

População, Espaço e Ambiente

Abordagens Espaciais em Estudos de População: Métodos Analíticos e Técnicas de Representação

Parte IV – Integração de Dados

Introdução aos interpoladores espaciais para geração de Superfícies de Densidade de População

Silvana Amaral

Antonio Miguel V. Monteiro

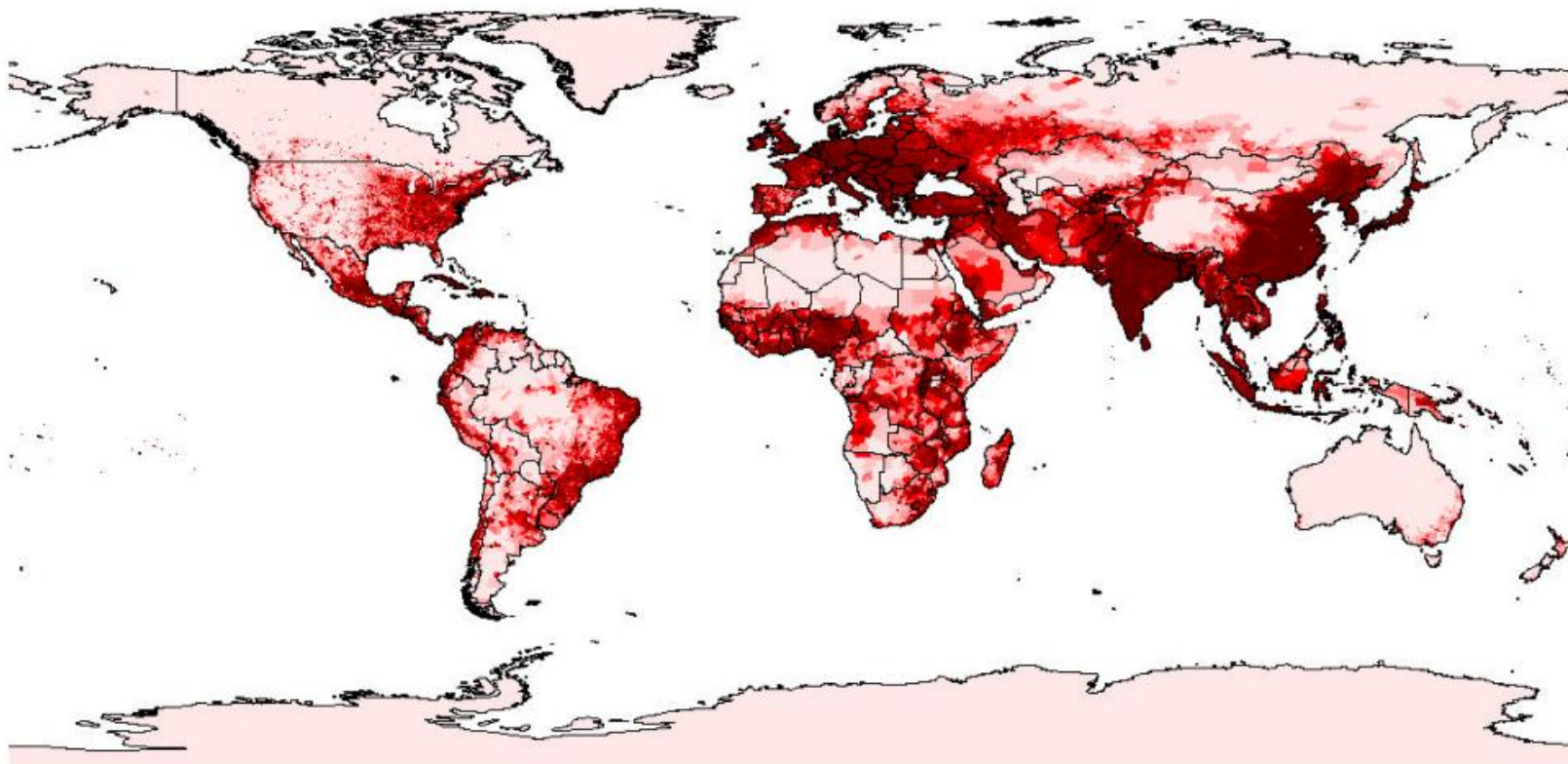
{silvana@dpi.inpe.br, miguel@dpi.inpe.br}



De Censos a Superfícies



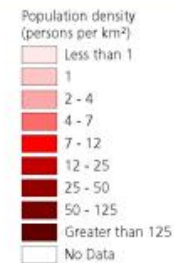
Global Population Density



Map Projection: Geographic

Citation: WRI, 2000 based on CIESIN 2000

Notes:

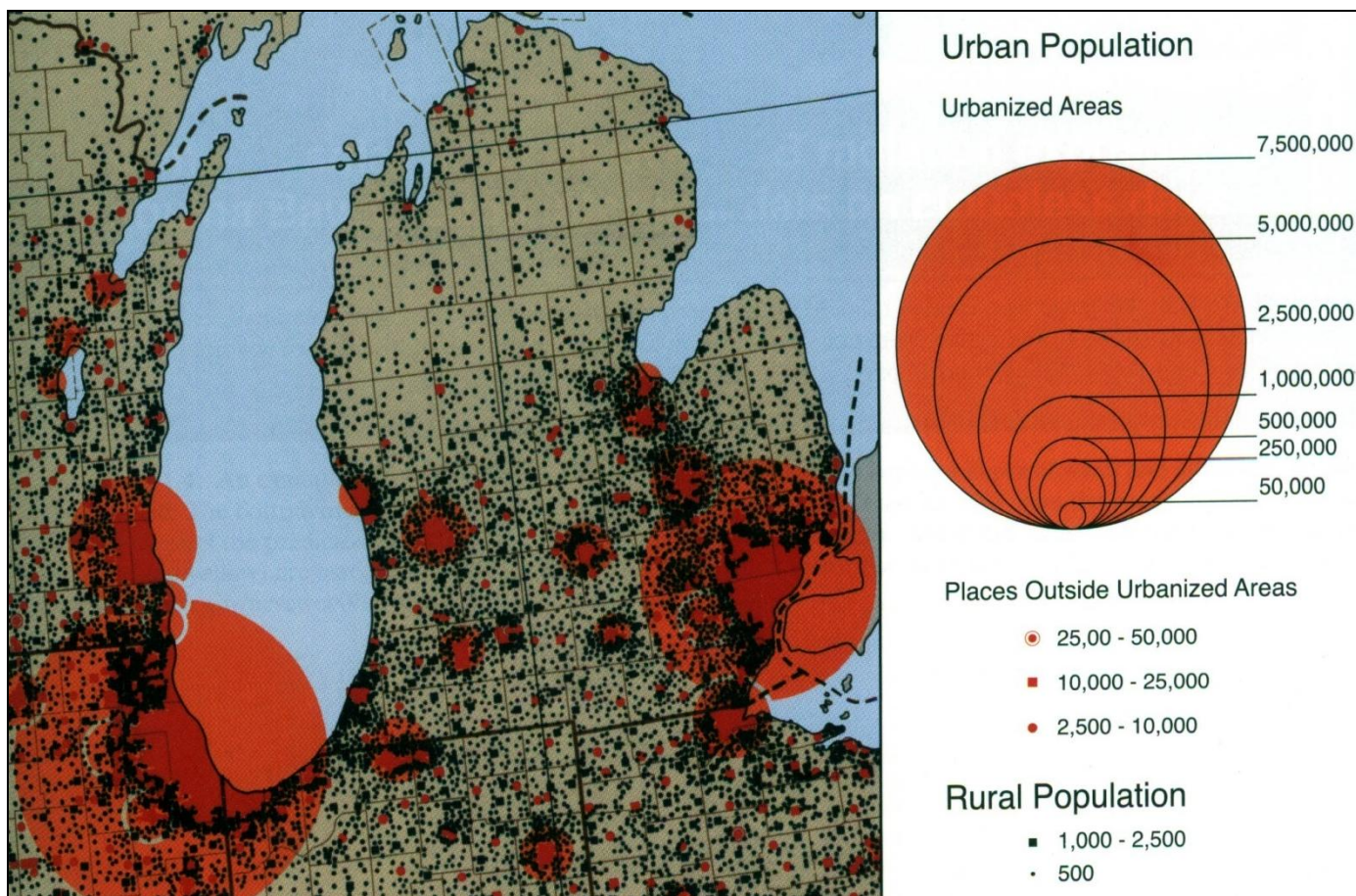


Introdução

- Dados Populacionais – representações cartográficas
 - Mapas de pontos – *dots* e símbolos
 - Mapas coropléticos - temáticos
 - tamanho e forma da unidade territorial, número de classes e limites entre as classes

Representações Cartográficas

- Mapas pontos - símbolos

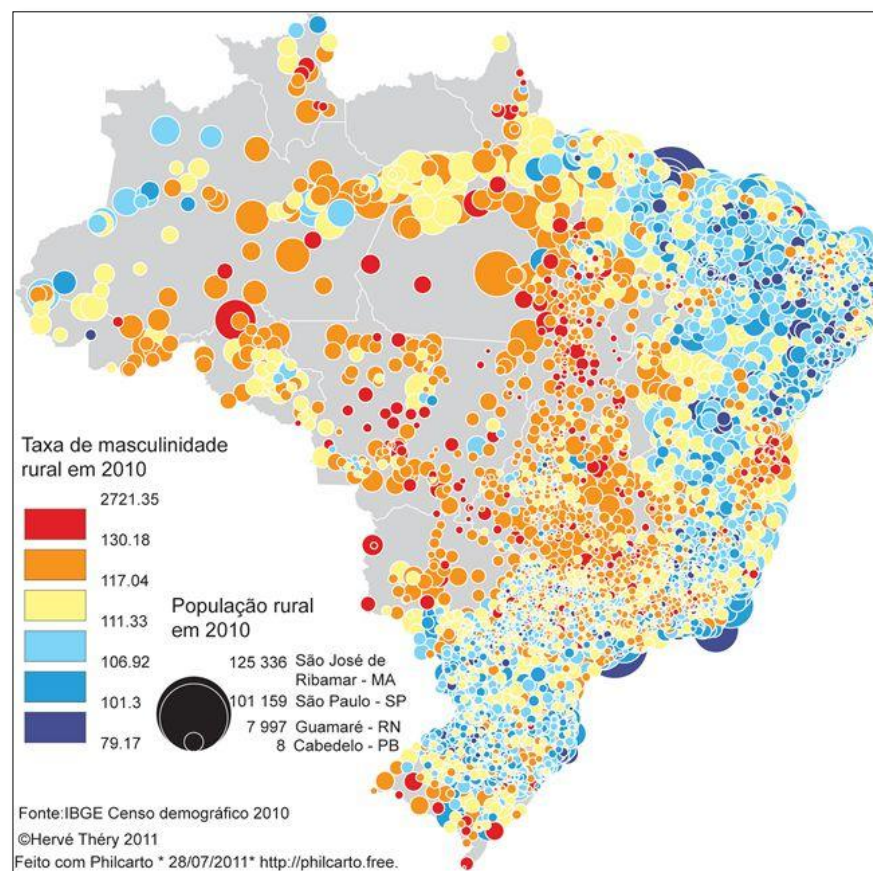
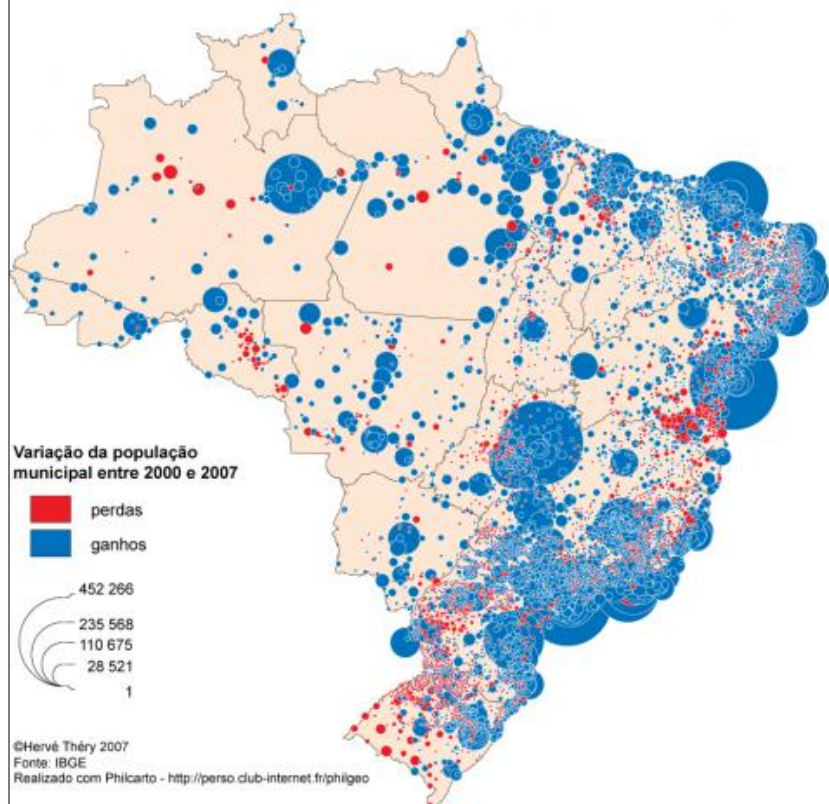


Representações Cartográficas

Mapas de símbolos pontuais proporcionais

- Diferenciação da intensidade do fenômeno em cada área
- Combinação de diferentes variáveis
- Não apresenta distribuição de quantidade no espaço

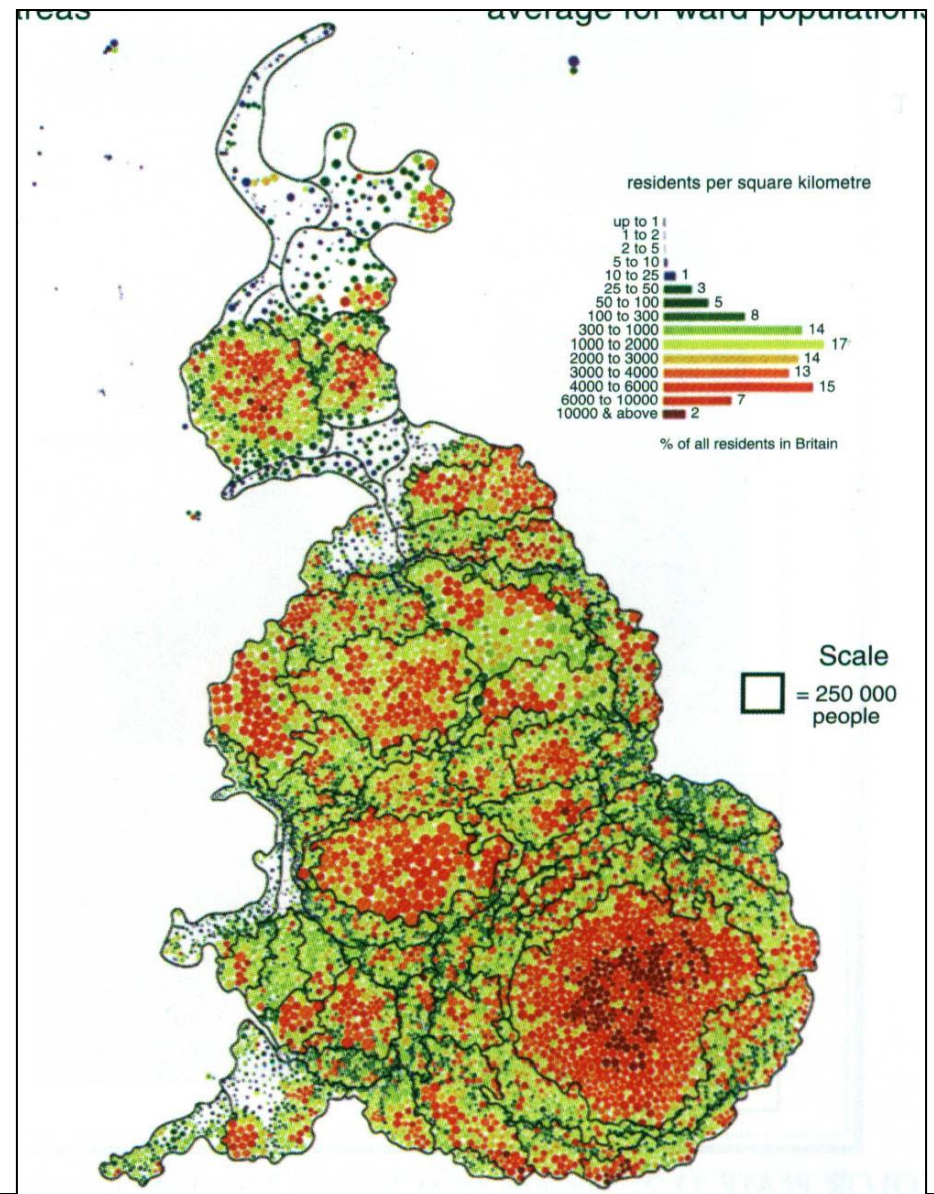
Evolução da população brasileira entre 2000 e 2007



Representações Cartográficas



- Mapas de pontos - *dots*
- Representar fenômenos discretos com conotação pontual. Ilustra a densidade espacial;
- Objetivo: facilitar a comunicação cartográfica
- É um tipo especial de mapas de símbolos proporcionais
- Cada ponto = mesma quantidade
 - Superposição em áreas densas
 - Valores agregados
1pto= x pessoas



Distribuição da População

Repres

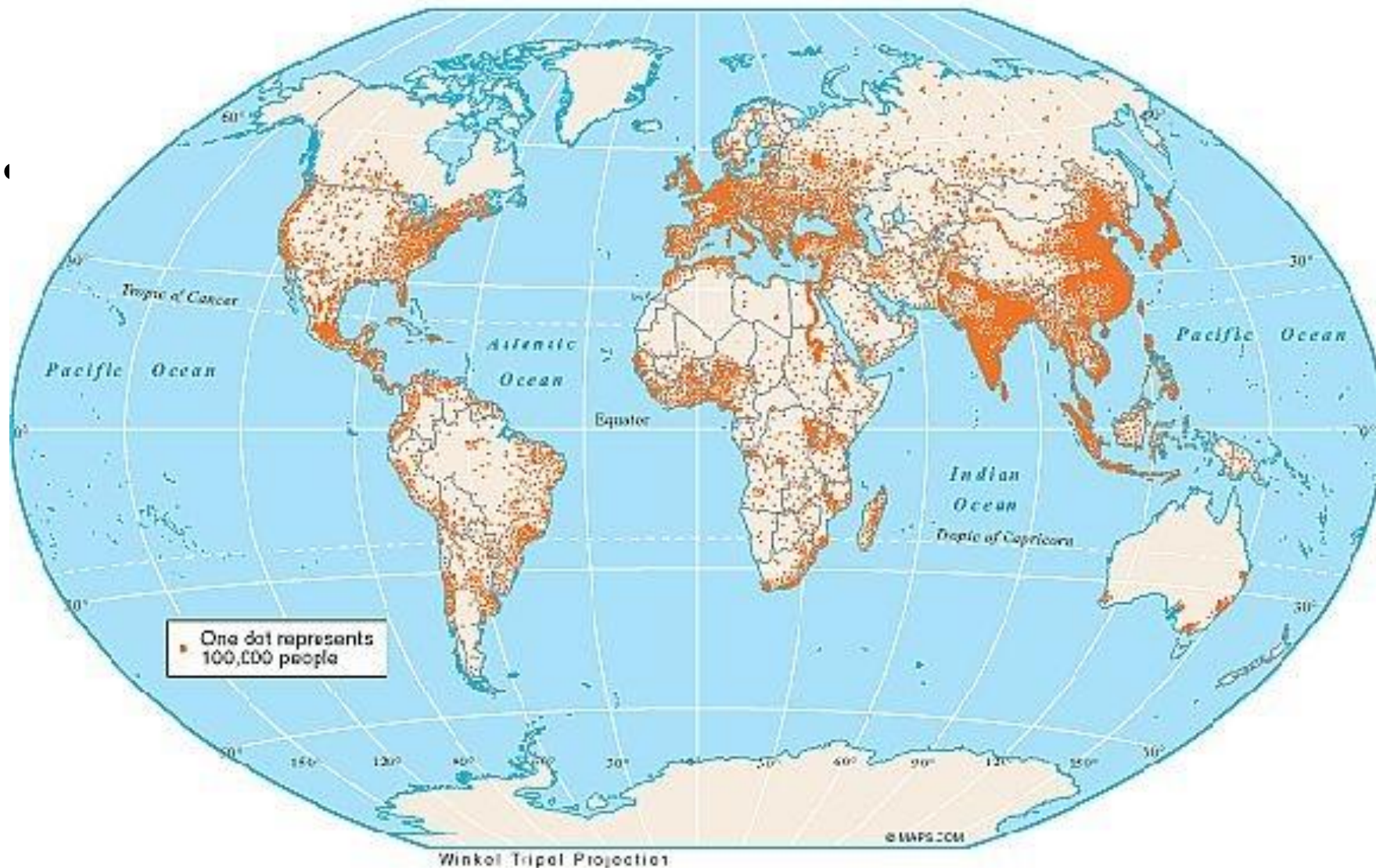
Mapa de pontos



Representações Cartográficas



World Population Distribution, 1998

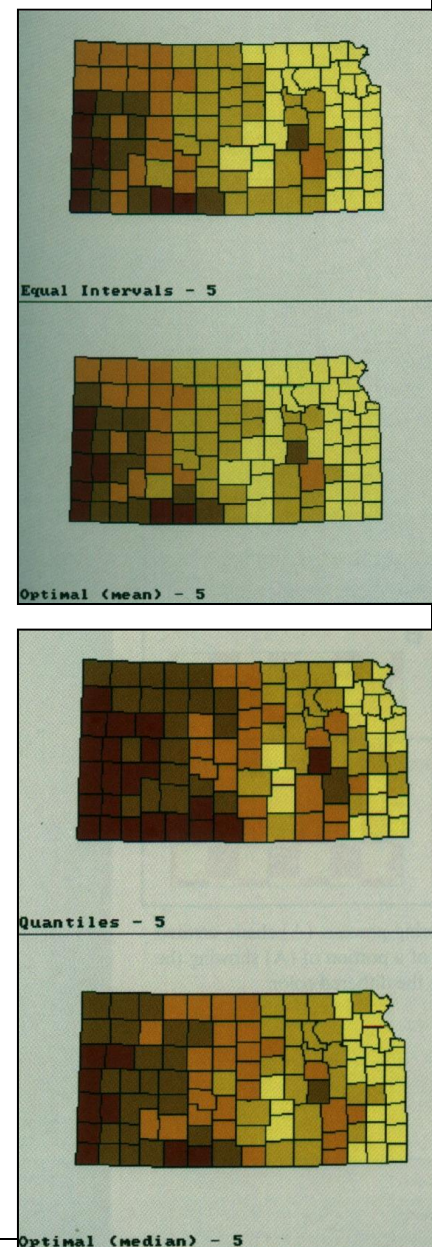


Representações Cartográficas



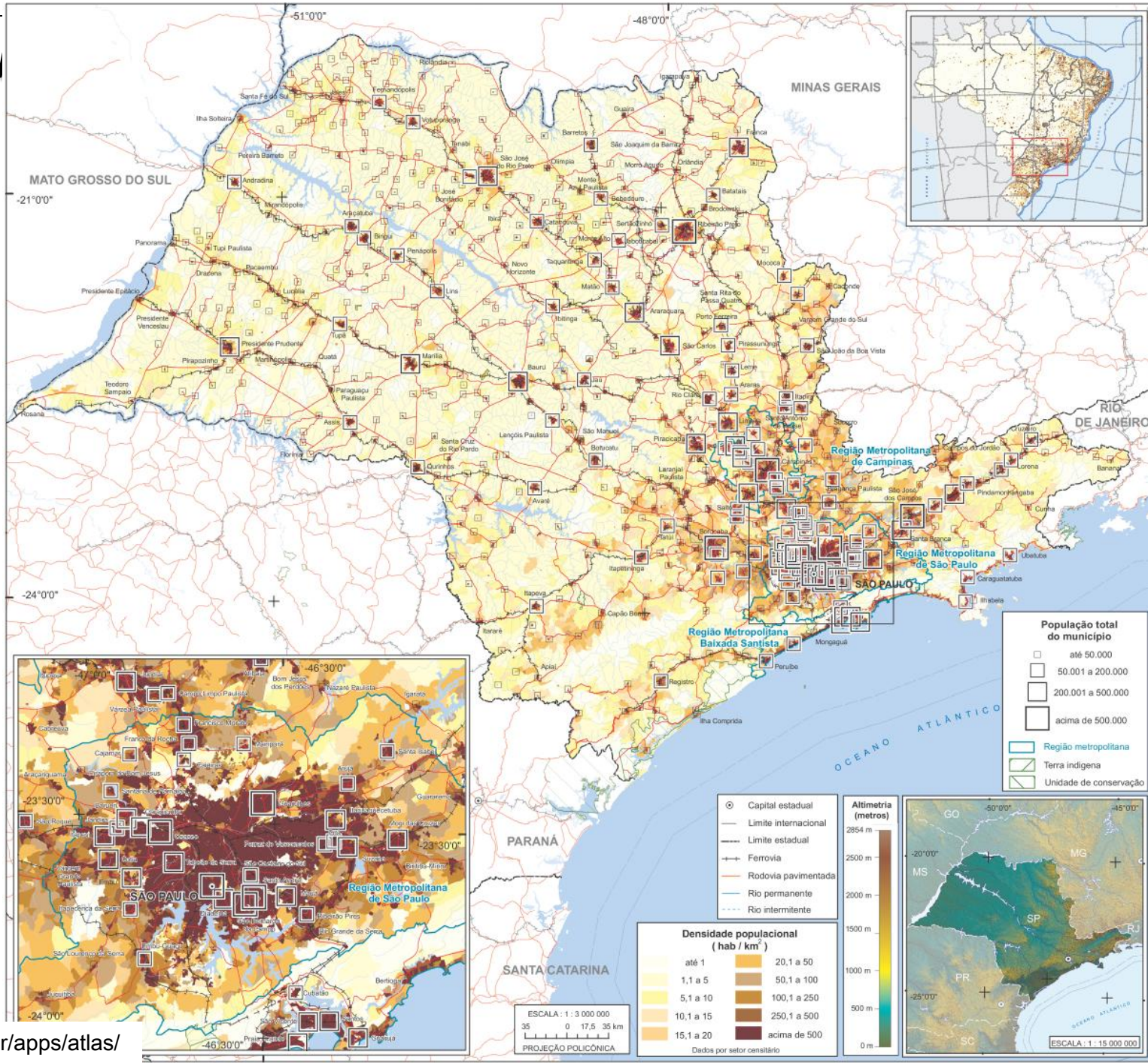
Mapas Coropléticos – representação cartográfica para traduzir valores para áreas

- Valores apresentados devem ser razões ou proporções. Evitar valores absolutos
- Luminosidade e saturação da cor, indicam diferenças ordenadas em classes distintas
- Dois tipos:
 - Densidade – razões, n pessoas /km²
 - Porcentagem, ex %habitantes/ total populacional
- Mapas temáticos
- Valores associados a áreas e diferenças de cores - intensidade do fenômeno
- Representa estrutura hierárquica dos dados
- Tamanho, forma das unidades, número de classes e definição de limites

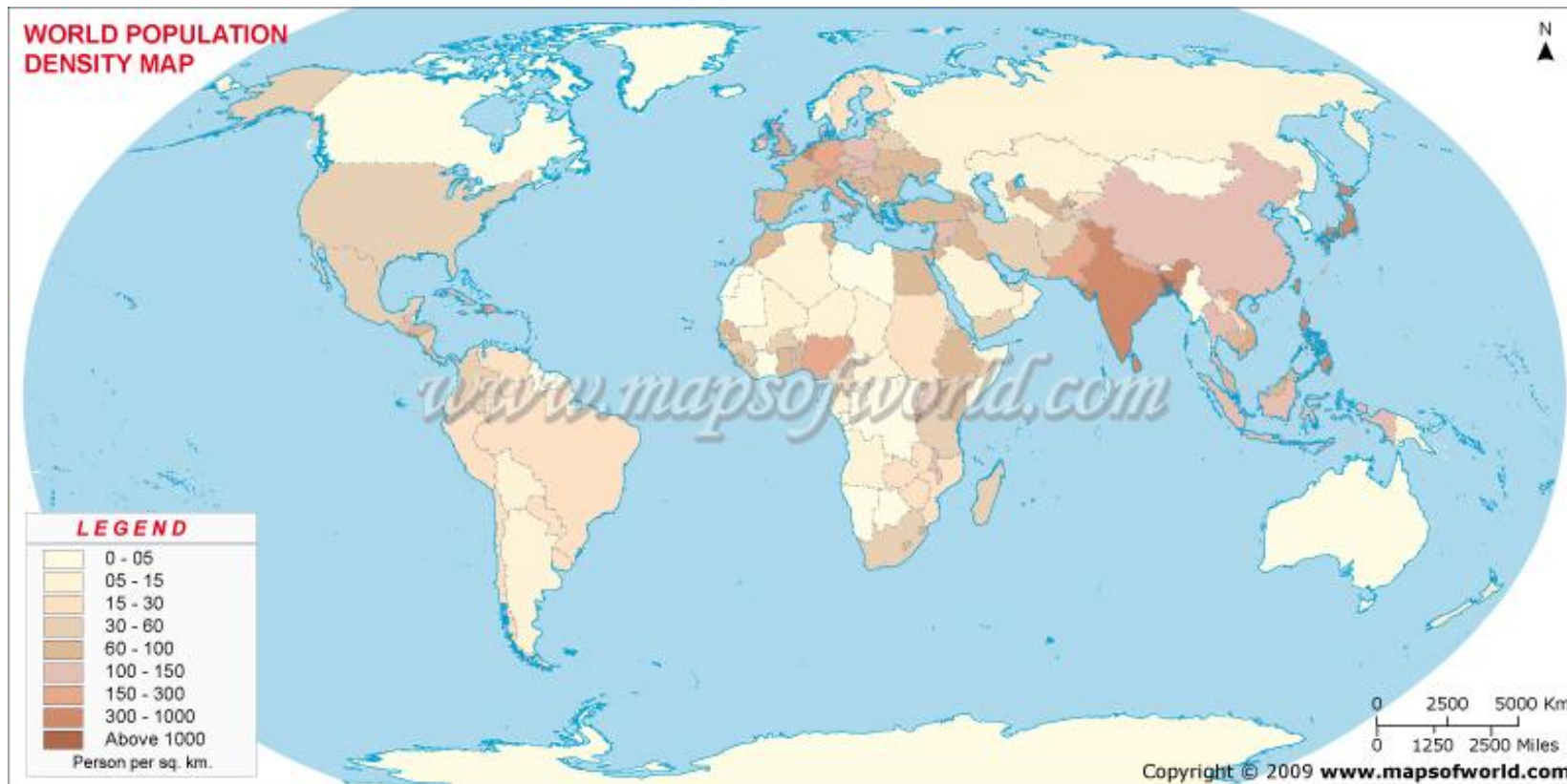


República

Mapa Coroplético



Representações Cartográficas



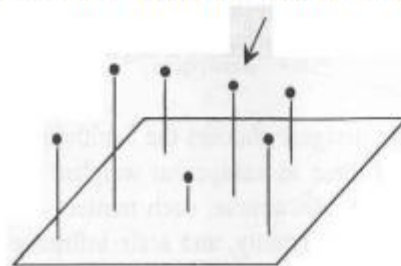
Representações Cartográficas

Um mapa isaritmico é a representação plana através de isolinhas de uma superfície tridimensional, que pode ser uma superfície real, como o relevo, ou uma superfície abstrata ou conceitual. Classificados em isométricos ou isopléticos, de acordo com a natureza dos dados:

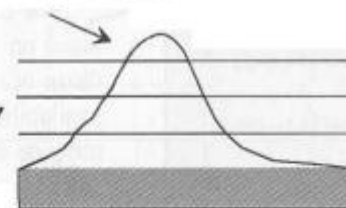
- Mapas isométricos, os dados foram coletados para localizações **pontuais** específicas, os pontos de controle, como dados de temperatura em estações meteorológicas.

- Mapas isopléticos, os dados referem-se a **unidades de enumeração de áreas geográficas**, como densidade demográfica ou taxa de produção agrícola

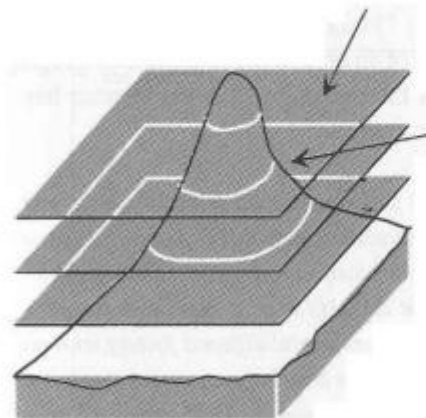
1. Representação proporcional ao valor do fenômeno.



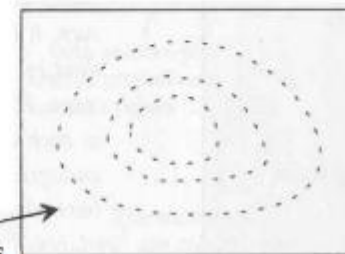
2. Superfície obtida pela interpolação dos valores.



3. Planos hipotéticos atravessam o volume.



4. O traço desenhado pela interseção dos planos com a superfície formam as isolinhas.



5. O desenho planimétrico dos traços formam o mapa isaritmico.

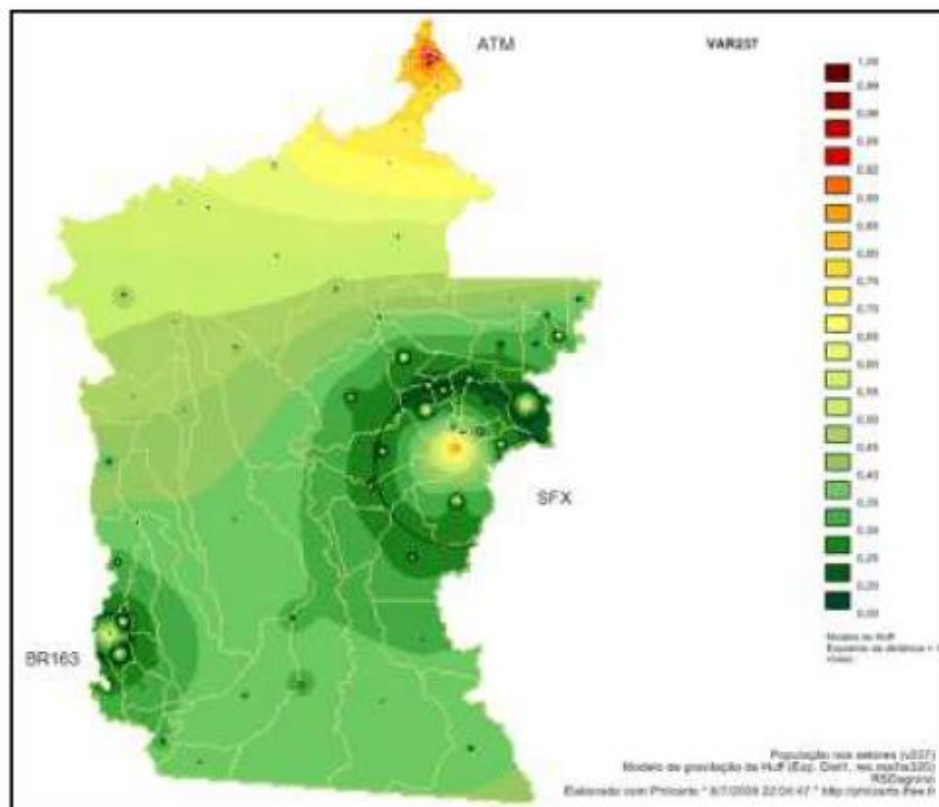
Modificado por Pantaleão (2003)

Representações Cartográficas

Mapa Isoplético ou de Linhas

- A partir de pontos, linhas indicam valor semelhante
- Pode ser usado para representar densidade populacional - há um valor para cada localização na superfície geográfica

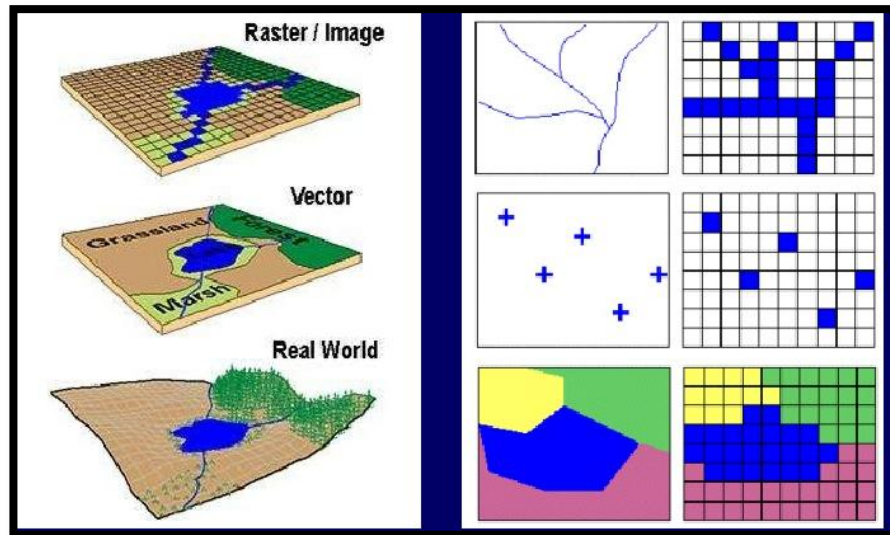
População residente nos setores (modelo de gravitação de Huff)



Representação Computacional



- **Representação digital** de dados geográficos – modelos matricial e vetorial

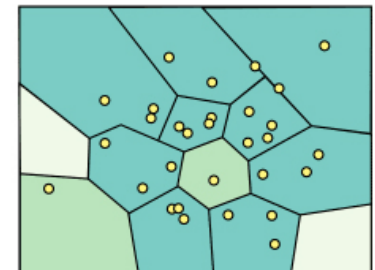
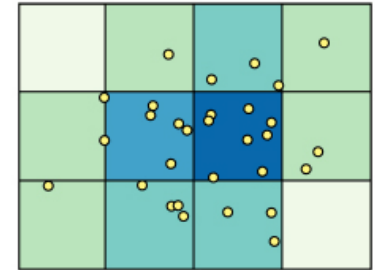


- Modelos de representação computacional para dados sócio-econômicos:
 - Indivíduo
 - Área
 - Fenômeno contínuo no espaço

Representação Computacional



- Área – fenômeno em regiões discretas, unidades homogêneas;
- Unidades – arbitrárias como setor censitário, não representam distribuição espacial da variável
- Problema de unidade de área modificável (MAUP)?
 - Escala (agregação) e Zona (agrupamento)
- Séries temporais ??



MAUP

Zona (agrupamento)

Superfície de Densidade



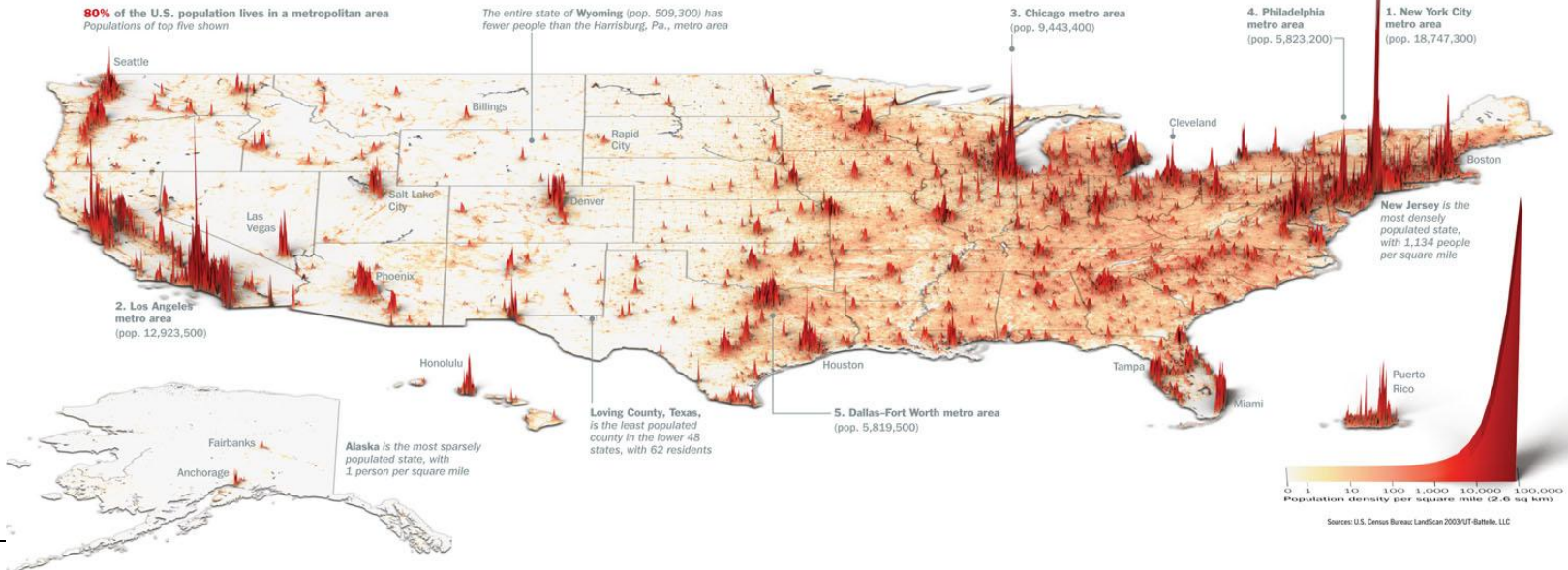
- Modelos de Superfície - alternativa a restrição das áreas
 - Densidade demográfica – fenômeno contínuo
 - Objetivo: estimar distribuição no maior detalhe possível
 - Vantagem: manipulação e análise independentes de área
 - Armazenar e disponibilizar dados em BD Globais

Where We Live...

Unlike many developed countries, the U.S. keeps growing. We are also moving south and west. But compared with China or India, the nation is a vast prairie

Our families are getting smaller—with one vital exception. Compared with those of Europe and Japan, the U.S. population is younger and more colorful because of the continued arrival of immigrants and their higher-than-average birthrates. Of the 100 million Americans who will join us in the next 37 years, half will be immigrants or their children. In the next few decades, 97% of the world's population growth will occur in the developing world; the U.S. is the largest developed country in the world that is still growing at a healthy clip. That matters, strategically, economical-

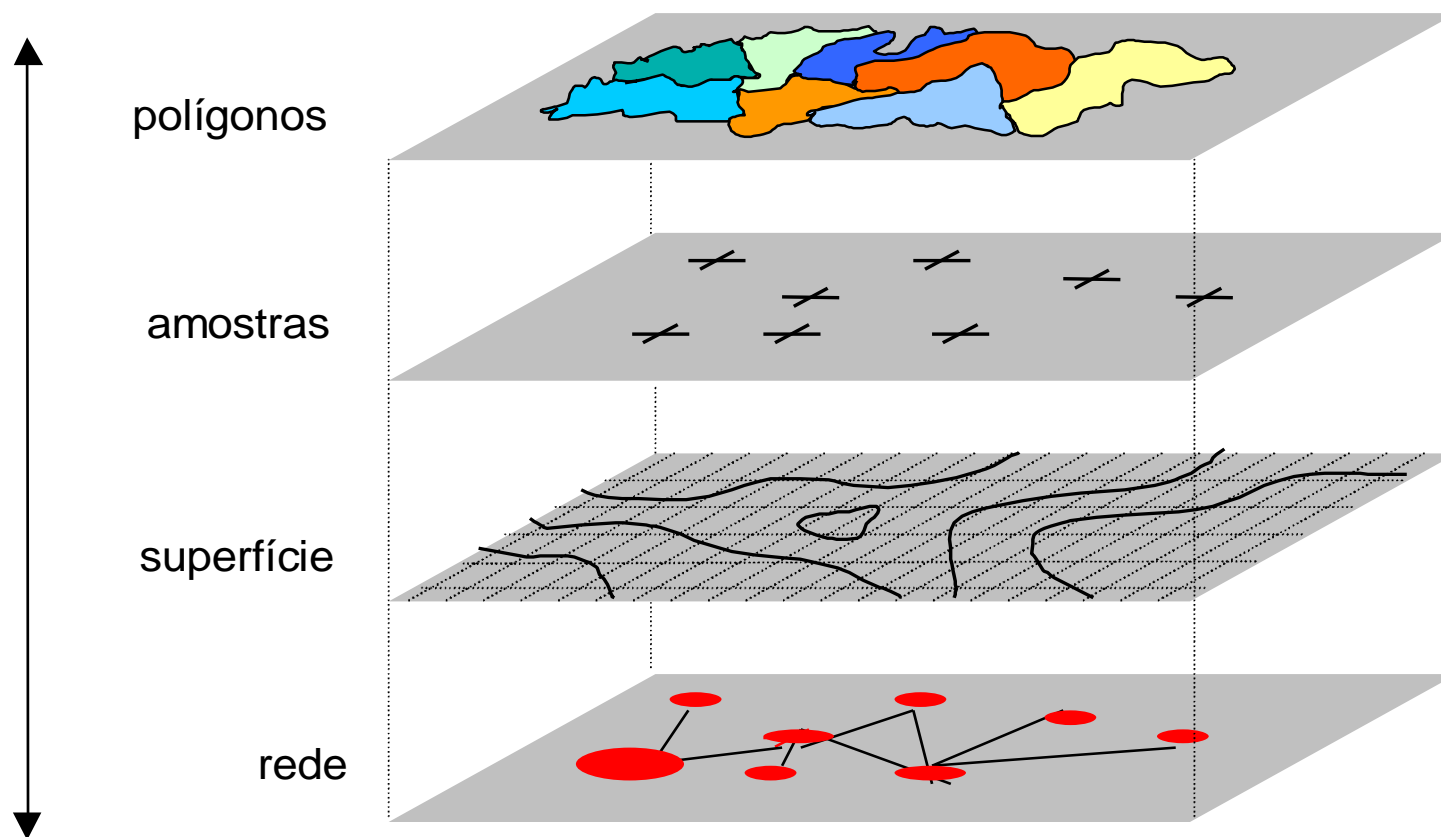
Ala.; Possum Trot, Ky; or Lonelyville, N.Y. But they are all probably close to someone's idea of paradise. —By Nancy Gibbs



Representação Computacional



- Representação digital de dados geográficos – modelos matricial e vetorial



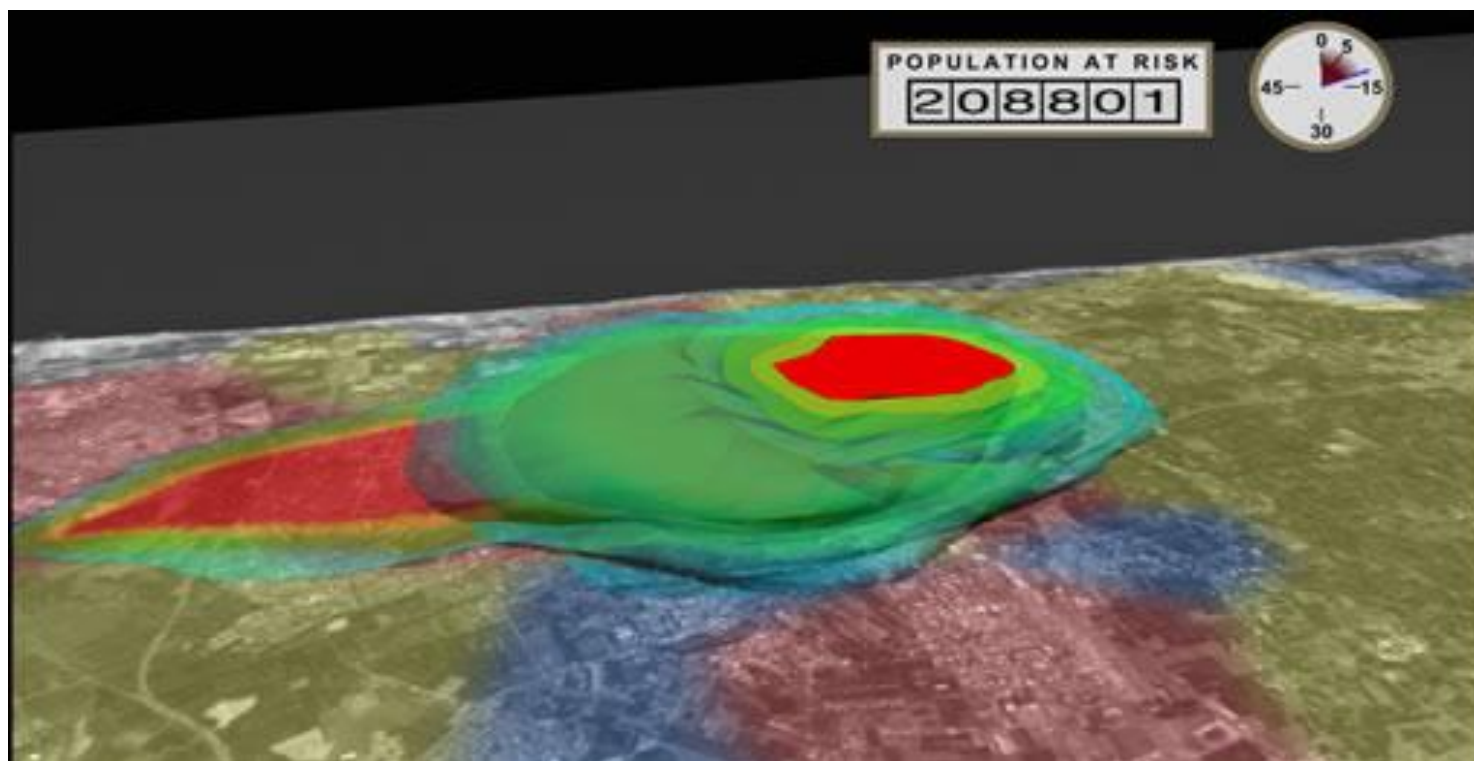
Representação Computacional



- ESCALA
- Modelos de representação computacional para dados sócio-econômicos:
 - Indivíduo
 - Área
 - Fenômeno contínuo no espaço
- Área – fenômeno em regiões discretas, unidades homogêneas
- Unidades – arbitrárias como setor censitário, não representam distribuição espacial da variável
- Problema de unidade de área modificável (MAUP) – séries temporais?

Superfície de Densidade

- Modelos de Superfície - alternativa a restrição das áreas



Superfície de Densidade

- Modelos de Superfície - alternativa a restrição das áreas
 - Densidade demográfica – fenômeno contínuo
 - Objetivo: estimar distribuição no maior detalhe possível
 - Vantagem: manipulação e análise independentes de área
 - Armazenar e disponibilizar dados em BD Globais

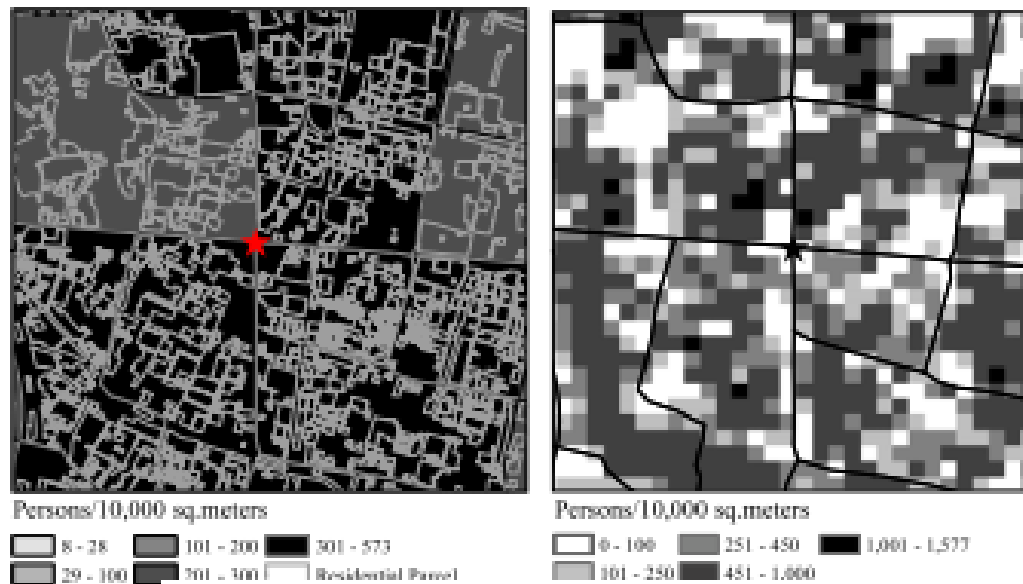


Figure 5 Comparison between vector-based subdistrict population distribution (left) and raster population surface (right) for CBD area

Luo (2005)

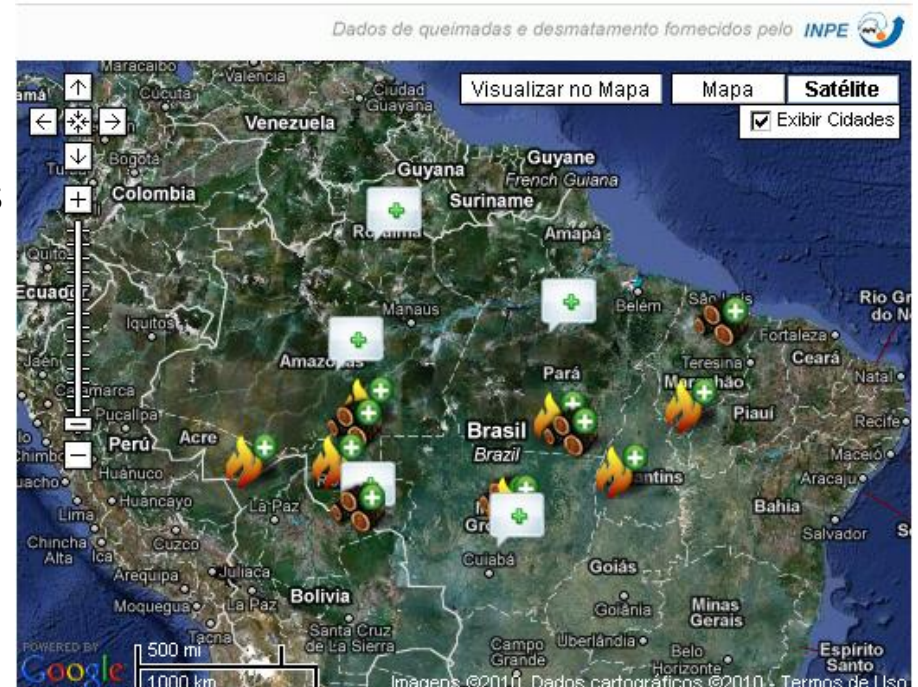
Para representar População na Amazônia

- Dados Censitários: Municípios ou Setor Censitário (Área)

EXEMPLO – “vazio demográfico”?!

- Evolução de uso e cobertura na Amazônia

- divisões territoriais
- grade regular para modelos espaciais
- Pressão da população
- gradiente da densidade populacional



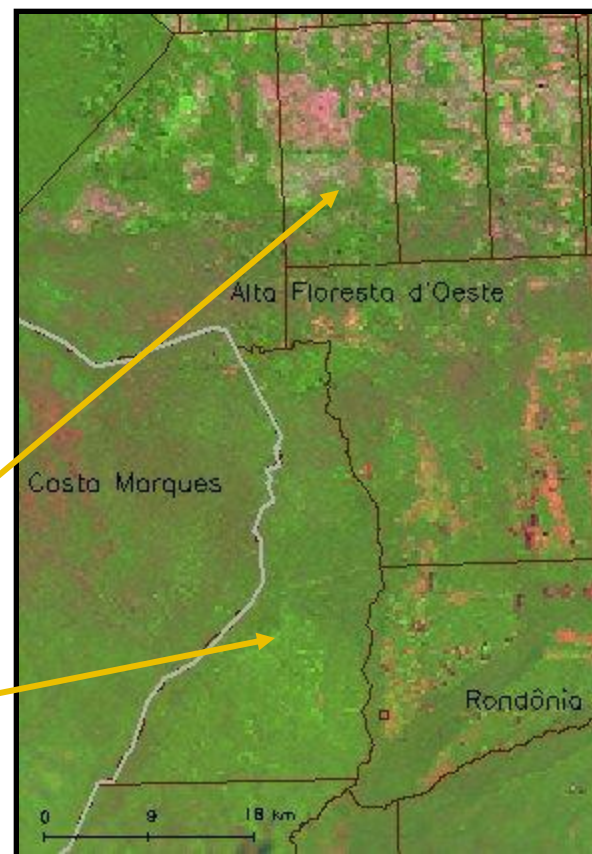
Para representar População na Amazônia

- Disponibilidade de dados
 - Dados censitários (10 anos)
 - Contagens por amostragem intercensos
 - Estimativas estatísticas – PNAD – UF, região metropolitana, no N somente para população urbana
- Indexação Espacial
 - Limite municipal até censo 2000, dados coletados por setor censitário (mapas analógicos), limite territorial oficial IBGE - malha municipal
 - Censo 2000 – setor censitário digital (apenas para área urbana – mun>25000). Atualmente disponíveis para 2000

Para representar População na Amazônia...

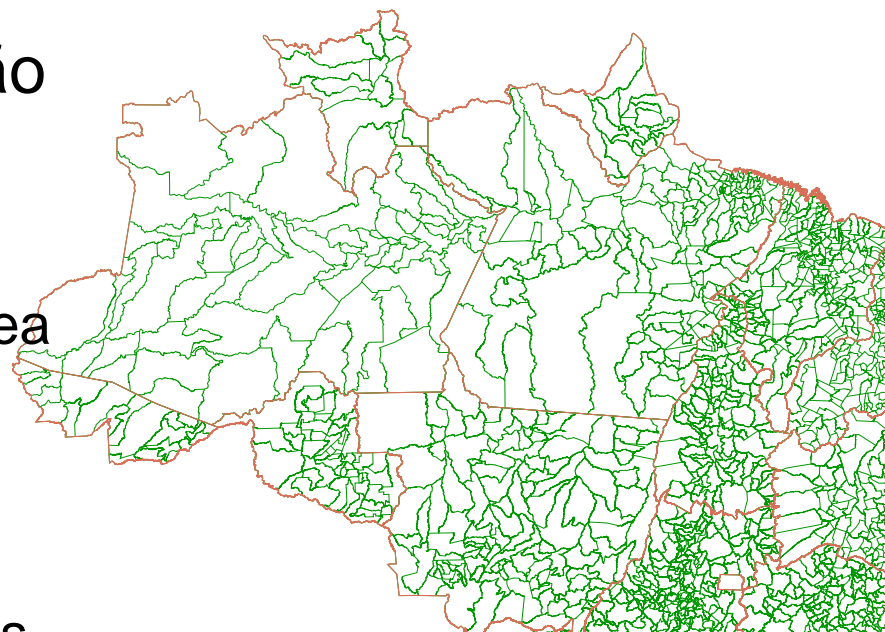
• Setor Censitário

- Área visitada em 1 mês – 250 domicílios rurais, 350 urbanos
- Na Amazônia – áreas extensas e heterogêneas
- Alta Floresta d'Oeste (RO)
 - 165 km² e limites regulares – assentamentos
 - 435 km² em áreas de floresta



Para representar População na Amazônia...

- Heterogeneidade da região
 - Tamanho de município: 64 km² – Raposa (MA), 160.000 km² – Altamira (PA)
 - Área dos municípios: 6.770 km² - média, 14.000 km² - desvio padrão
- Heterogeneidade da região
 - RO – 52 municípios com área média de 4.600 km²
 - AM - 62 municípios com área média de 25.800 km²
 - Área dos municípios determinam a dimensão dos setores censitários



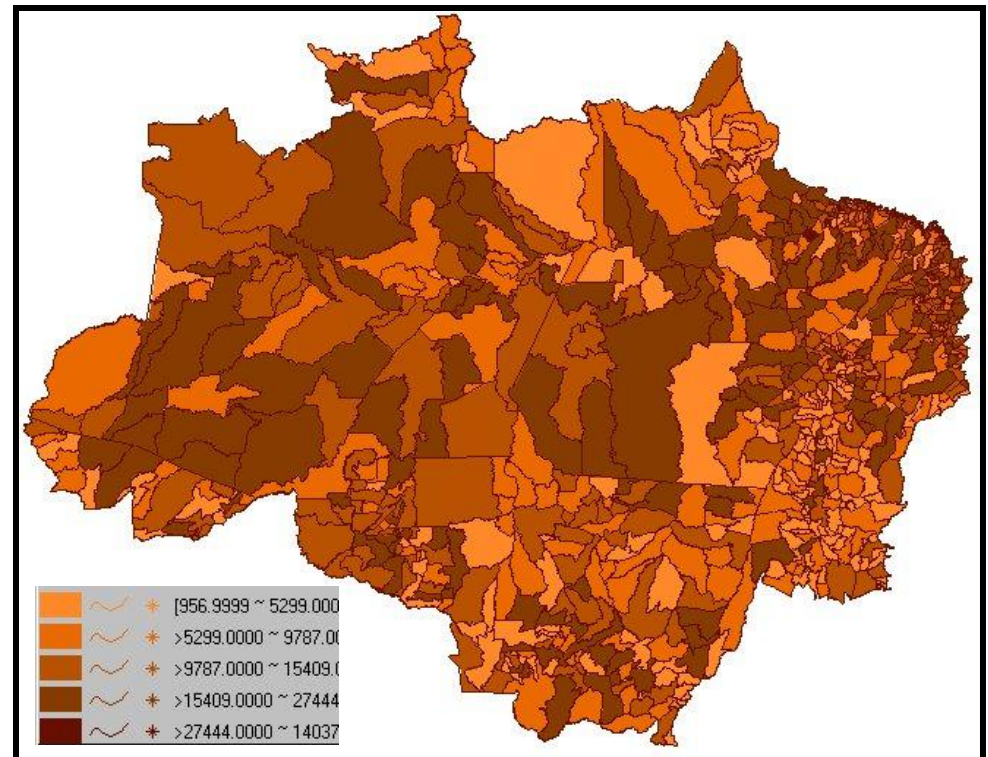
Para representar População na Amazônia...

- **Complexidade dos processos -> distribuição espacial**

- Rondônia: migrantes, assentamentos INCRA, núcleos urbanos ao longo de eixos viários e população na zona rural.
- Amazonas: menor densidade de núcleos urbanos, concentração em Manaus.
- Tendências:
 - desconcentração das metrópoles,
 - maior participação relativa das cidades de até 100.000 habitantes
 - crescimento da população nos núcleos de 20.000 habitantes
- População dispersa na zona rural e ao longo dos rios
- Contínuos de florestas – vazios demográficos

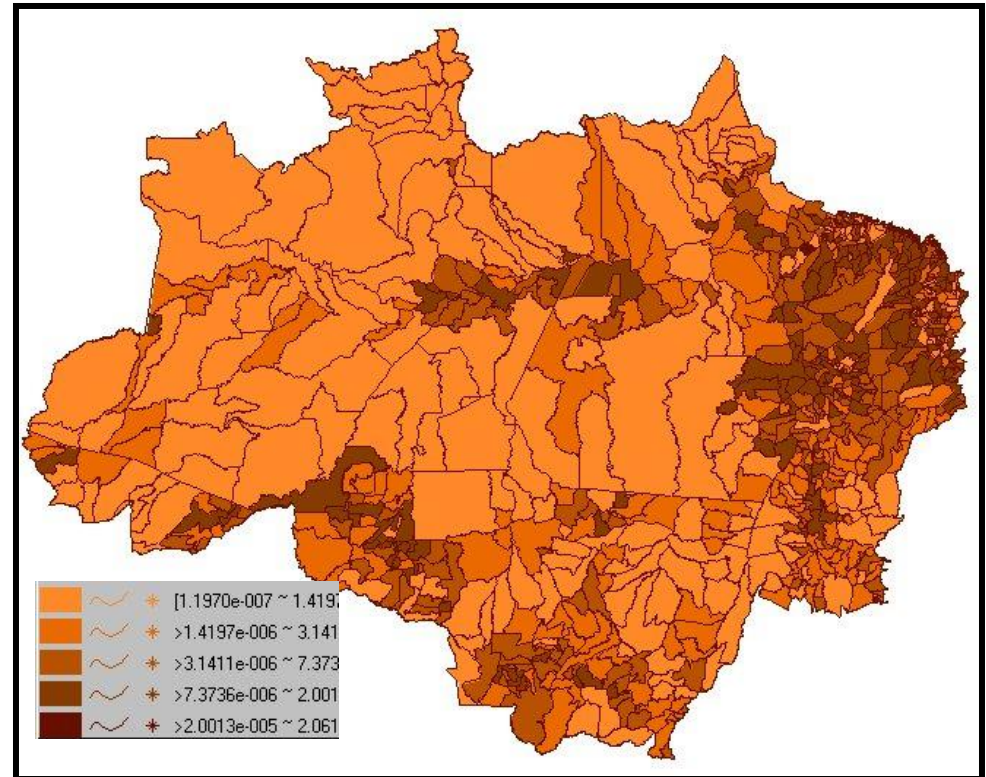
Modelos de População

- Dimensão Humana – importância nos projetos da região como LBA e LUCC....
- Representação mais frequente: mapas temáticos



Modelos de População

- Densidade Demográfica ao invés de população total 2000
- Importância dos intervalos e critério para visualização
- Destaque das regiões mais populosas e vazios demográficos ??



Técnicas de “interpolação” de Superfícies classificadas em dois grupos:

Interpoladores que consideram **apenas a variável população**:

- Superfície de Tendência, Ponderado pela área, Krigagem, Picnofilático de Tobler e Centróides Populacionais de Martin

Interpoladores que **utilizam variáveis auxiliares**, indicadoras da presença humana para distribuir população nas superfícies

- método Dasimétrico, Interpoladores Inteligentes e variantes

Modelos de População em Superfícies

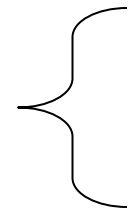
Baseados em Pontos – *Point-based*

Baseados em Áreas – *Areal-Based*

Informação Auxiliar

Conhecimento dos Processos

População



Preservar Volume
Restrições

Base Territorial - Censos 2000 e 2007

27 Unidades da Federação

162.769 Setores Coleta

92.685 Urbanos

70.084 Rurais

249.068 215.811 (*)

Setores Censitários

5.564 5.507 (*)

Municípios

10.088 9.848 (*)

Distritos

471 408 (*)

Subdistritos

9.588 7.459 (*)

Bairros

4.086 3.905 (*)

Aglomerados Subnormais

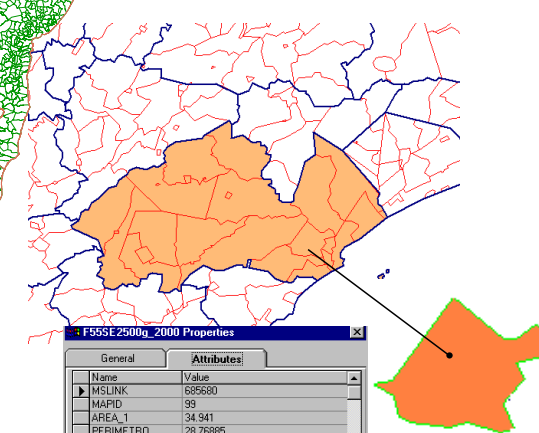
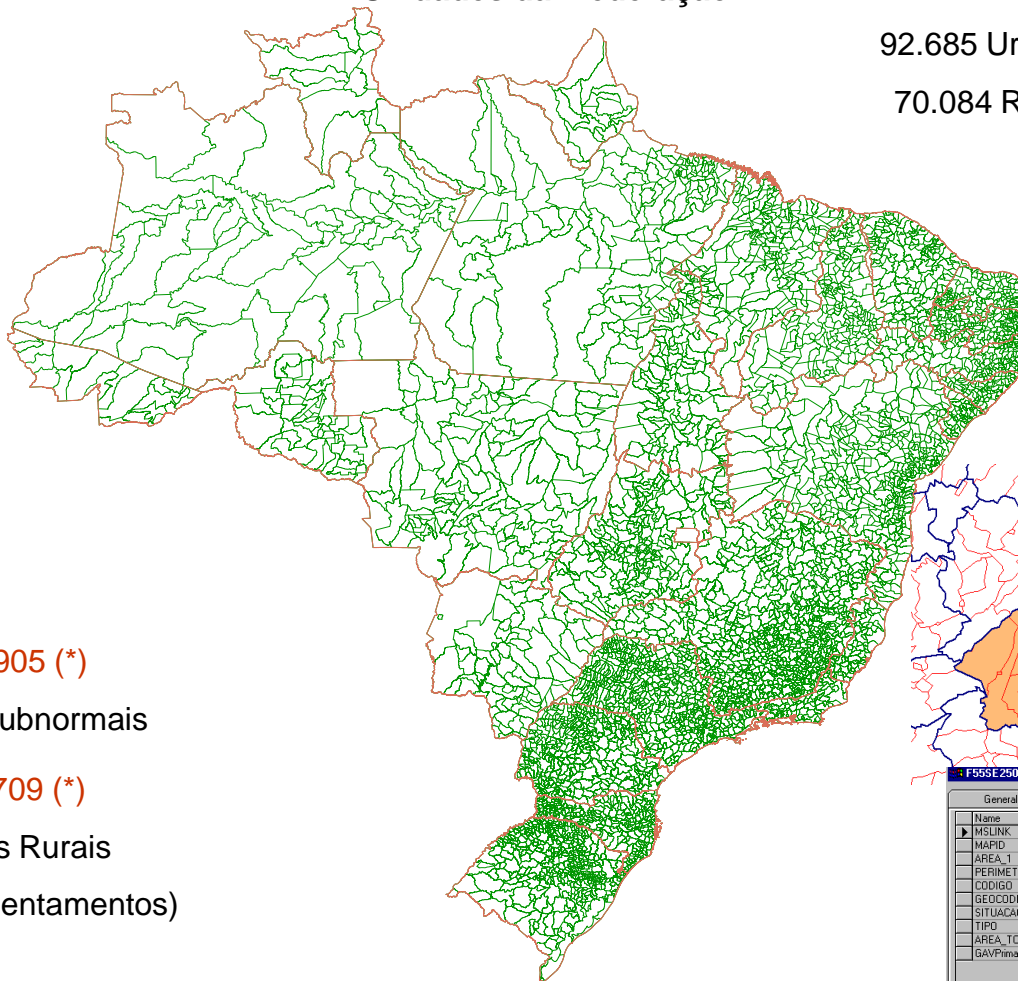
9.245 6.709 (*)

Aglomerados Rurais

(sendo 1.671 Assentamentos)

1.047 TI e UCA

872 (*) AESP

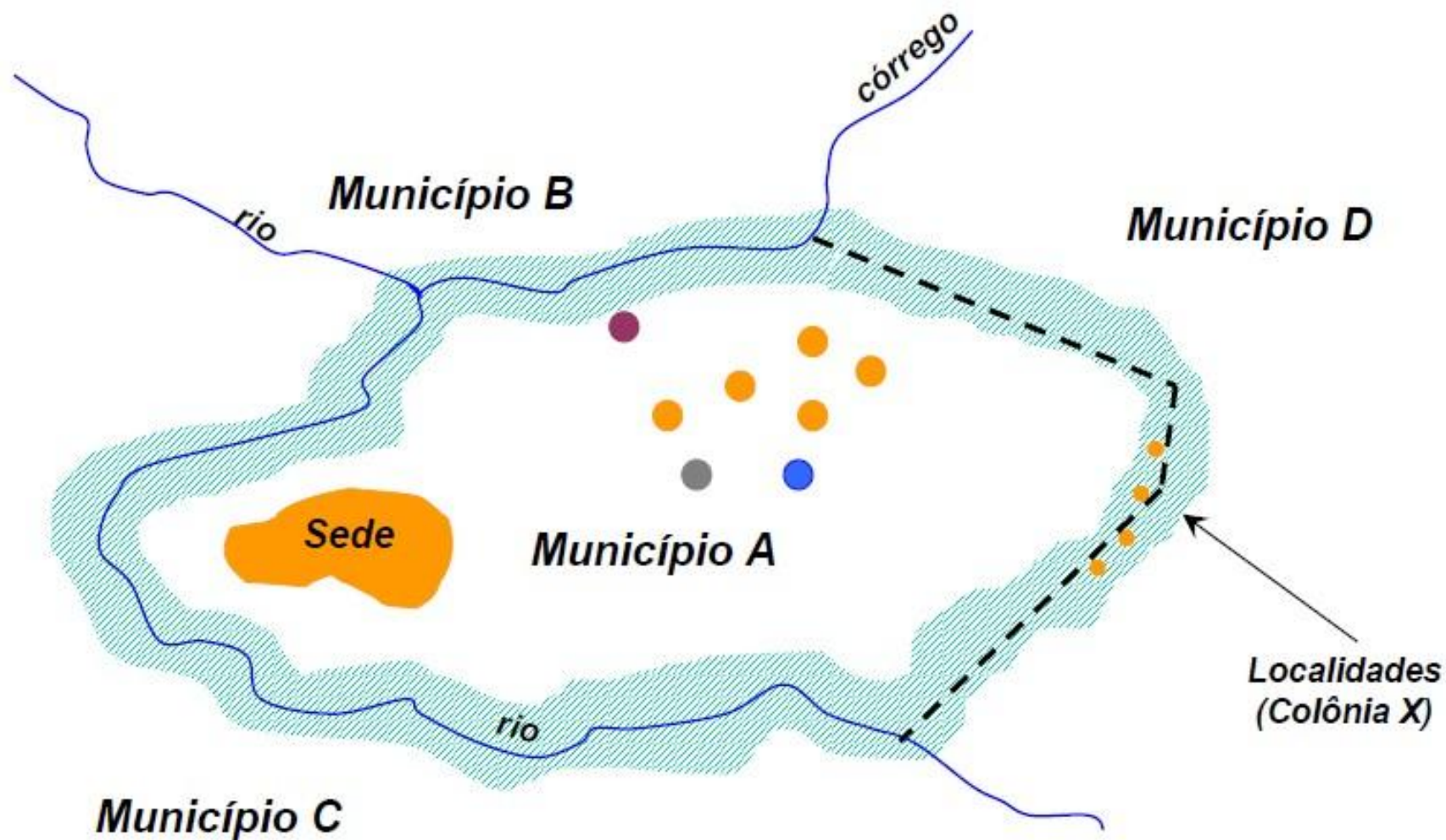


F55SE2500g_2000 Propriedades	
General	Attributes
Name	Value
MSLINK	686680
MAPID	99
AREA_1	34.941
PERIMETRO	28.76385
CODIGO	330240305000003
GECCODIGO	330240305000003
SITUACAO	8
TIPO	0
AREA_TOT_G	34.941
BAVPrimaryKey	32322

Como o Dado é Disseminado ?

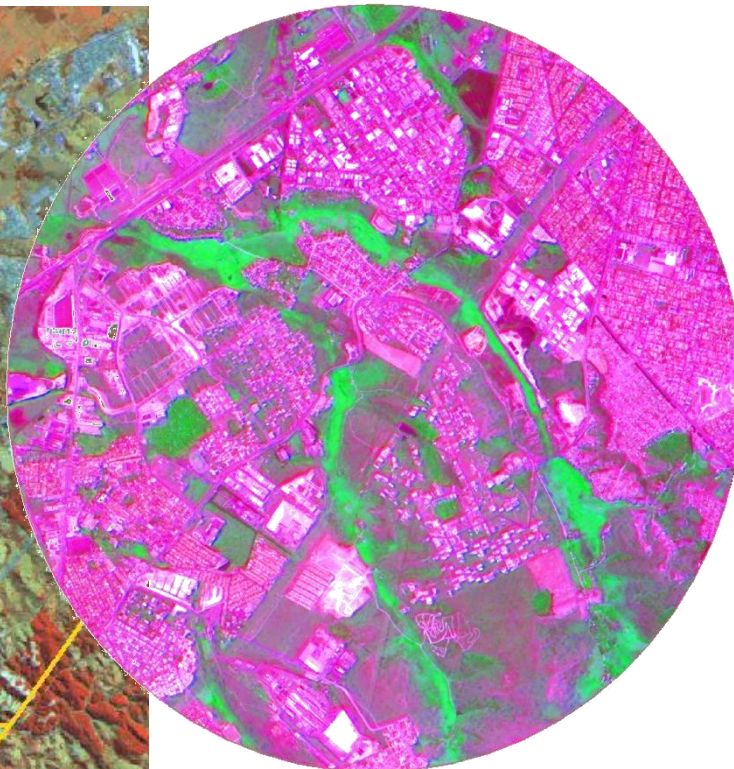
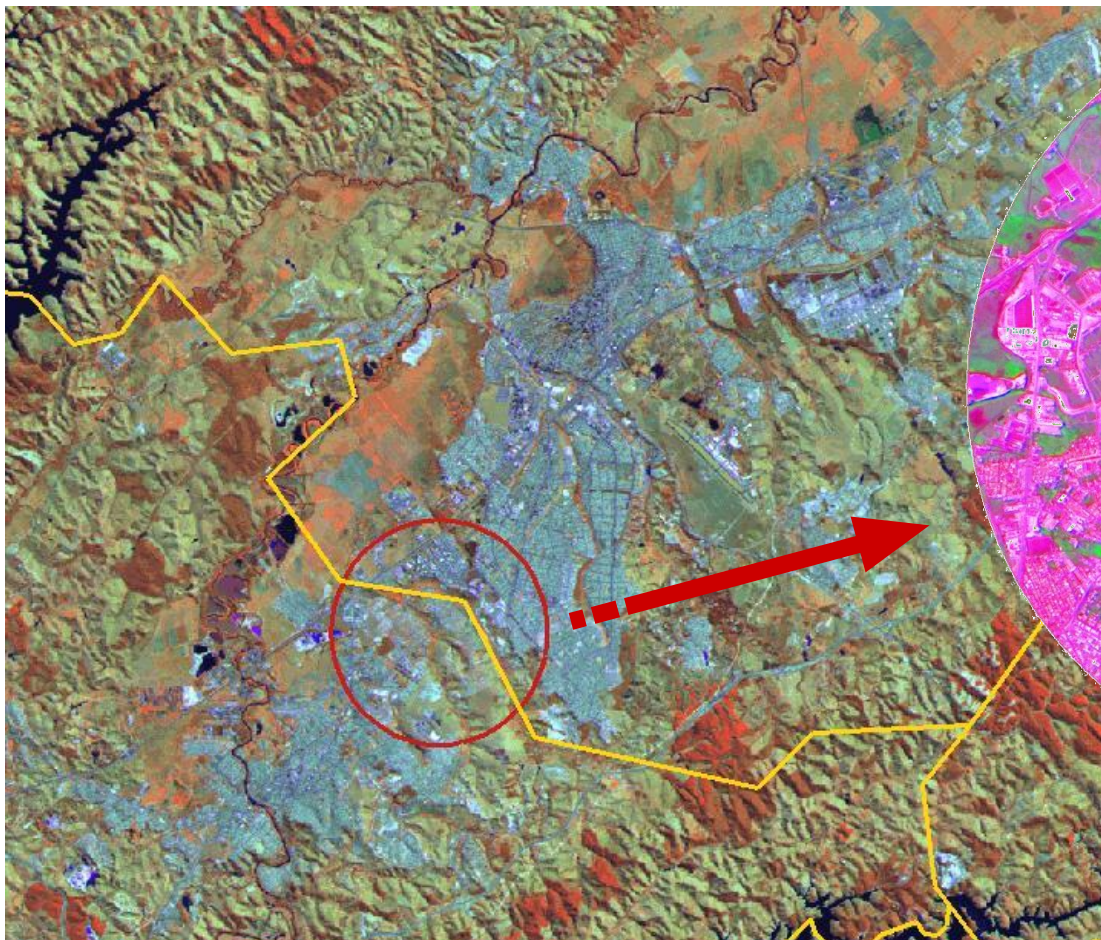
Malha Setorial do Município de Macaé

(*) Números da Base Territorial 2000

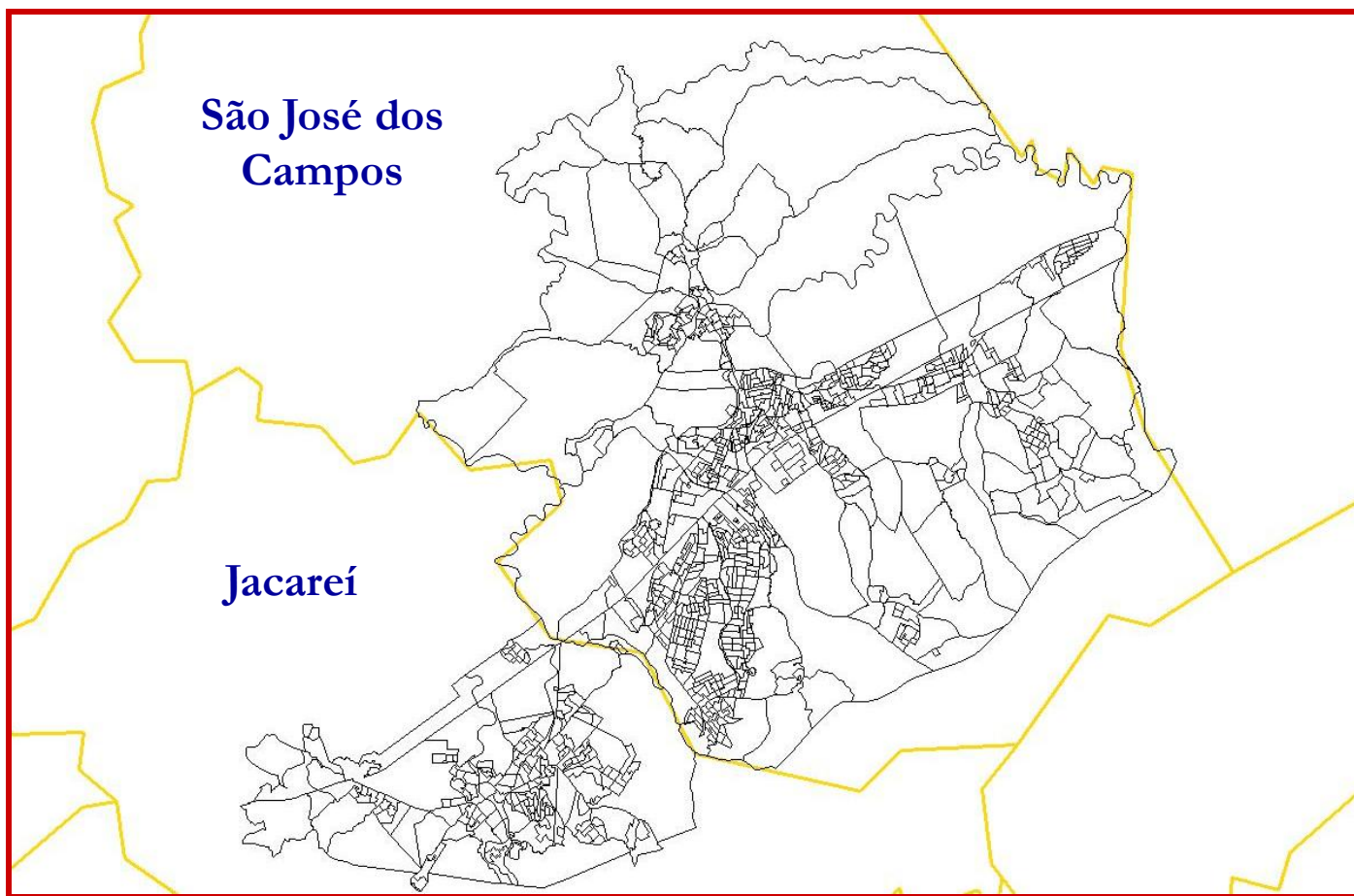




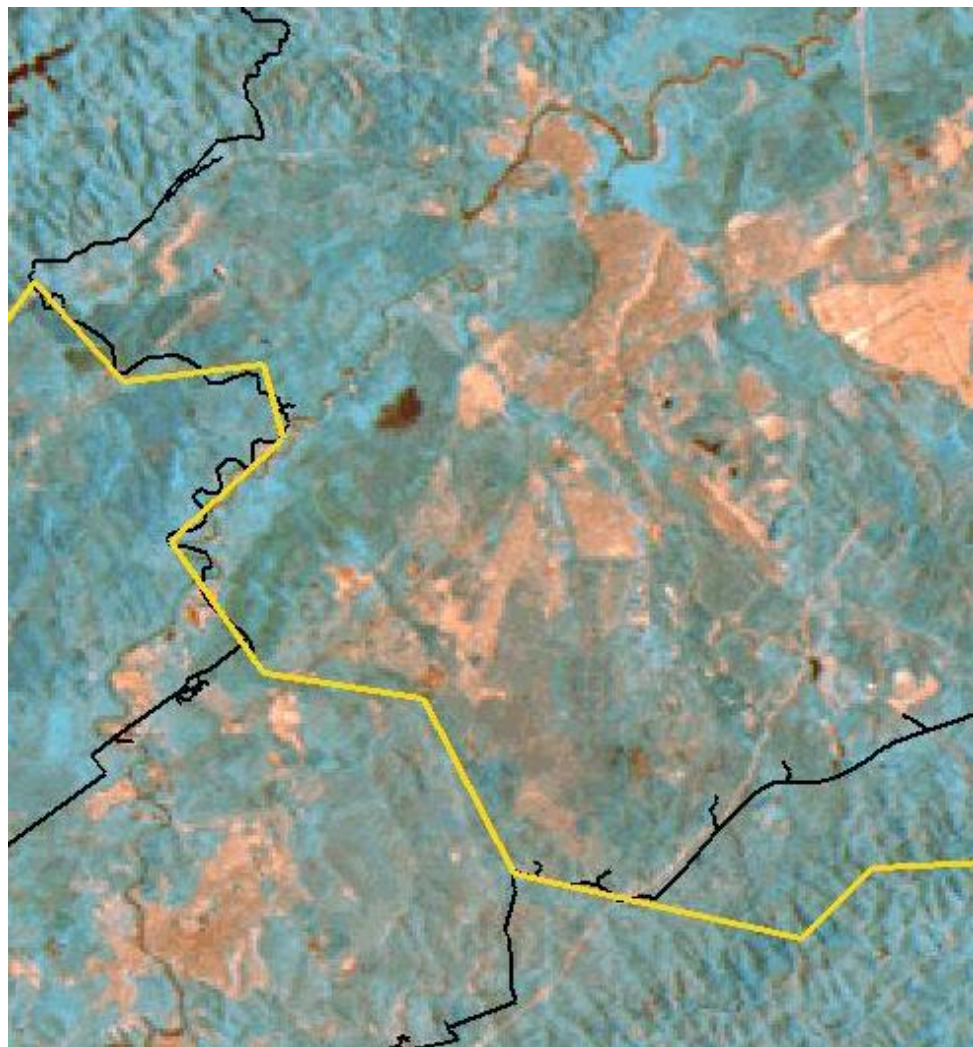
Conurbação

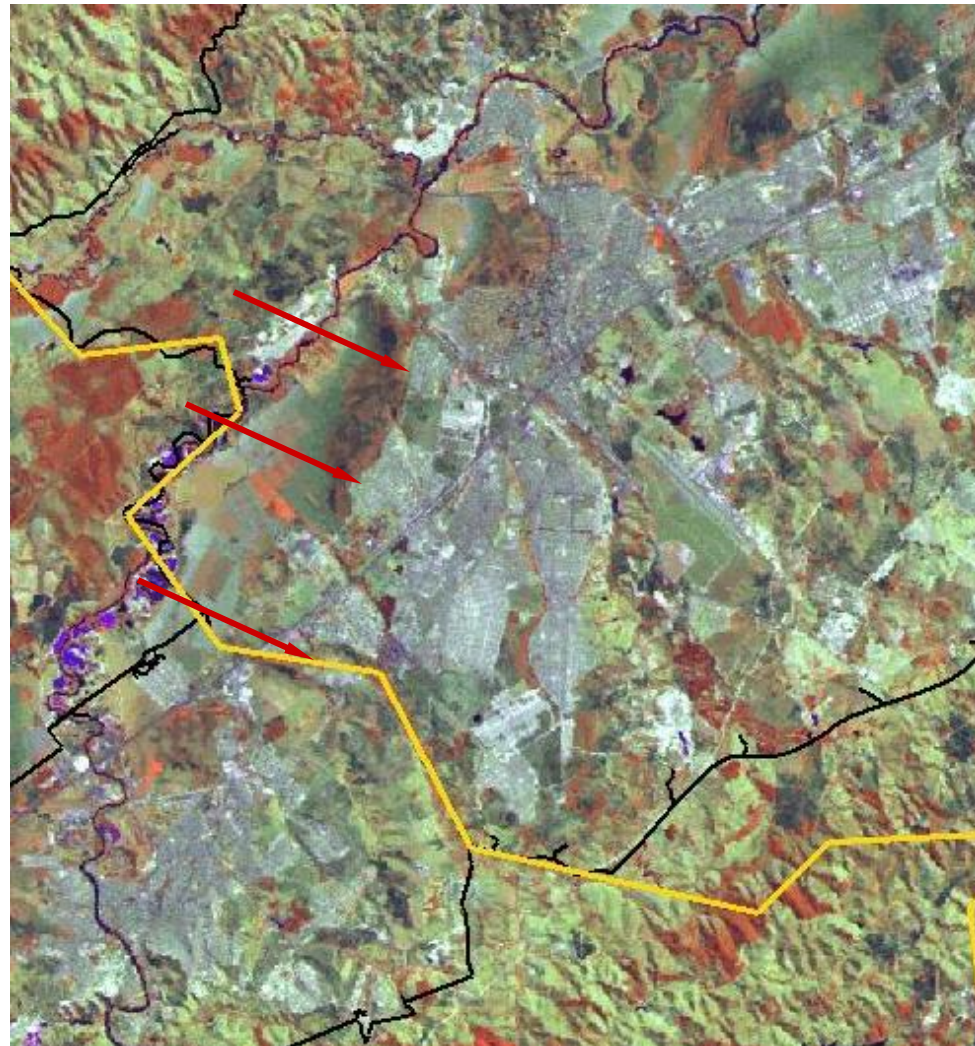


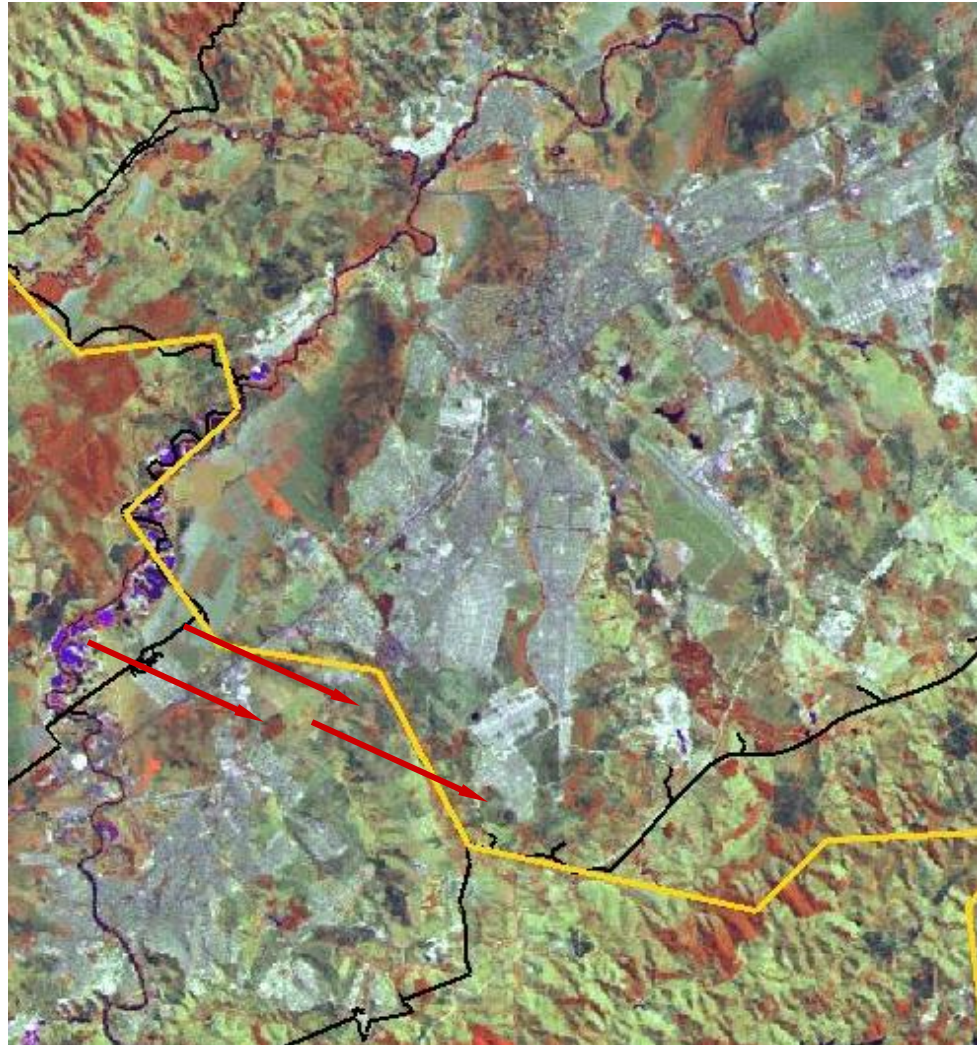
Caso: São José dos Campos - Jacareí

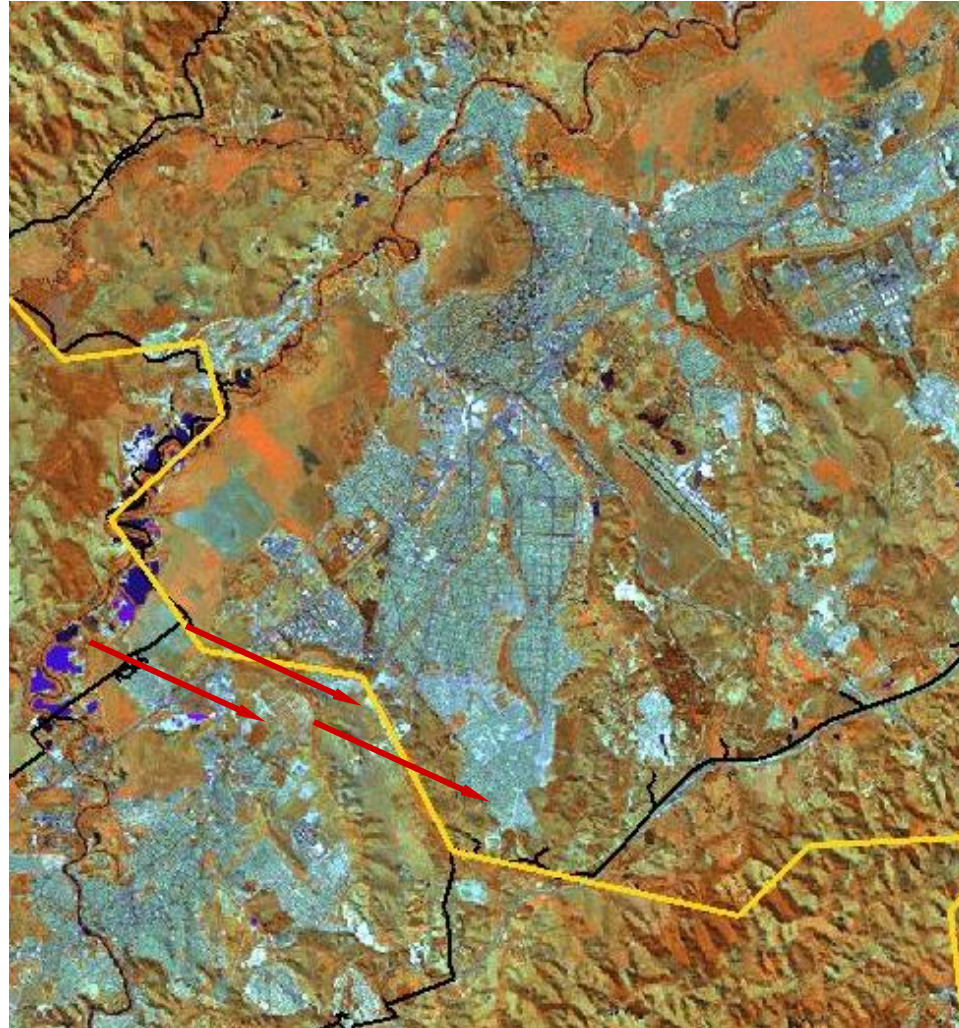


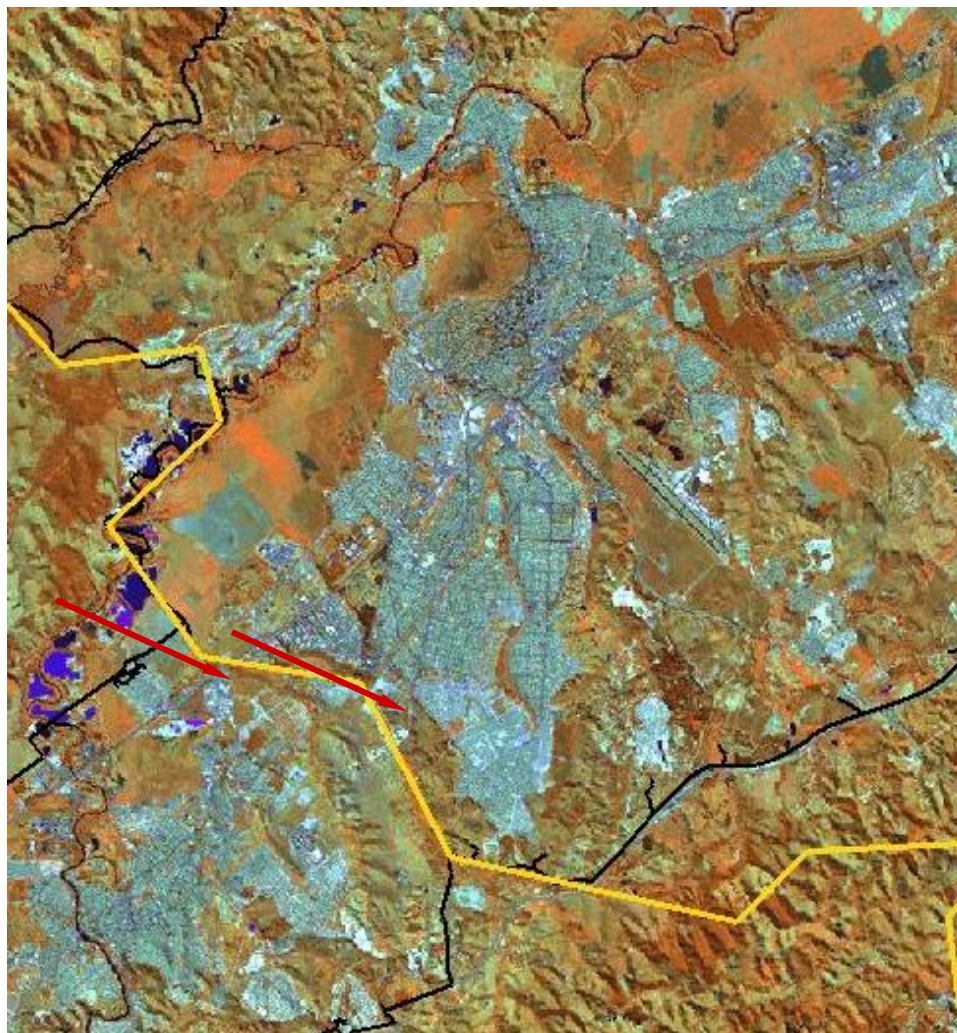
Setores censitários não rurais (2000)

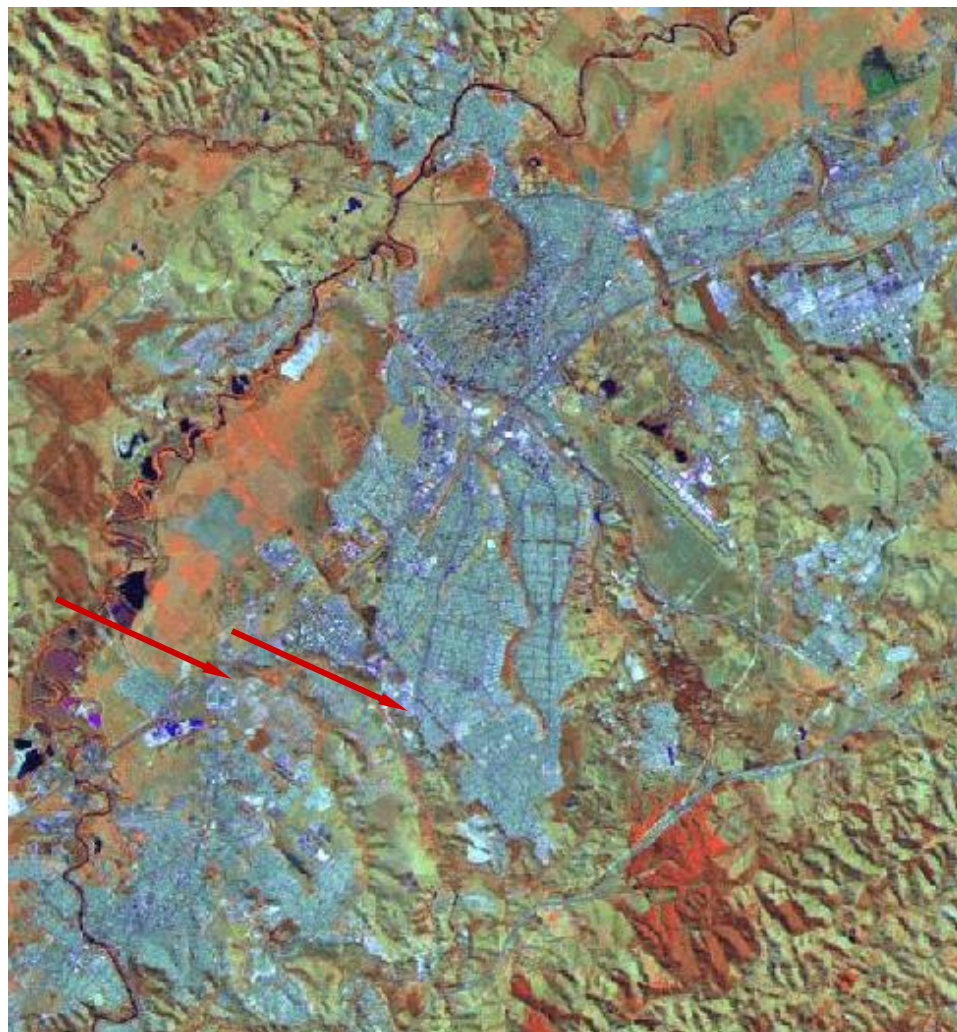


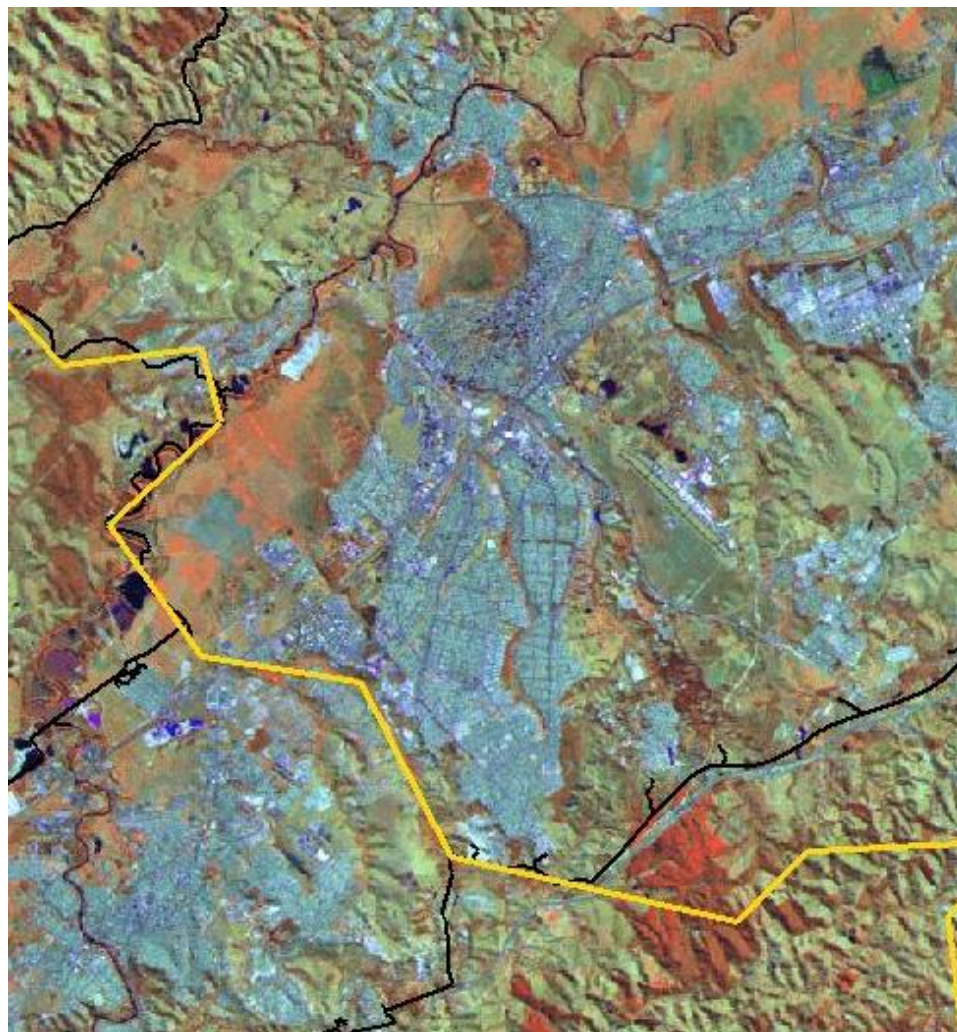


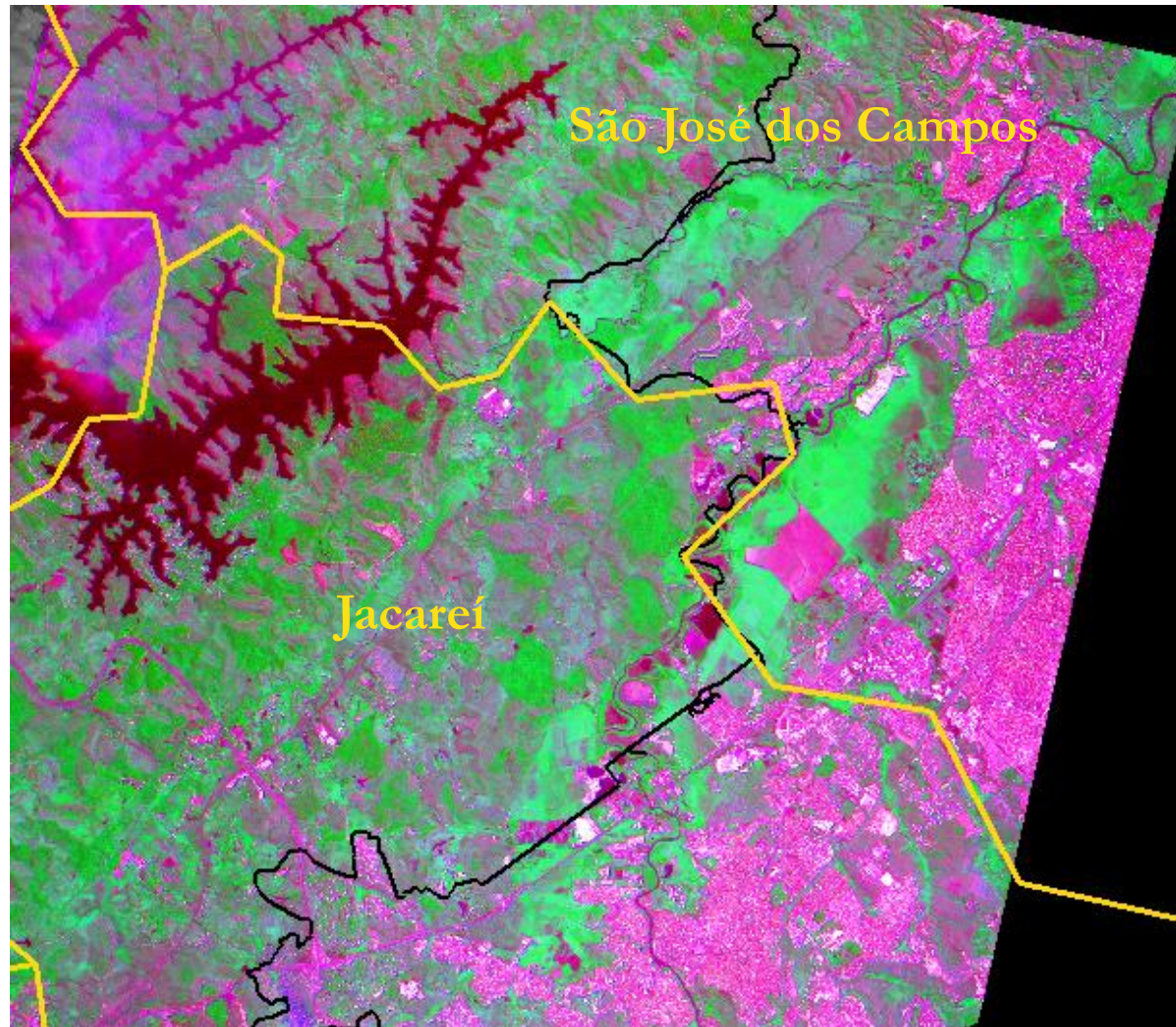


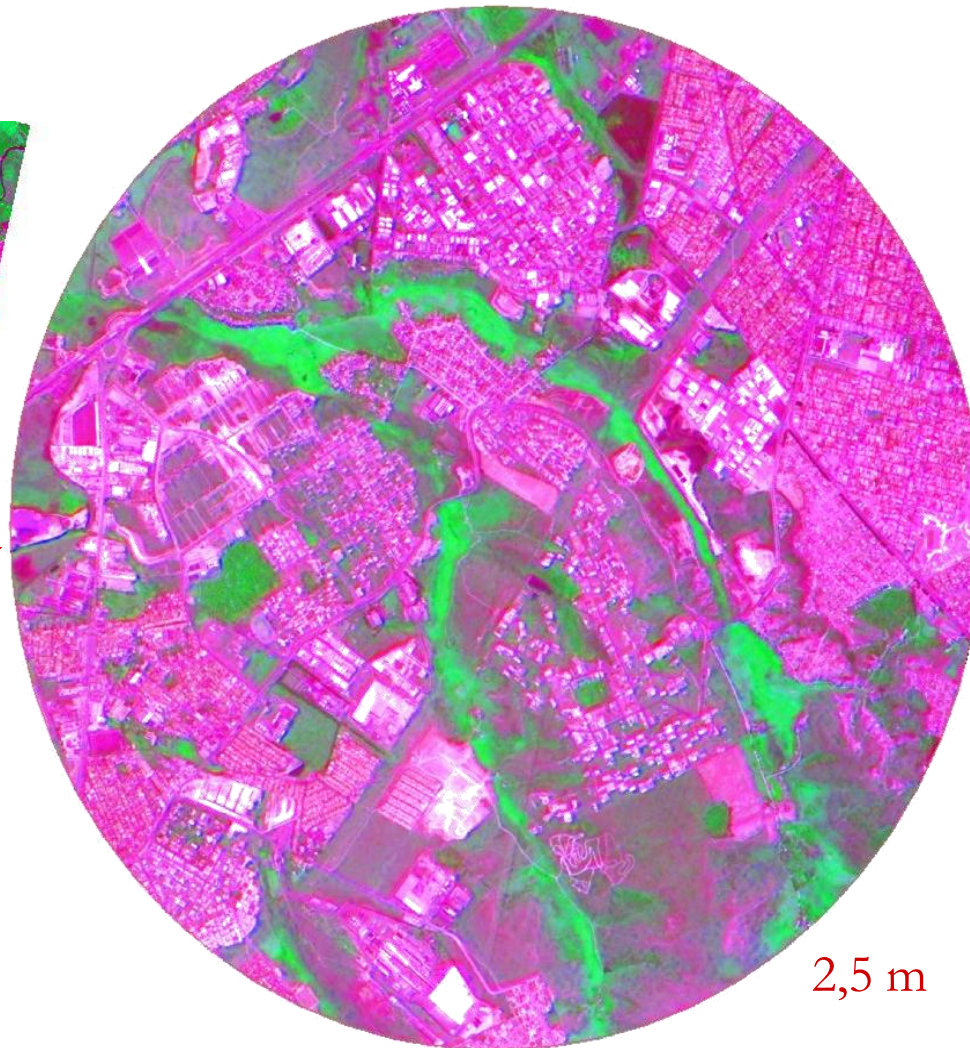
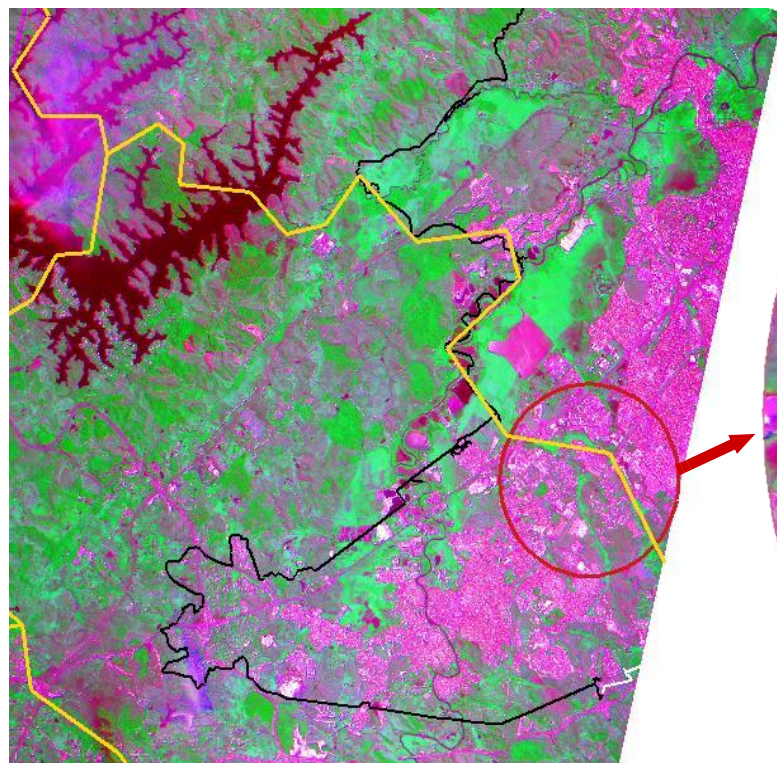






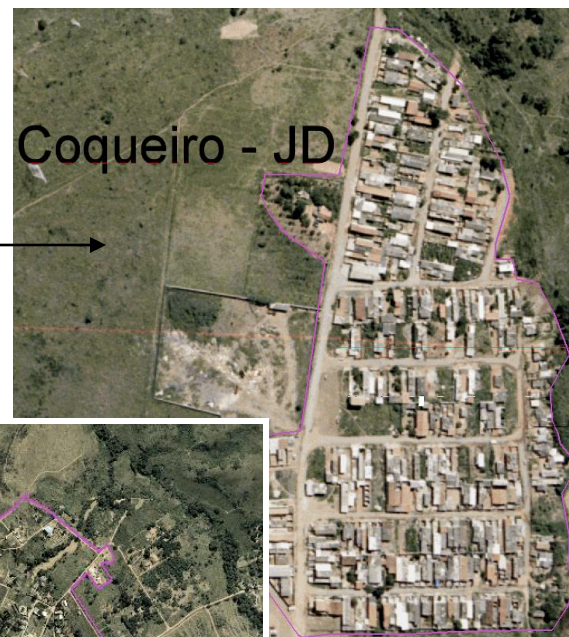
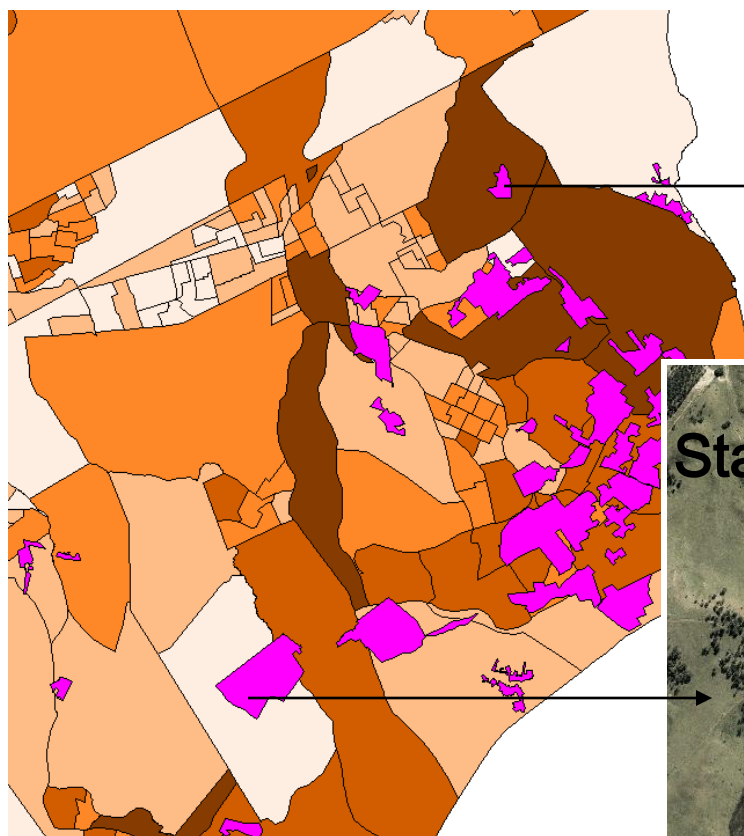






2,5 m





Interpoladores *Geográficos*

Modelos de População em Superfícies

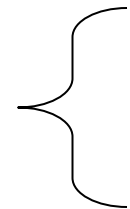
Baseados em Pontos – *Point-based*

Baseados em Áreas – *Areal-Based*

Informação Auxiliar

Conhecimento dos Processos

População



Preservar Volume
Restrições

População, Espaço e Ambiente

Abordagens Espaciais em Estudos de População: Métodos Analíticos e Técnicas de Representação

Parte IV – Integração de Dados

4. De pontos a superfícies: geoestatística

Silvana Amaral

Antonio Miguel V. Monteiro

{silvana@dpi.inpe.br, miguel@dpi.inpe.br}



Referência

Tobler, W. R. Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association*, 74, 519-530, 1979

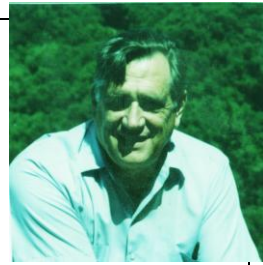
Geographical Interpolation

Waldo Tobler
Professor Emeritus
Geography Department
University of California
Santa Barbara, CA 93106-4060
<http://www.geog.ucsb.edu/~tobler>

- CSISS Summer Institute, 2001



Interpolação Geográfica



- Interpolação requer pressupostos sobre a natureza do fenômeno a ser interpolado...
...seguido por um “intelligent” guessing
- Por exemplo:

Amostragem aleatória seguida por interpolação seriam eficientes para representar esta superfície ??



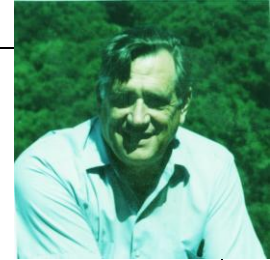
Outro desafio para interpolação



Amostras de elevação
em localidades
aleatórias em uma
cidade e depois
interpolação.

O que teríamos?





Interpolação de áreas é um caso especial de Interpolação Geográfica

Observações numéricas dadas em localizações pontuais

- Objetivo é produzir um mapa de contornos (isolinhas)
- Há vasta literatura para interpolação a partir de dados pontuais
- Clássicos: Kriging, inverso da distância, splining, etc.

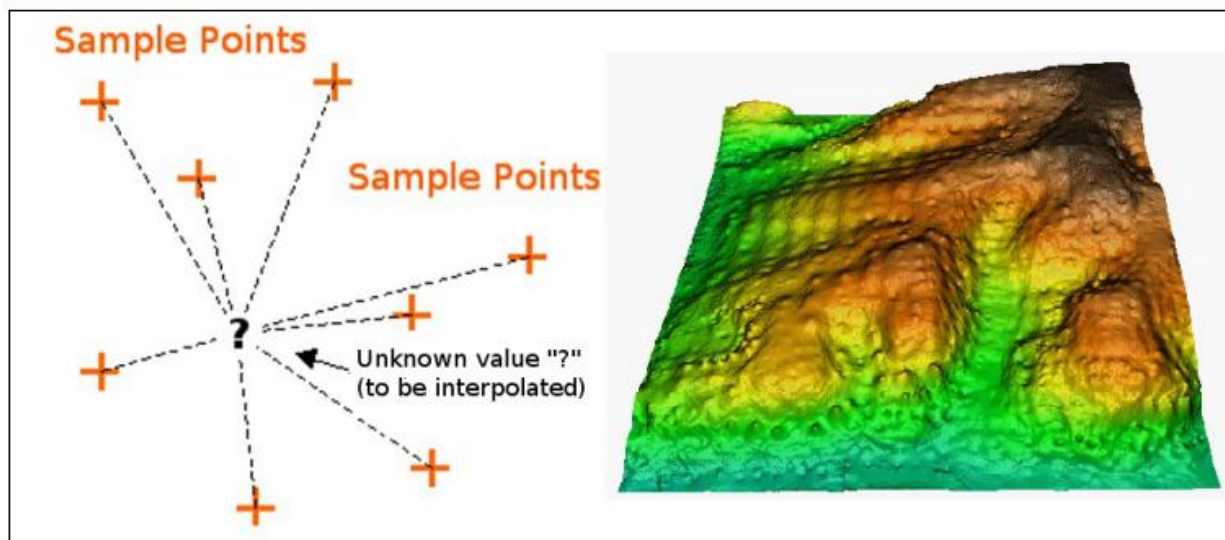


Illustration 2: Inverse Distance Weighted interpolation based on weighted sample point distance (left). Interpolated IDW surface from elevation vector points (right). Image Source: Mitas, L., Mitasova, H. (1999)).

Referência



Geoestatística para geoprocessamento

Eduardo C. G. Camargo

Geoestatística para geoprocessamento

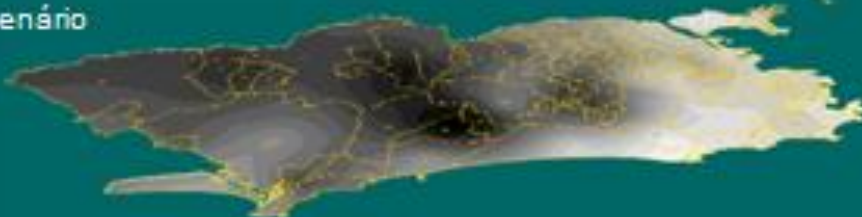
realidade



análise estrutural



cenário



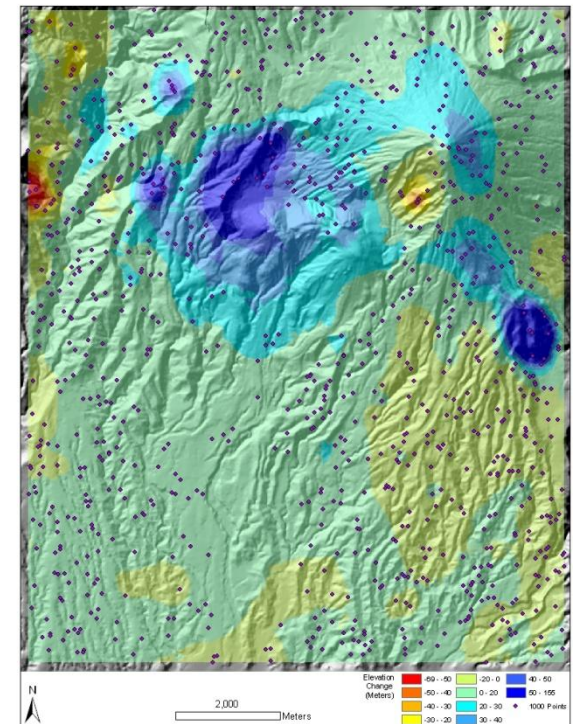
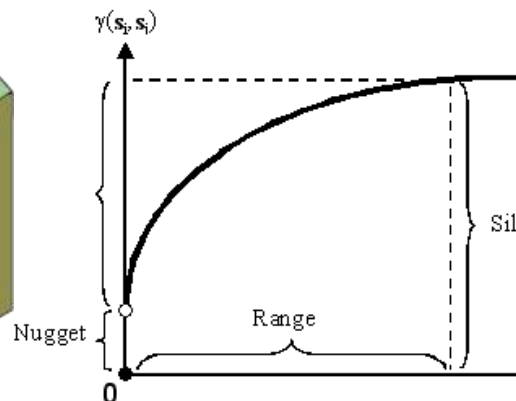
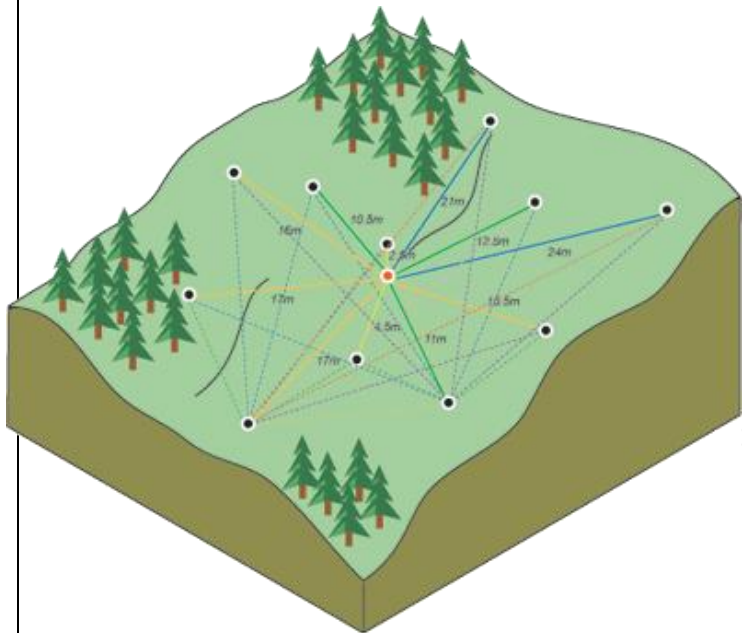
Palavras-chave

- variáveis regionalizadas
- semivariograma empírico
- efeito pepita, alcance e patamar
- análise estrutural
- isotropia e anisotropia
- validação cruzada
- krigagem

Modelos de População “Univariados”

- Krigeagem

- Interpolação para processos espaciais aleatórios, estima a ocorrência de um processo a partir da observação em outros locais
- Os valores da variável dependem da distância entre elas, uma função explica esta distribuição



O que é geoestatística?



É uma abordagem PROBABILÍSTICA de modelagem, que engloba um conjunto de métodos estatísticos, para a análise e mapeamento de dados distribuídos no espaço e/ou no tempo.

Requer o conhecimento de alguns conceitos básicos:

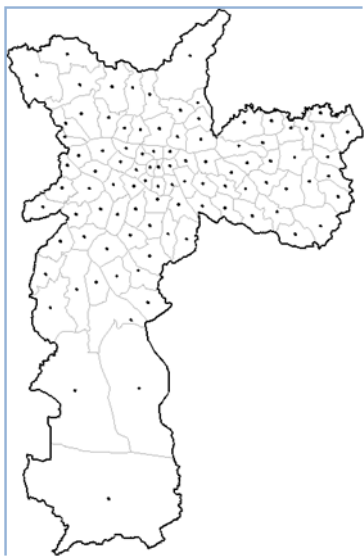
- Variável aleatória (V.A.)
- Momentos da V.A. Exs: $E[X]$, $C[X,Y]$;
- Função densidade de probabilidade (FDP);
- Função de Distribuição Acumulada (FDA): univariada e bivariada;
- Função aleatória (FA), etc.

- 1) Análise: objetiva descrever a variabilidade espacial do fenômeno em estudo, denominada de análise estrutural ou modelagem do semivariograma.
- 2) Inferência: objetiva estimar valores de uma variável distribuída no espaço em locais não amostrados, denominada de krigeagem.
- 3) Simulação: objetiva construir um conjunto de realizações equiprováveis ou igualmente representativa do fenômeno em estudo.

Etapas da modelagem geoestatística

Etapas da modelagem geoestatística

Região de estudo

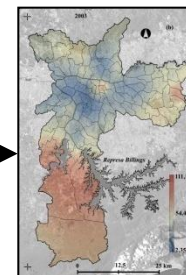


análise exploratória

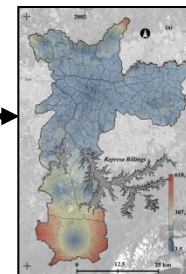
análise estrutural

interpolação
krigeagem

simulação
condicionada

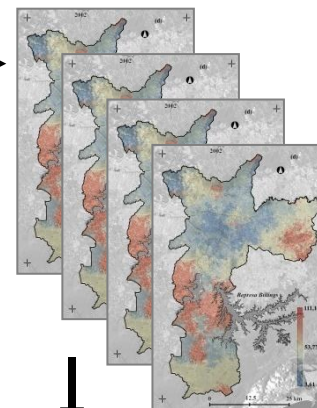


Superfície estimada
do fenômeno
investigado



Superfície da variância
da estimativa

Realizações equiprováveis



- Construção de cenários
- Mapas de incerteza

Porque usar geoestatística?



Interpoladores

Procedimentos determinísticos

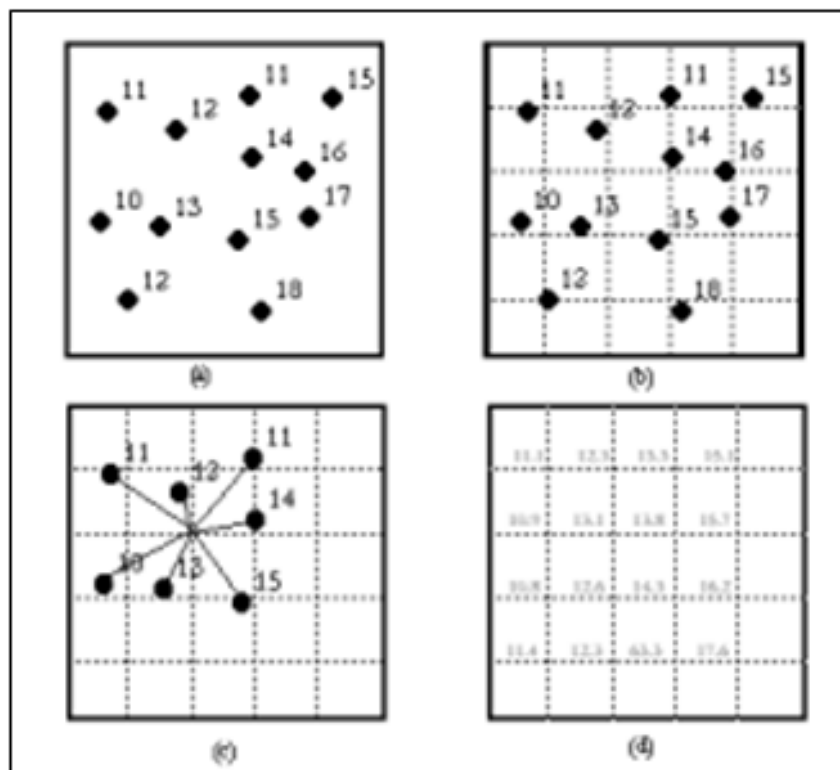


Figura 3-3 Ilustração do processo de interpolação por estimador local: (a) configuração original de amostras; (b) grade regular superposta às amostras; (c) interpolação de um valor a partir dos vizinhos; (d) grade regular resultante

Porque usar geoestatística?



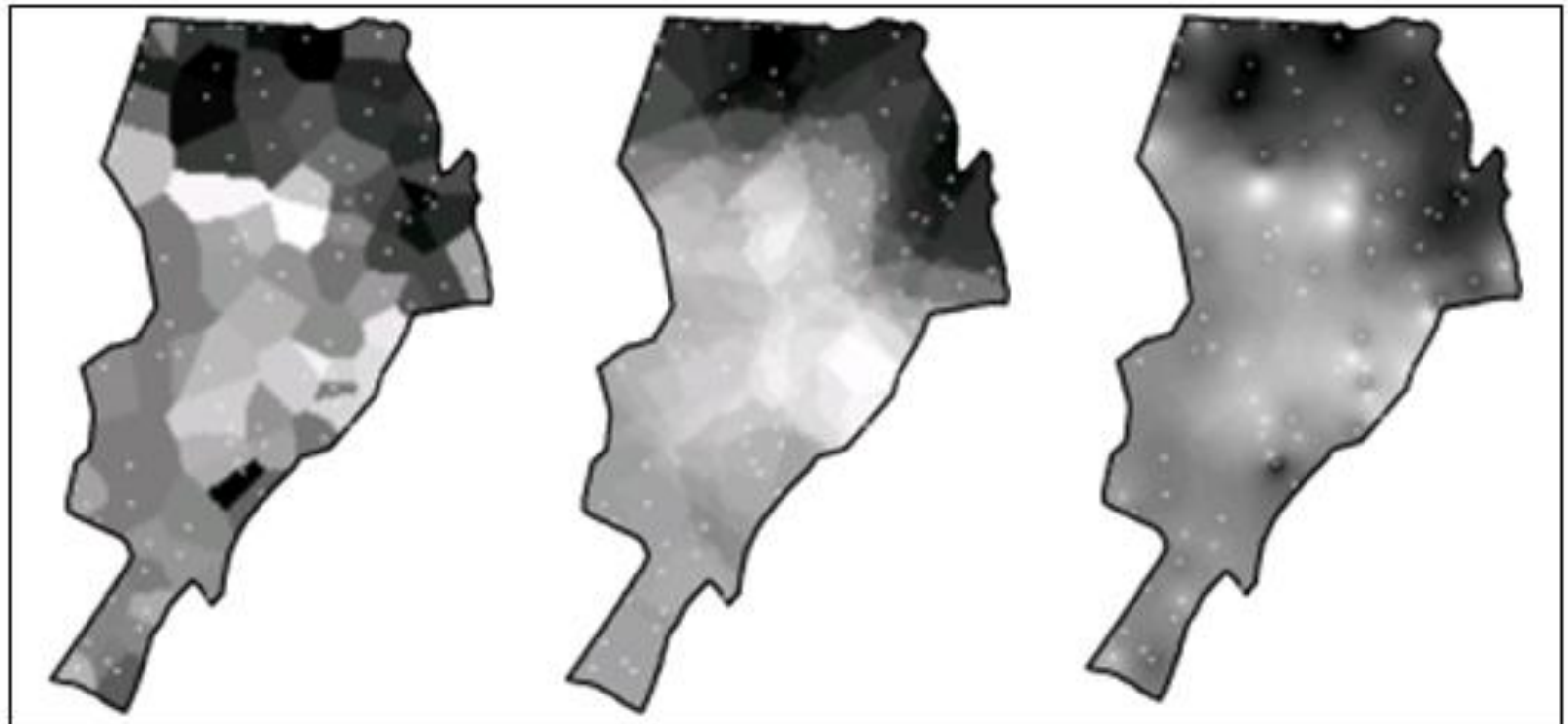
Interpoladores

Procedimentos determinísticos

Vizinho mais Próximo

Média Simples

Inverso quadrado distância

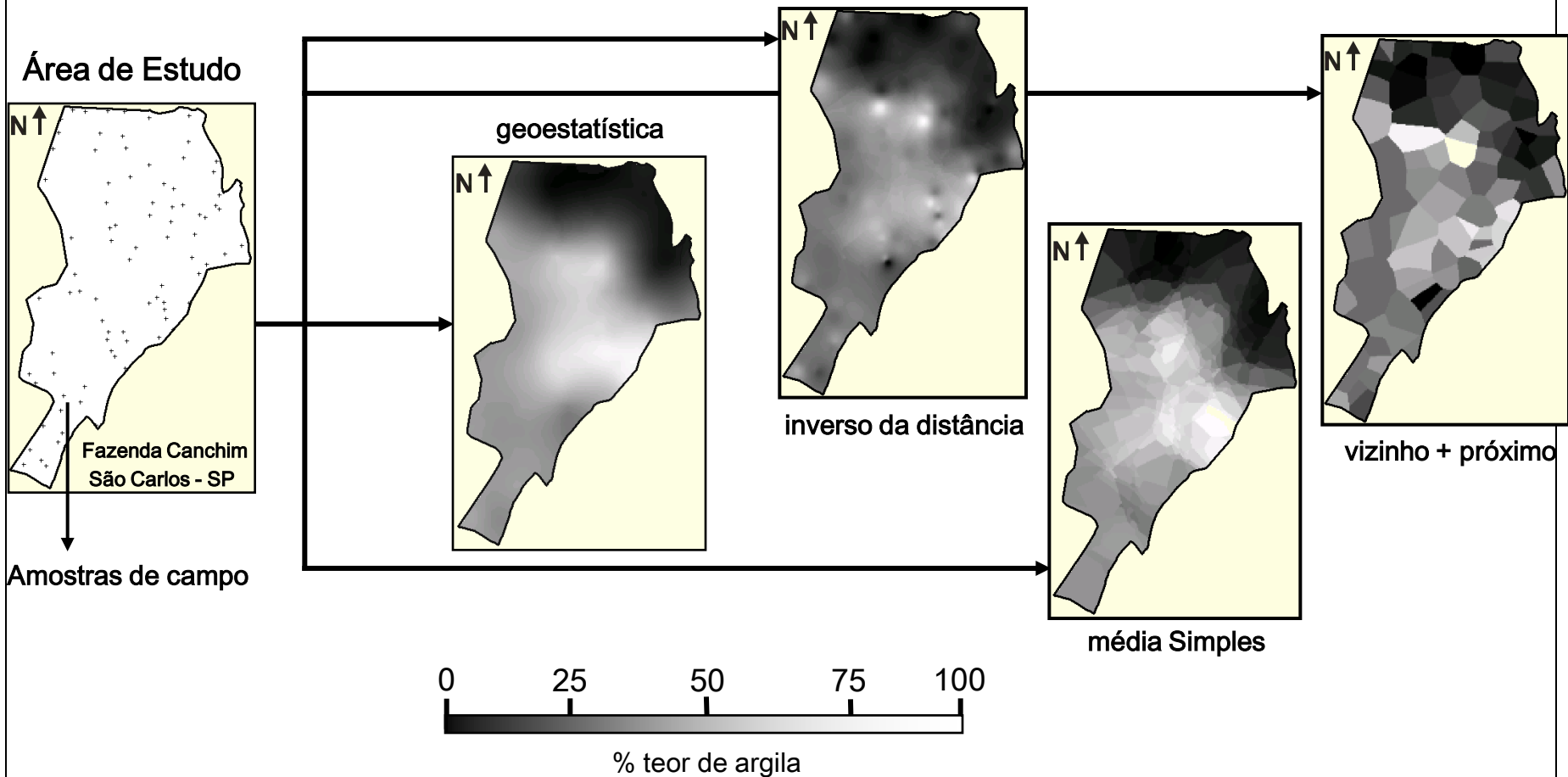


Porque usar geoestatística?



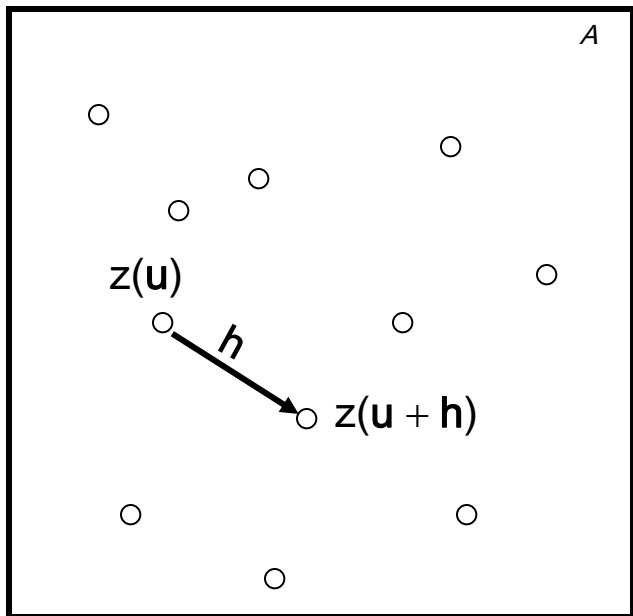
Interpoladores

Procedimentos determinísticos



Variograma $2\gamma(\mathbf{h})$

O variograma é uma ferramenta básica de suporte às técnicas de geoestatística, que permite representar quantitativamente a variação de um fenômeno regionalizado no espaço (Huijbregts, 1975).



$2\gamma(\mathbf{h})$

- mede o grau de dissimilaridade entre pares de observação separados pelo vetor distância \mathbf{h} ;
- é função do vetor distância \mathbf{h} ;
- depende da geometria de amostragem.

Variograma $2\gamma(\mathbf{h})$

Definição: esperança matemática (E) do quadrado da diferença entre os valores de pontos no espaço separados pelo vetor distância \mathbf{h} .

$$2\gamma(\mathbf{h}) = E\{[z(\mathbf{u}) - z(\mathbf{u} + \mathbf{h})]^2\}$$

Através de um conjunto amostral, $\{z(\mathbf{u}_1), z(\mathbf{u}_2), \dots, z(\mathbf{u}_N)\}$, o variograma pode ser estimado por:

$$\hat{2\gamma}(\mathbf{h}) = \frac{1}{N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{u}_i) - z(\mathbf{u}_i + \mathbf{h})]^2$$

em que:

$\hat{2\gamma}(\mathbf{h})$: é o estimador de variograma;

\mathbf{h} : é o vetor distância (modulo e direção) entre pares de observação;

$N(\mathbf{h})$: é o número de pares, $z(\mathbf{u}_i)$ e $z(\mathbf{u}_i + \mathbf{h})$, separados por \mathbf{h} ;

$z(\mathbf{u}_i)$ e $z(\mathbf{u}_i + \mathbf{h})$: são valores observados nas localizações \mathbf{u}_i e $\mathbf{u}_i + \mathbf{h}$.

Semivariograma $\gamma(\mathbf{h})$

Definição: metade da esperança matemática (E) do quadrado da diferença entre os valores de pontos no espaço separados pelo vetor distância \mathbf{h} .

$$\gamma(\mathbf{h}) = \frac{1}{2} E\{[z(\mathbf{u}) - z(\mathbf{u} + \mathbf{h})]^2\}$$

Através de um conjunto amostral, $\{z(\mathbf{u}_1), z(\mathbf{u}_2), \dots, z(\mathbf{u}_N)\}$, o semivariograma pode ser estimado por:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{u}_i) - z(\mathbf{u}_i + \mathbf{h})]^2$$

em que:

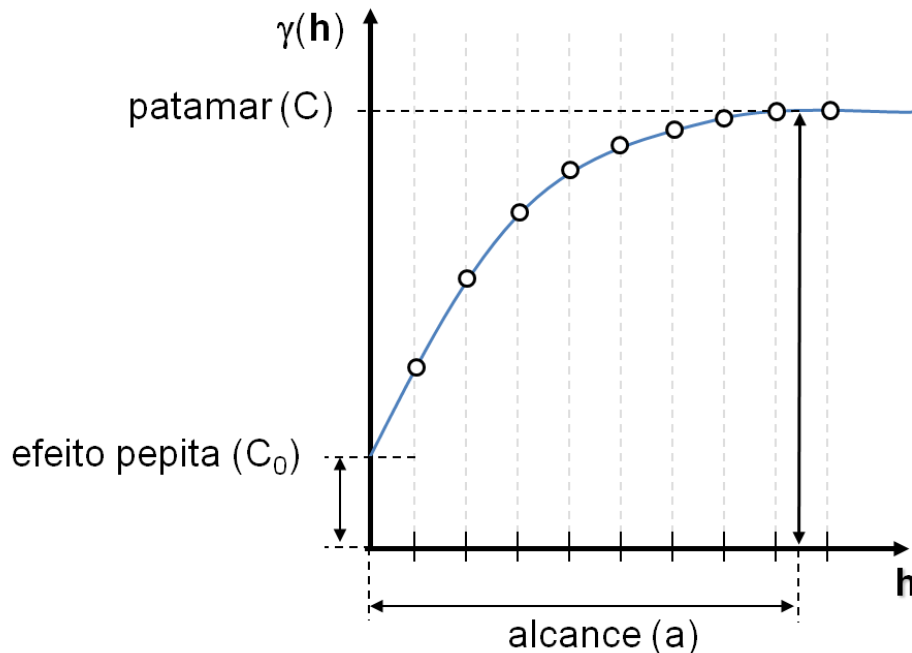
$\hat{\gamma}(\mathbf{h})$: é o estimador de semivariograma;

\mathbf{h} , $N(\mathbf{h})$, $z(\mathbf{u}_i)$ e $z(\mathbf{u}_i + \mathbf{h})$: conforme definidos anteriormente.

Semivariograma $\gamma(h)$

Definição: metade da esperança matemática (E) do quadrado da diferença entre os valores de pontos no espaço separados pelo vetor distância h .

A figura ilustra um semivariograma empírico (ou experimental) com características muito próximas do ideal.



- Alcance (a): distância dentro da qual as amostras apresentam-se correlacionadas espacialmente.

- Patamar (C): é o valor do semivariograma correspondente a seu alcance (a).

Deste ponto em diante, considera-se que não existe mais dependência espacial entre as amostras, porque a variância da diferença entre pares de amostras ($\text{Var}[Z(x) - Z(x+h)]$) torna-se aproximadamente constante.

- Efeito Pepita (C_0): idealmente, $\gamma(0)=0$.

Entretanto, na prática, à medida que h tende para zero, $\gamma(h)$ se aproxima de um valor positivo chamado Efeito Pepita (C_0), que revela a descontinuidade do semivariograma para distâncias menores do que a menor distância entre as amostras.

O efeito pepita é o valor da semivariância para a distância zero e representa a componente da variabilidade espacial que não pode ser relacionado com uma causa específica (variabilidade ao acaso)

Semivariograma $\gamma(h)$

A figura ilustra um semivariograma empírico (ou experimental) com características muito próximas do ideal.

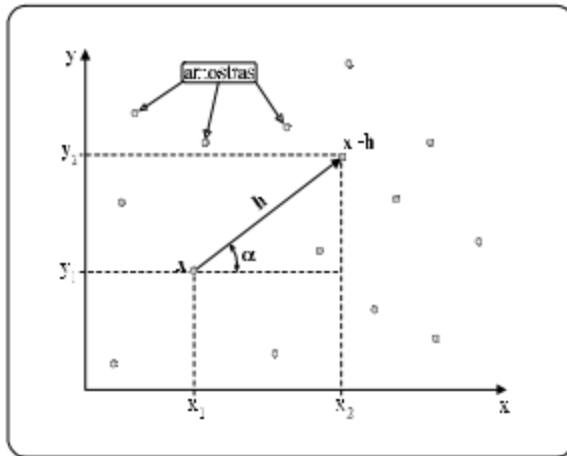


Fig. 2.2 - Amostragem em duas dimensões.

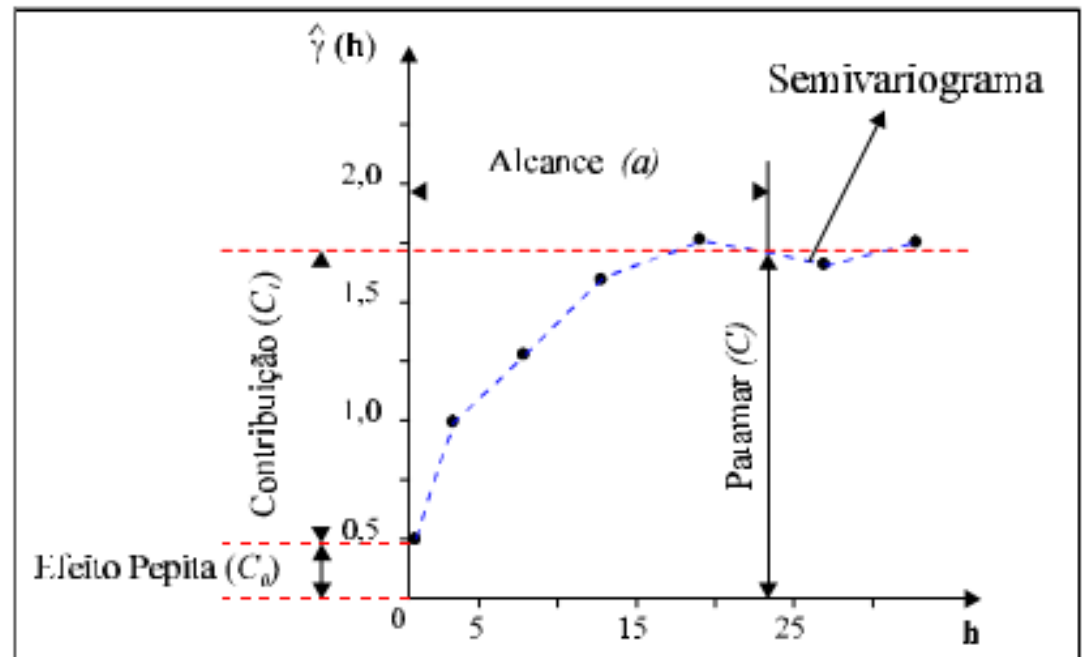
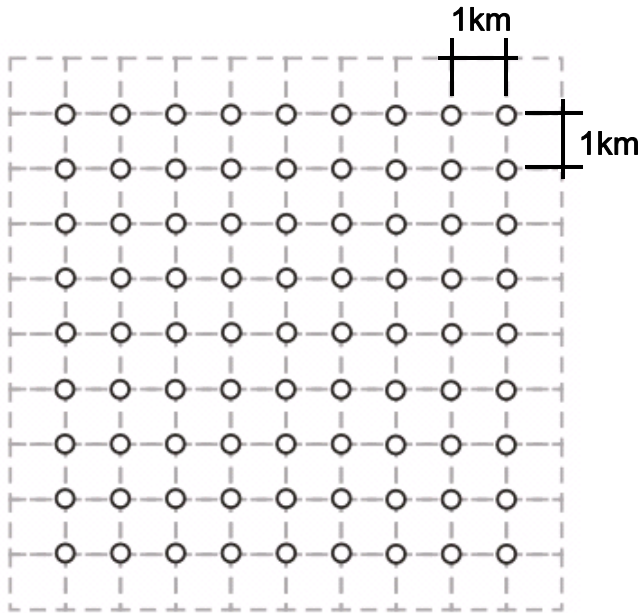


Fig. 2.3 - Exemplo de semivariograma.

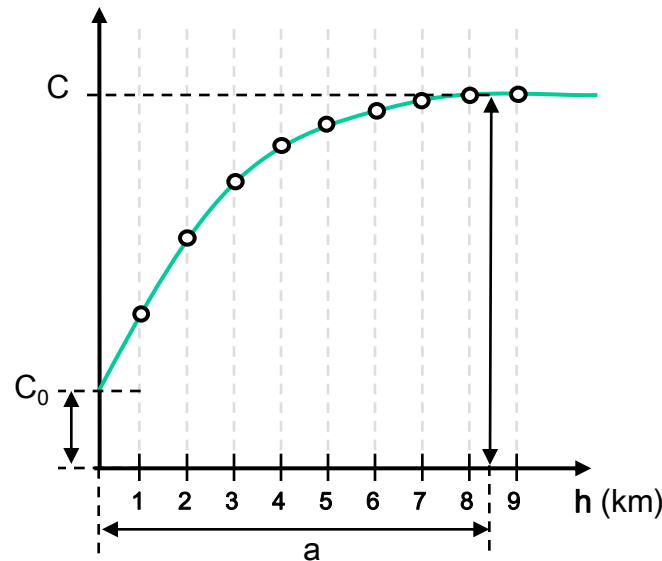
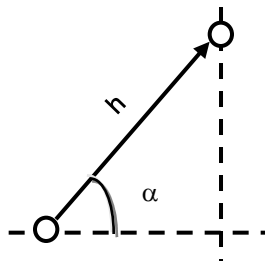
Semivariograma $\gamma(h)$

Cálculo do semivariograma a partir de amostras regularmente espaçadas.

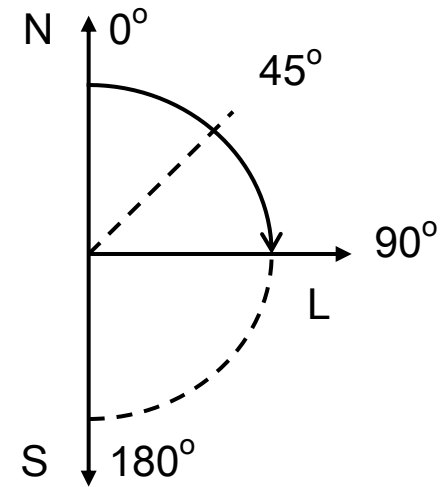


$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(\mathbf{u}_i) - z(\mathbf{u}_i + \mathbf{h})]^2$$

vetor distância h



direções de análise



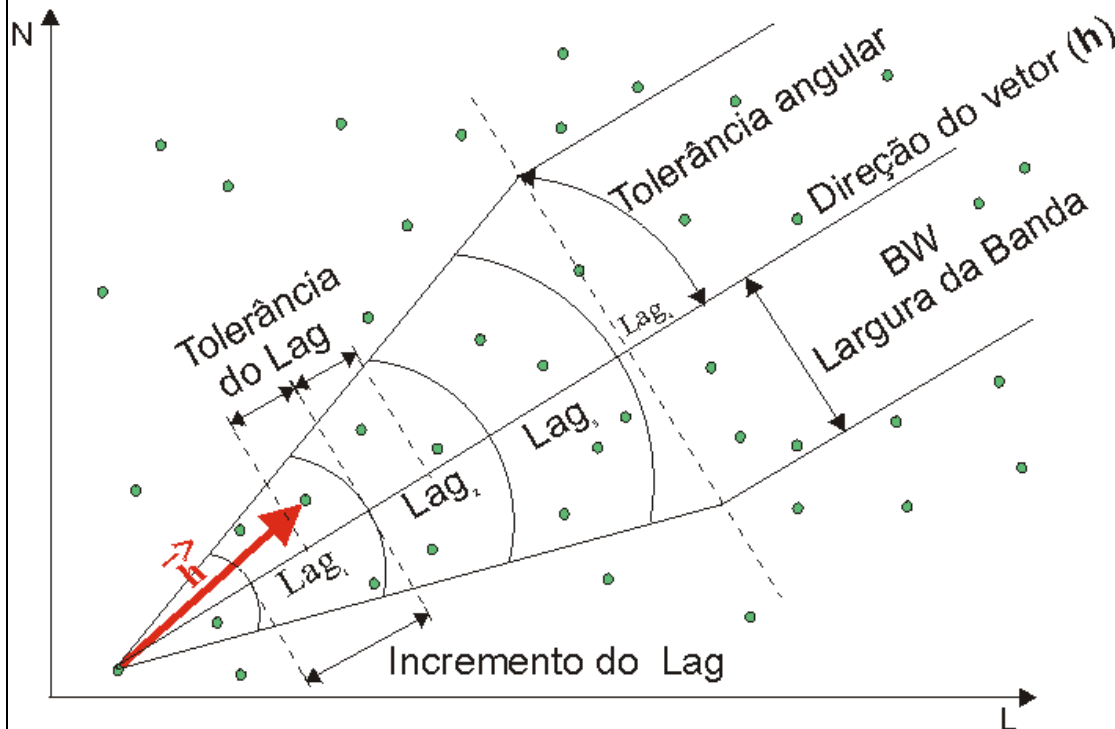
função simétrica

$$\gamma(h) = -\gamma(h)$$

Semivariograma $\gamma(\mathbf{h})$

Cálculo do semivariograma a partir de amostras irregularmente espaçadas.

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{u}_i) - z(\mathbf{u}_i + \mathbf{h})]^2$$



parâmetros adicionais

- tolerância do incremento (lag)
 - tolerância angular
 - largura de banda

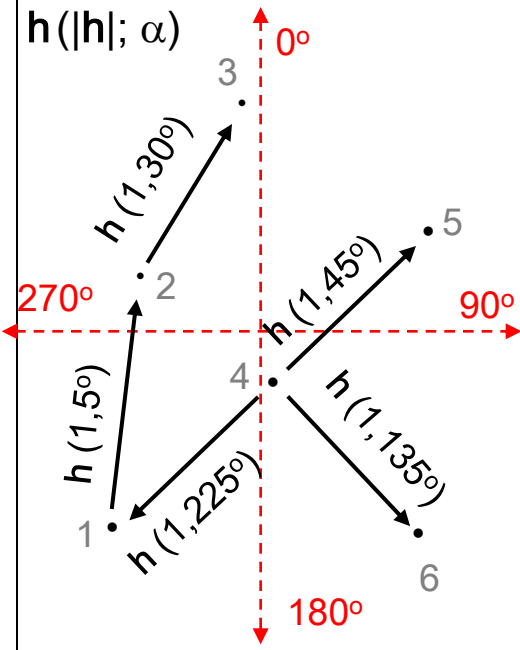
Semivariograma $\gamma(\mathbf{h})$

Cálculo do semivariograma a partir de amostras irregularmente espaçadas.

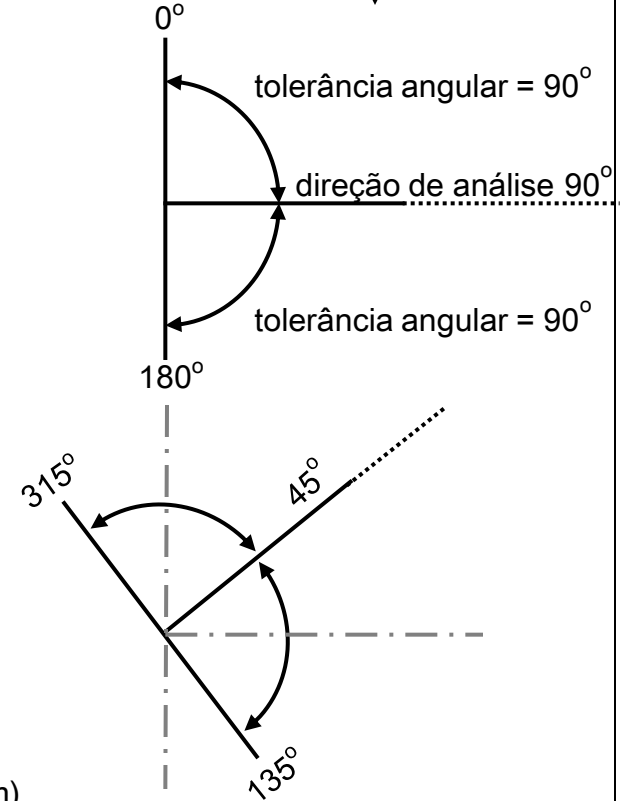
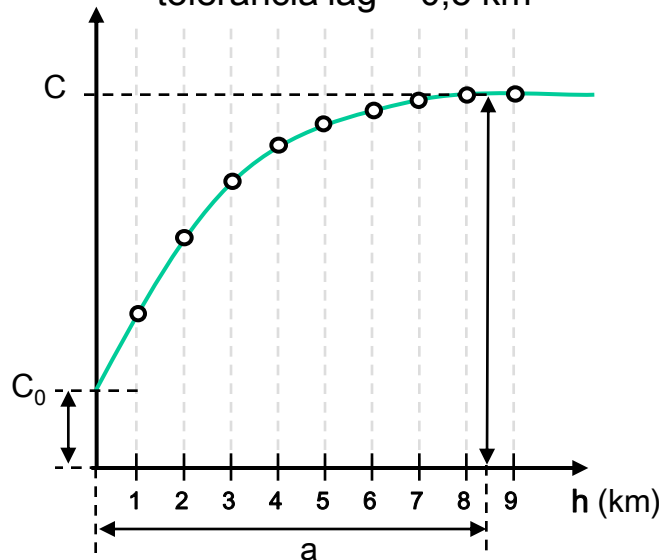
Semivariograma omnidirecional => tolerância angular = 90°

direção de análise (do vetor \mathbf{h}) não importa.

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{u}_i) - z(\mathbf{u}_i + \mathbf{h})]^2$$



Exemplo:
 incremento (lag) = 1 km
 tolerância lag = 0,5 km

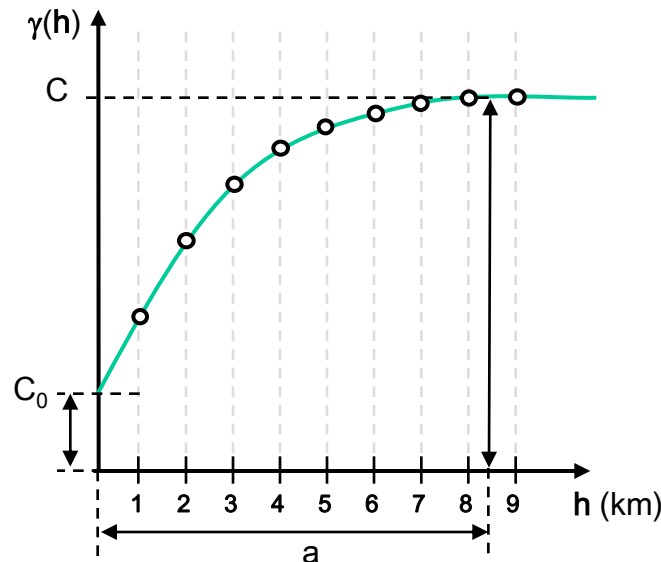
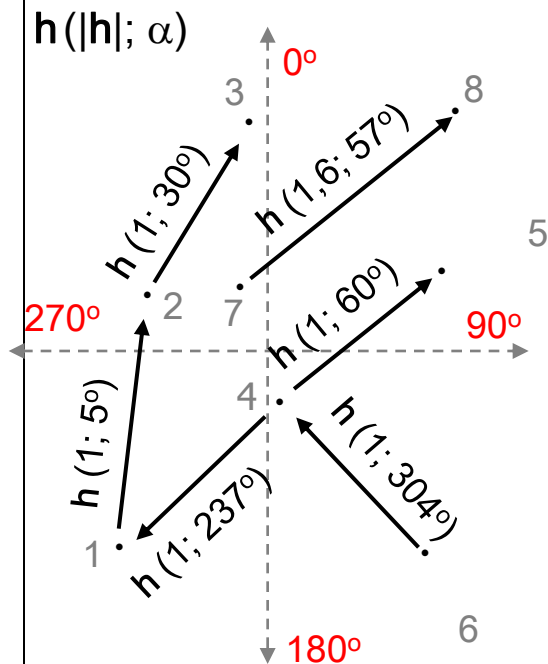
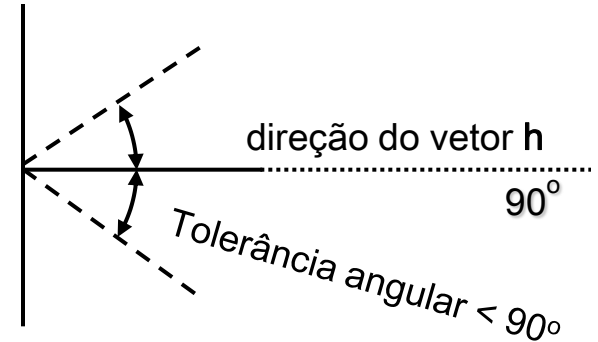


Semivariograma $\gamma(\mathbf{h})$

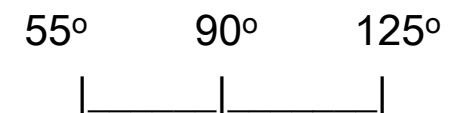
Cálculo do semivariograma a partir de amostras irregularmente espaçadas.

Semivariograma direcional => tolerância angular $< 90^\circ$

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(\mathbf{u}_i) - z(\mathbf{u}_i + \mathbf{h})]^2$$

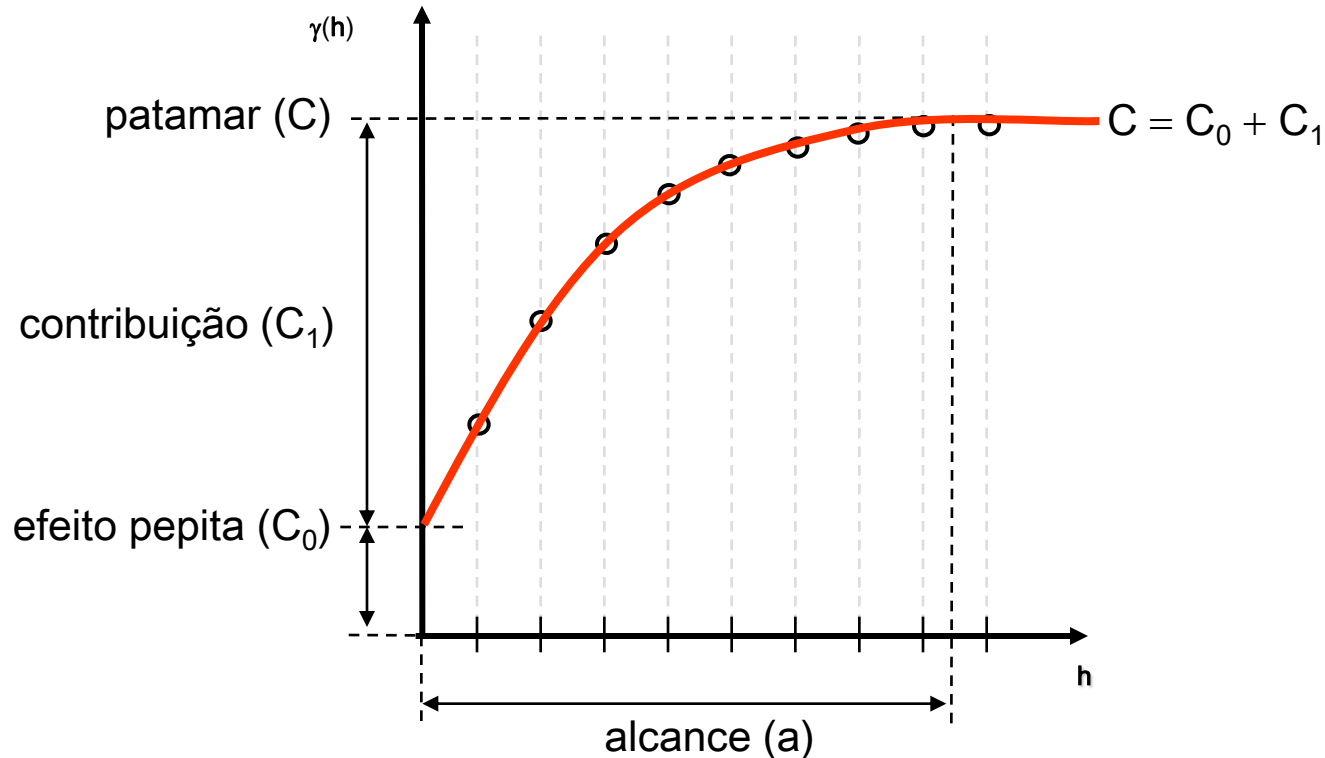


Exemplo:
 incremento (lag) = 1 km
 tolerância lag = 0,5 km
 direção de análise = 90°
 tolerância angular = 35°



Modelos teóricos de semivariograma

O gráfico do semivariograma empírico estimado por $\hat{\gamma}(h)$ é formado por uma série de valores, sobre os quais se objetiva ajustar uma função.



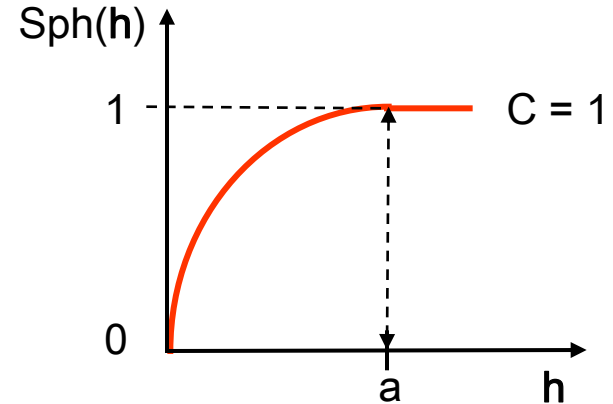
O modelo de ajuste deve representar o melhor possível o comportamento de $\gamma(h)$.

Modelos teóricos de semivariograma

Modelo de ajuste esférico

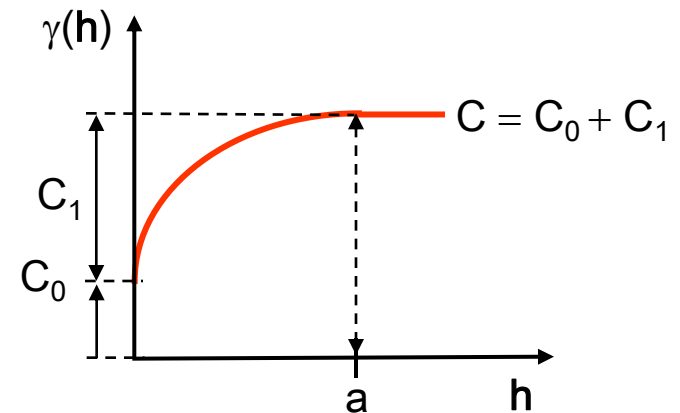
- Normalizado

$$Sph(\mathbf{h}) = \begin{cases} 0 & , |\mathbf{h}|=0 \\ 1,5\left(\frac{|\mathbf{h}|}{a}\right) - 0,5\left(\frac{|\mathbf{h}|}{a}\right)^3 & , 0 < |\mathbf{h}| \leq a \\ 1 & , |\mathbf{h}| > a \end{cases}$$



- Na prática: $C_0 > 0$ e $C_1 > 1$

$$\gamma(\mathbf{h}) = \begin{cases} 0, & C_0 \\ C_0 + C_1 \left[\frac{3}{2}\left(\frac{|\mathbf{h}|}{a}\right) - \frac{1}{2}\left(\frac{|\mathbf{h}|}{a}\right)^3 \right] = C_0 + C_1 [Sph(\mathbf{h})], & 0 < |\mathbf{h}| \leq a \\ C_0 + C_1, & |\mathbf{h}| > a \end{cases}$$

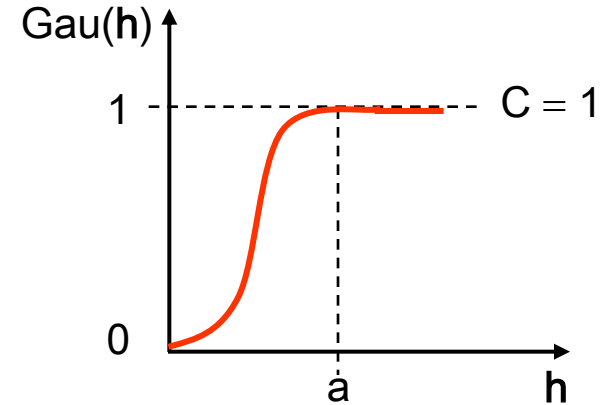


Modelos teóricos de semivariograma

Modelo de ajuste gaussiano

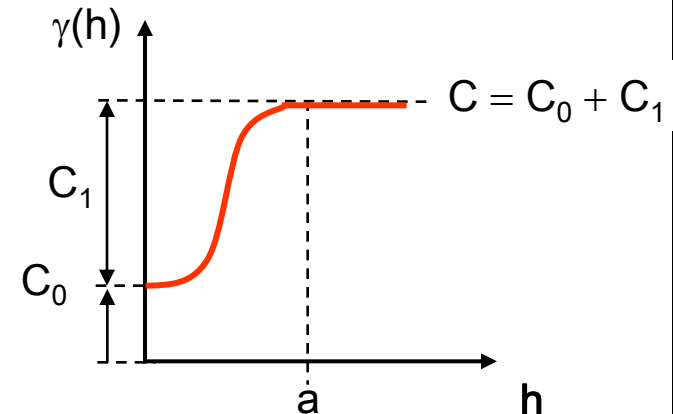
- Normalizado

$$\text{Gau}(\mathbf{h}) = \begin{cases} 0 & ,|\mathbf{h}|=0 \\ 1 - \exp\left(-\frac{|\mathbf{h}|^2}{a}\right) & ,|\mathbf{h}| \neq 0 \end{cases}$$



- Na prática: $C_0 > 0$ e $C_1 > 1$

$$\gamma(\mathbf{h}) = \begin{cases} 0, & \\ C_0 + C_1 \left[1 - \exp\left(-\frac{|\mathbf{h}|^2}{a}\right) \right] = C_0 + C_1[\text{Gau}(\mathbf{h})], & 0 < |\mathbf{h}| \leq a \\ C_0 + C_1, & |\mathbf{h}| > a \end{cases}$$

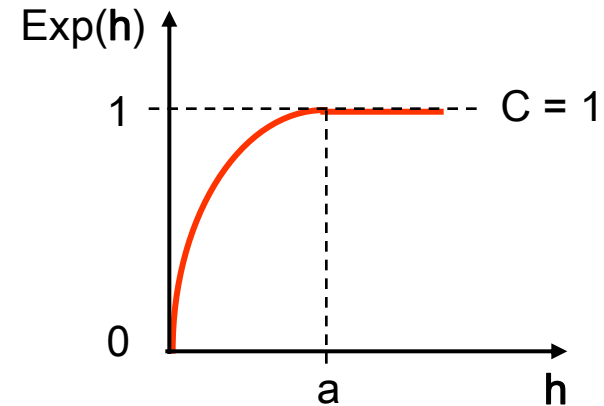


Modelos teóricos de semivariograma

Modelo de ajuste exponencial

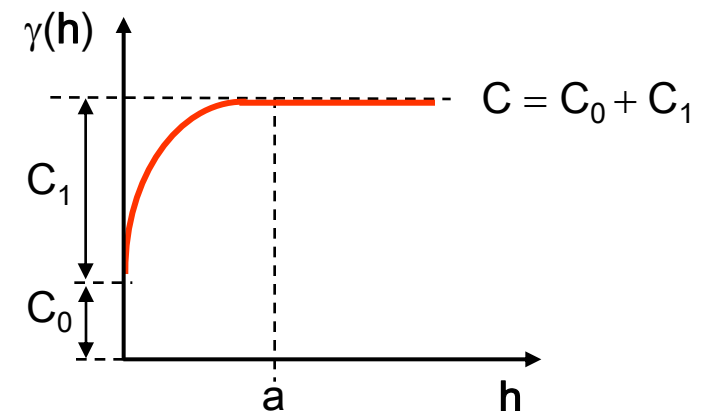
- Normalizado

$$\text{Exp}(\mathbf{h}) = \begin{cases} 0 & , |\mathbf{h}| = 0 \\ 1 - \exp\left(-\frac{|\mathbf{h}|}{a}\right) & , |\mathbf{h}| \neq 0 \end{cases}$$



- Na prática: $C_0 > 0$ e $C_1 > 1$

$$\gamma(\mathbf{h}) = \begin{cases} 0, & C_0 \\ C_0 + C_1 \left[1 - \exp\left(-\frac{|\mathbf{h}|}{a}\right) \right] = C_0 + C_1 [\text{Exp}(\mathbf{h})], & 0 < |\mathbf{h}| \leq a \\ C_0 + C_1, & |\mathbf{h}| > a \end{cases}$$

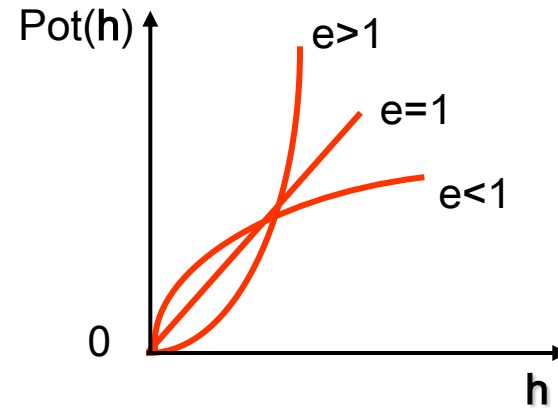


Modelos teóricos de semivariograma

Modelo de ajuste potência

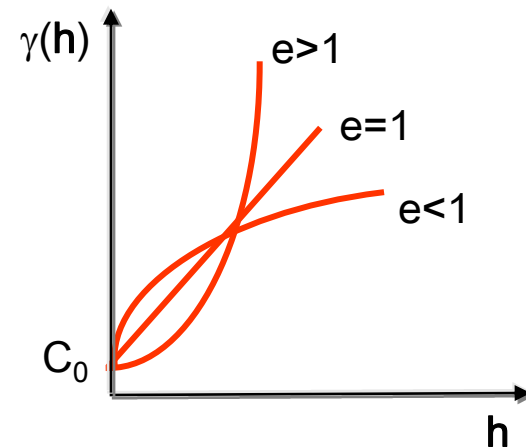
- Normalizado

$$\text{Pot}(\mathbf{h}) = \begin{cases} 0 & , |\mathbf{h}| = 0 \\ c \cdot |\mathbf{h}|^e & , |\mathbf{h}| \neq 0 \end{cases}$$



- Na prática: $C_0 > 0$ e $C_1 > 1$

$$\gamma(\mathbf{h}) = \begin{cases} 0 & , C_0 \\ C_0 + c \cdot |\mathbf{h}|^e = C_0 + \text{Pot}(|\mathbf{h}|) & , |\mathbf{h}| > 0 \end{cases}$$



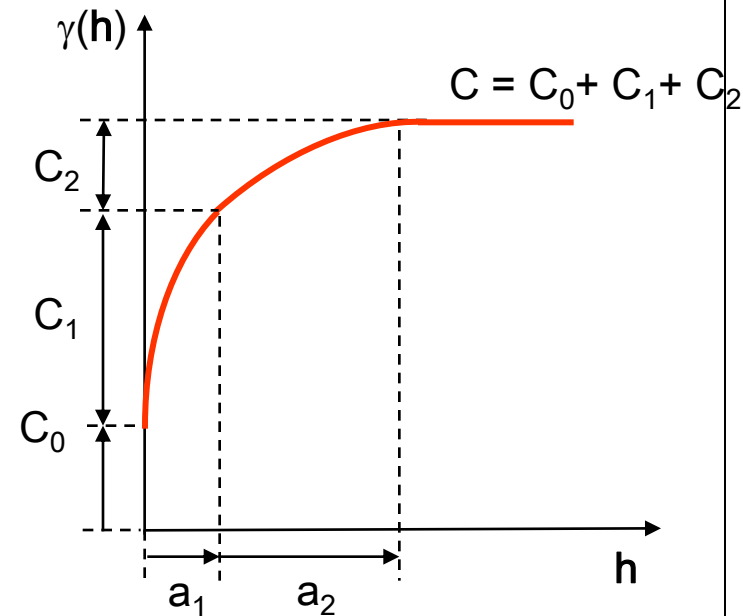
Modelos teóricos de semivariograma

Modelo de ajuste aninhados

Existem determinados fenômenos em que são necessários modelos mais complexos de semivariograma para explicar suas variações espaciais. Estes modelos são combinações de modelos simples, denominados *aninhados*.

Ex: Modelo aninhado duplo esférico

$$\gamma(\mathbf{h}) = \begin{cases} 0, & C_0 \\ C_0 + C_1 \left[\frac{3}{2} \left(\frac{|\mathbf{h}|}{a_1} \right) - \frac{1}{2} \left(\frac{|\mathbf{h}|}{a_1} \right)^3 \right] = \gamma_1(\mathbf{h}), & 0 < |\mathbf{h}| \leq a_1 \\ C_0 + C_2 \left[\frac{3}{2} \left(\frac{|\mathbf{h}|}{a_2} \right) - \frac{1}{2} \left(\frac{|\mathbf{h}|}{a_2} \right)^3 \right] = \gamma_2(\mathbf{h}), & a_1 < |\mathbf{h}| \leq a_2 \\ C_0 + C_1 + C_2, & |\mathbf{h}| > a_2 \end{cases}$$



Isotropia

Quando a variabilidade espacial de um fenômeno em estudo é a mesma em todas as direções, diz-se que o fenômeno é *isotrópico*.

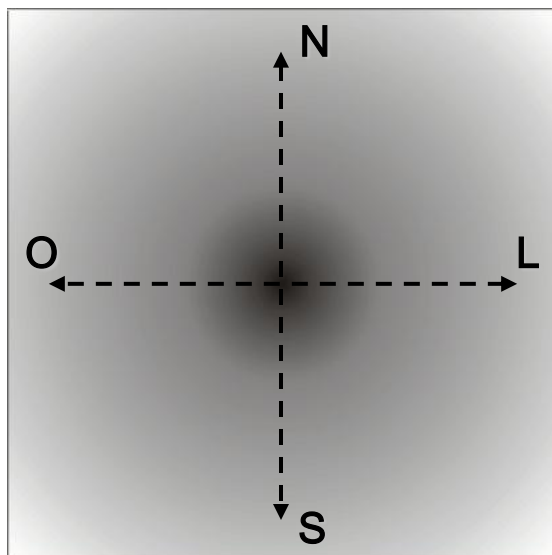
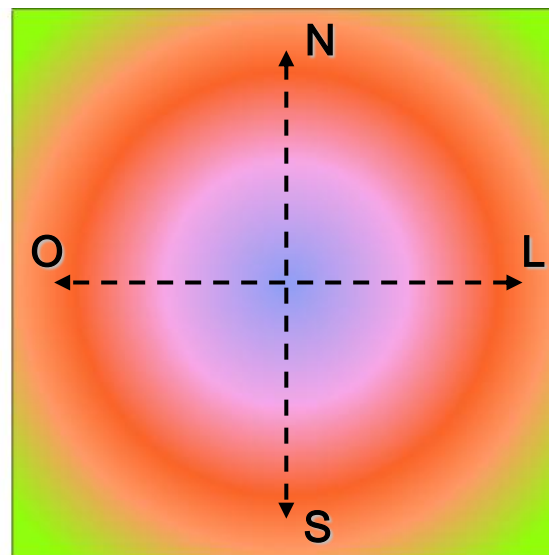


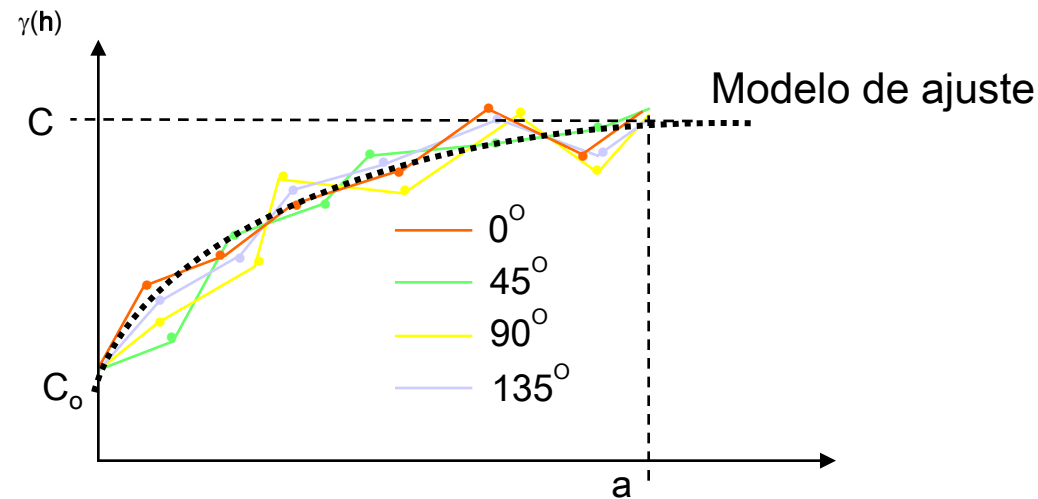
Imagem nível de cinza



Composição Colorida

Isotropia

Considere os semivariogramas ilustrados na figura abaixo



Esta é a representação de um caso simples e menos frequente, em que a distribuição espacial do fenômeno é denominada *isotrópica*.

Neste caso, um único modelo é suficiente para descrever a variabilidade espacial do fenômeno em estudo.

Anisotropia

Quando a variabilidade espacial de um fenômeno em estudo não é a mesma em todas as direções, diz-se que o fenômeno é *anisotrópico*.

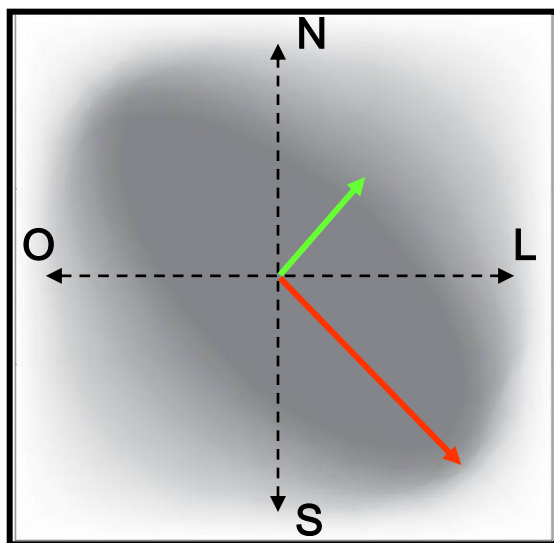
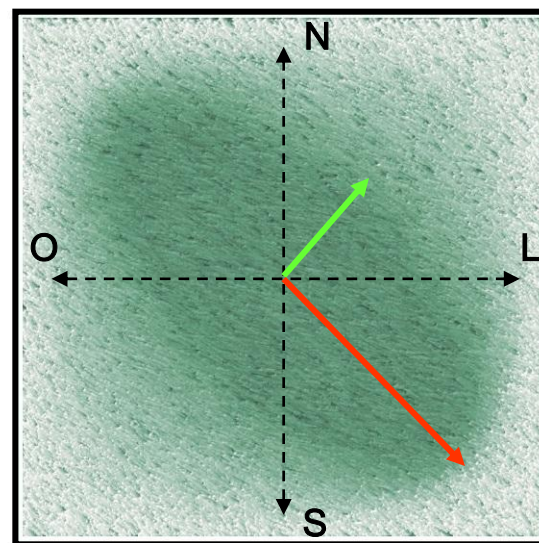


Imagem nível de cinza



Composição Colorida

direções de continuidade espacial

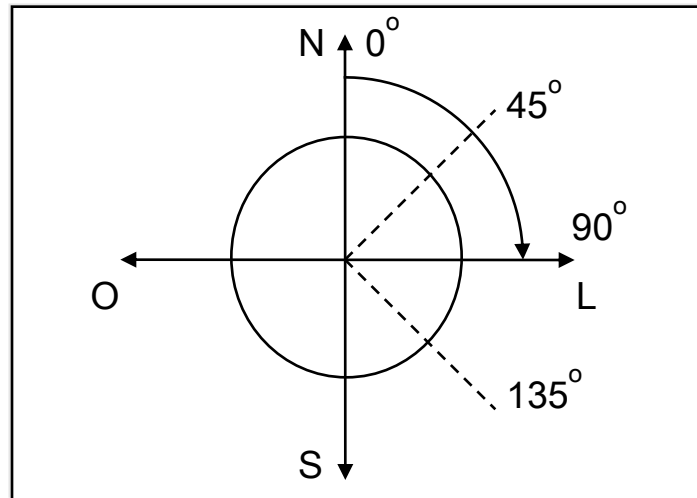
— maior

— menor

Anisotropia

A análise da *anisotropia* objetiva detectar as direções de maior e menor continuidade espacial do fenômeno investigado.

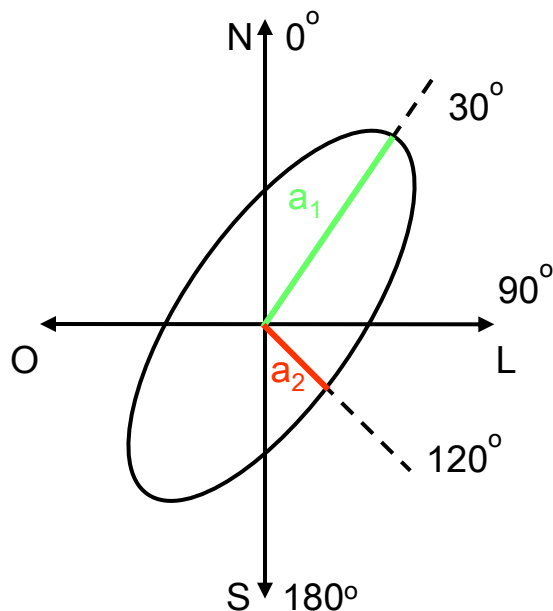
Convenções direcionais usadas na geoestatística



Uma forma de detectar a *anisotropia* é através da observação dos semivariogramas obtidos para diferentes direções.

Anisotropia

Um modo direto de visualizar e calcular os parâmetros (fator e ângulo) da anisotropia é através do esboço gráfico de uma elipse (ou *diagrama de rosa*).



Parâmetros da anisotropia

Fator de anisotropia (Fa)

$$Fa = a_2 / a_1$$

Ângulo de anisotropia (Aa)

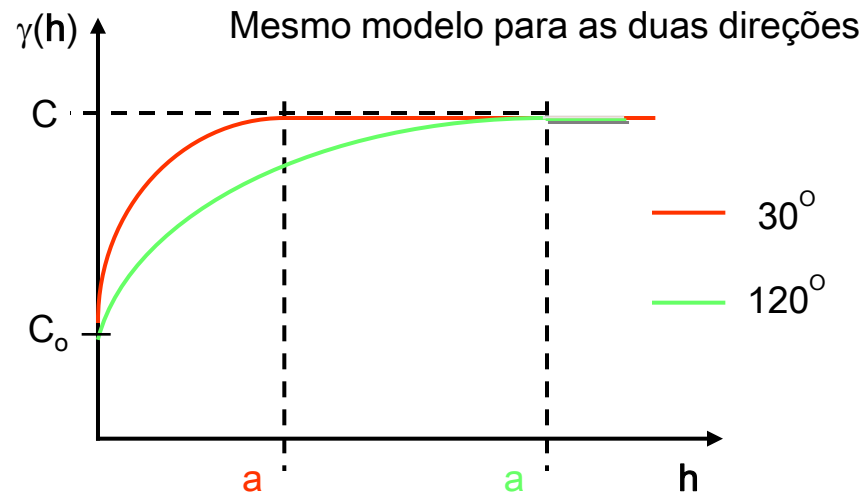
Aa = tomado da direção Norte para o eixo de maior continuidade. No exemplo = 30°.

Tipos de anisotropia: geométrica, zonal e combinada.

Anisotropia

Anisotropia geométrica

Neste caso, os semivariogramas apresentam o mesmo patamar (C) com diferentes alcances (a) para o mesmo modelo.

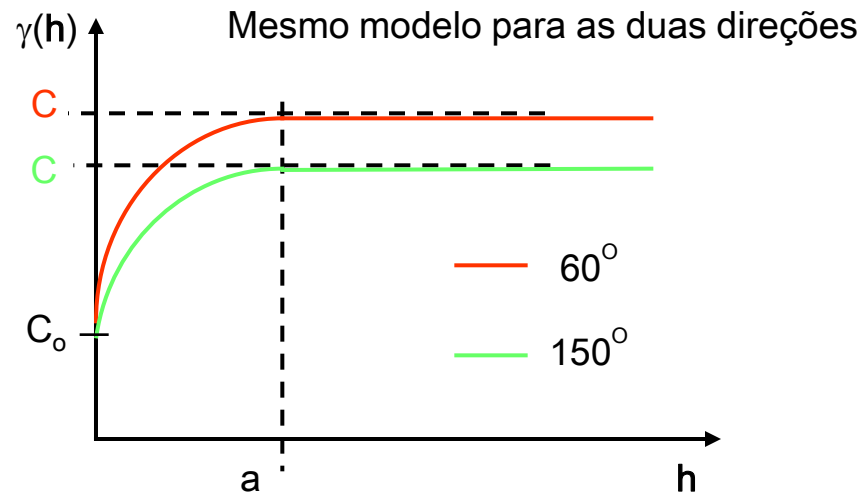


Anisotropia

Anisotropia zonal

Neste caso, os semivariogramas apresentam diferentes patamares (C) com mesmo alcance (a) para o mesmo modelo.

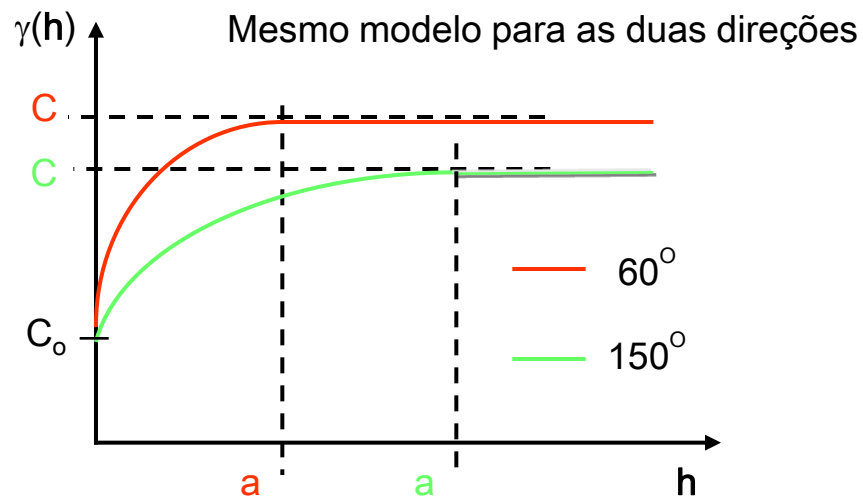
Como a isotropia, a **anisotropia zonal** é um caso menos frequente presente nos fenômenos naturais.



Anisotropia

Anisotropia combinada (geométrica + zonal)

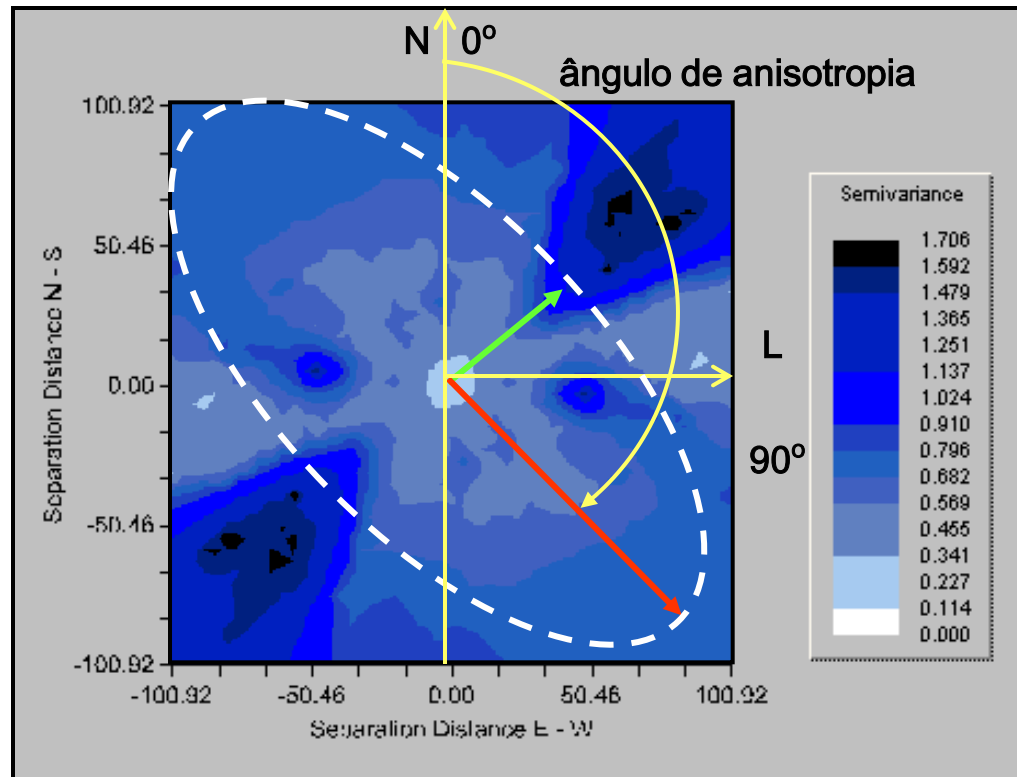
Neste caso, os semivariogramas apresentam diferentes patamares (C) e diferentes alcances (a) para o mesmo modelo. Pode apresentar também diferentes efeitos pepita.



Semivariograma de superfície

É um gráfico 2D que fornece uma visão geral da variabilidade espacial do fenômeno em estudo. Também conhecido como *Mapa de Semivariograma*.

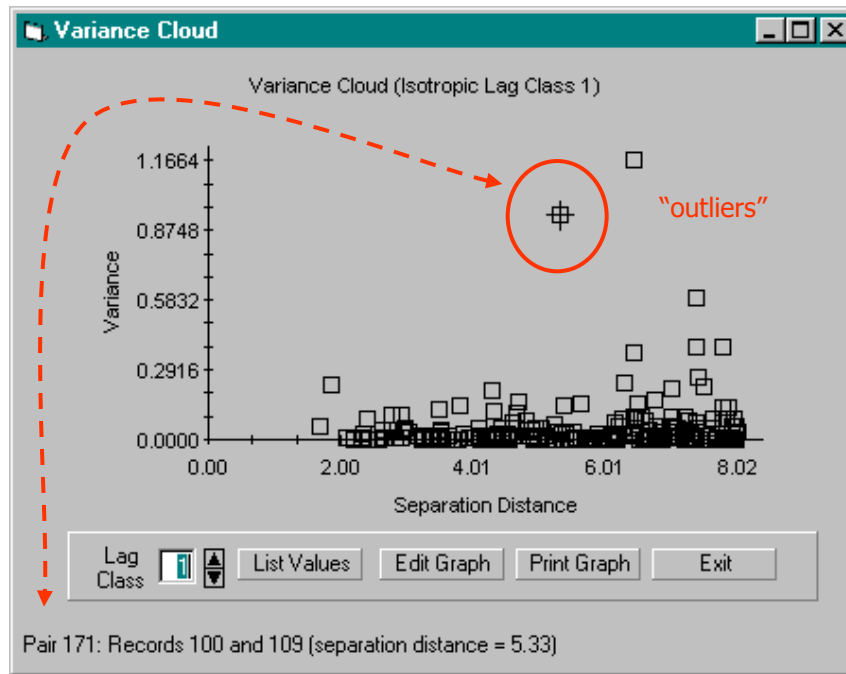
Utilizado para detectar os eixos de *Anisotropia* (direções de **maior** e **menor** continuidade espacial).



Semivariograma de nuvem

É um gráfico das semivariâncias de todos os pares de pontos tomados para um determinado lag (distância).

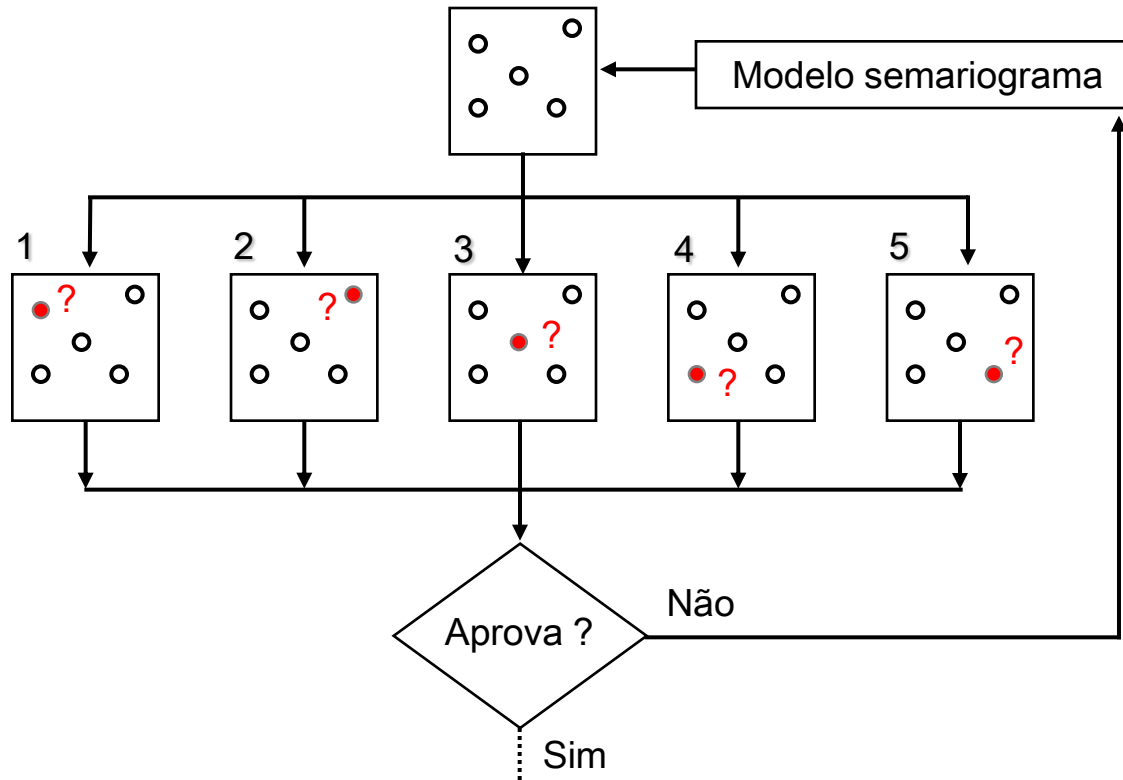
O variograma de nuvem é útil para detectar a presença de “outliers”.



Pair	Variance	Distance	Point 1 Record #	Point 2 Record #
1	0.0013	9.02	1	7
2	0.0002	15.43	1	10
3	0.0088	9.26	1	16
4	0.0030	10.18	1	18
5	0.0030	15.44	1	19
6	0.0230	12.74	1	25
7	0.0522	14.99	1	26
8	0.0343	12.52	1	27
9	0.0126	15.16	2	4
10	0.0084	13.70	2	9
11	0.0219	13.63	2	11
12	0.0042	8.78	2	19
13	0.0314	9.49	2	20
14	0.0010	12.30	2	21
15	0.0160	14.34	2	28
16	0.0126	12.16	3	4
17	0.0219	10.90	3	11
18	0.0042	11.32	3	19
19	0.0314	10.00	3	20

Validação cruzada

É um procedimento para verificar a adequação do modelo de ajuste ao semivariograma



Análises

- estatísticas do erro
- histograma do erro
- diagrama espacial do erro
- diagrama de valores observados versus estimados

Validação cruzada

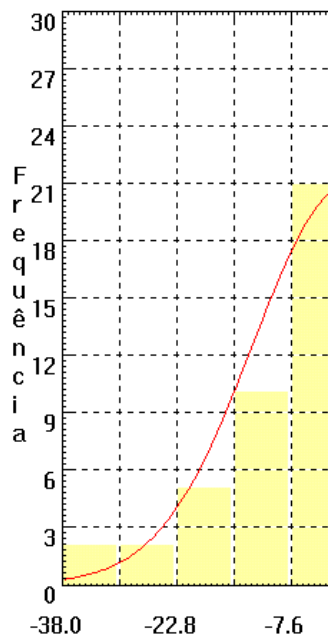
Análise de resultados



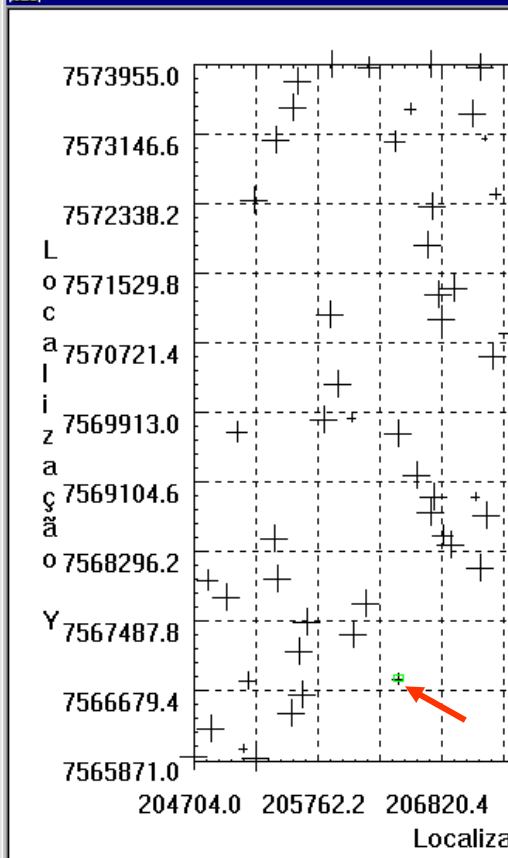
Histograma do Erro

Exibir Executar

Número de Classes: 10



Distribuição espacial do erro

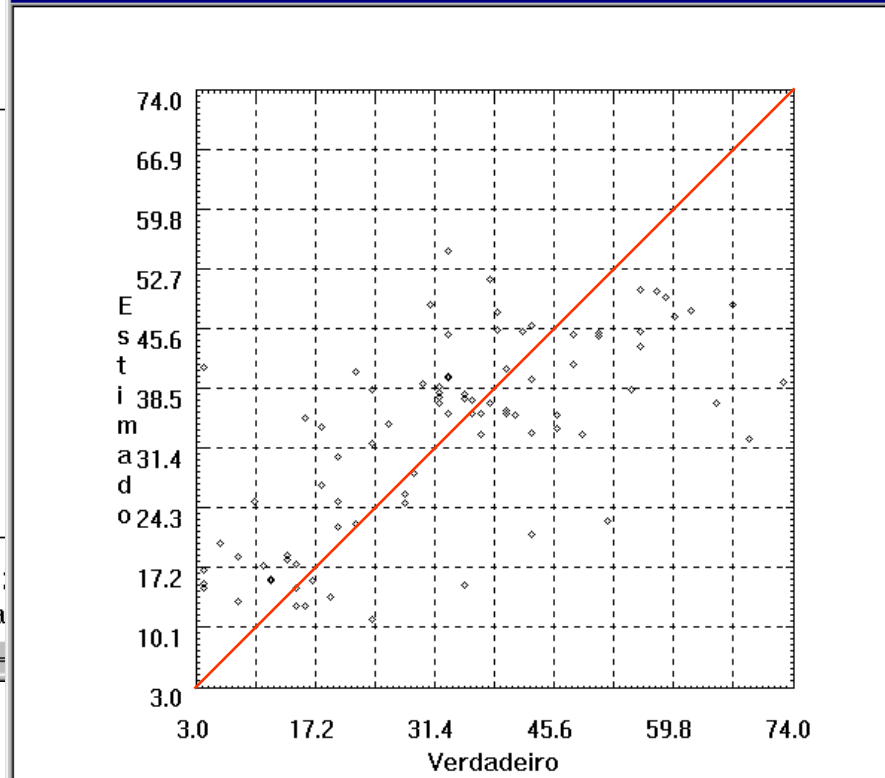


Relatório de Dados

CLASS DO ERRO

Class	Frequency
argila	85
	-0.252
	153.789
	12.401

Diagrama Verdadeiros X Estimados



Krigeagem

O termo **krigeagem** é derivado do nome *Daniel G. Krige*

A **krigeagem** é um estimador estocástico que depende da análise de correlação espacial baseada em semivariograma.

Áreas de Aplicações:

- mapeamento geológico (*Verly et al., 1984*)
- mapeamento solo (*Burgess e Webster, 1980*)
- mapeamento hidrológico (*Kitanidis et. al., 1983*)
- mapeamento atmosférico (*Lajaunie, 1984*)

A krigeagem engloba um conjunto de métodos de estimação que podem ser estacionários* ou não-estacionários:

- krigeagem Simples (*)
- krigeagem Ordinária (*)
- krigeagem Universal
- co-krigeagem
- krigeagem por indicação , outros



Krigeagem

A **krigeagem** é um estimador estocástico que depende da análise de correlação espacial baseada em semivariograma.

Na Krigeagem, o procedimento é semelhante ao de interpolação **por média móvel ponderada**, exceto que aqui os pesos são determinados a partir de uma **análise espacial, baseada no semivariograma experimental**.

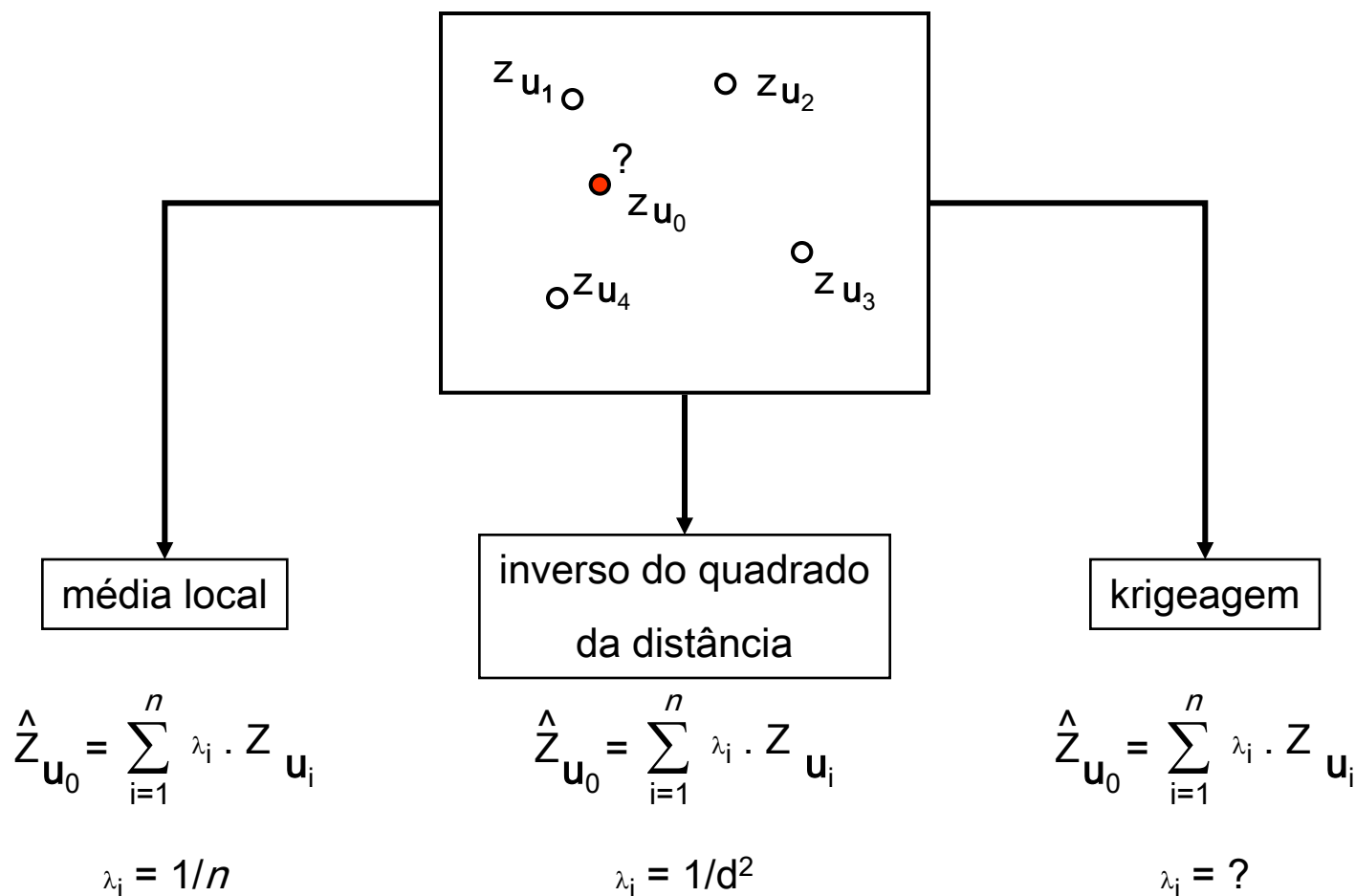
Além disso, a krigagem fornece, em média, estimativas não tendenciosas e com variância mínima

Estimativas não tendenciosas significam que, em média, a diferença entre valores estimados e observados para o mesmo ponto deve ser nula;

Variância mínima significa que estes estimadores possuem a menor variância dentre todos os estimadores não tendenciosos.

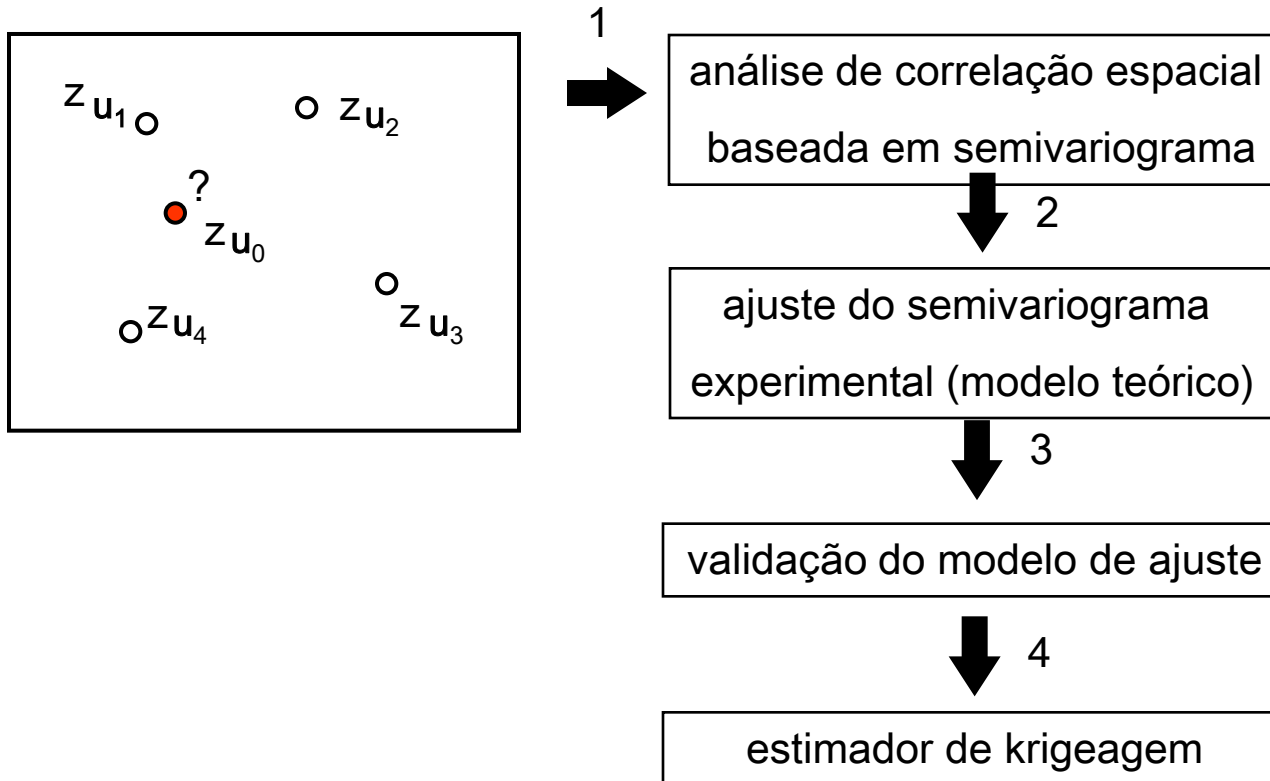
Krigeagem

Envolve uma combinação linear de n valores em pontos vizinhos.



Krigeagem

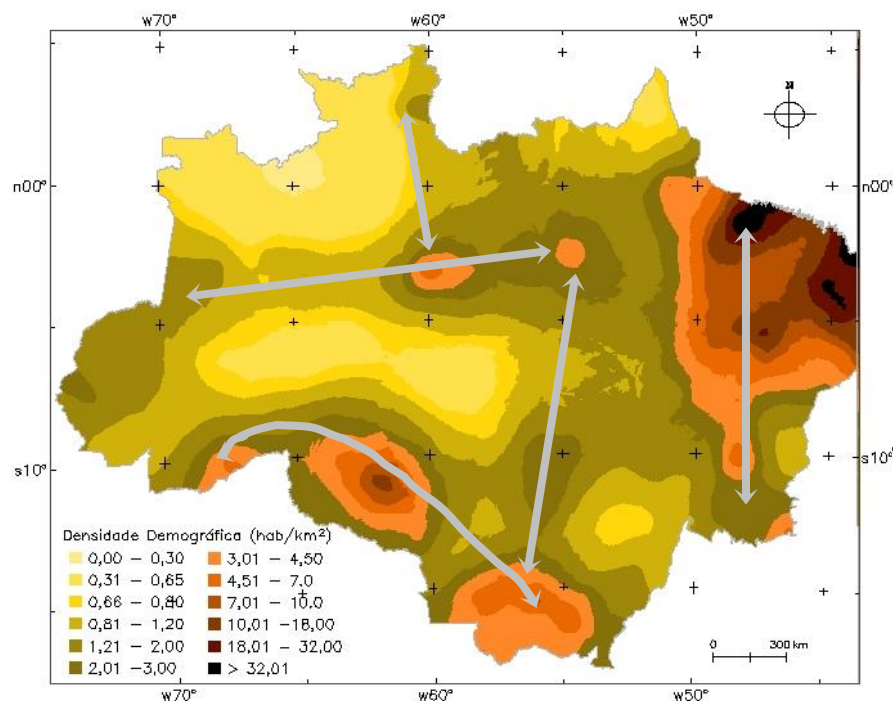
Os pesos são calculados considerando a estrutura de correlação espacial imposta pelo semivariograma



População na AmzL

Superfícies de Densidade - Krigeagem

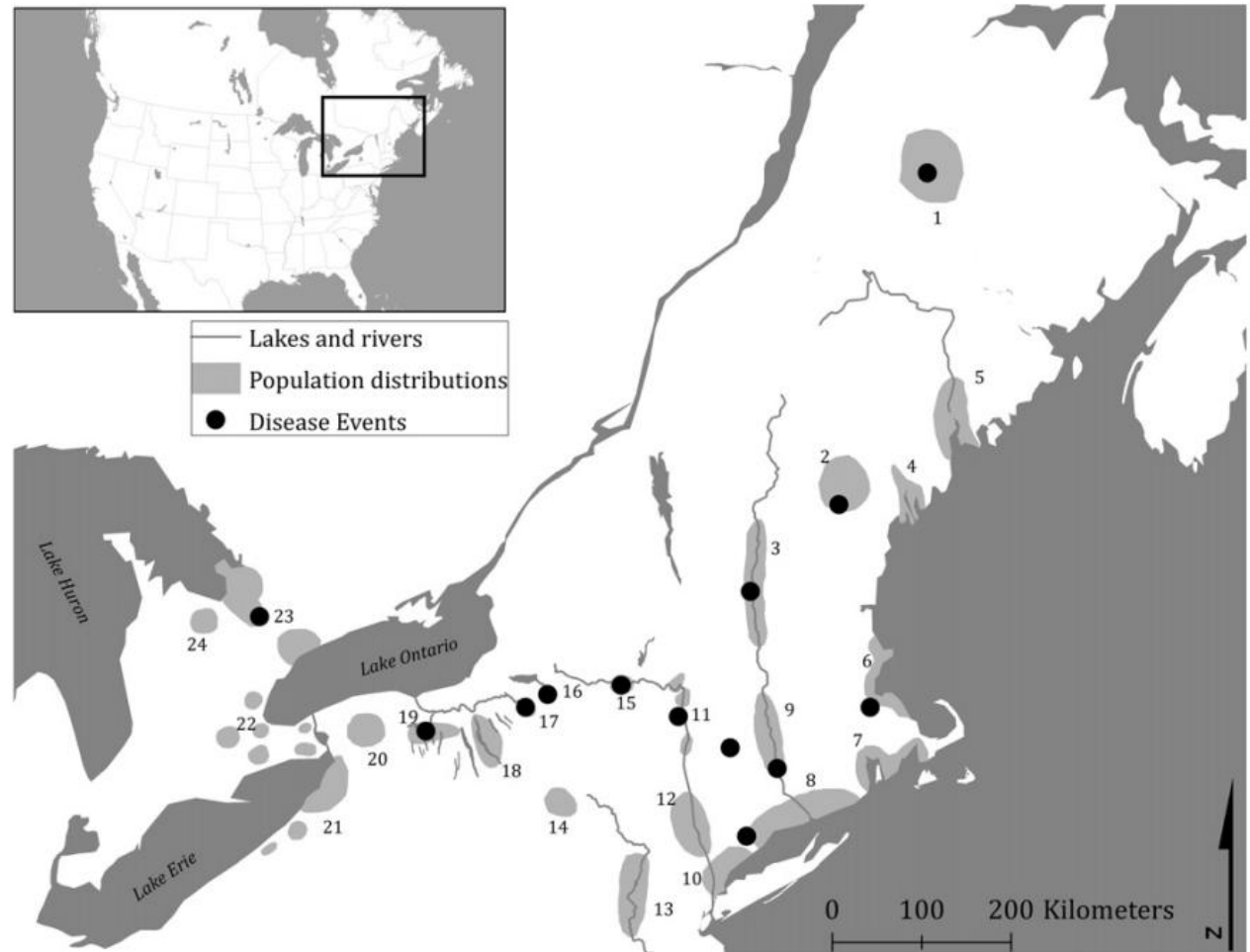
- Densidade de População Total – sedes de município
- Das superfícies-> Mapas temáticos para análise
- Mapas de tendências, imprecisos e generalizados
- Evolução das superfícies - > tendências demográficas



Densidade de População - 2000

Using spatial analysis to estimate depopulation for Native American populations in northeastern North America, AD 1616–1645

Eric E. Jones^{a,*}, Sharon N. DeWitte^b



Spatial analysis to estimate depopulation for Native American populations



Objetivos:

- mapear o conhecimento sobre a distribuição da população e do despovoamento (% pop perdida) das sociedades nativas americanas do sec 17 (NE US).
 - Padrões espaciais de despovoamento (*depopulation*);
- Usar interpolação espacial para determinar a *depopulation* dos grupos nativos para os quais não se tem info de população
 - Preencher lacunas de informação para entender o processo de *depopulation*
- Final goal: discutir fatores culturais, biológicos e históricos da *depopulation*

Spatial analysis to estimate depopulation for Native American populations

Métodos:

- Dados de distribuição espacial de publicações existentes e surveys arqueológicos
- Mapeamento prévio dos autores



Table 1

The 17th-century Native American cultures of the Northeast used in this study and the source of their geographic information. Those in italics previously had no population information but have depopulation percentage estimates from our results below.

Culture	Sources on geographic distributions
Maliseet-Passamaquoddy	Snow (1980, p. 341)
Mohegan-Pequot	Dincauze (1990)
<i>Kennebec</i>	Piotrowski (2002)
Mahican	Bradley (2007)
Eastern Abenaki	Snow (1980, p. 62)
Western Abenaki	Snow (1980, p. 342)
Massachusetts	Dincauze (1990) and Bragdon (1996)
Pocumtuc	Dincauze (1990) and Piotrowski (2002)
Quiripi-Unquachog	Dincauze (1990)
<i>Narragansett-Pokanoket</i>	Dincauze (1990)
Mohawk	Snow (1995a)
Oneida	Jones (2010a, 2010c)
Onondaga	Jones (2010a, 2010c)
Cayuga	Jones (2010c)
Seneca	Wray et al. (1987) and Jones (2010c)
Wendat	Engelbrecht (2003) and Warrick (2008)
<i>Neutral</i>	Snow (1994) and Engelbrecht (2003)
<i>Wenro</i>	Snow (1994) and Engelbrecht (2003)
<i>Erie</i>	Snow (1994) and Engelbrecht (2003)
<i>Munsee</i>	Dincauze (1990)
<i>Susquehannock</i>	Snow (1994) and Engelbrecht (2003)
Minisink	Kraft (1991)

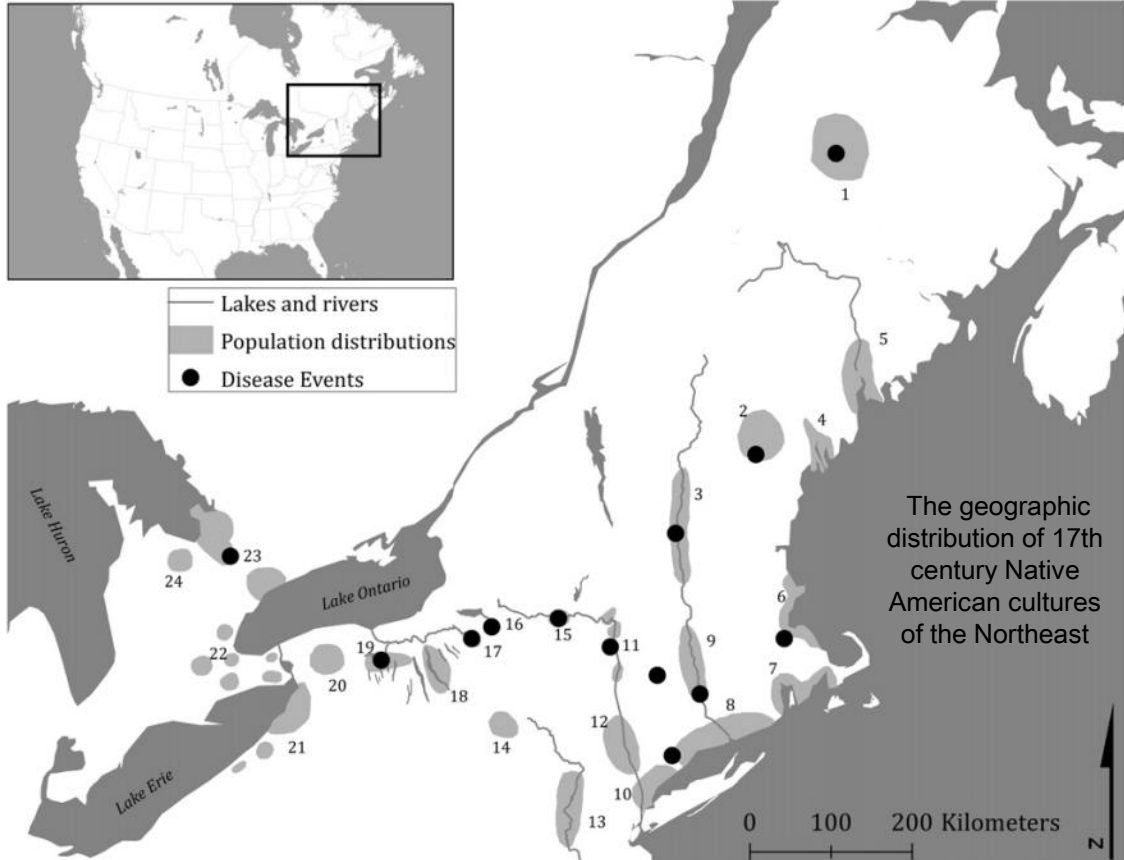
- Digitalizaram informações dos mapas para um GIS
- Usando localizações dos sites ocupados no período, criaram *shapefiles* das extensões geográficas que os continham
- Mapa de distribuição conhecida (1616-1645) baseada nos registros (figura)

Spatial analysis to estimate depopulation

Métodos:

Estimativas de população pré e pós contato também de trabalhos anteriores; (diferentes acurácias)

Tabela 2 - mais precisos



The geographic distribution of 17th century Native American cultures of the Northeast

Table 2
Population data for the 17th-century Native American populations used in this study, including the year of the depopulation event being studied, beginning and ending populations for the depopulation event during the indicated year, percentage of total population that died, and sources.

Population	Date	Initial population	Resulting population	Percent lost	References
Maliseet-Passamaquoddy	1633	7600	2500	67	Snow and Lanphear (1988, p. 24)
Mohegan-Pequot	1633	16,000	3000	81	Bradford (1952, p. 260)
Mahican	1633	6400	500	92	Snow and Lanphear (1988, p. 24)
Eastern Abenaki	1633	13,800	3000	78	Snow and Lanphear (1988, p. 24)
Western Abenaki	1633	12,000	250	98	Snow and Lanphear (1988, p. 24)
Massachusetts	1633	44,000	6400	86	Snow and Lanphear (1988, p. 24)
Pocumtuc	1633	18,400	920	95	Snow and Lanphear (1988, p. 24)
Quiripi-Unquachog	1633	29,900	1500	95	Snow and Lanphear (1988, p. 24)
Mohawk	1633	8000	1750	78	Snow (1995a, 1995b), Jones (2010a), and Snow and Lanphear (1988, p. 24)
Seneca	1634	4850	3150	35	Jones (2010b)
Oneida	1635	1500	400	73	Jones (2010a)
Wendat	1639	20,000	10,000	50	Heidenreich (1976), Trigger (1976, pp. 588–589), and Warrick (2008)
Onondaga	1645	2700	1150	57	Jones (2010a)

Spatial analysis to estimate depopulation for Native American populations

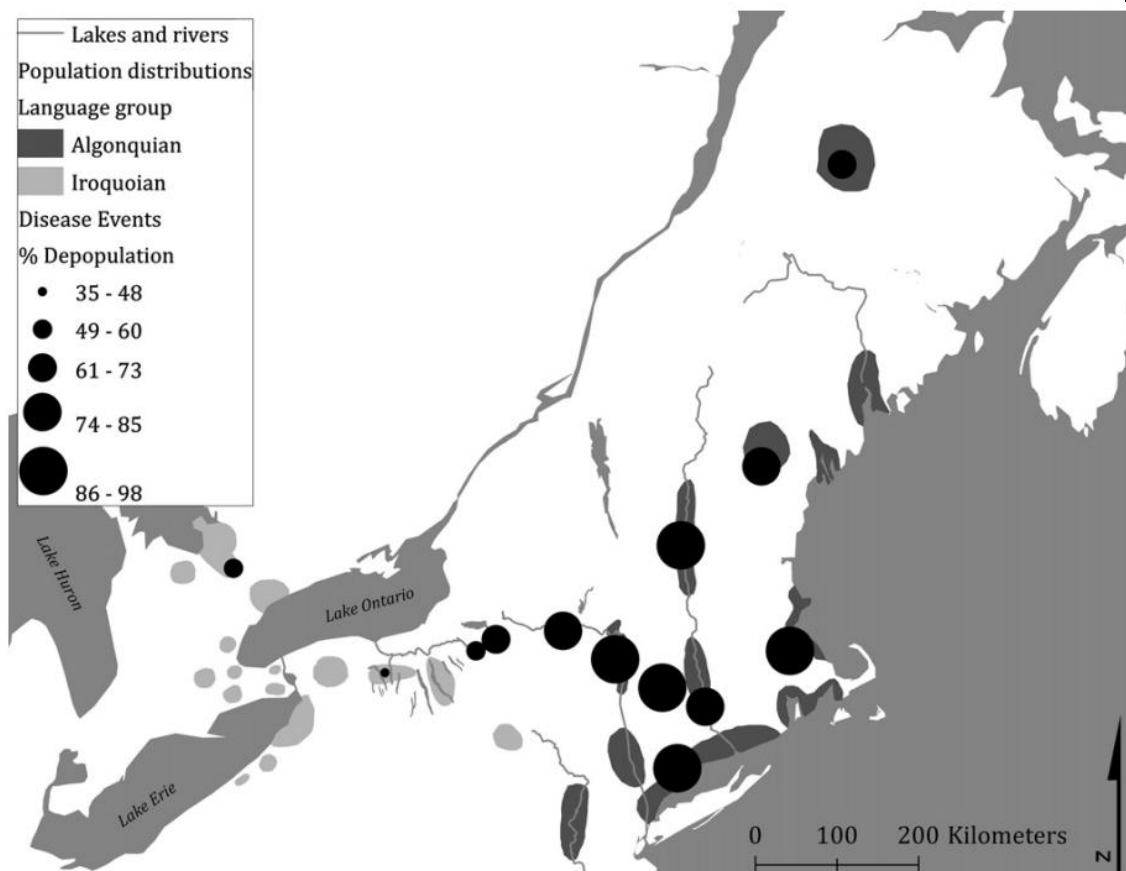
Métodos:

- Despovoamento (*depopulation*): % perda entre estimativas pré e pós contato (valores da literatura)
- Georeferenciou *depopulation* associando ao ponto central da distribuição inicial mapeada (PTOs)
 - krigagem da *depopulation* para estimar as taxas de *depopulation* para os grupos sem dados (e.g. doenças dependem da proximidade espacial...)

“It is important to note that kriging can only produce a surface that is an approximation of the spatial variation in the dataset”

- krigagem ordinária, modelos esférico

The percentage of population lost in each disease event overlaid on the distribution of Iroquoian-speaking and Algonquian-speaking population distributions.



Spatial analysis to estimate depopulation for Native American populations

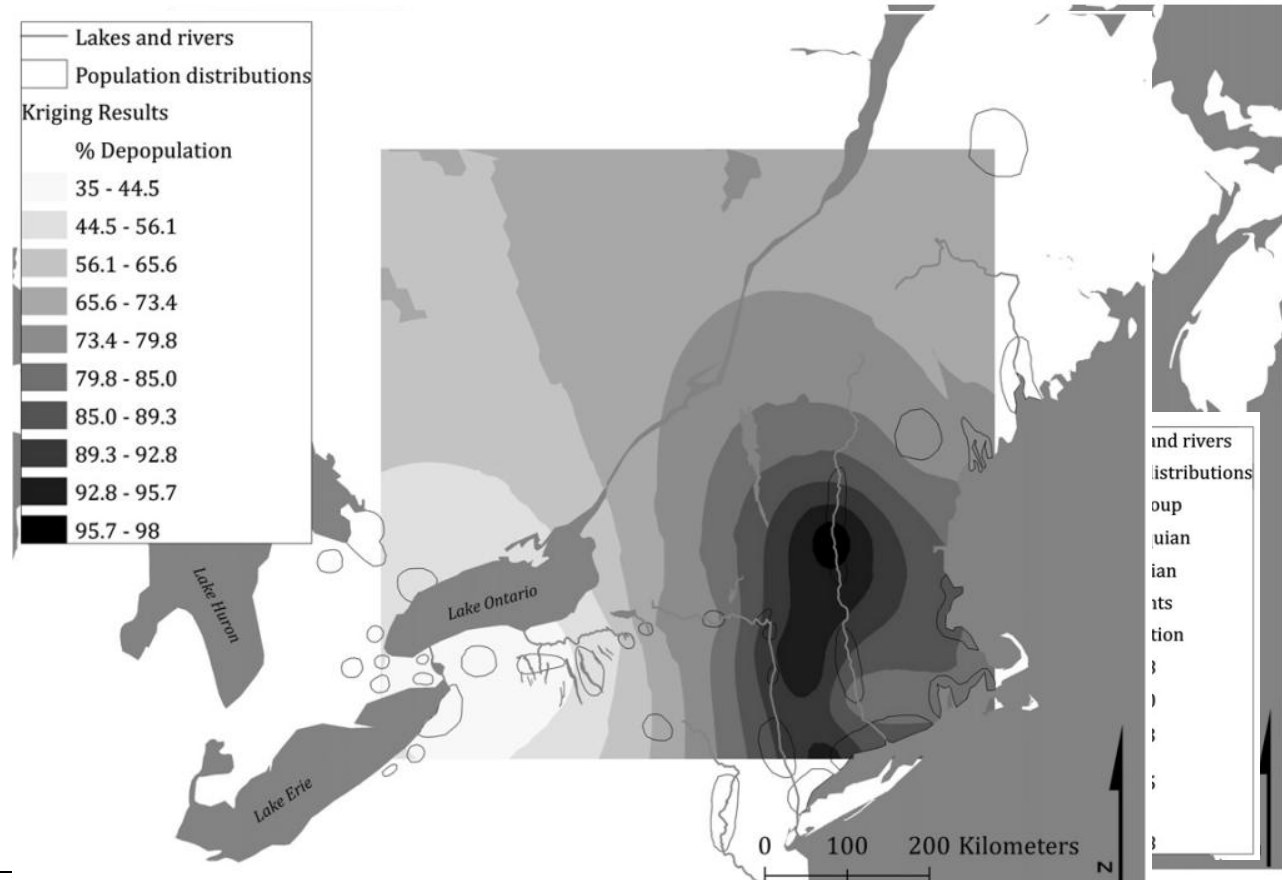
Sobreposição do mapa de distribuição geográfica das culturas nativas do sec 17 com a superfície de % de *depopulation*

Buscando entender os padrões para os grupos para os quais não se tinha informação

Discussão das causas:

- Doenças múltiplas e impacto na competência imunológica
- Tipo e Local de assentamento
- Variabilidade genética,
- práticas culturais.
- Outras /especulações...

Results of kriging based on existing depopulation data.



Referências



Básicas – Krigeagem:

- Camargo, E.C.G.,; Fucks, S. ; Camara, G. Análise Espacial de Superfícies <http://www.dpi.inpe.br/gilberto/livro/analise/cap3-superficies.pdf>
- Tobler, W.R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association*, 74, 519-530.
- Tobler, W.R., Deichmann, U., Gottsegen, J.& Maloy, K. (1995). *The Global Demography Project*. Santa Barbara, CA: National Center for Geographic Information and Analysis.
- Wolf-Dieter Rase, 2001, “Volume-preserving interpolation of a smooth surface from polygon-related data”, *J. Geograph. Syst.*, 3: 199-213

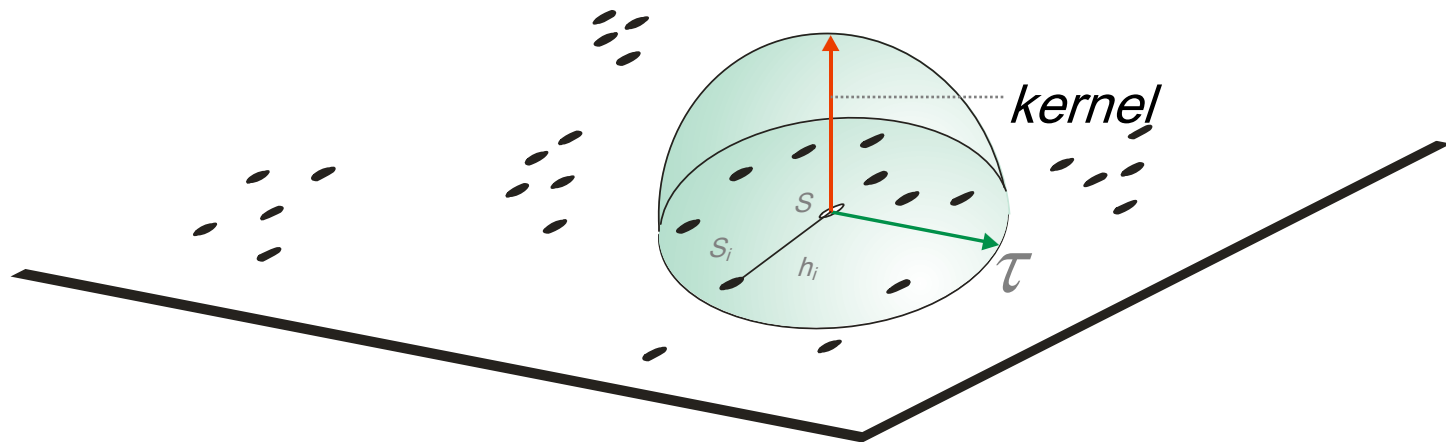
Caracterização de Distribuição de Pontos

- Processo de análise de pontos pode ser descritos em termos de :
 - **Efeitos de Primeira Ordem**
 - considerados globais ou de grande escala.
 - correspondem a variações no valor médio do processo.
 - Neste caso estamos interessados na intensidade do processo (Nº Eventos / Unidade de Área).
Estimador de Intensidade (“Kernel Estimation”)
 - **Efeitos de Segunda Ordem**
 - denominados locais ou de pequena escala.
 - representam a dependência espacial no processo
Vizinho mais Próximo e Função K
- A maior parte das técnicas de análise de distribuição de pontos supõe um comportamento isotrópico.

Estimador de Intensidade (“Kernel Estimation”)



$$\hat{\lambda}_{\tau}(s) = \sum_{i=1}^n \frac{3}{\pi\tau^2} I \left(1 - \frac{h_i^2}{\tau^2} \right)^2$$

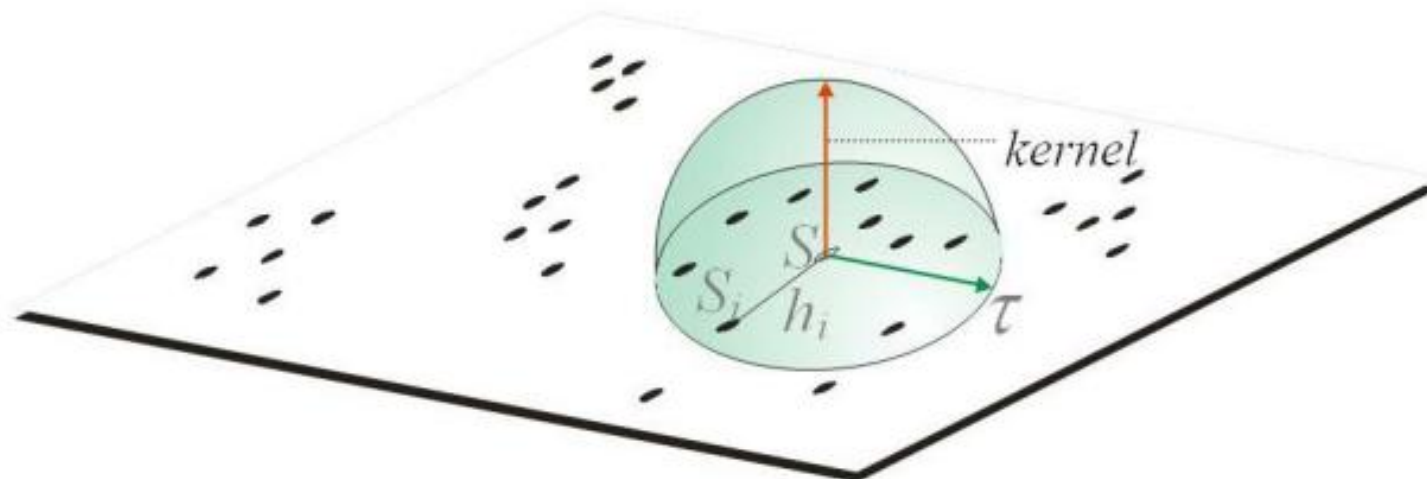


Estimador de Intensidade (“Kernel Estimation”)

- Kernel estimator é um estimador cujos parâmetros básicos são:
 - (a) um raio de influência que define a vizinhança do ponto a ser interpolado;
 - (b) uma função de estimação com propriedades “convenientes” de suavização do fenômeno.

Para toda posição z_i cujo valor queremos estimar:

- o estimador de intensidade será computado a partir dos valores das amostras $\{z_1, \dots, z_n\}$ contidos num raio de tamanho τ , e da distância euclidiana d_{ij} entre a i -ésima posição e a j -ésima amostra



Estimador de Intensidade (“Kernel Estimation”)

- Ex Funções

$$k(x, y, \tau) = \frac{1}{2\pi\tau} \exp\left(-\frac{d_{ij}^2}{2\tau^2}\right)$$

- Kernel gaussiano

- Kernel de Quarta ordem

$$k(x, y, \tau) = \frac{3}{\pi\tau^2} \left(1 - \frac{d_{ij}^2}{\tau^2}\right)^2$$

- Resultado com raios de busca de 500 e 1500m

- No primeiro mapa predominam os efeitos locais, pelo uso de um raio de busca reduzido; o segundo mapa evidencia melhor a distribuição do fenômeno, pelo uso de um raio mais apropriado aos dados.

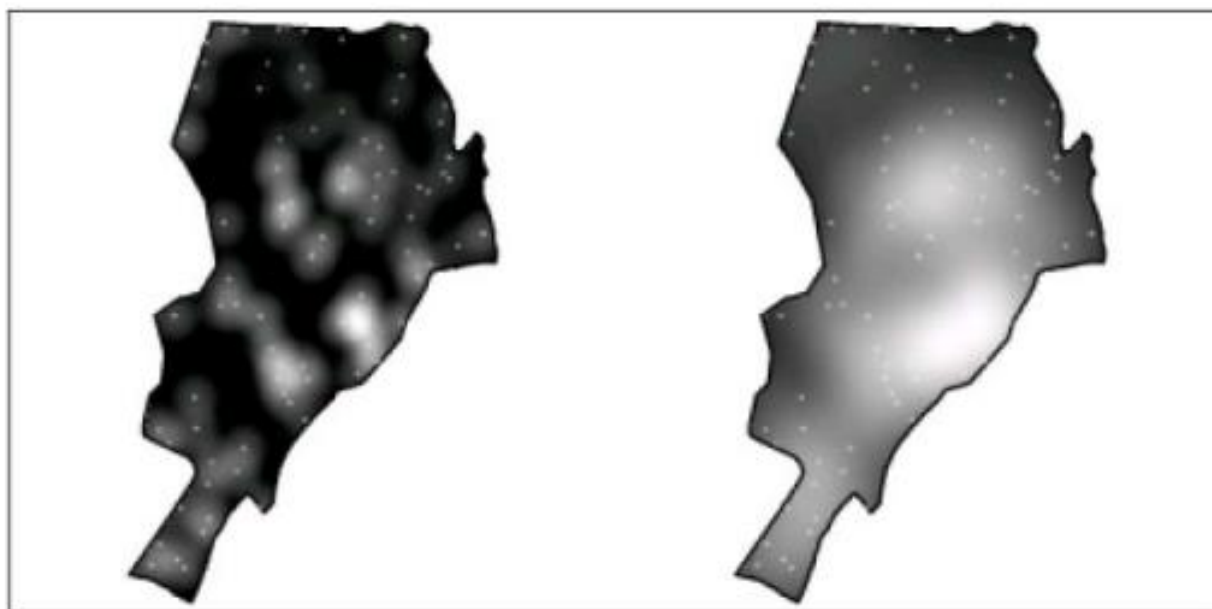


Figura 3-5- Superfícies de teor de argila interpoladas por kernel de quarta ordem. À esquerda, raio de busca de 500m; à direita, raio de busca de 1500m.

Estimador de Intensidade (“Kernel Estimation”)

- Segundo (Bailey e Gatrell, 1995):

$$\hat{\lambda}_{\tau}(s) = \sum_{i=1}^n \frac{1}{\tau^2} I\left(\frac{(s - s_i)}{\tau}\right)$$

Onde:

- A função $I(\cdot)$ -> FDP, escolhida de forma adequada para construir uma superfície contínua sobre os dados.
- O parâmetro τ denominado “largura de faixa”, controla o amaciamento da superfície gerada.
- s representa uma localização qualquer na área de estudo e s_i são as localizações dos eventos observados.
- n representa o número de eventos.

Estimador de Intensidade (“Kernel Estimation”)

- Uma função muito utilizada para $I()$ é:

$$I(h) = \frac{3}{\pi} (1 - h^2)^2$$

onde:

- h representa a distância entre a localização em que desejamos calcular a função e os eventos observados.
- Assim o estimador de intensidade pode ser expresso como:

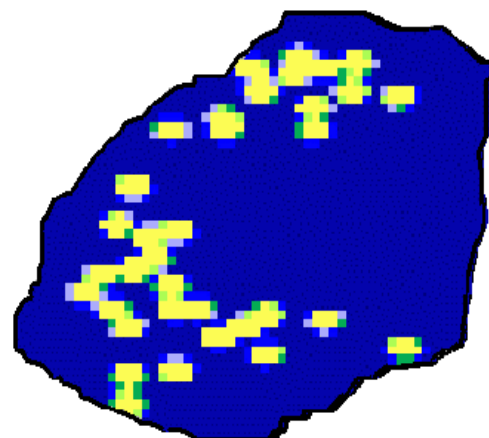
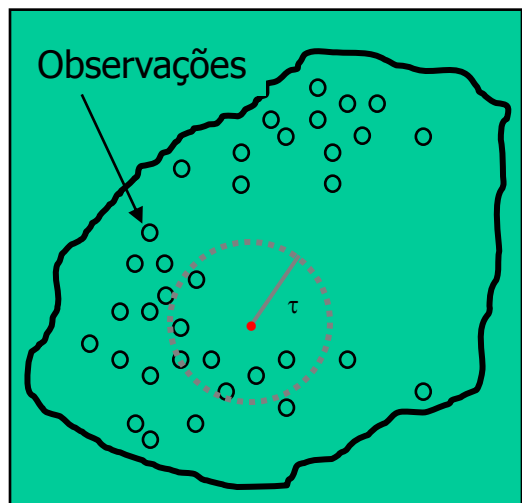
$$\hat{\lambda}_\tau(s) = \sum_{i=1}^n \frac{3}{\pi\tau^2} I\left(1 - \frac{h_i^2}{\tau^2}\right)^2$$

onde:

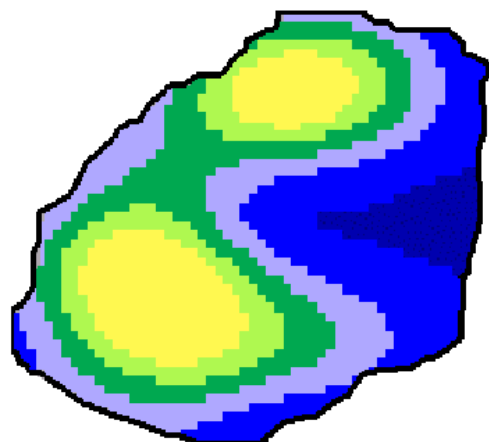
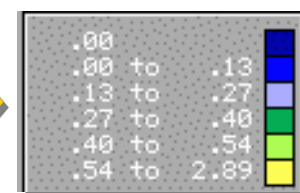
- h_i é a distância entre o ponto a calcular S e o valor observado S_i .

Estimador de Intensidade (“Kernel Estimation”)

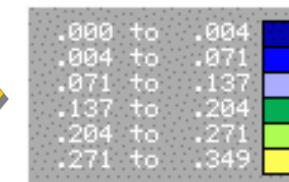
Efeito da Largura de Faixa (τ)



Banda estreita



Banda larga



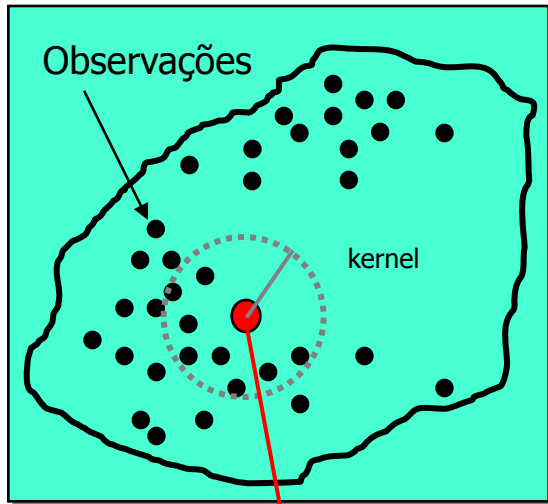
Estimador de Intensidade (“Kernel Estimation”)



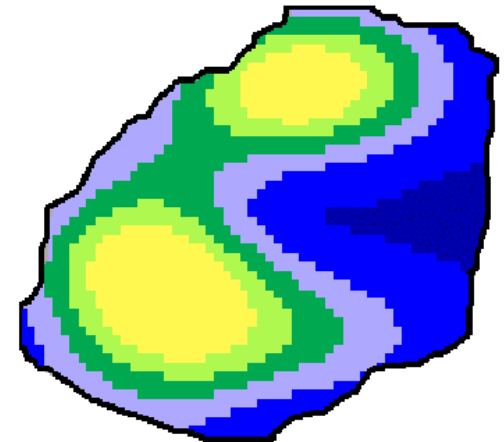
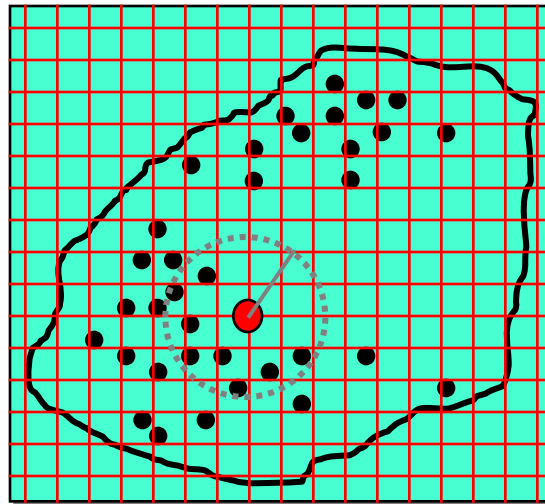
Plano de Informação (PI)

Grade de Intensidade

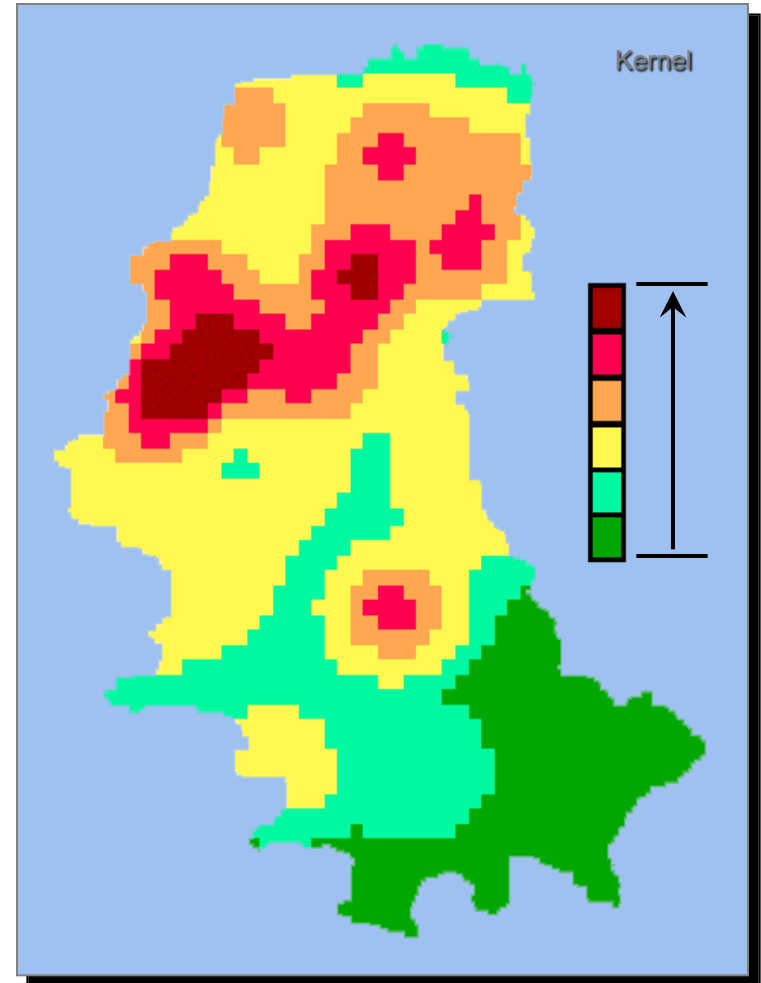
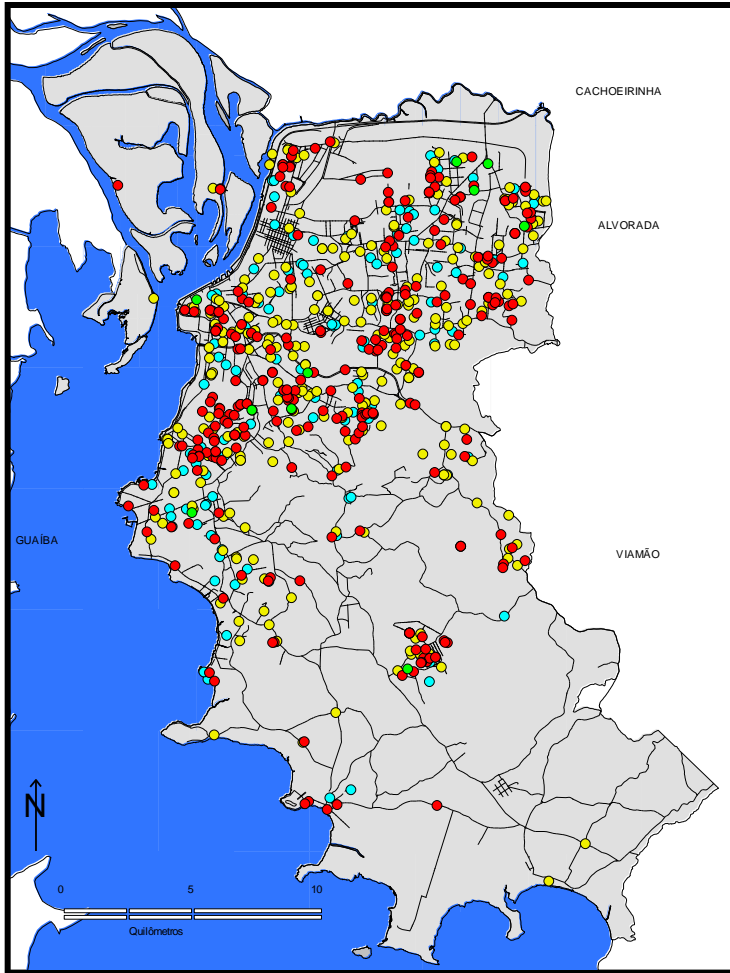
Superfície de saída



Ponto a ser estimado



Estimador de Intensidade (“Kernel Estimation”)



Legenda: ● Homicídios / ● Acidentes de transporte / ● Suicídios

Santos, S.M., 1999

Estimador de Intensidade (“Kernel Estimation”)



Em resumo:

- *Kernel estimators* são uma alternativa viável a métodos mais sofisticados de interpolação, pois não requerem a parametrização da estrutura de correlação espacial (como no caso da geoestatística).
- As superfícies interpoladas são suaves e aproximam muitos fenômenos naturais e socioeconômicos.
- As desvantagens destes estimadores são a forte dependência no raio de busca e a excessiva suavização da superfície, que pode em alguns casos esconder variações locais importantes.

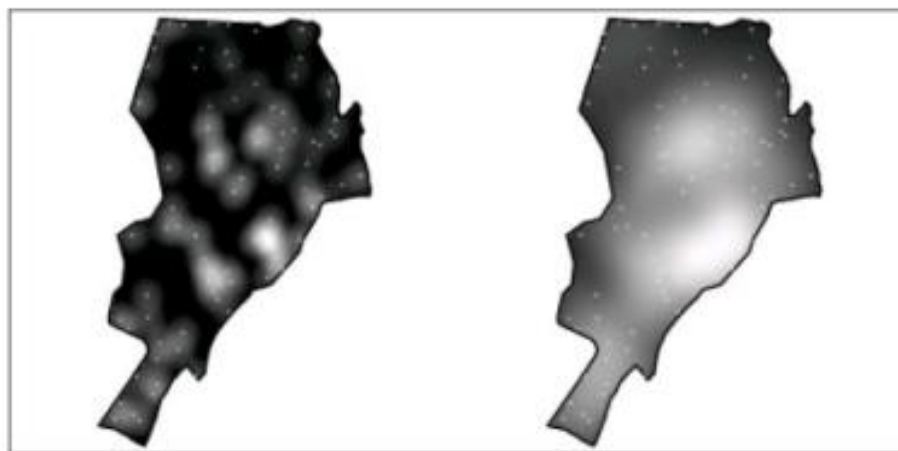


Figura 3-5- Superfícies de teor de argila interpoladas por kernel de quarta ordem. À esquerda, raio de busca de 500m; à direita, raio de busca de 1500m.

Mapa de Kernel da população residente em 2010 em setores censitários
Municípios do Projeto URBISAmazônia - Pará - Brasil

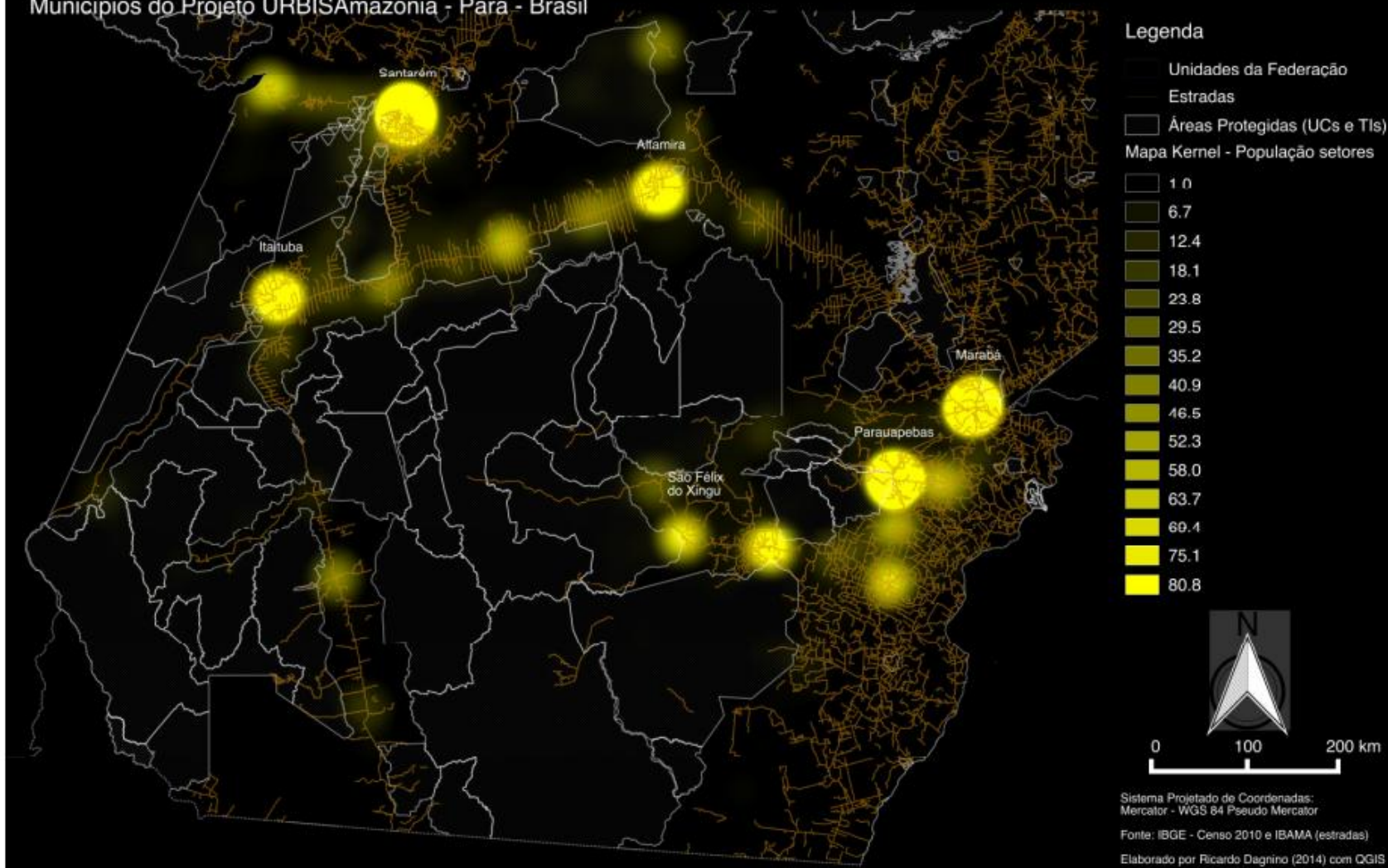


Figura 2 – Municípios do Projeto URBISAmazônia: (a) setores censitários classificados pela situação urbana ou rural, e sem população (em branco com hachuras); (b) Mapa de Kernel da distribuição da população nos setores censitários;

David Martin

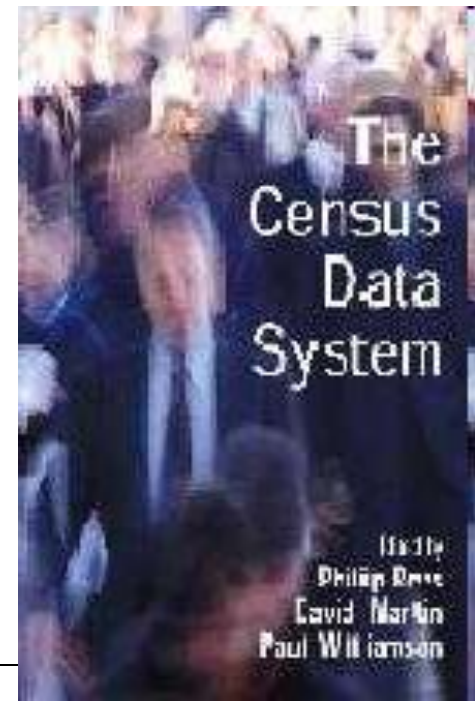
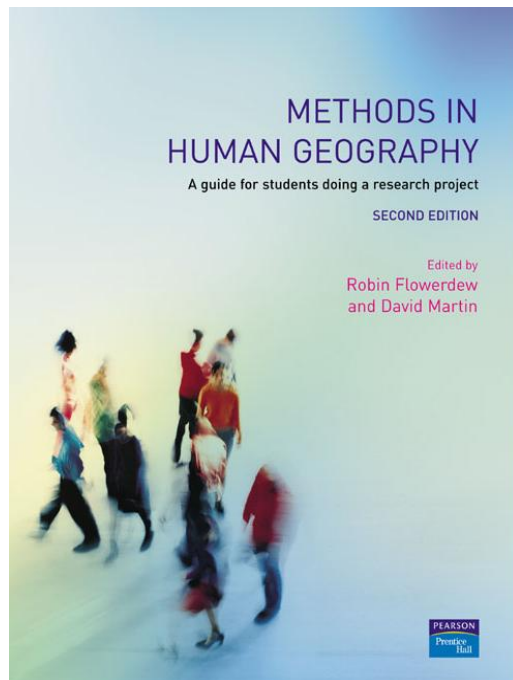
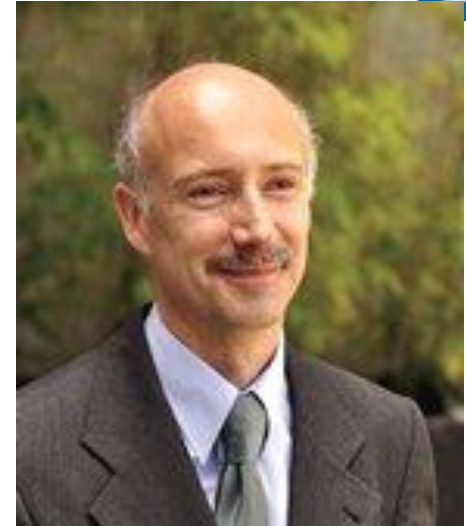
Professor na School of Geography, University of Southampton

Coordenador do ESRC/JISC 2001 Census of Population Programme

Co-Diretor do ESRC National Centre for Research Methods.

<http://www.public.geog.soton.ac.uk/users/martindj/>

<http://www.southampton.ac.uk/geography/about/staff/djm1.page?>



Grid modelling principles (1)

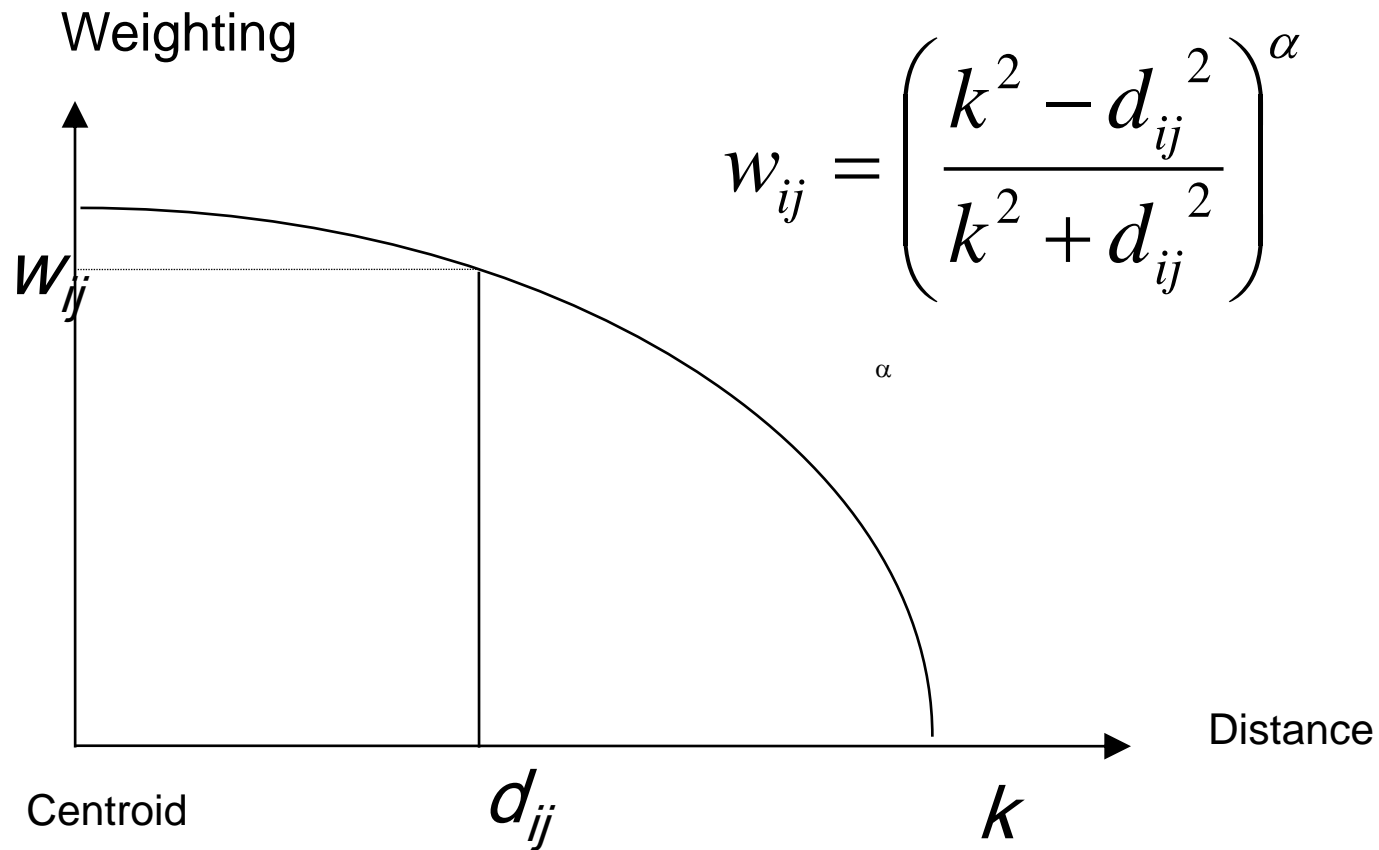


David Martin,
School of Geography
University of Southampton

- Population-weighted centroid(s) as summary points of local distribution
- Locally adaptive kernel estimation based on inter-centroid distances
- Redistribution of centroid counts into grid
- Could use dasymetric and other methods
- *This is NOT interpolation*

Grid modelling principles (2)

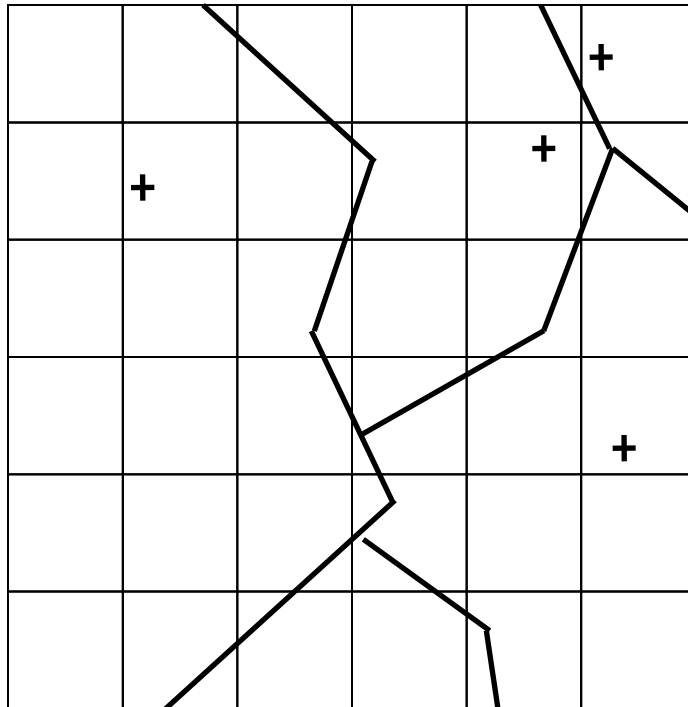
David Martin,
 School of Geography
 University of Southampton



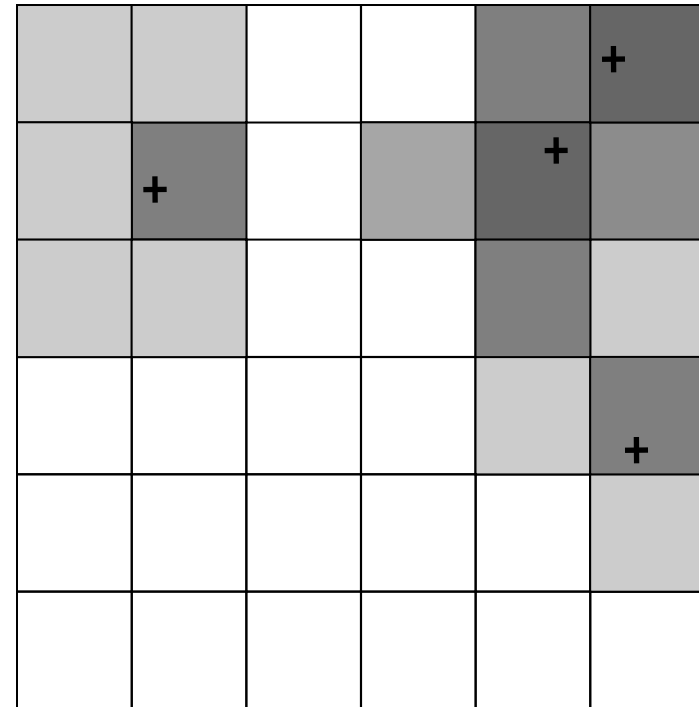
Grid modelling principles (3)

David Martin,
School of Geography
University of Southampton

Centroids and boundaries



Gridded population model



Centróides Ponderados de Martin



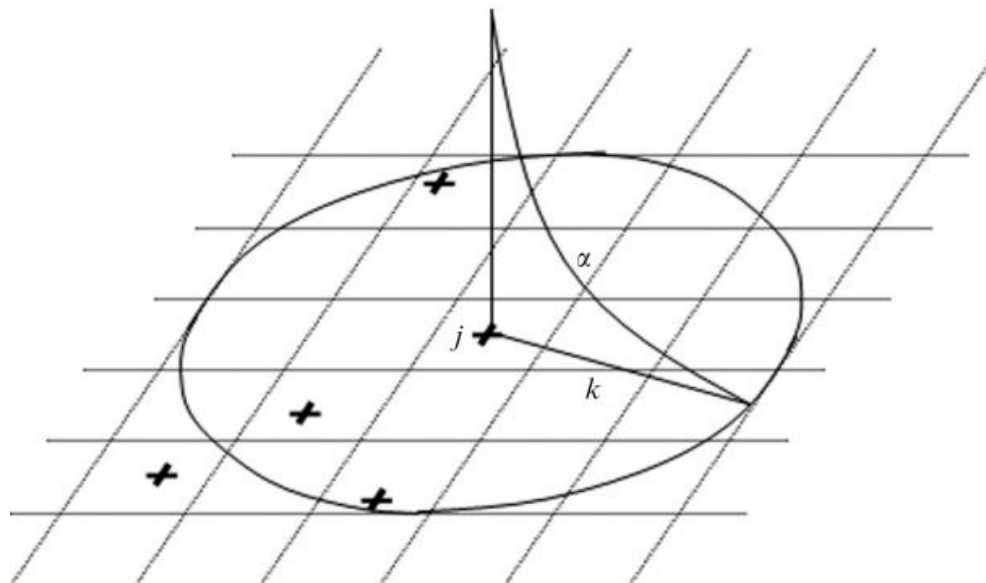
Martin, D. Mapping population data from zone centroid locations. *Transactions of the Institute of British Geographers NS*, 14, 90-97, 1989.

Bracken, I. and Martin, D. The generation of spatial population distributions from census centroid data. *Environment and Planning A*, 21, 537-543, 1989.

Martin, D. and Bracken, I. Techniques for modelling population-related raster databases. *Environment and Planning A*, 23, 1065-1079, 1991

Bracken I, Martin D. Linkage of the 1981 and 1991 UK Censuses using surface modelling concepts" *Environment and Planning A* 27 379-390, 1995.

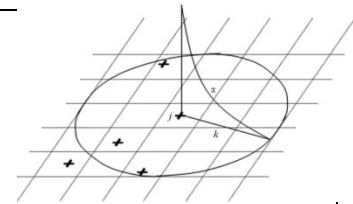
Centróides ponderados de Martin



Redistribuição da contagem populacional a partir do centroide j para as células dentro da largura de kernel k , usando a função de decaimento com a distância α .

Martin, Lloyd & Shuttleworth, 2011. Evaluation of gridded population models using 2001 Northern Ireland Census data. *Environment and Planning A* (43): 196-1980

Centróides ponderados de Martin



Cada centróide é considerado um ponto de soma local para uma distribuição populacional atual mas desconhecida;

Uma Zona pode ser representada por mais de um centróide (caso haja)

O modelo processa cada centróide por vez, estimando uma função de peso local para a redistribuição

Processo é definido por um tamanho de kernel inicial, definido pelo usuário, no qual a distância média entre centróides é determinada;

A largura do kernel é então adaptada localmente para ser igual a estes peso e distância atribuídos para as células locais de acordo com a função de decaimento com a distância:

$$w_{ij} = \left(\frac{k^2 - d_{ij}^2}{k^2 + d_{ij}^2} \right)^\alpha$$

D.Martin

$$w_{ij} = \left(\frac{k^2 - d_{ij}^2}{k^2 + d_{ij}^2} \right)^\alpha$$

k - é a largura de kernel ;

d_{ij} - é a distância entre o centro da célula i ao centroide j .

Quando $d > k$, todos os pesos são zero $w_{ij} = 0$

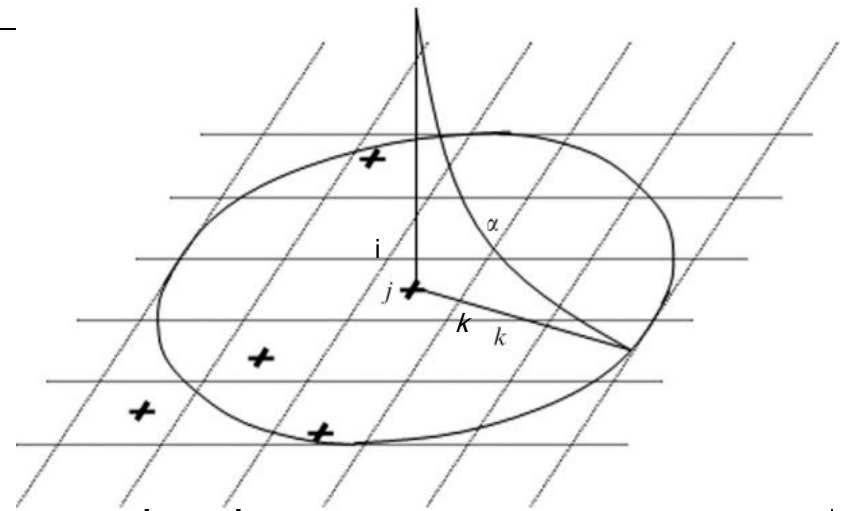
α - expoente que controla a forma da função de decaimento com a distância, dentro da extensão do kernel.

Permite variação da inclinação só variando α :

$\alpha=1$ declínio uniforme do centroide para o limite do kernel

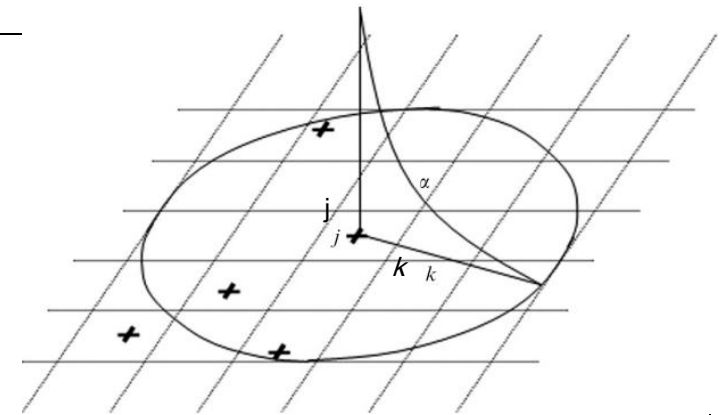
$\alpha < 1$ produzem kernels mais planos com declínio mais forte próximo ao limite do kernel;

$\alpha > 1$ produzem kernels de funções mais agudas com rápido declínio em direção ao limite do kernel



D.Martin

O total de população de cada centroide é redistribuído em função dos pesos atribuídos para as células dentro do kernel.



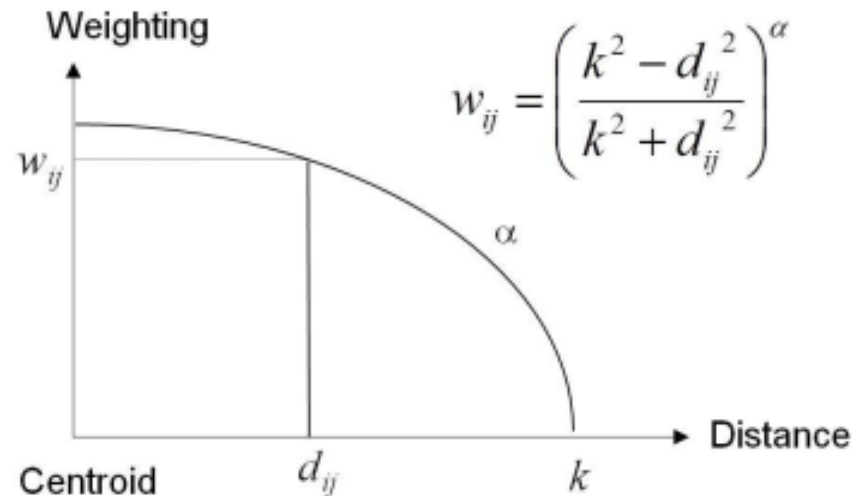
Todos os outros centróides são processados da mesma forma....então:

A população total de célula *i* é a soma das populações ponderadas a partir de todos os centroides:

$$\hat{P}_i = \sum_{j=1}^N P_j w_{ij}$$

N : total de centroides

P_j : é a população no centroide *j*



Martin et al 2011



Objetivos:

- Impacto dos parâmetros α e k e o tamanho da célula na estimativa de população em células (grades);
- impacto do uso de centroides em diferentes escalas espaciais;
- performance do modelo em áreas urbanas x rurais

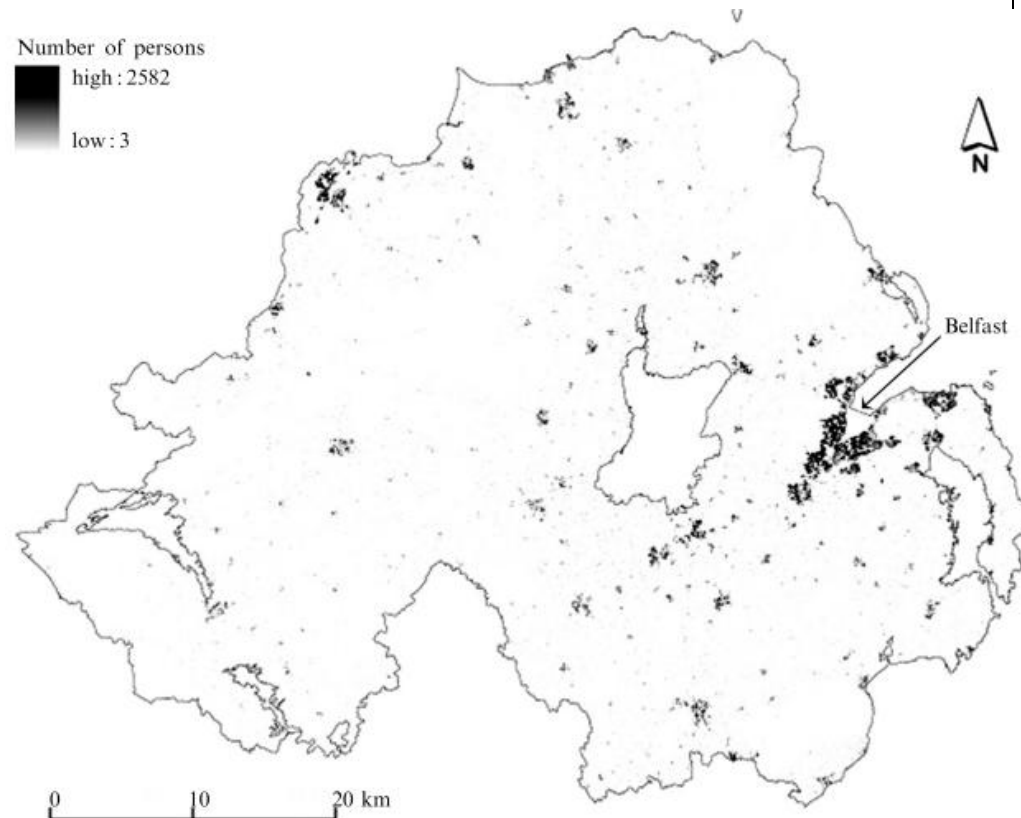
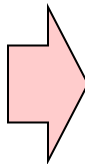
(Distrito de Belfast - Irlanda do Norte)

Células de 100m e 1km :

centroides \rightarrow modelo de grade
(superfície)

X

“True population counts”
(NI Grid Square Product)



Modelos

Table 1. Summary of model runs reported in tables 2–7.

Table	Coverage ^a	Input data ^b	Parameters held constant	Parameters varied	Outcome measures
2	NI	OA	none	mass preservation,	presence/absence of population
3	NI	postcode		search radius, distance decay, cell size	
4	NI	OA	cell size (50 m)	mass preservation,	mean error, standard deviation, root mean square error
5	Belfast	OA	cell size (50 m)	search radius,	
6	Belfast	OA	cell size (25 m)	distance decay	
7	NI	postcode	cell size (50 m)		

Note. In tables 2–7 local mass preservation refers to the constraint of population redistribution within the same zone; in global mass preservation, totals are constrained only within the entire study area.

^a NI—Northern Ireland.

^b OA—output area.

Células de 100m – Presença/ausência de população - centroides OA

Table 2. Presence or absence of population in 100 m cells from output area (OA) centroids.

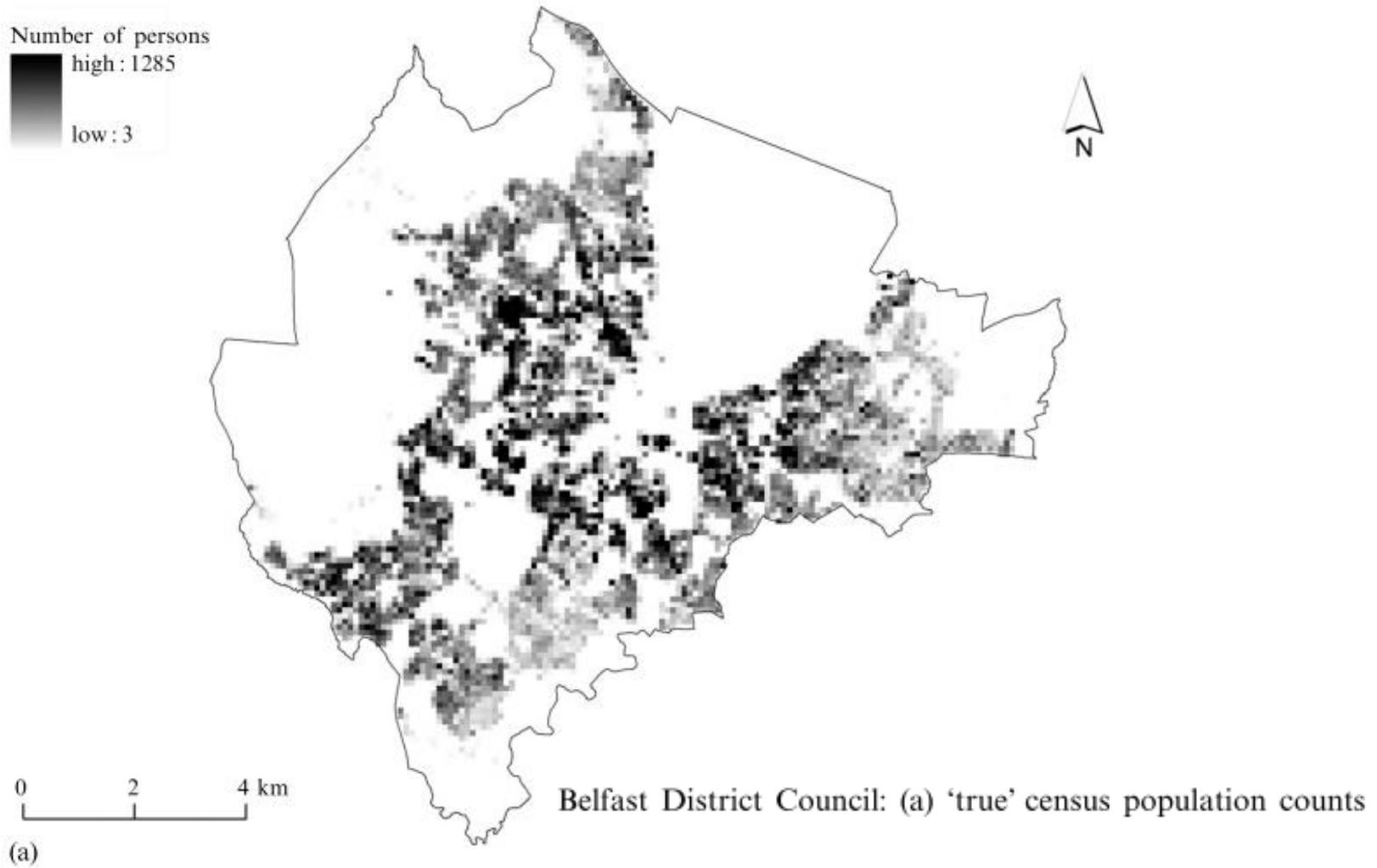
Model	Mass preservation	Search radius, k (m)	Distance decay, α	Cell size (m)	Percentage correctly	
					populated	unpopulated
1	OA centroids—global	na	na	100	3.70	99.9
2	OA zones—global	na	na	na	56.8	79.1
3	global	250	2.00	25	30.8	97.3
4	local	250	2.00	25	30.2	97.4
5	global	500	1.00	50	42.2	91.5
6	global	500	2.00	50	39.9	93.0
7	local	500	1.00	50	40.8	91.8
8	local	500	2.00	50	38.9	93.2

Células de 100m – Presença/ausência de população – centroides códigos postais

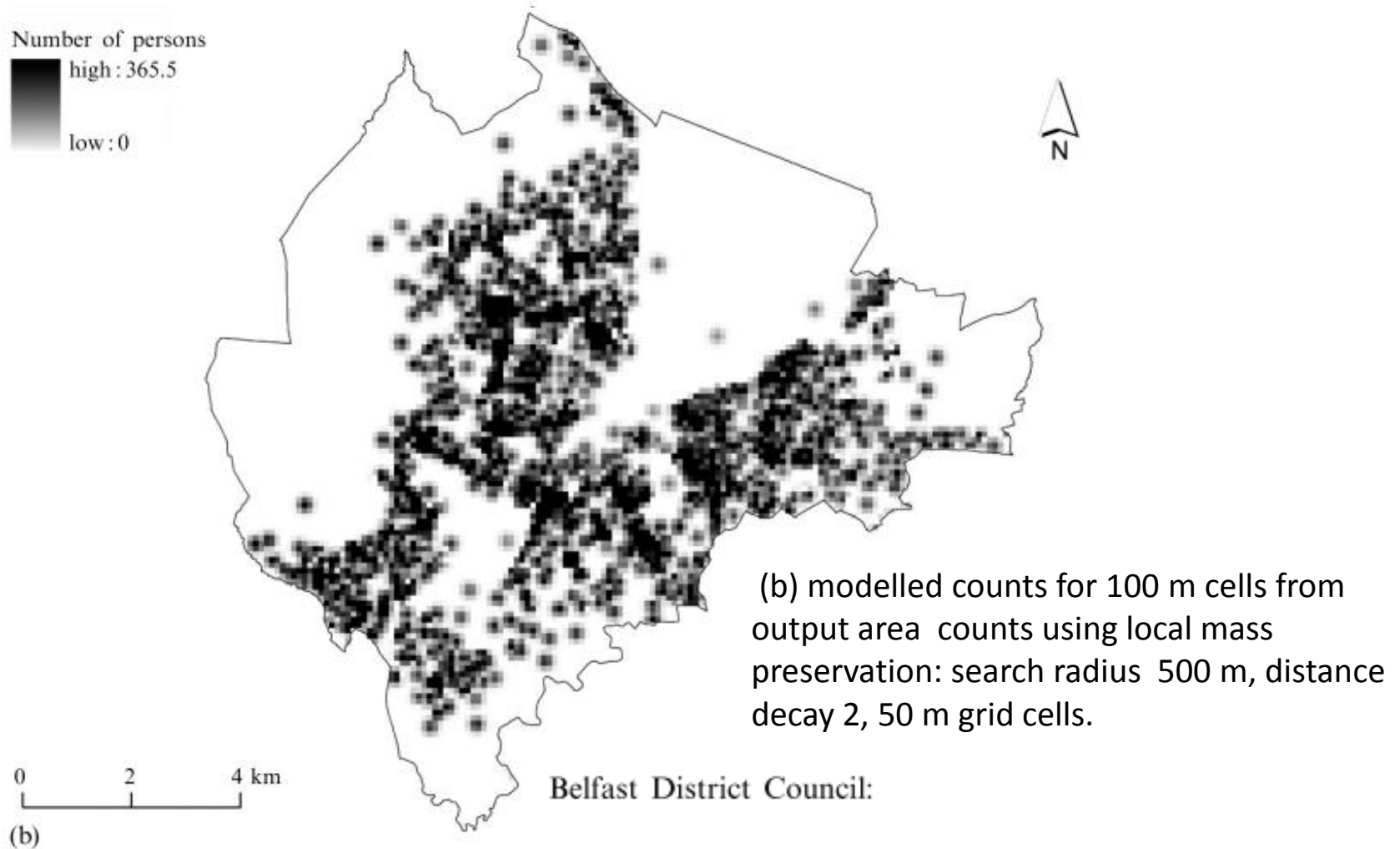
Table 3. Presence or absence of population in 100 m cells from postcode centroids.

Model	Mass preservation	Search radius, k (m)	Distance decay, α	Cell size (m)	Percentage correctly	
					populated	unpopulated
1	global	500	0.25	50	66.3	79.9
2	global	500	1.00	50	67.3	81.5
3	global	500	2.00	50	65.8	83.7
4	local	500	0.25	50	65.8	80.4
5	local	500	1.00	50	66.3	82.1
6	local	500	2.00	50	64.9	84.4

Células de 100m – Presença/ausência de população



Células de 100m – Presença/ausência de população



Contagem de População

Table 4. Prediction errors (prediction – observed) for Northern Ireland: prediction to 50 m cells and aggregation to 100 m cells.

Model	Mass preservation	Search radius, k (m)	Distance decay, α	Cell size (m)	Mean error	Standard deviation	RMSE ^a
1	global	250	0.25	50	-7.12	36.77	37.46
2	global	250	1.00	50	-6.20	38.07	38.57
3	global	250	2.00	50	-5.49	41.52	41.88
4	global	500	0.25	50	-10.58	32.96	34.61
5	global	500	1.00	50	-9.60	32.34	33.74
6	global	500	2.00	50	-8.76	32.27	33.43
7	local	250	0.25	50	-6.82	35.71	36.36
8	local	250	1.00	50	-5.81	37.18	37.63
9	local	250	2.00	50	-5.06	40.41	40.73
10	local	500	0.25	50	-10.26	32.56	34.13
11	local	500	1.00	50	-9.54	32.29	33.67
12	local	500	2.00	50	-9.89	32.36	33.56

^a RMSE—root mean square error.

Conclusões:

- Algumas propriedades são inerentes às grades , e.g. Independência da mudança dos limites
- Modelos de grades não são comuns, tem de ser construído em muitos países
- Preservar as vantagens de áreas sem população e padrões de ocupação dependem do dado de entrada e dos parâmetros de modelo utilizados, tais como tamanho da célula e extensão da função espacial de redistribuição.
- Só pôde fazer análise do modelo pq em NI o dado em grade é adquirido regularmente
- Modelos de grade resultante não foram muitos sensíveis aos parâmetros dos modelos, (apenas a algumas variações nos RMSE).
- **Raio de busca (k) e tamanho de célula foram mais importantes que a função de decaimento (α)**

Conclusões:

- Há indicações que a preservação de massa total é melhor que a global em áreas densamente populadas → modelos de pop em grade apresentam melhor performance com diferentes parâmetros em áreas urbanas que rurais.

-Contudo, a resolução espacial dos centroides de entrada pareceu ser mais relevante que a parametrização do modelo

-Escala espacial – em termos geoestatísticos – se mostrou mais importante que a seleção dos parâmetros do modelo.

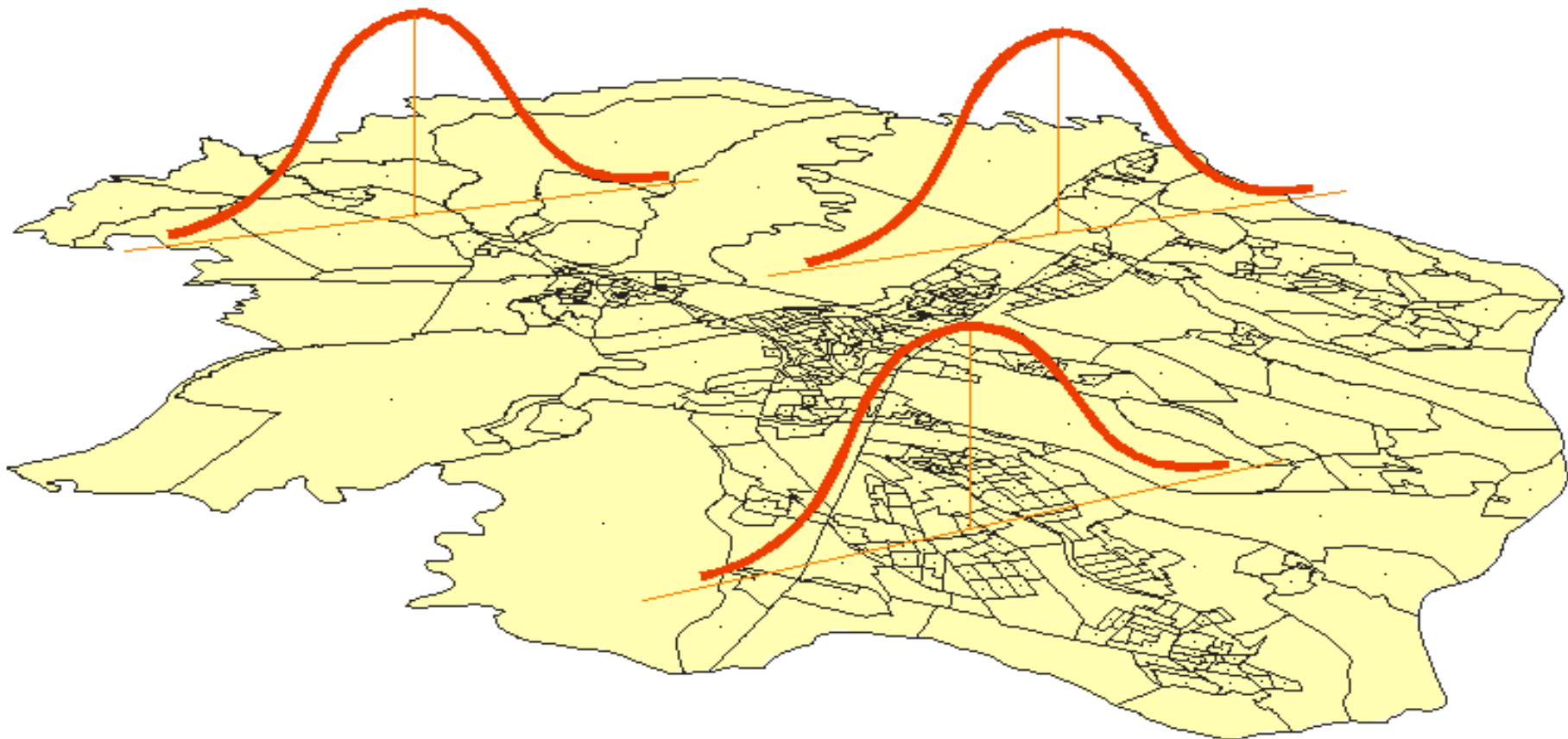
Conclusões:

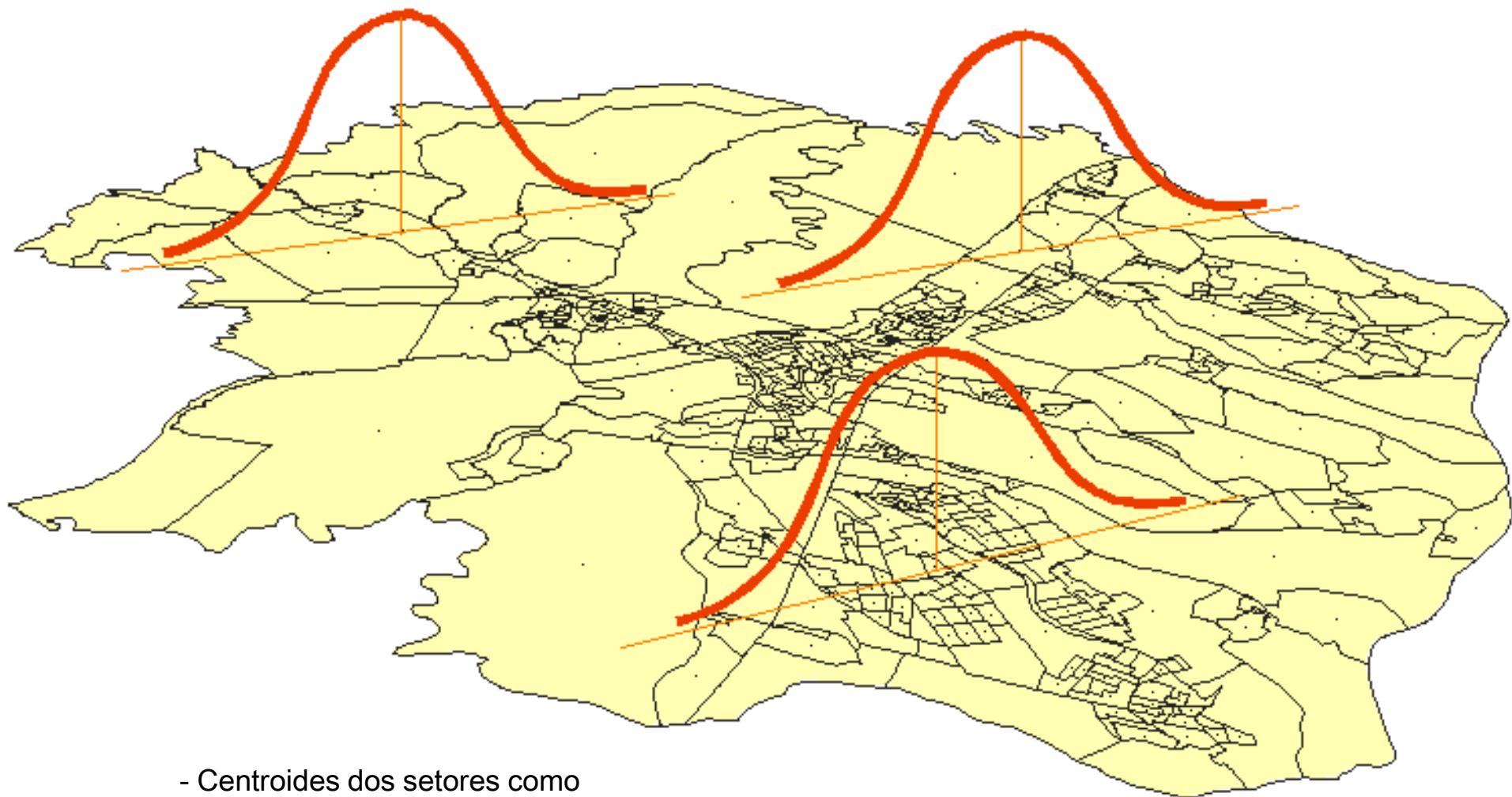
- Escala espacial – em termos geoestatísticos – se mostrou mais importante que a seleção dos parâmetros do modelo.
- Importância do lugar e estrutura espacial da população:
 - Interesse nas áreas urbanas – tamanho da população nas células ao invés de presença/ausência (rurais)
- Maiores Desafios: áreas com extensas zonas geográficas e população esparsa
- Ao invés de regras – observar contexto geográfico, e validar comparando com outros métodos

Intensidade populacional *local* da unidade de área j

- ❑ Considera a população de j e a população das unidades vizinhas a j
- ❑ Soma da população de todas as áreas, ponderadas por sua proximidade à área j
- ❑ Conceito de proximidade é definido por função escolhida pelo usuário – Função de Decaimento (com a distância)

Fonte: Flavia Feitosa

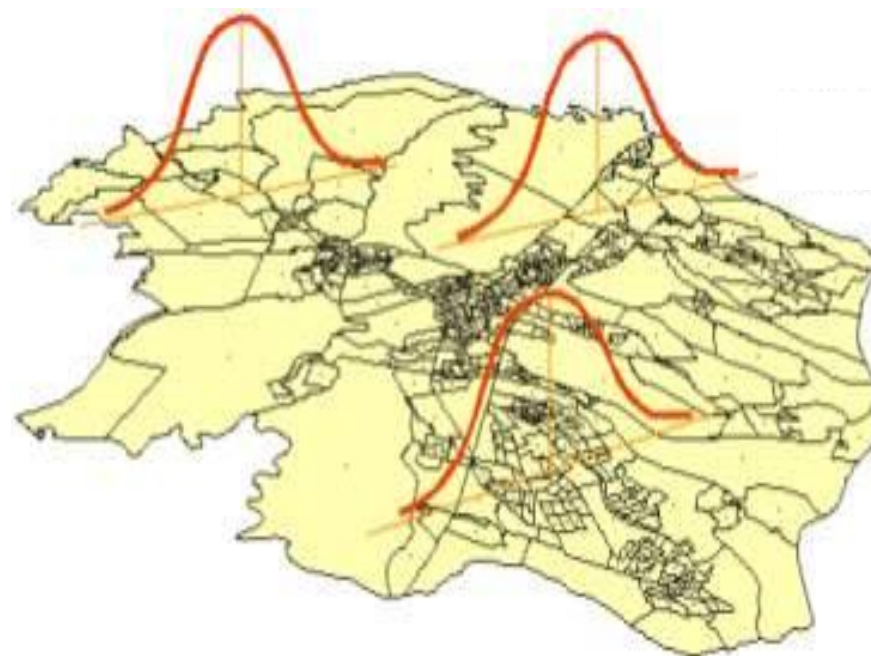
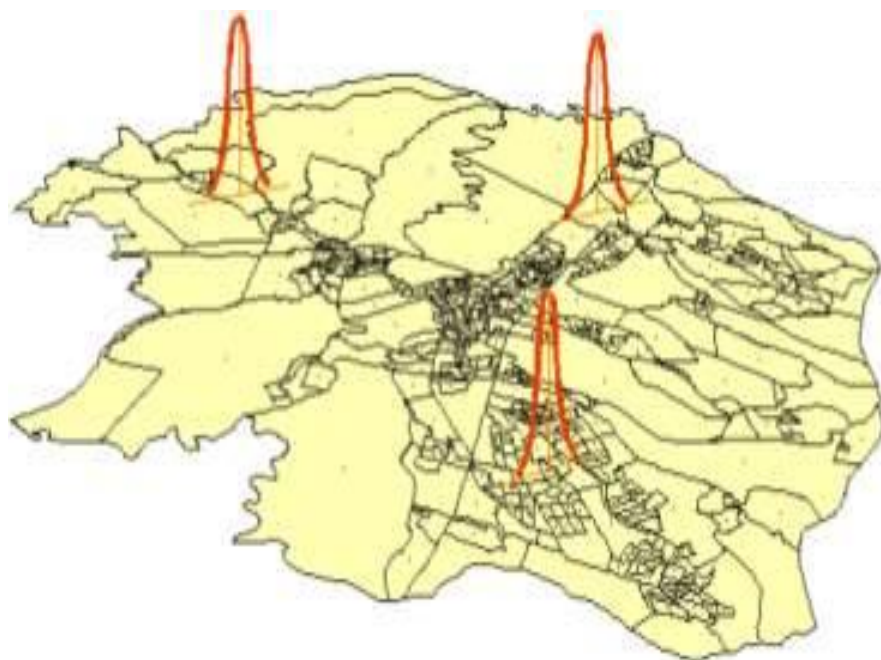




- Centroides dos setores como
Pontos de referência para
cálculo das distâncias

$$\hat{P}_i = \sum_{j=1}^N P_j W_{ij}$$

Testes de larguras de banda: cada largura adequada a uma questão → larguras mais largas para testes terias como o aumento da macrosegregação



$$W_{ij} = \left\{ \frac{d^2 - s^2}{d^2 + s^2} \right\}^\alpha$$

d - distância média entre os centróides da janela amostral

s_{ij} - distância entre a célula i e o centróide j .

Afetam os resultados da superfície (modelo) resultante:

- Resolução da grade;
- Tamanho da janela
- Parâmetro de decaimento

E

- seleção das localizações dos centróides

Este procedimento – abordagem de ajuste com preservação de volume - tem **inconvenientes significativos** onde:

- a geografia da região de estudo é muito variável e/ou
- onde os atributos não variam de modo similar (não exibem forte covariância positiva)

Fronteiras urbano-rural e em áreas rurais → abordagens com informações auxiliares, como cobertura da terra, densidade da rede de estradas e imagens de SR → melhor ajuste para atribuir localização.

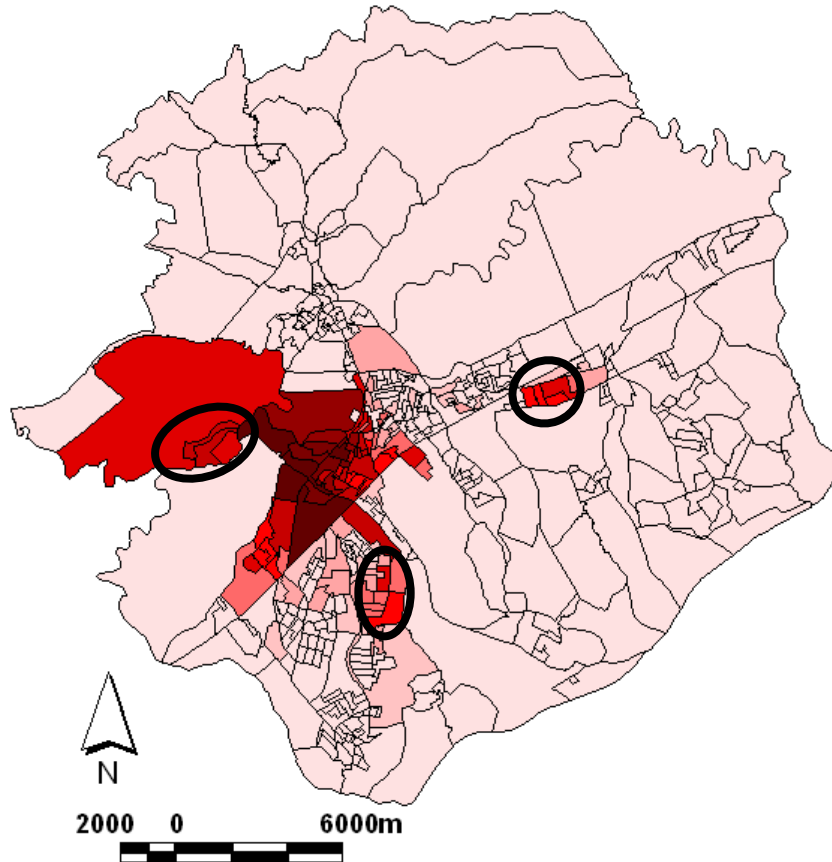
- Dados auxiliares podem ser úteis para ponderação

Dados Discretos X Contínuos

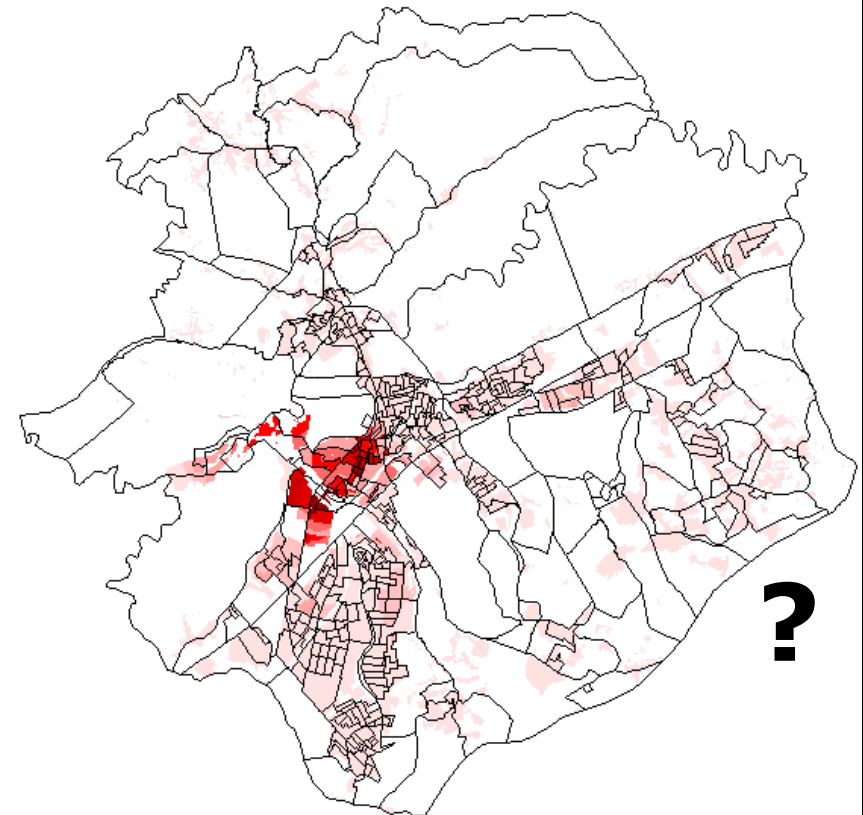
(função gaussiana – largura de banda = 400 m)

Áreas reconhecidamente segregadas, como os condomínios Quinta das Flores, o Cond. Floresta e o bairro Urbanova (onde predominam condomínios fechados) não são realçados

Dados Agregados



Densidade Populacional



Dados Discretos X Contínuos

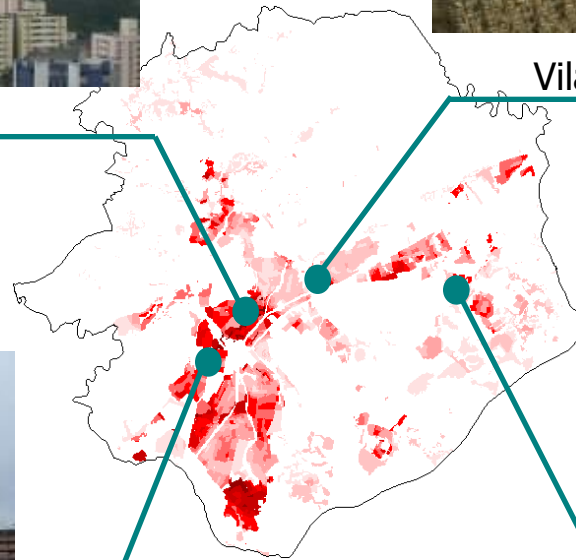
100 Células com as mais elevados índices locais de dissimilaridade

→ Ocupação vertical de alta densidade (mesmo no centro)



Bairros próximos ao Centro

Vila Industrial: Conjuntos Residenciais



Parque Residencial Aquarius

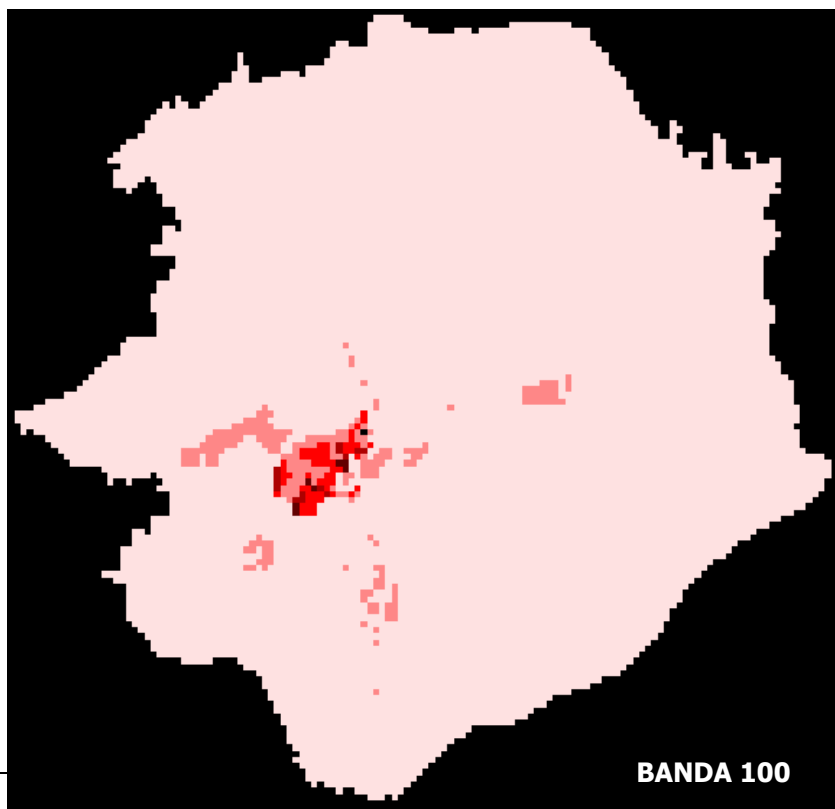


Jardim Nova Michigan:
Conjunto Residencial

Testes com superfícies (grid 200m): função gaussiana, bandas 100, 500 e 2000.

Índices globais de isolamento do grupo mais de 20SM: 8,29 (banda 100); 6,8 (banda 500) e 4,02 (banda 2000).

Quanto maior a largura de banda, menor o isolamento do grupo.



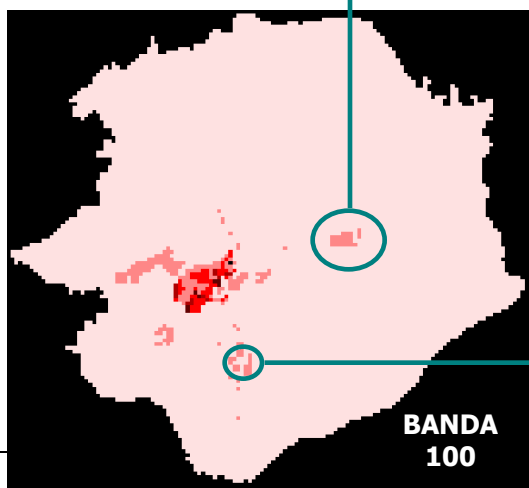
BANDA 100

A diminuição da largura de banda possibilita o aparecimento de núcleos de isolamento mais “locais”.

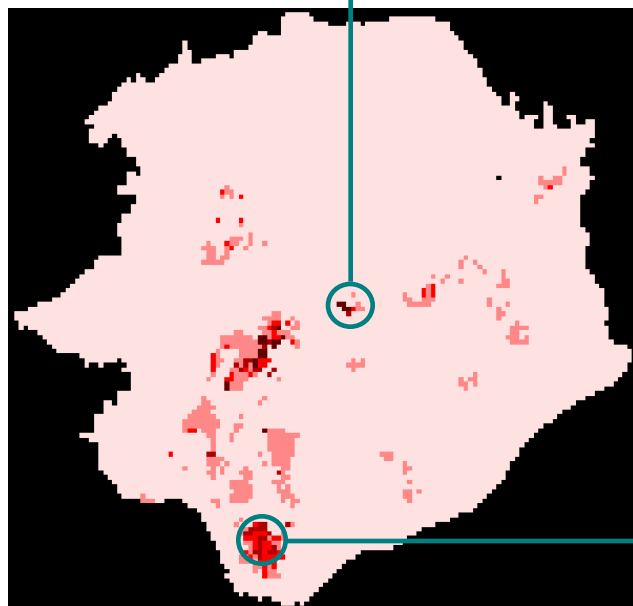
Fonte: Flavia Feitosa



Cond. Floresta/Vista Verde



Bosque dos Eucaliptos/Quinta das Flores

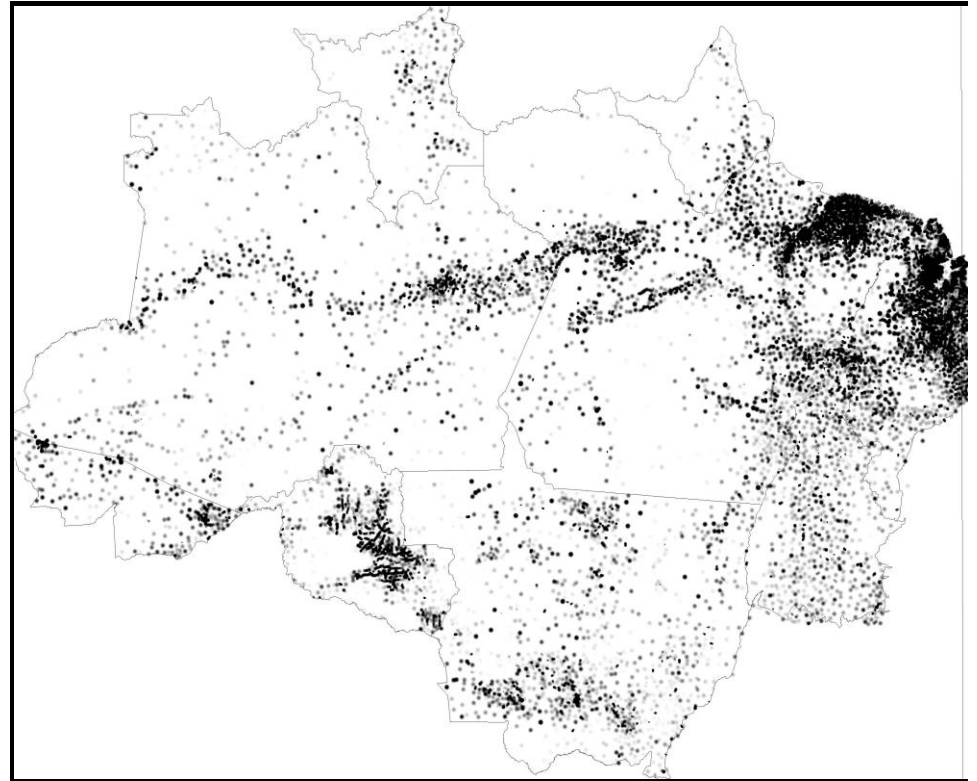


População na AmzL



Centróides Ponderados Martin

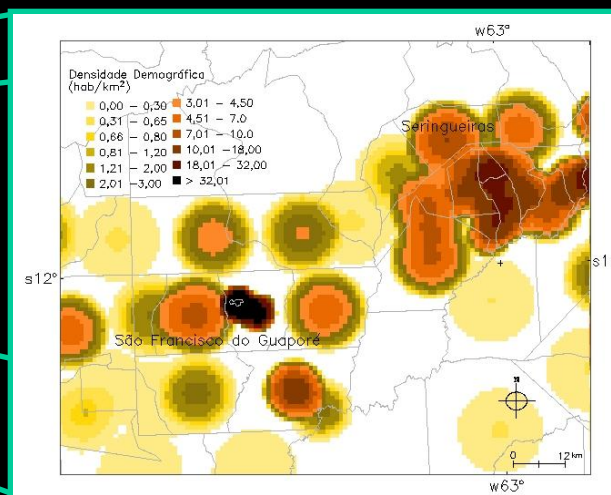
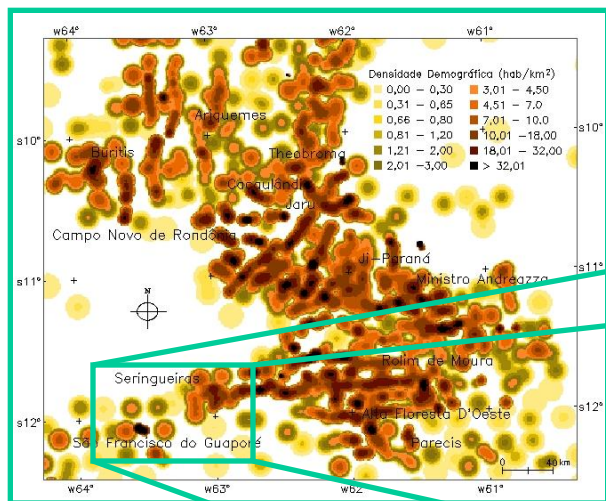
- Preserva volumes e vazios demográficos
- Kernel adaptativo a concentração de centróides
- População total distritos 2000
- Superfície “negativa”
- Depende da densidade de centróides
- Por setor censitário 2000



População Setores Censitários 2000

População na AmzL

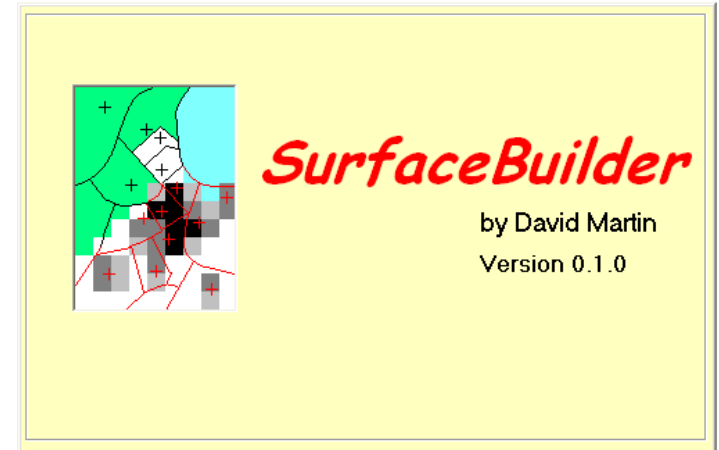
Centróides Ponderados Martin



SurfaceBuilder

David Martin,
School of Geography
University of Southampton

- Download and install VB program
- Download X,Y,Z centroid data file
- Specify surface parameters
- Run and preview model
- Export to GIS



<https://www2.geog.soton.ac.uk/users/martind/davehome/software.htm>

SurfaceBuilder sequence

Microsoft Excel - allhantspop.csv

	A1	Easting						
	A	B	C	D	E	F	G	H
1	Easting	Northing	ks0010001	Zone Code				
2	469292	151528.4	446	24UBHP0001				
3	470847.7	153939.6	147	24UBHP0002				
4	470275.2	154636.9	152	24UBHP0003				
5	470385.4	153128.2	214	24UBHP0004				
6	466758.7	152419	369	24UBHP0005				
7	465774.9	153865.2	298	24UBHP0006				
8	466013.1	153419.7	348	24UBHP0007				
9	467134.6	152254.2	321	24UBHP0008				
10	466689.9	151966.7	340	24UBHP0009				
11	466970.1	152061.8	305	24UBHP0010				
12	466806.7	154574.1	382	24UBHP0011				
13	466098	154054.1	386	24UBHP0012				
14	466260.2	153902	292	24UBHP0013				
15	466313.2	153233.5	355	24UBHP0014				
16	465894.8	153756.1	265	24UBHP0015				
17	465848.5	153643.2	255	24UBHP0016				
18	465572.8	153545.8	264	24UBHP0017				
19	465788.8	153488.2	295	24UBHP0018				
20	468265.6	154073.8	359	24UBHP0019				
21	465625.5	153248.5	352	24UBHP0020				
22	465508.4	152857.2	303	24UBHP0021				
23	467076.4	152860.3	224	24UBHP0022				
24	466782.8	152992.8	284	24UBHP0023				
25	466341.7	151348.3	289	24UBHP0024				
26	467278.2	152771.2	342	24UBHP0025				
27	467465.1	152308.8	298	24UBHP0026				

SurfaceBuilder

File Tools Window Help

New Surface 1

Add Data... Add Mask Build Export...

Data source
D:\Presentations\ESRC RMF July 06\allhantspop.csv
Delete Data
Number of selected files 1

Mask source
No mask selected
Delete Mask

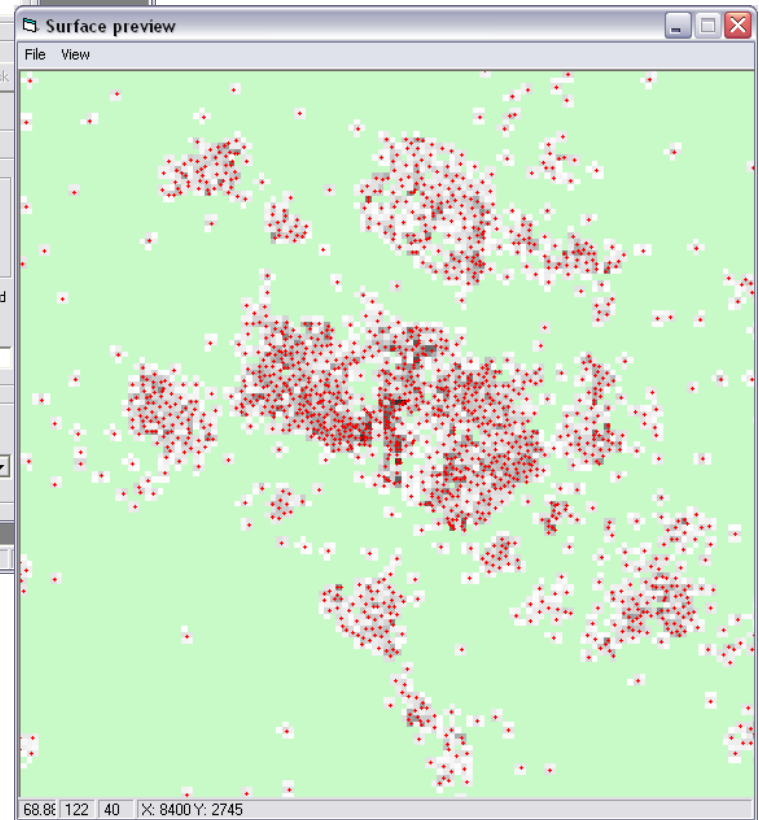
Surface parameters

Area to be modelled
Low High
X 430000 455000 Input records: 4816 Get from files
Y 100000 125000 Z total: 1457548 Whole cells

Cell size 200 Decay function Cressman Draw on build
Search radius 250 Decay parameter 1 Use mask
Description

Output options
Output file format Numeric data type Output order Nodata value
ArcInfo ASCII Real Row prim NW 0

20/06/2006



David Martin,

School of Geography

University of Southampton

Pode ser útil??

Disasters, 2012, 36(3): 477–494. © 2012 The Author(s). Journal compilation © Overseas Development Institute, 2012
Published by Blackwell Publishing, 9600 Garsington Road, Oxford, OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

Inequalities in exposure and awareness of flood risk in England and Wales

Jane L. Fielding Senior Lecturer in Quantitative Methods, Department of Sociology, University of Surrey, United Kingdom

This paper explores the environmental inequalities of living in the floodplains of England and Wales and the differences in flood awareness of those ‘at risk’. An area comparison is made between an etic, objective flood risk exposure, and an emic, subjective perception of that risk by social class. In all areas except the Midlands, the working classes were more likely to reside in the floodplains; the greatest exposure inequality is seen in the North East and Anglian regions. Flood awareness in the Anglian regions was much lower than average, but there were no significant class differences. In the Thames region, despite equal flood risk exposure between classes, the most deprived displayed the least awareness of flood risk. In the North East, inequalities in the distribution of flood risk exposure accompanied inequalities in perception, resulting in the least aware and most deprived experiencing the greatest flood risk.

Keywords: environmental justice, flood awareness, flood risk, inequalities



Pode ser útil??

Disasters, 2012, 36(3): 477–494. © 2012 The Author(s). Journal compilation © Overseas Development Institute, 2012
Published by Blackwell Publishing, 9600 Garsington Road, Oxford, OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

Inequalities in exposure and awareness of flood risk in England and Wales

Jane L. Fielding Senior Lecturer in Quantitative Methods, Department of Sociology, University of Surrey, United Kingdom

This paper explores the environmental inequalities of living in the floodplains of England and Wales and the differences in flood awareness of those 'at risk'. An area comparison is made between an etic, objective flood risk exposure, and an emic, subjective perception of that risk by social class. In all areas except the Midlands, the working classes were more likely to reside in the

Antropologia e ciências sociais, 2 tipos de pesquisa de campo – em função da óptica de observação:

Ética – percepção do observador X

Êmica - percepção do grupo social, ou do objeto de estudo

Keywords: environmental justice, flood awareness, flood risk, inequalities



Inequalities in exposure and awareness of flood risk in England and Wales

Para explorar medidas de desigualdade entre as planícies de inundação...

- [Fielding and Burningham \(2005\)](#) usaram *Surface Builder* para redistribuir características da população, a partir do censo UK 1991 – em superfícies (Martin and Brackan, 1999; Surface Builder, n.d.).
- Caracterizaram esta população por *classes sociais*, e as classificaram como população em risco de inundação a partir GIS e mapas de inundação.
- Observaram que as classes sociais mais baixas e os desempregados experimentavam um risco de inundação maior (Fielding and Burningham, 2005);
- Sem distinção se era risco fluvial ou de maré ...

Inequalities in exposure and awareness of flood risk in England and Wales

- Fielding (2007): fez distinção entre risco fluvial e de maré
- Usando **surface population models e análise de regressão logística**, observaram:
 - **Desigualdade significativa** entre as classes médias e a classe dos trabalhadores, e também entre as classes médias e os inativos, nos fatores de risco associados com emergências enchentes em todas as regiões da EA da Inglaterra e País de Gales, exceto a região de Midlands

Table 1 Social class groupings

Middle class	1. Higher and lower managerial and professional
	2. Intermediate occupations
	3. Small employers and own account workers
Working class	4. Lower supervisory and technical occupations
	5. Semi-routine and routine occupations
Inactive	6. Never worked and long-term unemployed/unclassified

Inequalities in exposure and awareness of flood risk in England and Wales



OBJETIVO

Explorar a distribuição de risco - dentro de um framework baseado em resultados para Inglaterra e País de Gales.

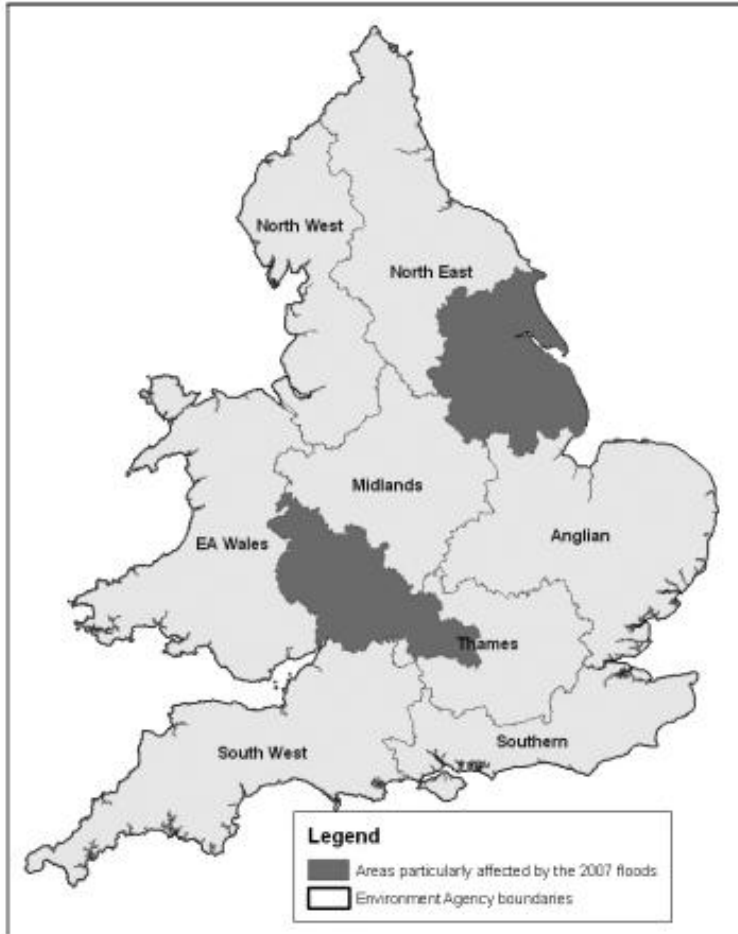
Relações entre desigualdades na distribuição do risco de inundação e a capacidade daqueles em maior risco de lidar com os riscos – RESILIENCIA

- Um aspecto da RESILIÊNCIA → Estar ciente dos riscos

DISCUSSÃO Ética x Êmica

Inequalities in exposure and awareness of flood risk in England and Wales

Figure 1 Environment Agency regions and the impact of the 2007 floods



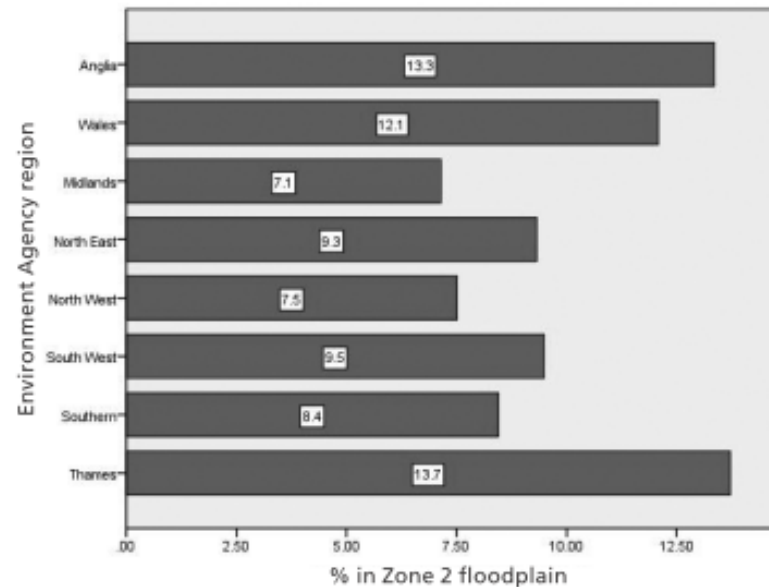
Source: Chatterton et al. (2010)

Metodologia

Surface Builder – estatística de distribuição de grandes áreas → grades de 200 m , baseadas na população dos centroides de cada área;

- Análises sobre as superfícies em GIS (Burningham (2005) e Fielding, (2007))

Figure 2 Proportion of households within the Zone 2 floodplains, by region



Source: author's analysis of ONS (2001)

Inequalities in exposure and awareness of flood risk in England and Wales

Metodologia que nos interessa:

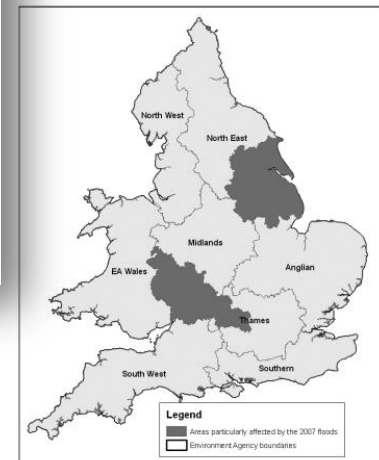
- Estatísticas de área para classes sociais foram acessadas do Censo 2001 (NS-SEC of Household Reference Persons (HRP) aged 16–74 in England and Wales)
- SPSS – registro das classes sociais → dados / pontos (csv)
- *Surface Buidar* – criou 6 grades uma para cada classe social → 3 grupos: classe média, trabalhadores e inativos

Table 1 Social class groupings

Middle class	1. Higher and lower managerial and professional
	2. Intermediate occupations
	3. Small employers and own account workers
Working class	4. Lower supervisory and technical occupations
	5. Semi-routine and routine occupations
Inactive	6. Never worked and long-term unemployed/unclassified

- Trouxe a estatística de áreas maiores para Superfície – grade de 200m

Figure 1 Environment Agency regions and the impact of the 2007 floods



Source: Chatterton et al. (2010)

Inequalities in exposure and awareness of flood risk in England and Wales

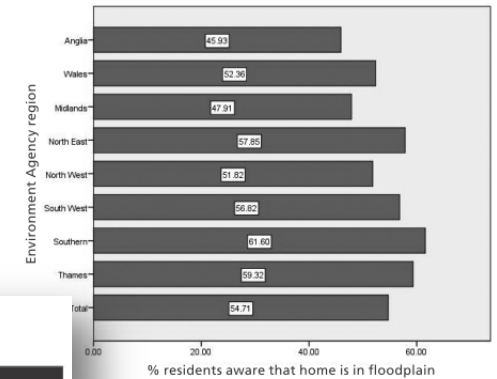
Resultados

Table 2 Proportion in Zone 2 floodplains, by region and class

Region	No. of households	Percentage of class within Zone 2 floodplains			
		Middle class	Working class	Inactive	Total
1. Anglia	2,090,562	11.7%	15.2%	14.7%	13.3%
2. Wales	1,135,195	11.4%	12.7%	12.4%	12.1%
3. Midlands	3,090,295	7.3%	7.1%	6.9%	7.1%
4. North East	2,629,608	8.4%	10.3%	9.4%	9.3%
5. North West	2,470,377	7.3%	7.8%	7.6%	7.5%
6. South West	1,513,869	9.2%	10.1%	9.3%	9.5%
7. Southern	1,657,054	7.9%	9.3%	8.8%	8.4%
8. Thames	4,444,968	13.3%	13.9%	14.6%	13.7%
Percentage overall in floodplains		10.1%	10.6%	10.5%	10.3%
Total households in floodplains		920,805	557,400	482,379	1,960,584

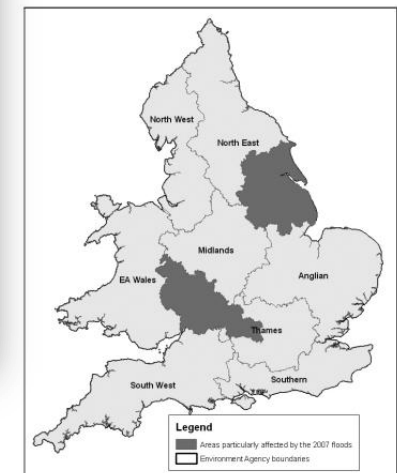
Source: author's analysis of ONS (2001)

Figure 3 Flood risk awareness in 2007–8, by region



Source: author's analysis of Ipsos MORI (2008)

Figure 1 Environment Agency regions and the impact of the 2007 floods



Source: Chatterton et al. (2010)

Inequalities in exposure and awareness of flood risk in England and Wales

Resultados – percepção do risco

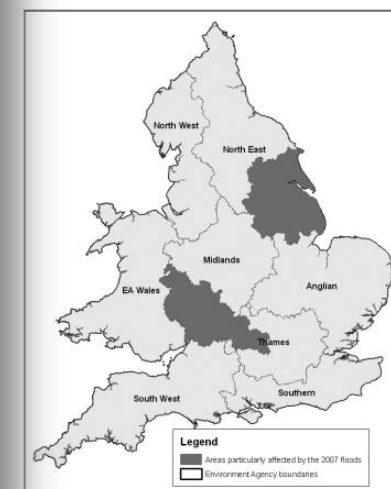
Table 3 Percentage of at-risk residents aware of flood risk, by region and class

Region	No.	Percentage of 'at risk' residents aware of flood risk				Significance	% flooded in 2007
		Middle class	Working class	Inactive	Total		
1. Anglia	159	46.4%	47.1%	37.5%	45.9%		1.2%
2. Wales	73	55.9%	48.4%	50.0%	52.1%		0.0%
3. Midlands	97	61.1%	48.0%	37.9%	47.4%		7.6%
4. North East	151	63.3%	46.3%	75.0%	57.6%	*	11.0%
5. North West	102	57.5%	46.0%	58.3%	52.0%		0.0%
6. South West	76	56.8%	58.3%	50.0%	56.6%		0.0%
7. Southern	79	55.8%	65.5%	85.7%	62.0%		0.0%
8. Thames	255	68.4%	39.7%	51.4%	59.3%	**	2.0%
Percentage aware		61.0%	47.7%	53.2%	54.6%	**	3.1%
Total respondents		482	411	141	1,034		

Notes: Chi-square significance: * $p < 0.01$; ** $p < 0.001$.

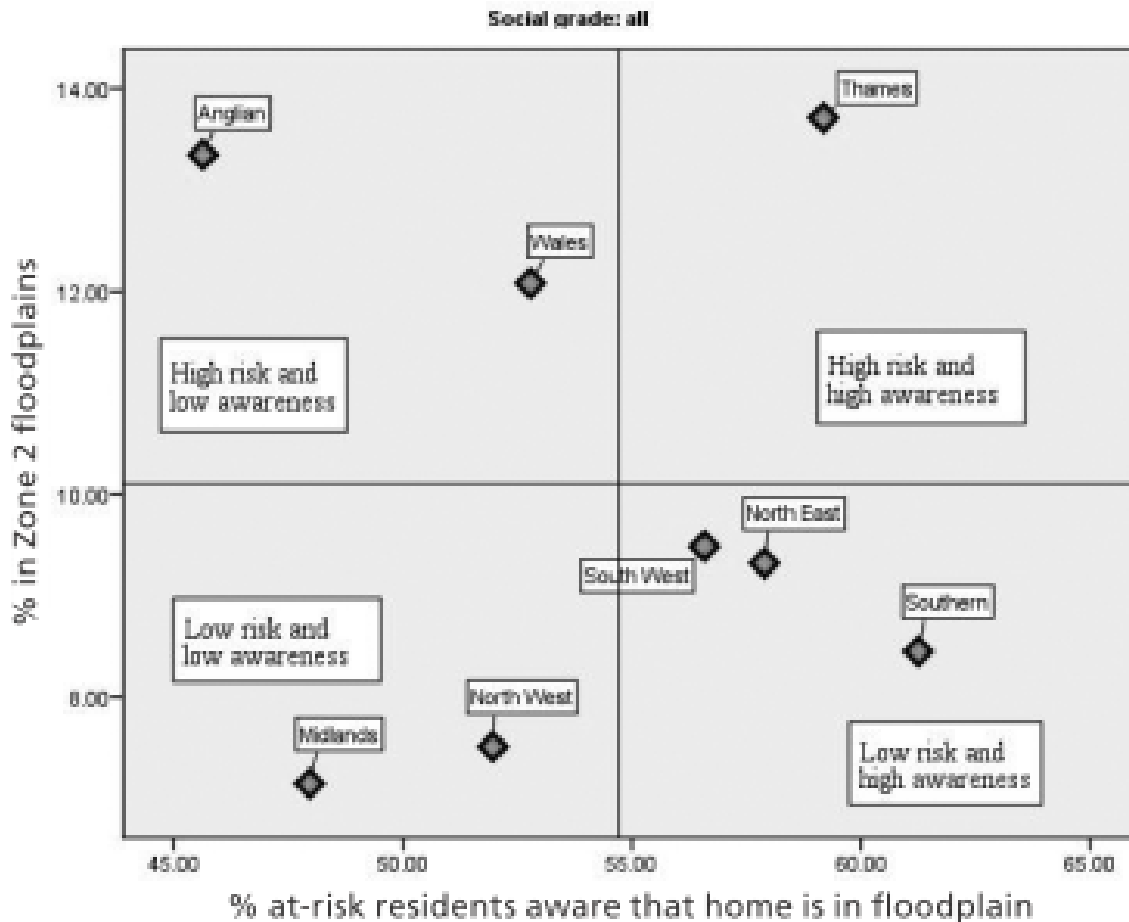
Source: author's analysis of Ipsos MORI (2008)

Figure 1 Environment Agency regions and the impact of the 2007 floods



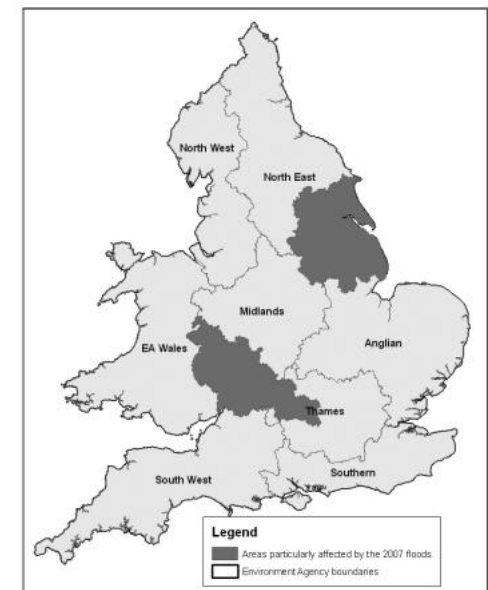
Source: Chatterton et al. (2010)

Figure 4 Overall flood risk and flood awareness in 2007-8



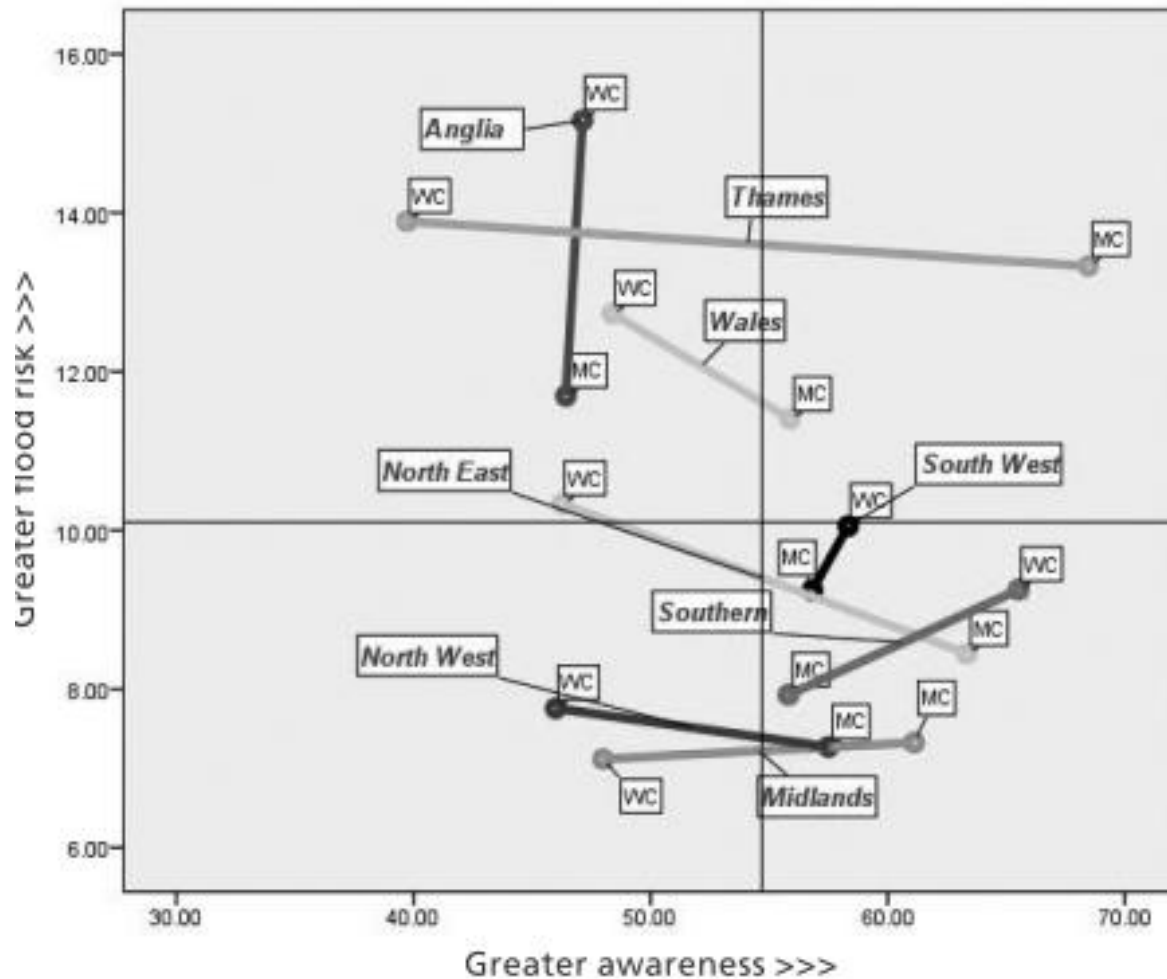
Source: author's analysis of Ipsos MORI (2008); ONS (2001)

Figure 1 Environment Agency regions and the impact of the 2007 floods



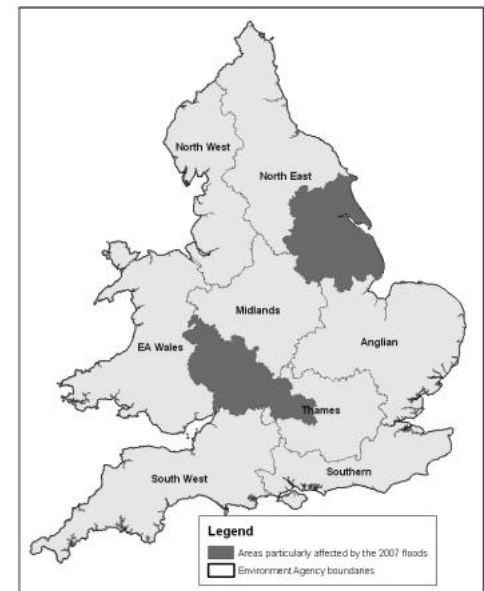
Source: Chatterton et al. (2010)

Figure 5 Visualization of inequalities between flood risks and flood awareness in 2007–8



Notes: MC=middle class; WC=working class.

Figure 1 Environment Agency regions and the impact of the 2007 floods



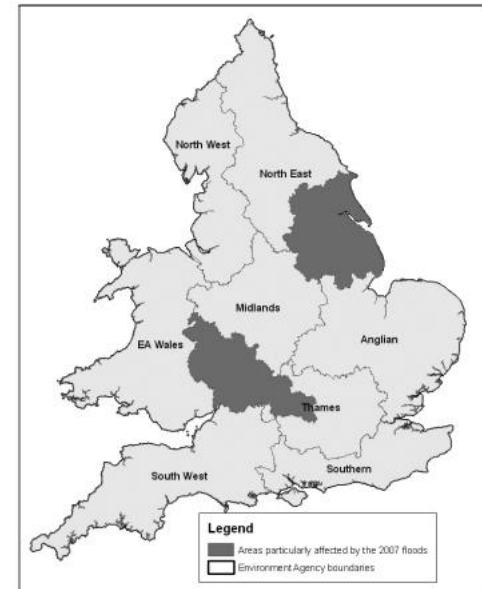
Source: Chatterton et al. (2010)

Inequalities in exposure and awareness of flood risk in England and Wales

Resultados – percepção do risco

- Muitas diferenças entre probabilidade de inundação (*flood risk likelihood*) e a percepção de risco (*flood awarness*) na Inglaterra e País de Gales
- Estas diferenças dependem da classe social dos residentes
working class – maior probabilidade de residir em áreas de inundação
- Maiores diferenças (*inequalities*) nas regiões NE e Anglian
- Fazer distinção entre o risco ético (mapas de inundação) do êmicos (percepção dos moradores)
- Regiões como Thames – *working class* estão menos atentos riscos que os *middle-classes*...
- Discussão das políticas e trabalhos futuros...

Figure 1 Environment Agency regions and the impact of the 2007 floods





- Testaram método para aumentar a resolução de dados censitários e apresentá-los agregados em uma grade regular
- Metodologia:
 - agrega domicílios (coordenadas geográficas) -Contagem da População do Instituto Brasileiro de Geografia e Estatística (IBGE) 2007
 - desagrega por setor censitário - proporcionalidade de área.
- Estima pop residente em 114 UCs na AmzLegal Bras

D'Antona, A.O.; Bueno, M.C. Dagnino, R.S. (2013) Estimativa da população em unidades de conservação na Amazônia Legal brasileira – uma aplicação de grades regulares a partir da Contagem 2007. REBEP (30) 2, p. 401-428.



STATISTICS FINLAND, 2010- dissemina dados em grades

- Validade no uso de grades regulares

(+) prescinde dos limites político-administrativos para apresentar as informações

(-) dificuldade de determinar e corrigir erros espaciais na coleta, como os decorrentes do registro incorreto de coordenadas geográficas ou até mesmo da falha na captura das coordenadas dos domicílios → novo campo(?)

Métodos de construção:

- técnicas de desagregação (*top-down*)
 - Dados agregados em áreas e métodos estatísticos para desagregar na grade (krigagem)
 - Variáveis auxiliares para a desagregação
 - Problemas conhecidos: não preserva volume e não representa espaços sem população
 - Outras técnicas – referências conhecidas

Métodos de construção:

- agregação (*bottom-up*)
 - quantificação dos dados discretos referentes à localização das residências, em células de uma grade regular que abrange a região de estudo
 - tem resultados superiores a qualquer método estatístico utilizado para desagregação dos volumes populacionais (óbvio, não?)

Brasil, agregação viável a partir de 2007 - uso sistemático e integrado de geotecnologias - espacialização do Cadastro de Endereços para Fins Estatísticos e seu relacionamento com os Mapas Censitários nas operações de 2010.



Contagem e Censo Agropecuário (2007) e Censo 2010 →
coordenadas geográficas das unidades situadas nas áreas rurais, sendo as mesmas complementares aos endereços, já que nessas áreas não existe um endereçamento formal como nas áreas urbanas (IBGE, 2009).

As coordenadas dos domicílios visitados é a base para a técnica de agregação em grades regulares, possibilitando a conversão de objetos discretos – os **pontos** representando os domicílios – em pequenas unidades de **áreas, as células**, que agregam as informações individuais.

METODOLOGIA

1) Agregação:

- da localização dos domicílios visitados na Contagem 2007 + Cadastro de Endereços para Fins Estatísticos (áreas exclusivamente rurais), exclusive as áreas de aglomerado rural– e relativos a domicílios particulares.

2) Desagregação:

- Malha vetorial de setores censitários - aglomerados rurais (povoados e vilas) e urbanos, associando-se os dados demográficos dos setores com as células da grade por meio de equivalência de área.
- Células da grade = 1min e 15 seg ou $\sim 5,3 \text{ km}^2$.
- A Amazônia Legal = 953.383 células. (testes)

- Operação espacial – atribuiu ao ponto a localização associada a grade (Junção espacial da grade vetorial e pontos)
- A estimativa do volume da população por domicílio = média de moradores no setor censitário onde o mesmo se localiza.

$$\begin{array}{l} \text{Média de} \\ \text{moradores por} \\ \text{domicílio em} \\ \text{cada setor} \\ \text{censitário} \end{array} = \frac{\text{população no setor censitário, residente em} \\ \text{domicílios permanentes ou improvisados}}{\text{número de domicílios, permanentes ou} \\ \text{improvisados, existentes no setor.}}$$

- Posteriormente, esse valor foi agregado para cada célula.



- Subnumeração relativa ao registro das coordenadas geográficas da Contagem 2007.
 - Não é erro de cobertura censitária, mas perda de dados quando da realização da espacialização.
 - O total de domicílios na grade regular (877.771) foi inferior ao Total do universo da pesquisa (1.113.258).
 - A “subnumeração espacial” - o erro não se distribui de modo uniforme no espaço - varia entre os setores censitário
 - Precisaria de campo para corrigir....

Então....



- nas células que **apresentam população**, os dados foram ajustados aplicando-se o percentual de erro calculado para o setor onde a mesma se encontra;
- nos casos em que a célula localizava-se **em mais de um setor censitário**, foi considerado o erro do setor com maior área dentro da célula;
- nas células onde a metodologia apontou **ausência de população**, o erro relativo à falta de domicílios com coordenadas não foi corrigido, assumindo-se que a população não localizada por meio de coordenadas da Contagem 2007 esteja totalmente distribuída por áreas onde a presença de população foi detectada.



Nas células de aglomerados rurais e urbano (polígonos)

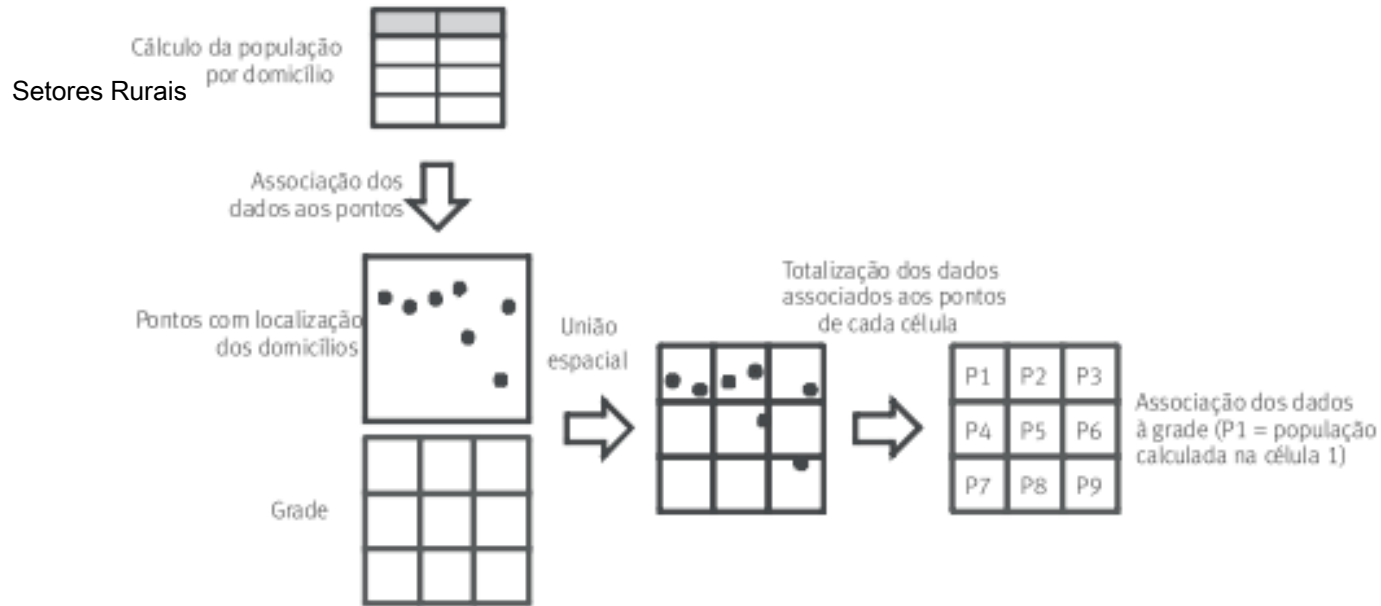
- metodologia de proporcionalidade de área,
 - supõe que a população esteja distribuída homogeneamente dentro do setor censitário.

Assim, a parcela de população equivalente à área do setor censitário que se encontra dentro de uma célula da grade é agregada a esta célula.

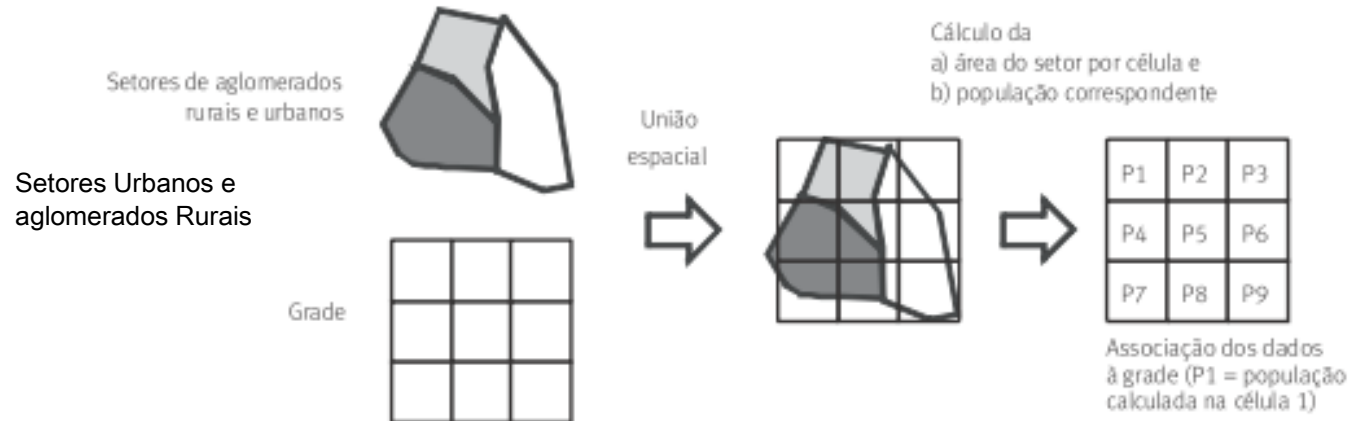
* pressuposto de homogeneidade

D'Antona, A.O.; Bueno, M.C. Dagnino, R.S. (2013) Estimativa da população em unidades de conservação na Amazônia Legal brasileira – uma aplicação de grades regulares a partir da Contagem 2007. REBEP (30) 2, p. 401-428.

A – Geração de dados por agregação (*bottom-up*)



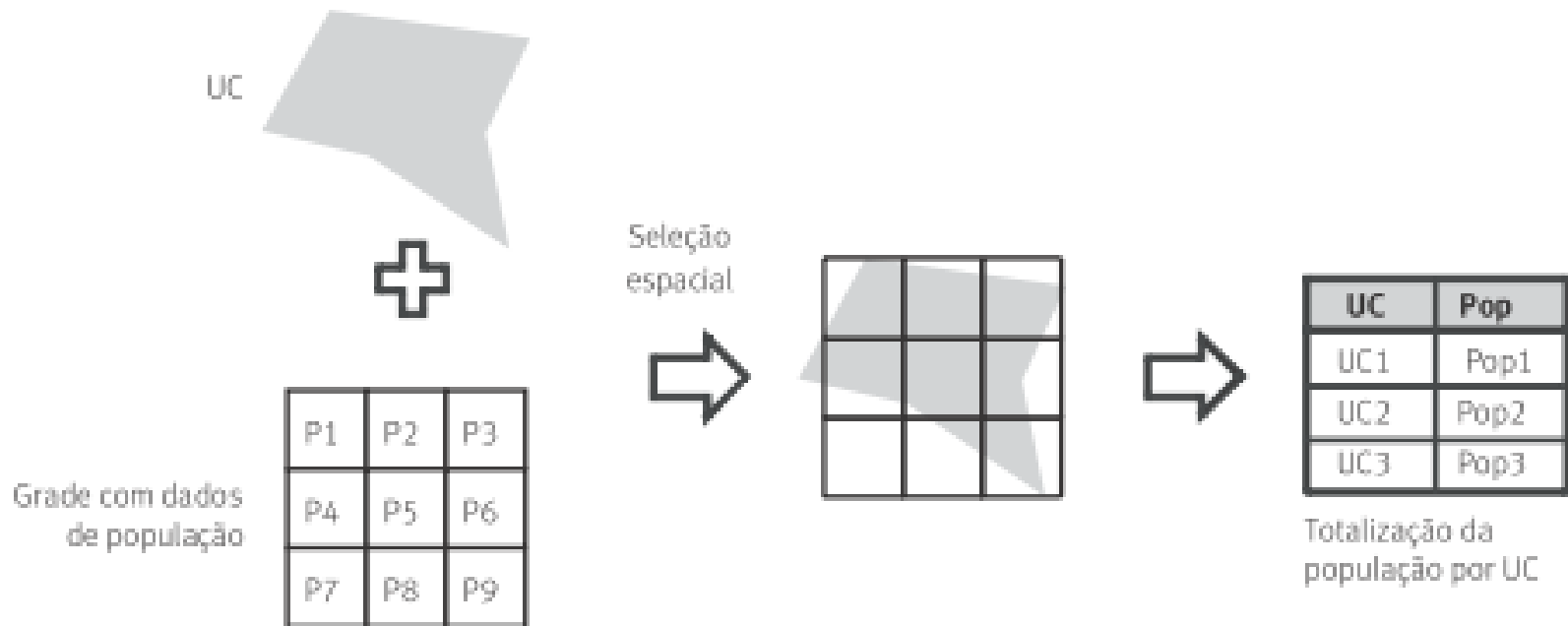
B – Geração de dados por desagregação (*top-down*)



Sobre a grade consolidada – limites dos polígonos das Ucs

- 121 Ucs federais (43 UCPI e 70 UCUS)
- Buffer 10km para computar pop no entorno tb;

C – Estimativa da população por Unidade de Conservação (UC)



D'Antona, A.O.; Bueno, M.C. Dagnino, R.S. (2013) Estimativa da população em unidades de conservação na Amazônia Legal brasileira – uma aplicação de grades regulares a partir da Contagem 2007. REBEP (30) 2, p. 401-428.

TABELA 1
População estimada, segundo tipos e categorias de unidades de conservação federais
Amazônia Legal brasileira – 2007

Tipos e categorias de UCs	Unidades de conservação (UC)						Entorno (1)			
	Qtde.	Área (km ²)	População (hab.) (2)				População (hab.) (2)			
			Rural (3)	Urbana (4)	Total (5)	Total (ajustada) (6)	Rural (3)	Urbana (4)	Total	Total (ajustada) (6)
Proteção integral	40	277.173	23.593	1.283	24.877	27.705	42.353	50.867	93.220	96.738
Estações Ecológicas	11	55.908	679	0	679	905	6.866	14.425	21.291	21.771
Parques Nacionais	21	185.819	15.937	1.085	17.022	19.228	26.770	35.927	62.697	65.082
Reservas Biológicas	8	35.447	6.978	198	7.176	7.572	8.717	515	9.232	9.886
Uso sustentável	74	261.239	119.347	164.293	283.640	297.693	237.791	657.148	894.938	923.499
Áreas de Proteção Ambiental	2	20.800	2.766	615	3.381	4.013	8.257	33.059	41.316	41.701
Áreas de Relevante Interesse Ecológico	3	200	213	0	213	296	4.861	9.568	14.429	14.629
Florestas Nacionais	29	146.429	28.905	34.163	63.068	65.196	48.673	120.035	168.708	174.699
Reservas de Desenvolvimento Sustentável	1	673	1.375	0	1.375	1.499	1.966	47	2.013	2.506
Reservas Extrativistas	39	93.137	86.088	129.515	215.603	226.689	174.033	494.439	668.472	689.964
Total	114	538.412	142.940	165.577	308.517	325.398	280.144	708.015	988.158	1.020.237

Fonte: IBGE. Microdados da Contagem Populacional de 2007.

(1) Definido pela delimitação de uma faixa de 10 km de largura ao redor de cada uma das UCs consideradas.

(2) Estimativa a partir da Contagem Populacional 2007.

(3) População em domicílios com localização registrada em setores censitários rurais, incluindo os aglomerados rurais.

(4) População em domicílios com localização registrada em setores censitários urbanos.

(5) População total (rural + urbana). Sujeito à subenumeração, conforme o setor censitário.

(6) População total ajustada para minimizar a subenumeração.

D'Antona, A.O.; Bueno, M.C. Dagnino, R.S. (2013) Estimativa da população em unidades de conservação na Amazônia Legal brasileira – uma aplicação de grades regulares a partir da Contagem 2007. REBEP (30) 2, p. 401-428.

(continuação)

Unidades de conservação	Unidade de conservação							Entorno (1)			
	Área (km ²)	Células		População (hab.) (2)				População (hab.) (2)			
		Total	Ocup. (%) (3)	Rural (4)	Urbana (5)	Total (6)	Total (ajustada) (7)	Rural (4)	Urbana (5)	Total (6)	Total (ajustada) (7)
Mata Grande	125	40	80,00	1.315	15.317	16.632	16.682	8.565	215.639	224.204	224.481
Médio Juruá	2.536	572	5,59	1.699	0	1.699	1.713	1.043	0	1.043	1.072
Quilombo Flexal	87	24	45,83	679	1.483	2.162	2.174	6.526	7.212	13.738	14.267
Rio Cajari	5.201	1.062	10,36	2.430	0	2.430	2.824	1.344	0	1.344	1.490
Rio do Cautário	743	151	8,61	111	0	111	114	134	0	134	138
Rio Iriri	3.992	897	0,67	29	0	29	45	27	0	27	30
Rio Jutai	2.741	602	5,81	1.537	0	1.537	1.548	241	0	241	245
Rio Ouro Preto	2.037	479	12,53	486	0	486	490	436	0	436	461
Rio Unini	8.416	1.863	2,79	1.126	0	1.126	1.172	1.102	0	1.102	1.113
Riozinho da Liberdade	3.492	767	3,26	1.112	0	1.112	1.396	1.161	0	1.161	1.405
Riozinho do Anfrísio	7.351	1.511	1,26	187	0	187	250	505	0	505	565
São João da Ponta	32	30	53,33	1.294	2.293	3.587	3.592	12.874	18.543	31.417	32.820
Terra Grande-Pracuúba	1.949	434	5,99	1.045	0	1.045	1.400	3.600	9.874	13.475	14.315
Verde Para Sempre	12.860	2.548	11,07	6.086	0	6.086	7.391	4.407	26.447	30.854	31.368
Total	538.412	113.447	3,96	142.940	165.577	308.517	325.398	280.144	708.015	988.158	1.020.237

Fonte: IBGE. Microdados da Contagem Populacional de 2007.

(1) Definido pela delimitação de uma faixa de 10km de largura ao redor de cada uma das UCs consideradas.

(2) Estimativa a partir da Contagem Populacional 2007.

(3) Percentual de células com população em relação ao total de células no interior da UC.

(4) População em domicílios com localização registrada em setores censitários rurais, incluindo os aglomerados rurais.

(5) População em domicílios com localização registrada em setores censitários urbanos.

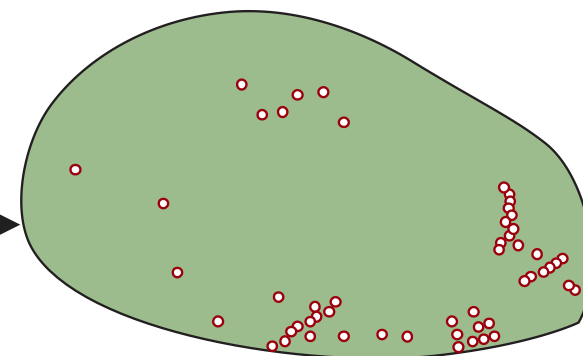
(6) População total (rural + urbana). Sujeito à subenumeração, conforme o setor censitário.

(7) População total ajustada para minimizar a subenumeração.

GEOCODIFICAÇÃO DO CADASTRO NACIONAL DE ENDEREÇOS PARA FINS ESTATÍSTICOS (CNEFE)



row.names	Codigo.do.setor	nome.mun	end
1	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
2	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
3	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
4	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
5	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
6	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
7	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
8	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
9	350315805000005	ARAPEÍ	ESTRADA SP 68, 250, ARAPEÍ - SP, 12870-000, Brazil
10	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
11	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
12	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
13	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
14	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
15	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
16	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
17	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
18	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
19	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
20	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
21	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
22	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
23	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
24	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
25	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
26	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
27	350315805000005	ARAPEÍ	ESTRADA JOSE ANTONIO DE PAULA MACHADO, 0, ARAPEÍ - S
28	350315805000005	ARAPEÍ	ESTRADA CAXAMBU, 0, ARAPEÍ - SP, 12870-000, Brazil
29	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
30	350315805000005	ARAPEÍ	ESTRADA SP 68, 0, ARAPEÍ - SP, 12870-000, Brazil
31	350350505000001	AREIAS	RUA RIACHUELO, 250, AREIAS - SP, 12870-000, Brazil
32	350350505000001	AREIAS	RUA RIACHUELO, 286, AREIAS - SP, 12870-000, Brazil

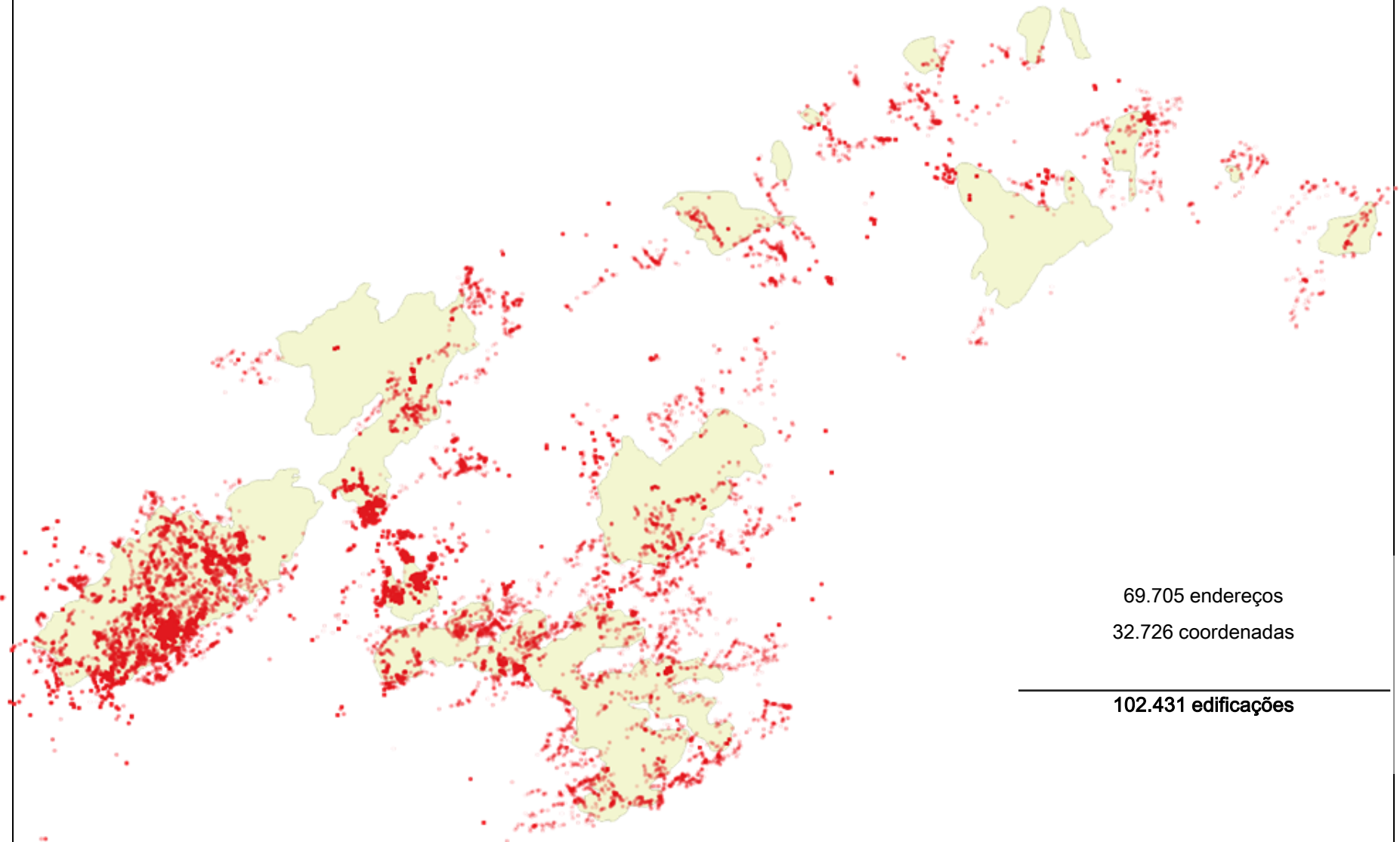


69.705 endereços

32.726 coordenadas

102.431 edificações

GEOCODIFICAÇÃO DO CADASTRO NACIONAL DE ENDEREÇOS PARA FINS ESTATÍSTICOS (CNEFE)



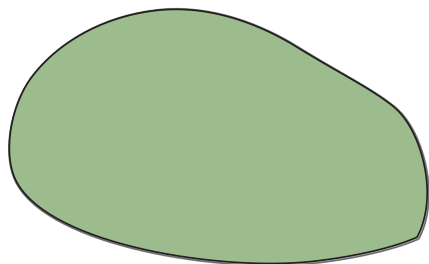
69.705 endereços
32.726 coordenadas

102.431 edificações

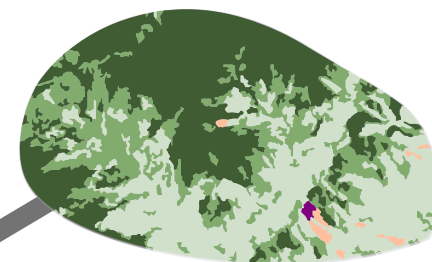
INTEGRAÇÃO DOS DADOS NO ESPAÇO CELULAR



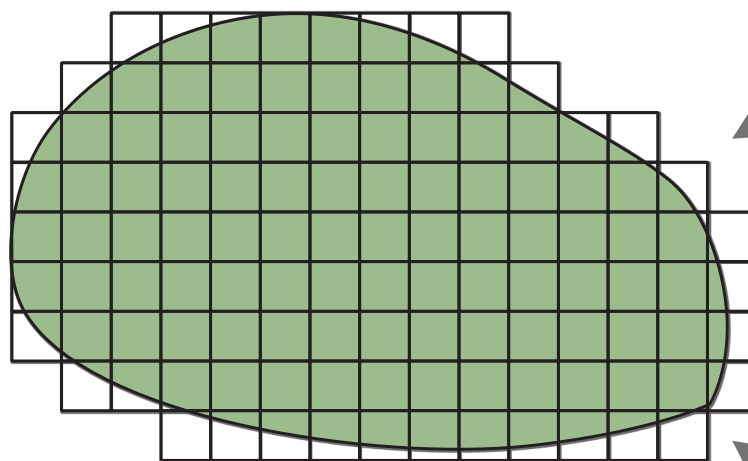
APA-MRPS



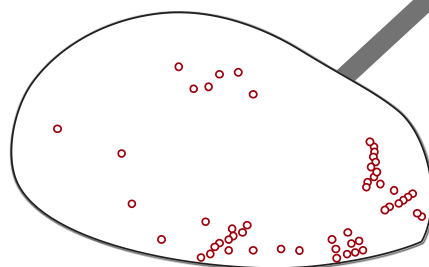
Uso e cobertura da terra



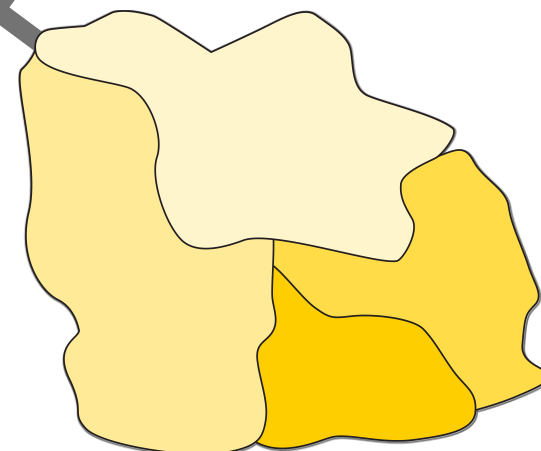
Espaço celular



CNEFE



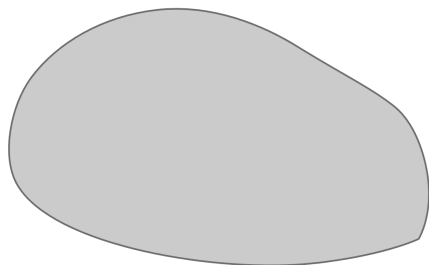
Setores censitários



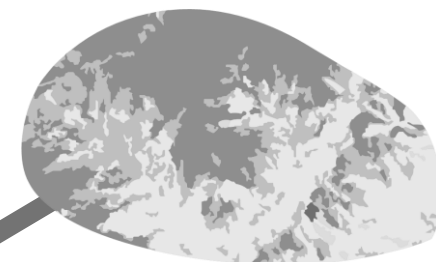
INTEGRAÇÃO DOS DADOS NO ESPAÇO CELULAR



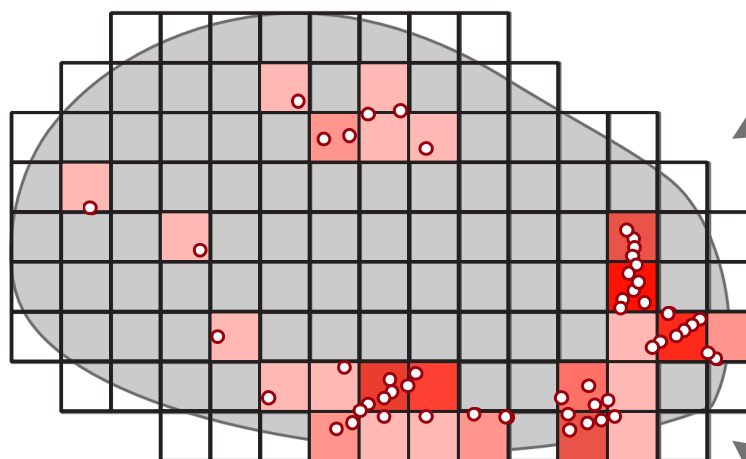
APA-MRPS



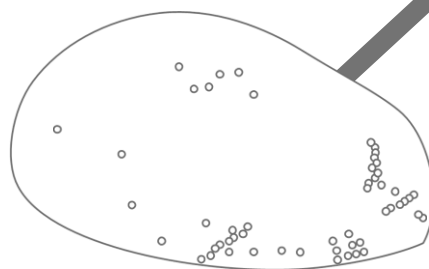
Uso e cobertura da terra



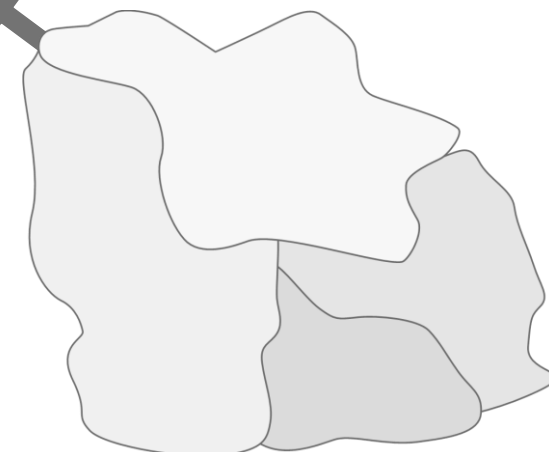
Espaço celular



CNEFE



Setores censitários





Principais conceitos teóricos

$$2\gamma(\mathbf{h}) = 2C(0) - 2C(\mathbf{h}) \text{ ou } \gamma(\mathbf{h}) = C(0) - C(\mathbf{h})$$

onde:

$\gamma(h)$ representa o **semivariograma**, que é metade do **variograma**.

A relação indica que sob a hipótese de estacionariedade de 2a ordem, que a covariância e o semivariograma são formas alternativas de caracterizar a autocorrelação dos pares $Z(x)$ e $Z(x+h)$ separados pelo vetor h .

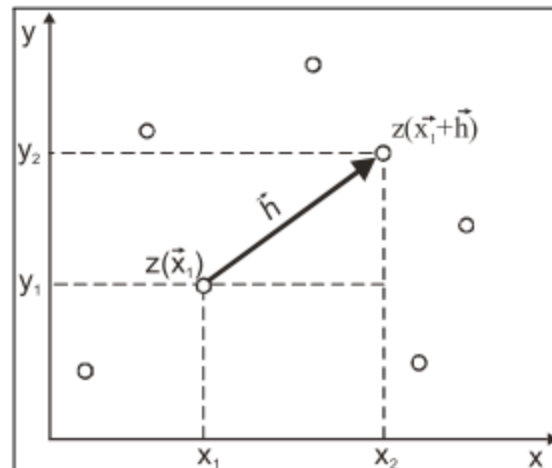


Figura 3-8 – Amostragem em duas dimensões.

Principais conceitos teóricos

relação entre as funções semivariograma e covariância

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

