Análise geoestatística

Paulo Justiniano Ribeiro Junior *

LEG - Laboratório de Estatística e Geoinformação
Departamento de Estatística
Universidade Federal do Paraná

Mini-curso apresentado na
VI Semana de Estatística

Universidade Estadual de Maringá
Maringá, PR
23–25 Novembro 2009

Endereço para correspondência: Departamento de Estatística, Universidade Federal do Paraná, E-mail: paulojus@ufpr.br

INTRODUÇÃO À ESTATÍSTICA ESPACIAL

1. **Exemplos Básicos de dados espaciais**

2. **Terminologia para estatística espacial**

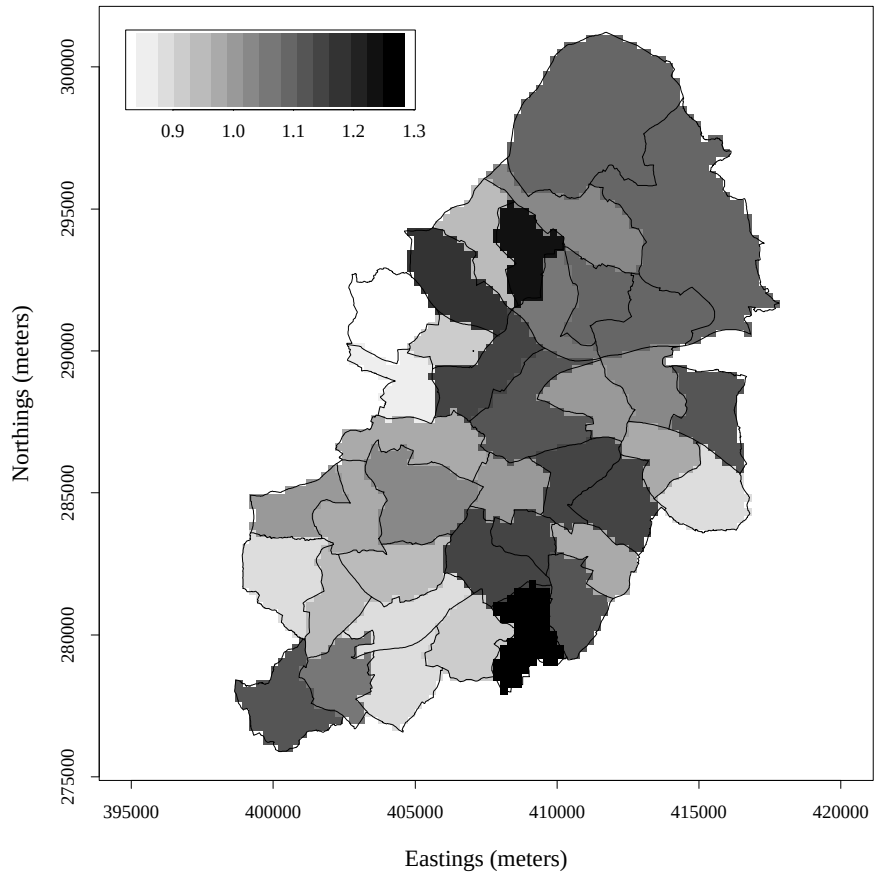
3. **Modelos e problemas em estatística espacial**

4. **Um estudo de caso: doença de citrus**

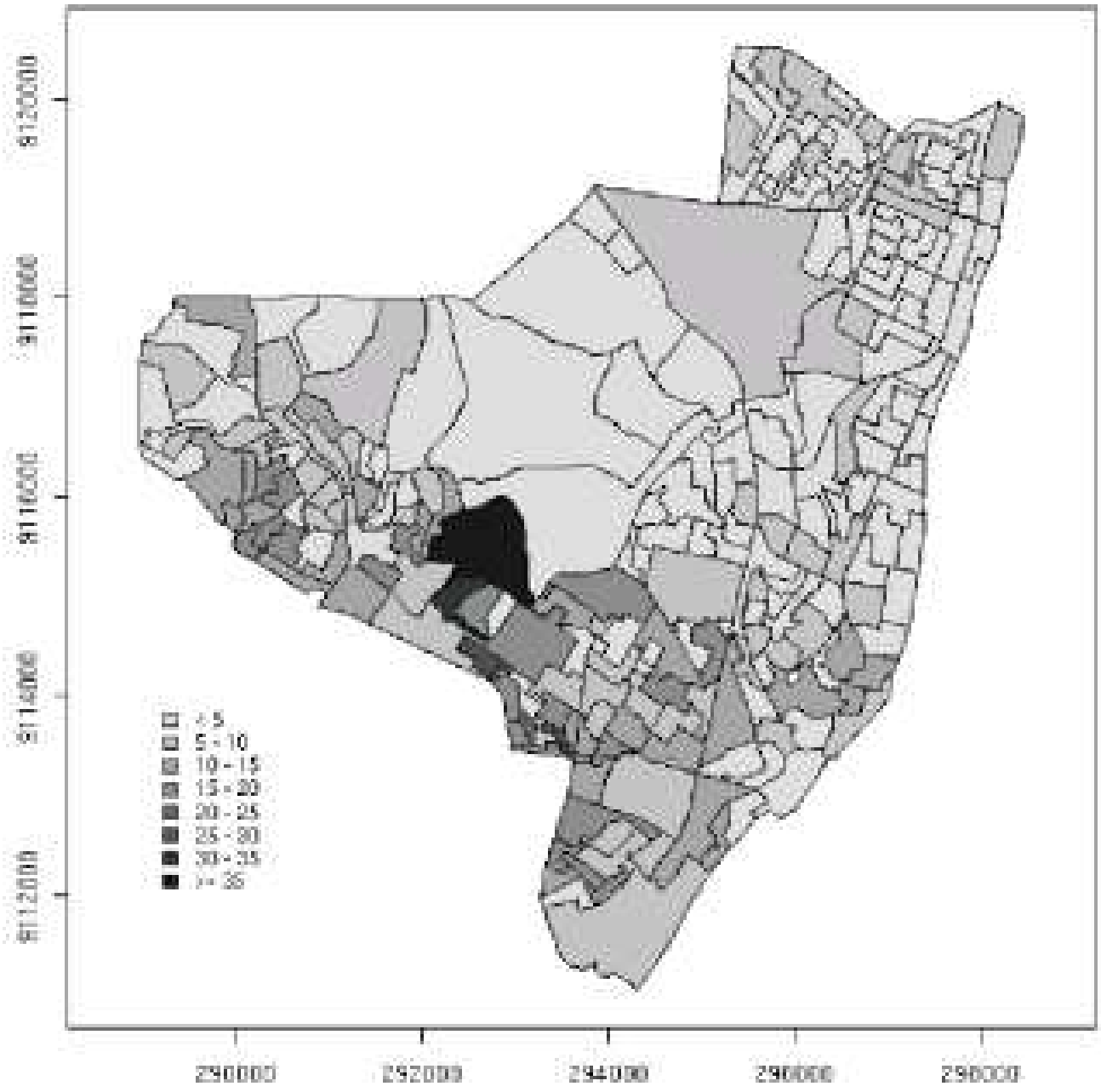
1. Estatística Espacial: Alguns Exemplos Básicos

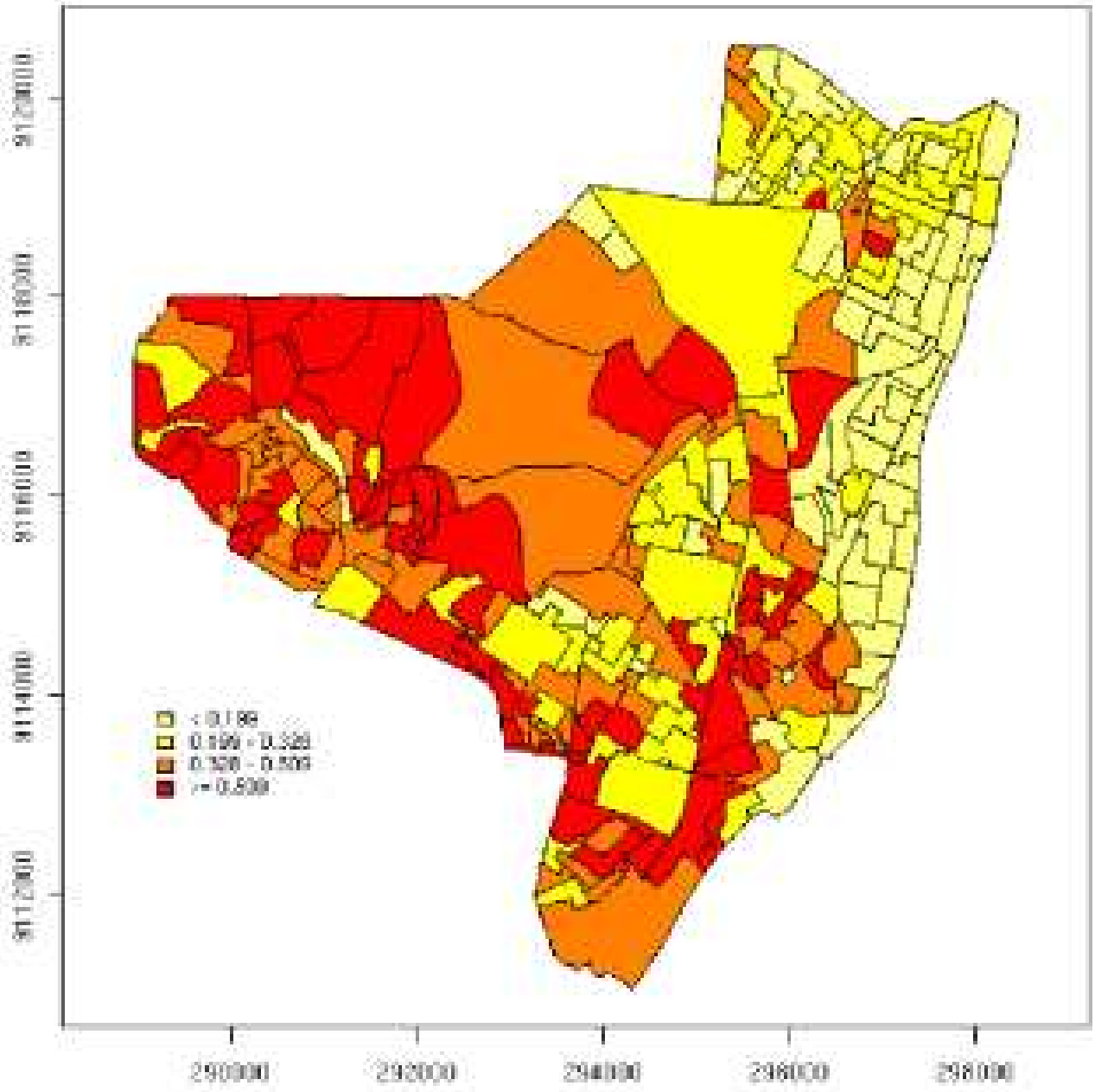
(a) Taxas de câncer por regiões administrativas

tons de cinza correspondem à variação estimada do risco relativo de câncer coloretal em 36 zonas eleitorais da cidade de Birmingham, UK.



(b) Hanseníase em Olinda




< 0.189
0.189 - 0.329
0.329 - 0.509
>= 0.509
290000
292000
294000
296000
298000

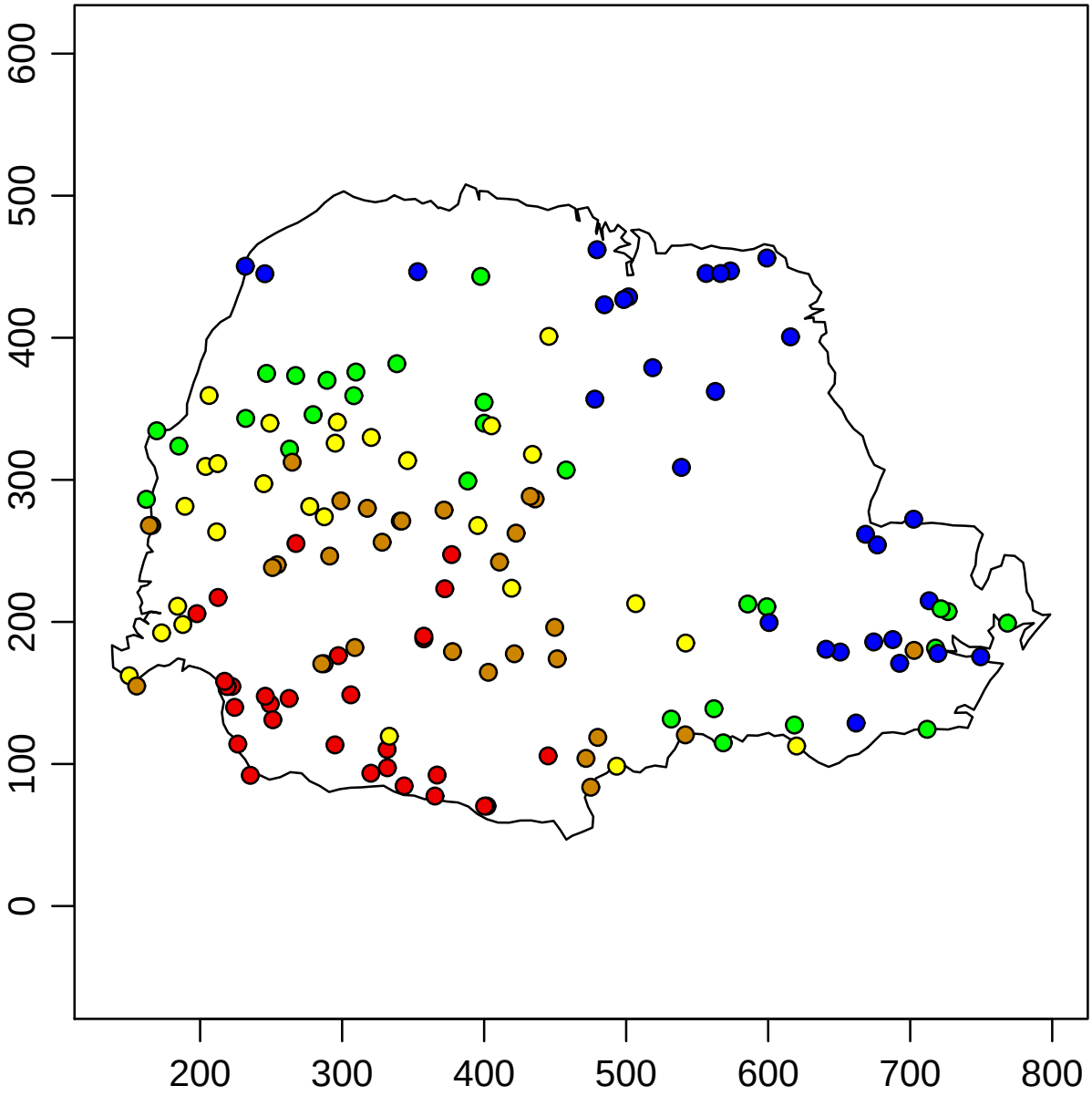
Alguns problemas com estrutura de dados semelhante

- Índices de criminalidade por bairros
- Mortalidade infantil (e/ou outros indicadores) por municípios de um estado
- Experimentos agrícolas de campo
- Análise de imagens

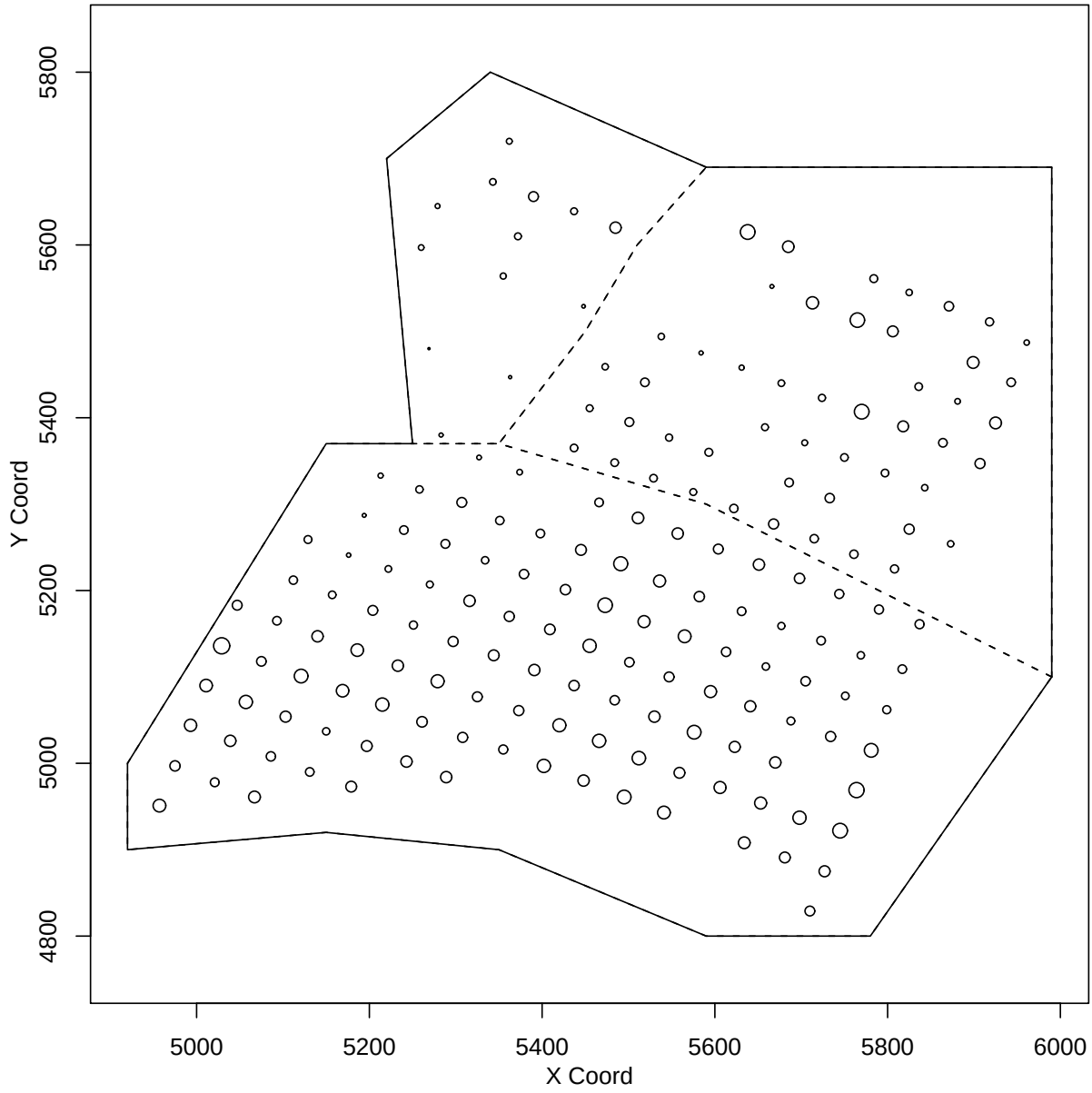
(c) Precipitação no Estado do Paraná

Medidas de chuva em 143 postos meteorológicos.

Médias históricas para o período de Maio-Junho (estação seca).



(d) Teores de Cálcio em um solo agrícola

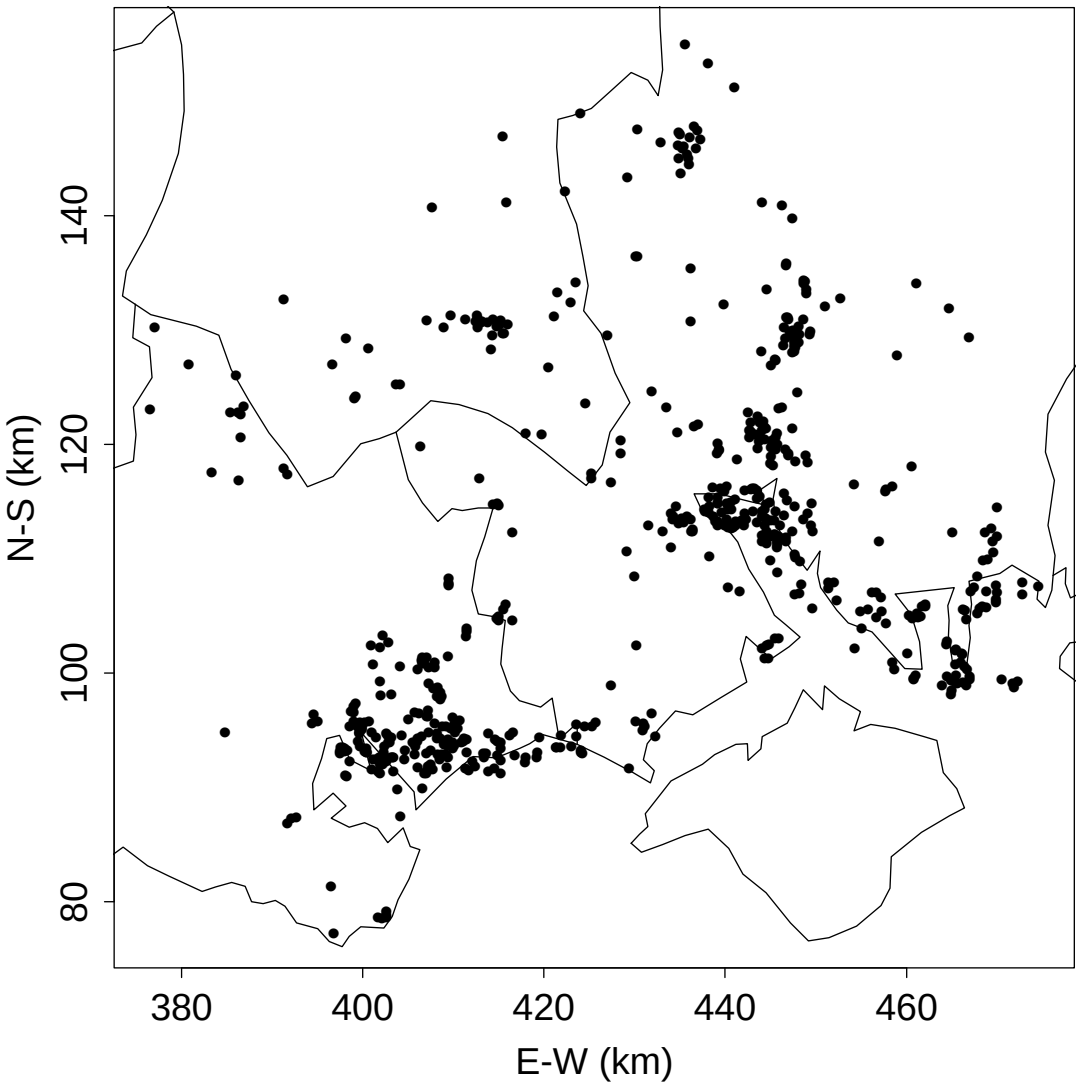


Alguns problemas com estrutura de dados semelhante

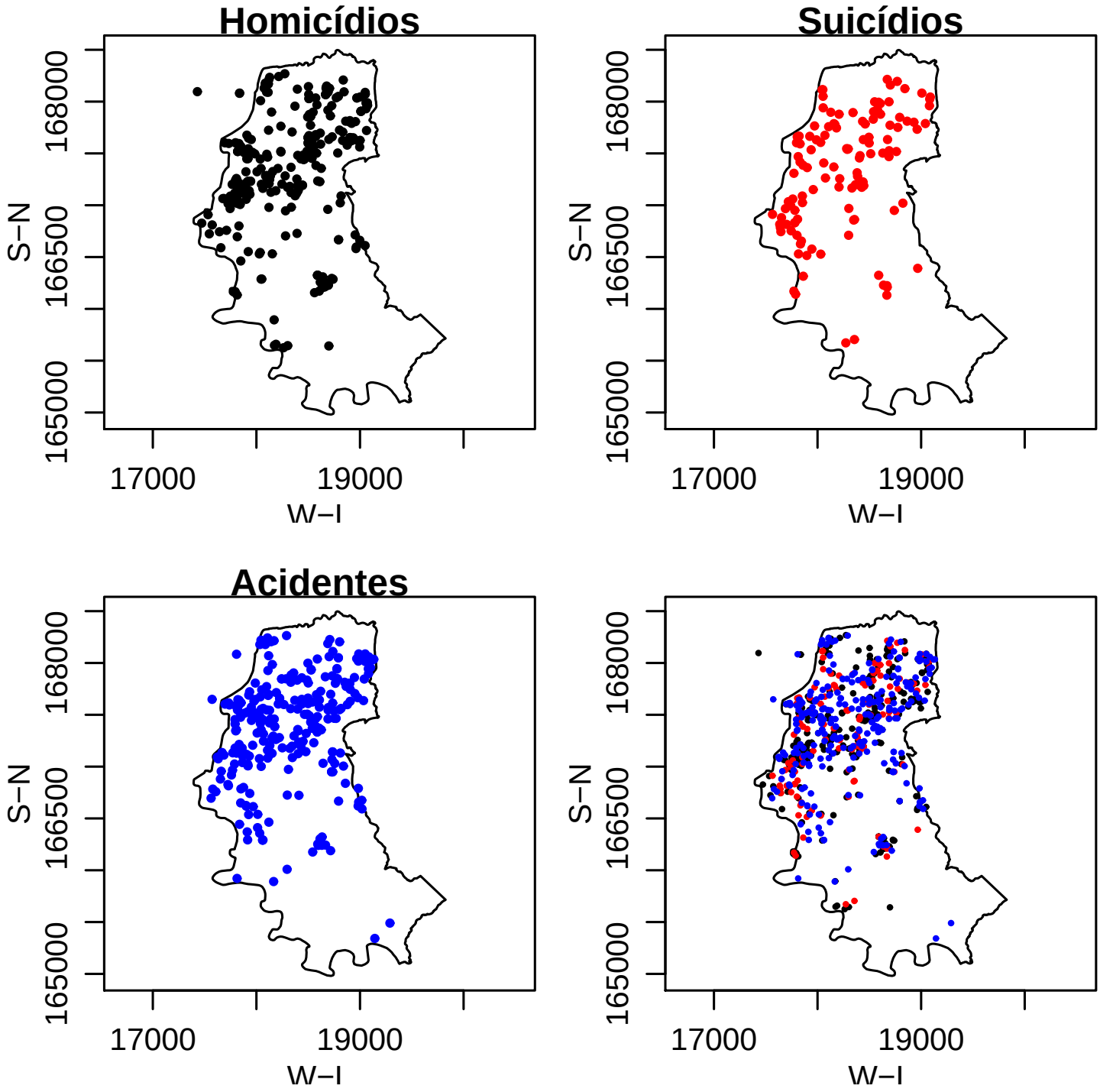
- Teores de elementos minerais em uma jazida
- Níveis de poluição do ar medidos em estações de monitoramento
- Estoque de peixes em uma certa área marítima

(e) Infecções bacterianas no sul da Inglaterra

Localizações das residências de 651 casos notificados num período de 1 ano na região central do sul da Inglaterra.



(f) Ocorrências em Porto Alegre



Alguns problemas com estrutura de dados semelhante

- Localização de árvores de certa espécie em uma area de floresta natural

- Pontos de ocorrência de crimes em uma cidade

- Posições de ninhos de certo pássaro em uma região

2. Terminologia e questões para estatística espacial

(a) Variação espacial discreta

Estrutura básica. $Y_i : i = 1, ..., n$

- raramente ocorre naturalmente
- útil como estratégia pragmática
- modelos são tipicamente definidos indiretamente a partir de condicionais $[Y_i|Y_j, \forall j \neq i]$

i. Medidas de agregação comumente utilizadas (por ex. *I de Moran*)

ii. Diversas opções de modelos, entre eles:

iii. a. Regressão ponderada geográficamente

iv. b. Modelos CAR (auto-regressivo condicional) e SAR

v. Definição de vizinhança pode depender do problema

(b) Variação espacial contínua

Estrutura básica. $Y(x) : x \in \mathbb{R}^2$

- dados $(y_i, x_i) : i = 1, ..., n$, localizações x_i podem ser:
  - não estocástica (ex. grade cobrindo a região em estudo A) ou estocástica, *porém independente do processo* $Y(x)$

i. em geral (mas nem sempre!) o objetivo é de predição
ii. a predição pode ser do processo subjacente ou um funcional deste
iii. possíveis relações com covariáveis
iv. modelos comumente utilizados podem ser vistos como MLG, com efeito aleatório espacialmente estruturado
v. grande número de métodos/algorítmos "ad-hoc" na literatura de geoestatística

(c) Processo pontual espacial

Estrutura básica. Conjunto contável de pontos $x_i \in \mathbb{R}^2$, gerados estocásticamente.

- às vezes dados são agregados em regiões

i. questão chave é dizer se processo é aleatório, agrupado ou regular
ii. modelagem básica para superfície de intensidade do processo pontual
iii. vários modelos disponíveis, "fácil" de simular, difícil de estimar
iv. conhecimento do processo subjacente pode guiar escolha de modelos
v. alguns pontos importantes: correções de borda, correções para população sob risco, estudos de caso-controle, etc

Estatística espacial é a seleção de métodos estatísticos nos quais a localização espacial tem papel explícito na análise dos dados.

Temas estratégicos

- não confundir *formato dos dados* com o *processo subjacente*.
- a escolha do modelo pode ser influenciada pelos objetivos científicos do estudo
- problemas reais não necessariamente se encaixam em um dos tipos básicos, podem ser abordados de difrentes formas ou conterem elementos de cada um dos tipos. A divisão é puramente didática
- há outras possibilidades tais comos processos pontuais marcados, processos espaço temporais, etc
- Estatística espacial e Sistemas de Informação Geográfica (SIG)
- Estatística espacial e Geoestatística
- Problemas espaço-temporais

3. Estudo de caso – Doença de citrus

Problema consiste em quantificar a importância relativa dos dois mecanismos de reprodução do fungo causador a Pinta Preta dos citrus.

Conjectura-se que o padrão espacial da fase assexuada deve ser agragado enquanto que o da fase sexuada deve ser aleatório.

Análises ilustram o uso de métodos associados a cada um dos "tipos básicos" a um mesmo problema.

4. **Palavras finais**

NÃO ANALISE DADOS . . . . . .

. . . . . ANALISE PROBLEMAS !

PROJETO SAUDAVEL
http://saudavel.dpi.inpe.br

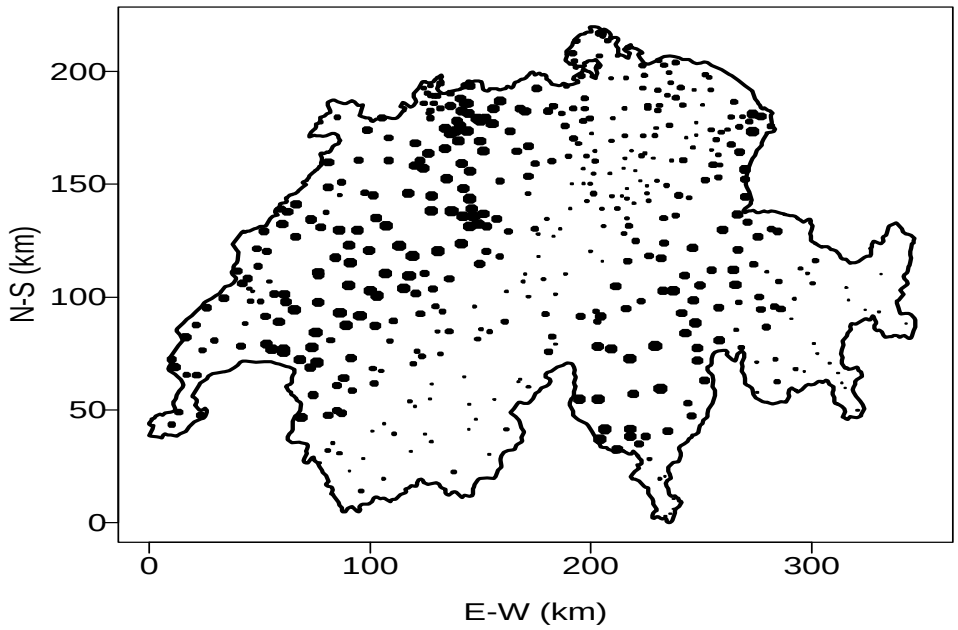
LEG – Laboratório de Estatística de Geoinformação
http://www.est.ufpr.br/leg

Geoestatística

1. Geoestatística: outros exemplos
2. Questões Centrais
3. Modelo Geoestatístico
4. O modelo Gaussiano
5. Predição
6. Estudo de Caso

1. Outros Exemplos de Problemas Geoestatísticos

(a) Dados de chuva na Suíça

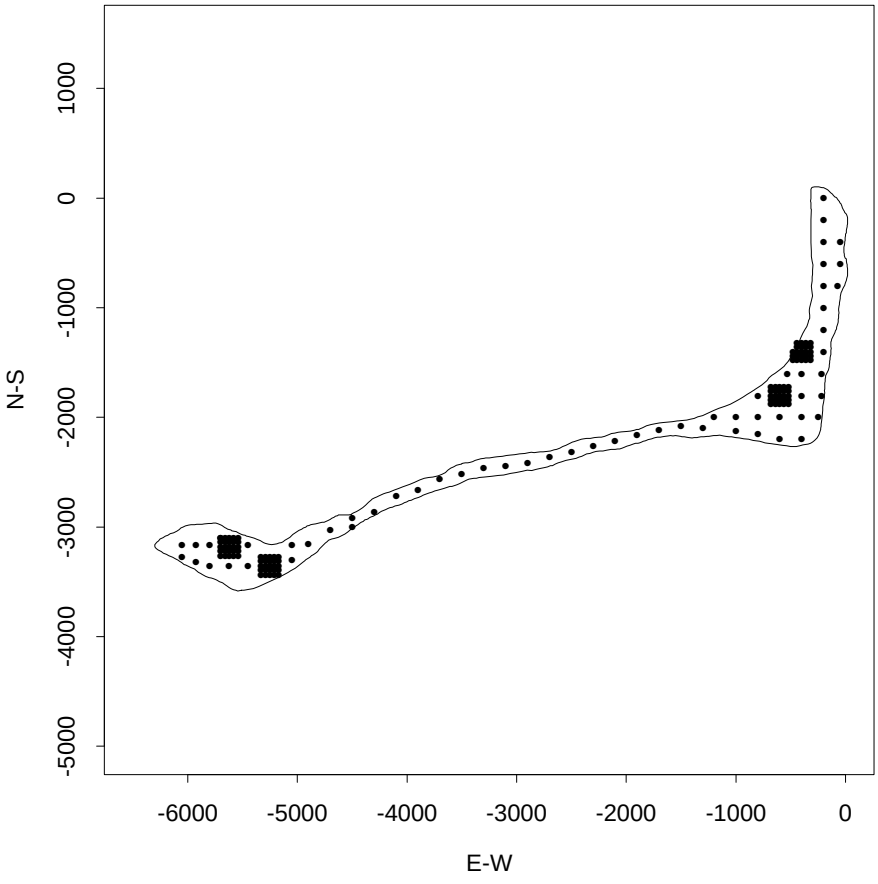


Localizações com tamanhos dos pontos proporcionais aos valores observados de precipitação

- 467 postos na Suíça
- medidas diárias de chuva em 8/05/1986
- dados do projeto:
  Spatial Interpolation Comparison 97
  `ftp://ftp.geog.uwo.ca/SIC97/.`

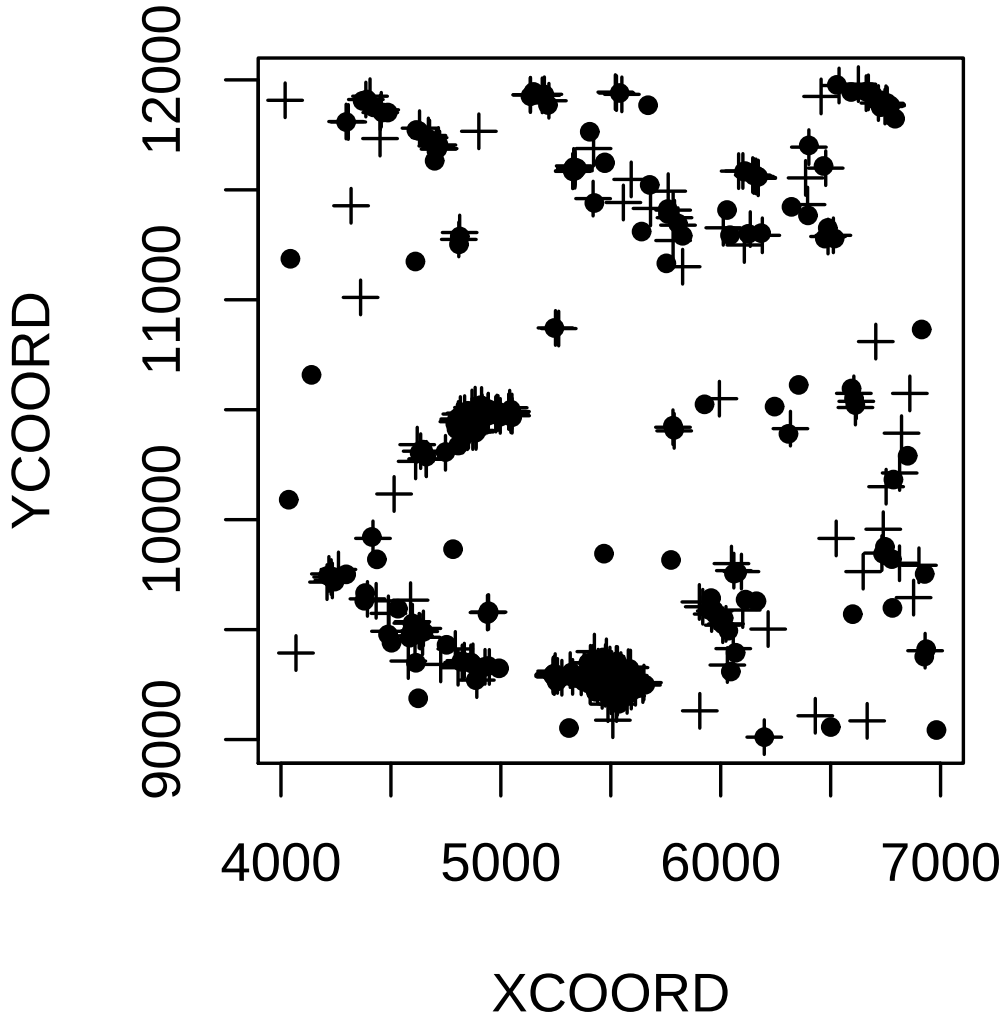
(b) Ilha de Rongelap

- estudo do resíduo de contaminação decorrente de testes de armas nucleares durante a década de 50
- ilha evacuada em 1985. Segura para reocupação
- pesquisa produz medidas com ruído Y_i de concentração de césio radioativo
- particular interesse em níveis máximos de concentração de césio



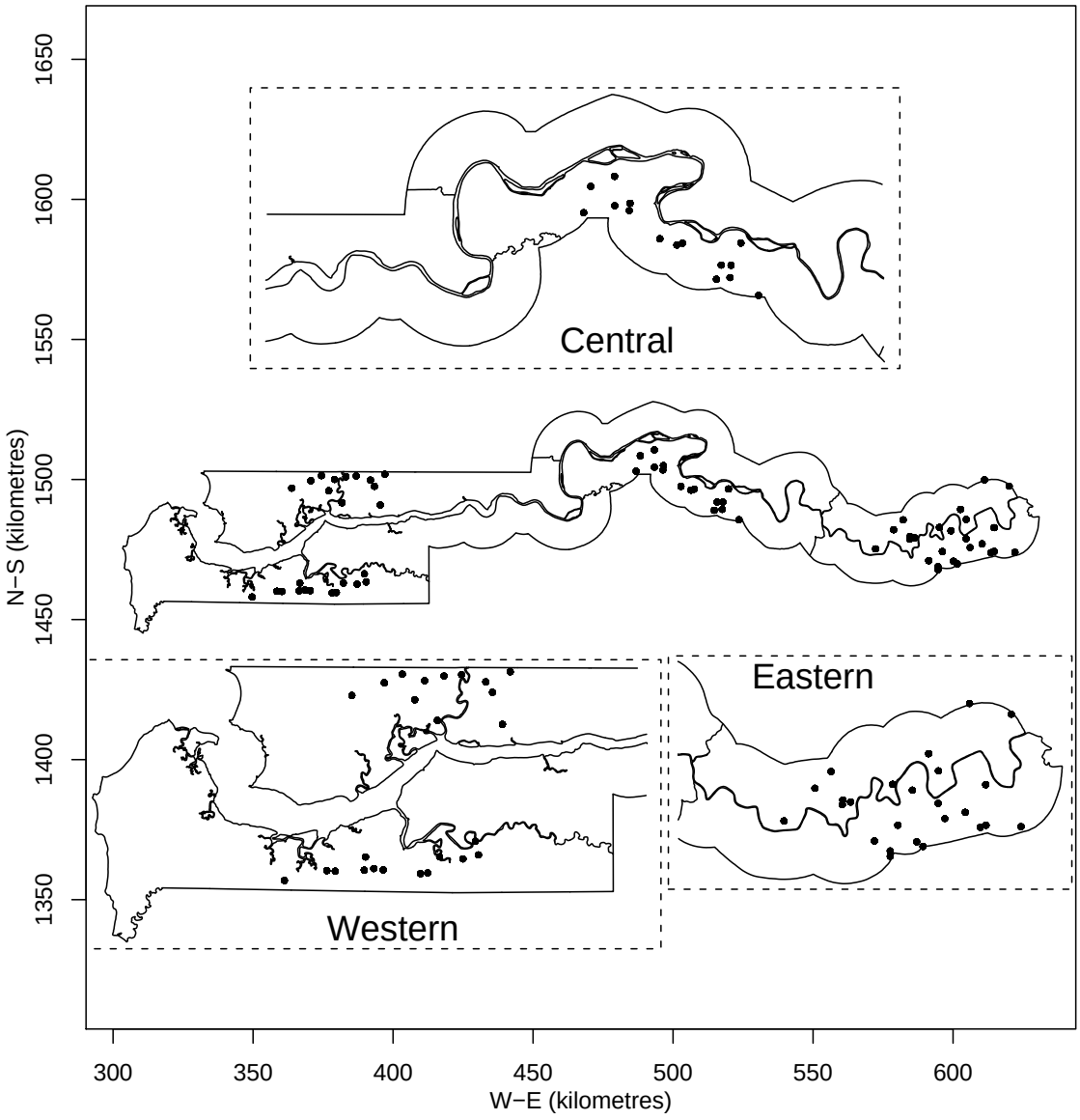
(c) Espécies de líquens

- fatores associados a distribuição espacial da presença de líquens em troncos de ávores
- resposta 0/1: presença ou ausência
- covariáveis: diâmetro, umidade, sombreamento, cobertura do tronco, viva



(d) Malária em Gambia

- na vila i, dado $Y_{ij} = 0/1$ denota ausência ou presença de malária no sangue da criança j
- covariáveis ao nível de vilas:
  - localização (coordenadas), presença de centro de saúde, índice de vegetação derivado de satélite
- covariáveis ao nível de crianças:
  - idade, uso e tratamento de mosquiteiro
- interesses: efeito das covariáveis e padrão espacial da variação residual



2. Características Principais dos Problemas Geoestatísticos

- dados consistem em **respostas** Y_i associadas com **localizações** x_i

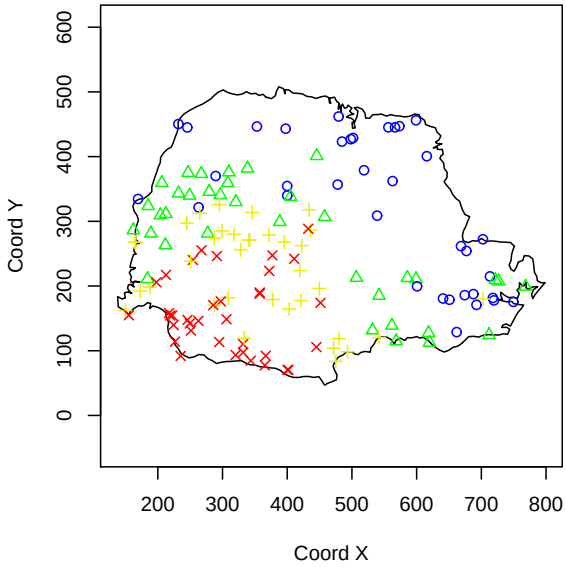
- em princípio, Y pode ser determinado em qualquer localização x dentro da região espacialmente contínua A

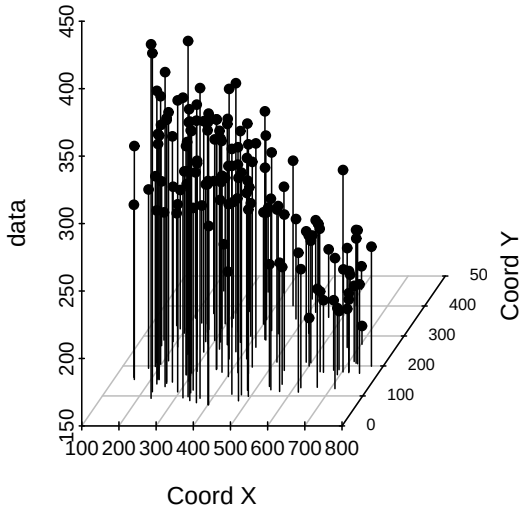
- assume-se que $\{Y(x) : x \in A\}$ é um processo estocástico

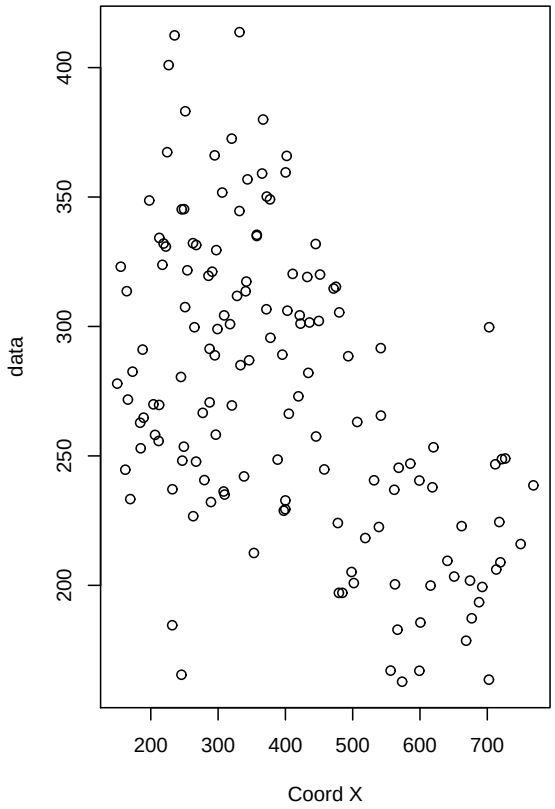
- x_i é tipicamente fixo. Se as localizações x_i são geradas por um processo estocástico pontual, assume-se que este processo é independente de $Y(x)$

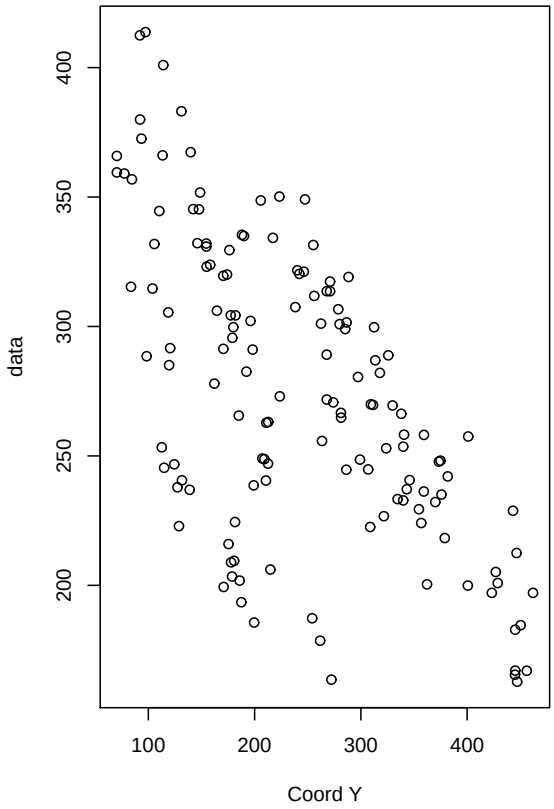
- objetivos científicos incluem a predição de um ou mais funcionais de processo (sem ruído) $\{S(x) : x \in A\}$

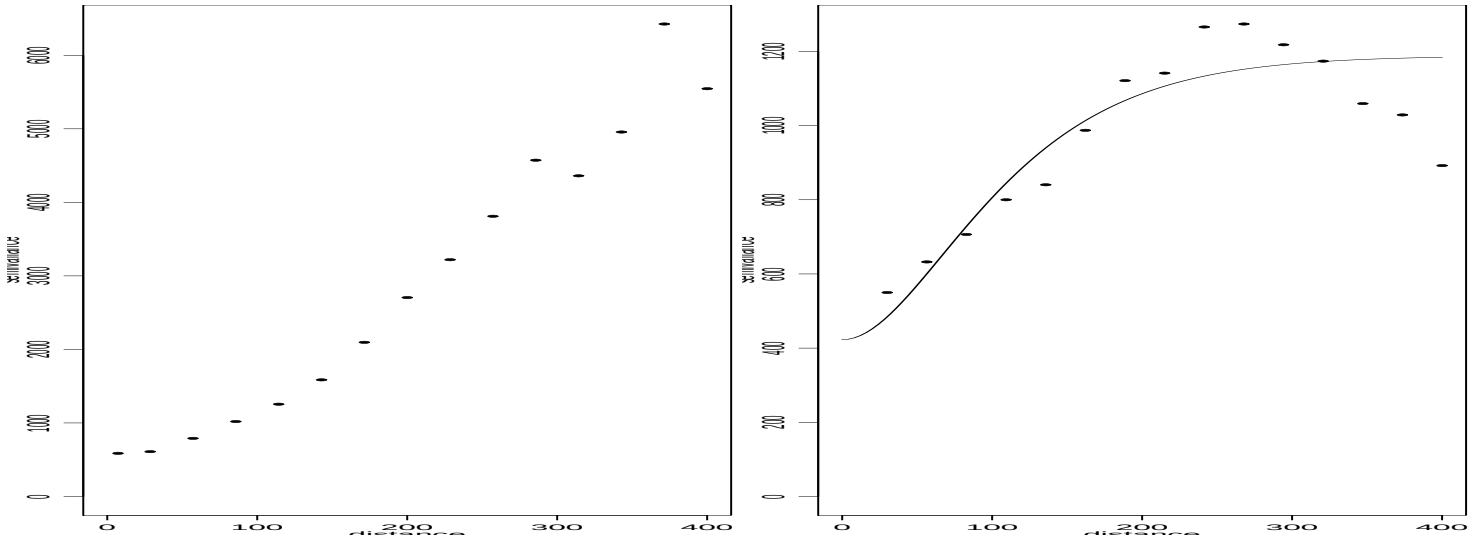
Exemplo básico: chuva no Paraná



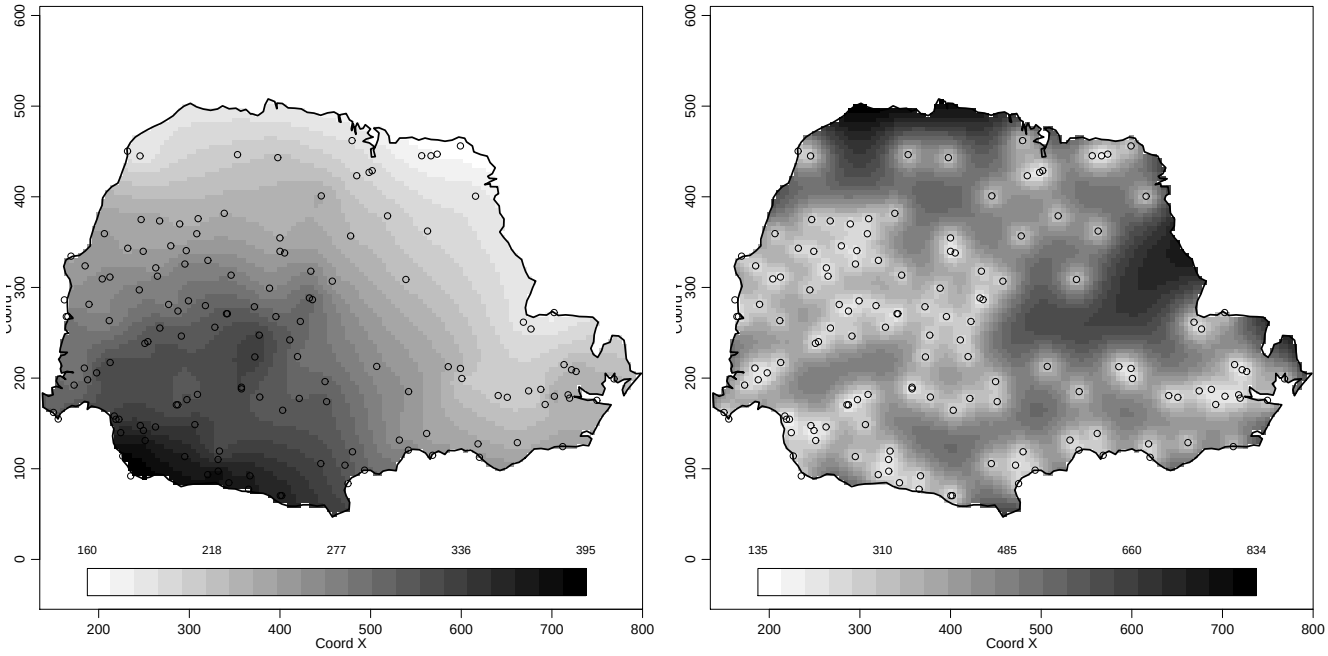








variogramas para dados originais (esquerda) e após retirada de tendência, com modelo ajustado (direita).



Krigagem: mapas de valores preditos (esquerda) e variâncias de predição (direita).

3. Questões Centrais

- **Delineamento**
  - quantas localizações?
  - quantas medidas?
  - configuração das localizações?
  - o que deve-se medir em cada localização?

- **Modelagem**
  - modelo probabilístico para o sinal $[S]$
  - modelo de probabilidade condicional para as medidas, $[Y|S]$

- **Estimação**
  - valores para parâmetros desconhecidos do modelo
  - inferências sobre os parâmetros ou funções destes

- **Predição**
  - avalia-se $[T|Y]$, a distribuição condicional aos dados do objetivo de predição

4. "Geoestatística baseada em modelos"

O termo "Geoestatística baseada em modelos" significa que adotamos um enfoque baseado em modelos para esta classe de problemas, o que quer dizer que começamos com um modelo estocástico explícito e derivamos métodos de estimação de parâmetros, interpolação e suavização através da aplicação de princípios gerais de estatística.

Notação

$$(Y_i, x_i) : i = 1, ..., n$$

- $\{x_i : i = 1, ..., n\}$ é o **plano amostral**
- $\{Y(x) : x \in A\}$ é o **processo de medida**
- $\{S(x) : x \in A\}$ é o **processo do sinal**
- $T = \mathcal{F}(S)$ é o **objetivo de predição**
- $[S, Y] = [S][Y|S]$ é o **modelo geoestatístico**

Modelo linear generalizado linear clássico

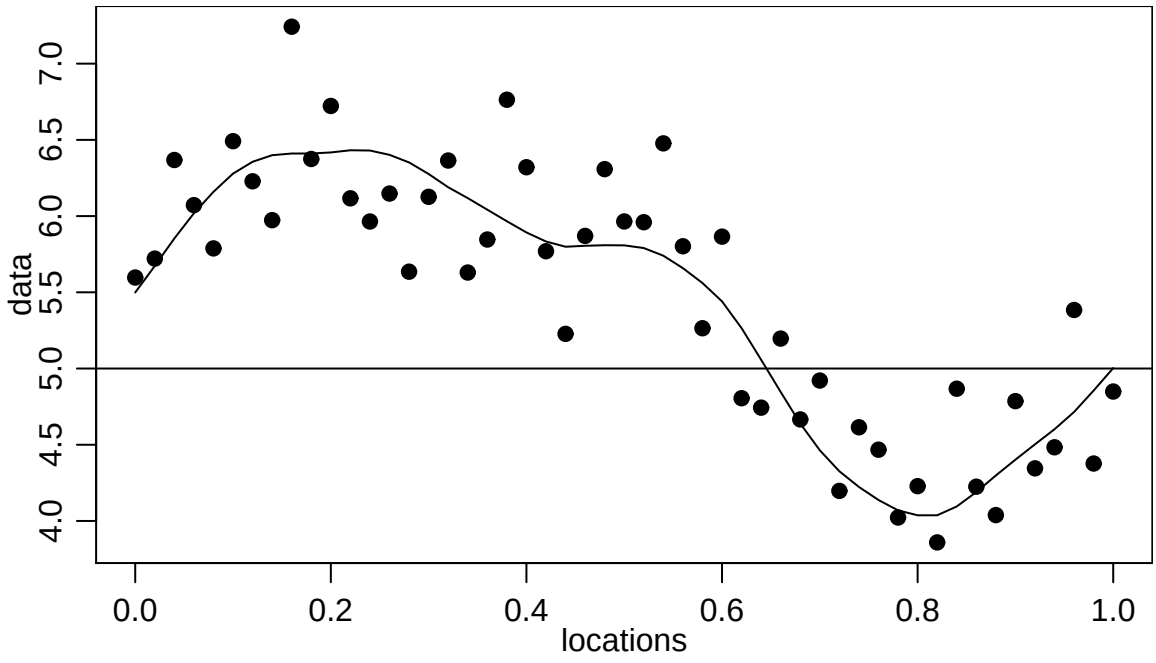
- $Y_i : i = 1, ..., n$
  mutuamente independentes, com $\mu_i = \mathrm{E}[Y_i]$
- $h(\mu_i) = \sum_{j=1}^{k} f_{ij}\beta_j$, com função de ligação conhecida $h(\cdot)$.

Modelo Linear Generalizado Mixto

- $Y_i : i = 1, ..., n$
  mutuamente independentes, com $\mu_i = \mathrm{E}[Y_i]$, conditional às realizações de um conjunto de de variáveis aleatórias latentes U_i,
- $h(\mu_i) = \sum_{j=1}^{k} f_{ij}\beta_j + U_i$,
  para uma função de ligação conhecida $h(\cdot)$.

A modelo espacial (geoestatístico)

- $Y_i : i = 1, ..., n$
  mutuamente independentes, com $\mu_i = \mathrm{E}[Y_i]$, conditional às realizações de um conjunto de de variáveis aleatórias latentes U_i,
- $h(\mu_i) = U_i + \sum_{j=1}^{p} f_{ij}\beta_j$,
  para uma função de ligação conhecida $h(\cdot)$,
- $U_i = S(x_i)$
  onde $\{S(x) : x \in \mathbb{R}^2\}$ é um processo estocástico espacial.
- $h(\mu_i) = \sum_{j=1}^{k} f_{ij}\beta_j + U_i$,



simulação ilustrando os componentes do modelo: dados $Y(x_i)$ (pontos), sinal $S(x)$ (linha curva) e média μ (linha horizontal).

5. O Modelo Gaussiano

(a) $S(\cdot)$ é um processo Gaussiano estacionário com

i. $\mathrm{E}[S(x)] = 0$,

ii. $\mathrm{Var}\{S(x)\} = \sigma^2$

iii. $\rho(u) = \mathrm{Corr}\{S(x), S(x-u)\}$;

(b) a distribuição condicional de Y_i dado $S(\cdot)$ é Gaussiana com média $\mu + S(x_i)$ e variância τ^2;

(c) $Y_i : i = 1, ..., n$ são mutuamente independentes, condicional à $S(\cdot)$.

Uma formulação equivalente para o modelo Gaussiano:

$$Y_i = \mu + S(x_i) + Z_i : i = 1, ..., n.$$

onde $Z_i : i = 1, ..., n$ são mutuamente independentes e identicamente distribuídos com $Z_i \sim \mathrm{N}(0, \tau^2)$.

Desta forma a distribuição conjunta de Y é multivariada Normal,

$$Y \sim \mathrm{MVN}(\mu \mathbf{1}, \sigma^2 R + \tau^2 I)$$

onde:

$\mathbf{1}$ denota um vetor de 1's com n elementos

I é matrix identidade $n \times n$

R é uma matrix $n \times n$ com $(i, j)^{th}$ elemento $\rho(u_{ij})$ onde $u_{ij} = ||x_i - x_j||$, é distancia Euclideana entre x_i e x_j.

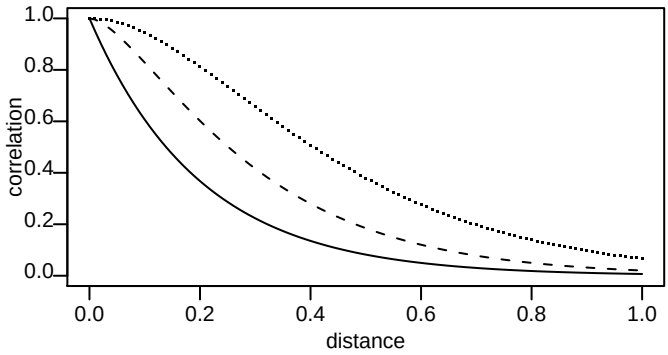
6. Especificação da função de correlação

A família de Matérn

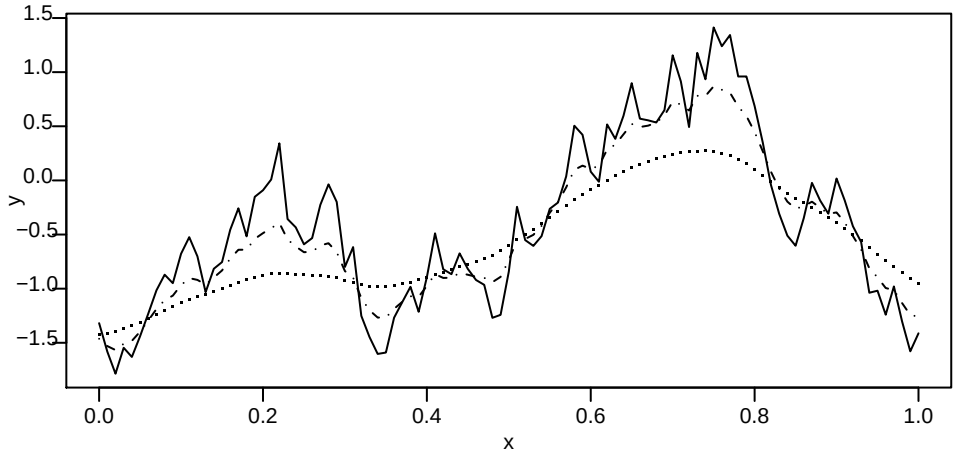
Função de correlação dada por

$$\rho(u) = \{2^{\kappa-1}\Gamma(\kappa)\}^{-1}(u/\phi)^{\kappa}K_{\kappa}(x/\phi)$$

- κ e ϕ são parâmetros
- $K_{\kappa}(\cdot)$ denota função de Bessel de ordem κ
- válida para $\phi > 0$ e $\kappa > 0$.
- $\kappa = 0.5$: *modelo exponencial*
- $\kappa \to \infty$: *modelo Gaussiano*
- $S(x)$ é $\lceil \kappa - 1$ vezes diferenciável



Três exemplos de funções de Matérn com $\phi = 0.2$ and $\kappa = 1$ (linha sólida), $\kappa = 1.5$ (linha interrompida) and $\kappa = 2$ (pontos).



simulações de processos em 1-D com funções de correlação de de Matérn com $\phi = 0.2$ e $\kappa = 0.5$ (linha sólida), $\kappa = 1$ (linha interrompida) and $\kappa = 2$ (linha pontilhada).

7. Extensões do modelo básico

(a) Modelos Gaussianos transformados

- O modelo Gaussiano é claramente inapropriado para distribuições assimétricas.

- Certos dados podem indicar relações entre média e variância, que violam o modelo Gaussiano.

- Parâmetro extra λ da transformação Box-Cox introduz certa flexibilidade.

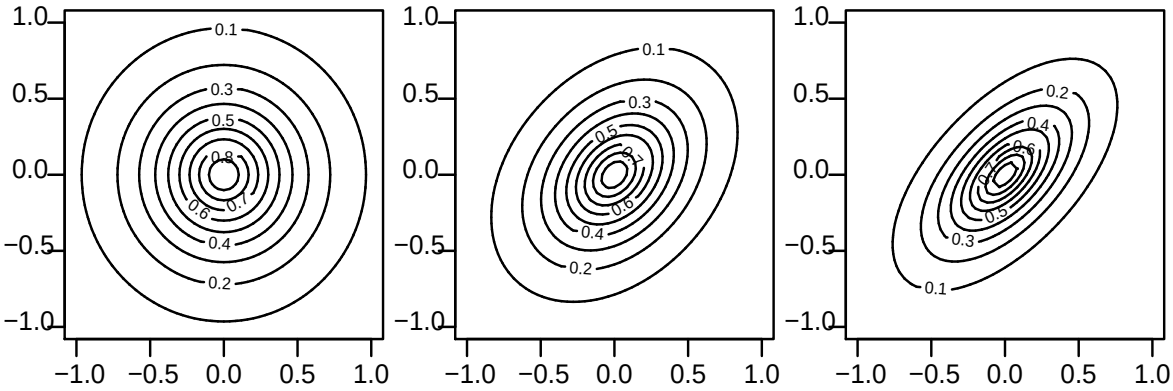
- O modelo fica então definido da forma:
  - assume-se $Y^* \sim MVN(F\beta, \sigma^2 V)$

  - dados $y = (y_1, ..., y_n)$, são gerados por uma transformação do modelo linear Gaussiano $Y = h_\lambda^{-1}(Y^*)$ tal que:

$$Y_i^* = h_\lambda(Y) = \begin{cases} \frac{(y_i)^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \log(y_i) & \text{if } \lambda = 0 \end{cases}$$

(b) Efeitos Direcionais

- Condições ambientais podem induzir efeitos direcionais (vento, formação do solo, etc)
- como consequência a correlação espacial pode variar com a direção

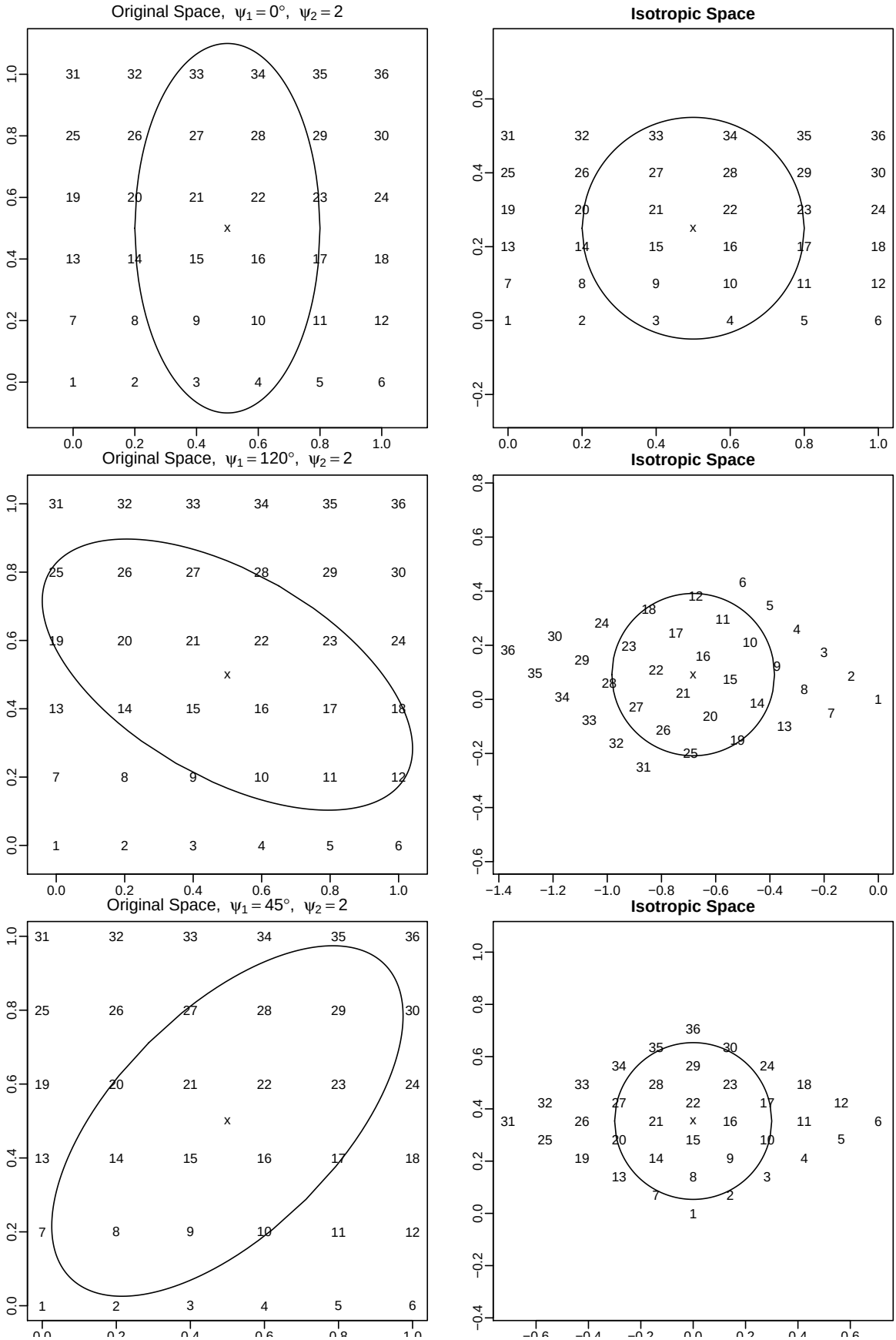


contornos de correlação para modelo isotrópico (esq.) e dois modelos anisotrópicos (centro e dir.)

- *anisotropia geométrica*: possível (e simples) abordagem.
- dois parâmetros extra: *ângulo de anisotropia* ψ_A e *razão de anisotropia* ψ_R.
- rotação e contração/expansão das coordenadas originais:

$$(x_1\prime, x_2\prime) = (x_1, x_2) \begin{pmatrix} \cos(\psi_A) & -\sin(\psi_A) \\ \sin(\psi_A) & \cos(\psi_A) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\psi_R} \end{pmatrix}$$

“Correção” de anisotropia geométrica



(c) Modelos não estacionários

- *Modelos com médias não constantes (ou, incluindo covariáveis)*

  Substituir a média constante μ por

  $$\mu(x) = F\beta = \sum_{j=1}^{k} \beta_j f_j(x)$$

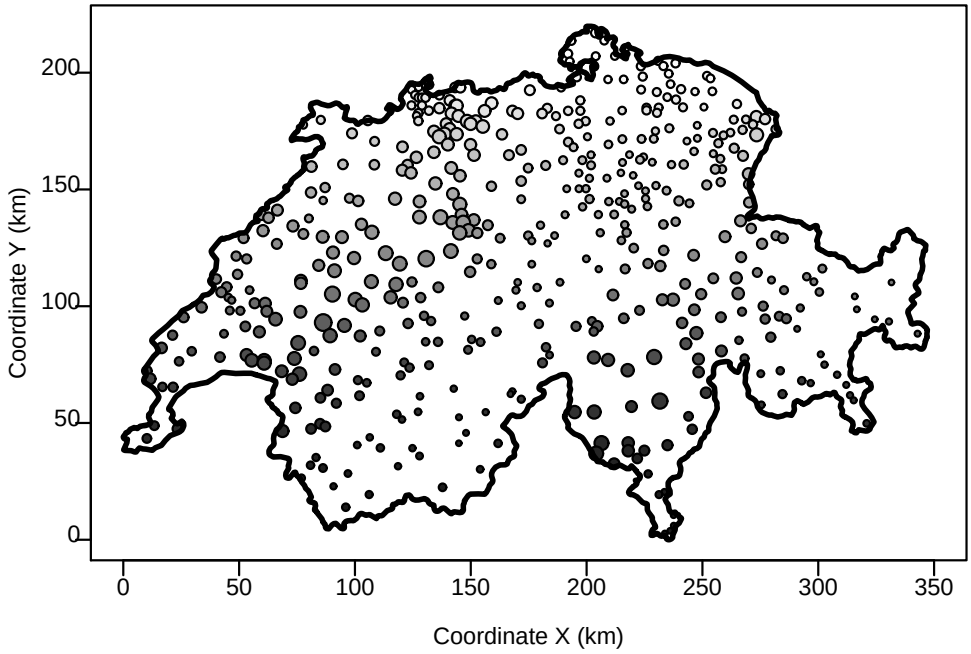
  para medidas $f_j(x)$ das covariáveis (lineares ou não lineares).

- *Variação aleatória não estacionária*

  Variabilidade **intrínsica**: pressuposto mais fraco de estacionaridade (processo com incrementos estacionários, como passeios aleatórios em séries temporais), largamente utilizados como modelo padrão para variação espacial discreta (Besag, York and Molié, 1991).

  Métodos de **deformação espacial** (Sampson and Guttorp, 1992) buscam estacionaridade por transformações (complexas) do espaço geográfico, x.
  É preciso ter em mente o balanço entre a o aumento da flexibilidade de modelos mais gerais contra a sobre-modelagem de dados esparsos, que leva a pobre identificação dos parâmetros.

8. Estudo de caso: chuva na Suíça



Localizações com tamanho dos pontos proporcional aos valores observados. Distâncias em kilometros

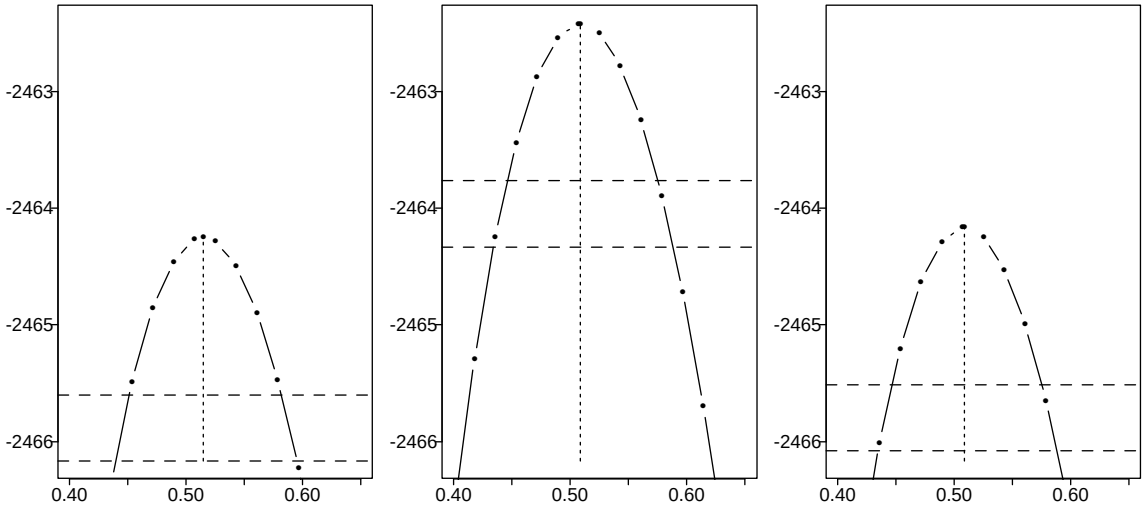
- 467 localizações
- medidas de precipitação em 8 de Maio 1986
- dados são valores inteiros com unidade de medida igual á 1/10 mm
- 5 localizações com valores iguais à zero.

chuva na Suíça (cont.)

Estimação parâmetros de transformação e suavidade (modelo de Matérn)

κ	$\hat{\lambda}$	$\log \hat{L}$
0.5	0.514	-2464.246
1	0.508	-2462.413
2	0.508	-2464.160

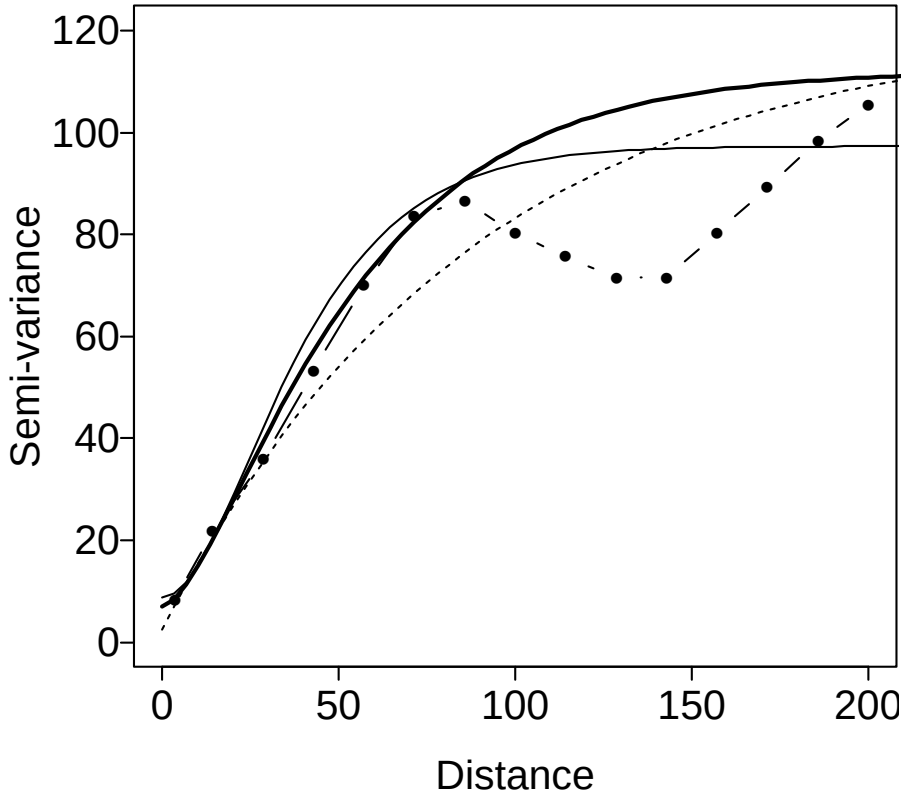
Estimativas de MV de $\hat{\lambda}$ e valores da log-verossimilhança $\log \hat{L}$ para diferentes valores de κ.



Verossimilhanças prefilhadas para λ. esquerda: $\kappa = 0.5$, meio: $\kappa = 1$, direita: $\kappa = 2$.

transformação logarítimica ou nãotransformação são claramente NÃO indicadas!

chuva na Suíça (cont.)



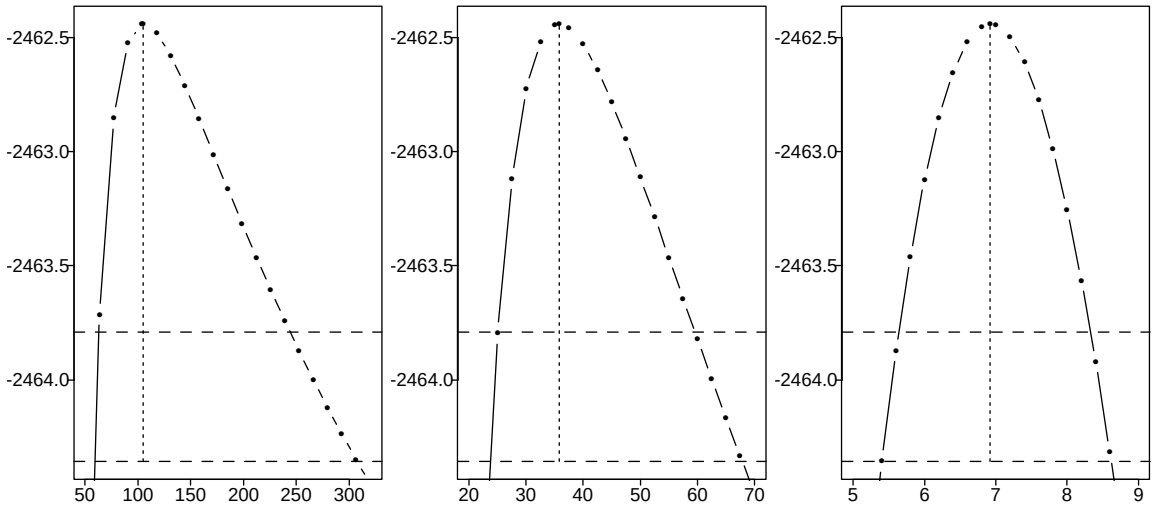
semivariograma empírico para dados transformados e variogramas teóricos com estimativas de MV para $\kappa = 0.5$ (linha interrompida), $\kappa = 1$ (linha grossa), $\kappa = 2$ (linha fina).

chuva na Suíça (cont.)

Estimativas para modelo com $\lambda = 0.5$

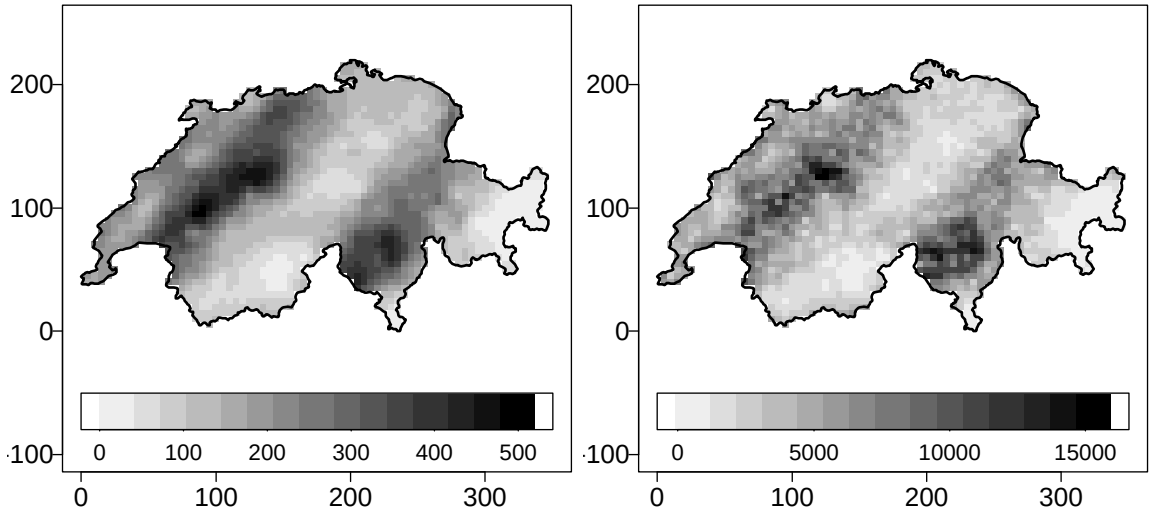
κ	$\hat{\beta}$	$\hat{\sigma}^2$	$\hat{\phi}$	$\hat{\tau}^2$	$\log \hat{L}$
0.5	18.36	118.82	87.97	2.48	-2464.315
1	20.13	105.06	35.79	6.92	-2462.438
2	21.36	88.58	17.73	8.72	-2464.185

Maximum likelihood estimates $\hat{\beta}$, $\hat{\phi}$, $\hat{\sigma}$, $\hat{\tau}$ and the corresponding value of the likelihood function $\log \hat{L}$ for different values of the Matérn parameter κ, for $\lambda = 0.5$



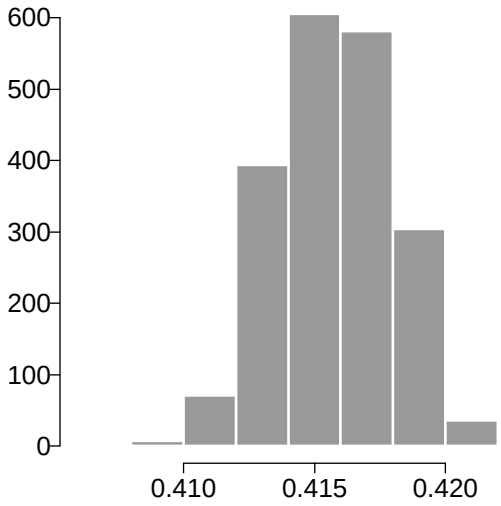
Verossimilhança perfilhada para parâmetros de covariância $\kappa = 1$ and $\lambda = 0.5$. esquerda: σ^2, meio: ϕ, direita: τ^2.

chuva na Suíça (cont.)



Mapas com predições (esquerda) e variâncias de predição (direita).

Predição da percentagem da área onde $Y(x) \geq 200$: $\tilde{A}_{200}$ é de 0.4157



Amostras da preditiva de $\tilde{A}_{200}$.

PARTE IV:

ESTIMAÇÃO DE PARÂMETROS

1. Propriedades do Segundo Momento

2. Estimação usando Variogramas

3. Estimação por Verossimilhança

4. Predição “plug-in”

5. Estudo de Caso

6. Comentários e Extensões

1. Propriedades do segundo momento

- o **variograma** é a função

$$V(x, x') = \frac{1}{2}\mathrm{Var}\{Y(x) - Y(x')\}$$

- para $u = ||x - x'||$,

$$V(u) = \tau^2 + \sigma^2\{1 - \rho(u)\}$$

- os paramêtros estruturais básicos são
  - *efeito pepita ("nugget")*: τ^2
  - *patamar ("sill")*: $\tau^2 + \sigma^2 = \mathrm{Var}\{Y(x)\}$
  - *o alcance ("range")*: ϕ, tal que $\rho(u) = \rho_0(u/\phi)$

- implicações práticas:
  - qualquer versão razoável do modelo linear Gaussiano tem pelo menos três parâmetros de covariância
  - um volume da dados substancial pode ser necessário para estimar maior número de parâmetros
  - a família **Matérn** possui um parâmetro extra para determinar a suavidade do processo $S(x)$

Paradigmas para estimação

- **Métodos "Ad-hoc" (baseados em variogramas)**
  - calcule o variograma empírico
  - ajuste um modelo teórico de variograma

- **Métodos baseados na verossimilhança**
  - tipicamente sob pressupostos de Gaussianidade
  - Ótimos sob as condições declaradas
  - maior demanda computacional
  - podem não ser robustos

- **Implementação Bayesiana**,
  - estimação e predição combinadas
  - cada vez mais aceitos

2. Estimação usando Variogramas

- O variograma é definido por

$$V(x, x') = \frac{1}{2}\mathrm{Var}\{Y(x) - Y(x')\}$$

- se $Y(x)$ é estacionário,

$$V(x, x') = V(u) = \frac{1}{2}\mathrm{E}[\{Y(x) - Y(x')\}^2]$$

  onde $u = ||x - x'||$

- sugere uma estimativa empírica para $V(u)$:

$$\hat{V}(u) = \mathrm{average}\{[y(x_i) - y(x_j)]^2\}$$

  onde cada média é tomada entre todos os pares $[y(x_i), y(x_j)]$ tal que $||x_i - x_j|| \approx u$

- para processo com média não constante a tendência pode ser removida:
  - defina $r_i = Y_i - \hat{\mu}(x_i)$
  - defina $\hat{V}(u) = \mathrm{average}\{(r_i - r_j)^2\}$,
    onde cada média é tomada entre todos os pares (r_i, r_j)

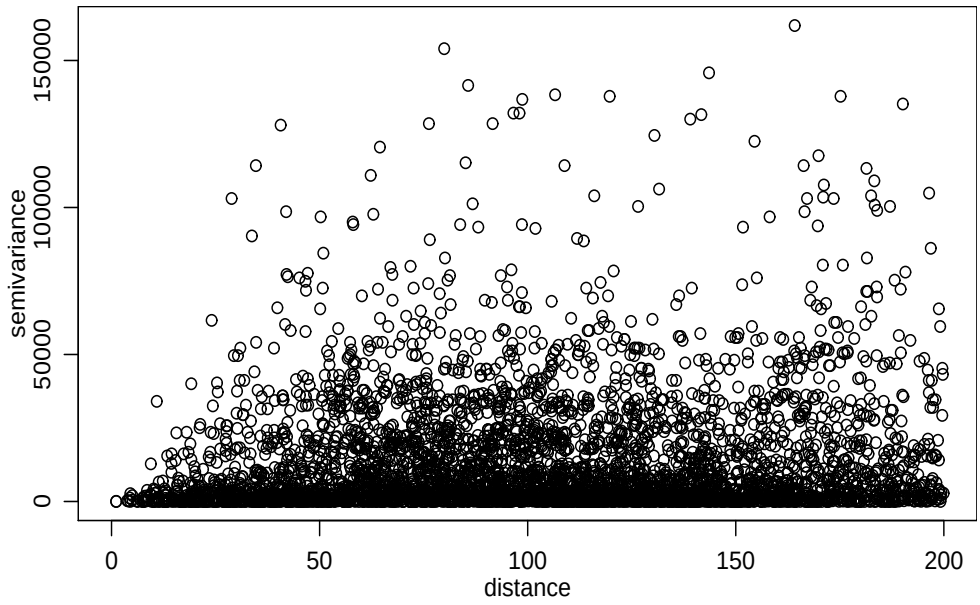
(a) A nuvem variográfica

- defina as quantidades:

$$\begin{aligned} r_i &= Y_i - \hat{\mu}(x_i) \\ u_{ij} &= ||x_i - x_j|| \\ v_{ij} &= \frac{(r_i - r_j)^2}{2} \end{aligned}$$

- a **nuvem de variograma** é um gráfico de pontos (u_{ij}, v_{ij})

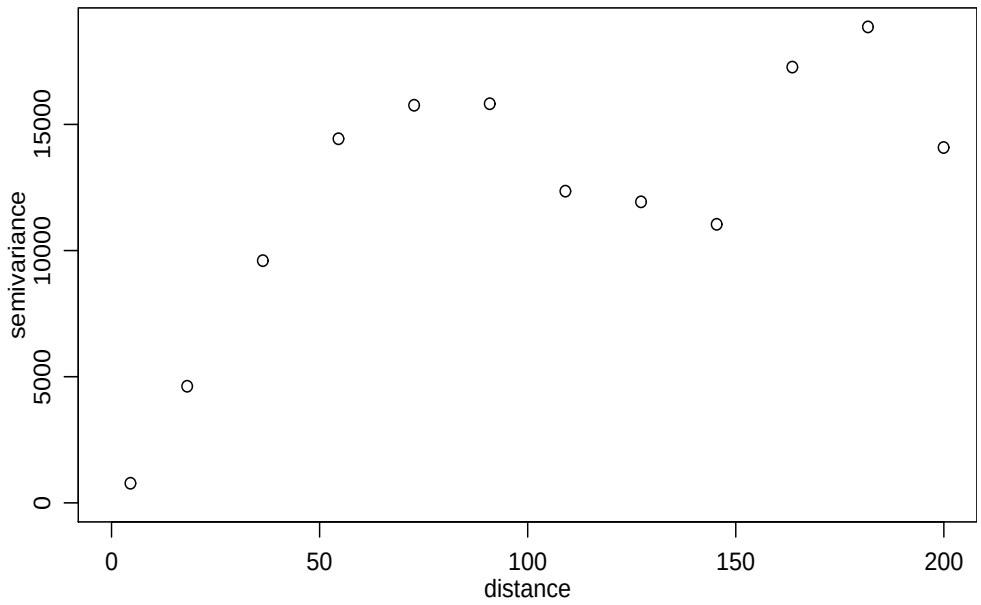
Exemplo: Dados de chuva na Suíça



(b) O variograma empírico

- obtido a partir da nuvem de variograma tomando médias de classes de distância: $u - h/2 \leq u_{ij} < u + h/2$
- forma k classes de distâncias, cada uma média de n_k pares,
- muito sensível à especificação de média do processo $\mu(x)$

Exemplo: Dados de chuva na Suíça



Variograma empírico

(c) O ajuste de variogramas

Estime os parâmetros $\tilde{\theta}$ minimizando um particular critério

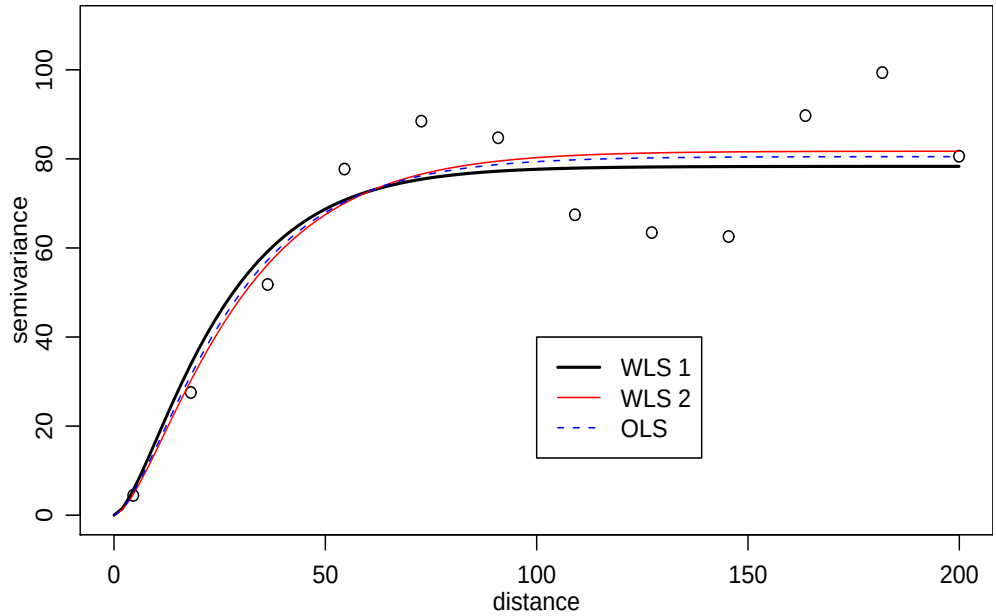
por exemplo, mínimos quadrados generalizados (Cressie, 1993)

$$S(\theta) = \sum_k n_k \{[\bar{V}_k - V(u_k;\theta)]/V(u_{ij};\theta)\}^2$$

onde $\bar{V}_k$ é a média das n_k ordenadas v_{ij} do variograma.

Outros critérios: OLS, WLS com diferentes pesos, GLS, quasi-verossimilhança.

Exemplo: Dados de chuva na Suíça

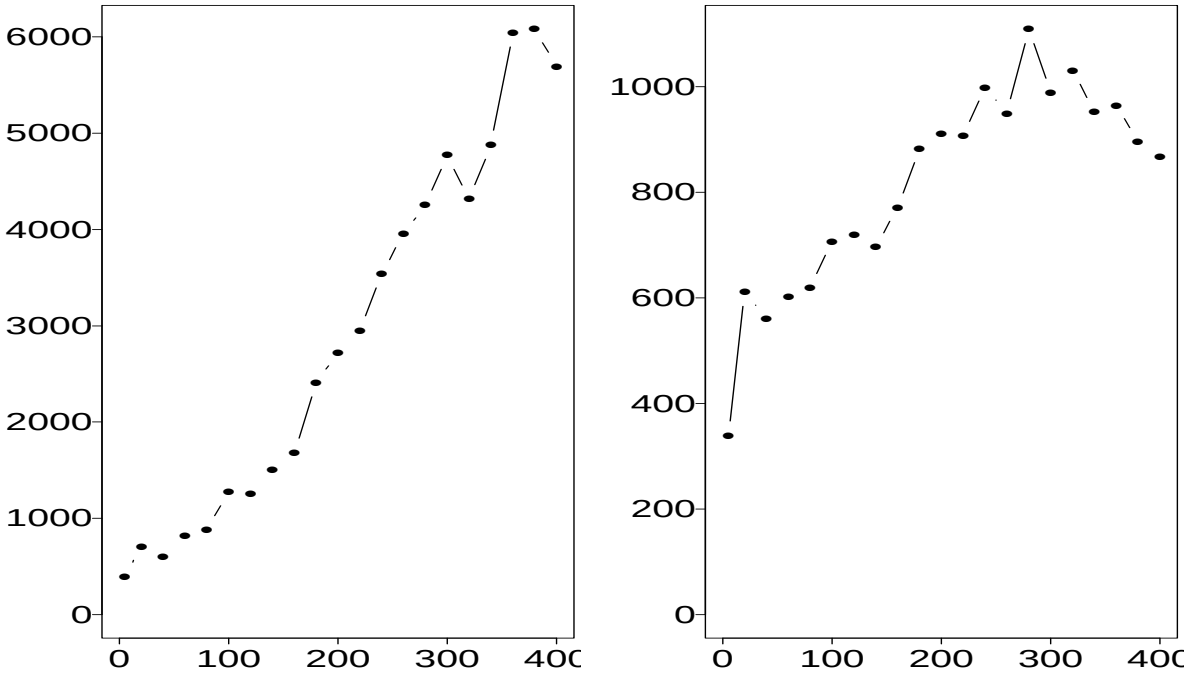


Ajuste de variograma empírico por WLS com pesos dados por n_k apenas (linha grossa), WLS com pesos sugeridos pro Cressie (linha cheia) o OLS (linha pontilhada)

(d) Comentários sobre uso de variogramas para inferência

i. Variogramas dos dados originais e resíduos podem ser muito diferentes

Exemplo: Dados do Paraná



variogramas empíricos dos dados originais (esq.) e resíduos após regressão em latitude, longitude e altitude (dir.)

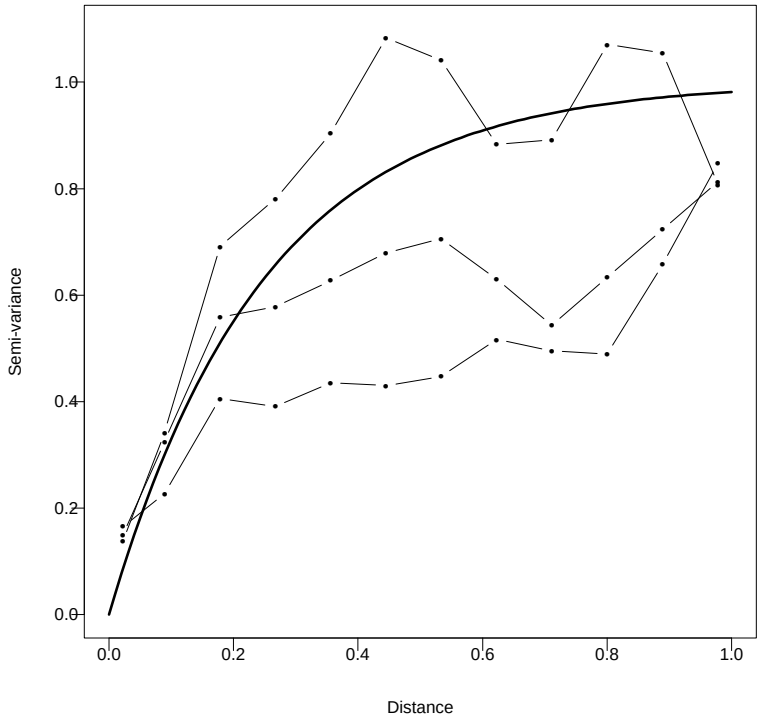
- variograma de dados originais reflete (inclui) variação de tendência geográfica de larga escala.

- variograma de resíduos elimina esta fonte de variação

ii. Quão instáveis são os variogramas empíricos?

- sob o modelo linear Gaussiano:
  - $v_{ij} \sim V(u_{ij})\chi^2_1$
  - the v_{ij} são correlacionados

iii. variogramas de diferentes realizações do mesmo processo podem ser bem diferentes



- linha sólida mostra o variograma verdadeiro do processo
- linha finas mostram variogramas empíricos de três realizações do mesmo processo (modelo)
- as altas correlações entre $\hat{V}(u)$ para sucessivos u conferem uma suavidade "enganadora"

iv. Ajuste de modelos por mínimos quadrados
   - corresponde a um sistema de equações de estimação que produz estimativas viciadas de θ,
   - mesmo assim é largamente utilizado na prática
   - potencialmente "perigoso" devido as correlações inerentes aos sucessivos $\bar{V}_k$'s

v. estimação baseada em objeto que é por si estimado

vi. é possivel ajustar o modelo diretamente aos **dados** e não ao variograma.

3. Estimação por verossimilhança

O modelo Gaussiano é dado por:

$$Y_i|S \sim \mathrm{N}(S(x_i), \tau^2)$$

- $S(x_i) = \mu(x_i) + S_c(x_i)$
- $S_c(\cdot)$ é um processo estocástico Gaussiano com parâmetros de covariância (σ^2, ϕ, κ),
- $\mu(x_i) = F\beta = \sum_{j=1}^{k} f_k(x_i)\beta_k$, onde $f_k(x_i)$ é vetor de covariáveis na localização x_i

$$Y \sim \mathrm{MVN}(F\beta, \sigma^2 R + \tau^2 I)$$

e a função de verossimilhaça é:

$$\begin{aligned}\ell(\beta, \tau, \sigma, \phi, \kappa) \propto \quad & -0.5\{\log|(\sigma^2 R + \tau^2 I)| + \\ & (y - F\beta)'(\sigma^2 R + \tau^2 I)^{-1}(y - F\beta)\}.\end{aligned}$$

para qual maximização (numérica) produz as estimativas de máxima verossimilhança

Modelos Gaussianos transformados

A log-verossimilhança é:

$$\begin{aligned}\ell(\beta, \theta, \lambda) = \; - \; & \frac{1}{2}\{\log|\sigma^2 V| \\ & +(h_\lambda(y) - F\beta)'\{\sigma^2 V\}^{-1}(h_\lambda(y) - F\beta)\} \\ & + \sum_{i=1}^{n} \log\left((y_i)^{\lambda-1}\right)\end{aligned}$$

para $\theta = (\tau, \sigma, \phi, \kappa)$.

ML: tipicamente usada sob pressupostos de normalidade

- estimativas ótimas sob os pressupostos declarados

- porém computacionalmente caros e podem não ser robustos

- dificuldades computacionais para grande número de dados

- Implementação Bayesiana combinando estimação e predição tem sido cada vez mais aceita (ao menos entre estatísticos!).

- Para família de Matérn considere tomar κ em um conjunto discreto $\{0.5, 1, 2, 3, ..., N\}$

4. Predição "plug-in"

Em geral o interesse está em predizer

- o valor da realização do processo $S(\cdot)$ em um ponto
- ou a média de $S(\cdot)$ em uma região

$$T = |B|^{-1} \int_B S(x)dx$$

onde $|B|$ denota a área da região B.

Para o modelo Gaussiano o preditor de minímos quadrados de $T = S(x)$ é:

$$\hat{T} = \mu + \sigma^2 \mathbf{r}'(\tau^2 I + \sigma^2 R)^{-1}(Y - \mu \mathbf{1})$$

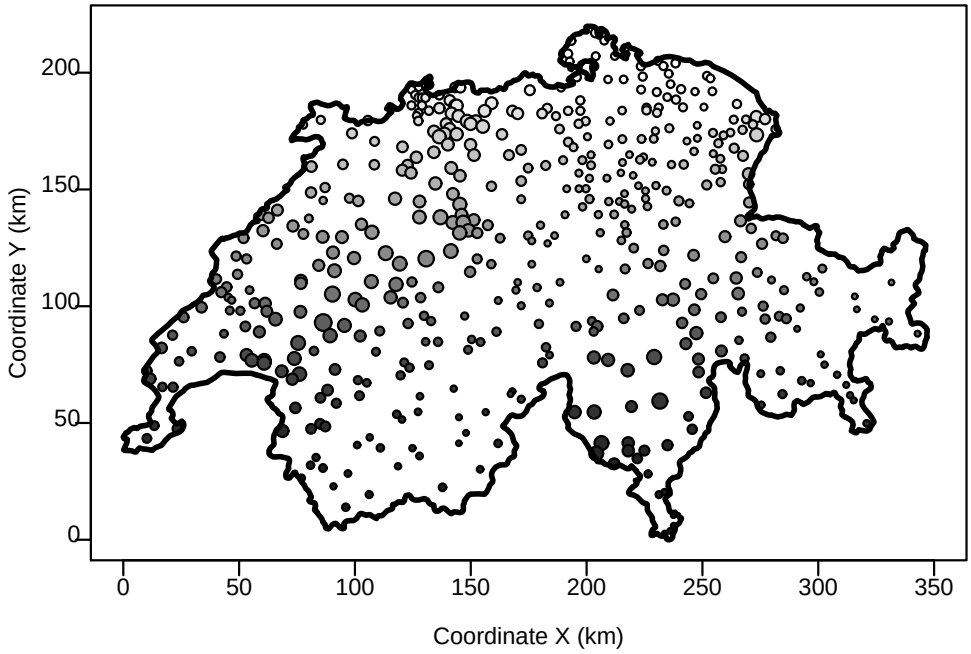
e a variância de predição

$$\mathrm{Var}(T|Y) = \sigma^2 - \sigma^2 \mathbf{r}'(\tau^2 I + \sigma^2 R)^{-1}\sigma^2 \mathbf{r}$$

onde os únicos termos desconhecidos são os parâmetros do modelo

A **predição "plug-in"** consiste em substituir os parâmetros por suas estimativas.

5. Estudo de caso: chuva na Suíça



Localizações com tamanho dos pontos proporcional aos valores observados. Distâncias em kilometros

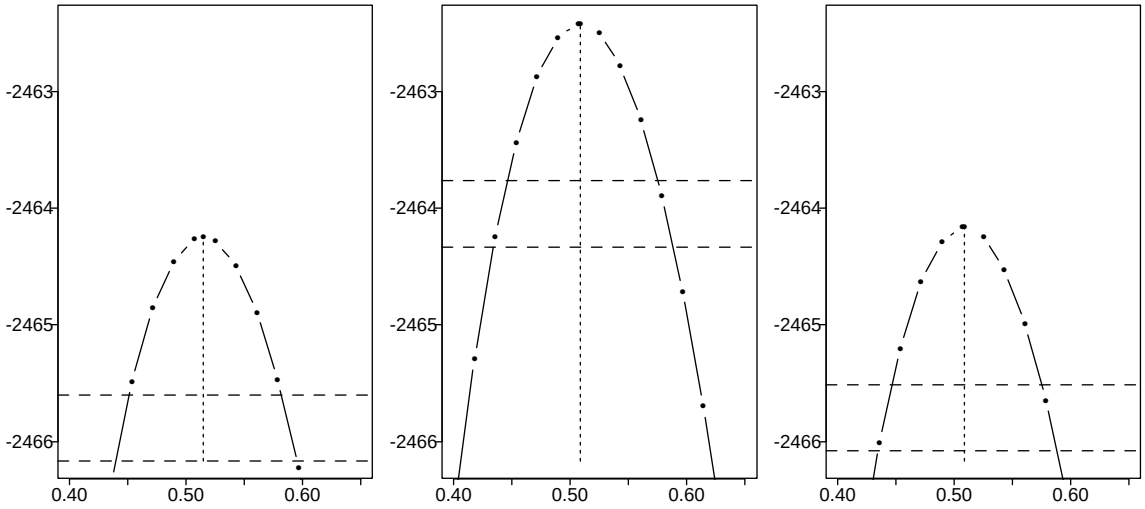
- 467 localizações
- medidas de precipitação em 8 de Maio 1986
- dados são valores inteiros com unidade de medida igual á $1/10$ mm
- 5 localizações com valores iguais à zero.

chuva na Suíça (cont.)

Estimação parâmetros de transformação e suavidade (modelo de Matérn)

κ	$\hat{\lambda}$	$\log \hat{L}$
0.5	0.514	-2464.246
1	0.508	-2462.413
2	0.508	-2464.160

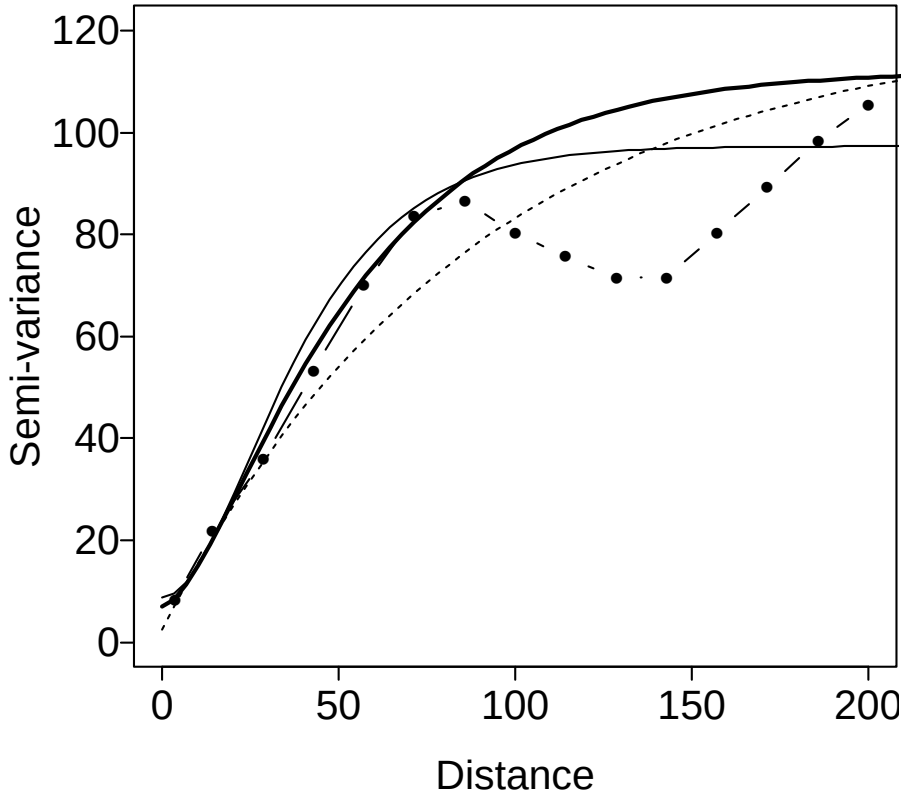
Estimativas de MV de $\hat{\lambda}$ e valores da log-verossimilhança $\log \hat{L}$ para diferentes valores de κ.



Verossimilhanças prefilhadas para λ. esquerda: $\kappa = 0.5$, meio: $\kappa = 1$, direita: $\kappa = 2$.

transformação logarítimica ou nãotransformação são claramente NÃO indicadas!

chuva na Suíça (cont.)



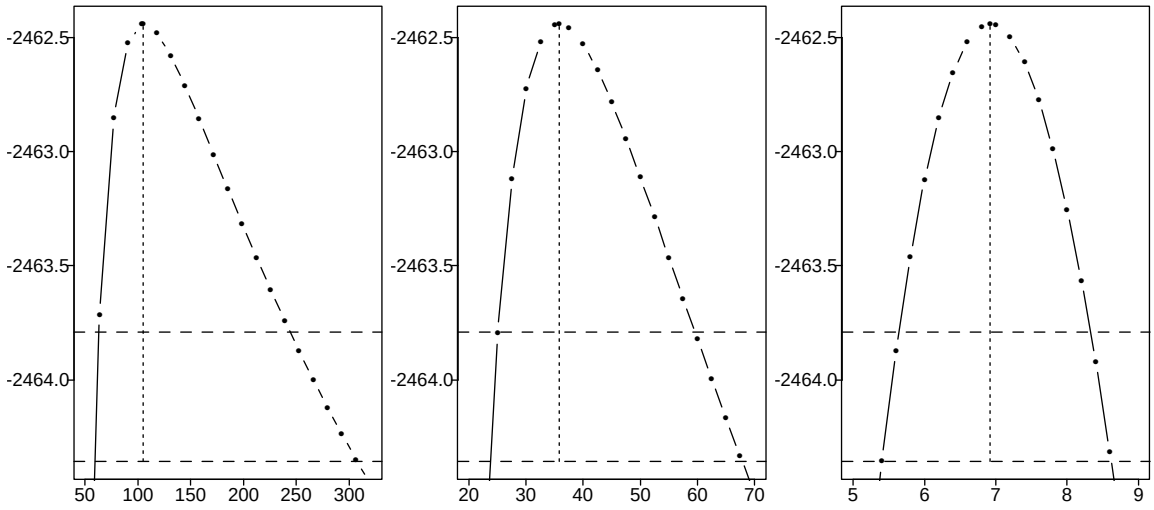
semivariograma empírico para dados transformados e variogramas teóricos com estimativas de MV para $\kappa = 0.5$ (linha interrompida), $\kappa = 1$ (linha grossa), $\kappa = 2$ (linha fina).

chuva na Suíça (cont.)

Estimativas para modelo com $\lambda = 0.5$

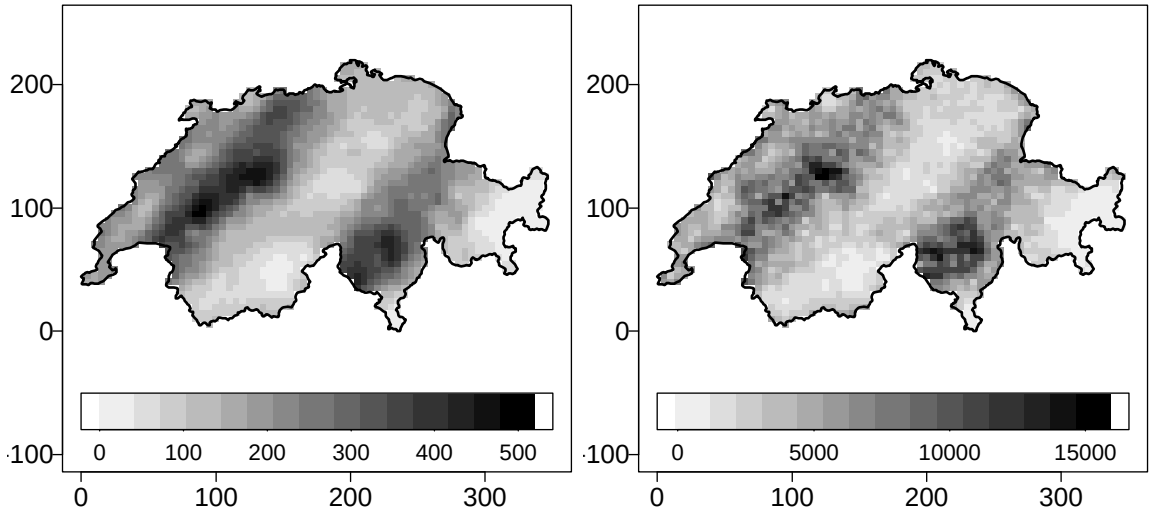
κ	$\hat{\beta}$	$\hat{\sigma}^2$	$\hat{\phi}$	$\hat{\tau}^2$	$\log \hat{L}$
0.5	18.36	118.82	87.97	2.48	-2464.315
1	20.13	105.06	35.79	6.92	-2462.438
2	21.36	88.58	17.73	8.72	-2464.185

Maximum likelihood estimates $\hat{\beta}$, $\hat{\phi}$, $\hat{\sigma}$, $\hat{\tau}$ and the corresponding value of the likelihood function $\log \hat{L}$ for different values of the Matérn parameter κ, for $\lambda = 0.5$



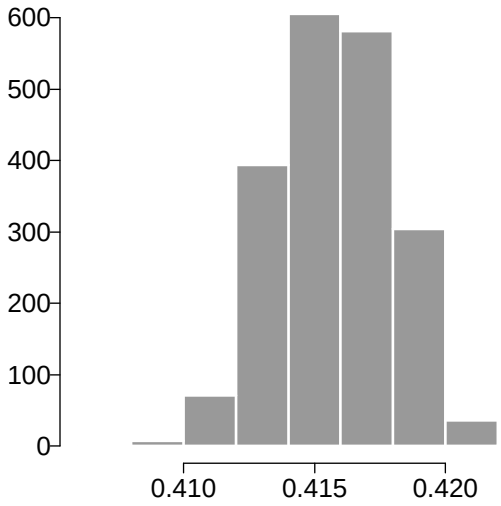
Verossimilhança perfilhada para parâmetros de covariância $\kappa = 1$ and $\lambda = 0.5$. esquerda: σ^2, meio: ϕ, direita: τ^2.

chuva na Suíça (cont.)



Mapas com predições (esquerda) e variâncias de predição (direita).

Predição da percentagem da área onde $Y(x) \geq 200$: $\tilde{A}_{200}$ é de 0.4157



Amostras da preditiva de $\tilde{A}_{200}$.

6. Outros tópicos e extensões

- Máxima verossimilhança restrita (REML)
- Verossimilhanças perfilhadas
- Estimação de modelos anisotrópicos
- Validação dos modelos
- Modelos não estacionários
  - tendências e/ou covariáveis
    (a) Predição *ad-hoc*:
    (b) estima β por OLS, $\tilde{\beta} = (F'F)^{-1}F'Y$, e calcula-se resíduos $Z = Y - F\tilde{\beta}$.
    (c) calcula-se o variograma empírico (dos resíduos) que é utilizado para formulação do modelo e estimação de parâmetros
    (d) reestima-se β por GLS e usa-se modelo ajustado para predição

    Nota: **krigagem universal** ou **krigagem com tendência externa**

  - Relações funcionais entre médias e variâncias
  - variação aleatória não estacionária
    Intrínseca Campos aleatórios de Markov (Besag, York and Molié, 1991).
    Deformações espaciais Sampson and Guttorp, 1992 tentam obter estacionaridade através de transformações não lineares do espaço geográfico x.
    Ver tese de Alexandra Smith (2001).
  - flexibilidade *vs* identificabilidade

- papel dos variogramas empíricos
  - diagnóstico (abordagem "model-based")
  - ferramenta de inferência (abordagem tradicional)

- ambas abordagens anexam estimativas dos parametros ao modelo como se fosse valores verdadeiros.
  Predição "plug-in"
  - usualmente produz boas estimativas pontuais de $T = S(x)$
  - em geral sub-estima variância de predição
  - pode produzir estimativas inacuradas de outras quantidades objetivo T

PARTE V:

INFERÊNCIA BAYESIANA PARA O MODELO GAUSSIANO

1. Inferência Bayesiana

2. Resultados para o modelo Gaussiano

3. Estudo de Caso: Dados da Suíça

1. Análise Bayesiana - Conceitos Básicos

Inferência Bayesiana considera a incerteza nos parâmetros tratando-os como variáveis aleatórias, e expressando inferências sobre parâmetros em termos de sua distribuição condicional, dadas todos os dados observados.

Para inferências sobre parâmetros do modelo, a especificação do modelo deve agora incluir os parâmetros.

$$[Y, \theta] = [\theta][Y|\theta]$$

O teorema de Bayes nos permite calcular:

$$[Y, \theta] = [Y|\theta][\theta] = [Y][\theta|Y]$$

Portanto

$$[\theta|Y] = [Y|\theta][\theta]/[Y]$$

é a *distribuição a posteriori* onde

$$[Y] = \int [Y|\theta][\theta]d\theta.$$

O paradigma Bayesiano

(a) **Modelo**

- a especificação completa do modelo consiste de
  $[Y, \theta] = [Y|\theta][\theta]$.
- formular um modelo para variável observada Y.
- este modelo define $[Y|\theta]$ (e então a xpressão para a log-verossimilhança $\ell(\theta; Y)$)

(b) **Priori**

- antes de observar Y, a marginal $[\theta]$ expressa a incerteza sobre θ
- chama-se $[\theta]$ a *distribuição a priori* para θ

(c) **Posteriori**

- tendo observado Y, este não é mais uma quantidade desconhecida (variando aleatoriamente)
- desta forma revisamos a incerteza acerca de θ condicionando nos valores observados de Y
- chama-se $[\theta|Y]$ *distribuição a posteriori* para θ, que é usada para inferências

NOTA: a verossimilhança tem papel central em ambas, inferência clássica e Bayesiana

Predição

Tratando θ como v.a. a inferência Bayesiana não faz distinção formal entre problemas de estimação de parâmetros e predição e desta forma fornece uma forma natural de incorporar a incerteza sobre os parâmetros na inferência preditiva

A idéia geral é formular um modelo para

$$[Y, T, \theta] = [Y, T|\theta][\theta]$$

e fazer inferências baseadas na distribuição condicional

$$\begin{aligned}[T|Y] &= \int [T, \theta|Y] d\theta \\ &= \int [\theta|Y][T|Y, \theta] d\theta\end{aligned}$$

Comparando "plug-in" e Bayesiana

- a predição "plug-in" corresponde a inferências sobre $[T|Y, \hat{\theta}]$

- a predição Bayesiana é uma media ponderada de predições "plug-in", com diferentes valores de θ ponderados de acordo com suas probabilidades condicionais aos dados observados.

A predição Bayesiana usualmente mais cautelosa que a "plug-in", ou seja,

- incorparando-se a incerteza sobre os parâmetros usualmente resulta em intervalos de predição mais largos.

Notas:

(a) Até recentemente a necessidade de avaliar a integral que define $[Y]$ representava o maior obstáculo a aplicações práricas

(b) Desenvolvimento de métodos de Cadeias de Markov via Monte Carlo (*MCMC*) methods transformou esta situação

(c) entretanto, para problemas geoestatísticos imlementações confiáveis de MCMC não são triviais. Modelos geoestatísticos não possuem a estrutura Markoviana natural com a qual os algoritmos funcionam bem.

(d) em paticular para modelo Gaussiano outros algorítmos podem ser usados

2. Resultados para o modelo Gaussiano

Incerteza apenas no parâmetro de média

Assuma que parâmetro de média β é considerado aleatório e com priori (conjugada)

$$\beta \sim \mathrm{N}\left(m_\beta \,;\, \sigma^2 V_\beta\right)$$

A posteriori é dada por

$$\begin{aligned}
[\beta|Y] \;&\sim\; \mathrm{N}((V_\beta^{-1} + F'R^{-1}F)^{-1}(V_\beta^{-1}m_\beta + F'R^{-1}y)\,; \\
&\qquad \sigma^2\,(V_\beta^{-1} + F'R^{-1}F)^{-1}) \\
&\sim\; \mathrm{N}\left(\hat{\beta}\,;\, \sigma^2\, V_{\hat{\beta}}\right)
\end{aligned}$$

E a distribuição preditiva é:

$$p(S^*|Y,\sigma^2,\phi) \;=\; \int p(S^*|Y,\beta,\sigma^2,\phi)\, p(\beta|Y,\sigma^2,\phi)\, d\beta.$$

com média e variância dadas por

$$\begin{aligned}
E[S^*|Y] \;&=\; (F_0 - r'V^{-1}F)(V_\beta^{-1} + F'V^{-1}F)^{-1}V_\beta^{-1}m_\beta + \\
&\qquad \left[r'V^{-1} + (F_0 - r'V^{-1}F)(V_\beta^{-1} + F'V^{-1}F)^{-1}F'V^{-1}\right] Y \\
\mathrm{Var}[S^*|Y] \;&=\; \sigma^2\left[V_0 - r'V^{-1}r + \right. \\
&\qquad \left. (F_0 - r'V^{-1}F)(V_\beta^{-1} + F'V^{-1}F)^{-1}(F_0 - r'V^{-1}F)'\right].
\end{aligned}$$

Os componentes da variância da preditiva são interpretáveis como: a variância a priori, a redução na variância devida aos dados e a incerteza sobre a média.

$V_\beta \to \infty$ corresponde à krigagem ordinária (ou ainda universal e com tendência externa para médias não constantes)

Incerteza em todos os parâmetros

Assuma inicialmente o modelo sem erro de medida com priori priori $p(\beta, \sigma^2, \phi) \propto \frac{1}{\sigma^2} p(\phi)$.

A distribuição a posteriori é:

$$p(\beta, \sigma^2, \phi|y) \;=\; p(\beta, \sigma^2|y, \phi)\; p(\phi|y)$$

$$pr(\phi|y) \;\propto\; pr(\phi)\; |V_{\hat{\beta}}|^{\frac{1}{2}}\; |R_y|^{-\frac{1}{2}}\; (S^2)^{-\frac{n-p}{2}}.$$

Algorítmo 1:

(a) Discretizar $[\phi|y]$, i.e. escolher uma sequencia de valores (suporte discreto) razoável para ϕ a atribua uma distribuição priori discreta,

(b) calcule os valores da posteriori para estes valores de ϕ $\tilde{pr}(\phi|y)$,

(c) amostra um valor de ϕ da posteriori $\tilde{pr}(\phi|y)$.

(d) com o valor amostrado de ϕ calcule o valor dos parâmetros de $[\beta, \sigma^2|y, \phi]$ e amostra desta distribuição,

(e) repetir os passos (3) e (4) quantas vezes desejar. As trincas amostradas (β, σ^2, ϕ) compõe uma amostra da distribuição posteriori conjunta.

A distribuição preditiva é:

$$\begin{aligned} p(S^*|Y) &= \int\int\int p(S^*, \beta, \sigma^2, \phi|Y)\, d\beta\, d\sigma^2\, d\phi \\ &= \int\int\int p\big(s^*, \beta, \sigma^2|y, \phi\big)\ d\beta\, d\sigma^2\, pr(\phi|y)\ d\phi \\ &= \int p(S^*|Y, \phi)\ p(\phi|y)\, d\phi. \end{aligned}$$

Para amostrar desta distribuição:

Algorítmo 2:

(a) Discretizar $[\phi|Y]$, como no Algoritimo 1,

(b) calcular os valores da posteriori $\tilde{pr}(\phi|y)$ em cada ponto do suporte discreto.

(c) amostrar ϕ de $\tilde{pr}(\phi|y)$,

(d) tomar o valor amostrado de ϕ e obter $[s^*|y, \phi]$ que pode ser amostrada fornecendo amostras da preditiva,

(e) repetir passos (3) e (4) quantas vezes desejado para obter uma amostra da distribuição preditiva.

Notas:

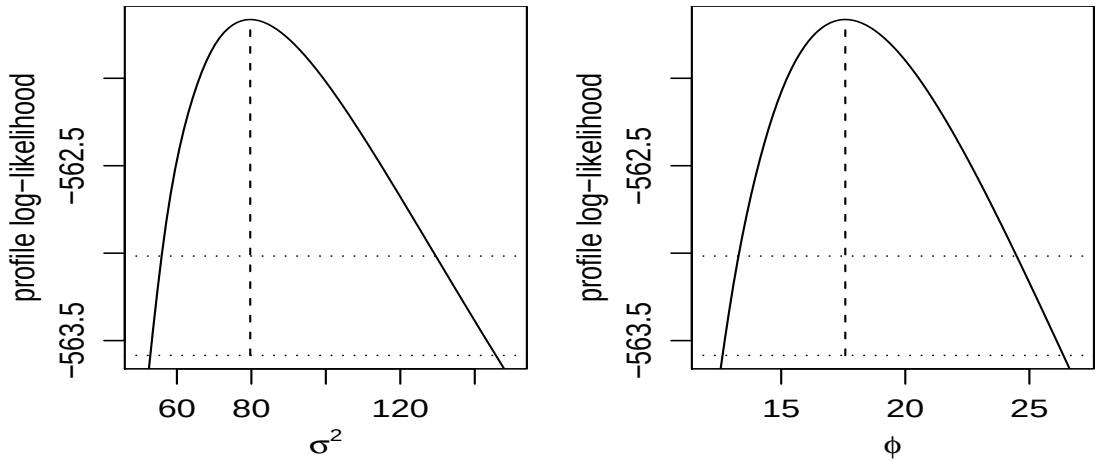
(a) Algorítmos similares podem ser usados incorporando τ e/ou κ como parâmetros desconhecidos.

(b) neste caso especifica-se uma priori discreta em um grid de valores multi-dimensional

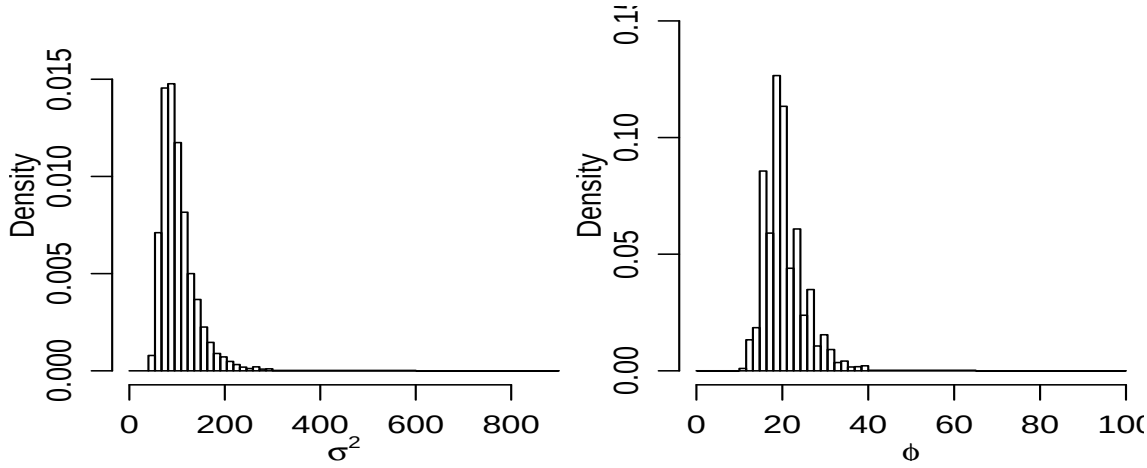
(c) isto acarreta uma grande acréscimo no custo computacional (mas sob os mesmos princípios)

(d) a estratégia de discretização torna-se então inviável para modelos de estrutura mais complexa com grande número de parâmetros de correlação.

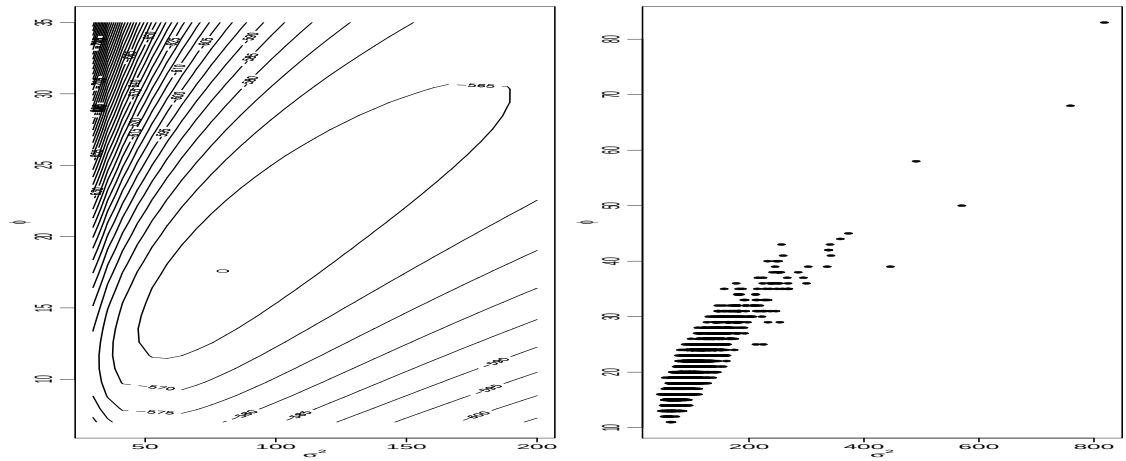
3. Estudo de caso: precipitação na Suiça, 100 dados



log-verossimilhanças perfilhadas dos parâmetros de correlação: σ^2; ϕ

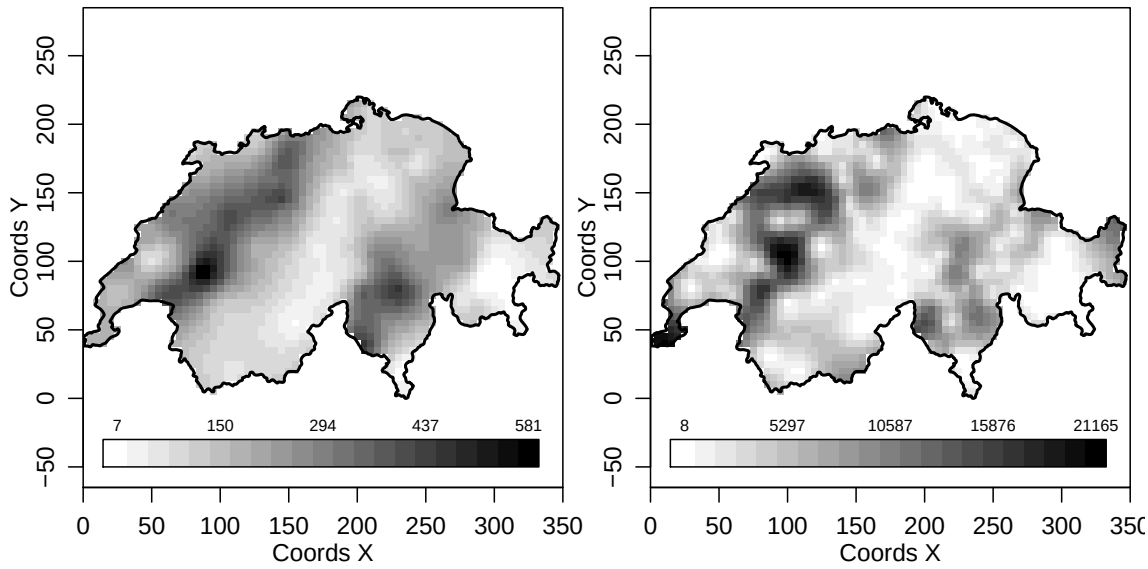


Posterioris dos parâmetros de correlação:: σ^2; ϕ

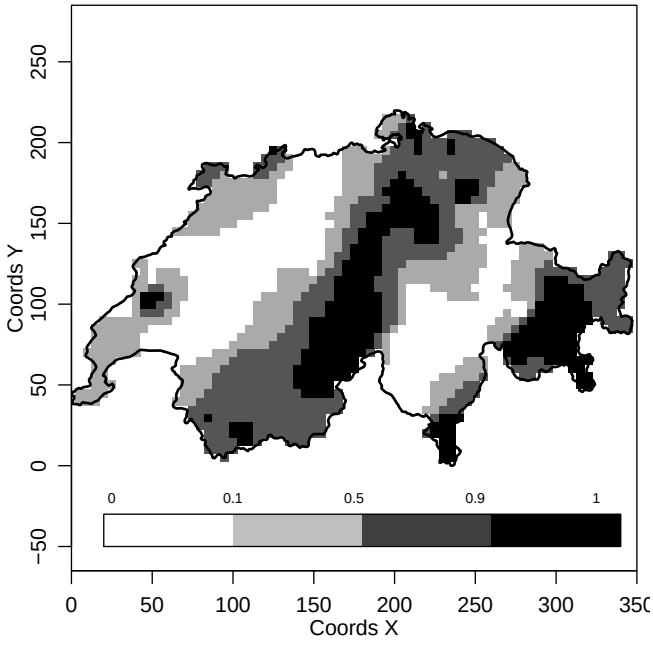


log-verossimilhança perfilhada 2-D (esquerda) a amostras das posterioris (direita) para parâmetros σ^2 e ϕ

Precipitação na Suiça: resultados de predição

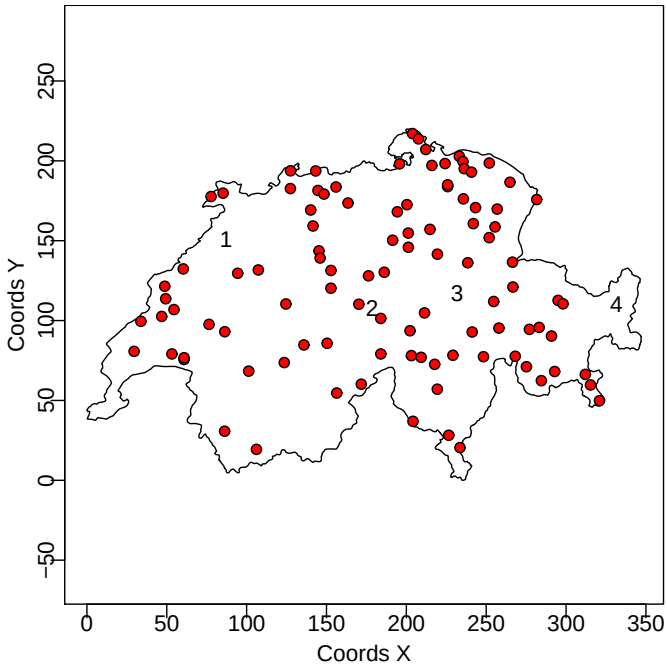


Superfície de predição e as medidas de precisão associadas: (a) médias da posteriori; (b) variância da posteriori

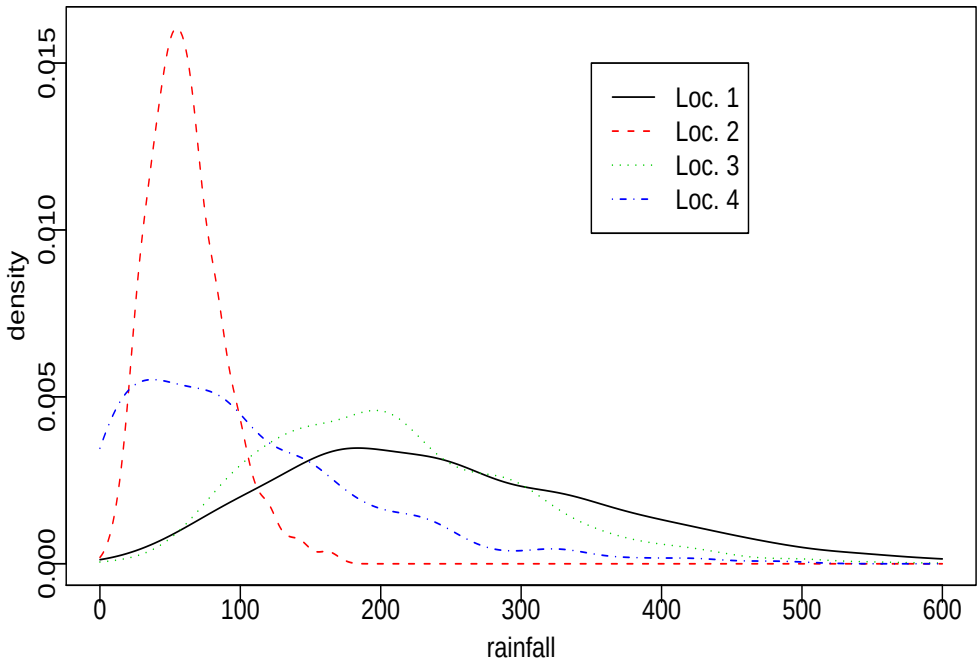


Contornos de probabilidade predititos para níveis 0.10, 0.50 and 0.90 para o conjunto $T = \{x : S(x) < 150\}$

Precipitação na Suiça: resultados de predição (cont.)



Estações meteorológicas e pontos selecionados para predição (1 a 4)



Distribuição preditiva nos pontos selecionados.

PARTE VI:

MODELOS LINEARES GENERALIZADOS ESPACIAIS

1. Modelos lineares Generalizados Espaciais

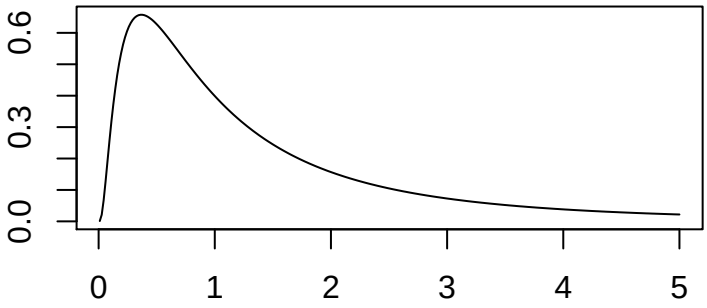
2. Inferência via MCMC

3. estudo de caso: Ilha de Rongelap

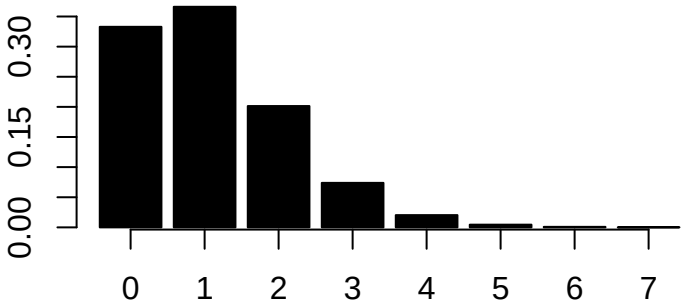
4. estudo de caso: Malária em Gâmbia

1. Modelos lineares Generalizados Espaciais

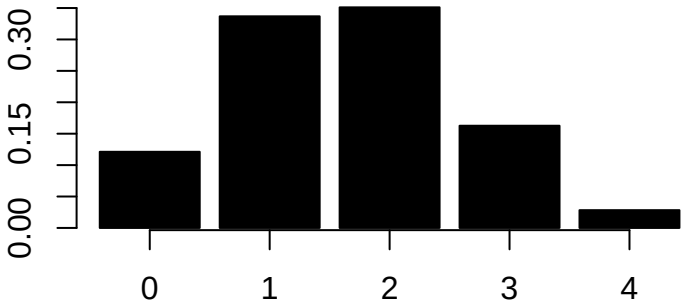
Dados Positivos:



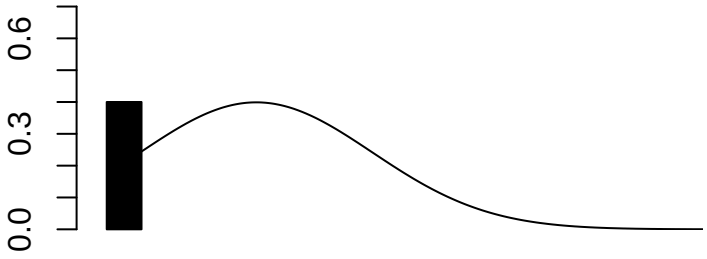
Dados Contagem:



Dados Binomiais:



Dados Positivos com Zeros:



Exemplos de Modelos

$x_1, \ldots, x_n$ posições com observações

Poisson-log

- $[Y(x_i) \mid S(x_i)]$ é Poisson com densidade

$$f(z;\mu) = \exp(-\mu)\mu^z/z! \quad z = 0, 1, 2, \ldots$$

- ligação: $E[Y(x_i) \mid S(x_i)] = \mu_i = \exp(S(x_i))$

Binomial-logit

- $[Y(x_i) \mid S(x_i)]$ é binomial com densidade

$$f(z;\mu) = \binom{r}{z}(\mu/r)^z(1-\mu/r)^{r-\mu} \quad z = 0, 1, \ldots, r$$

- ligação: $\mu_i = E[Y(x_i) \mid S(x_i)]$, $S(x_i) = \log(\mu_i/(r-\mu_i))$

Função de Verossimilhança

$$L(\theta) = \int_{\mathbb{R}^n} \prod_i^n f(y_i; h^{-1}(s_i)) f(s \mid \theta) ds_1, \ldots, ds_n$$

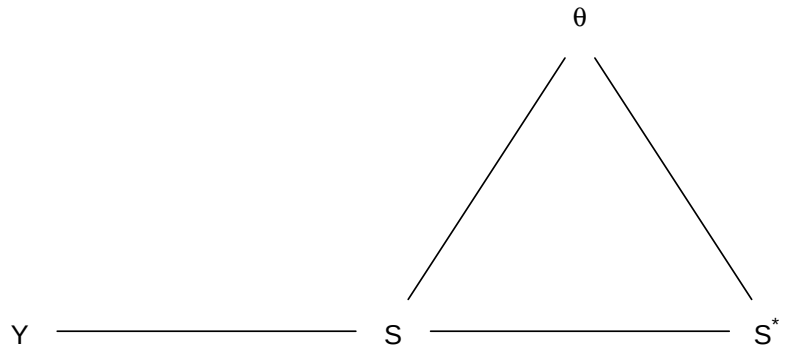
Integral de alta dimensionalidade !!!

2. Inferência para o modelo geoestatístico linear generalizado

- avaliacão da verossimilhança envolve integracão numérica de multidimensional
- métodos aproximados (ex Breslow and Clayton, 1993) tem acurácia duvidosa
- MCMC é possível embora não rotineira

Esquemas para MCMC

- Ingredientes
  - · Prioris para os parâmetros de regressão β e de covariância θ
  - · Dados: $Y = (Y_1, ..., Y_n)$
  - · $S = (S(x_1), ..., S(x_n))$
  - · $S^* =$ todos outros $S(x)$
- Estrutura de independência condicional



- use resultados das cadeias para contruir declarações à posteriori sobre $[T|Y]$, onde $T = \mathcal{F}(S^*)$

3. Estudo de caso: Ilha Rongelap

- **Ilha de Rongelap**
  - aproximadamente a 2500 milhas sudoeste do Hawaii
  - contaminada por testes de armas nucleares em 1950's
  - evacuada em 1985
  - segura para re-assentamento?

- **Problemas estatísticos**
  - delineamento e medidas de campo de ^{137}Cs
  - estimar variacão espacial da radiotividade de ^{137}Cs
  - comparacão com padrões de segurança

O modelo Poisson

- Medidas básicas são contagens Y_i em intervalos de tempo t_i nas localizações x_i $(i = 1, ..., n)$

- estrutura dos dados sugere o modelo:

  · $S(x) : x \in R^2$ processo estacionário Gaussiano (radioatividade local)

  · $Y_i|\{S(\cdot)\} \sim$ Poisson(μ_i)

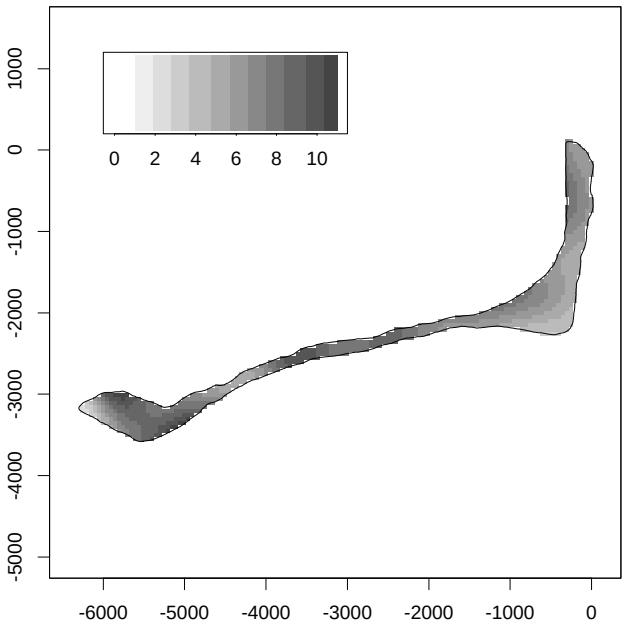
  · $\mu_i = t_i\lambda(x_i) = t_i \exp\{S(x_i)\}$.

- Objetivos:

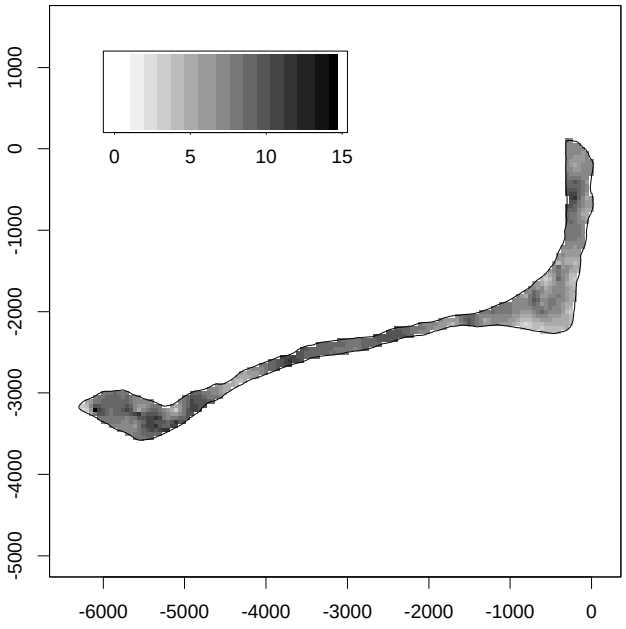
  · predizer $\lambda(x)$ sobre toda ilha

  · max $\lambda(x)$

  · arg(max $\lambda(x)$)

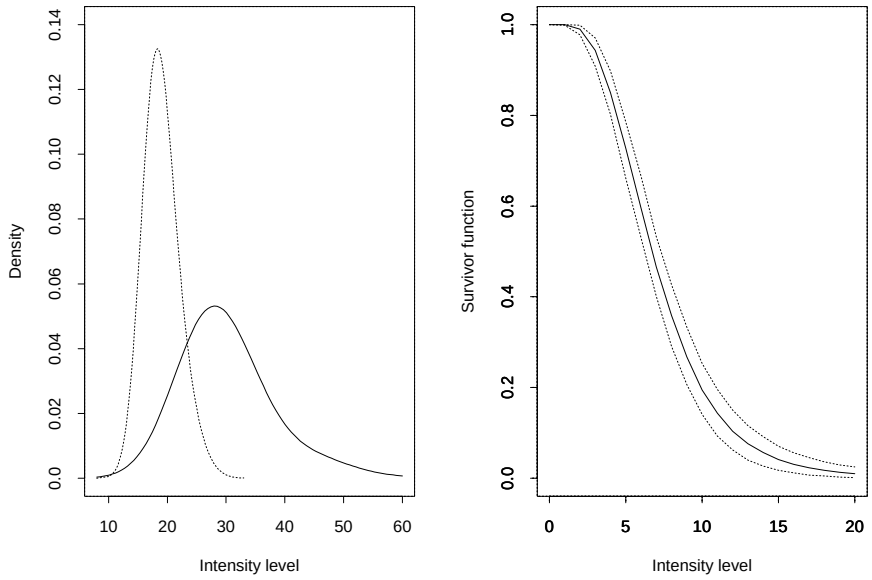


superfície de radiotividade predita utilizando krigagem logarítmica



superfície de radiotividade predita utilizando o modelo log-linear Poisson com processo latentre Gaussiano

Predicão Bayesiana de funcionais não lineares da superfície de radiacão



The left-hand panel shows the predictive distribution of maximum radioactivity, contrasting the effects of allowing for (solid line) or ignoring (dotted line) parameter uncertainty; the right-hand panel shows 95% pointwise credible intervals for the proportion of the island over which radioactivity exceeds a given threshold.

4. Estudo de caso: Malária em Gâmbia

- Neste exemplo a variacão espacial é de interesse científico secundário.

- O objetivo primário e descrever a dependência entre a prevalência de parasitas de malária e as covariáveis medidas
  - em vilas
  - em indivíduos

- Particular interesse em saber se o índice de vegetacão derivado de medidas de satélite pode ser utilizado como preditor da prevalência de malária.
  Isto ajudaria profissionais de saúde a alocar melhor os recursos que são escarsos.

Estrutura dos dados

- 2039 crianças em 65 vilas
- cada uma testada para presença de parasitas de malária no sangue

Covariáveis das crianças

- idade (dias)
- sexo (F/M)
- uso de mosquiteiro (nenhum, não tratado e tratado)

Covariáveis das vilas:

- localizacão
- indice de vegetacão (satélite)
- presença de centro de saúde na vila

Modelo de regressão logística

- $Y_{ij} = 0/1$ presença ou ausência de parasitas de malária na jth criança da ith vila
- $f_{ij} =$ covariável da criança
- $w_i =$ covariável da vila
- $\text{logit}(P(Y_{ij} = 1|S(\cdot))) = f'_{ij}\beta_1 + w_i{}'\beta_2 + S(x_i)$

É razoável assumir infeções condicionalmente independentes na mesma vila?

Caso não, o modelo deve ser extendido para permitir variabilidade extra-binomial não-espacial

- $U_i \sim \text{N}(0, \nu^2)$
- $\text{logit}P(Y_{ij} = 1|S(\cdot), U) = f'_{ij}\beta_1 + w_i{}'\beta_2 + U_i + S(x_i)$

Análise exploratória

- ajuste modelo logístico padrão sem $S(x)$ e/ou U

- calcule para cada vila:

  $N_i = \sum_{j=1}^{n_i} Y_{ij}$

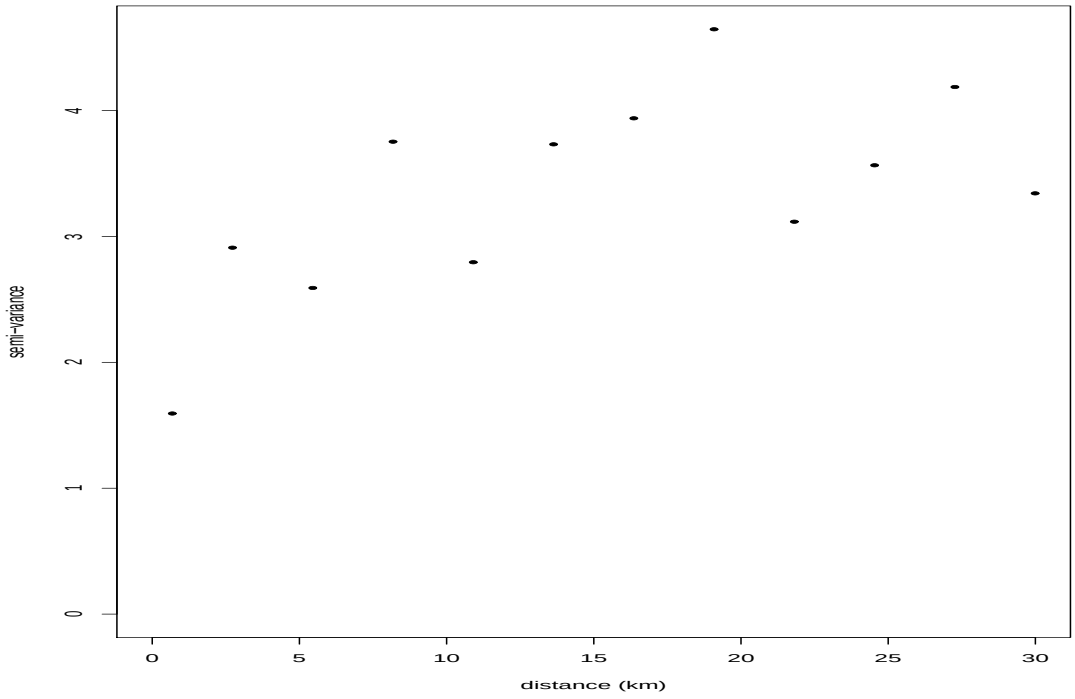
  $\mu_i = \sum_{j=1}^{n_i} \hat{P}_{ij}$

  $\sigma_i^2 = \sum_{j=1}^{n_i} \hat{P}_{ij}(1 - \hat{P}_{ij})$

- resíduos de vila, $r_i = (N_i - \mu_i)/\sigma_i$

- derivar dados r_i

- ajuste de parâmetros de covariância



Variograma do resíduos de vilas

Análise "model-based"

$\alpha =$ intercepto

$\beta_1 =$ coeficiente para idade

$\beta_2 =$ coeficiente uso de mosquiteiro

$\beta_3 =$ coeficiente para mosquiteiro tratado

$\beta_3 =$ coeficiente para indice de verde

$\beta_4 =$ coeficiente para presença de centro de saúde

$\nu^2 =$ variância do efeito aleatório não espacial U_i

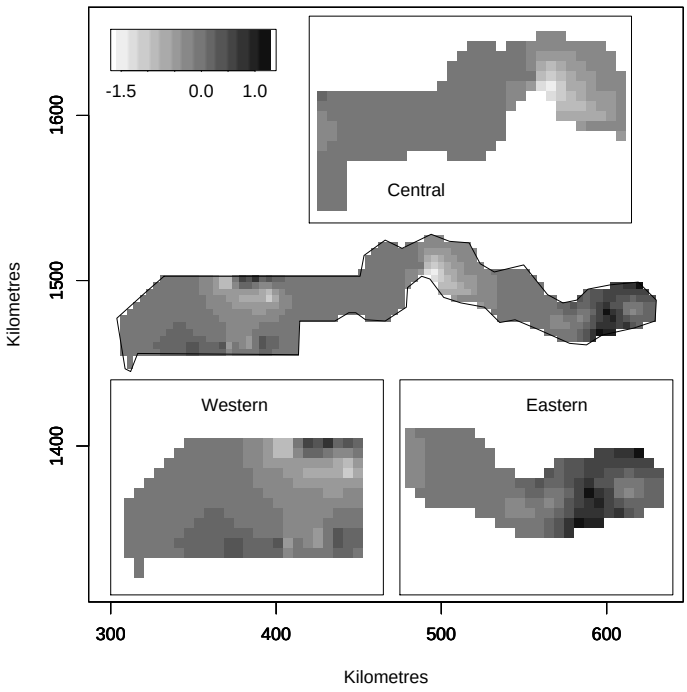
$\sigma^2 =$ variância do preocesso espacial $S(x)$

$\phi =$ parâmetro de decaimento da correlacão

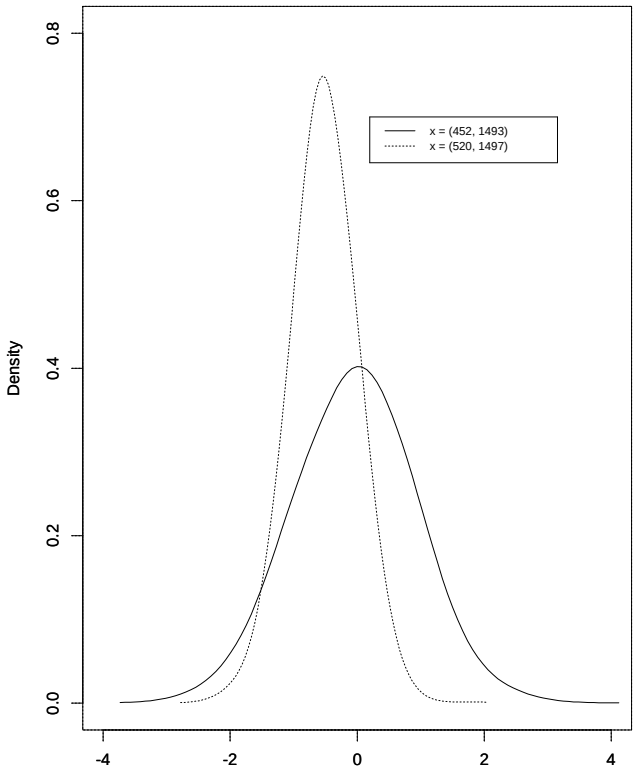
$\kappa =$ parâmetro de suavidade

Param.	2.5% Qt.	97.5% Qt.	Mean	Median
α	-4.232073	1.114734	-1.664353	-1.696228
β_1	0.000442	0.000918	0.000677	0.000676
β_2	-0.684407	-0.083811	-0.383750	-0.385772
β_3	-0.778149	0.054543	-0.355655	-0.355632
β_4	-0.039706	0.071505	0.018833	0.020079
β_5	-0.791741	0.180737	-0.324738	-0.322760
ν^2	0.000002	0.515847	0.117876	0.018630
σ^2	0.240826	1.662284	0.793031	0.740790
ϕ	1.242164	53.351207	11.653717	7.032258
κ	0.150735	1.955524	0.935064	0.830548

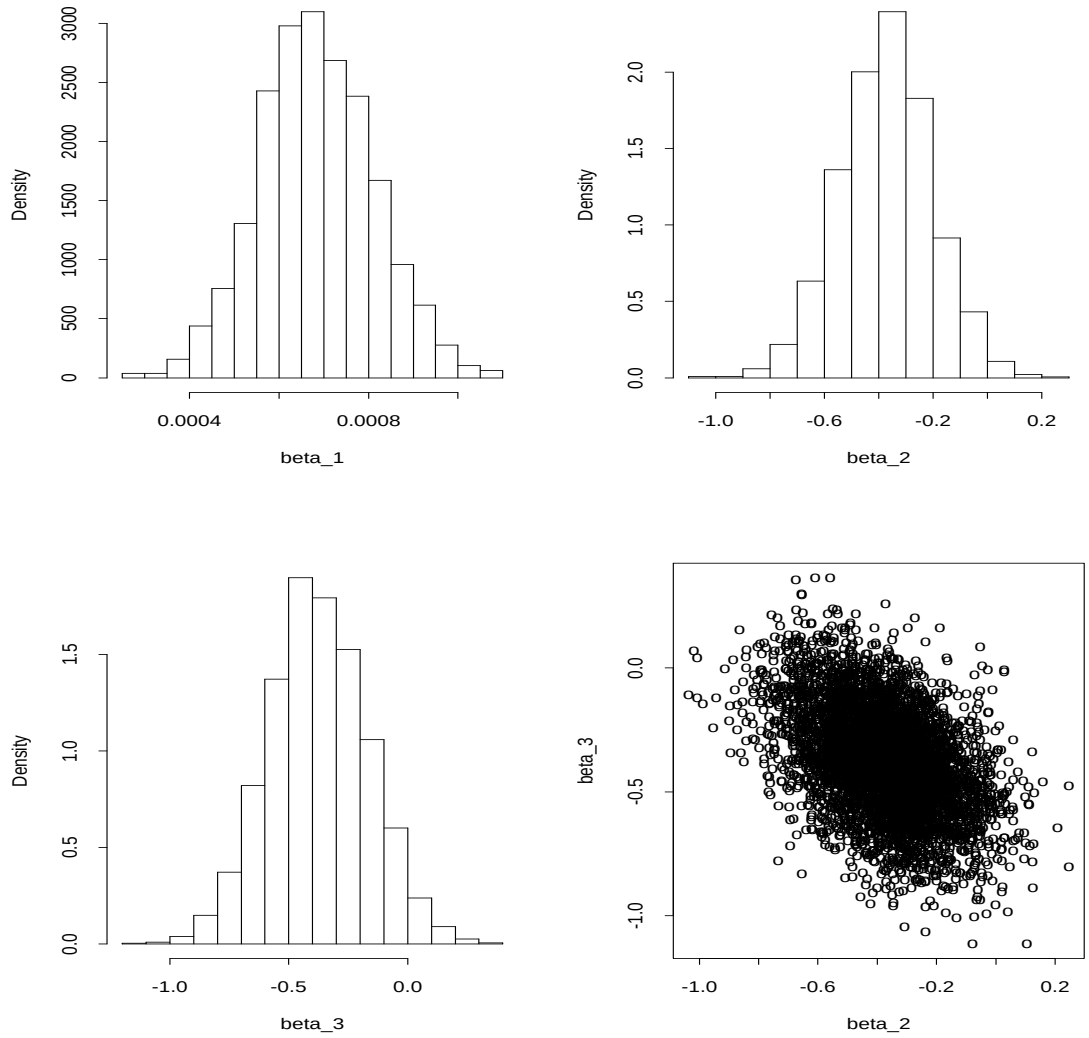
- ν^2 próximo de zero



superfície predita $\hat{S}(x)$ (média à posteriori)



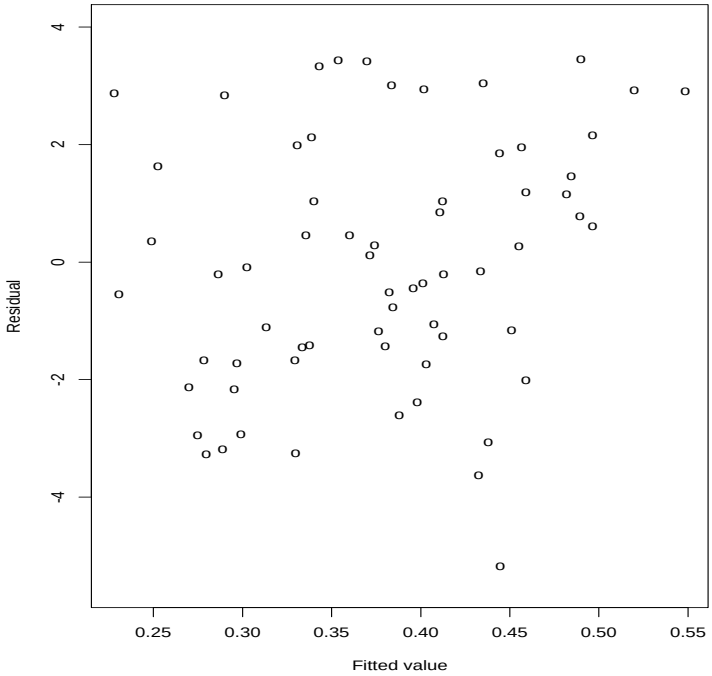
Posterioris para $S(x)$ em dias localizacoes, linha sólida – remota (452, 1493), linha interrompida – central (520, 1497)



posterioris para os parâmetros de regressão

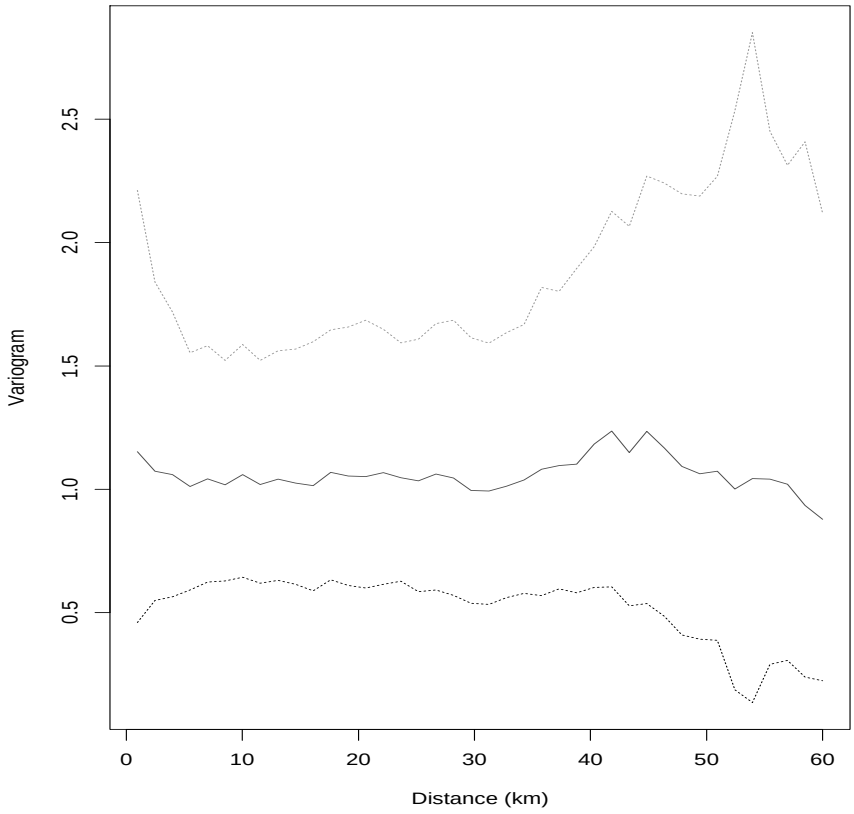
- $\beta_1 =$ efeito de idade
- $\beta_2 =$ efeito de mosquiteiro não tratado
- $\beta_3 =$ efeito adicional de tratamento de mosquiteiro

Qualidade do ajusto do modelo



resíduos de vila *vs* valores ajustados

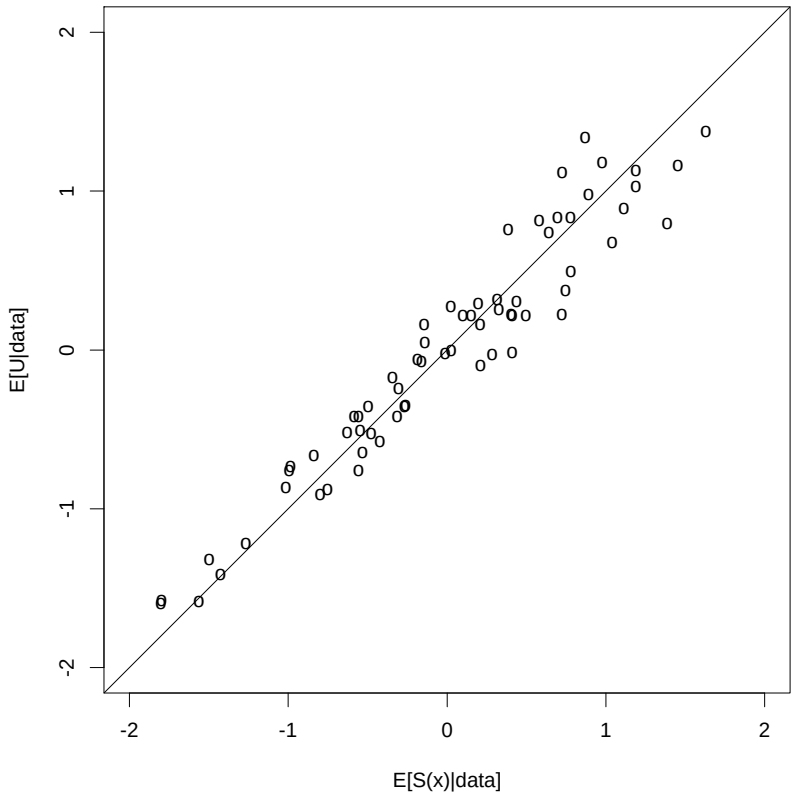
- $r_{ij} = (Y_{ij} - \hat{p}_{ij})/\sqrt{\{\hat{p}_{ij}(1 - \hat{p}_{ij})\}}$
- $r_i = \sum r_{ij}/\sqrt{n_i}$
- checa adequacidade do modelo para p_{ij}



variograma empírico de resíduos padronizados com intervalos de confiança (95%) construídos a partir de simulações do modelo ajustado

- $r_{ij} = (Y_{ij} - \hat{p}^*_{ij})/\sqrt{\{\hat{p}^*_{ij}(1 - \hat{p}^*_{ij})\}}$
- $r_i = \sum r_{ij}/\sqrt{n_i}$
- $\text{logit}(p^*_{ij}) = \hat{\alpha} + z'_{ij}\hat{\beta}\hat{S}(x_i)$
- checa adequacidade do modelo para $S(x)$

O modelo geostatístico é mesmo necessário?



média da posteriori para os efeitos aleatórios $\hat{U}_i$ de um GLMM não espacial contra médias a posteriori de $\hat{S}(x_i)$ nas localizações observadas no modelo geoestatístico

- alta correlacão evidencia depêndencia espacial

GEE: uma alternativa para problemas onde a ênfase está nas covariáveis?