)
Ch=-t(Delta)%*%solve(L)%*%DCh
lCh= Ch/sum(diag(t(Ch)%*%Ch))
plot(abs(lCh),xlab="Ordem",ylim=c(0,3.5),
  ylab=expression(l[Ch]),pch=19)
identify(abs(lCh),offset=0.3)

## Gráfico de diagnóstico l[p]
bor<-read.table("C:/Análise Covariáveis/Bor.txt",
                 head=FALSE)
plot(bor)
coords=P$coords
n=length(P$data)
s0<-coefs[20,]
s0<-matrix(s0,1,2)
points(s0)

```

```

##PREDIÇÃO DE R1 EM s0
## Algoritmo EM para R1
cR1=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=5)
Z=as.matrix((cR1$data),47,1)
beta<-1.40
fi1<-0.2
fi2<-0.1
fi3<-70
Cov<-cov.spatial(D,cov.model="exp", cov.par=c(fi1+fi2,fi3))
A<-(g1/(g1-2))*Cov

K<-function (i){
  if (i==1){
    Kx<-as.matrix(A[(i+1):n,(i+1):n])
  }
  else if (i==n){
    Kx<-as.matrix(A[(1:(n-1)),(1:(n-1))])
  }
  else if (i==2){
    K1<-as.matrix(A[1,1])
    K2<-as.matrix(A[(i+1):n,i])
    K3<-as.matrix(t(A[1,(i+1):n]))
    K4<-as.matrix(A[(i+1):n,(i+1):n])
    Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
  }
  else if (i==(n-1)){
    K1<-as.matrix(A[1:(i-1),1:(i-1)])
    K2<-as.matrix(t(A[n,1:(i-1)]))
    K3<-as.matrix(A[1:(i-1),n])
    K4<-as.matrix(A[n,n])
    Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
  }
  else {
    K1<-as.matrix(A[(1:(i-1)),(1:(i-1))])
    K2<-as.matrix(A[(i+1):n,1:(i-1)])
    K3<-as.matrix(A[1:(i-1),(i+1):n])
    K4<-as.matrix(A[(i+1):n,(i+1):n])
    Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
  }
}
K<-K(6)

Li<-function (i){
  if (i==1){
    Li<-as.matrix(A[i,(1+i):n])
    Li<-rbind(Li,1)
    Lix<-Li
  }
  else if (i==n){
    Li<-as.matrix(A[n,1:(n-1)])
    Li<-rbind(Li,1)
    Lix<-Li
  }
  else {
    Li1<-as.matrix(A[i,1:(i-1)])
    Li2<-as.matrix(A[i,(i+1):n])
    Li<-rbind(Li1,Li2,1)
    Lix<-Li
  }
}
Li.<-Li(6)

T<-as.matrix(rep(1,n-1),n-1,1)
K<-cbind(K,T)
T<-t(rbind(T,0))
K<-rbind(K,T)
Lamb<-solve(K)/*%Li.
Lamb1<-Lamb[1:(n-1),]
Lamb1<-as.matrix(Lamb1,(n-1),1)

Z.<-function (i{
  Z.x<-matrix(0,n-1,1)
  if (i==1){
    Z.x<-as.matrix(Z[(i+1):n,1])
    return(Z.x)
  }
}

```

```

else if (i==n){
Z.x<-as.matrix(Z[1:(i-1),1])
return(Z.x)
}
else {
Z.1<-as.matrix(Z[1:(i-1),1])
Z.2<-as.matrix(Z[(i+1):n,1])
Z.x<-rbind(Z.1,Z.2)
return(Z.x)
}
}
ZR1.<-Z.(6)
ZR1<-t(Lamb1)%%ZR1.
ZR1

##PREDIÇÃO DE R2 EM s0
cR2=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=6)
Z=as.matrix((cR2$data),47,1)
beta=mean(Z)
f11=0.05
f12=0.15
f13=70
teta=c(beta,f11,f12,f13)
ZR2.<-Z.(6)
ZR2<-t(Lamb1)%%ZR2.
ZR2

##PREDIÇÃO DE R3 EM s0
cR3=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=7)
Z<-as.matrix(cR3$data)
beta=mean(Z)
f11=0
f12=0.25
f13=100
teta=c(beta,f11,f12,f13)
ZR3.<-Z.(6)
ZR3<-t(Lamb1)%%ZR3.
ZR3

##PREDIÇÃO DE D1 EM s0
cD1=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=8)
Z<-as.matrix(cD1$data)
beta=mean(Z)
f11=0.004
f12=0.004
f13=100
teta=c(beta,f11,f12,f13)

ZD1.<-Z.(6)
ZD1<-t(Lamb1)%%ZD1.
ZD1

##PREDIÇÃO DE D2 EM s0
cD2=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=9)
Z<-as.matrix(cD2$data)

beta=mean(Z)
f11=0.006
f12=0.004
f13=120
teta=c(beta,f11,f12,f13)

ZD2.<-Z.(6)
ZD2<-t(Lamb1)%%ZD2.
ZD2

##PREDIÇÃO DE D3 EM s0
cD3=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=10)
Z<-as.matrix(cD3$data)
beta=mean(Z)
f11=0.004
f12=0.002

```

```

fi3=40 #corresponde a 1/3 do alcance
teta=c(beta,fi1,fi2,fi3)
ZD3.<-Z.(6)
ZD3<-t(Lamb1)%%ZD3.
ZD3
x0=t(cbind(1,ZR1,ZR2,ZR3,ZD1,ZD2,ZD3))
fi1=0.0941
fi2=0.0367
fi3=147.14
dist_s0=sapply(1:n,function(p)
sqrt(((matrix(coords[p,])[1,])-s0[,1])^2+
((matrix(coords[p,])[2,])-s0[,2])^2))
dist_s0
C0=cov.spatial(dist_s0,cov.model="exp", cov.par=c(fi1+fi2,fi3))
cov_s0=gl/(gl-2)*C0
dcf1=0
dcf2=gl/(gl-2)*exp(-(dist_s0/ fi3))
dcf3=gl/(gl-2)*(dist_s0*fi2*exp(-(dist_s0/ fi3)))/(fi3)^2
pb=x0-t(X)%%IC%%CO
pf1=(dcf1-t(CO))%%IC%%D1)%%IC%%r
pf2=(dcf2-t(CO))%%IC%%D2)%%IC%%r
pf3=(dcf3-t(CO))%%IC%%D3)%%IC%%r
pf=cbind(pf1,pf2,pf3)
pf=matrix(pf)
pt=rbind(pb,pf)
p.=-t(Delta)%%solve(L)%%pt
Lp=abs(p./sqrt(sum(p.^2)))
plot(Lp,xlab="Ordem", ylab=expression(l[p]),
ylim=c(0,0.6),pch=19)
identify(Lp,offset=0.3)

##Gráfico de Diagnóstico LG vs i##
LbY=-1/2*v*t(X)%%IC+vi*t(X)%%IC%%r%%t(r)%%IC
Lf1Y=-v*t(r)%%IC%%D1%%IC+vi*t(r)%%IC%%D1%%IC%%r
(r)%%t(r)%%IC

Lf2Y=-v*t(r)%%IC%%D2%%IC+vi*t(r)%%IC%%D2%%IC%%r
(r)%%t(r)%%IC

Lf3Y=-v*t(r)%%IC%%D3%%IC+vi*t(r)%%IC%%D3%%IC%%r
(r)%%t(r)%%IC

LfY=rbind(Lf1Y,Lf2Y,Lf3Y)
LG1=X%%solve(Lbb+Lbf)%%solve(Lff)%%LbY
LG2=X%%solve(Lbb+Lbf)%%solve(Lff)%%Lbf)%%(-Lbf%%solve(Lff)%%LfY)
LG=LG1+LG2
G=abs(diag(LG))
plot(G,xlab="Ordem",ylab=expression(LG),ylim=c(0,60),
font.main=4, pch=19)
#abline(h=mean(G)+2*sd(G), col = "black", lty = 3)
identify(G,offset=0.3)

#####
##      DIAGNÓSTICO COM COVARIÁVEIS PARA AS ESTIMATIVAS    ##
##          PELA T-STUDENT "Q-deslocamento"                 ##
#####
beta<-matrix(beta,7,1)
teta=cbind(t(beta),fi1,fi2,fi3)
Y=Z
##Calculando a matriz de covariância com as estimativas
##dos parâmetros phi(f1, f2, f3)
IC<-solve(covariancia)

##Derivadas
D1<-diag(1,nr=n)
D2<-1/fi2*(covariancia-fi1*D1)
D3<-fi2*((D2*%D)/(fi3^2))
D11<-matrix(0,n,n)
D22<-matrix(0,n,n)
D33<-fi2*((D2*%D)/(fi3^3))%%((D/fi3)-2)
D23<-(D2*%D)/(fi3^2)

##Matriz delta beta e delta phi
r<-Y-X%%beta
delta<-(t(r))%%IC%%r
W<-as.numeric((gl+n)/(gl+delta))

```

```

Wl<-as.numeric((gl+n)/(gl+delta)^2)
Db<-W*t(X)%%IC -2*Wl*t(X)%%IC%%r%%t(r)%%IC
Dfi1<-W*t(r)%%IC%%D1-Wl/2*t(r)%%IC%%D1%%r%%t(r)%%IC
Dfi2<-W*t(r)%%IC%%D2-Wl/2*t(r)%%IC%%D2%%r%%t(r)%%IC
Dfi3<-W*t(r)%%IC%%D3-Wl/2*t(r)%%IC%%D3%%r%%t(r)%%IC

##Matriz delta
Df<-rbind(Dfi1,Dfi2,Dfi3)
Delta<-rbind(Db,Df)

##Matriz de informação observada (Q)
##Q com respeito a beta
Qbb<-W*t(X)%%IC%%X

##L com respeito a beta e phi
Qbfi1<-W*t(X)%%D1%%IC%%r
Qbfi2<-W*t(X)%%D2%%IC%%r
Qbfi3<-W*t(X)%%D3%%IC%%r
Qbf<-cbind(Qbfi1,Qbfi2,Qbfi3)
Qfb<-t(Qbf)

##Q com respeito a phii e phij
Qf1f1<-1/2*(sum(diag(IC%%(D11-D1%%IC%%D1)))+
  W/2*(t(r)%%IC%%(D11-D1%%IC%%D1-D1%%IC%%D1))%%IC%%(r))

Qf2f2<-1/2*(sum(diag(IC%%(D22-D2%%IC%%D2)))+
  W/2*(t(r)%%IC%%(D22-D2%%IC%%D2-D2%%IC%%D2))%%IC%%(r))

Qf3f3<-1/2*(sum(diag(IC%%(D33-D3%%IC%%D3)))+
  W/2*(t(r)%%IC%%(D33-D3%%IC%%D3-D3%%IC%%D3))%%IC%%(r))

Qf1f2<-1/2*(sum(diag(IC%%(D11-D2%%IC%%D1)))+
  W/2*(t(r)%%IC%%(D11-D2%%IC%%D1-D1%%IC%%D2))%%IC%%(r))

Qf1f3<-1/2*(sum(diag(IC%%(D11-D3%%IC%%D1)))+
  W/2*(t(r)%%IC%%(D11-D3%%IC%%D1-D1%%IC%%D3))%%IC%%(r))

Qf2f3<-1/2*(sum(diag(IC%%(D23-D3%%IC%%D2)))+
  W/2*(t(r)%%IC%%(D23-D3%%IC%%D2-D2%%IC%%D3))%%IC%%(r))

Qff<-rbind(cbind(Qf1f1,Qf1f2,Qf1f3),cbind(Qf1f2,Qf2f2,Qf2f3),
  cbind(Qf1f3,Qf2f3,Qf3f3))

Q1<-cbind(Qbb,Qbf)
Q2<-cbind(Qfb,Qff)
Q<-rbind(Q1,Q2)
Q

#####
# INFLUÊNCIA LOCAL #
#####

##Influência no afastamento da verossimilhança
Bq<-t(Delta)%%solve(Q)%%Delta
Afq<-(sum(diag(t(Bq)%%Bq)))^0.5
Bdq<-Bq/Afq
Bdq

##Gráfico de coeficientes CQi
Cqi<-2*abs(diag(Bdq))
plot(Cqi, xlab="Ordem", ylab=expression(C[Qi]),
  ylim=c(0,1), pch=19)
abline(h=mean(Ci)+2*sd(Ci), col = "black", lty = 3)
identify(Cqi,offset=0.3)

##Gráfico do autovalor |Qmax|
BAq<-eigen(Bdq,symmetric=TRUE, only.value=FALSE)
abs(BAq$values)
Qmax<-BAq$vector[,1]
Qmax<-abs(Qmax/sqrt(sum(Qmax^2)))
plot(Qmax, xlab="Ordem", ylab=expression(abs(Q[max])),
  ylim=c(0,1), pch=19)
abline(h=mean(Lmax)+2*sd(Lmax), col = "black", lty = 3)
identify(Qmax,offset=0.3)

## INFLUÊNCIA NA MATRIZ DE COVARIÂNCIA E INFLUÊNCIA NO
## PREDITOR LINEAR

```

```

## Gráfico de diagnóstico na matriz de covariância lC[Qh]
h=50
dCf1=0
dCf2=g1/(g1-2)*exp(-(h/fi3))
dCf3=g1/(g1-2)*(h*fi2*exp(-(h/fi3)))/(fi3)^2
DChq=t(cbind(0,0,0,0,0,0,dCf1,dCf2,dCf3))
DCqh=matrix(DChq)
Chq=-t(Delta) %*% solve(Q) %*% DChq
lChq= Chq/sum(diag(t(Chq)) %*% Chq)
plot(abs(lChq),xlab="Ordem",ylim=c(0,3.5),
     ylab=expression(Q[Ch]),pch=19)
identify(abs(lChq),offset=0.3)

## Gráfico de diagnóstico nno preditor linear lp[Q]
bor<-read.table("C:/Desktop/Análise Covariáveis/Bor.txt", head=FALSE)
plot(bor)
coords=P$coords
n=length(P$data)
s0<-coords[20,]
s0<-matrix(s0,1,2)
points(s0)

##PREDIÇÃO DE R1 EM s0
## Algoritmo EM para R1##
cR1=read.geodata("C:/Análise Covariáveis/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=5)
Y<-as.matrix(cR1$data)
X.=matrix(rep(1,n))
beta=mean(Z)
fi1=0.15
fi2=0.15
fi3=90
teta=c(beta,fi1,fi2,fi3)
g1=10
t=0
g<-function(num, tipo ) {
nr<-matrix(0,num,5)
while( t < num ) {
  t = t+1
  A<-cov.spatial(D,cov.model="exp", cov.par=c(fi1+fi2,fi3))
  E<-(g1/(g1-2))*A
  IE<-solve(E)
  beta.<-solve(t(X.) %*% IE %*% X.) %*% t(X.) %*% IE %*% Y
  beta.
  r<-(Y-X.%*%beta.)
  u<-(t(r)%*%IE%*%r)
  v<-as.numeric((g1+n)/(g1+u))
  vi<-as.numeric((g1+n)/(g1+u)^2)
  I<-(diag(1,nr=n))
  R<-1/fi2*(E-fi1*I)
  F<-fi2*R%*%D/(fi3^2)
  H<-fi2*R%*%(D/fi3^3)%*%((D/fi3)-2)
  S1<-v*(t(r)%*%IE%*%I%*%IE%*%r)
  S2<-v*(t(r)%*%IE%*%R%*%IE%*%r)
  S<-cbind(S1,S2)
  a11<-sum(diag(IE%*%I%*%IE))
  a12<-sum(diag(IE%*%I%*%IE%*%R))
  a21<-sum(diag(IE%*%R%*%IE))
  a22<-sum(diag(IE%*%R%*%IE%*%R))
  A<-matrix(c(a11,a21,a12,a22),2,2)
  Fi<-S%*%solve(A)
  fi1<-Fi[1,1]
  fi12<-Fi[1,2]
  Qfi3<- -1/2*(sum(diag(IE%*%F)))+1/2*v*(t(r)%*%IE%*%F%*%IE%*%r)
  Qfffi3<- -1/2*(sum(diag(IE%*%(H-F%*%IE%*%F))))+1/2*v*(t(r)%*%IE%*%
    (H-2*F%*%IE%*%F)%*%IE%*%r)
  fi33<-as.numeric(fi3-(solve(Qfffi3)%*%fi3))
  teta<-cbind(beta,fi1,fi2,fi3)
  teta1<-cbind(beta.,fi11,fi22,fi33)
  l<-((sum((teta-teta1)^2)^0.5))/((sum(teta1^2))^0.5)
  nr[t,]=c(teta1,l)
  beta<-beta.
  fi1<-fi11
  fi2<-fi22
  fi3<-fi33
}
}

```

```

return(nr)
}
nr1<-g(100,"exp")
nr1
min(nr1[,5])
EMR1<-matrix(nr1[1,])
beta<-EMR1[1,]
fi1<-EMR1[2,]
fi2<-EMR1[3,]
fi3<-EMR1[4,]
Cov<-cov.spatial(D,cov.model="exp", cov.par=c(fi1+fi2,fi3))
A<-(gl/(gl-2))*Cov

K<-function (i){
if (i==1){
Kx<-as.matrix(A[(i+1):n,(i+1):n])
}
else if (i==n){
Kx<-as.matrix(A[(1:(n-1)),(1:(n-1))])
}
else if (i==2){
K1<-as.matrix(A[1,1])
K2<-as.matrix(A[(i+1):n,1])
K3<-as.matrix(t(A[1,(i+1):n]))
K4<-as.matrix(A[(i+1):n,(i+1):n])
Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
}
else if (i==(n-1)){
K1<-as.matrix(A[1:(i-1),1:(i-1)])
K2<-as.matrix(t(A[n,1:(i-1)]))
K3<-as.matrix(A[1:(i-1),n])
K4<-as.matrix(A[n,n])
Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
}
else {
K1<-as.matrix(A[(1:(i-1)),(1:(i-1))])
K2<-as.matrix(A[(i+1):n,1:(i-1)])
K3<-as.matrix(A[1:(i-1),(i+1):n])
K4<-as.matrix(A[(i+1):n,(i+1):n])
Kx<-cbind(rbind(K1,K2),rbind(K3,K4))
}
}
}
K<-K(6)

Li<-function (i){
if (i==1){
Li<-as.matrix(A[i,(1+i):n])
Li<-rbind(Li,1)
Lix<-Li
}
else if (i==n){
Li<-as.matrix(A[n,1:(n-1)])
Li<-rbind(Li,1)
Lix<-Li
}
else {
Li1<-as.matrix(A[i,1:(i-1)])
Li2<-as.matrix(A[i,(i+1):n])
Li<-rbind(Li1,Li2,1)
Lix<-Li
}
}
}
Li.<-Li(6)

T<-as.matrix(rep(1,n-1),n-1,1)
K<-cbind(K,T)
T<-t(rbind(T,0))
K<-rbind(K,T)
Lamb<-solve(K)%*%L.
Lamb1<-Lamb[1:(n-1),]
Lamb1<-as.matrix(Lamb1,(n-1),1)

Z.<-function (i){
Z.x<-matrix(0,n-1,1)
if (i==1){
Z.x<-as.matrix(Z[(i+1):n,1])
return(Z.x)
}

```

```

}

else if (i==n){
Z.x<-as.matrix(Z[1:(i-1),1])
return(Z.x)
}
else {
Z.1<-as.matrix(Z[1:(i-1),1])
Z.2<-as.matrix(Z[(i+1):n,1])
Z.x<-rbind(Z.1,Z.2)
return(Z.x)
}
}

ZR1.<-Z.(6)
ZR1<-t(Lamb1)%%ZR1.
ZR1

##PREDIÇÃO DE R2 EM s0
cR2=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=6)
Z<-as.matrix(cR2$data)
g1=10
beta=mean(Z)
fi1=0.05
fi2=0.15
fi3=100 #corresponde a 1/3 do alcance
teta=c(beta,fi1,fi2,fi3)
nrR2<-g(100,"exp")
nrR2
min(nrR2[,5])
EMR2<-matrix(nrR2[,1])
beta<-EMR2[1,]
fi1<-abs(EMR2[2,])
fi2<-abs(EMR2[3,])
fi3<-abs(EMR2[4,])

ZR2.<-Z.(6)
ZR2<-t(Lamb1)%%ZR2.
ZR2

##PREDIÇÃO DE R3 EM s0
cR3=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=7)
Z<-as.matrix(cR3$data)
g1=3
beta=mean(Z)
fi1=0.05
fi2=0.15
fi3=100
teta=c(beta,fi1,fi2,fi3)
nrR3<-g(100,"exp")
nrR3
min(nrR3[,5])
EMR3<-matrix(nrR3[,10])
beta<-EMR3[1,]
fi1<-abs(EMR3[2,])
fi2<-abs(EMR3[3,])
fi3<-abs(EMR3[4,])
ZR3.<-Z.(6)
ZR3<-t(Lamb1)%%ZR3.
ZR3

##PREDIÇÃO DE D1 EM s0
cD1=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=8)

Z<-as.matrix(cD1$data)
g1=5
beta=mean(Z)
fi1=0.05
fi2=0.15
fi3=100 #corresponde a 1/3 do alcance
teta=c(beta,fi1,fi2,fi3)
nrD1<-g(100,"exp")
nrD1
min(nrD1[,5])
EMD1<-matrix(nrD1[,2])
beta<-EMD1[1,]

```

```

fi1<-abs(EMD1[2,])
fi2<-abs(EMD1[3,])
fi3<-abs(EMD1[4,])
ZD1.<-Z.(6)
ZD1<-t(Lamb1)%*%ZD1.
ZD1

##PREDIÇÃO DE D2 EM s0
cD2=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=9)
Z<-as.matrix(cD2$data)
g1=3
#
beta=mean(Z)
fi1=0.05
fi2=0.15
fi3=100 #corresponde a 1/3 do alcance
teta=c(beta,fi1,fi2,fi3)
nrD2<-g(100,"exp")
nrD2
min(nrD2[,5])
EMD2<-matrix(nrD2[10,])
beta<-EMD2[1,]
fi1<-abs(EMD2[2,])
fi2<-abs(EMD2[3,])
fi3<-abs(EMD2[4,])

ZD2.<-Z.(6)
ZD2<-t(Lamb1)%*%ZD2.
ZD2

##PREDIÇÃO DE D3 EM s0
cD3=read.geodata("C:/PCovariáveis.txt",head=FALSE,
                  coords.col=2:3,data.col=10)

Z<-as.matrix(cD3$data)
shapiro.test(Z)
g1=7
beta=mean(Z)
fi1=0.08
fi2=0.05
fi3=100
teta=c(beta,fi1,fi2,fi3)
nrD3<-g(100,"exp")
nrD3
min(nrD3[,5])
EMD3<-matrix(nrD3[10,])
beta<-EMD3[1,]
fi1<-abs(EMD3[2,])
fi2<-abs(EMD3[3,])
fi3<-abs(EMD3[4,])
ZD3.<-Z.(6)
ZD3<-t(Lamb1)%*%ZD3.
ZD3
x0=t(cbind(1,ZR1,ZR2,ZR3,ZD1,ZD2,ZD3))
x0
dist_s0=sapply(1:n,function(p)
  sqrt(((matrix(coords[p,])[1,])-s0[,1])^2+
    ((matrix(coords[p,])[2,])-s0[,2])^2))
dist_s0
CO=cov.spatial(dist_s0,cov.model="exp", cov.par=c(fi1+fi2,fi3))
cov_s0=g1/(g1-2)*CO
dcf1=0
dcf2=g1/(g1-2)*exp(-(dist_s0/fi3))
dcf3=g1/(g1-2)*(dist_s0*fi2*exp(-(dist_s0/fi3)))/(fi3)^2
pbq=x0-t(X)%*%IC%*%CO
pf1q=(dcf1-t(CO)%*%IC%*%D1)%*%IC%*%(r)
pf2q=(dcf2-t(CO)%*%IC%*%D2)%*%IC%*%(r)
pf3q=(dcf3-t(CO)%*%IC%*%D3)%*%IC%*%(r)
pfq=cbind(pf1q,pf2q,pf3q)
pfq=matrix(pfq)
ptq=rbind(pbq,pfq)
pq.=-t(Delta)%*%solve(Q)%*%ptq

Qp=abs(pq./sqrt(sum(pq.^2)))
plot(Qp,xlab="Ordem", ylab=expression(1[pQ]),ylim=c(0,0.6),pch=19)
identify(Qp,offset=0.3)

```

```

##Influência nos valores preditos LG[Q] vs i##
QbY=-1/2*W*t(X)%%IC+W1*t(X)%%IC%%*(r)%%t(r)%%IC
Qf1Y=-W*t(r)%%IC%%D1%%IC+W1*t(r)%%IC%%D1%%IC%%*(r)%%t(r)%%IC
Qf2Y=-W*t(r)%%IC%%D2%%IC+W1*t(r)%%IC%%D2%%IC%%*(r)%%t(r)%%IC
Qf3Y=-W*t(r)%%IC%%D3%%IC+W1*t(r)%%IC%%D3%%IC%%*(r)%%t(r)%%IC
QfY=rbind(Lf1Y,Lf2Y,Lf3Y)
QG1=X%%solve(Qbb+Qbf%%solve(Qff)%%Qfb)%%QbY
QG2=X%%solve(Qbb+Qbf%%solve(Qff)%%Qfb)%%(-Qbf%%solve(Qff)%%QfY)
LGQ=QG1+QG2
LGQ=abs(diag(LGQ))
plot(LGQ,xlab="Ordem",ylab=expression(LG[Q]),ylim=c(0,60),
      font.main=4, pch=19)
#abline(h=mean(G)+2*sd(G), col = "black", lty = 3)
identify(LGQ,offset=0.3)

```

ANEXO A – Dados

Tabela 37: Dados experimentais do nutriente potássio [$cmol\ dm^{-3}$]

X	Y	Potássio	X	Y	Potássio
239951.665	7236987.657	0.28	240283.754	7237341.329	0.20
240092.218	7237009.159	0.21	240348.018	7237403.298	0.23
240235.845	7237027.571	0.41	240475.431	7237347.402	0.41
240374.995	7237047.307	0.19	240615.159	7237366.532	0.23
240796.904	7237109.176	0.29	239606.675	7237364.627	0.31
240759.550	7237141.004	0.51	239751.242	7237384.528	0.21
240832.098	7237176.619	0.36	239892.630	7237409.102	0.48
240937.065	7237128.455	0.30	240032.762	7237425.172	0.29
239931.225	7237128.032	0.19	240174.769	7237444.025	0.21
240073.312	7237147.322	0.18	240313.072	7237464.558	0.21
240025.444	7237168.408	0.19	240456.952	7237486.299	0.25
240117.817	7237208.698	0.23	240595.900	7237504.457	0.26
240214.354	7237171.776	0.18	239590.183	7237506.509	0.40
240354.701	7237185.028	0.21	239633.922	7237484.156	0.29
240495.514	7237207.325	0.36	239729.180	7237522.164	0.25
240635.727	7237228.664	0.24	239873.179	7237546.209	0.27
240776.387	7237248.394	0.29	240011.794	7237563.824	0.21
239629.187	7237227.236	0.31	240435.290	7237625.247	0.32
239773.754	7237245.183	0.17	240390.188	7237644.645	0.33
239912.970	7237267.329	0.36	240478.499	7237564.051	0.29
240051.784	7237285.707	0.25	240575.156	7237643.964	0.18
240191.679	7237307.537	0.26	239711.752	7237660.776	0.28
240334.762	7237326.101	0.23	239848.988	7237685.976	0.30

Tabela 38: Dados experimentais da produtividade [t ha⁻¹]

Ordem	X	Y	Produtividade	Ordem	X	Y	Produtividade
1	40674	37587	2.84	25	40362	37255	3.46
2	40811	37409	2.86	26	40348	37150	3.47
3	40948	37231	3.64	27	40211	37328	3.30
4	40569	37601	3.15	28	40074	37506	3.19
5	40463	37615	3.59	29	39877	37639	3.29
6	40509	37555	2.83	30	39923	37579	2.87
7	40600	37437	3.54	31	40014	37461	3.15
8	40783	37199	3.00	32	40151	37283	3.49
9	40678	37213	2.79	33	40242	37164	3.65
10	40541	37391	4.00	34	40288	37105	3.28
11	40449	37510	3.26	35	40137	37178	3.55
12	40404	37569	3.14	36	39909	37475	3.07
13	40344	37524	2.09	37	39863	37534	2.64
14	40435	37405	3.90	38	39817	37593	2.94
15	40481	37346	3.33	39	39757	37548	2.80
16	40572	37227	3.63	40	39803	37489	3.10
17	40513	37182	3.19	41	39895	37370	3.08
18	40467	37241	3.63	42	39940	37311	2.73
19	40421	37300	3.44	43	40077	37132	3.65
20	40376	37360	3.02	44	40123	37073	3.49
21	40284	37478	2.90	45	39881	37265	3.38
22	40179	37492	3.11	46	39835	37324	3.02
23	40225	37433	3.02	47	39729	37338	3.04
24	40316	37314	4.09				

Tabela 39: Dados experimentais das covariáveis Resistência do solo a penetração [Mpa]

Ordem	$RSP_{0-0,1}$	$RSP_{0-0,1}$	$RSP_{0,2-0,3}$	Ordem	$RSP_{0-0,1}$	$RSP_{0-0,1}$	$RSP_{0,2-0,3}$
1	2.697	2.359	1.750	25	2.469	3.193	2.736
2	3.692	3.709	2.123	26	3.590	2.212	1.478
3	3.124	2.053	2.172	27	2.392	2.360	1.735
4	3.511	2.776	1.969	28	2.237	1.915	1.826
5	2.920	2.564	2.524	29	2.640	2.196	1.574
6	2.680	2.327	1.911	30	2.750	1.919	4.958
7	1.398	1.954	2.738	31	2.992	2.068	1.728
8	2.016	1.356	2.547	32	2.418	2.262	1.897
9	3.455	3.207	2.109	33	3.160	2.048	2.011
10	2.367	2.405	2.296	34	2.813	2.707	1.580
11	2.560	2.178	2.102	35	2.388	1.973	1.677
12	3.259	3.682	2.536	36	2.369	1.988	1.879
13	3.651	2.452	2.197	37	1.929	1.705	1.481
14	2.170	2.782	2.096	38	2.537	2.286	1.762
15	3.672	2.980	2.049	39	3.004	2.593	2.141
16	3.423	2.477	2.099	40	3.330	2.152	1.452
17	2.416	2.824	2.406	41	2.763	2.471	2.289
18	2.519	2.755	2.211	42	2.329	2.132	1.683
19	1.808	2.310	2.463	43	2.544	2.897	2.095
20	2.700	2.366	1.954	44	3.245	3.085	2.663
21	3.508	3.274	2.403	45	2.913	1.861	1.337
22	2.102	2.421	1.630	46	2.541	2.140	1.639
23	2.389	2.331	1.986	47	2.627	2.202	1.637
24	2.556	2.735	2.224				

Tabela 40: Dados experimentais das covariáveis Densidade [$mg\ m^{-3}$]

Ordem	$Des_{0-0,1}$	$Des_{0-0,1}$	$Des_{0,2-0,3}$	Ordem	$Des_{0-0,1}$	$Des_{0-0,1}$	$Des_{0,2-0,3}$
1	1.15	1.30	1.11	25	1.00	1.08	1.01
2	1.15	1.12	1.12	26	1.04	1.12	1.30
3	1.18	1.21	1.07	27	1.21	1.14	1.12
4	1.12	0.99	1.16	28	1.15	1.20	1.08
5	1.01	1.27	1.06	29	1.09	1.09	1.13
6	0.98	1.09	1.01	30	1.19	1.06	1.03
7	1.02	1.11	1.06	31	0.94	1.02	0.99
8	1.16	1.16	1.17	32	0.77	1.22	1.04
9	1.15	1.18	1.19	33	1.02	1.00	1.02
10	1.10	1.12	1.03	34	1.07	1.15	1.03
11	1.03	1.11	1.02	35	1.02	1.21	1.01
12	1.07	1.14	1.14	36	1.04	1.07	0.99
13	1.10	1.14	1.14	37	1.10	1.06	1.07
14	1.07	1.18	1.09	38	1.16	1.02	0.99
15	1.10	1.03	1.16	39	1.11	1.12	1.14
16	1.06	1.12	1.08	40	1.22	1.28	1.24
17	1.16	1.19	1.13	41	1.16	0.98	1.04
18	1.19	1.07	1.04	42	1.05	1.04	1.10
19	1.17	1.14	1.11	43	1.05	1.22	1.10
20	1.10	1.05	1.06	44	1.15	1.05	1.06
21	1.14	1.14	1.13	45	0.98	1.05	1.20
22	1.15	1.14	1.12	46	1.16	1.04	1.20
23	1.05	1.55	1.06	47	1.09	1.09	1.05
24	1.18	1.03	0.95				