

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

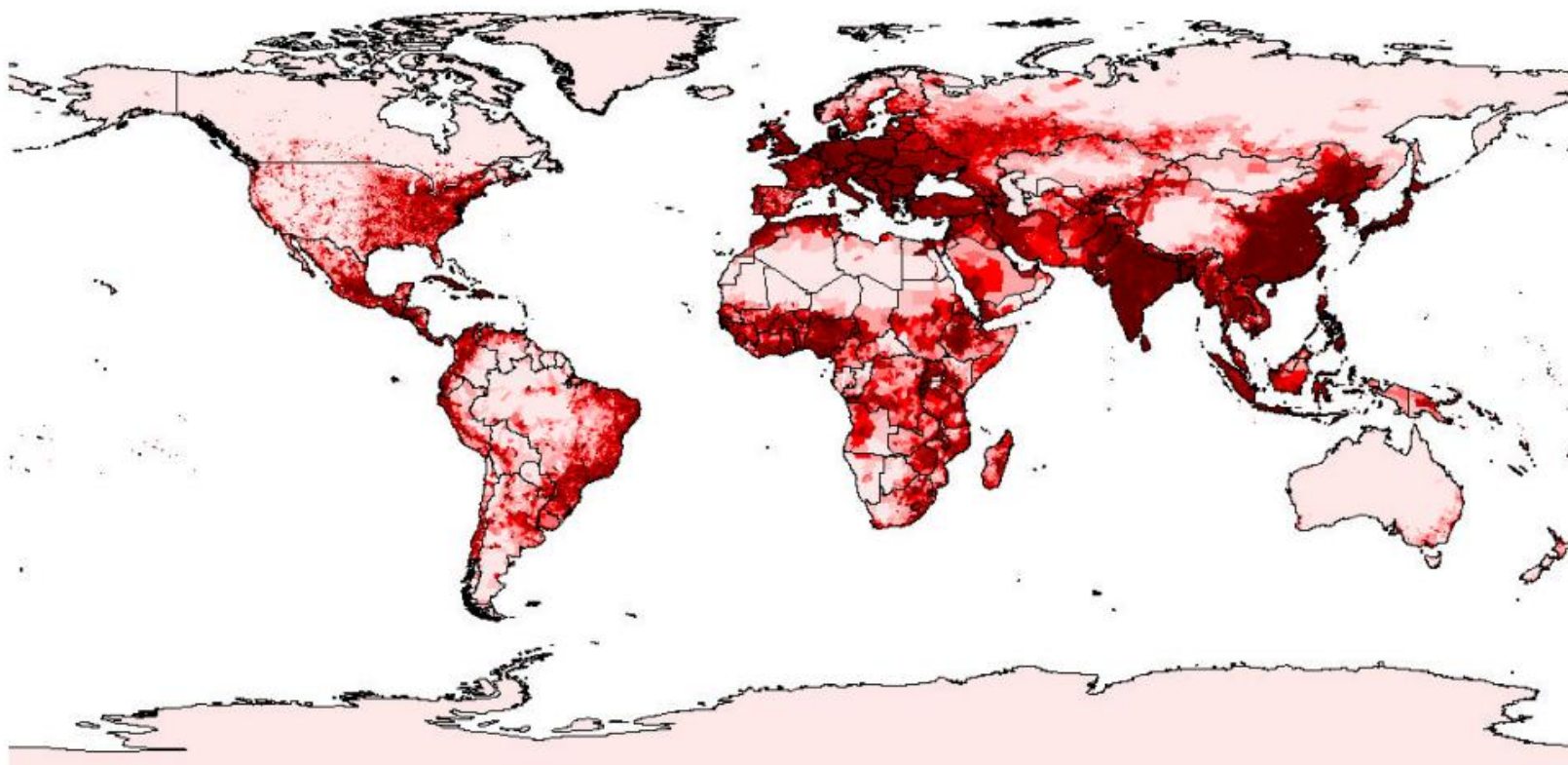
{silvana@dpi.inpe.br, miguel@dpi.inpe.br}



De Censos a Superfícies



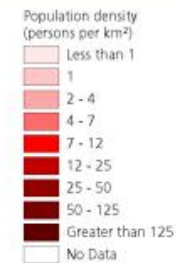
Global Population Density



Map Projection: Geographic

Citation: WRI, 2000 based on CIESIN 2000

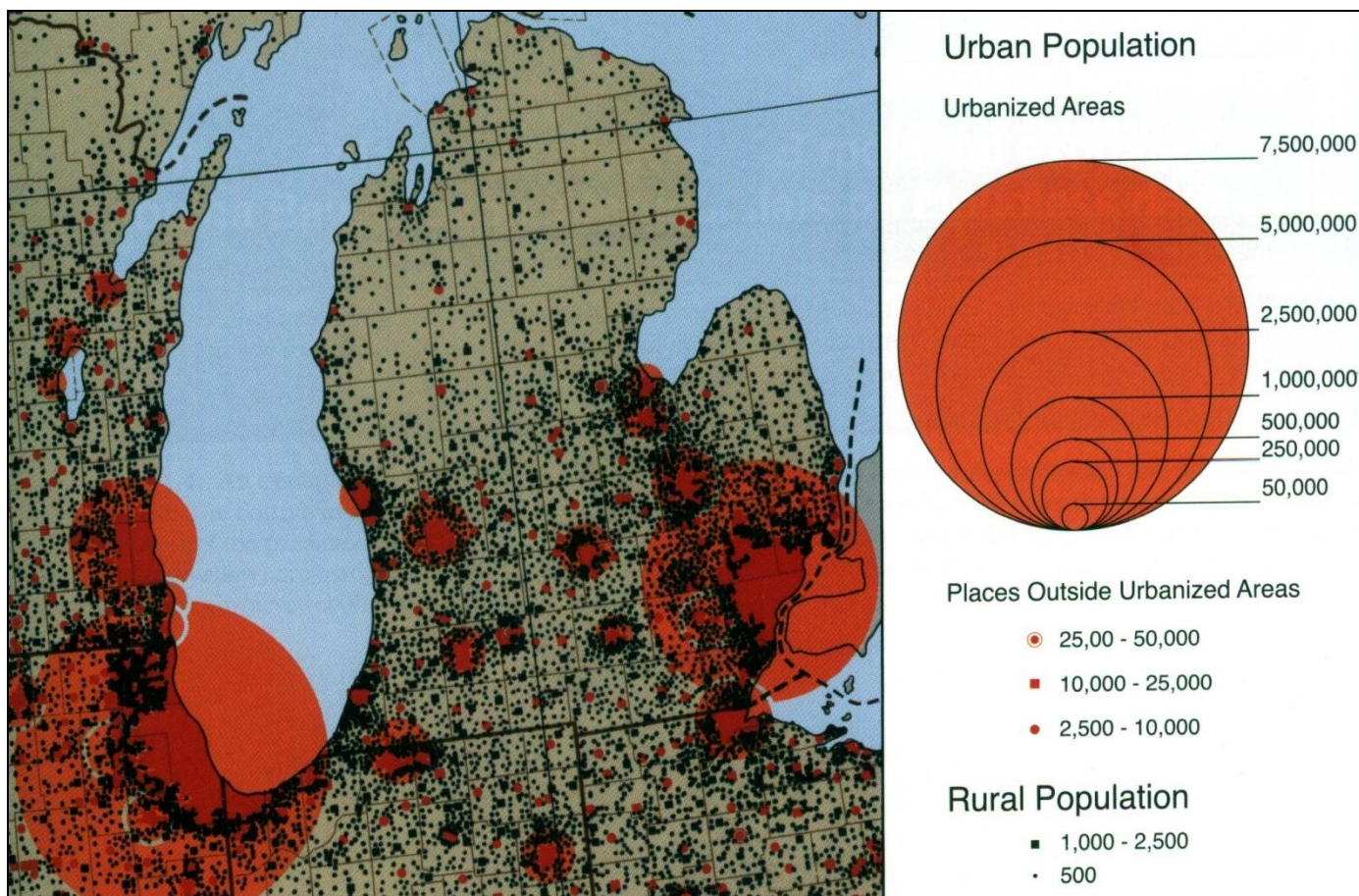
Notes:



- 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

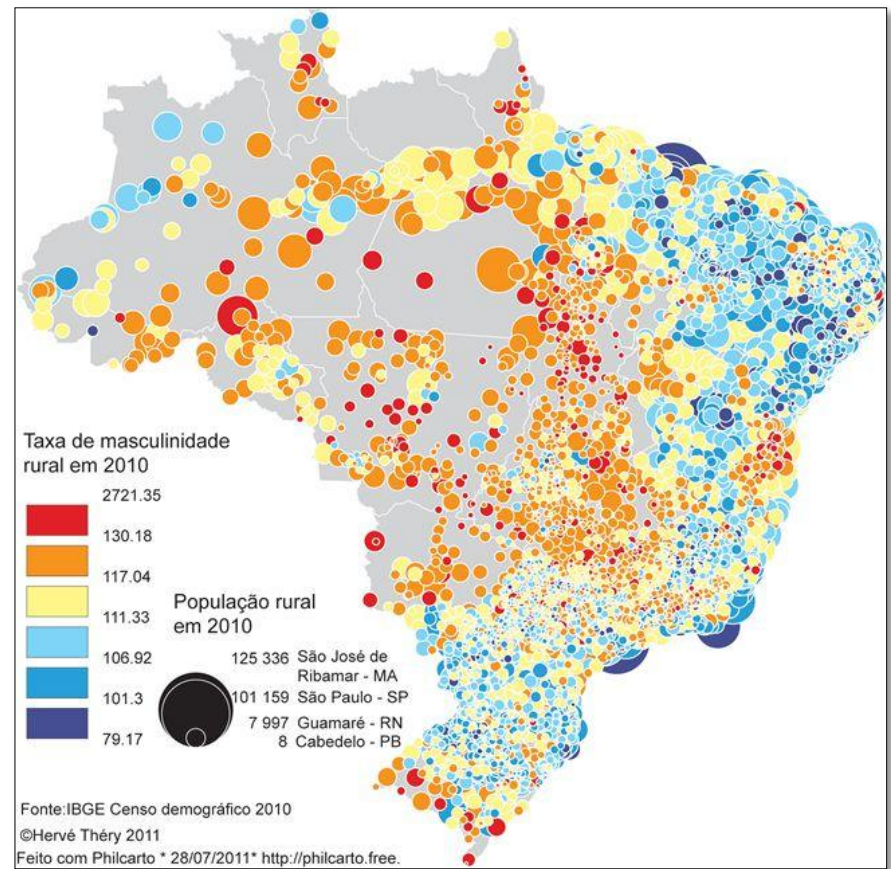
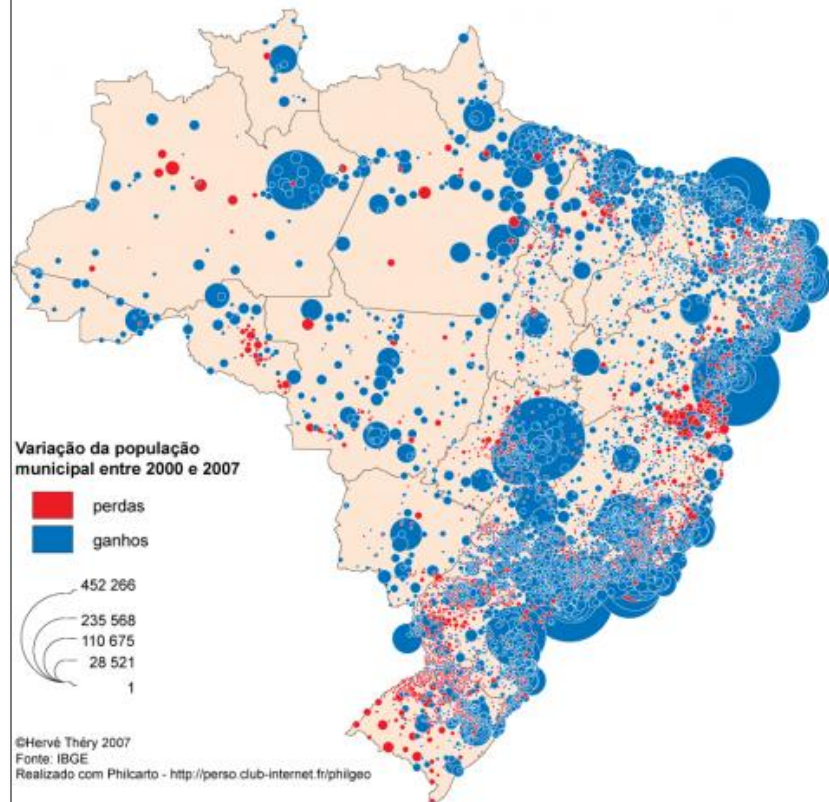


Representações Cartográficas

Mapas de símbolos pontuais proporcionais

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

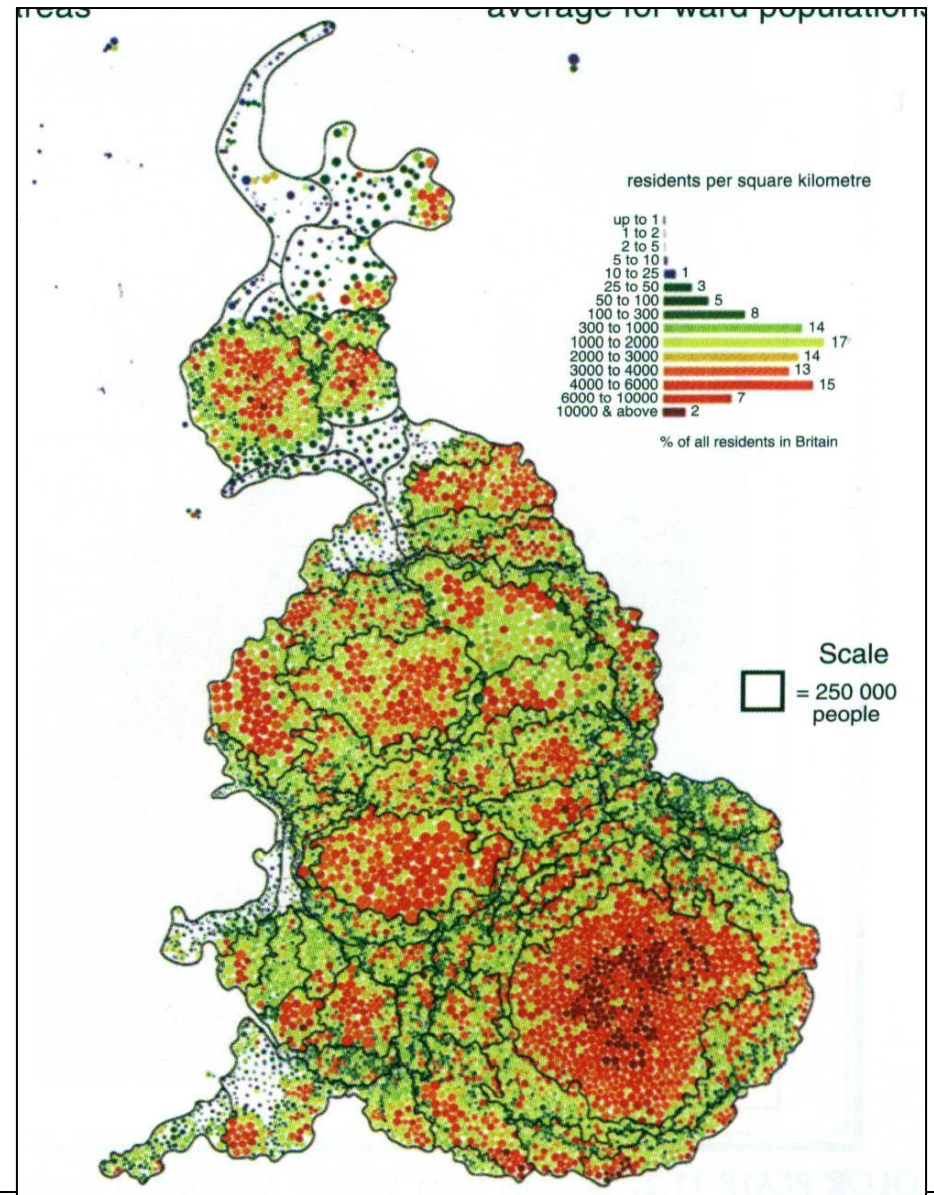
Evolução da população brasileira entre 2000 e 2007



Representações Cartográficas



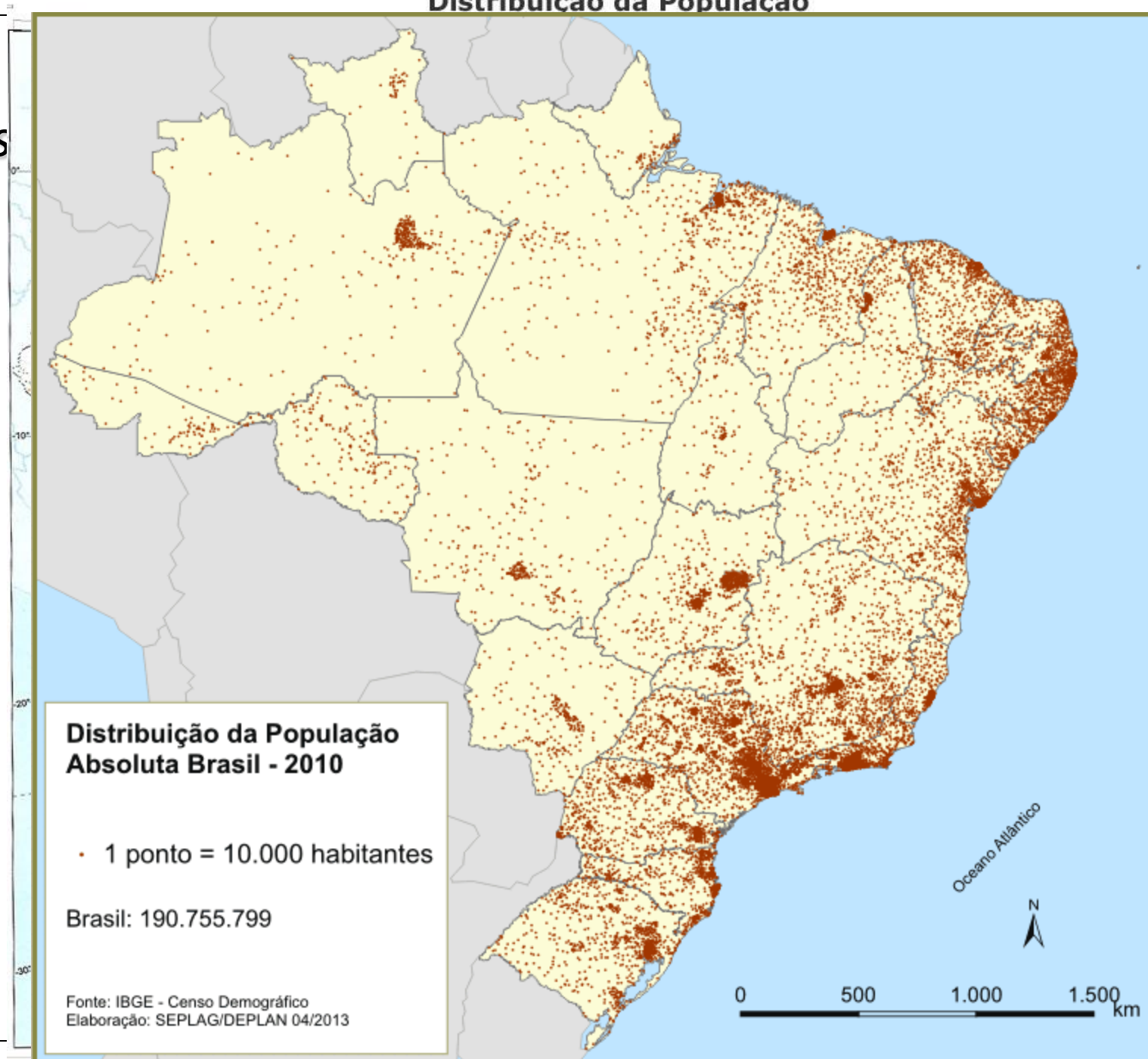
- Mapas de pontos - *dots*
- Representar fenômenos discretos com conotação pontual. Ilustra a densidade espacial;
- Objetivo: facilitar a comunicação cartográfica
- É um tipo especial de mapas de símbolos proporcionais
- Cada ponto = mesma quantidade
 - Superposição em áreas densas
 - Valores agregados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



Distribuição da População

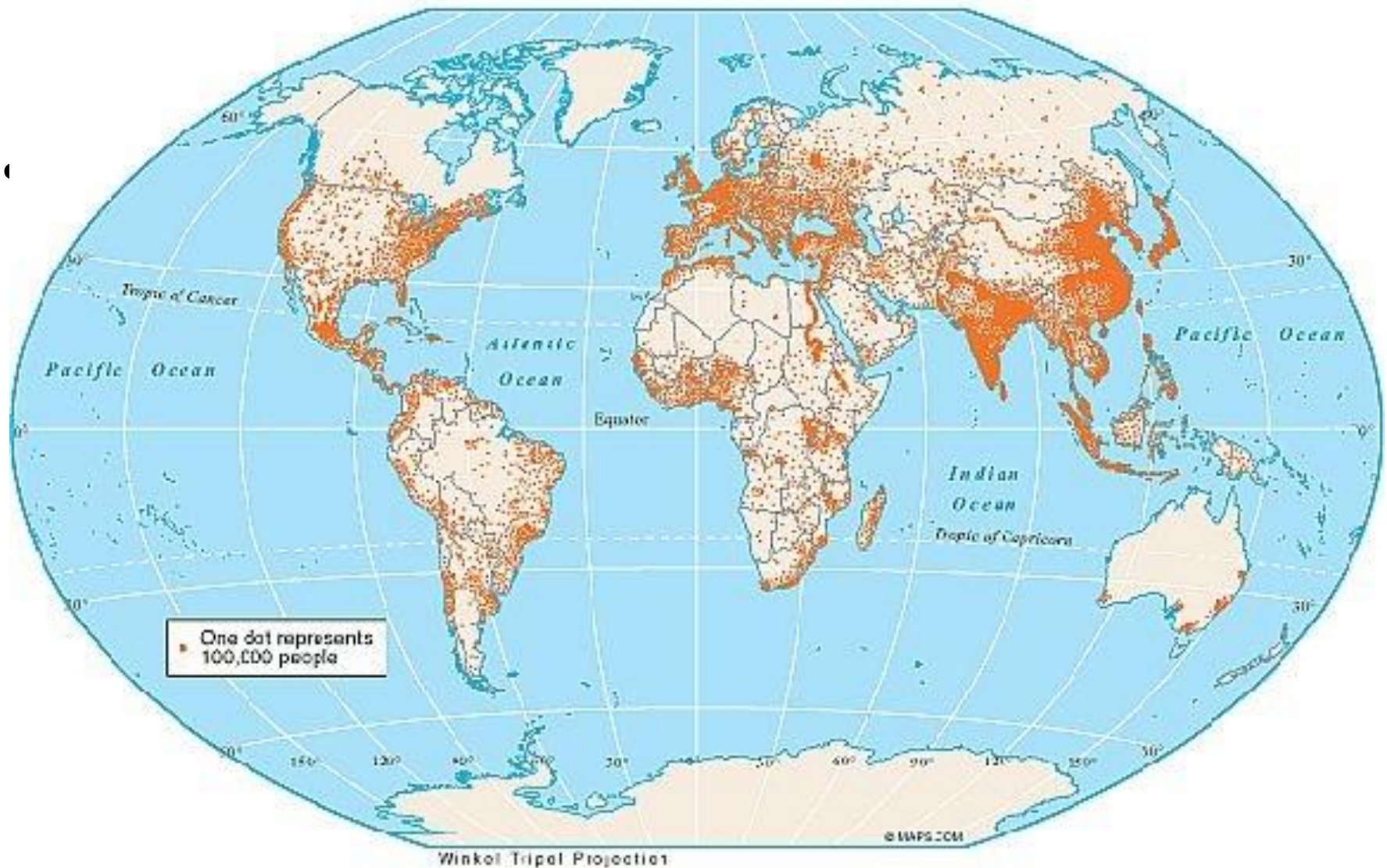
Repres

Mapa de
pontos



Representações Cartográficas

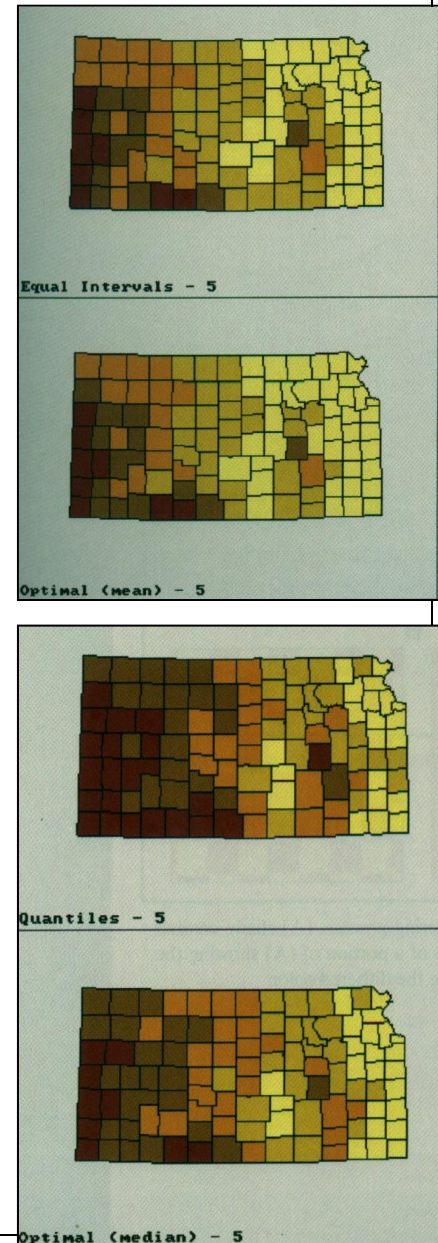
World Population Distribution, 1998



Representações Cartográficas

Mapas Coropléticos – representação cartográfica para traduzir valores para áreas

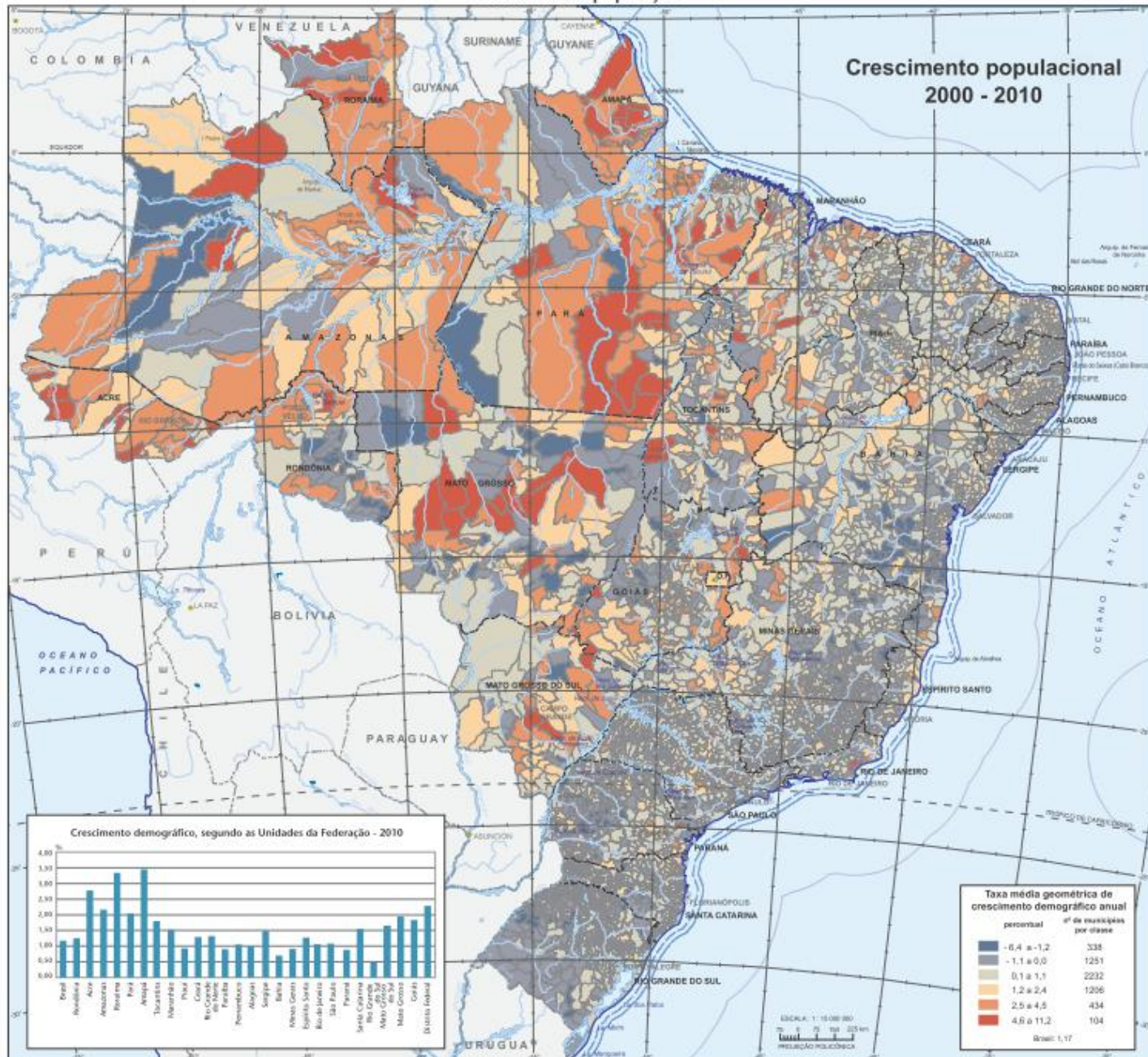
- Valores apresentados devem ser razões ou proporções. Evitar valores absolutos
- Luminosidade e saturação da cor, indicam diferenças ordenadas em classes distintas
- Dois tipos:
 - Densidade – razões, n pessoas /km²
 - Porcentagem, ex %habitantes/ total populacional
- Mapas temáticos
- Valores associados 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



Repres

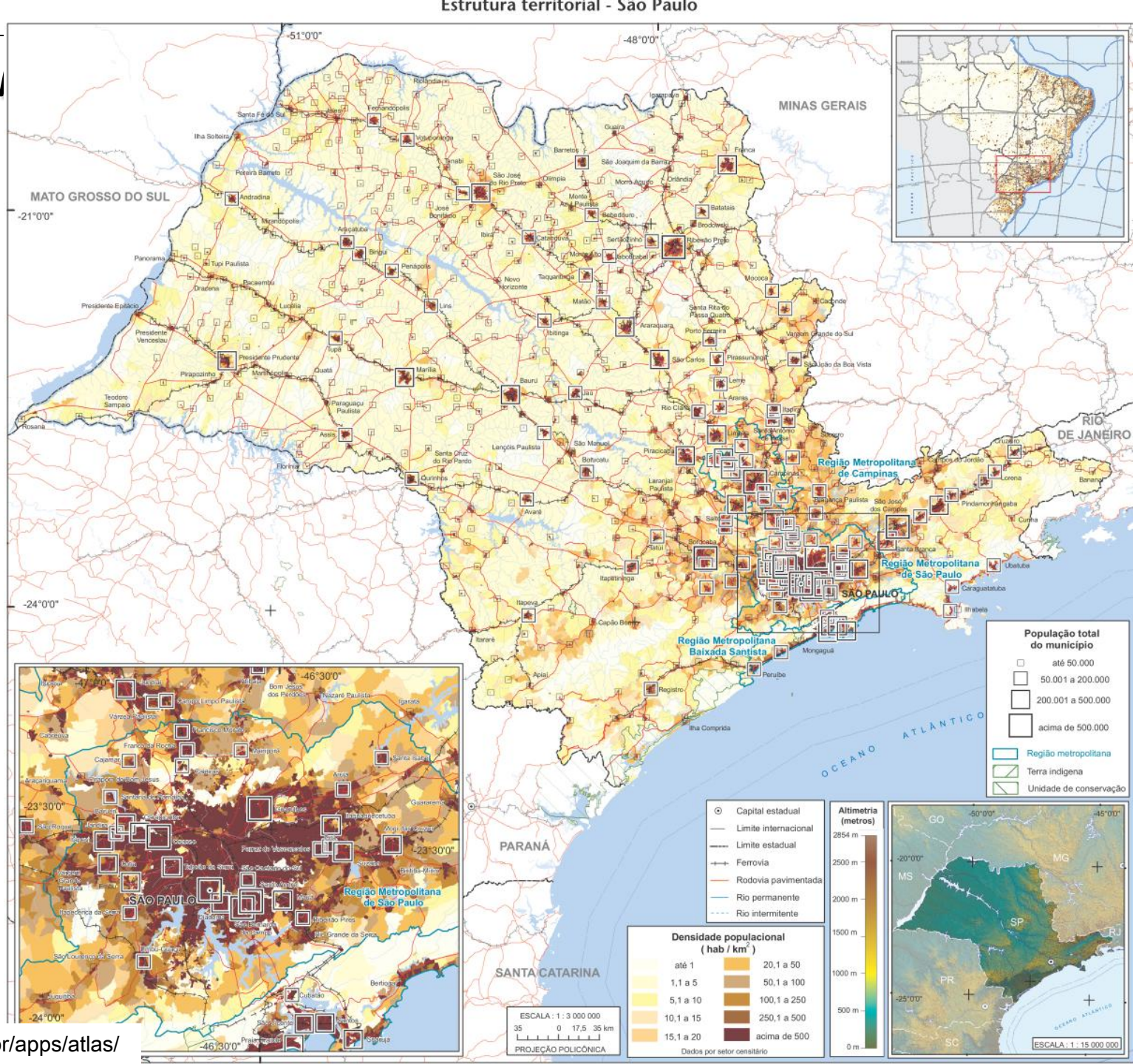
Mapa Coroplético

Dinâmica da população

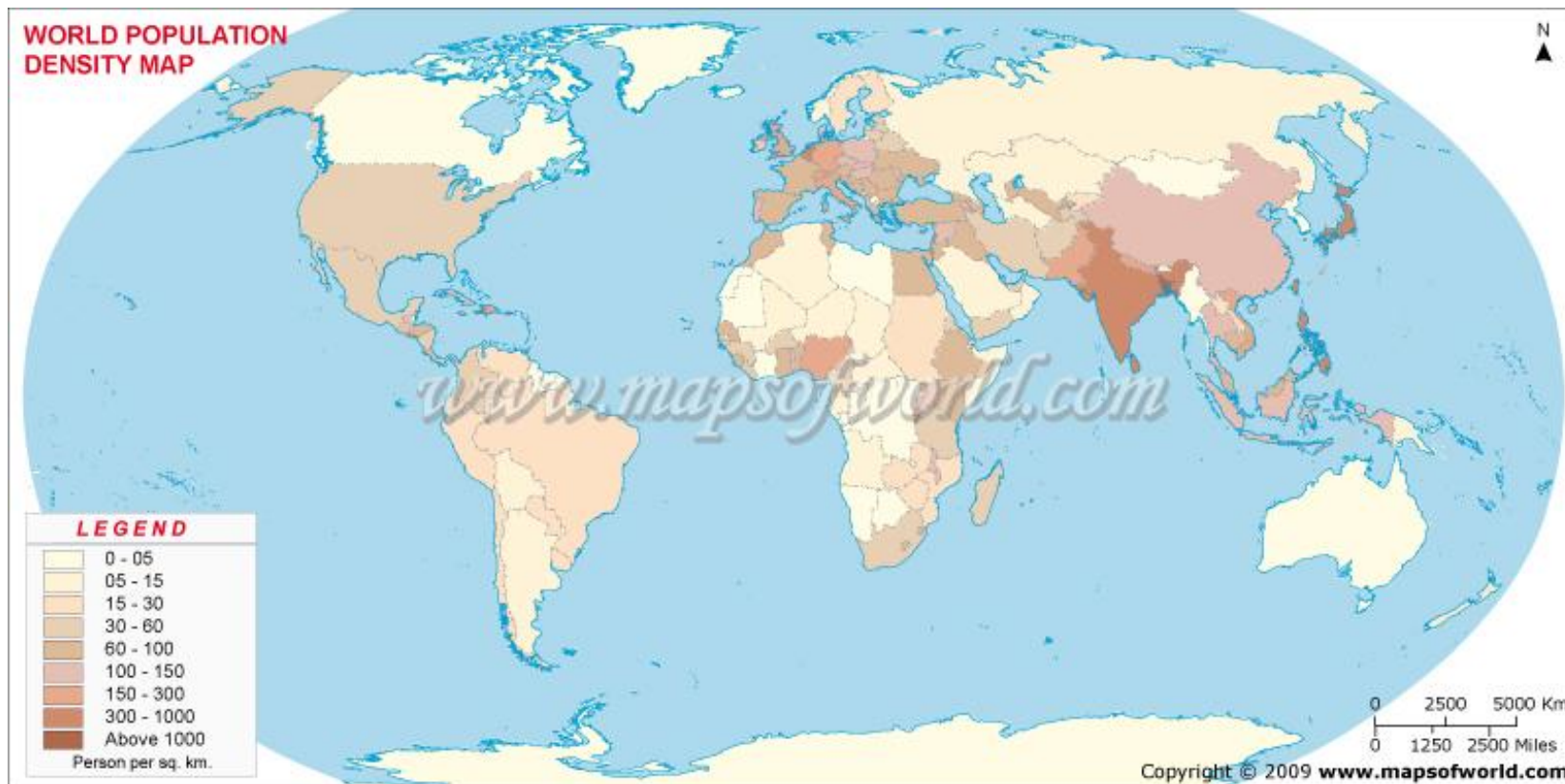
Crescimento populacional
2000 - 2010

República

Mapa Coroplético



Representações Cartográficas



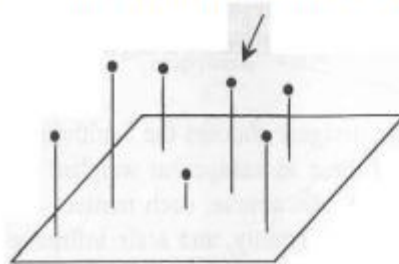
Representações Cartográficas

Um mapa isarítmico 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:

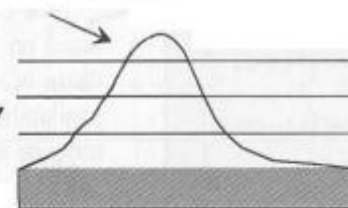
- Mapas isométricos, os dados foram coletados para localizações **pontuais** específicas, os pontos de controle, como dados de temperatura em estações meteorológicas.

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

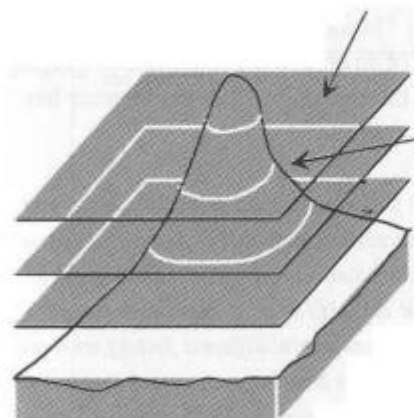
1. Representação proporcional ao valor do fenômeno.



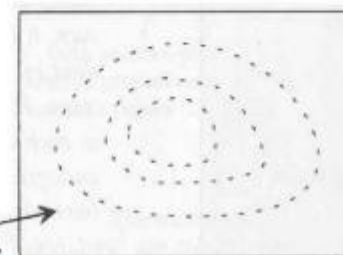
2. Superfície obtida pela interpolação dos valores.



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 isarítmico.

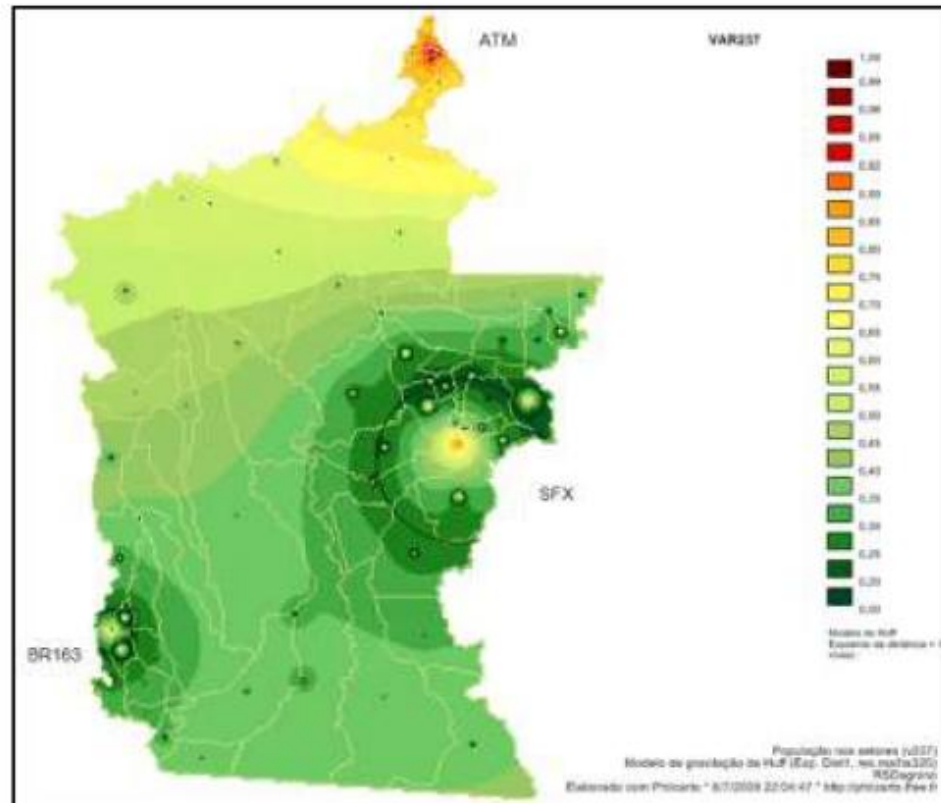
Modificado por Pantaleão (2003)

Representações Cartográficas

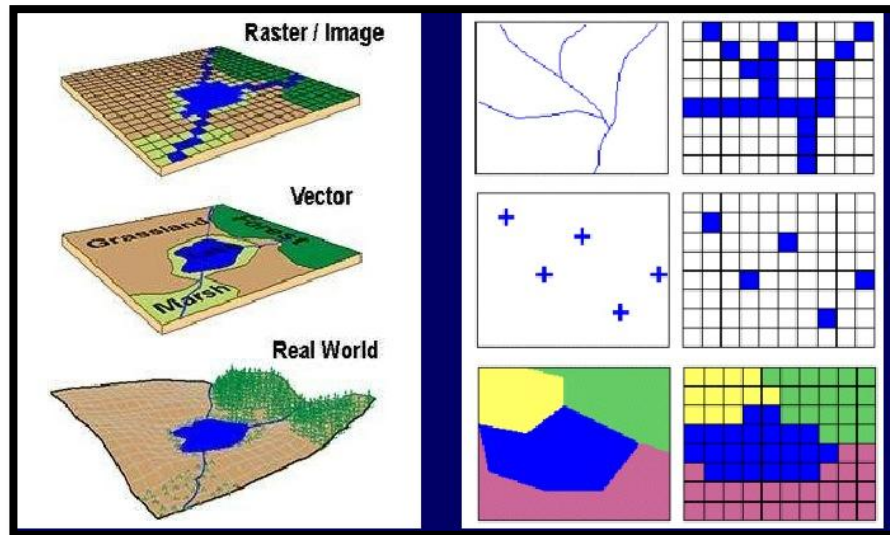
Mapa Isoplético ou de Linhas

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

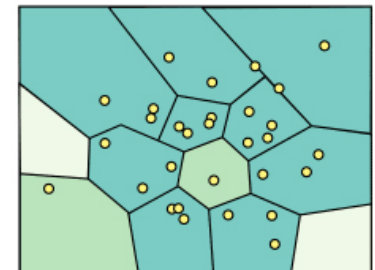
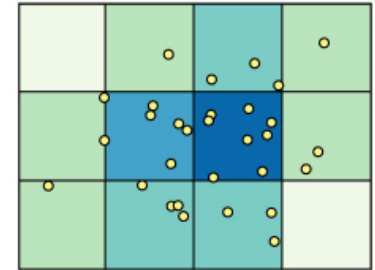


- Modelos de representação computacional para dados sócio-econômicos:
 - Indivíduo
 - Área
 - Fenômeno contínuo no espaço

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



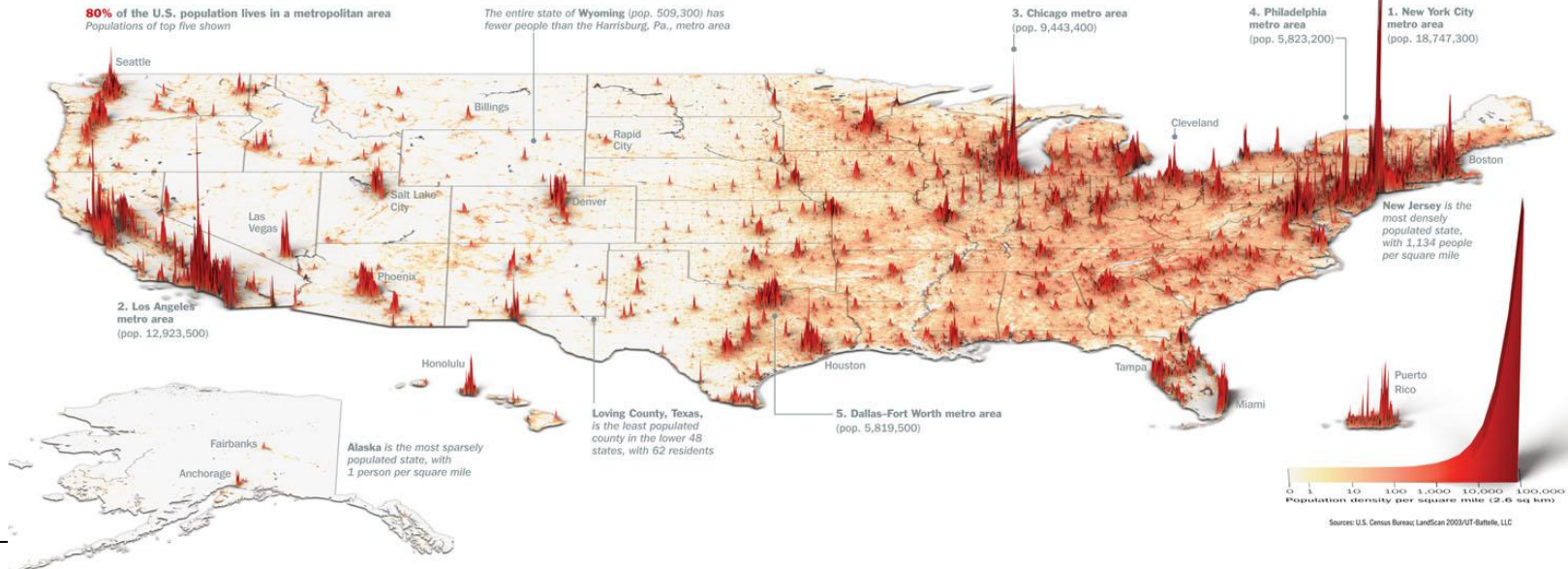
- Modelos de Superfície - alternativa a restrição das áreas
 - Densidade demográfica – fenômeno contínuo
 - Objetivo: estimar distribuição no maior detalhe possível
 - Vantagem: manipulação e análise independentes de área
 - Armazenar e disponibilizar dados em BD Globais

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-

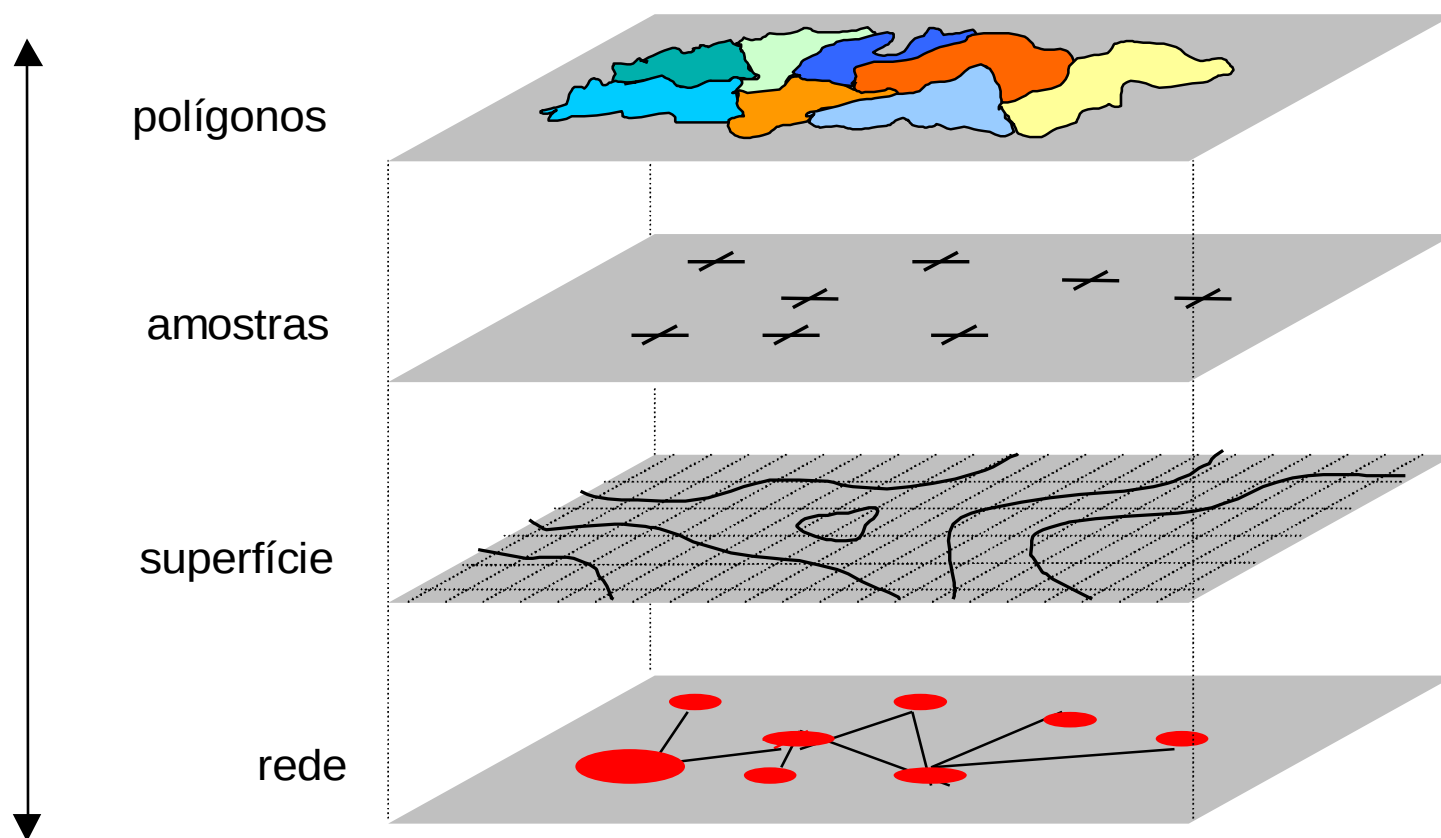
Ala.; Possum Trot, Ky.; or Lonelyville, N.Y. But they are all probably close to someone's idea of paradise. —By Nancy Gibbs



Representação Computacional



- Representação digital de dados geográficos – modelos matricial e vetorial



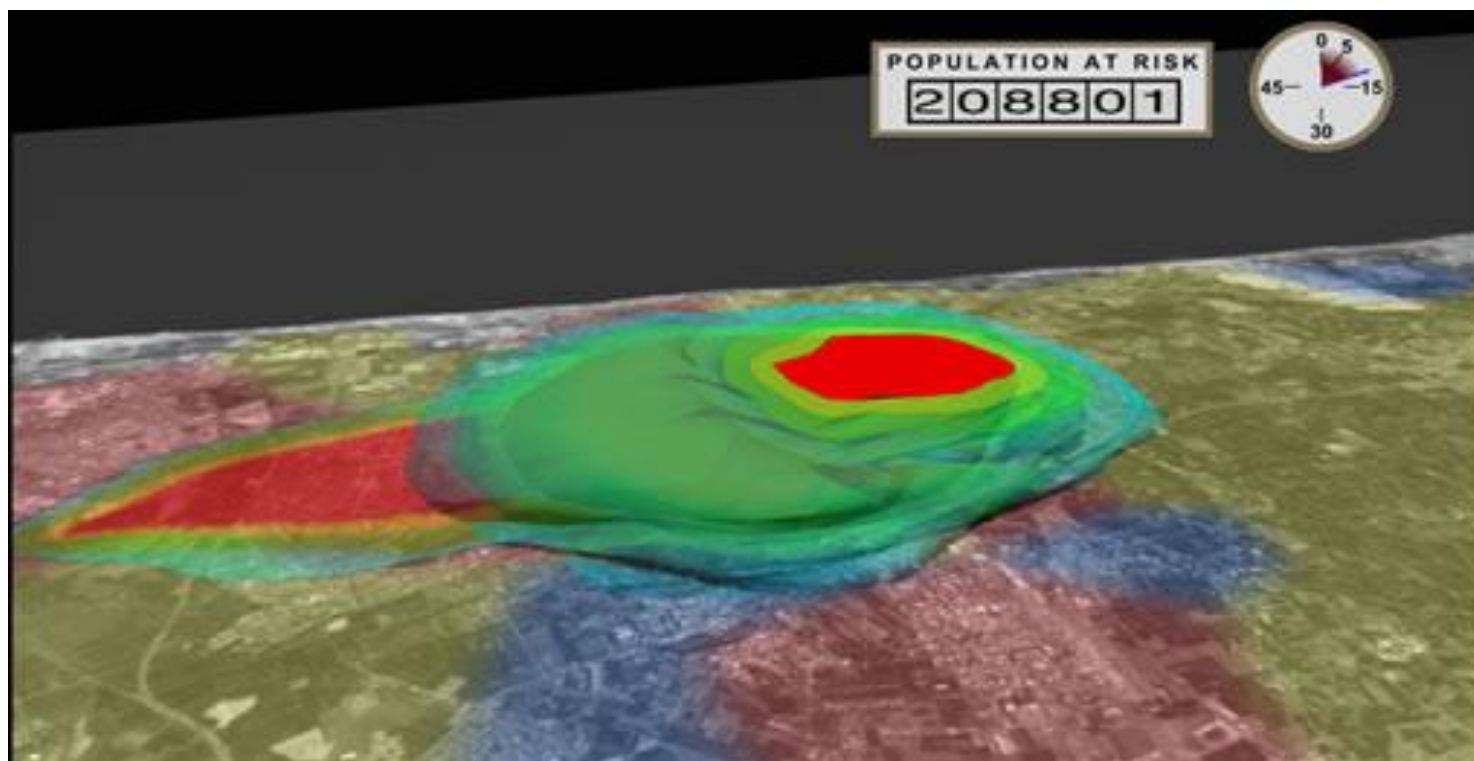
Representação Computacional



- 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

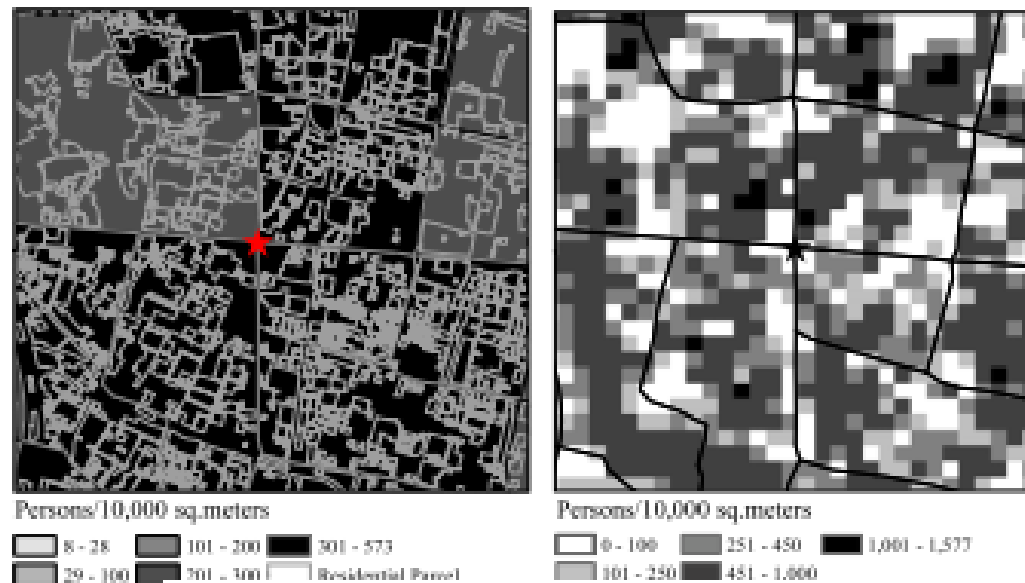


Figure 5 Comparison between vector-based subdistrict population distribution (left) and raster population surface (right) for CBD area

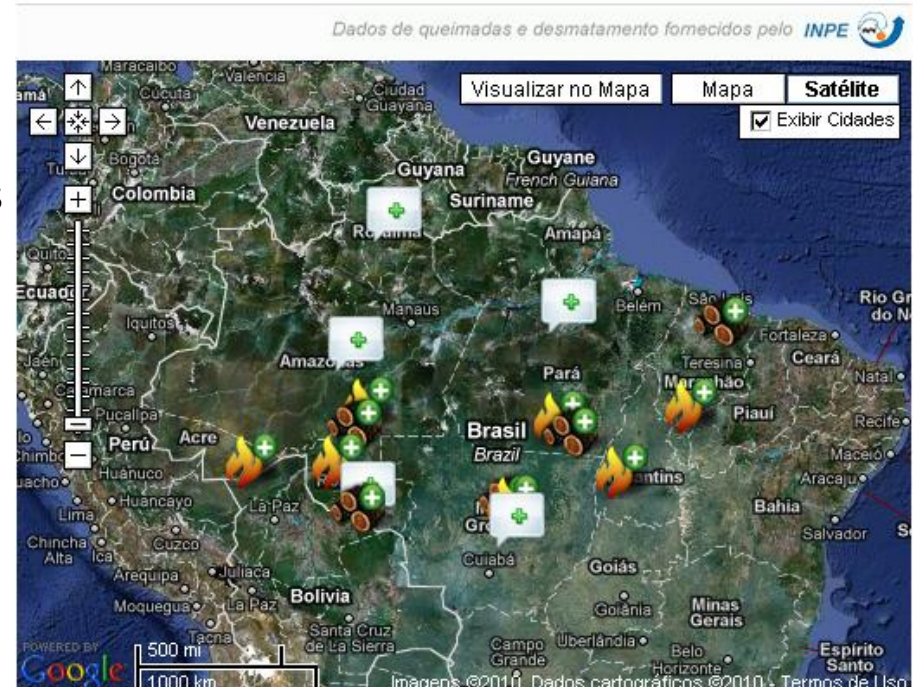
Luo (2005)

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



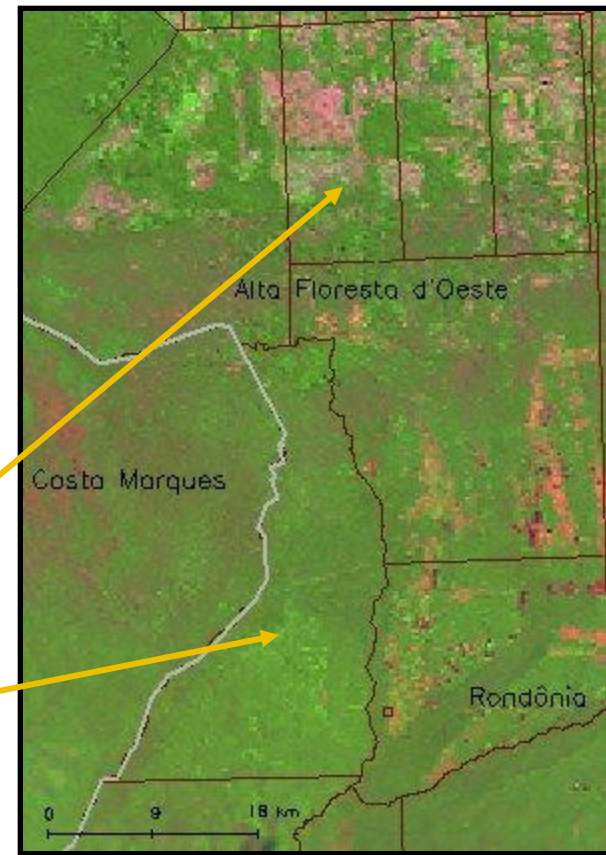
- Disponibilidade de dados
 - Dados censitários (10 anos)
 - Contagens por amostragem intercensos
 - Estimativas estatísticas – PNAD – UF, região metropolitana, no N somente para população urbana

- 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...

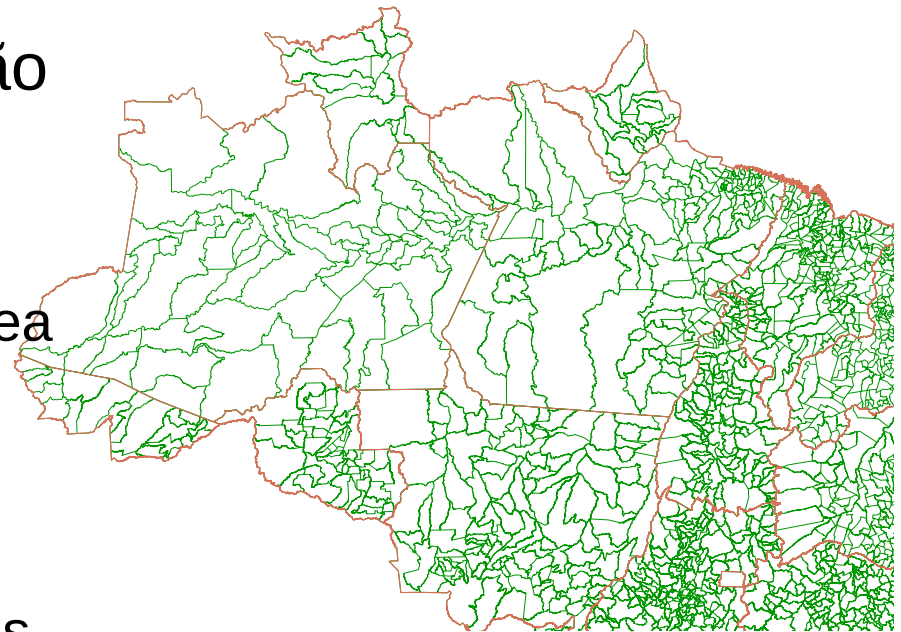
• Setor Censitário

- 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

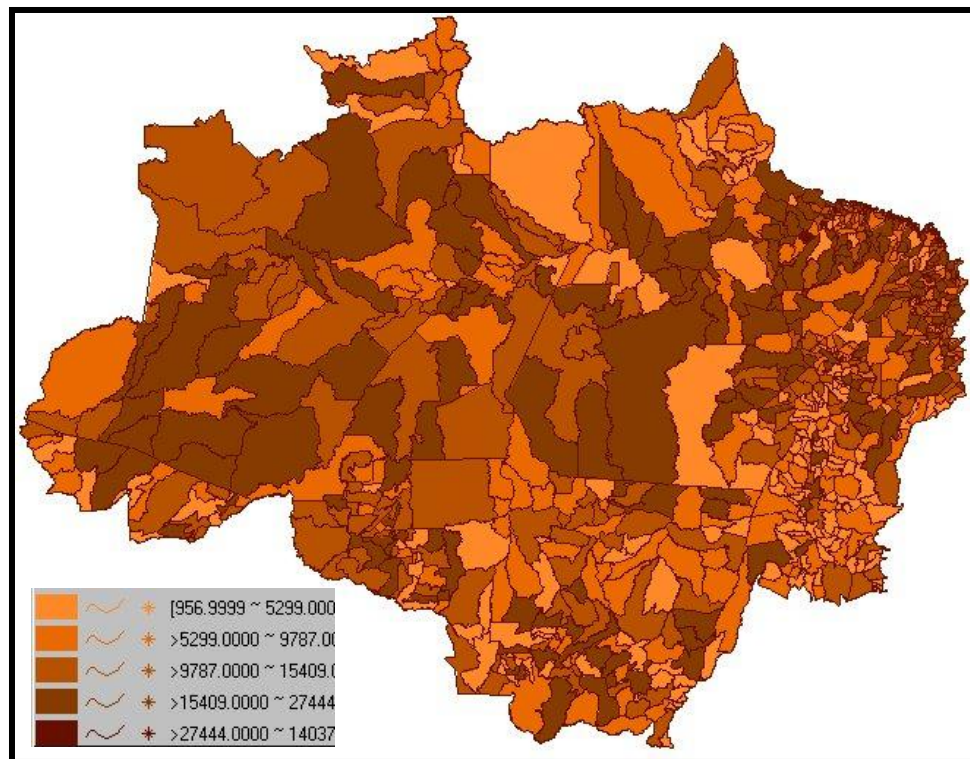


Para representar População na Amazônia...

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

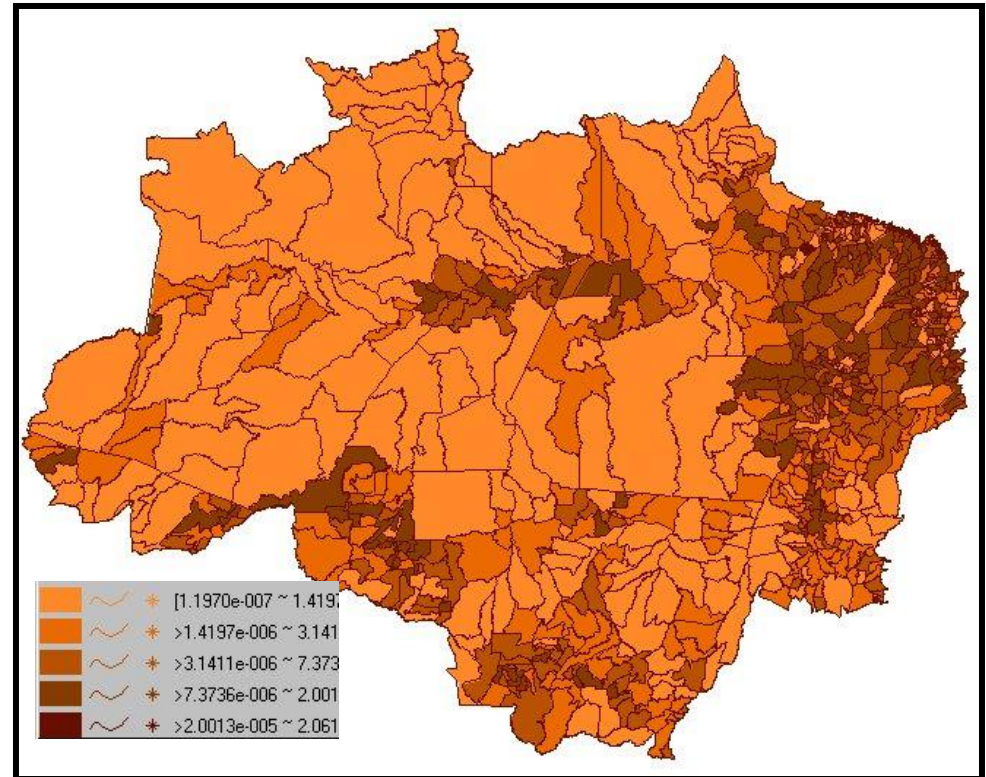
- Rondônia: migrantes, assentamentos INCRA, núcleos urbanos ao longo de eixos viários e população na zona rural.
- Amazonas: menor densidade de núcleos urbanos, concentração em Manaus.
- Tendências:
 - desconcentração das metrópoles,
 - maior participação relativa das cidades de até 100.000 habitantes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

- 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 vazias demográficas ??



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

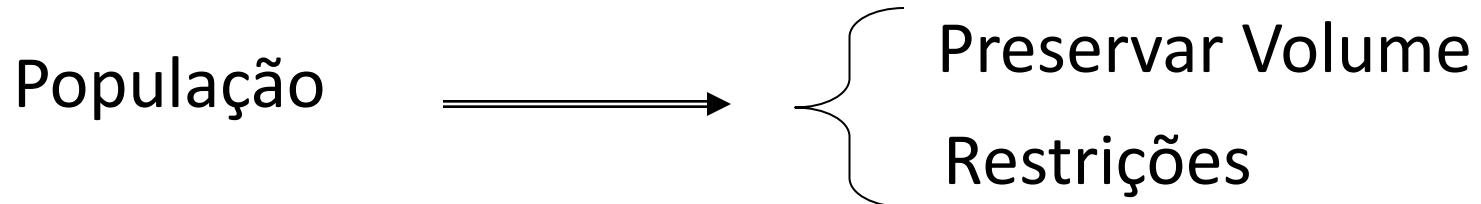
Modelos de População em Superfícies

Baseados em Pontos – *Point-based*

Baseados em Áreas – *Areal-Based*

Informação Auxiliar

Conhecimento dos Processos



Como o Dado é Coletado?

FAÇA OS ALGARISMOS CONFORME O MODELO: 0 1 2 3 4 5 6 7 8 9 NÃO RISQUE. SE NECESSÁRIO, APAGUE COM A APROPRIADA. ESCREVA SOMENTE COM A LÁPIS.

Censo Demográfico 2000
CD 1.01
Questionário Básico

MUNICÍPIO: _____
AGÊNCIA: _____

USO: _____

1 IDENTIFICAÇÃO 1

1.02 UF: _____ 1.03 MUNICÍPIO: _____ 1.04 DISTRITO: _____ 1.05 SUBDISTRITO: _____ 1.06 SETOR: _____

1.09 NA NA FOLHA DE DOMICÍLIO COLETIVO: _____ 1.10 TOTAL DE HOMENS: _____ 1.11 TOTAL DE MULHERES: _____ 1.12 QUESTIONÁRIO NO DOM: _____
1 - ÚNICO
2 - TEM OUTROS

LOCALIDADE: _____
LOGRADOURO: _____
(Endereço completo)

2 CARACTERÍSTICAS DO D

2.01 - ESPÉCIE
1 - PARTICULAR PERMANENTE → Siga questão 2.02
2 - PARTICULAR IMPROVISADO → Passe para a Lista de Moradores
3 - COLETIVO

2.02 - TIPO
1 - CASA 2 - APARTAMENTO 3 - CÔMODOS

2.03 - ESTE DOMICÍLIO É:
1 - PRÓPRIO - JÁ PAGO → Siga questão 2.04
2 - PRÓPRIO - AINDA PAGANDO → Siga questão 2.04
3 - ALUGADO → Siga questão 2.04
4 - CEDIDO POR EMPREGADOR → Passe para o questionário 2.05
5 - CEDIDO DE OUTRA FORMA → Passe para o questionário 2.05
6 - OUTRA CONDIÇÃO

2.04 - O TERRENO EM QUE SE LOCALIZA ESTE DOMICÍLIO É:
1 - PRÓPRIO 2 - CEDIDO 3 - OUTRA CONDIÇÃO

2.05 - A FORMA DE ABASTECIMENTO DE ÁGUA UTILIZADA NESTE DOMICÍLIO É:
1 - REDE GERAL
2 - POÇO OU NASCENTE (Na propriedade)
3 - OUTRA

2.06 - A ÁGUA UTILIZADA NESTE DOMICÍLIO CHEGA:
1 - CANALIZADA EM PELO MENOS UM CÔMODOS
2 - CANALIZADA SÓ NA PROPRIEDADE OU TERRENO
3 - NÃO CANALIZADA

FAÇA OS ALGARISMOS CONFORME O MODELO: 0 1 2 3 4 5 6 7 8 9 NÃO RISQUE. SE NECESSÁRIO, APAGUE COM A BORRACHA APROPRIADA. ESCREVA SOMENTE COM A LÁPIS. PREENCHA A QUADRICULA DESTA FORMA. X

Censo Demográfico 2000
CD 1.02
Questionário da Amostra

MUNICÍPIO: _____
AGÊNCIA: _____

1 IDENTIFICAÇÃO 1.01 NÚMERO DO QUESTIONÁRIO: _____

1.02 UF: _____ 1.03 MUNICÍPIO: _____ 1.04 DISTRITO: _____ 1.05 SUBDISTRITO: _____ 1.06 SETOR: _____ 1.07 NA DA PÁGINA DA FOLHA DE COLETA: _____ 1.08 NA NA FOLHA DE COLETA: _____

1.09 NA NA FOLHA DE DOMICÍLIO COLETIVO: _____ 1.10 TOTAL DE HOMENS: _____ 1.11 TOTAL DE MULHERES: _____ 1.12 QUESTIONÁRIO NO DOMICÍLIO: _____
1 - ÚNICO 3 - É CONTINUAÇÃO
2 - TEM OUTROS

1.13 TOTAL DE QUESTIONÁRIOS UTILIZADOS NO DOMICÍLIO: _____

LOCALIDADE: _____
LOGRADOURO: _____
(Endereço completo)

2 CARACTERÍSTICAS DO DOMICÍLIO

2.01 - ESPÉCIE
1 - PARTICULAR PERMANENTE → Siga questão 2.02
2 - PARTICULAR IMPROVISADO → Passe para a Lista de Moradores
3 - COLETIVO

2.02 - TIPO
1 - CASA 2 - APARTAMENTO 3 - CÔMODOS

2.03 - QUANTOS CÔMODOS EXISTEM NESTE DOMICÍLIO?

2.04 - QUANTOS CÔMODOS SERVEM DE DORMITÓRIO PARA OS MORADORES NESTE DOMICÍLIO?

Se 9 cômodos ou mais servindo de dormitório, registrar 9 (nove).

2.05 - ESTE DOMICÍLIO É:
1 - PRÓPRIO - JÁ PAGO → Siga questão 2.06
2 - PRÓPRIO - AINDA PAGANDO → Siga questão 2.06
3 - ALUGADO → Siga questão 2.06
4 - CEDIDO POR EMPREGADOR → Passe para o questionário 2.07
5 - CEDIDO DE OUTRA FORMA → Passe para o questionário 2.07
6 - OUTRA CONDIÇÃO

2.06 - O TERRENO EM QUE SE LOCALIZA ESTE DOMICÍLIO É:
1 - PRÓPRIO 2 - CEDIDO 3 - OUTRA CONDIÇÃO

2.07 - A FORMA DE ABASTECIMENTO DE ÁGUA UTILIZADA NESTE DOMICÍLIO É:
1 - REDE GERAL
2 - POÇO OU NASCENTE (Na propriedade)
3 - OUTRA

2.08 - A ÁGUA UTILIZADA NESTE DOMICÍLIO CHEGA:
1 - CANALIZADA EM PELO MENOS UM CÔMODOS
2 - CANALIZADA SÓ NA PROPRIEDADE OU TERRENO
3 - NÃO CANALIZADA

2.09 - QUANTOS BANHEIROS EXISTEM NESTE DOMICÍLIO?
(Considere somente os que contém chuveiro ou banheira e aparelho sanitário)

Se 0 (zero) siga questão 2.10. Caso contrário, passe para o questionário 2.11.
Se 9 banheiros ou mais registrar 9 (nove).

2.10 - NESTE DOMICÍLIO, TERRENO OU PROPRIEDADE EXISTE SANITÁRIO UTILIZADO PELOS MORADORES?
1 - SIM → Siga questão 2.11
2 - NÃO → Passe para o questionário 2.12

Base Territorial - Censos 2000 e 2007

27 Unidades da Federação

162.769 Setores Coleta

92.685 Urbanos

70.084 Rurais

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

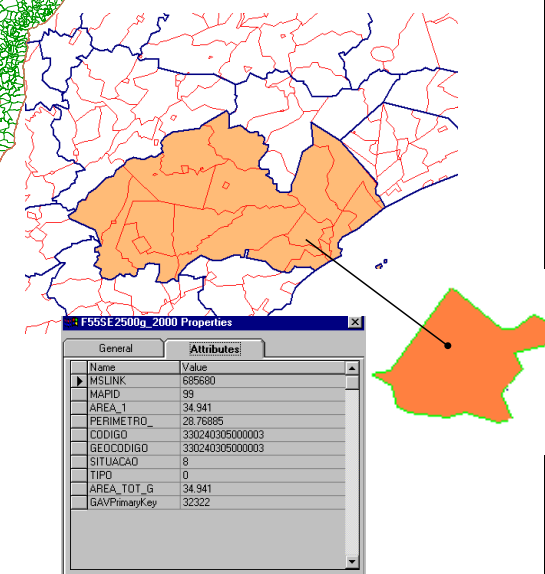
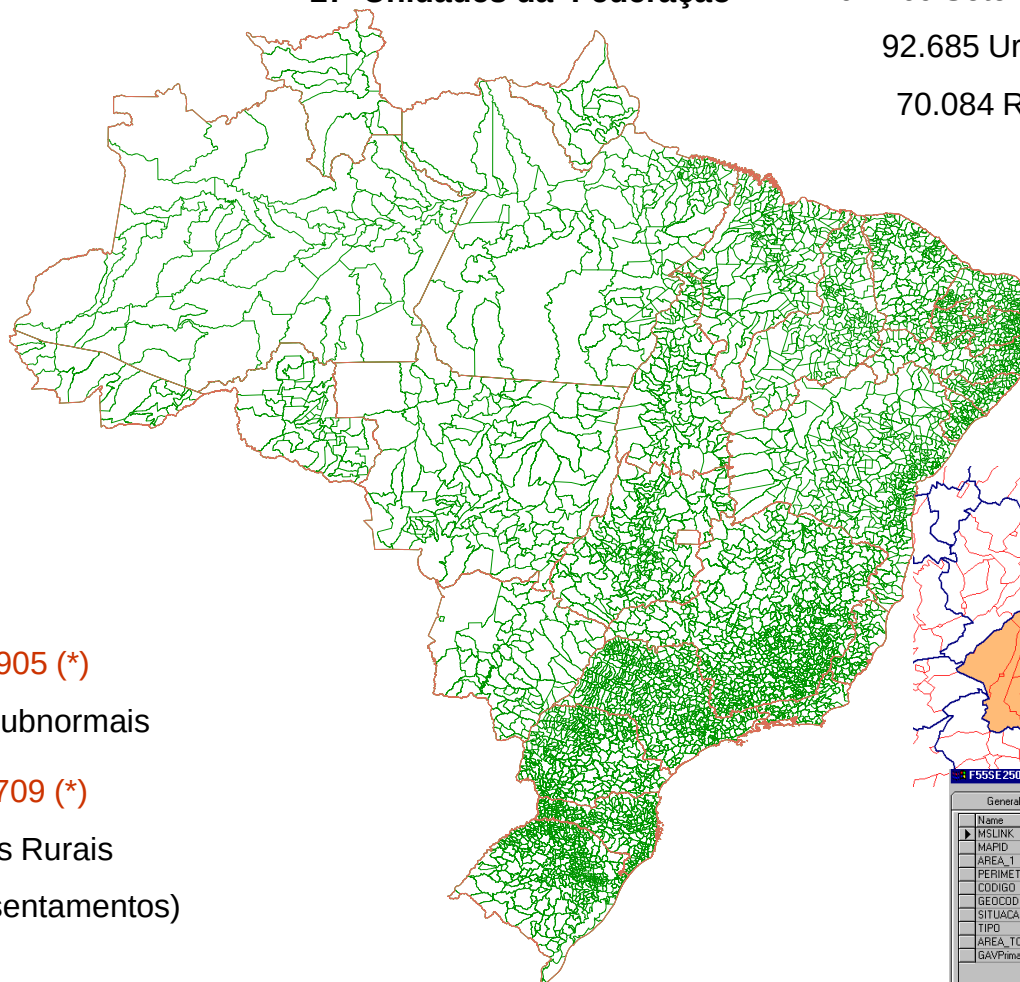
(sendo 1.671 Assentamentos)

1.047 TI e UCA

872 (*) AESP

249.068 215.811 (*)

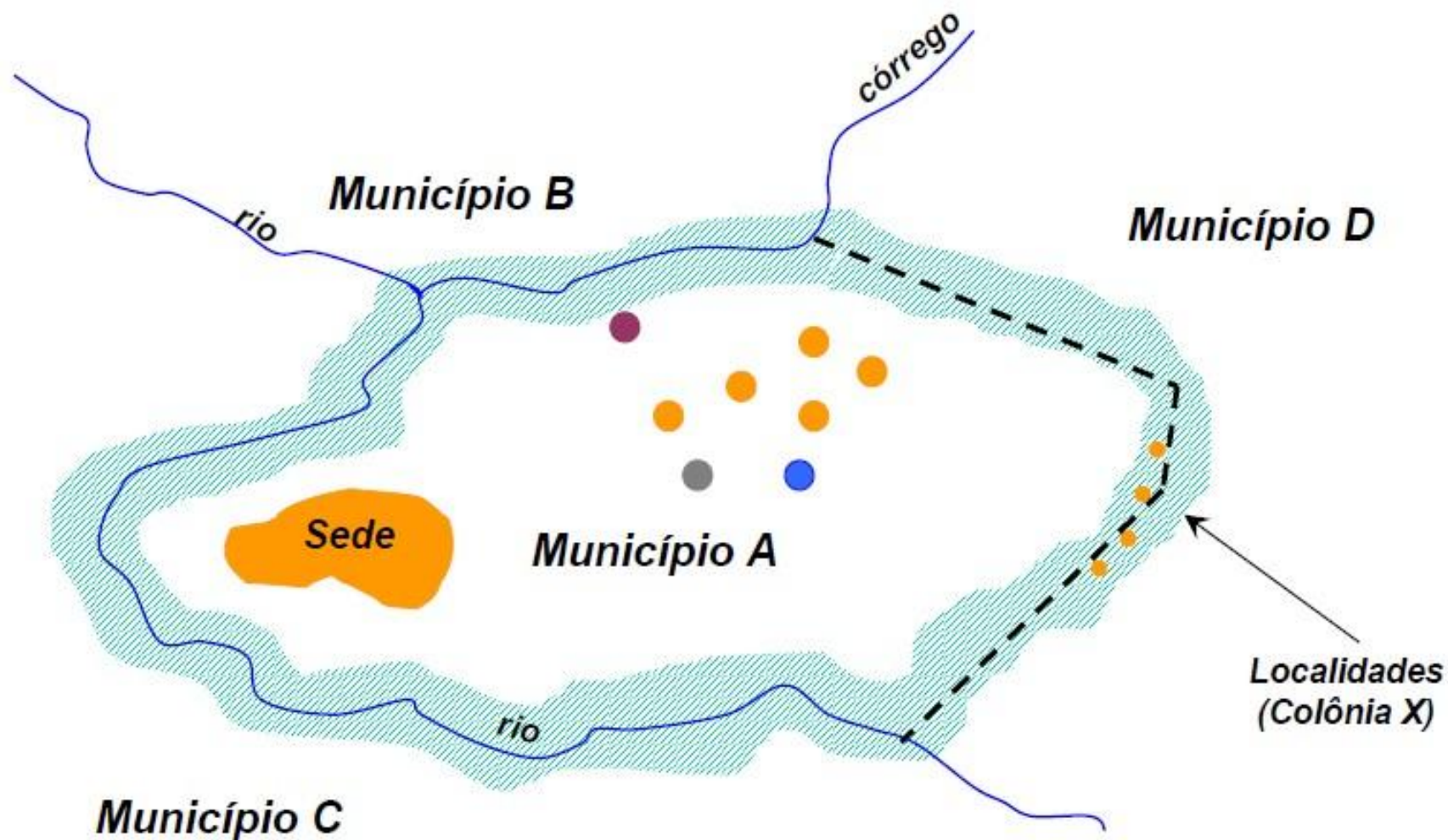
Setores Censitários



Como o Dado é Disseminado ?

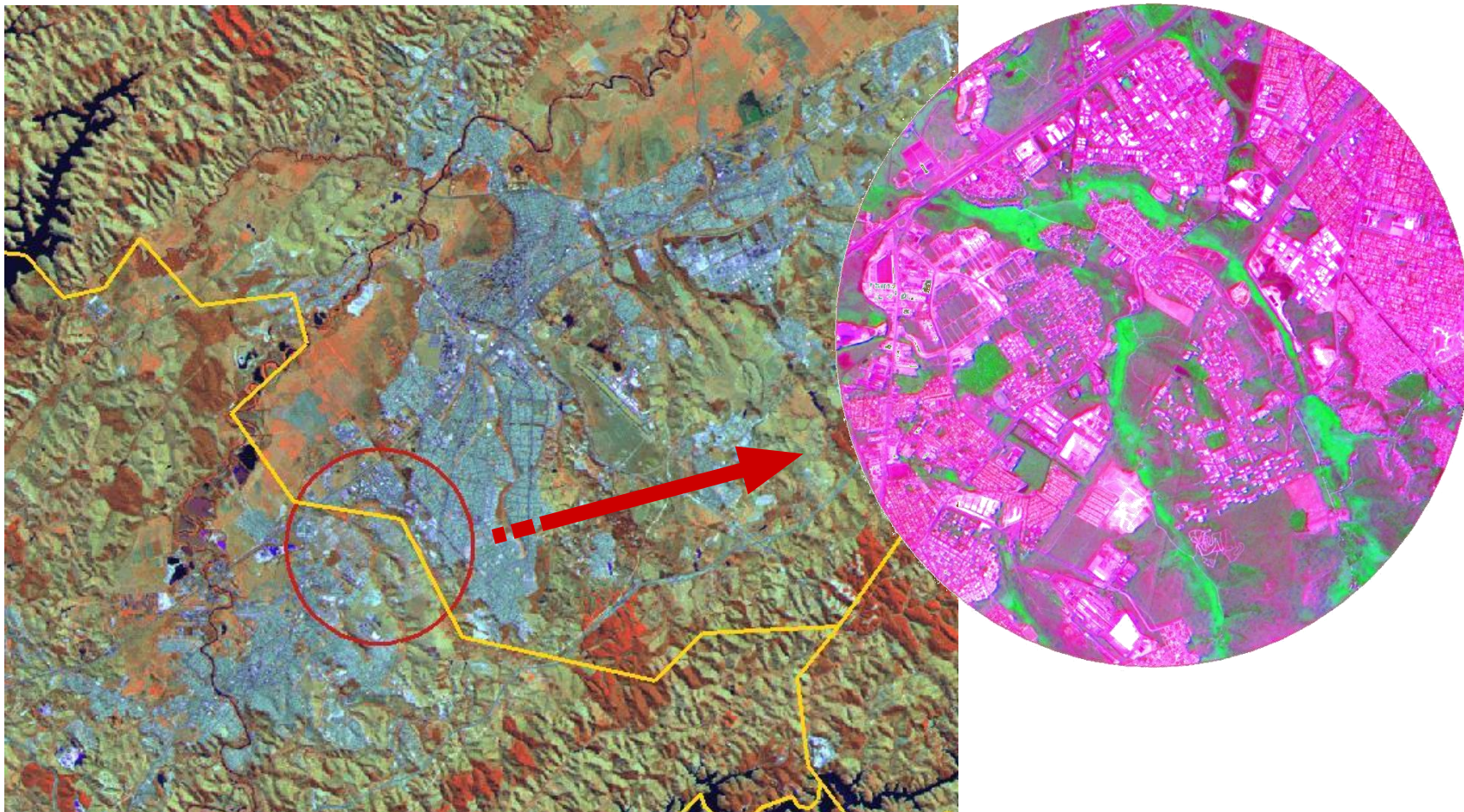
Malha Setorial do Município de Macaé

(*) Números da Base Territorial 2000

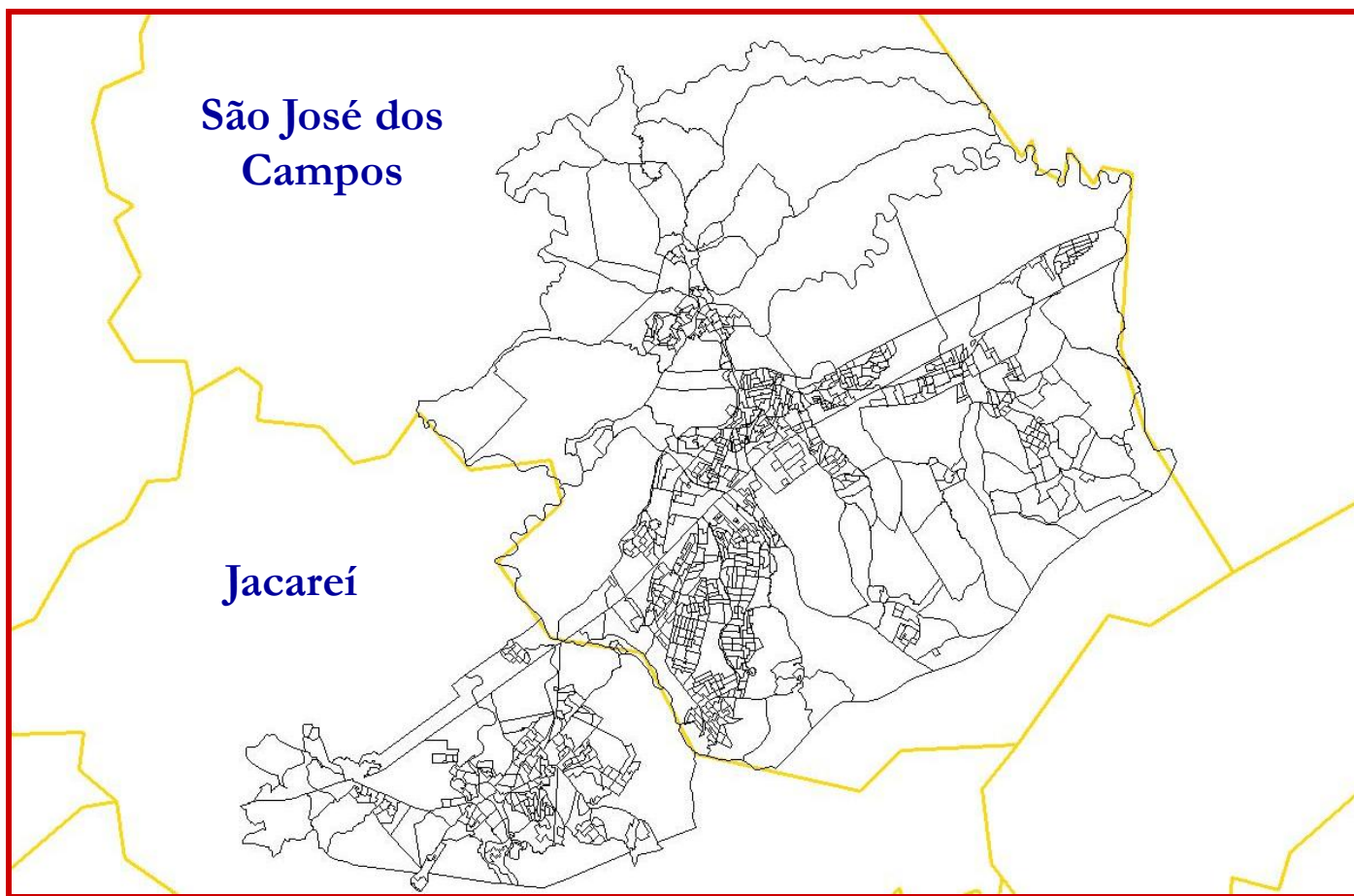




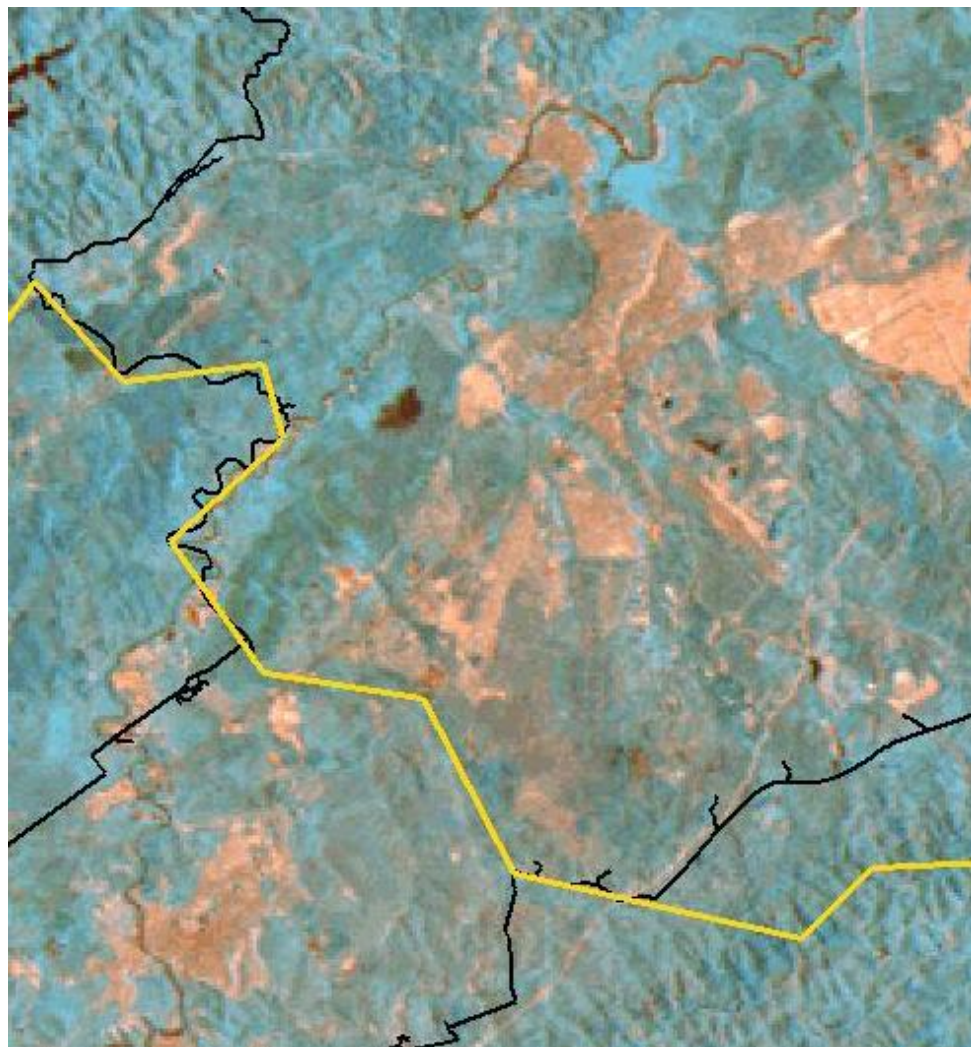
Conurbação

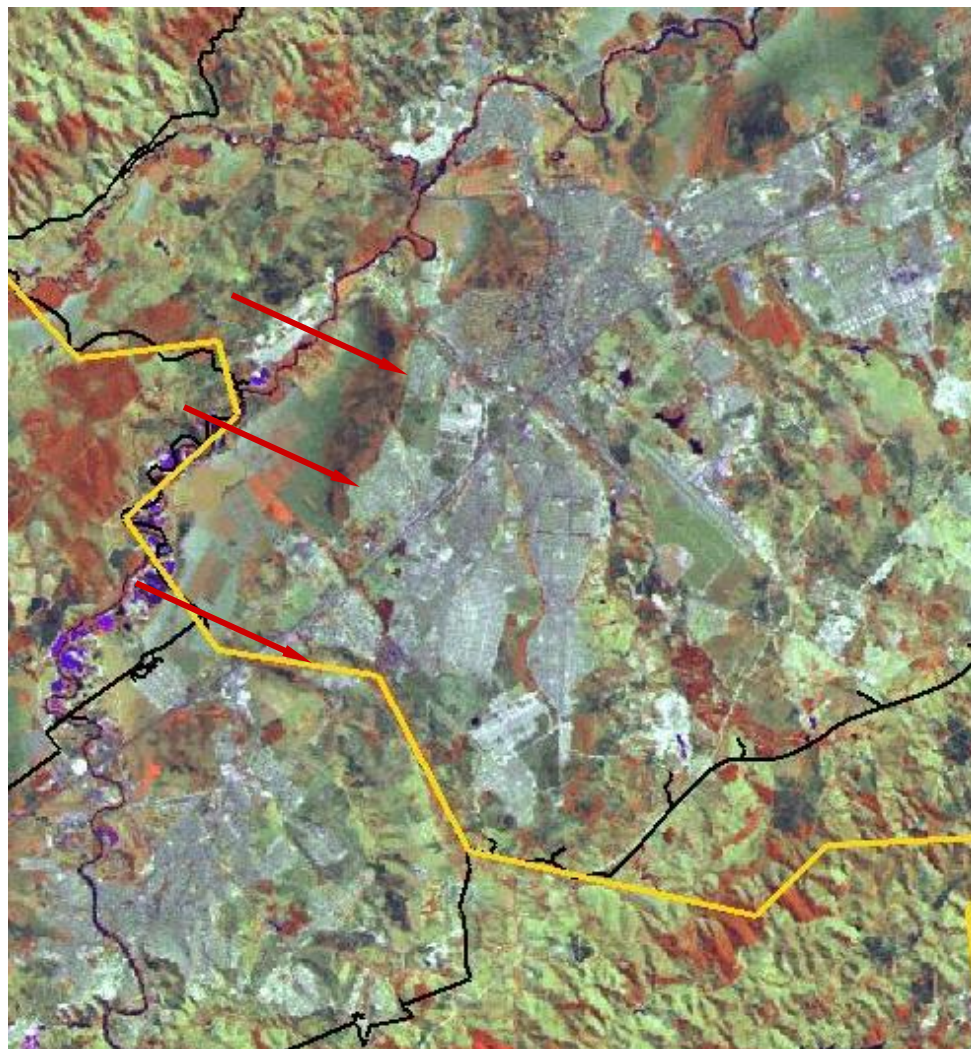


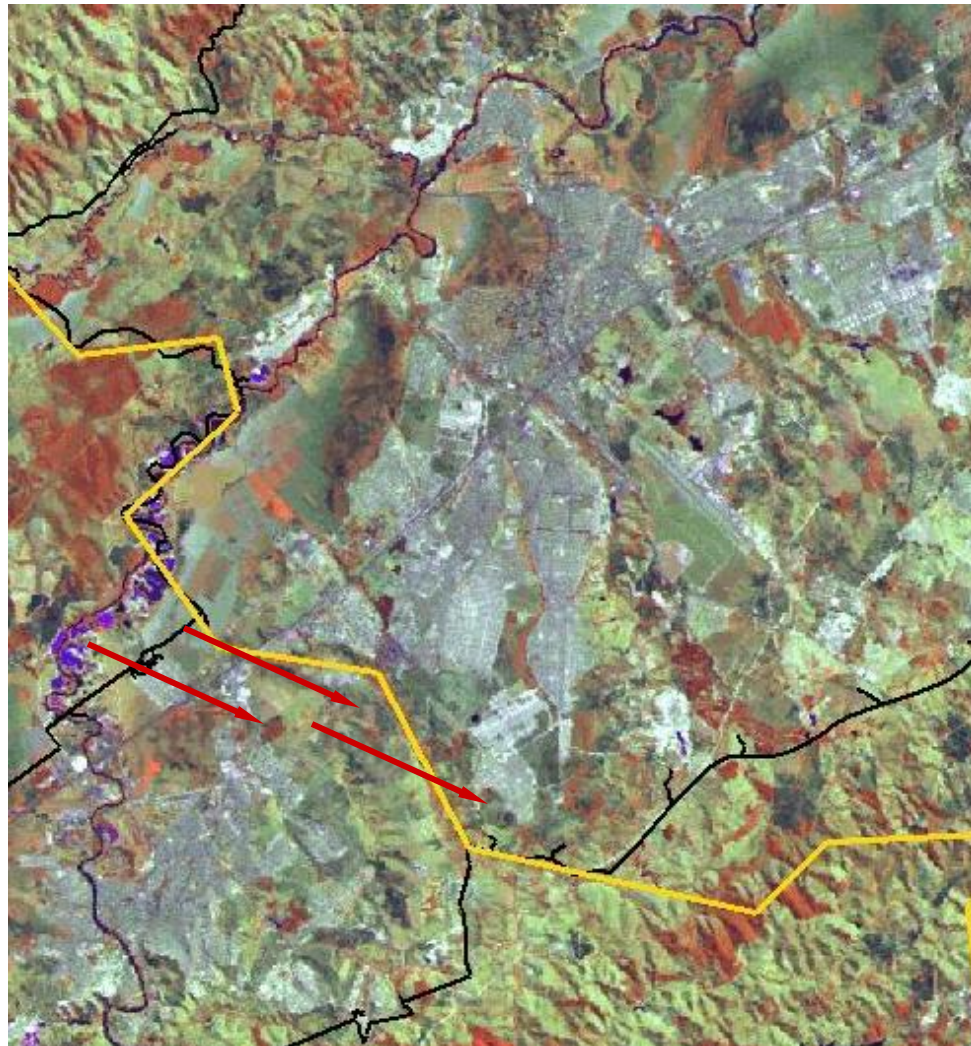
Caso: São José dos Campos - Jacareí

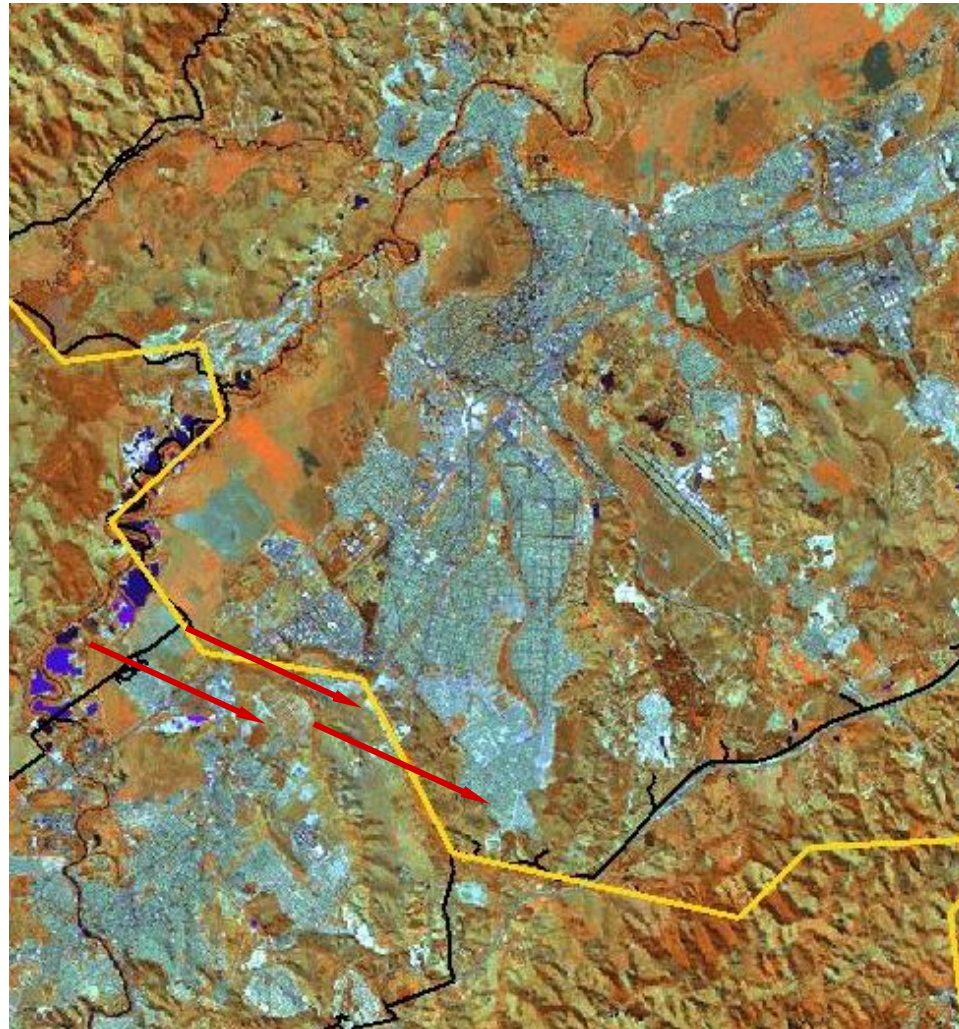


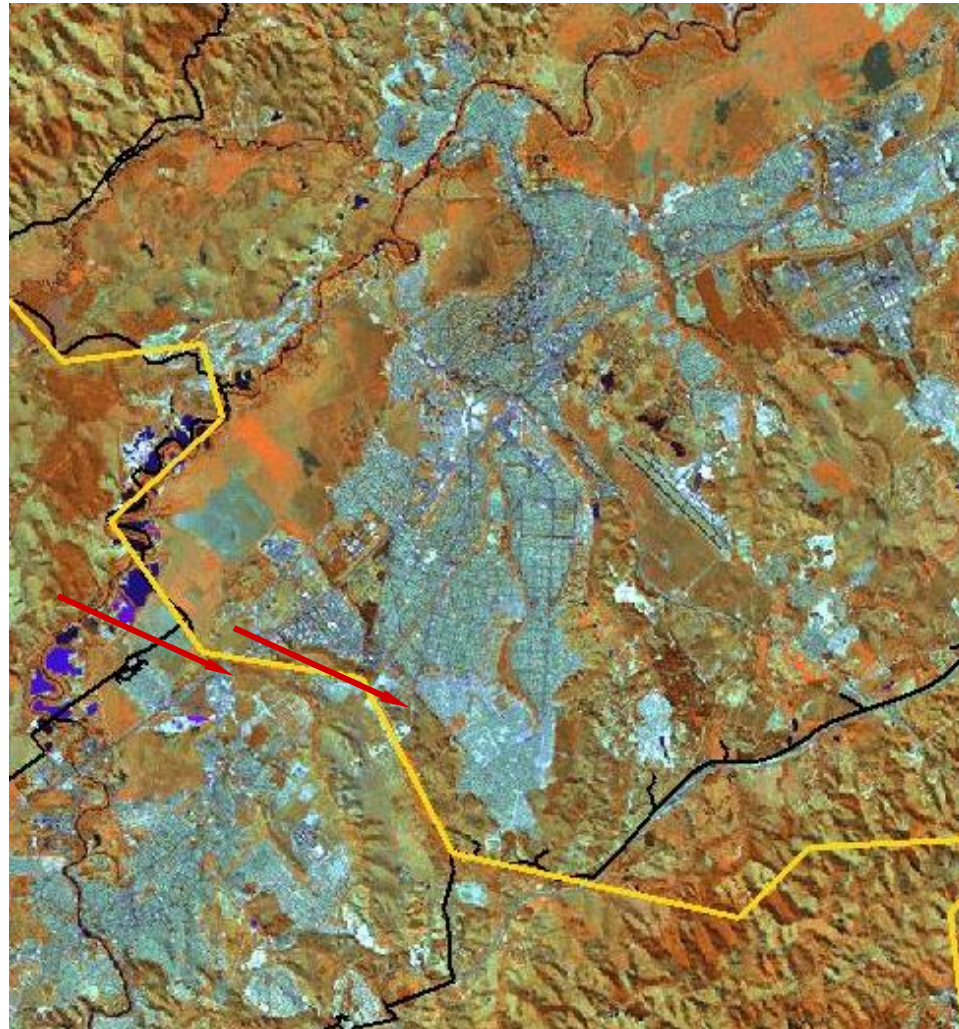
Setores censitários não rurais (2000)

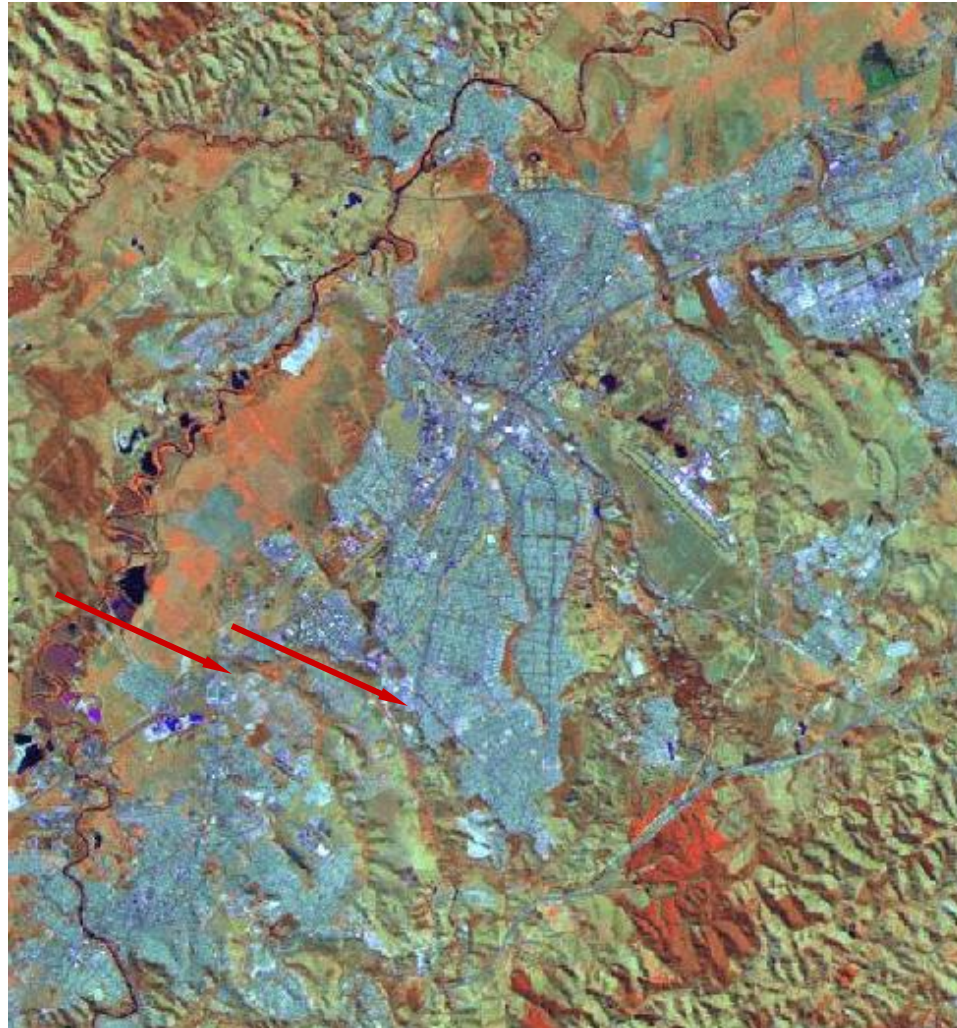


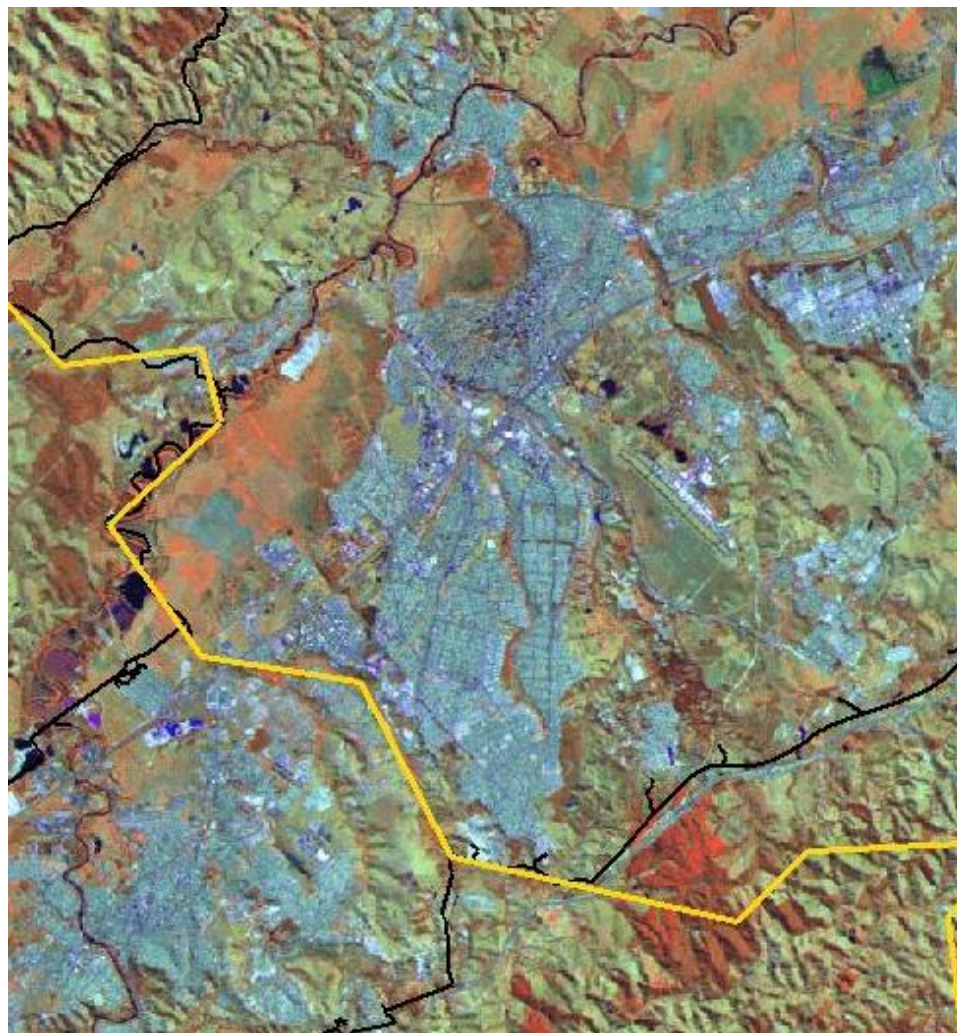


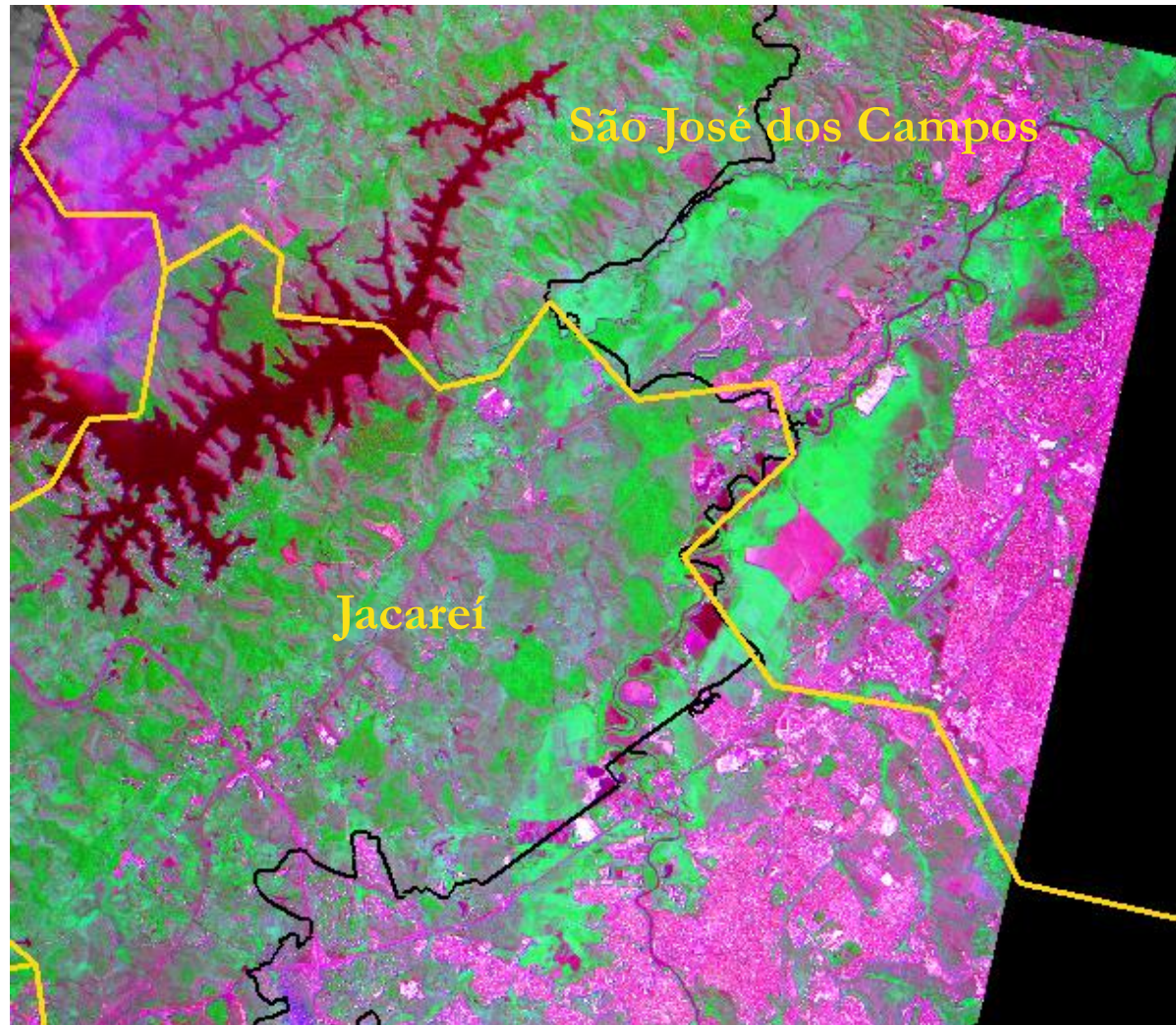


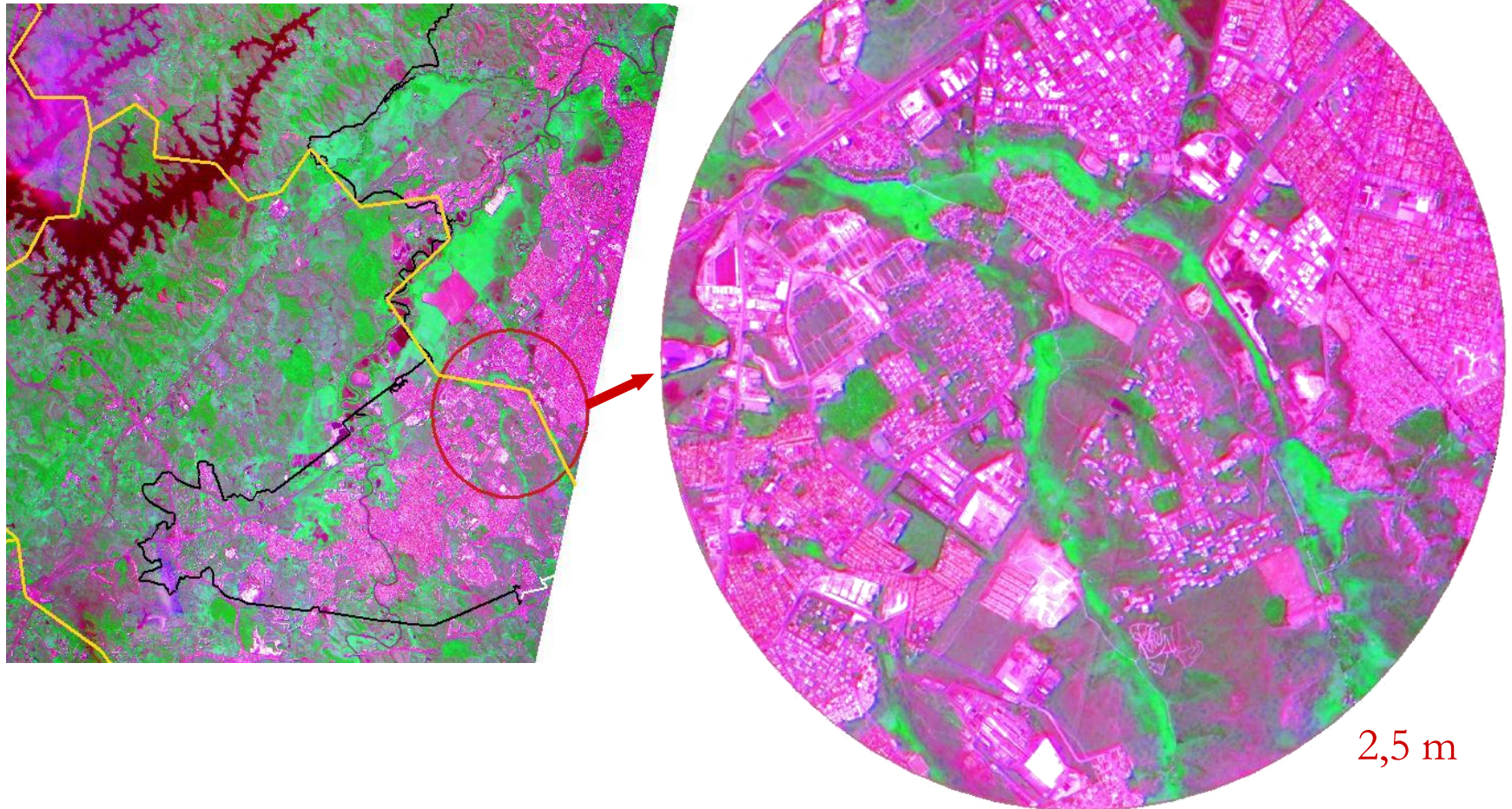




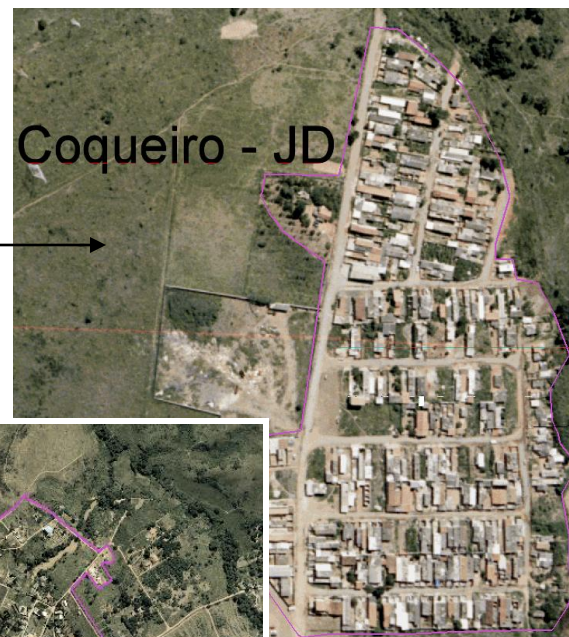
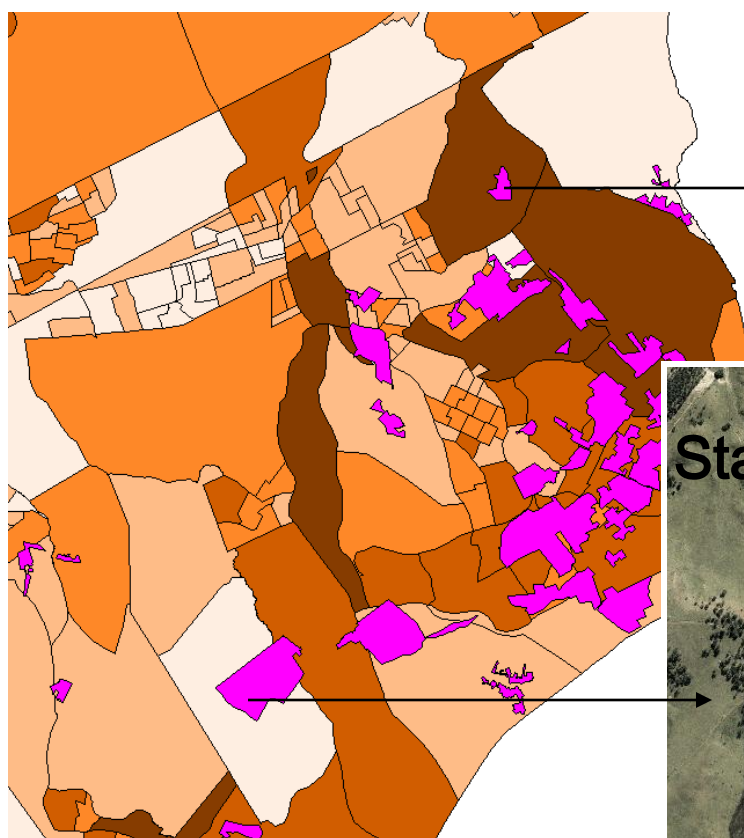












Interpoladores *Geográficos*



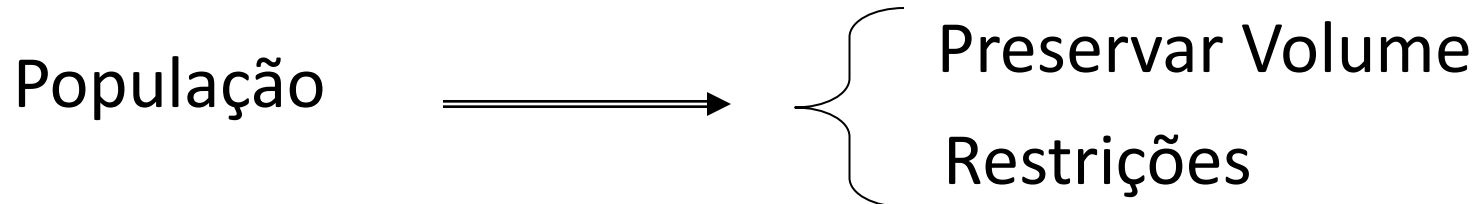
Modelos de População em Superfícies

Baseados em Pontos – *Point-based*

Baseados em Áreas – *Areal-Based*

Informação Auxiliar

Conhecimento dos Processos



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 ??



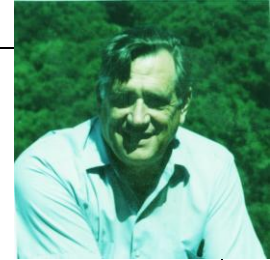
Outro desafio para interpolação



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.

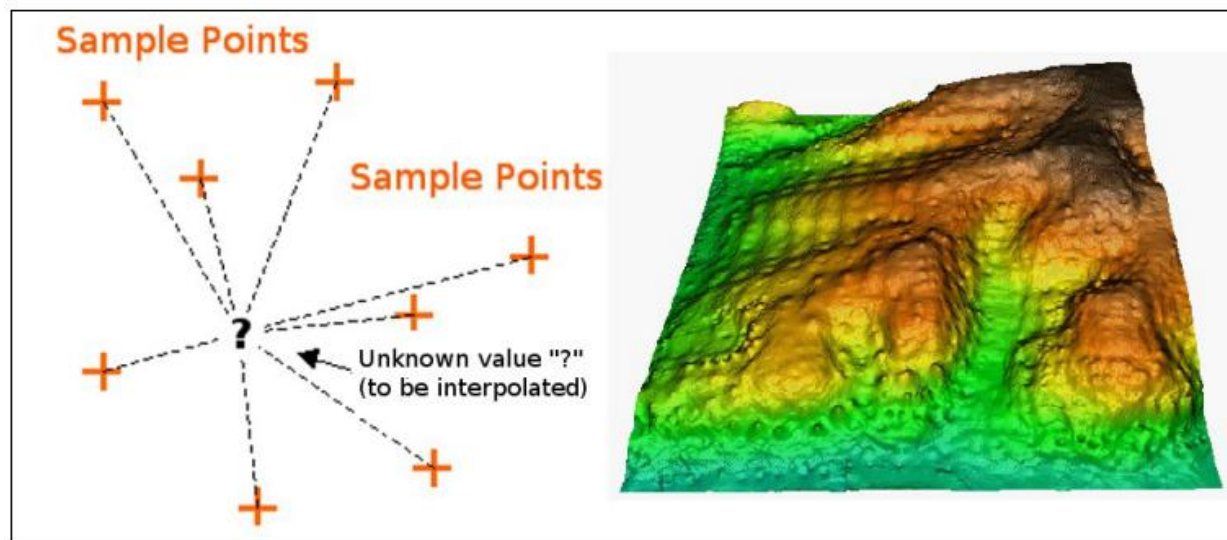


Illustration 2: Inverse Distance Weighted interpolation based on weighted sample point distance (left). Interpolated IDW surface from elevation vector points (right). Image Source: Mitas, L., Mitasova, H. (1999)).

Referência



Geoestatística para geoprocessamento

Eduardo C. G. Camargo

Geoestatística para geoprocessamento

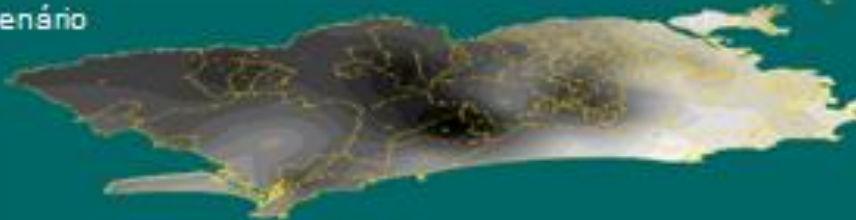
realidade



análise estrutural



cenário



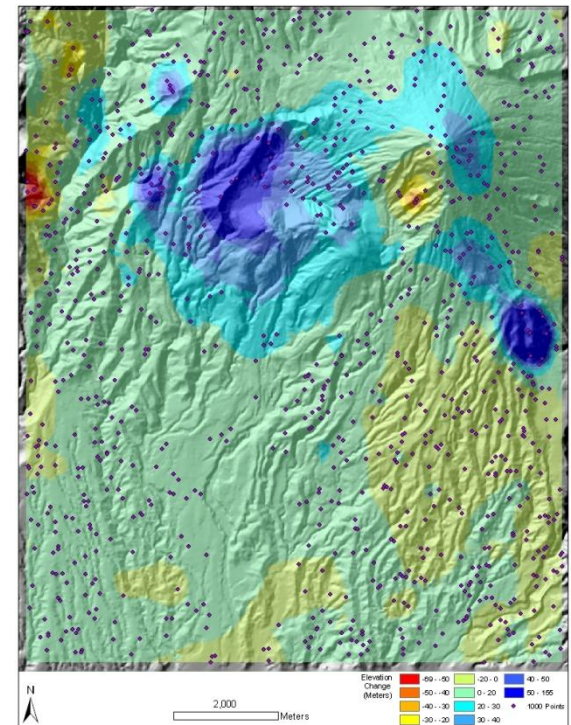
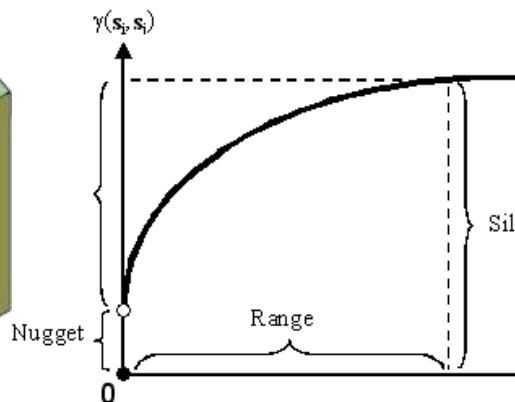
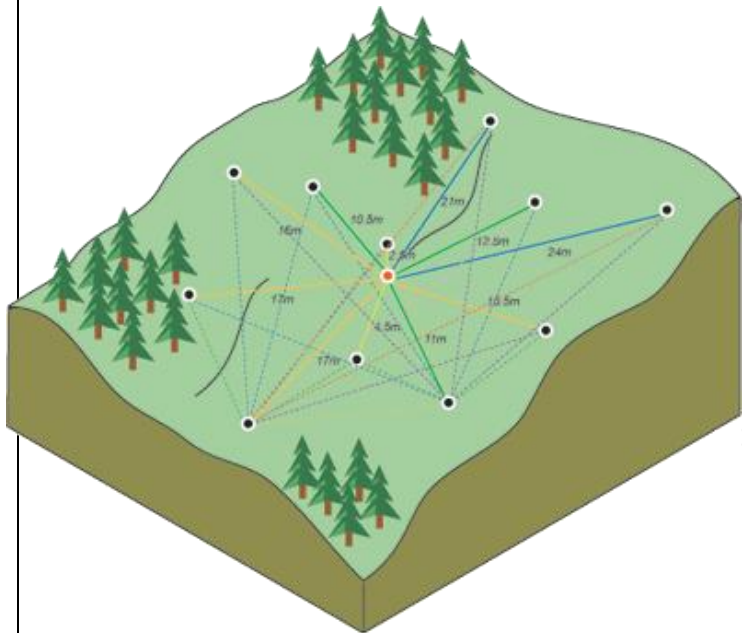
Palavras-chave

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

Modelos de População “Univariados”

• Krigagem

- 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.

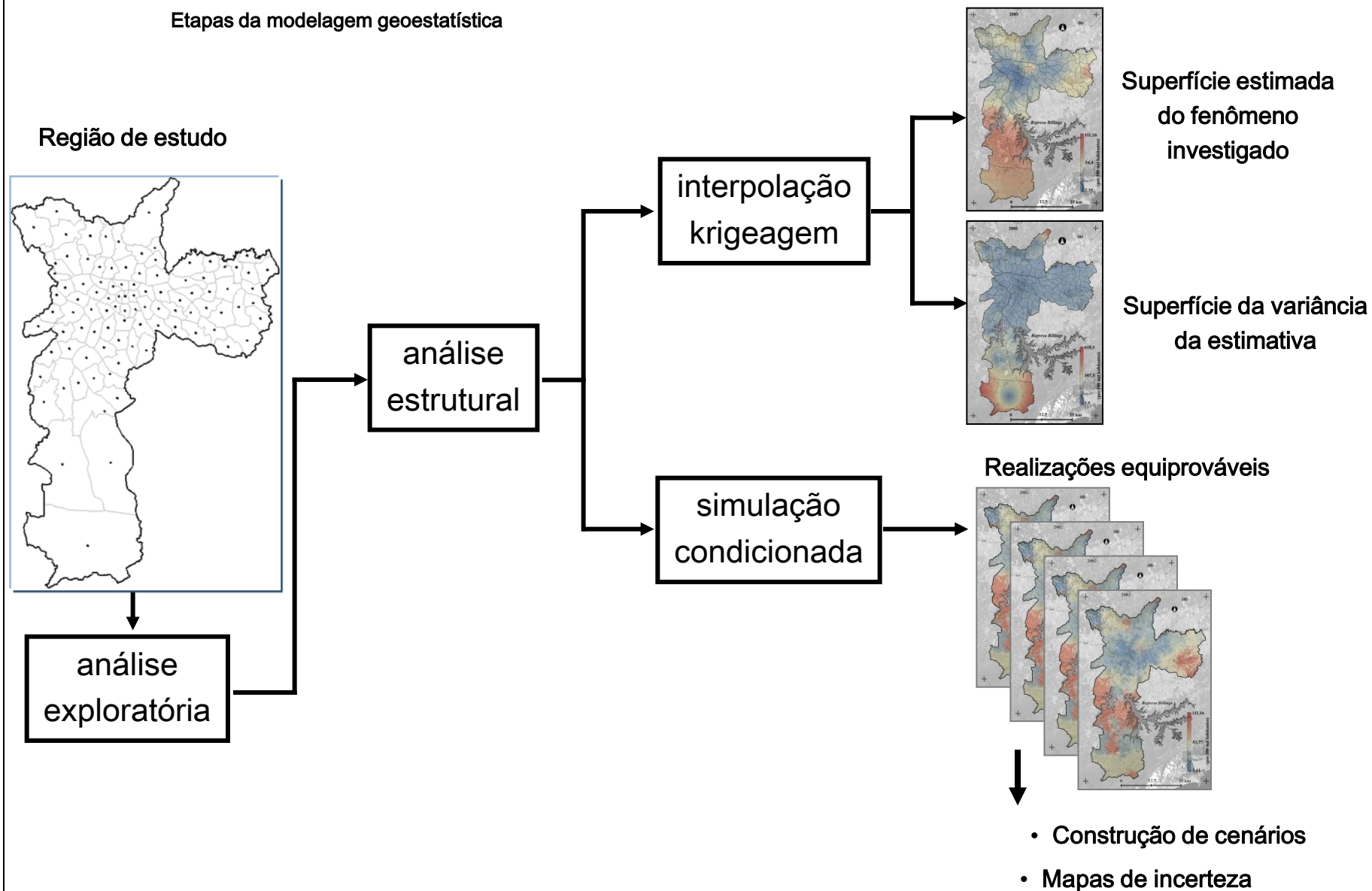
Requer o conhecimento de alguns conceitos básicos:

- 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



Porque usar geoestatística?

Interpoladores

Procedimentos determinísticos

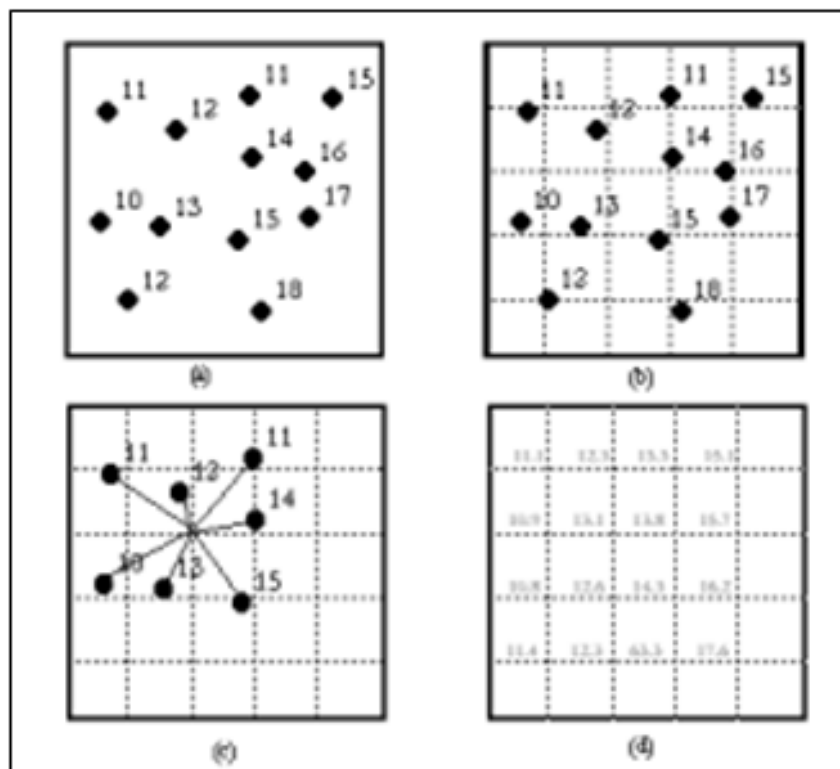


Figura 3-3 Ilustração do processo de interpolação por estimador local: (a) configuração original de amostras; (b) grade regular superposta às amostras; (c) interpolação de um valor a partir dos vizinhos; (d) grade regular resultante

Porque usar geoestatística?



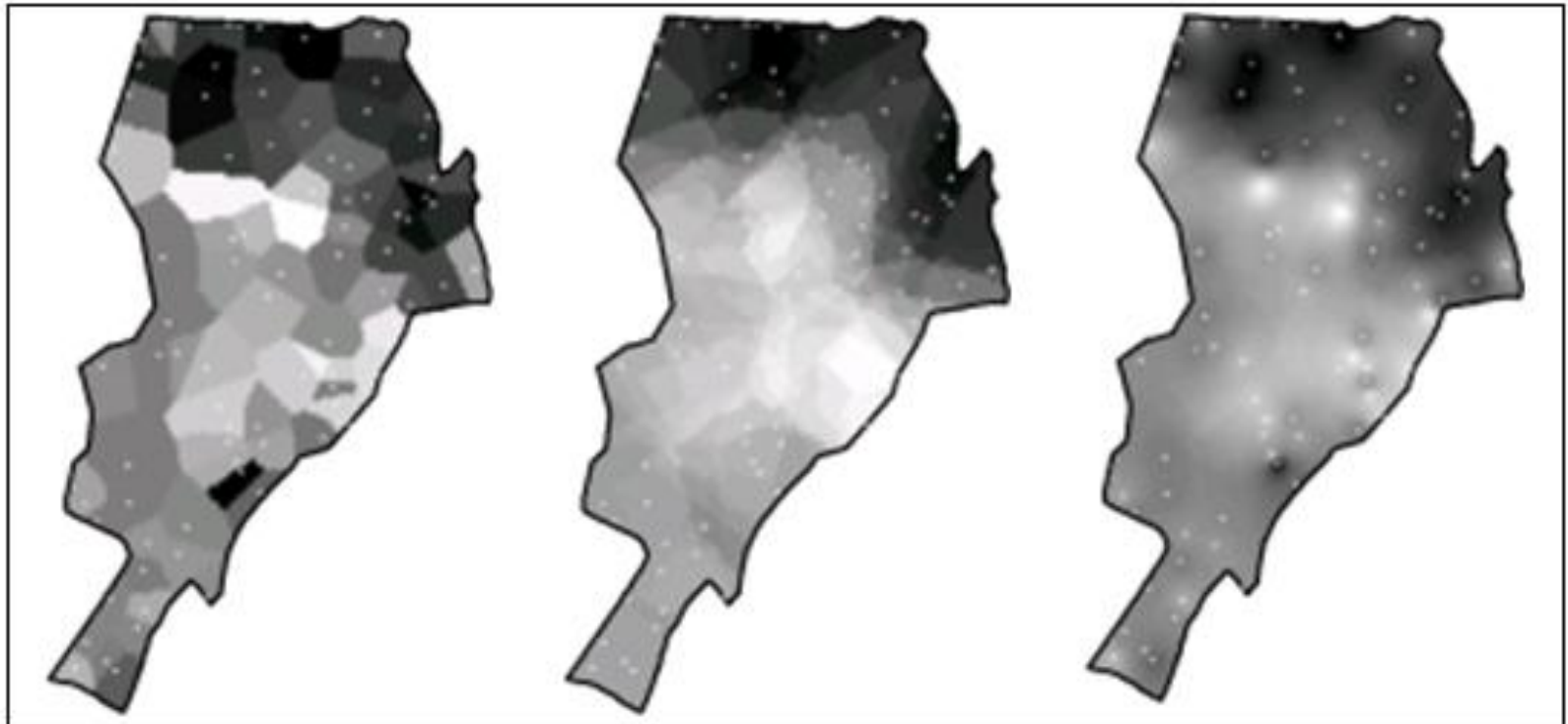
Interpoladores

Procedimentos determinísticos

Vizinho mais Próximo

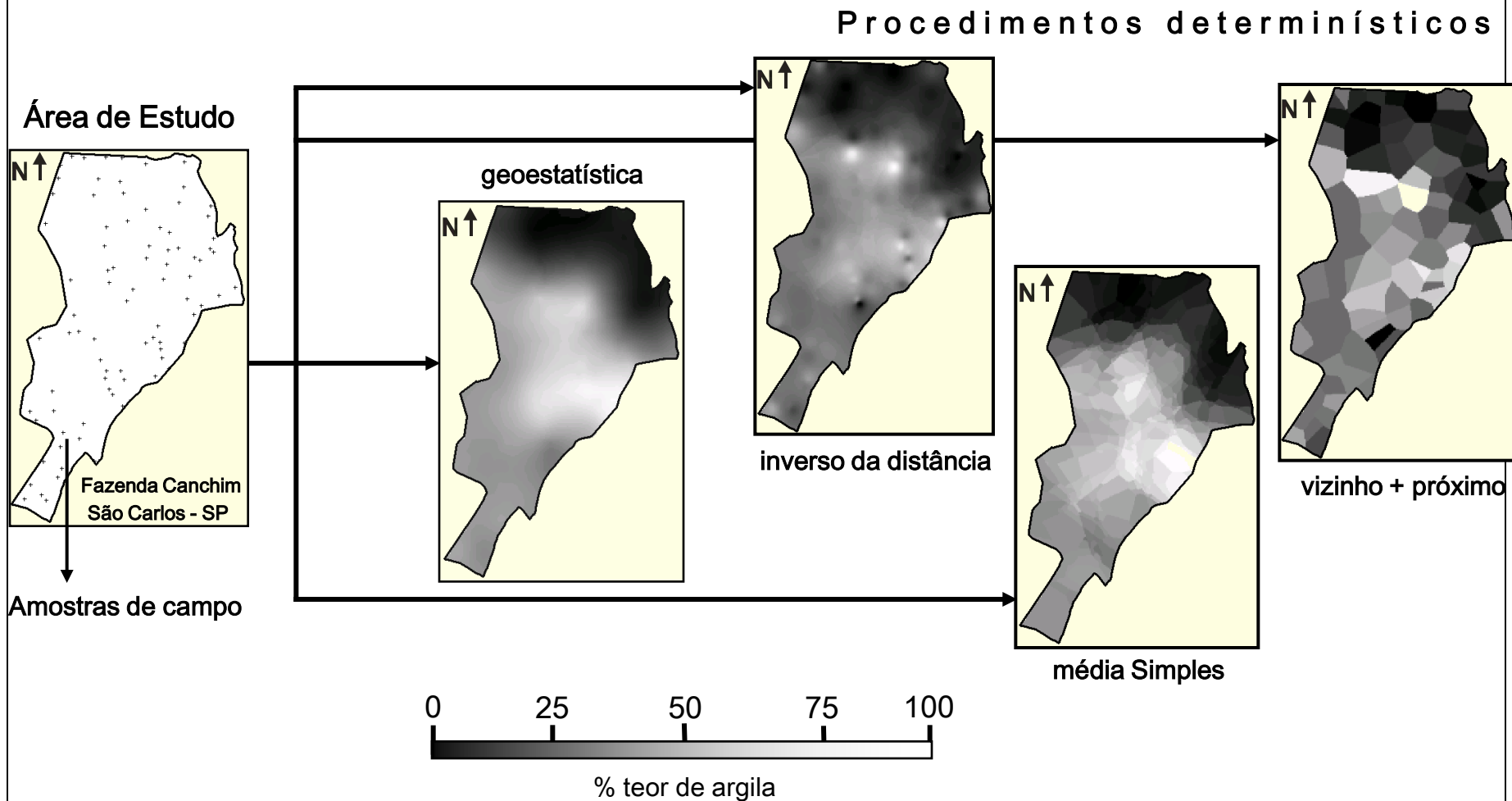
Média Simples

Inverso quadrado distância



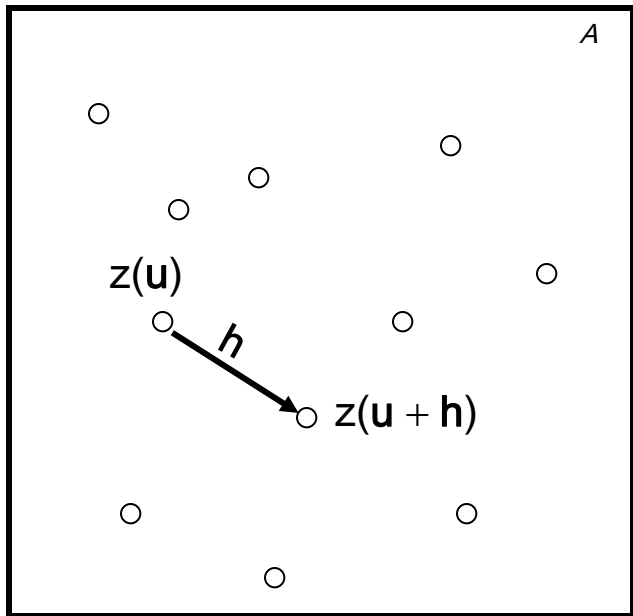
Porque usar geoestatística?

Interpoladores



Variograma $2\gamma(h)$

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).



$2\gamma(h)$

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

Variograma $2\gamma(h)$

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

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

Através de um conjunto amostral, $\{z(u_1), z(u_2), \dots, z(u_N)\}$, o variograma pode ser estimado por:

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

em que:

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

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

$N(h)$: é o número de pares, $z(u_i)$ e $z(u_i + h)$, separados por h ;

$z(u_i)$ e $z(u_i + h)$: são valores observados nas localizações u_i e $u_i + h$.

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:

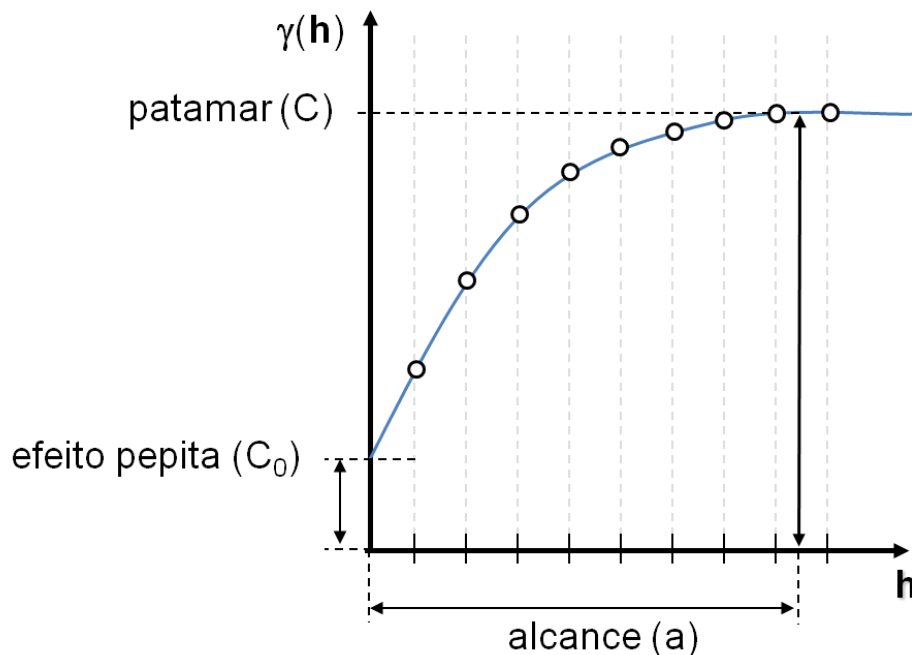
$\hat{\gamma}(\mathbf{h})$: é o estimador de semivariograma;

$\mathbf{h}$, $N(\mathbf{h})$, $z(\mathbf{u}_i)$ 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.

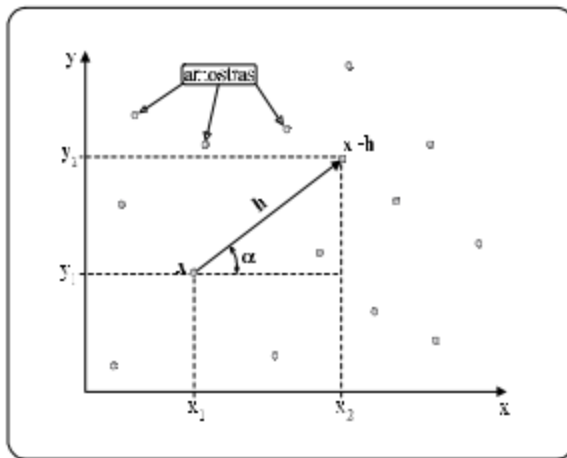


Fig. 2.2 - Amostragem em duas dimensões.

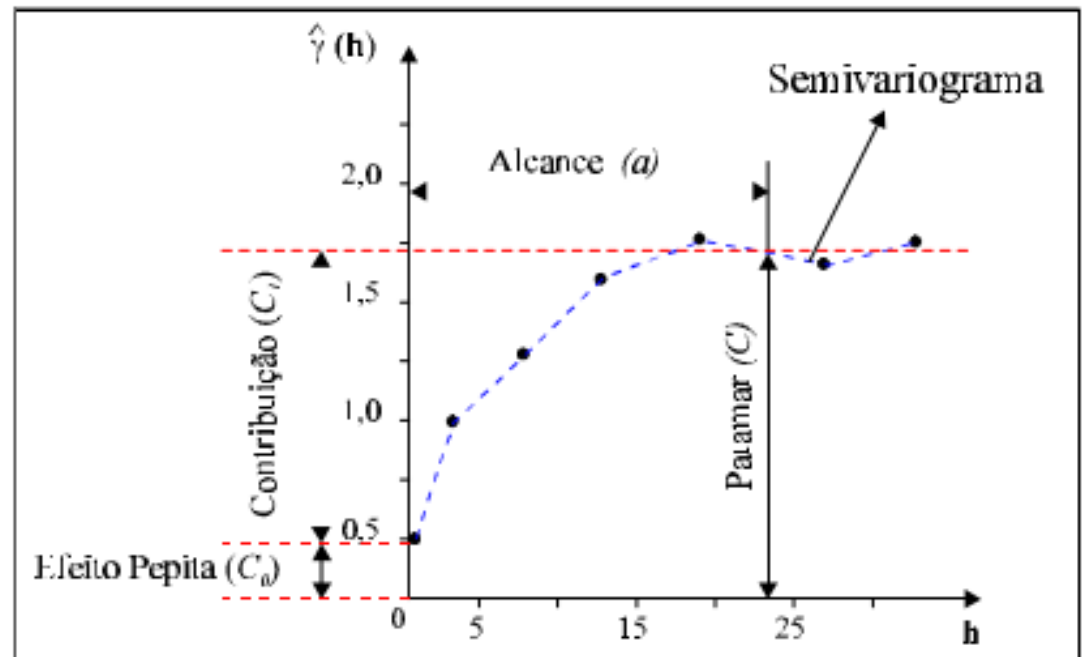
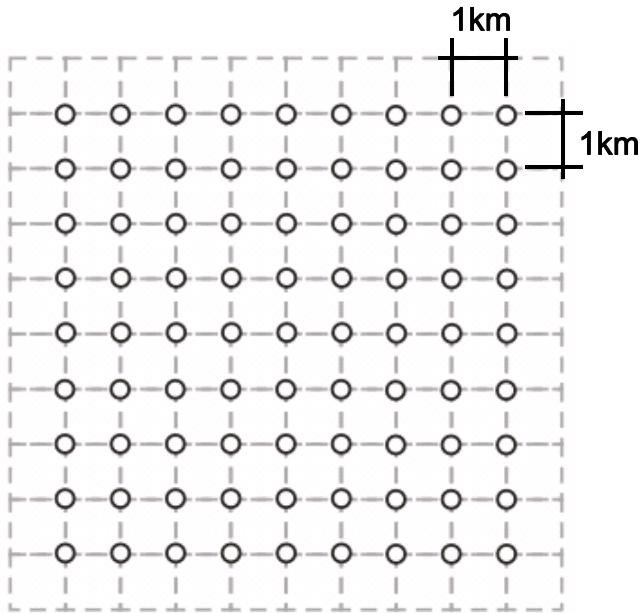


Fig. 2.3 - Exemplo de semivariograma.

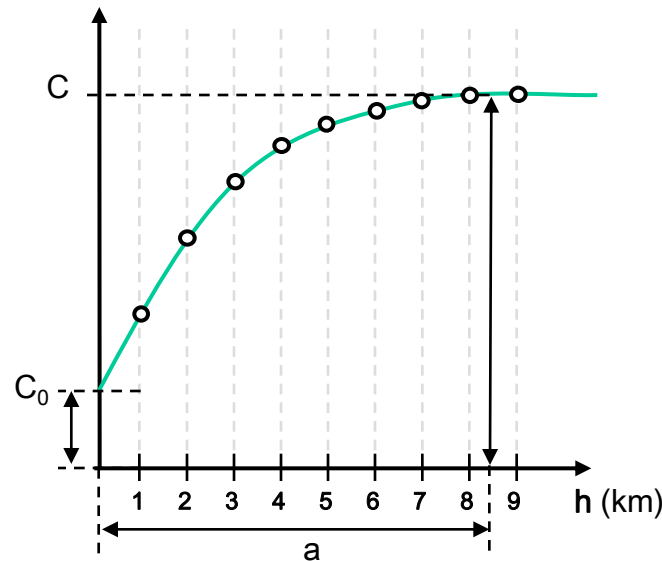
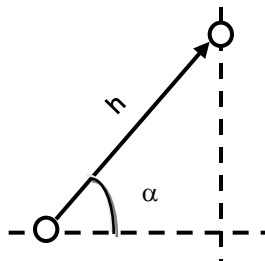
Semivariograma $\gamma(h)$

Cálculo do semivariograma a partir de amostras regularmente espaçadas.

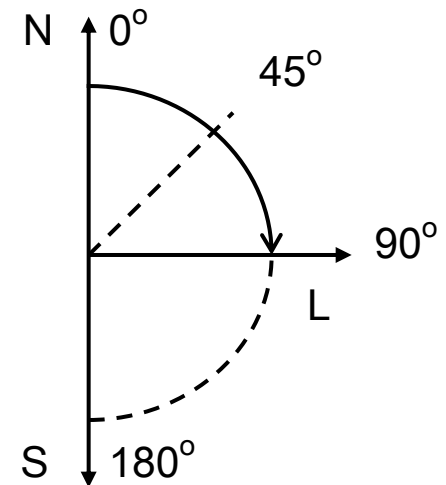


$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(u_i) - z(u_i + h)]^2$$

vetor distância h



direções de análise



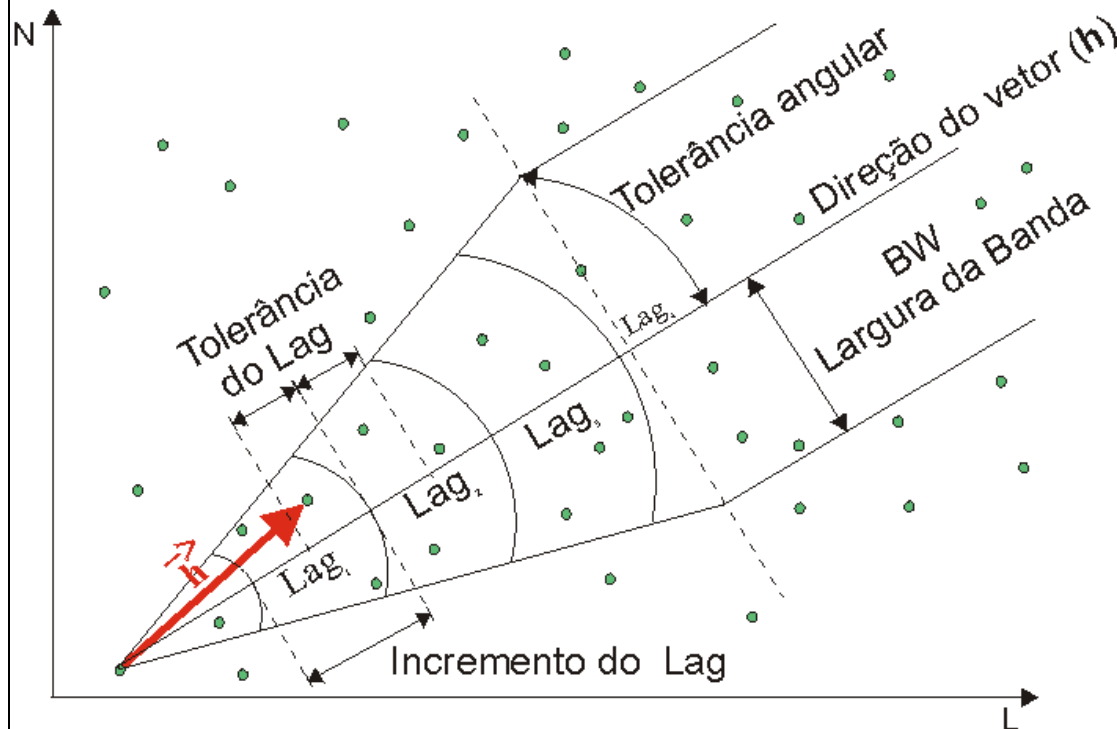
função simétrica

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

Semivariograma $\gamma(h)$

Cálculo do semivariograma a partir de amostras irregularmente espaçadas.

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



parâmetros adicionais

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

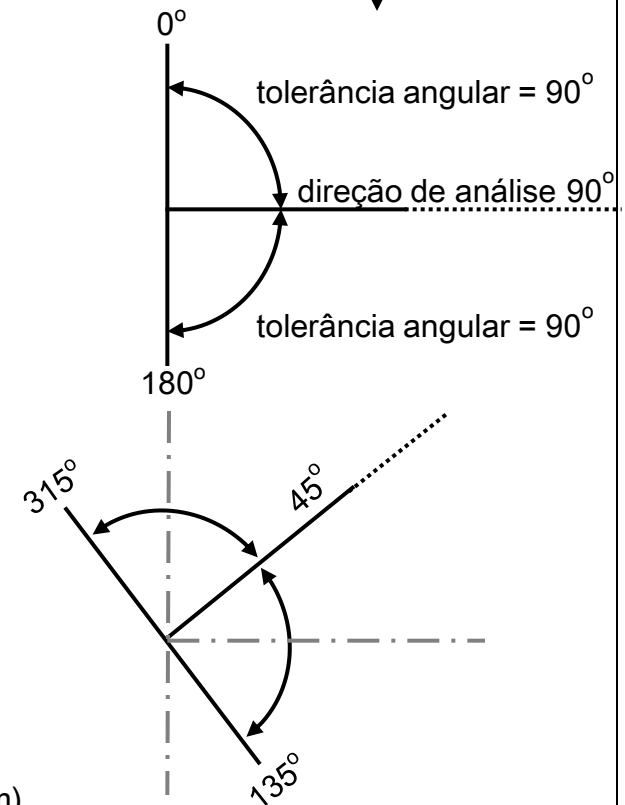
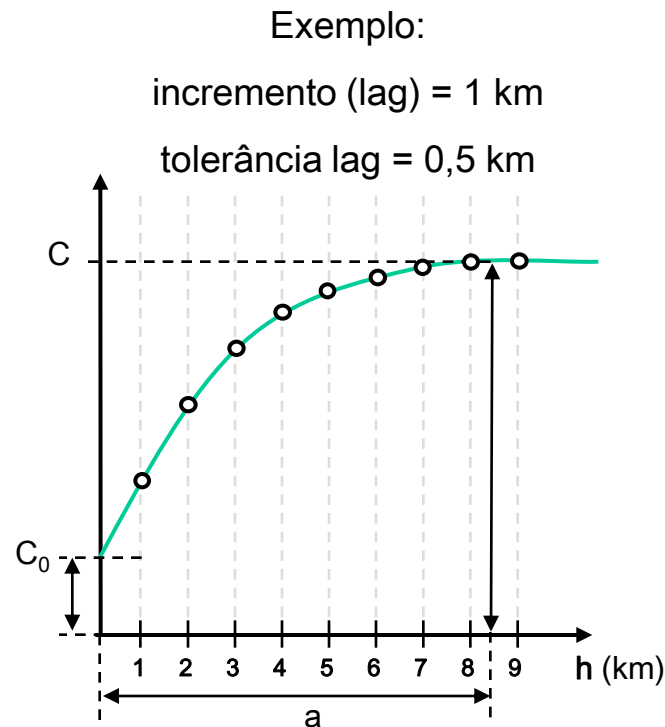
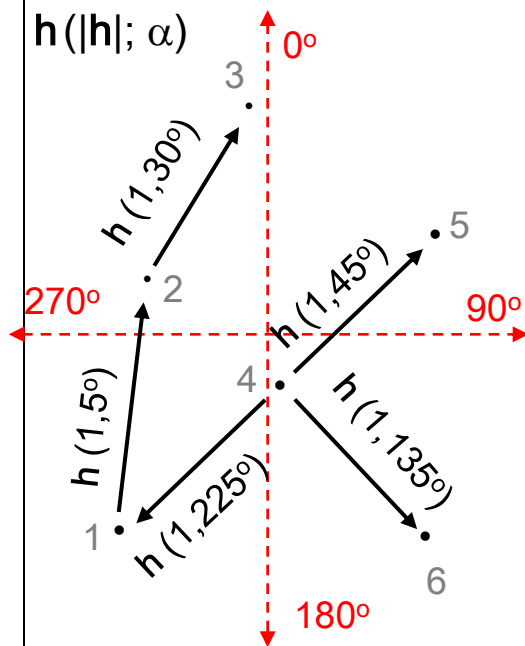
Semivariograma $\gamma(h)$

Cálculo do semivariograma a partir de amostras irregularmente espaçadas.

Semivariograma omnidirecional => tolerância angular = 90°

direção de análise (do vetor h) não importa.

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

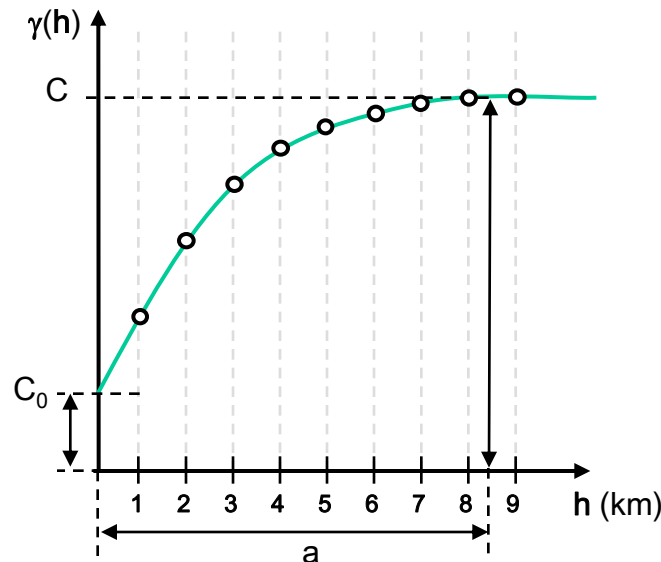
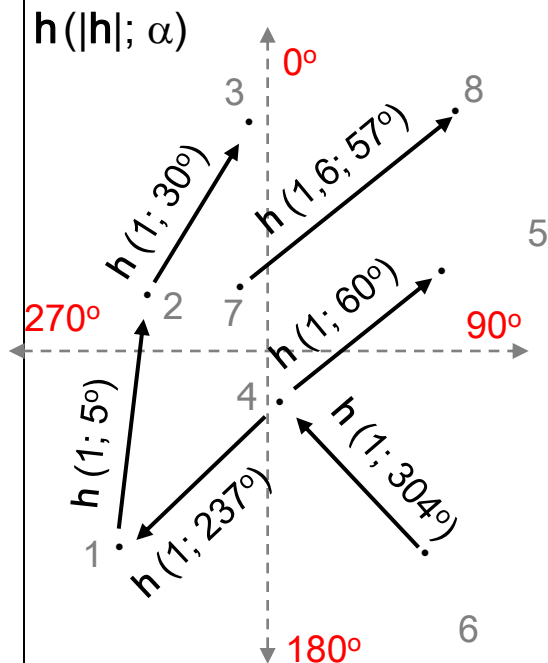
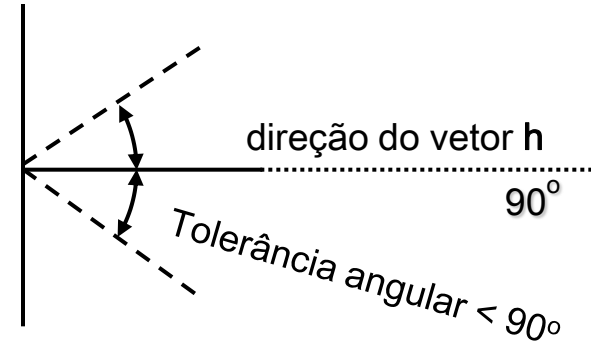


Semivariograma $\gamma(h)$

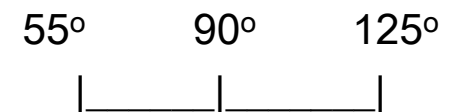
Cálculo do semivariograma a partir de amostras irregularmente espaçadas.

Semivariograma direcional $\Rightarrow$ tolerância angular $< 90^\circ$

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

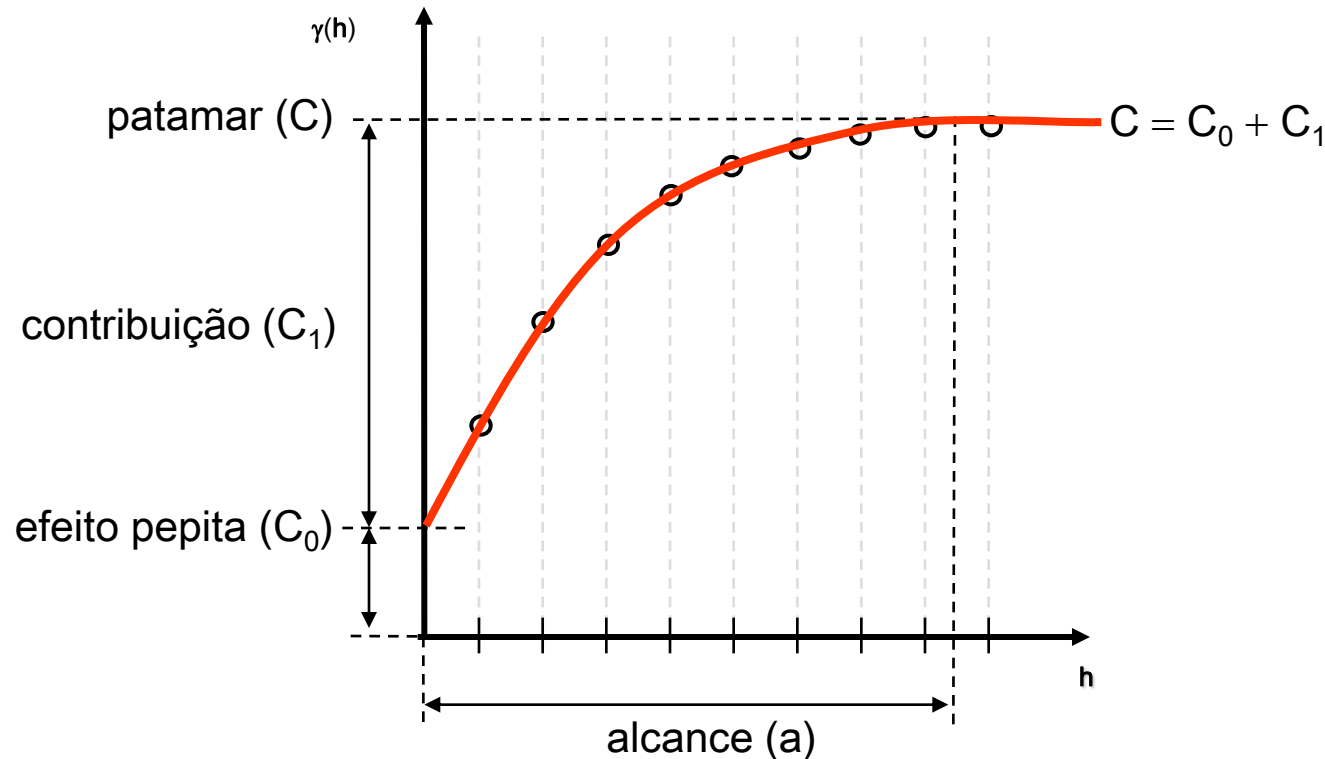


Exemplo:
 incremento (lag) = 1 km
 tolerância lag = 0,5 km
 direção de análise = 90°
 tolerância angular = 35°



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.



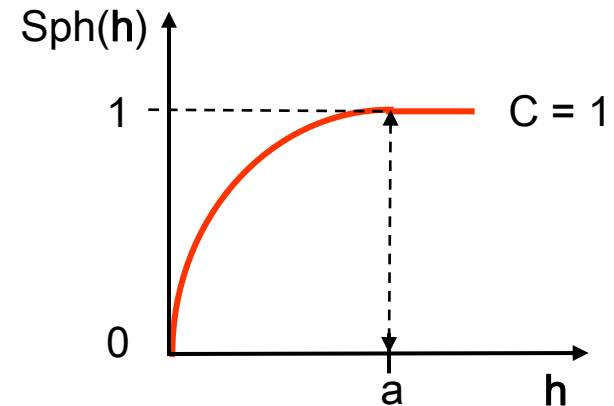
O modelo de ajuste deve representar o melhor possível o comportamento de $\gamma(h)$.

Modelos teóricos de semivariograma

Modelo de ajuste esférico

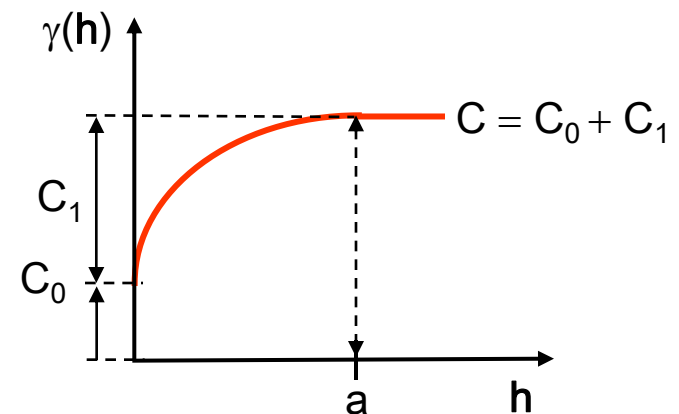
- Normalizado

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



- Na prática: $C_0 > 0$ e $C_1 > 1$

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

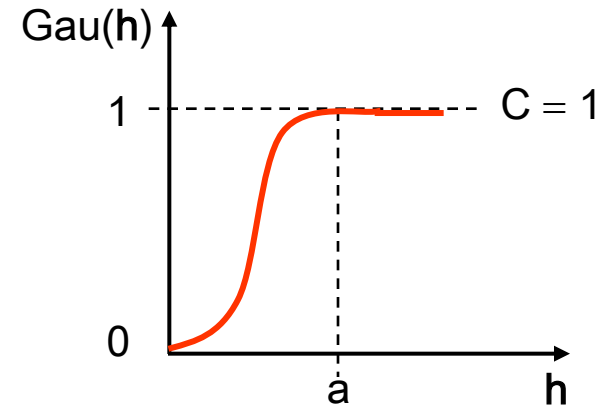


Modelos teóricos de semivariograma

Modelo de ajuste gaussiano

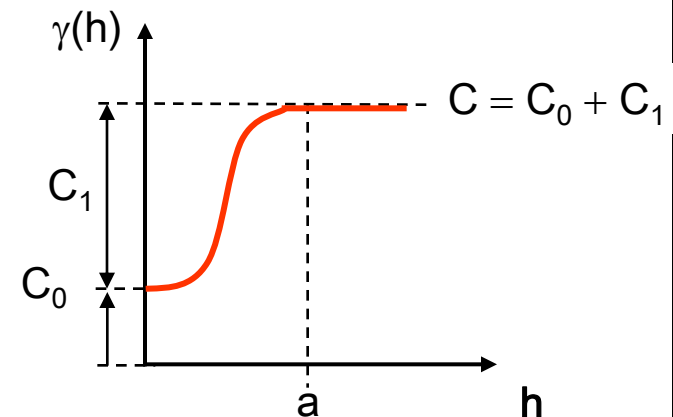
- Normalizado

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



- Na prática: $C_0 > 0$ e $C_1 > 1$

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

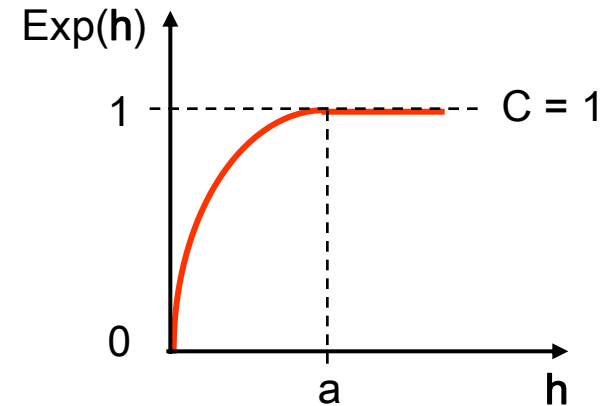


Modelos teóricos de semivariograma

Modelo de ajuste exponencial

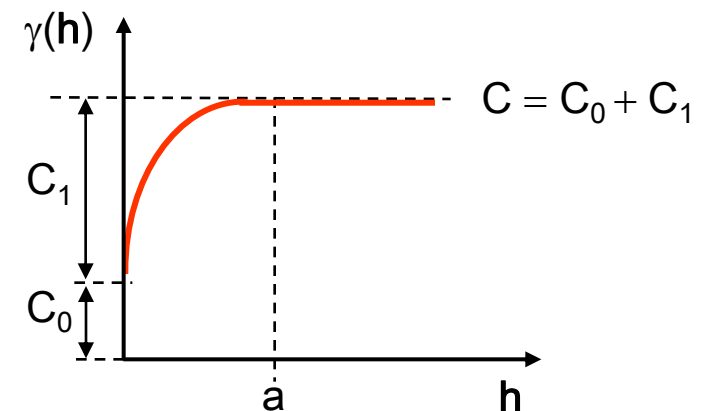
- Normalizado

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



- Na prática: $C_0 > 0$ e $C_1 > 1$

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

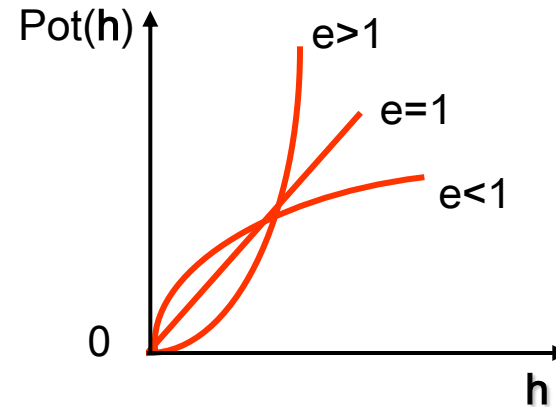


Modelos teóricos de semivariograma

Modelo de ajuste potência

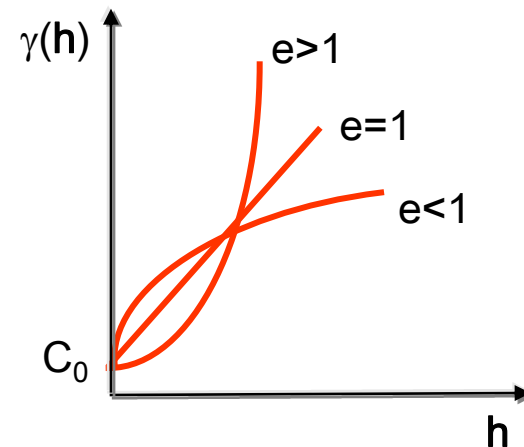
- Normalizado

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



- Na prática: $C_0 > 0$ e $C_1 > 1$

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



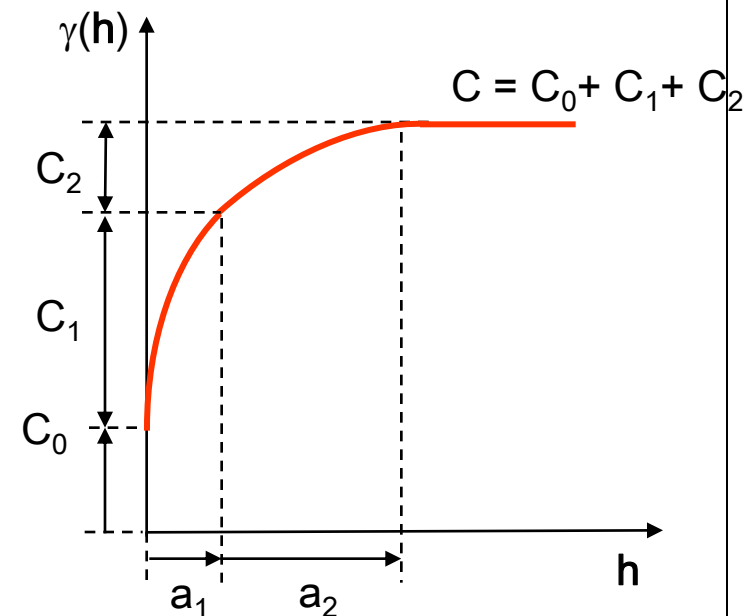
Modelos teóricos de semivariograma

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*.

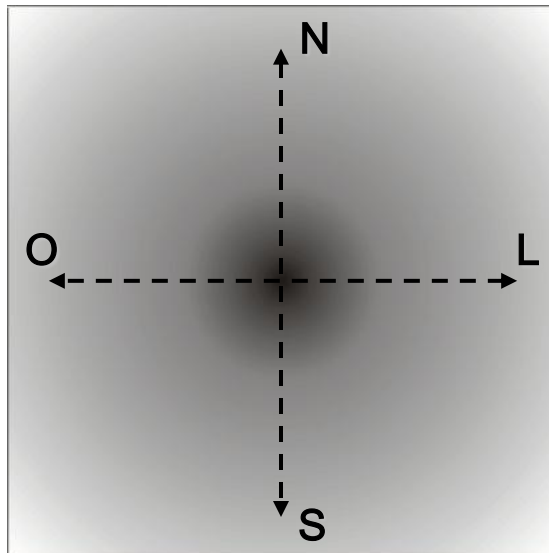
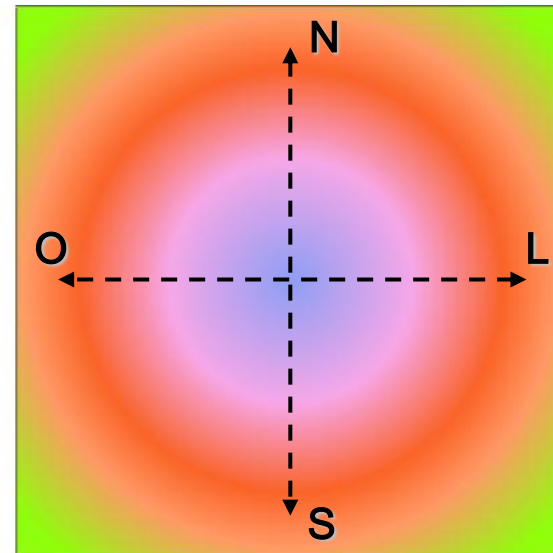


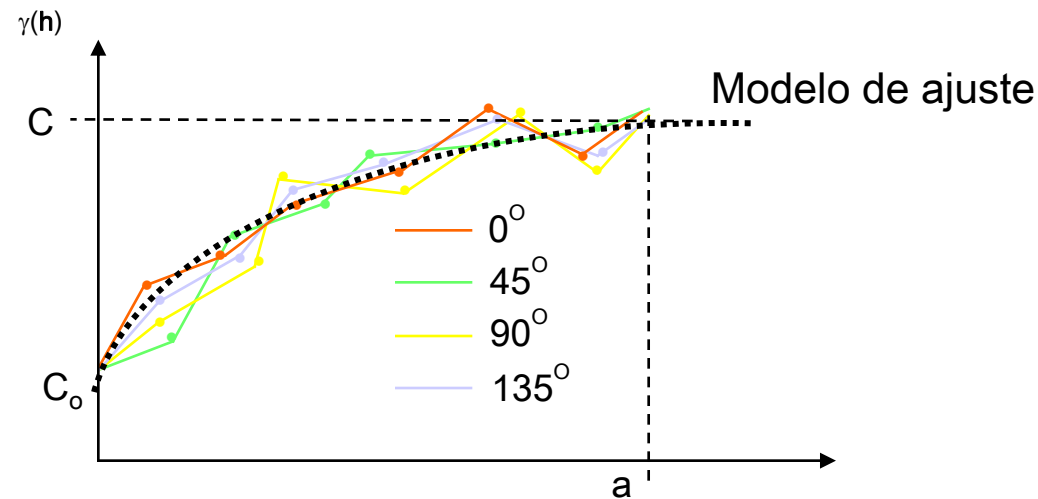
Imagem nível de cinza



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*.

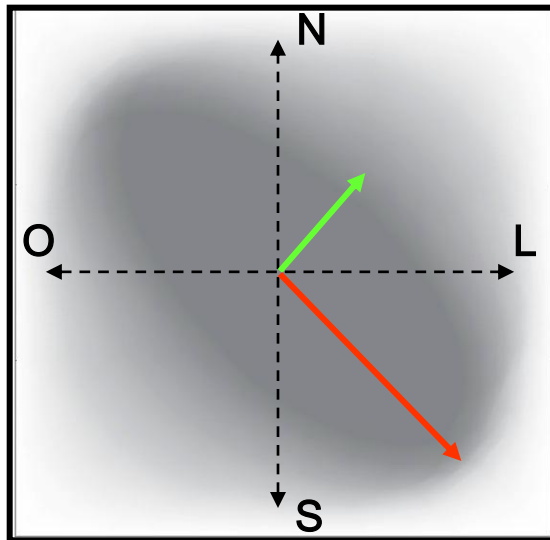
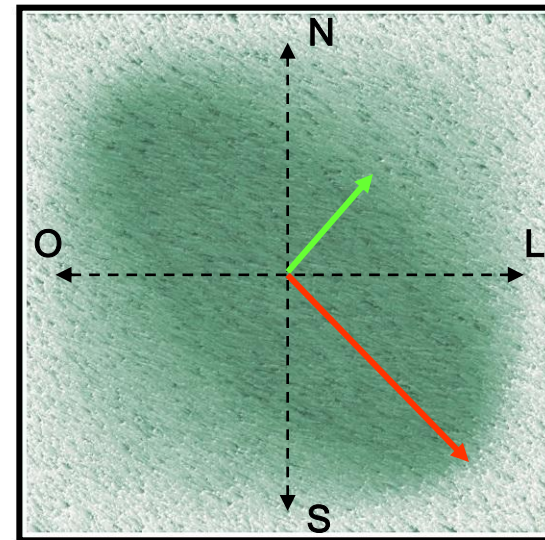


Imagem nível de cinza



Composição Colorida

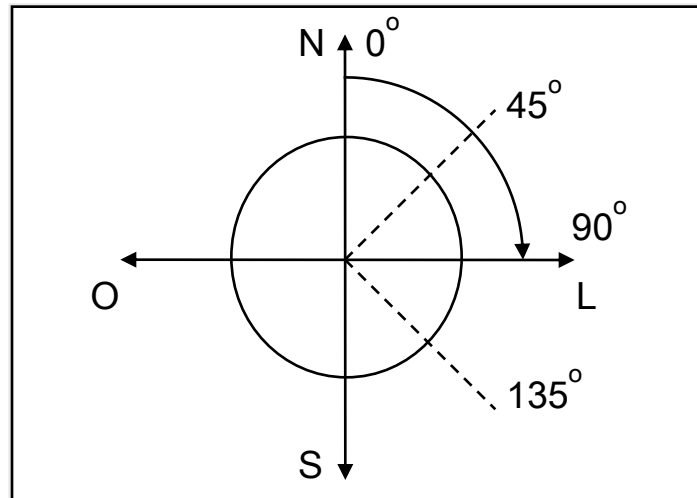
direções de continuidade espacial

— maior — menor

Anisotropia

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

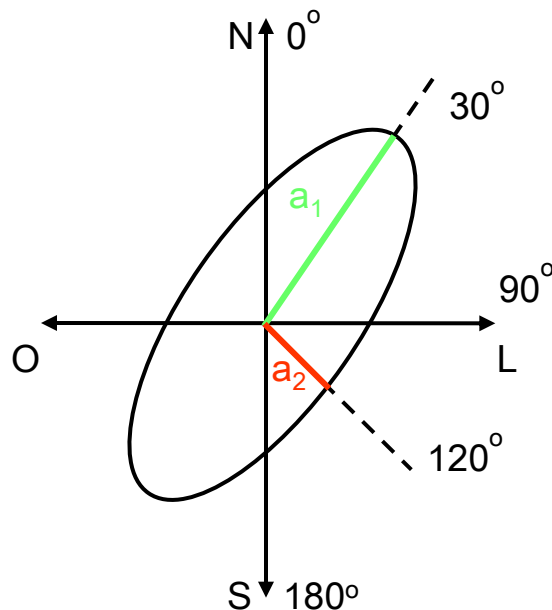
Convenções direcionais usadas na geoestatística



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 (F_a)

$$F_a = a_2 / a_1$$

Ângulo de anisotropia (A_a)

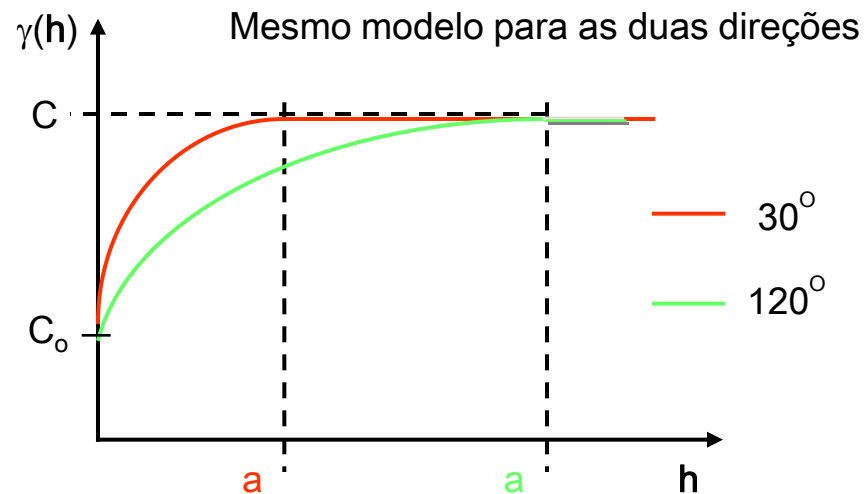
A_a = tomado da direção Norte para o eixo de maior continuidade. No exemplo = 30° .

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.

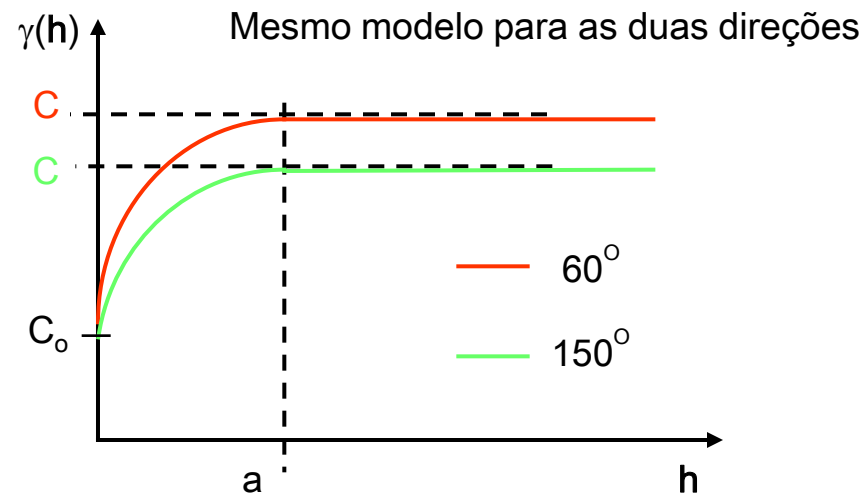


Anisotropia

Anisotropia zonal

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

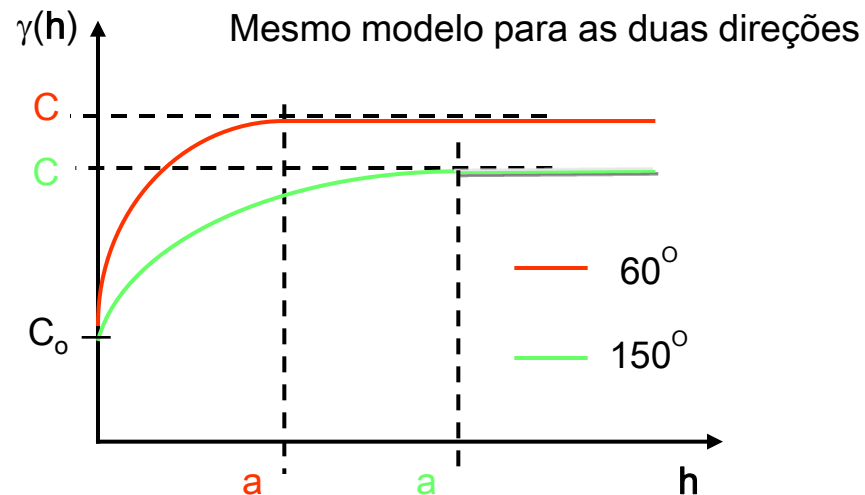
Como a isotropia, a **anisotropia zonal** é um caso menos frequente presente nos fenômenos naturais.



Anisotropia

Anisotropia combinada (geométrica + zonal)

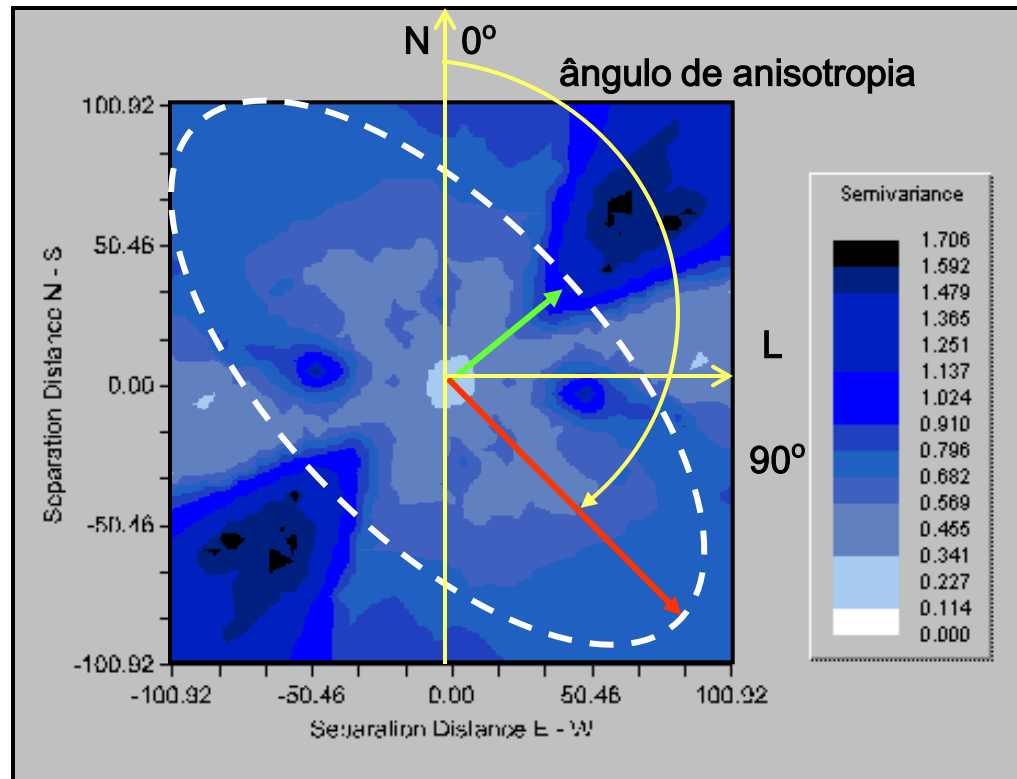
Neste caso, os semivariogramas apresentam diferentes patamares (C) e diferentes alcances (a) para o mesmo modelo. Pode apresentar também diferentes efeitos pepita.



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*.

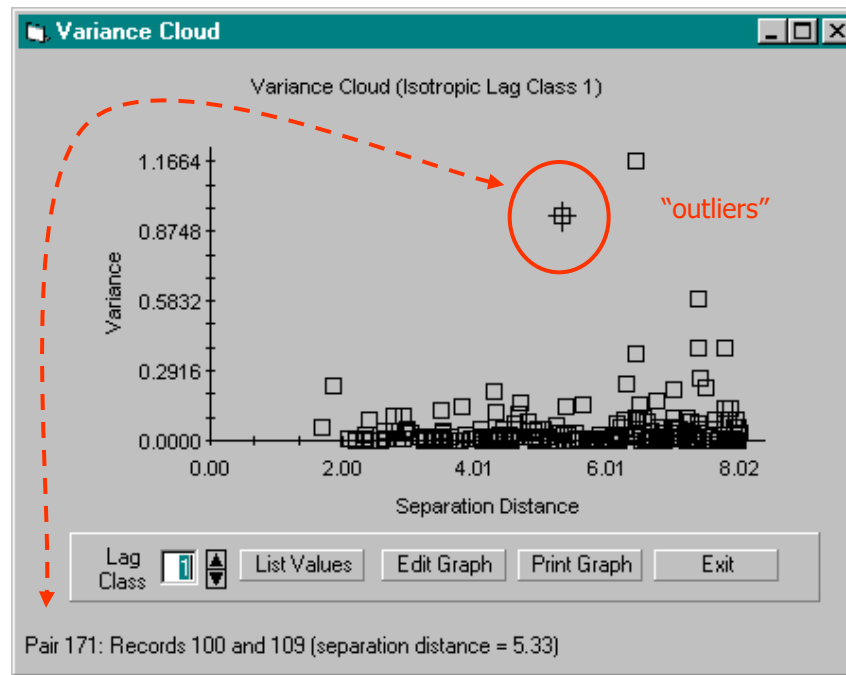
Utilizado para detectar os eixos de *Anisotropia* (direções de maior e menor continuidade espacial).



Semivariograma de nuvem

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

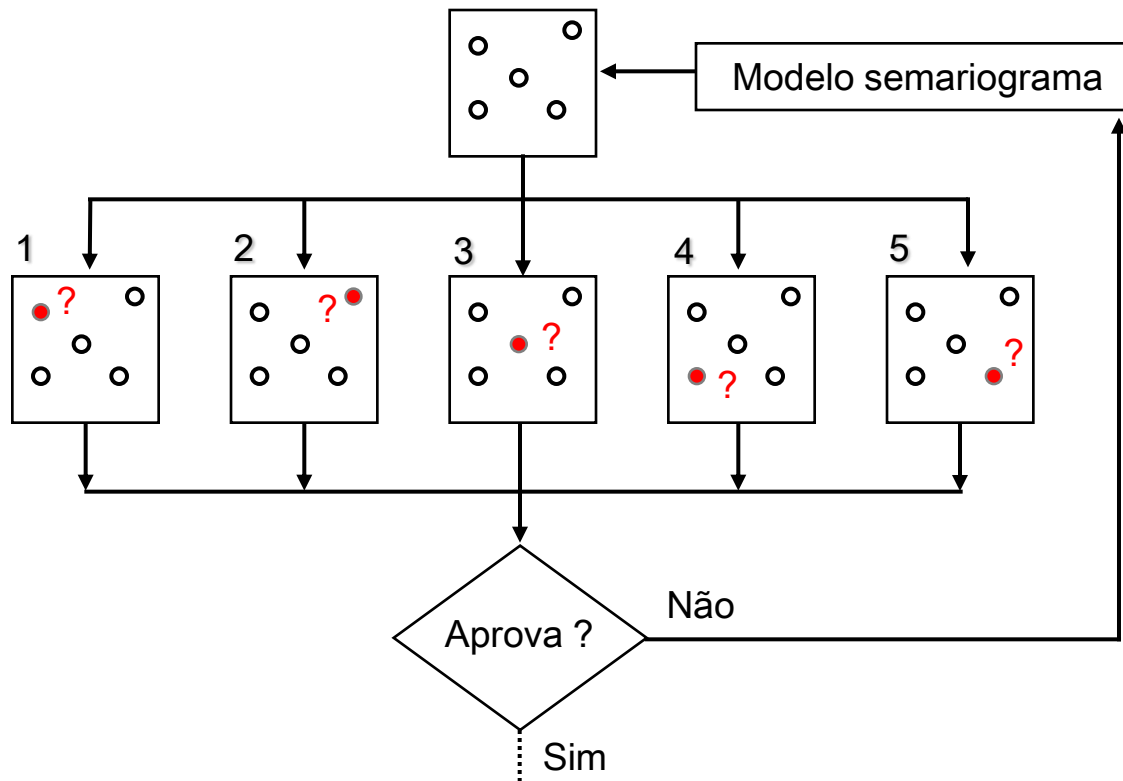
O variograma de nuvem é útil para detectar a presença de “outliers”.



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

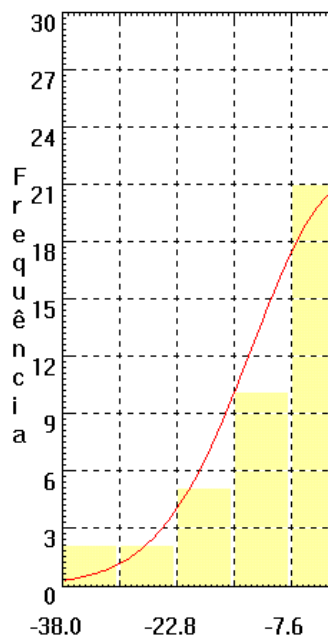
Análise de resultados



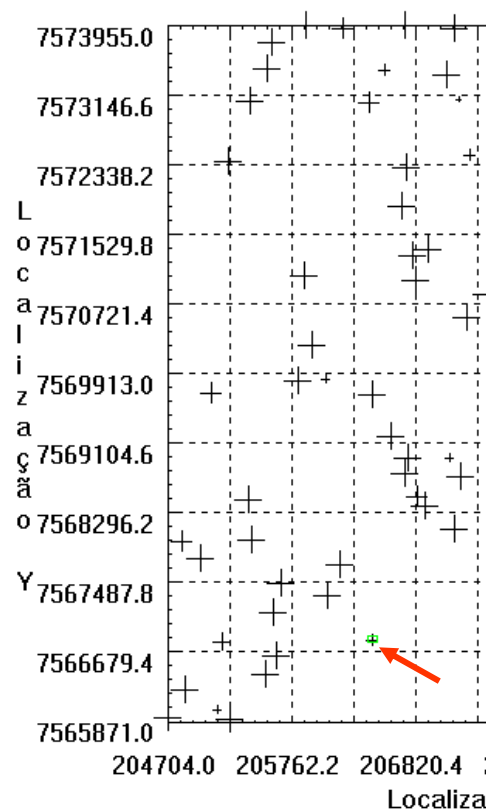
Histograma do Erro

Exibir Executar

Número de Classes: 10



Distribuição espacial do erro



N.pontos=85,Reestimados=85,Local.X=206455.094,Y=

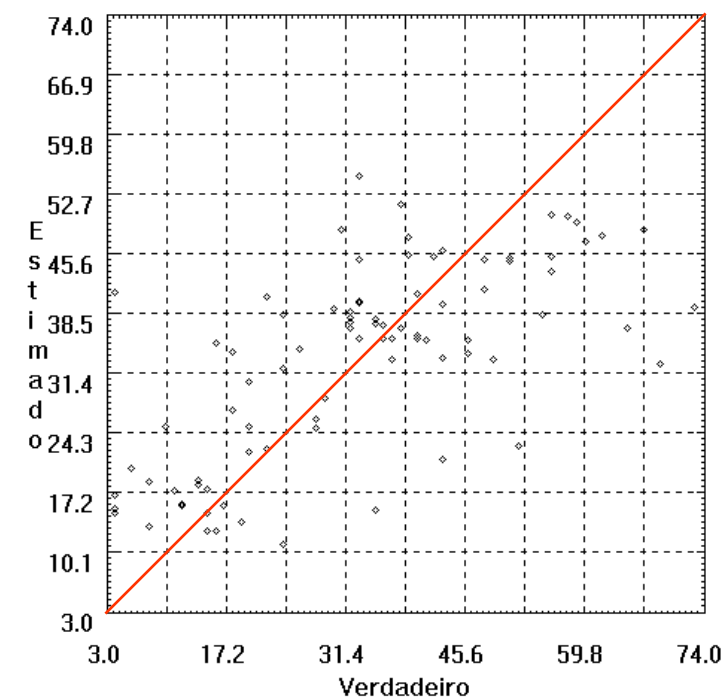
Relatório de Dados

CAS DO ERRO

argila

.....	85
.....	-0.252
.....	153.789
.....	12.401

Diagrama Verdadeiros X Estimados



N.pontos=85,Reestimados=85,Verd.=35.837,Estim.=35.483

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

- krigagem Simples (*)
- krigagem Ordinária (*)
- krigagem Universal
- co-krigagem
- krigagem 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**.

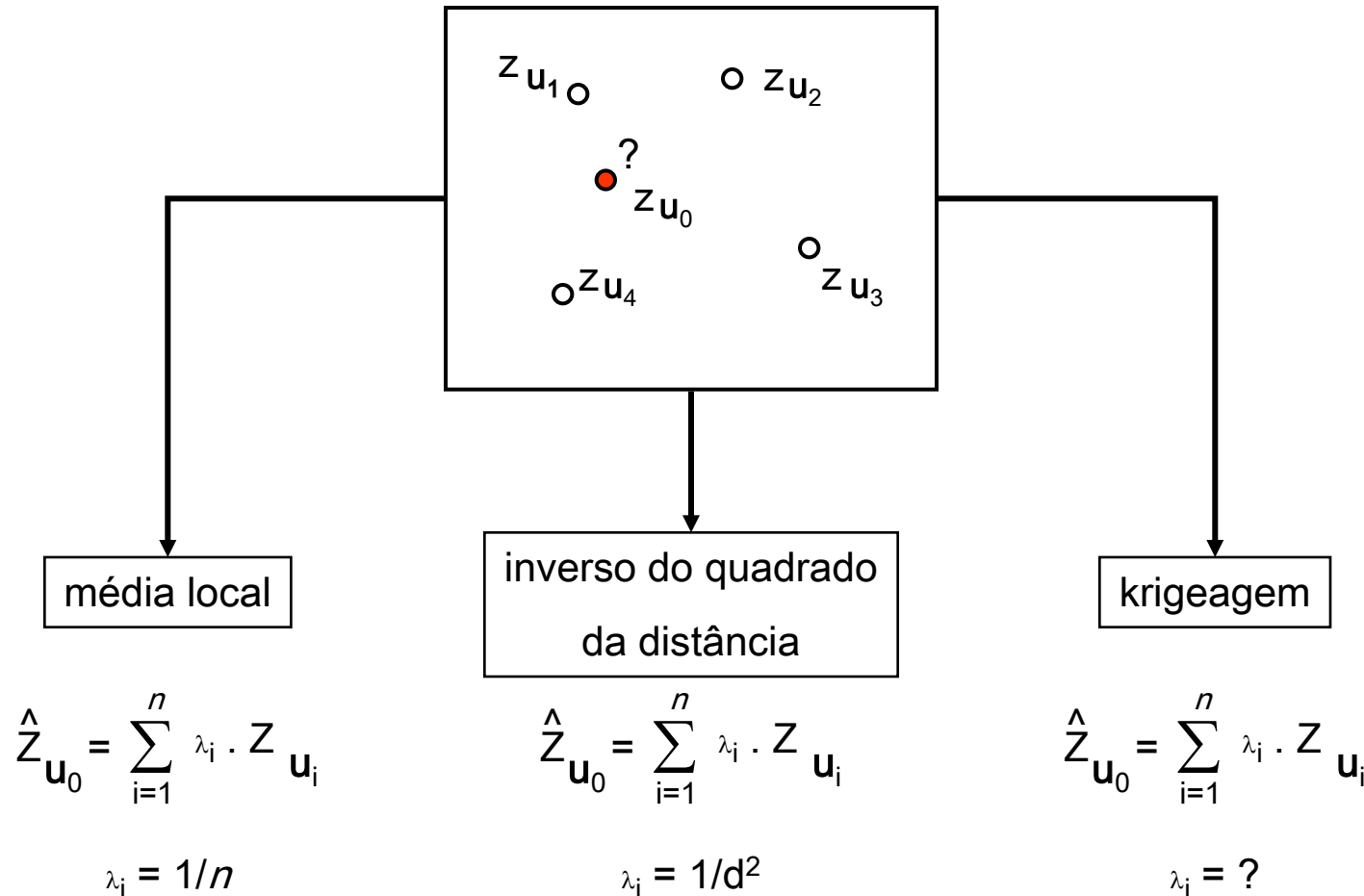
Além disso, a krigagem fornece, em média, estimativas não tendenciosas e com variância mínima

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.

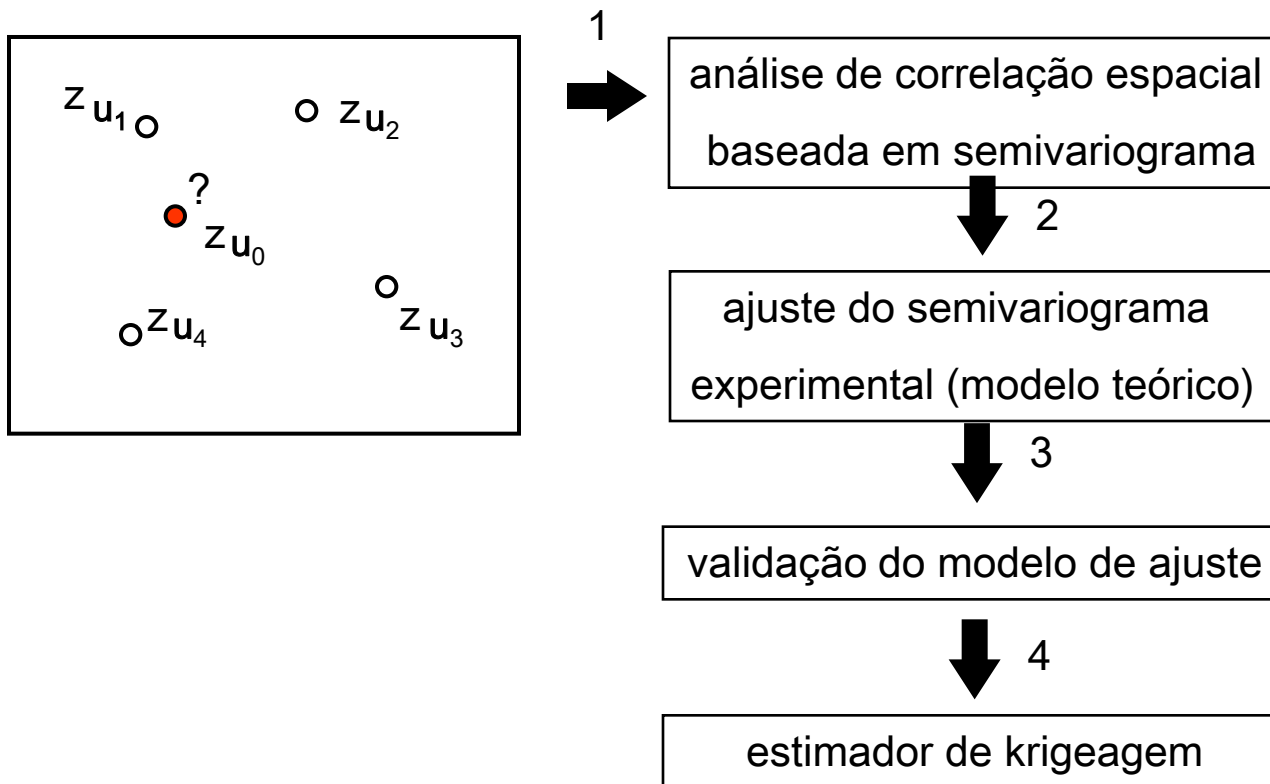
Krigeagem

Envolve uma combinação linear de n valores em pontos vizinhos.



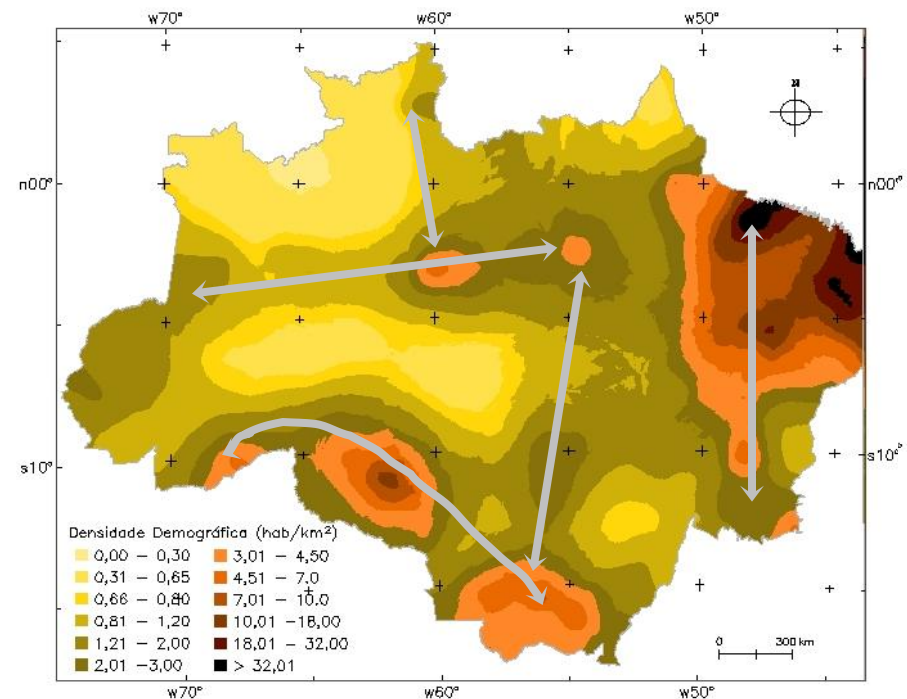
Krigeagem

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



Superfícies de Densidade - Krigeagem

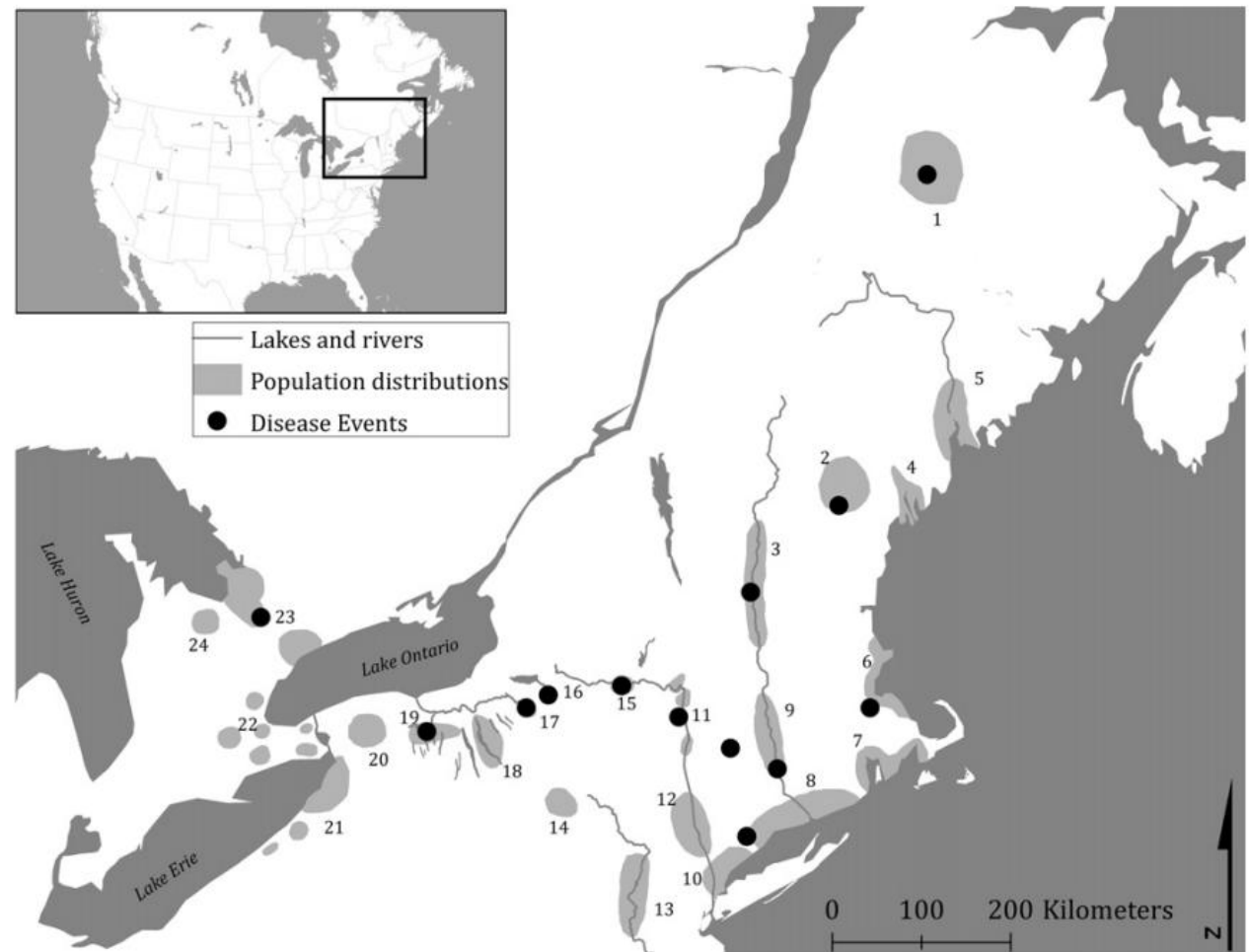
- Densidade de População Total – sedes de município
- Das superfícies-> Mapas temáticos para análise
- Mapas de tendências, imprecisos e generalizados
- Evolução das superfícies - > tendências demográficas



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

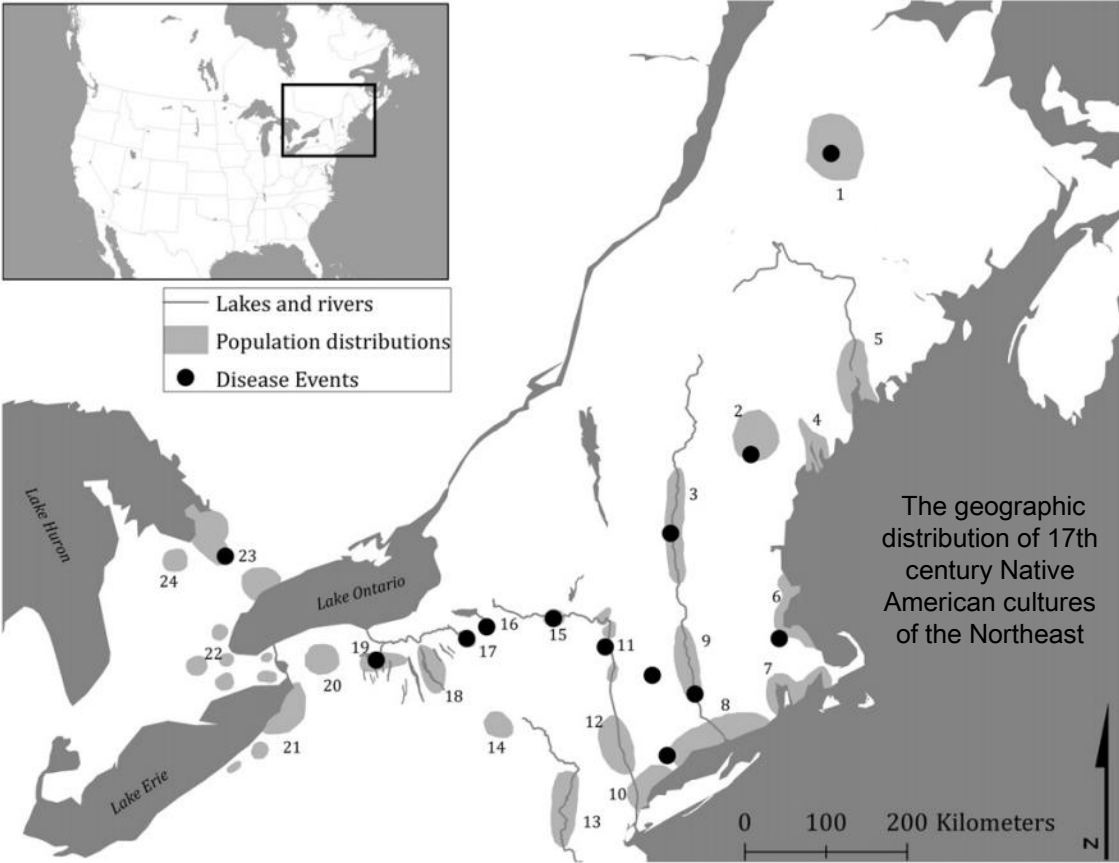


Table 2
Population data for the 17th-century Native American populations used in this study, including the year of the depopulation event being studied, beginning and ending populations for the depopulation event during the indicated year, percentage of total population that died, and sources.

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:

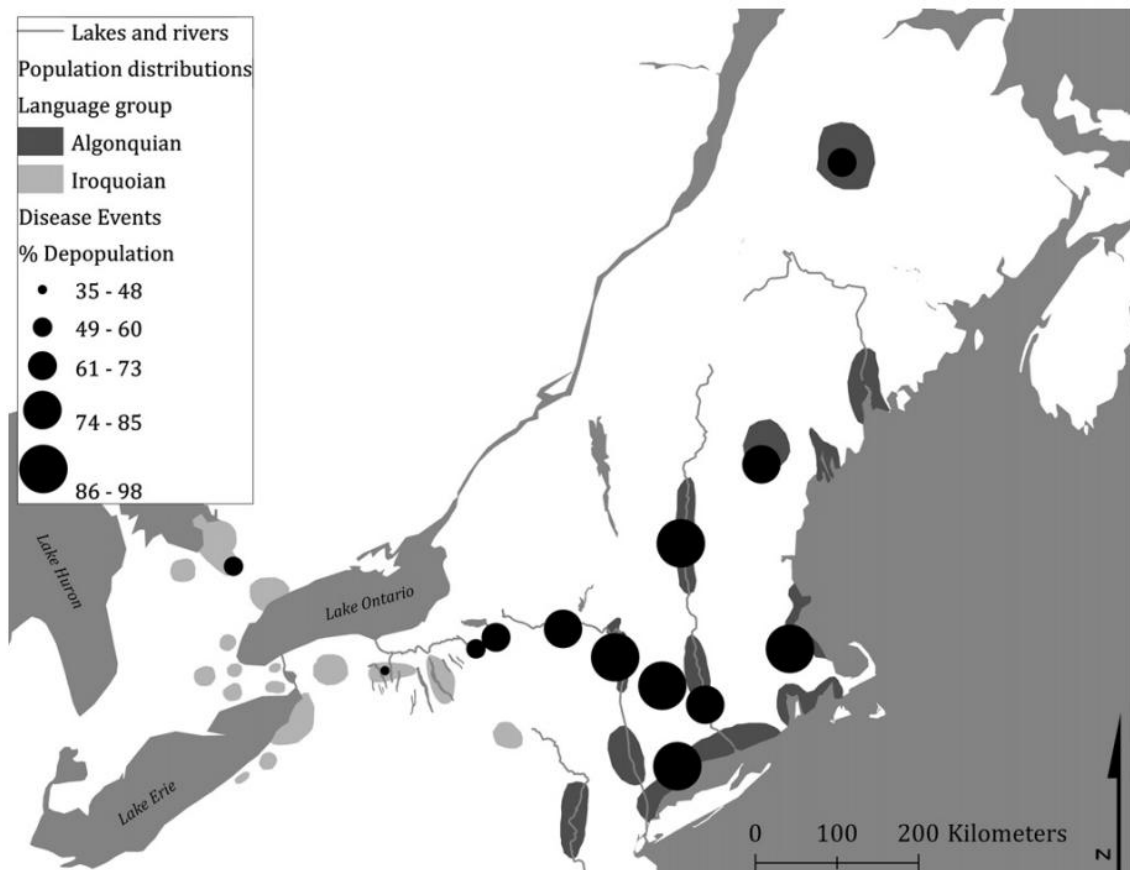
Despovoamento (*depopulation*): % perda entre estimativas pré e pós contato (valores da literatura)

- 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

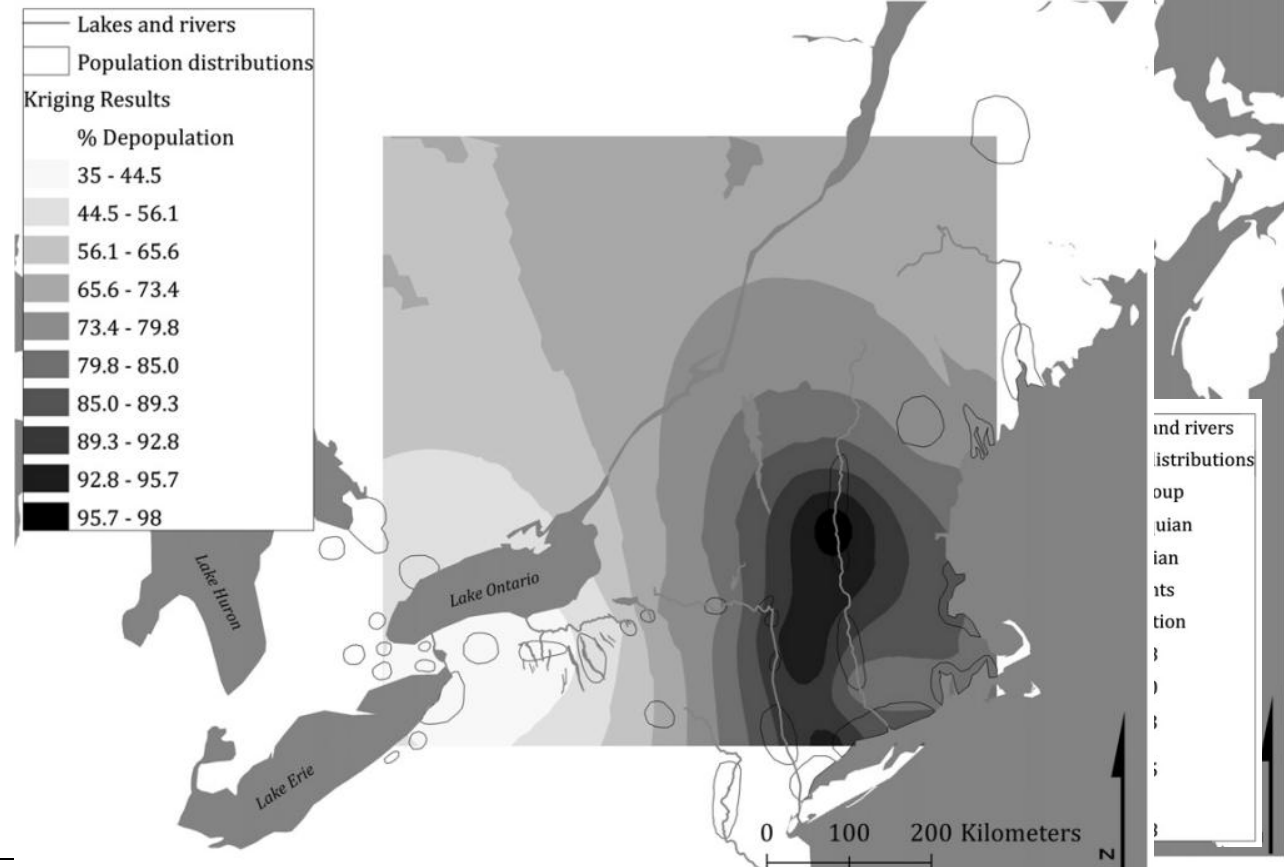
Sobreposição do mapa de distribuição geográfica das culturas nativas do sec 17 com a superfície de % de *depopulation*

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.



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