

CST-316
2011

Introdução à Análise Espacial

Laércio M. Namikawa
Jussara de Oliveira Ortiz



Waldo Tobler

Primeira lei da geografia

- Todas as coisas são parecidas, mas coisas próximas se parecem mais que coisas mais distantes

Introdução

- Sistemas de Informação Geográfica (SIG ou GIS) são as ferramentas computacionais usadas para Geoprocessamento
 - Característica forte é a capacidade de integrar dados geográficos vindos de diversas fontes em uma mesma base
- O SIG deve oferecer o conjunto mais amplo possível de estruturas de dados e algoritmos capazes de representar a grande diversidade de concepções do espaço
 - Do Ecólogo, do Sociólogo, do Geólogo...

O que há de especial com dados espaciais?

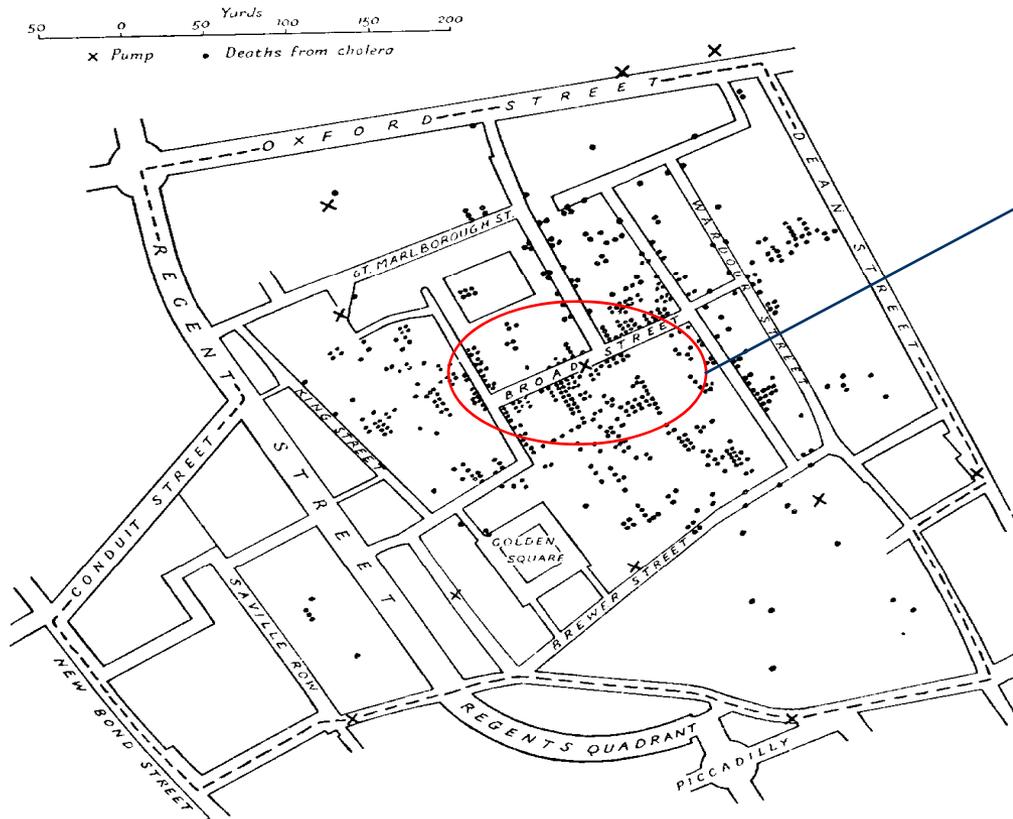
- Alguns tipos de análise que se deseja fazer sobre dados espaciais
 - Localização: Onde está...?
 - Quais as áreas com declividade acima de 20%?
 - Condição: Qual é...?
 - Qual a população desta cidade?
 - Tendência: O que mudou...?
 - Esta terra era produtiva há cinco anos atrás?

O que há de especial com dados espaciais?

- Roteamento: Qual o melhor caminho...?
 - Qual o melhor caminho para a linha do metrô?
- Padrões: Qual o padrão...?
 - Qual a distribuição da dengue em Fortaleza?
- Modelos: O que acontece se...?
 - O que acontece com o clima se desmatarmos a Amazônia?

O que há de especial com dados espaciais?

- Exemplo pioneiro e clássico do uso de informações espacializadas para tomada de decisão. Epidemia de cólera, em Londres, 1854



Fechamento do poço diminuiu a epidemia e comprovou a hipótese da transmissão pela água

O que há de especial com dados espaciais?

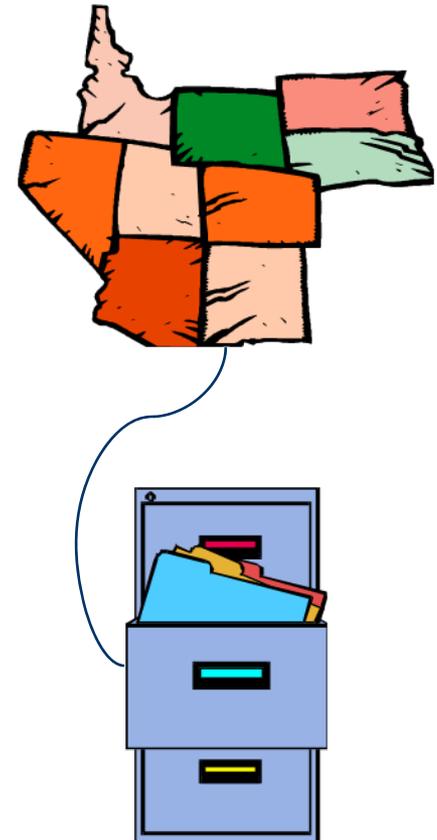
- Dados geográficos possuem uma natureza dual: localização geográfica e atributos descritivos

Localização geográfica

- representação da superfície terrestre
- processos no espaço geográfico

Atributos descritivos

- representação num banco de dados
- descrição do fenómeno



O que há de especial com dados espaciais?

- Espaço Geográfico
 - coleção de localizações sobre a superfície da Terra
 - define-se por suas coordenadas, altitude e posição relativa, o que o torna possível de ser cartografado
- Informação espacial
 - existência de objetos com propriedades: localização e relação com outros objetos
 - as relações entre os objetos podem ser topológicas, métricas e direcionais

O que há de especial com dados espaciais?

- Quando os **dados espaciais** são **organizados**, **analisados**, **interpretados** e **apresentados** de forma útil para um problema de **decisão** específico, se transformam em **informação espacial**.

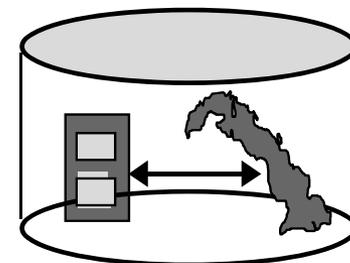


Geoprocessamento - SIG

Geoprocessamento

Existem diferentes visões sobre geoprocessamento

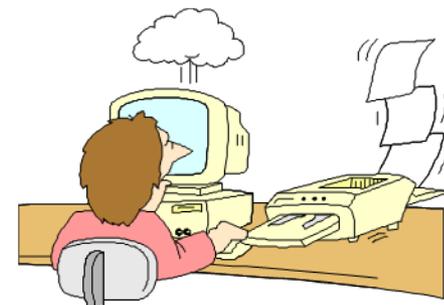
- Banco de Dados: armazenamento e recuperação de informação espacial



- Produção de Mapas: geração e visualização de dados espaciais



- Análise Espacial: combinação de informações espaciais



Análise Espacial - Introdução

Apresentar algumas técnicas de análise espacial que são empregadas a classes distintas de problemas. Exemplos:

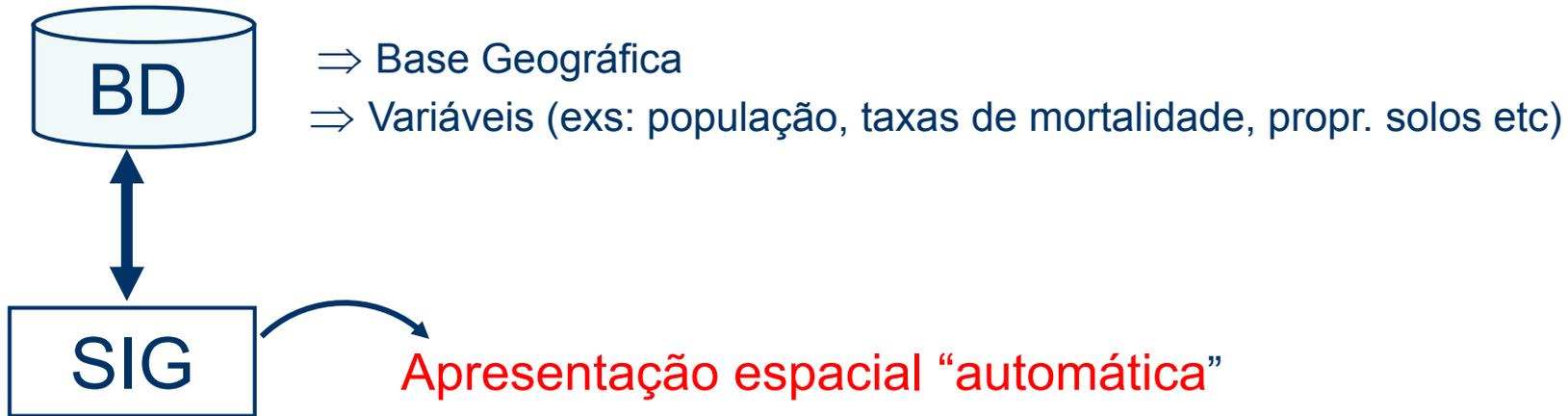
- saúde
- pedologia
- segurança pública
- biologia
- entre tantas outras...

Todos esses problemas fazem parte da análise espacial de dados geográficos.

ênfase: mensurar propriedades e relacionamentos, levando em conta a localização espacial do fenômeno em estudo de forma explícita. Ou seja, a idéia central é incorporar o espaço à análise que se deseja fazer.

Inferência: qual o mecanismo de funcionamento

Introdução à Análise Espacial



Mapas: temáticos e pontuais



A QUESTÃO É:

Grande parte dos usuários limita seu uso em SIG a essas operações de visualização, tirando conclusões intuitivas.

Introdução à Análise Espacial

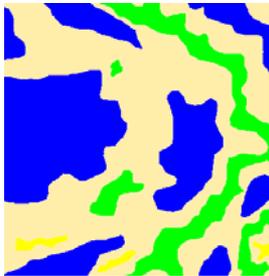
- É possível ir muito além...
- Quando visualizamos um padrão espacial, é muito útil traduzi-lo em considerações objetivas, como:
 - O padrão que observamos é aleatório ou apresenta uma agregação definida?
 - Esta distribuição pode ser associada a causas mensuráveis?
 - Como é o mecanismo de associação ?

Introdução à Análise Espacial

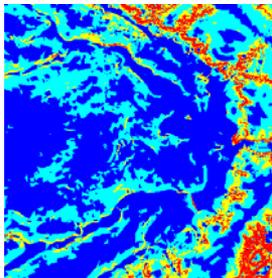
- Fenômeno espacial: **Discreto** ou **Contínuo**
- **Discreto (Objetos)** - espaço contém entidades do mundo real.
 - Exs: municípios, bairros, quadras, lotes, etc...
- **Contínuo (Campos)** - informação presente em todas as posições.
 - Exs: temperatura, pressão, teor de argila no solo, etc...

Análise espacial em SIG

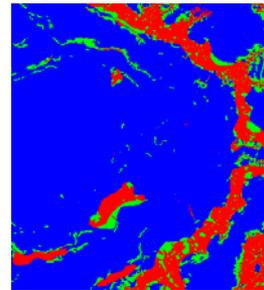
- Integram-se representações de atributos para se obter respostas combinadas, relatórios ou cenários que servirão de apoio a decisões sobre uma área em estudo



Mapa de Solos



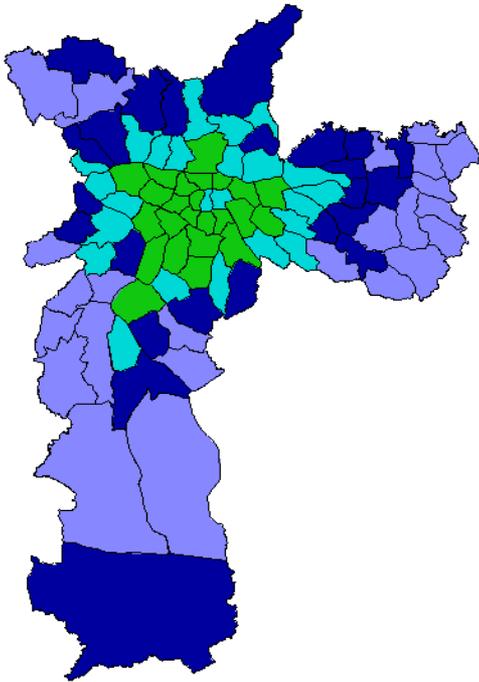
Mapa de Declividade



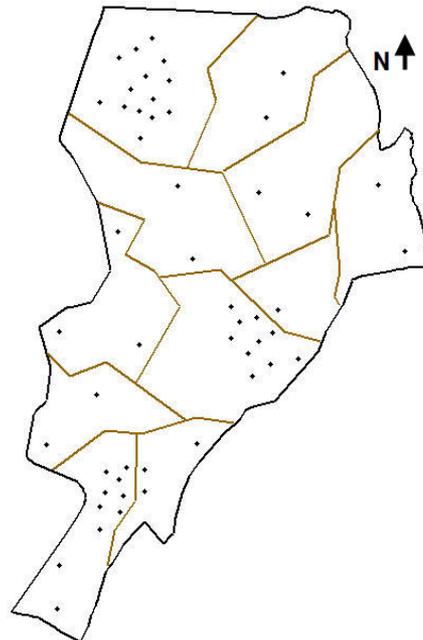
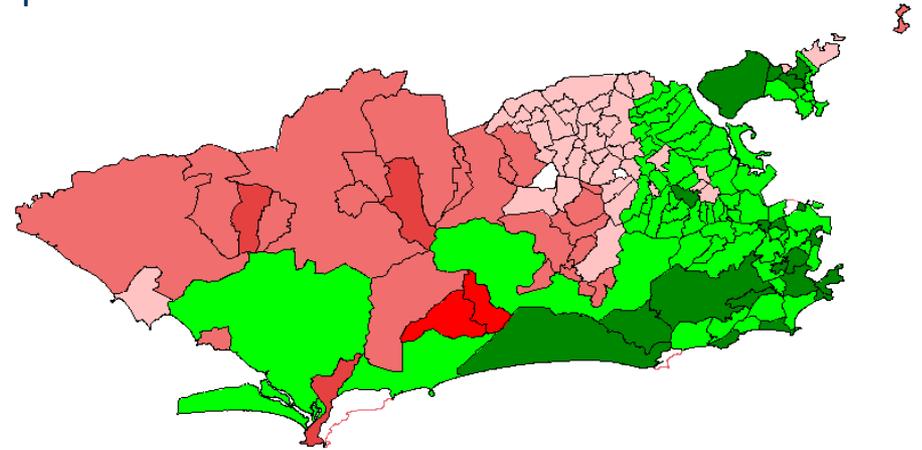
Mapa de Aptidão

Alguns Exemplos: Mapas Temáticos e Pontuais

A distribuição espacial do Percentual de Idosos por distritos na cidade São Paulo.



A distribuição espacial do Índice de Apgar por bairros na cidade do Rio de Janeiro.



Eventos de uma doença por setores de saúde em um município.

Análise Espacial - Técnicas

S Í N T E S E : C L A S S E S D E P R O B L E M A S

Tipos de Dados	Exemplos	Problemas Típicos	Técnicas
Eventos Localizados (Dados Pontuais)	Ocorrências de Crimes	Determinação de Padrões	Análise de Padrões Pontuais
Entidades e Atributos (Dados Zonais)	Dados Censitários	Relacionamento das entidades	Análise de Áreas
Amostras de Campo (Dados Contínuos)	Teor de Zinco no solo	Interpolação de Superfície	Métodos Geoestatísticos

Técnicas de ESDA (*Exploratory Spatial Data Analysis*)

- ESDA: “Coleção de técnicas para descrever e visualizar distribuições espaciais, identificar situações atípicas, descobrir padrões de associação espacial, clusters e sugerir regimes espaciais ou formas de heterogeneidade espacial” (Anselin).

1- Visualização de distribuição espacial

- técnicas convencionais de visualização cartográfica, estatísticas não-espaciais.

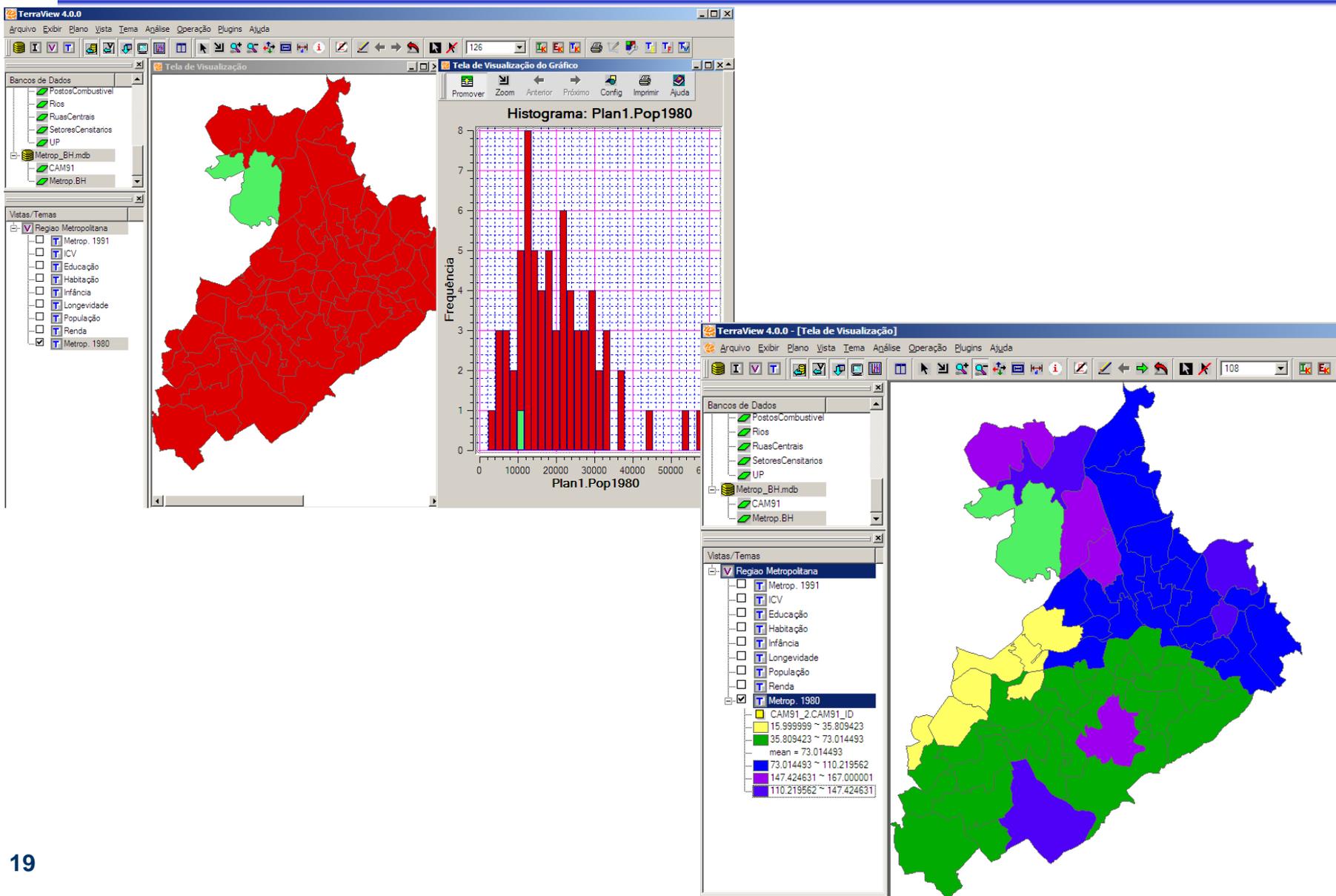
2- Indicadores Globais de Autocorrelação

- explorar a dependência espacial, mostrando como os valores estão correlacionados no espaço.
- O conceito utilizado é o de autocorrelação espacial.
- Ex. Indicadores Globais: Moran's I, Geary's C

3- Indicadores Locais de Associação Espacial (LISA)

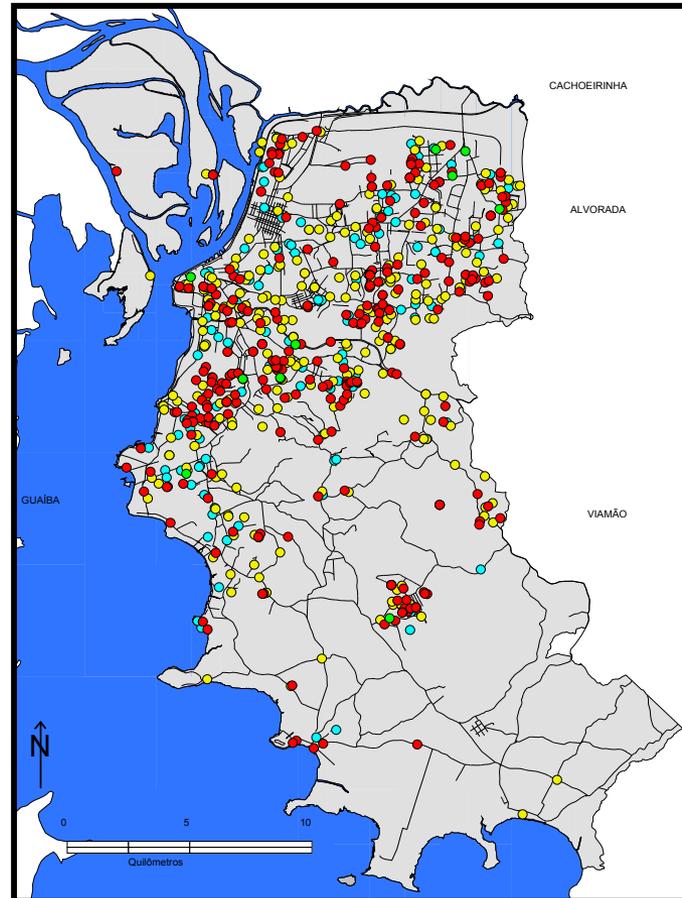
- Identificação de:
- “Clusters”: objetos com valores de atributos semelhantes,
 - “Outliers”: objetos anômalos,
 - A presença de mais de um regime espacial.
 - Ex. Indicadores Locais: Moran (I_i), Getis e Ord (G_i e G_i^*).

Visualização de distribuição espacial



Análise de Padrões de Distribuição de Pontos

- A violência que ocorrem em determinadas regiões formam um padrão no espaço? Estão correlacionados com características sócio econômicas?



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

Introdução- Análise de Padrões de Distribuição de Pontos

- Consideramos aqui fenômenos expressos através de ocorrências pontuais.

Representações pontuais podem corresponder a dados como:

- índices de mortalidades,
- ocorrências de doenças,
- localização de espécie vegetais etc.

- Objetivo
 - aumentar o entendimento do processo inferindo valores em locais sem observações.

ALGUNS EXEMPLOS

- Epidemiologia

- A distribuição dos casos de uma doença formam um padrão no espaço? Existe associação com alguma fonte de poluição? Evidência de contágio?

- Crime

- Roubos que ocorrem em determinadas áreas estão correlacionados com características sócio econômicas e/ou de infra estrutura básica?

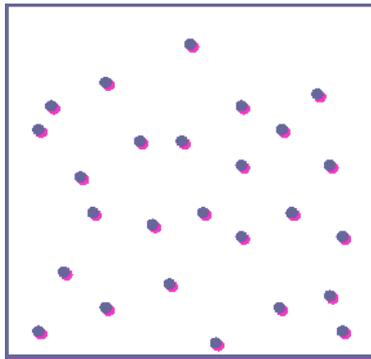
Distribuição de Pontos

- **Distribuição pontual:** é um conjunto de localizações pontuais, contidos numa área de estudo, que estão associados a eventos de interesses.

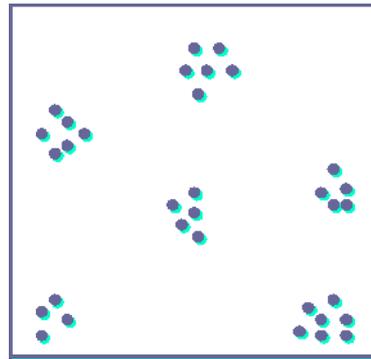


Distribuição de Pontos

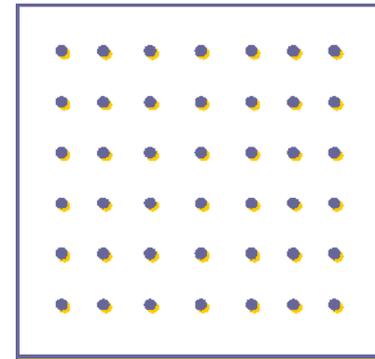
- Uma característica importante de um padrão espacial de pontos é a localização dos eventos e a relação entre eles (arranjos espaciais).
- Objetivo: verificar se os eventos observados apresentam algum tipo de padrão sistemático, ao invés de estarem distribuídos aleatoriamente.



Aleatório



Agrupado



Regular

- Na realidade o que se deseja é detectar padrões de aglomerados espaciais ("clusters").

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).
- ⇒ **Efeitos de Segunda Ordem**
 - denominados locais ou de pequena escala.
 - representam a dependência espacial no processo

Caracterização de Distribuição de Pontos

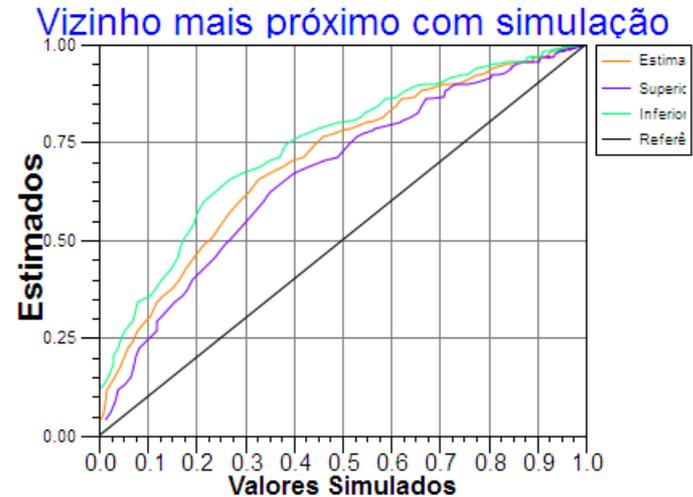
- Técnicas :
 - Para Efeitos de Primeira Ordem
 - Estimador de Intensidade (“Kernel Estimation”)
 - Para Efeitos de Segunda Ordem
 - Vizinho mais Próximo
 - Função K

Método - Vizinho mais Próximo

O método do vizinho mais próximo estima a função de distribuição cumulativa ($\hat{G}(w)$) baseado nas distâncias entre eventos em uma região de análise.

A plotagem dos resultados da função em relação as distâncias, pode ser utilizado como um **método exploratório** para verificar se existe evidência de interação entre os eventos.

Método - Vizinho mais Próximo



Crescimento rápido para pequenas distâncias → interação entre eventos → agrupamentos nestas escalas.

Pequenos valores no início → crescendo rápido para distâncias maiores → distribuição mais regular.

Aleatoriedade Espacial Completa → função próxima de uma linear a 45 graus.

Agrupamento → acima da reta de 45 graus.

Regularidade → abaixo da reta.



Função K

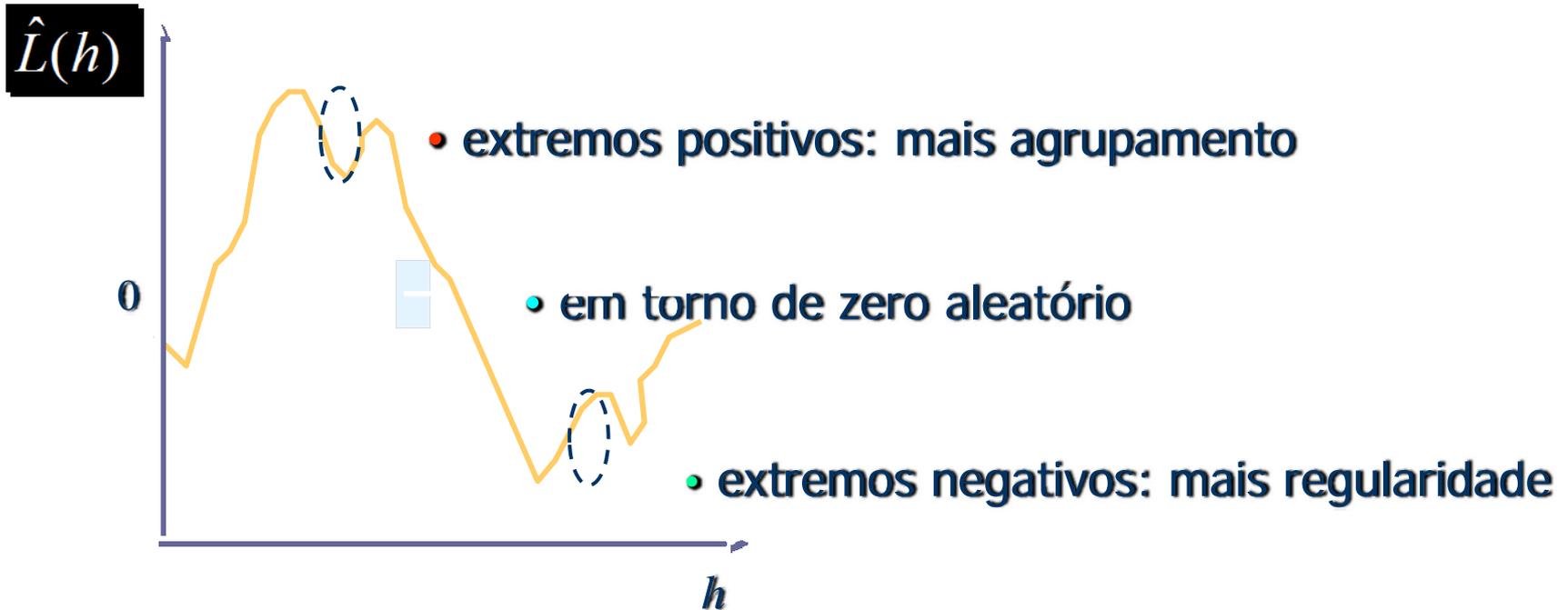
- Ao contrário do método do vizinho mais próximo, a função K considera escalas maiores para análise de padrões pontuais (Momento 2ª ordem).
- A função K se aplica a processos isotrópicos, isto é, supõe-se uma região de estudo homogênea.

Função K

Uma vez obtido, $\hat{K}(h)$ este pode ser plotado e examinado.

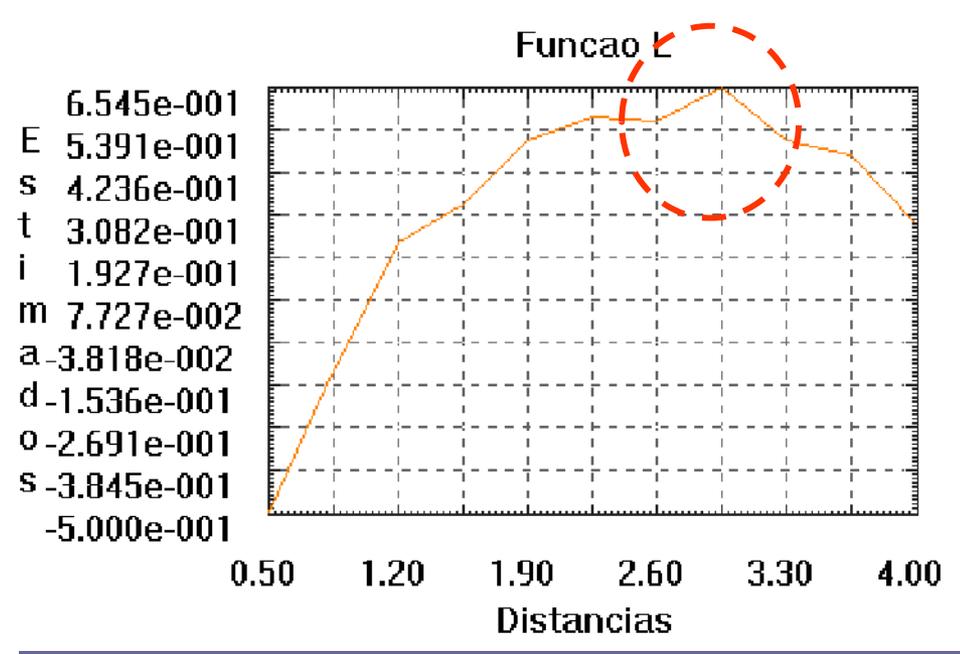
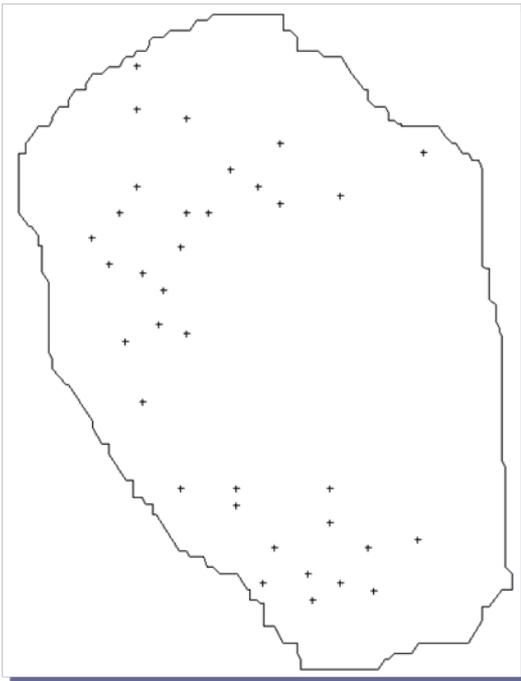
O gráfico da função K não é tão intuitivo quanto a do gráfico do vizinho mais próximo. Portanto, utiliza-se uma função auxiliar L, para facilitar a interpretação.

O estimador de L é $\hat{L}(h)$, e tem a seguinte interpretação da plotagem:



Função K

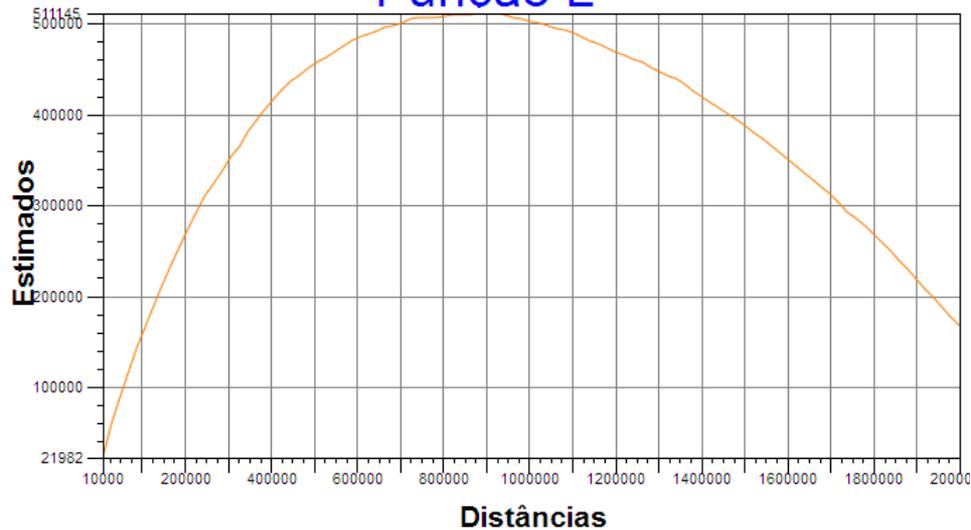
- Exemplo - município Bood Moor, situado no condado de Cornwall, Inglaterra. Dados geomorfológicos com 36 localizações de rochas de granito. Neste caso percebe-se que para distâncias entre aproximadamente 2.5 a 3 (extremo positivo do gráfico) há evidências de agrupamento.



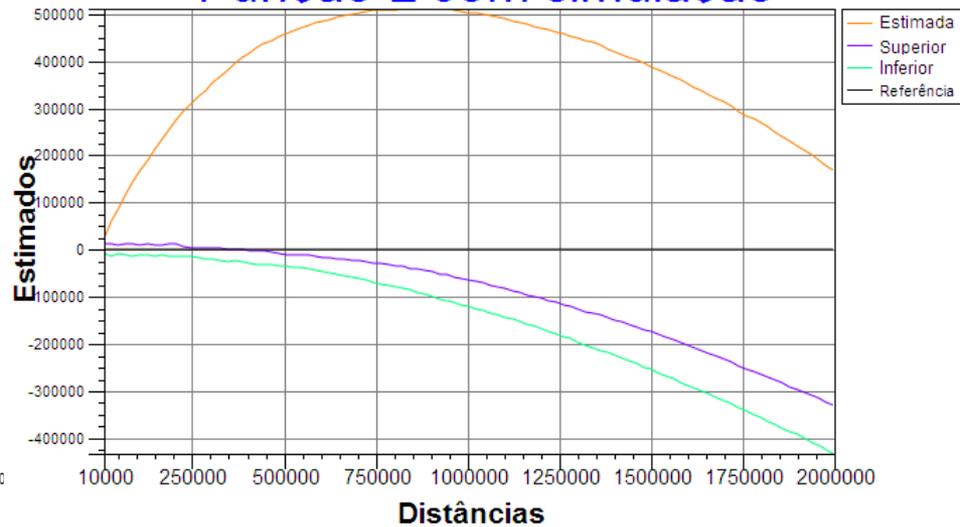
Distância: mín.=0.5 máx.=4 #intervalos=10

Função K

Função L



Função L com simulação

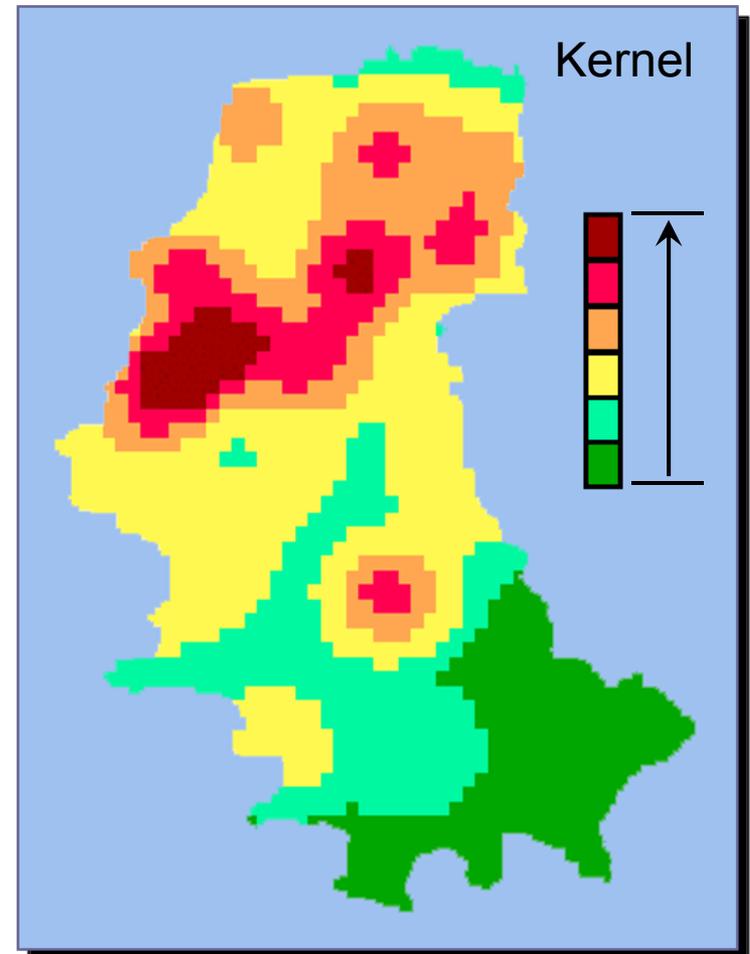
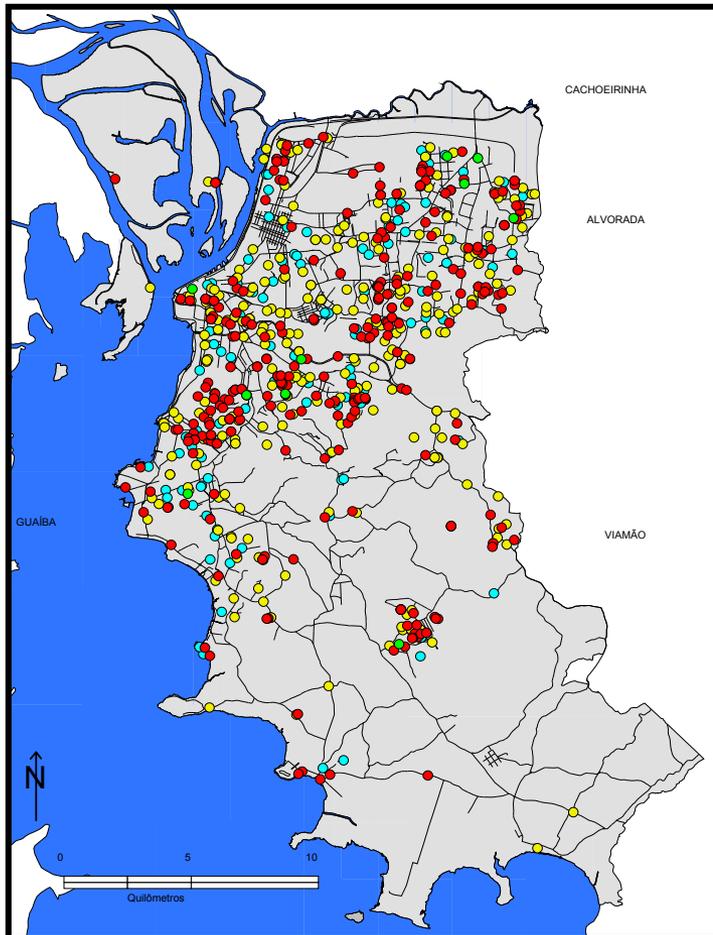


Agrupamento → valores positivos
Regularidade → valores negativos



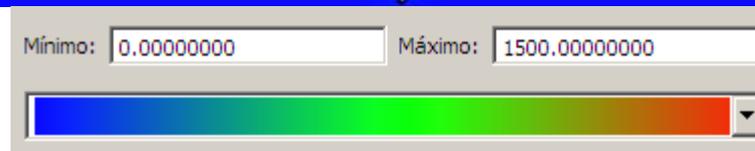
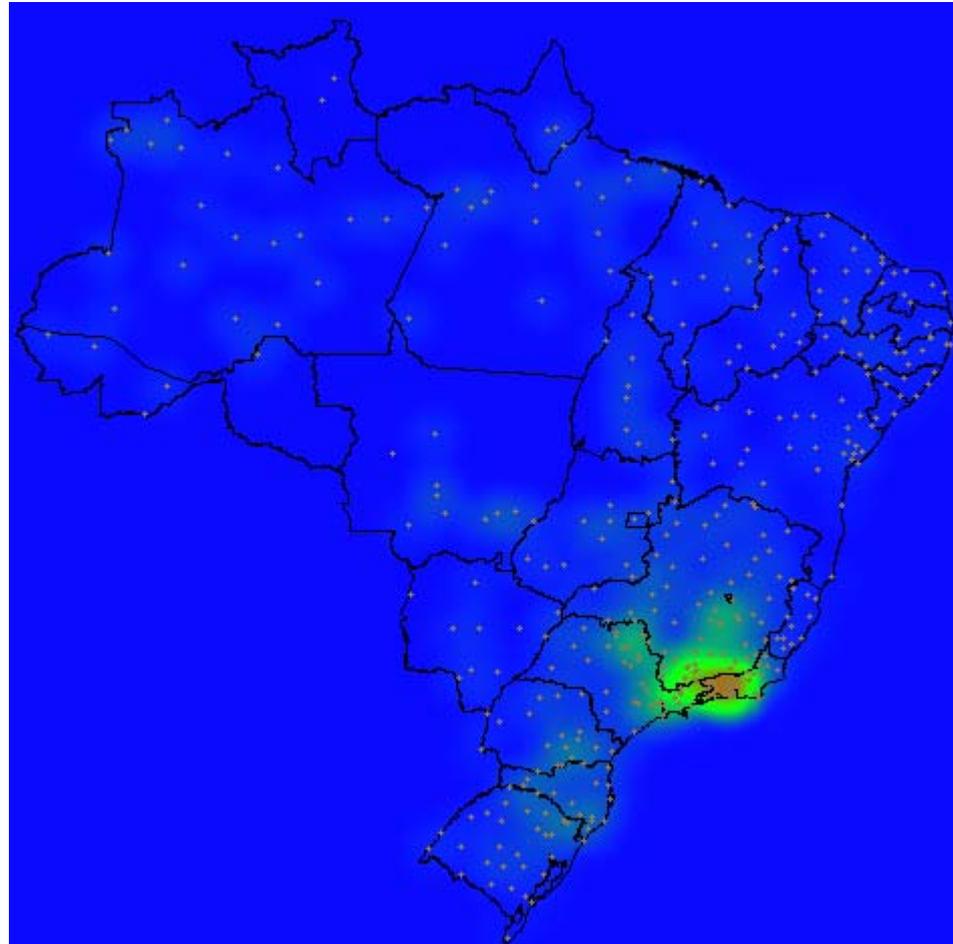
Estimador de Intensidade (“Kernel Estimation”)

- Exemplo: Mapeando a violência na cidade de Porto Alegre - RS.



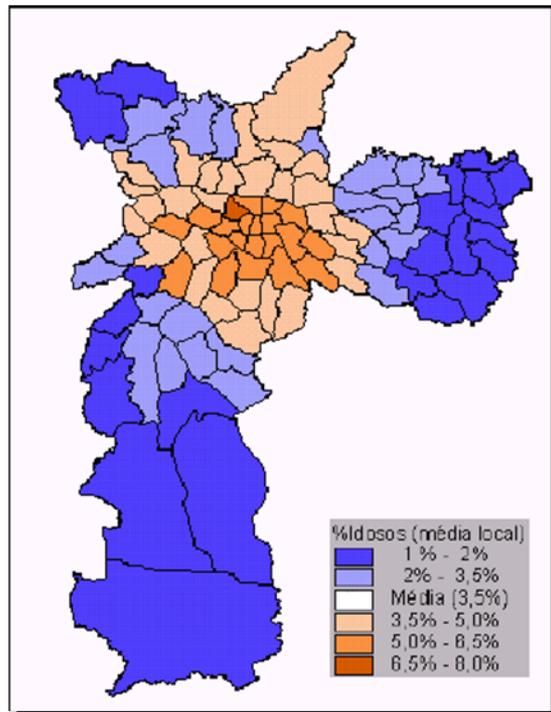
Legenda: ● Homicídios / ● Acidentes de transporte / ● Suicídios

Kernel – Número de Eventos



Análise de Padrões de Áreas

- No caso da análise de padrões de áreas, a distribuição de eventos está associada a áreas (polígonos).
- Objetivo de análise será determinar a existência de um padrão espacial nos valores agregados aos polígonos.



Disparidade Social

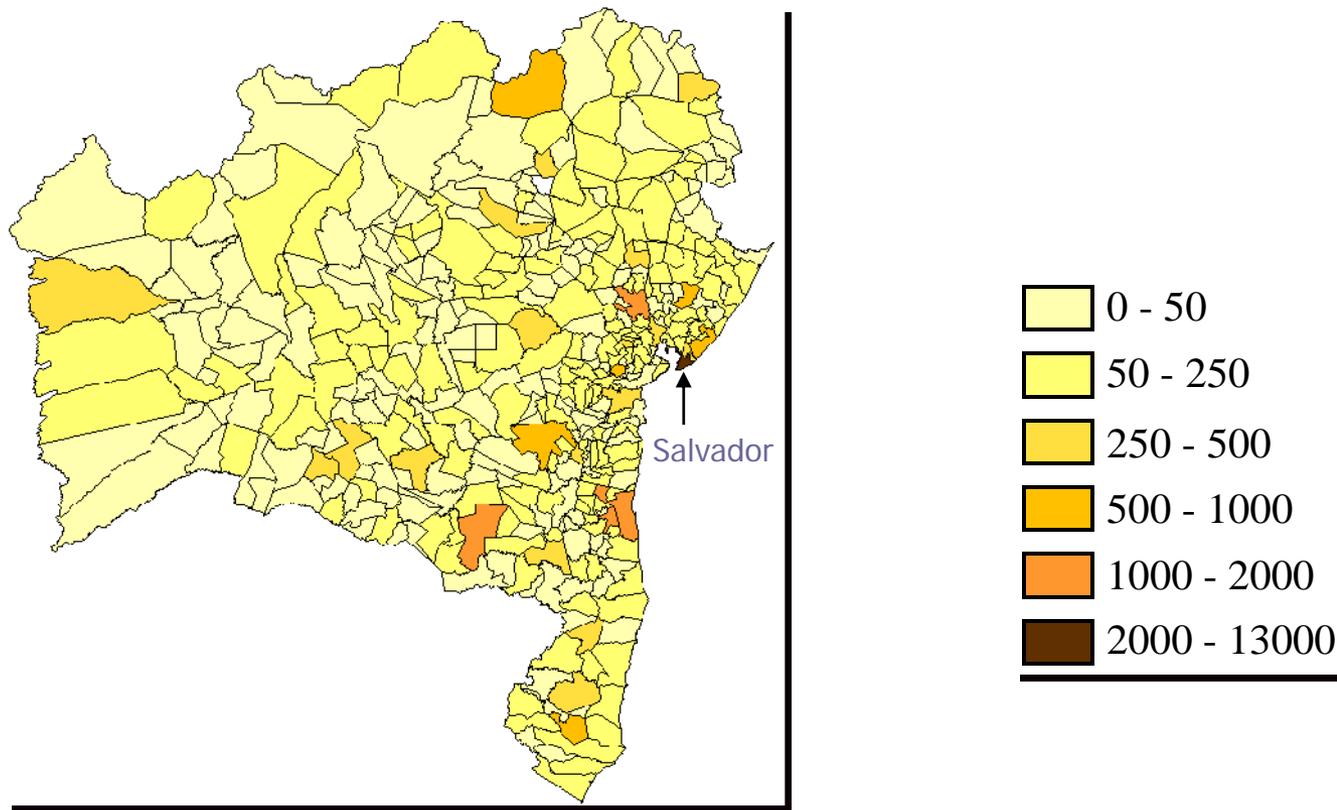
Percentual de Idosos na cidade de São Paulo.

Existe algum padrão espacial ?

Que fatores explicam essa distribuição ?

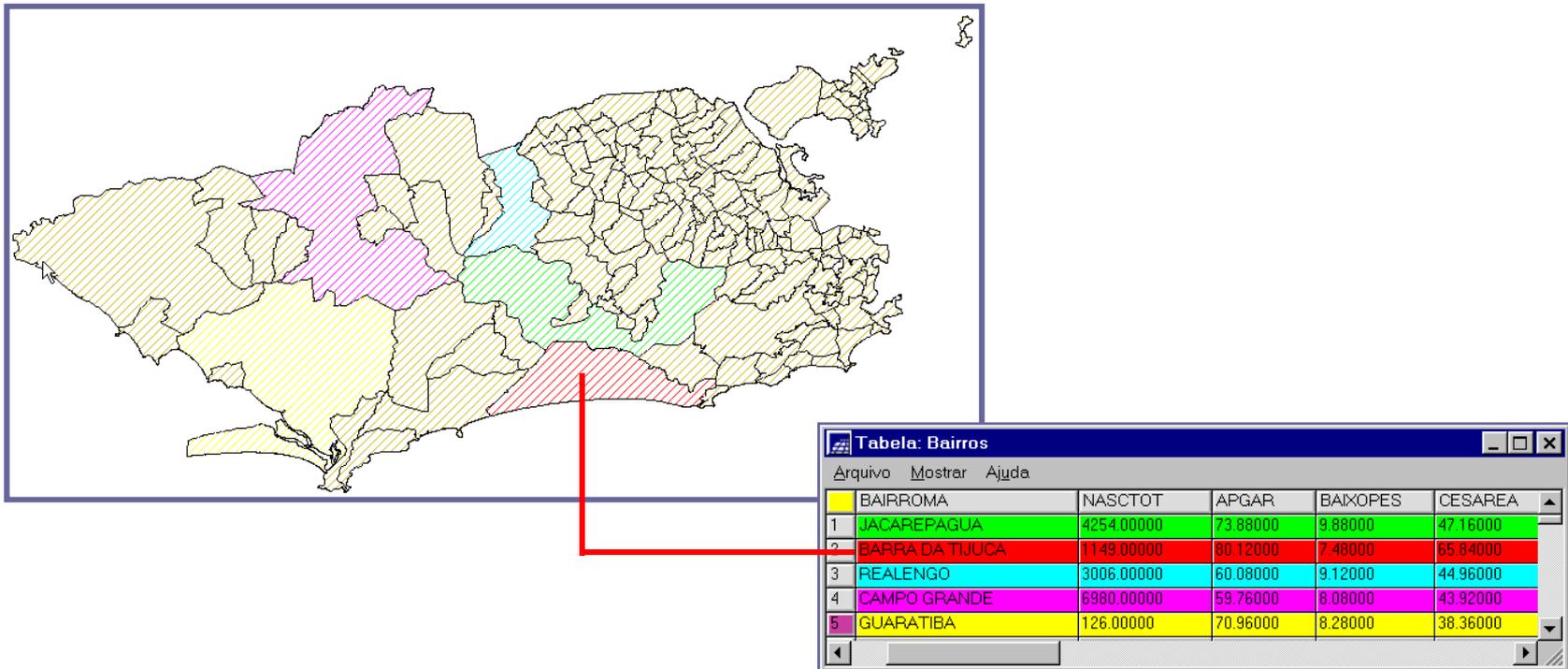
Análise de Padrões de Áreas

- A forma usual de apresentação dos padrões de áreas é através de mapas temáticos.



Análise de Padrões de Áreas

- Mapas Cadastrais (poligonais)
 - Objetos: entidades do mundo real (*Ex: Estados, Municipios, Bairros, etc...*)
 - Atributos: valores agregados aos objetos.



Explorando Dados de Área

- Efeitos de Primeira Ordem
 - Média Espacial Móvel
- Dependência Espacial Global
 - Efeitos de segunda ordem
 - Indicadores: Moran's I, Geary's C
- Dependência Espacial Local
 - LISA (Local Indicators of Spatial Association)
 - Indicadores: Moran Local I_i (Anselin), G_i (Getis)

Indicadores Globais de Autocorrelação Espacial

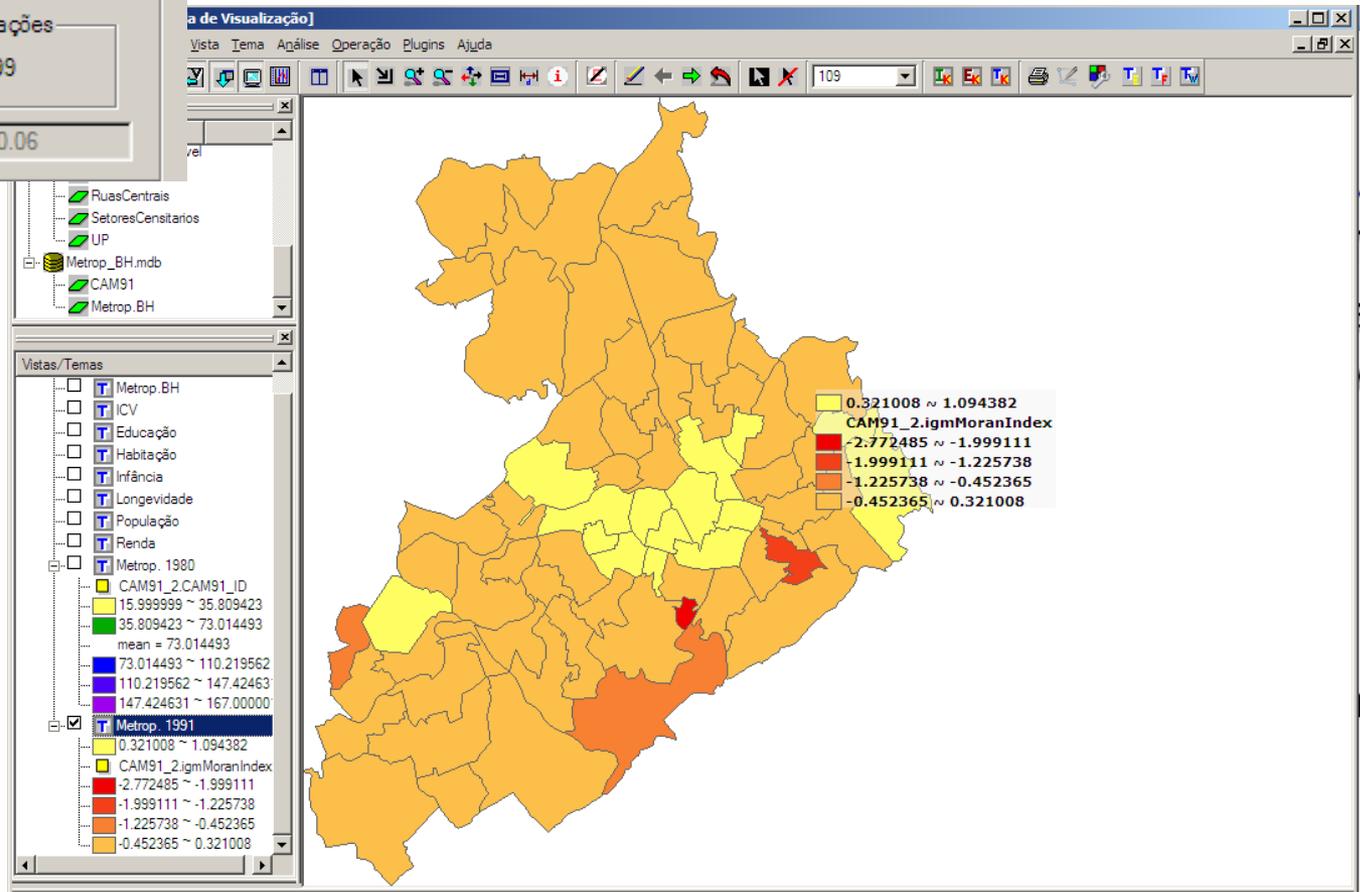
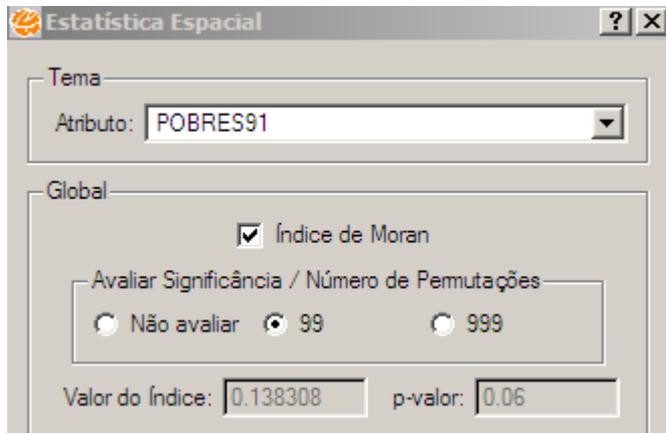
- Para muitos tipos de dados é importante explorar a dependência espacial, mostrando como os valores estão correlacionados no espaço.
- O conceito mais utilizado é o de autocorrelação espacial.
- Resumidamente a autocorrelação espacial mede o quanto o valor observado de um atributo numa região é independente dos valores desta mesma variável nas localizações vizinhas.
- Uma das formas de detecção de similaridade entre áreas é através do **índice global de Moran I** .

Indicadores Globais de Autocorrelação Espacial

- O significado do valor do índice global de Moran (I) é análogo ao coeficiente de correlação convencional, porque têm em seu numerador um termo que é produto de momento.
- Como um coeficiente de correlação, os valores de I também variam de:
 - 1 a +1, quantificando o grau de autocorrelação espacial existente.
 - ⇒ -1 autocorrelação espacial negativa ou inversa.
 - ⇒ 0 significa aleatoriedade
 - ⇒ +1 significa autocorrelação espacial positiva ou direta.

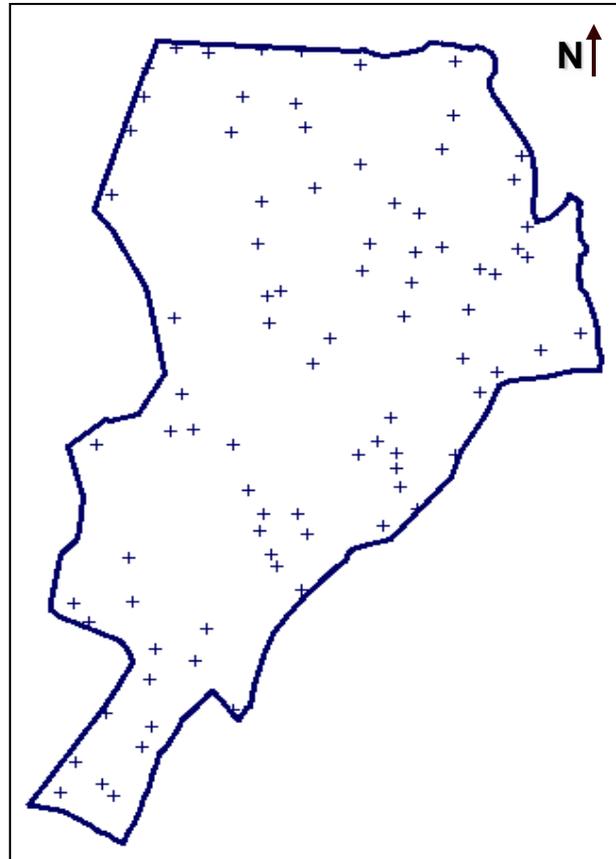
Índice de Moran

- ⇒ -1 autocorrelação espacial negativa
- ⇒ 0 significa aleatoriedade
- ⇒ +1 significa autocorrelação espacial positiva

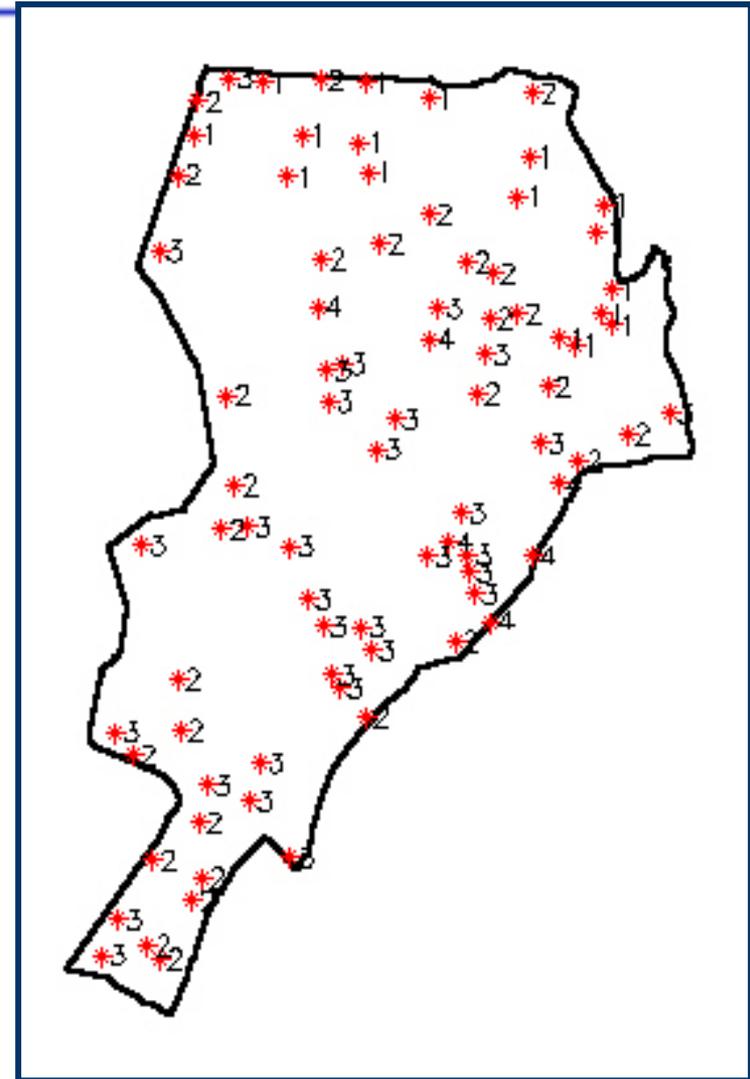


Geoestatística

- Solos: dado um conjunto de amostras, qual o comportamento espacial ou a extensão da variabilidade espacial de uma certa propriedade?



- *Características dos atributos de dados espaciais*
 - Amostrados pontualmente e com distribuição esparsa
 - Valor do Atributo:
 - **contínuo** (altimetria, temperatura, teor geoquímico, ...)
 - **categórico** (classe de vegetação, classe de textura do solo, ...)
 - *Dados + próximos tem valores de atributo + parecidos*



- *Modelagem dos atributos espaciais*
 - Soluções por procedimentos geoestatísticos
 - A *Geoestatística* fornece uma coleção de ferramentas estatísticas e determinísticas que são utilizadas para entendimento e modelagem da variabilidade espacial
 - Atributo z de um dado espacial, em posições não amostradas, é considerado como *Variável Aleatória* $Z(\mathbf{u})$.

• Estatística clássica

- variáveis independentes
- sem continuidade espacial

• Estatística espacial

- valores associados à localização no espaço e/ou no tempo
- distribuição contínua dos valores
- processos de estimativa para pontos não amostrados

Variável Regionalizada

Varia de um lugar a outro no espaço com certa continuidade. São variáveis cujos valores estão relacionados, de algum modo, com a posição espacial que ocupam.

Variograma

- Ferramenta básica, que permite descrever quantitativamente a variação no espaço de um fenômeno regionalizado.
- A natureza estrutural de um conjunto de dados, assumido pela variável regionalizada, é definida a partir da comparação de valores tomados simultaneamente em dois pontos, segundo uma determinada direção.
- A função variograma $2\gamma(h)$ é definida como sendo a esperança matemática do quadrado da diferença entre os valores de pontos no espaço, separados por uma distância h .

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

Variograma

- A interpretação do variograma permite obter parâmetros que descrevem o comportamento espacial das variáveis regionalizadas.
- Uma feição resultante da análise dos parâmetros do variograma experimental é a zona de influência: qualquer valor de $Z(x)$ estará correlacionado com outros valores $Z(x+h)$ que estiverem dentro de um raio "a" de x . Esta correlação, ou a influência de um valor em outro, decresce conforme $Z(x+h)$ aproxima-se de "a".
- O variograma é utilizado para calcular os valores de semivariância, para uma dada distância, os quais são necessários para a organização do sistema de equações de krigagem.

Variograma

- O variograma substitui a distância euclidiana “h” pela distância “ $2\gamma(h)$ ”, atributo específico do local em estudo.
- A distância dada pelo variograma mede o **grau médio de similaridade** entre um valor não amostrado e um valor conhecido vizinho.

- A **geoestatística** estuda o comportamento das chamadas **variáveis regionalizadas** e fundamentalmente baseia-se nos seguintes pressupostos:
 - **Estacionariedade**: na região em que se pretende fazer estimativas, o fenômeno é descrito como homogêneo dentro desse espaço;

Para a obtenção de um variograma, portanto, é suposto que a variável regionalizada tenha um comportamento fracamente estacionário, onde os valores esperados, assim como sua covariância espacial, sejam os mesmos por uma determinada área.

- ***Krigeagem***

é uma “coleção de técnicas de regressão linear generalizadas para *minimizar uma variância de estimação* a partir de um modelo de covariância definido a priori”, Journel, 1996.

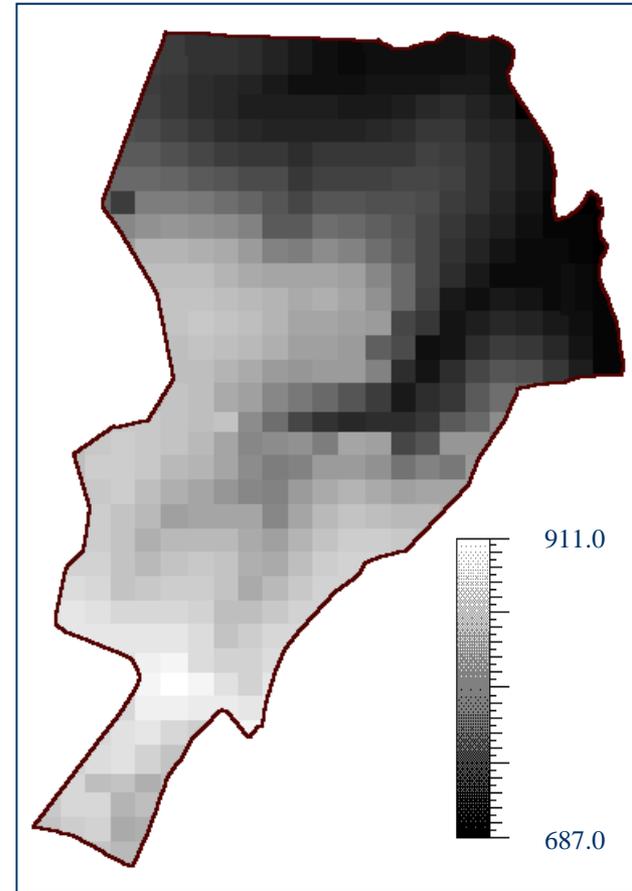
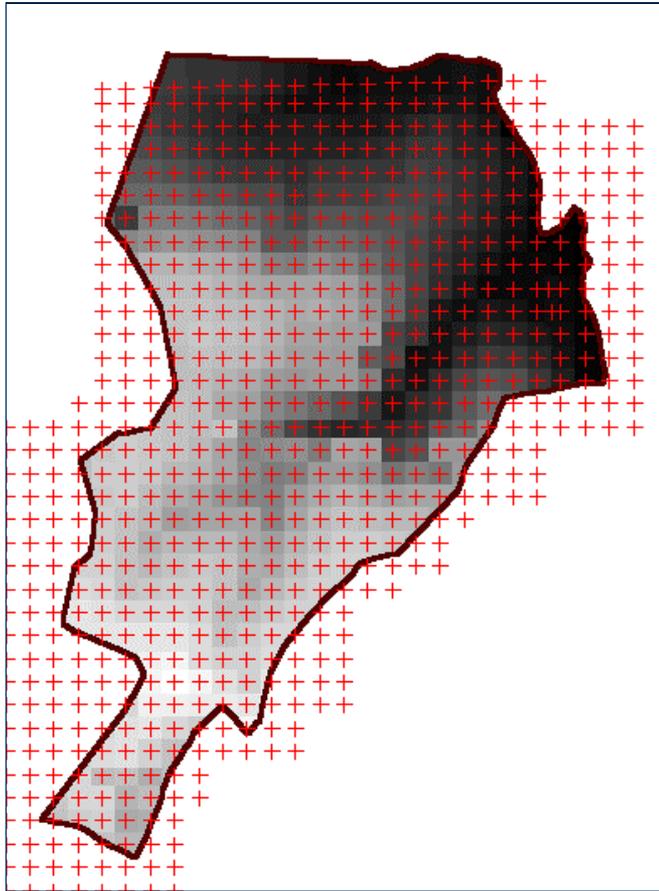
- A krigagem estima um valor não amostrado a partir de um conjunto de valores vizinhos $z(\mathbf{u}_\alpha)$, $\alpha = 1, \dots, n$. Considerando-se um modelo de FA estacionária com média μ e covariância $C(\mathbf{h})$, o estimador linear para ***Krigeagem Simples (SK)*** é definido por:

$$z^*_{SK}(\mathbf{u}) = \sum_{\alpha=1}^n \lambda_\alpha(\mathbf{u}) \cdot z(\mathbf{u}_\alpha) + \left[1 - \sum_{\alpha=1}^n \lambda_\alpha(\mathbf{u}) \right] \cdot \mu$$

- ***Krigeagem ordinária*** não depende do valor da média

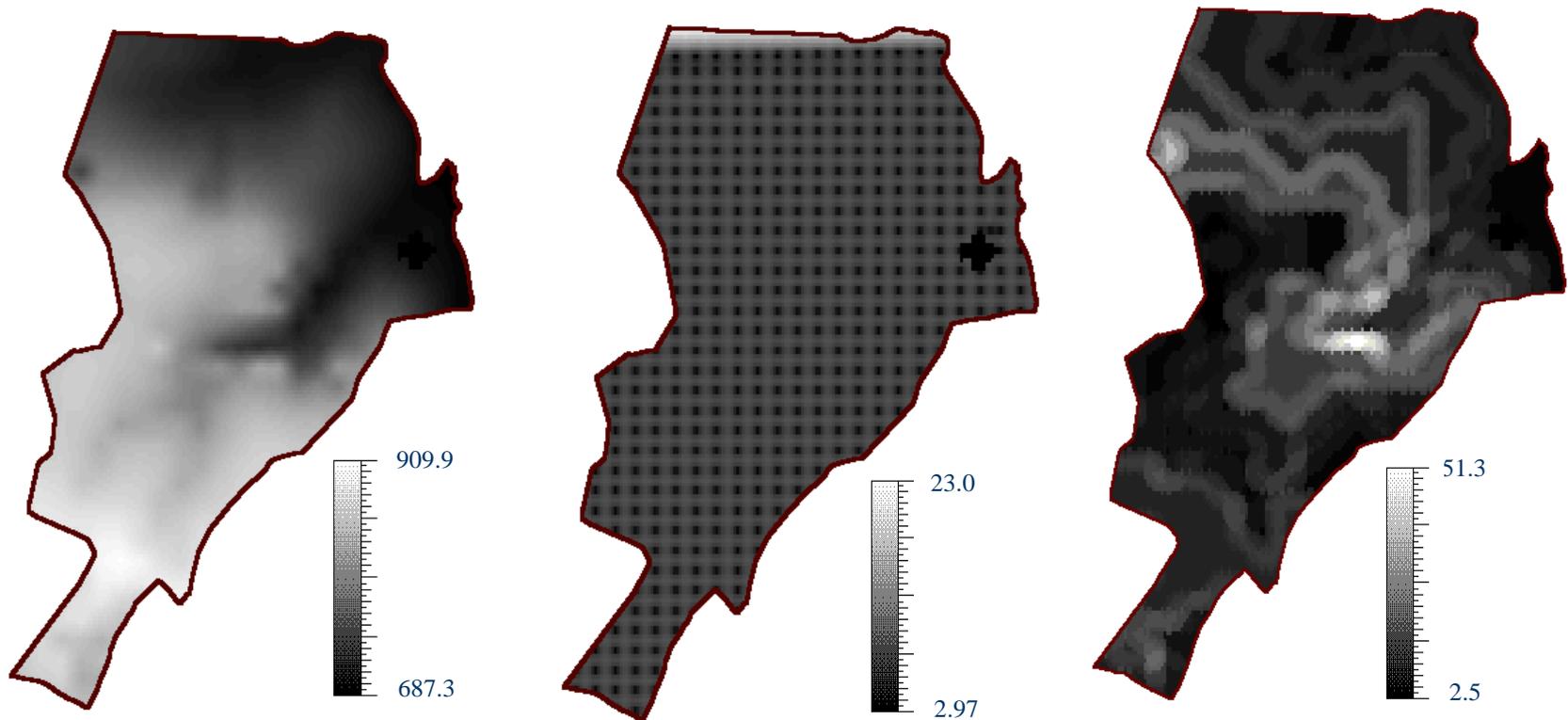
$$\sum_{\alpha=1}^n \lambda_\alpha(\mathbf{u}) = 1 \Rightarrow z^*_{SK}(\mathbf{u}) = \sum_{\alpha=1}^n \lambda_\alpha(\mathbf{u}) \cdot z(\mathbf{u}_\alpha)$$

- Exemplos - Modelagem de variáveis contínuas



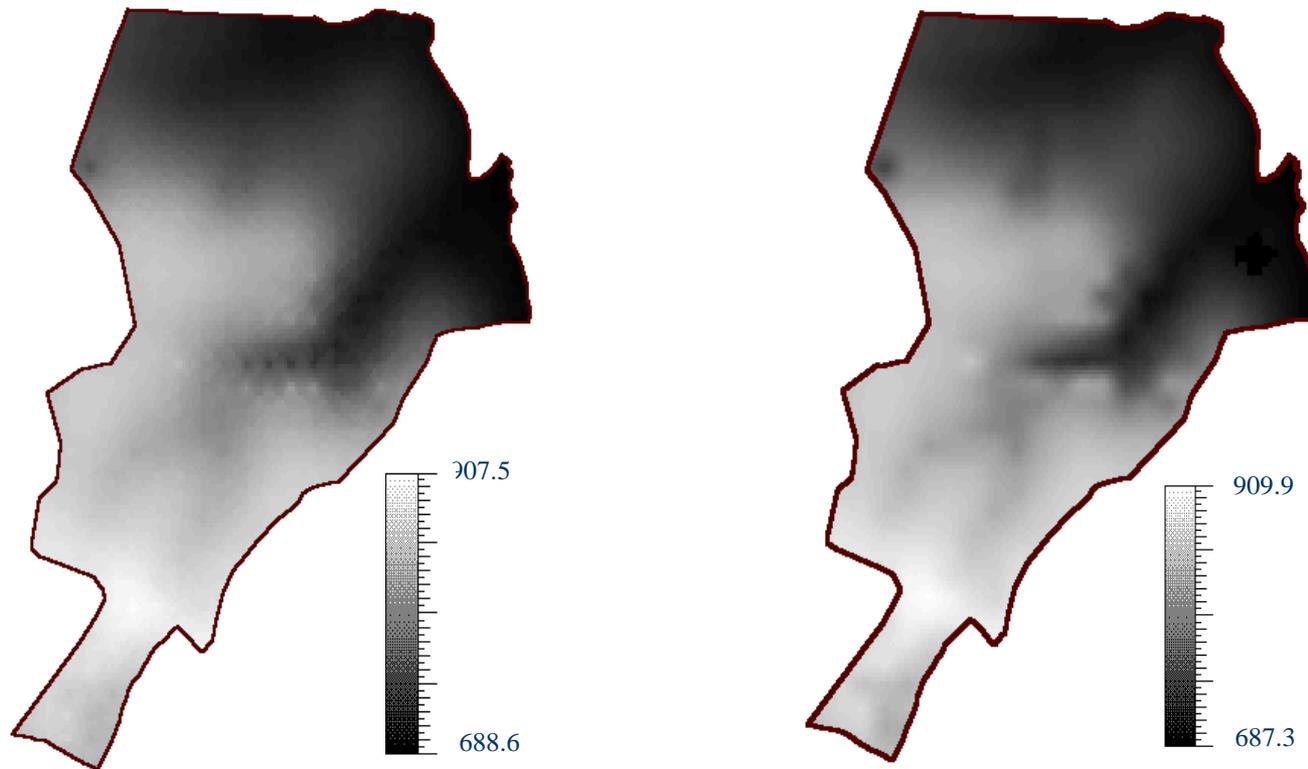
Figuras: Distribuição de amostras e interpolação por vizinho mais próximo.

- Variáveis contínuas



Figuras: Mapas de média (esquerda) e desvio padrão calculados a partir da variância de krigeagem ordinária e da fdac determinada pela krigeagem por indicação

- Variáveis contínuas



Figuras: Interpolação por média dos vizinhos por quadrante e por krigagem ordinária.

Análise Espacial - Sugestões para Literaturas de Apoio

geral

Câmara, G.; Druck, S.; Carvalho, M. S.; Monteiro, A. M. V.; Camargo, E. C. G.; Felgueira, C. A. **Análise espacial de dados geográficos**. Planaltina, DF: Embrapa Cerrados, 2004. 209 p.
disponível em: <http://www.dpi.inpe.br/livros.php>

ponto

Bailey, T. C.; Gatrell, A. C. **Interactive spatial data analysis**. New York: Wiley, 1995. 413 p.

área

Anselin, L. **SpaceStat tutorial: a workbook for using SpaceStat in the analysis of spatial data**. Santa Barbara, NCGIA (National Center for Geographic Information and Analysis), 1992.

Anselin, L. Local indicators of spatial association - LISA. **Geographical Analysis** v.27, p.91-115, 1995.

geoestatística

Issaks, M. e E. Srivastava (1989). **An Introduction to Applied Geostatistics**. New York, Oxford University Press, 1989.

Deutsch, C. e A. Journel (1992). **GSLIB: Geostatistical Software Library and user's guide**. New York, Oxford University Press.

Goovaerts, P. **Geostatistics for natural resources evaluation**. New York: Oxford University Press, 1997. 483 p.