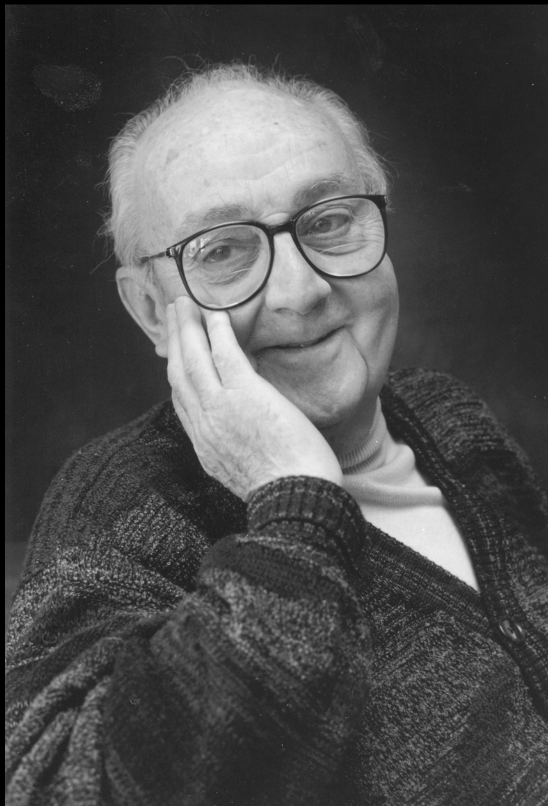
A very short introduction to Modeling in Landscape Archaeology

Daniel Knitter

EXCELLENCE CLUSTER TOPOI


By DavidMCEddy at en.wikipedia, CC BY-SA 3.0,
https://commons.wikimedia.org/w/index.php?curid=14941622

"Since all models are wrong
the scientist cannot obtain a 'correct' one
by excessive elaboration.

On the contrary following William of Occam
[s]he should seek an economical description
of natural phenomena.

Just as the ability to devise simple but evocative
models is the signature of the great scientist
so overelaboration and overparameterization
is often the mark of mediocrity" (Box 1976, 792).

Content

- Modeling and processes
- Why space matters: Potentials and Pitfalls of Spatial Data
- Point Pattern Analyses: What processes might have caused the distribution of findings?
    - First order processes (density based approaches)
    - Second order processes (distance based approaches)
- From single attributes to continuous surfaces: geostatistics
- Geomorphometry: What is the character of the landscape?
- Outlook

Content

- **Modeling and processes**
- Why space matters: Potentials and Pitfalls of Spatial Data
- Point Pattern Analyses: What processes might have caused the distribution of findings?
  - First order processes (density based approaches)
  - Second order processes (distance based approaches)
- From single attributes to continuous surfaces: geostatistics
- Geomorphometry: What is the character of the landscape?
- Outlook

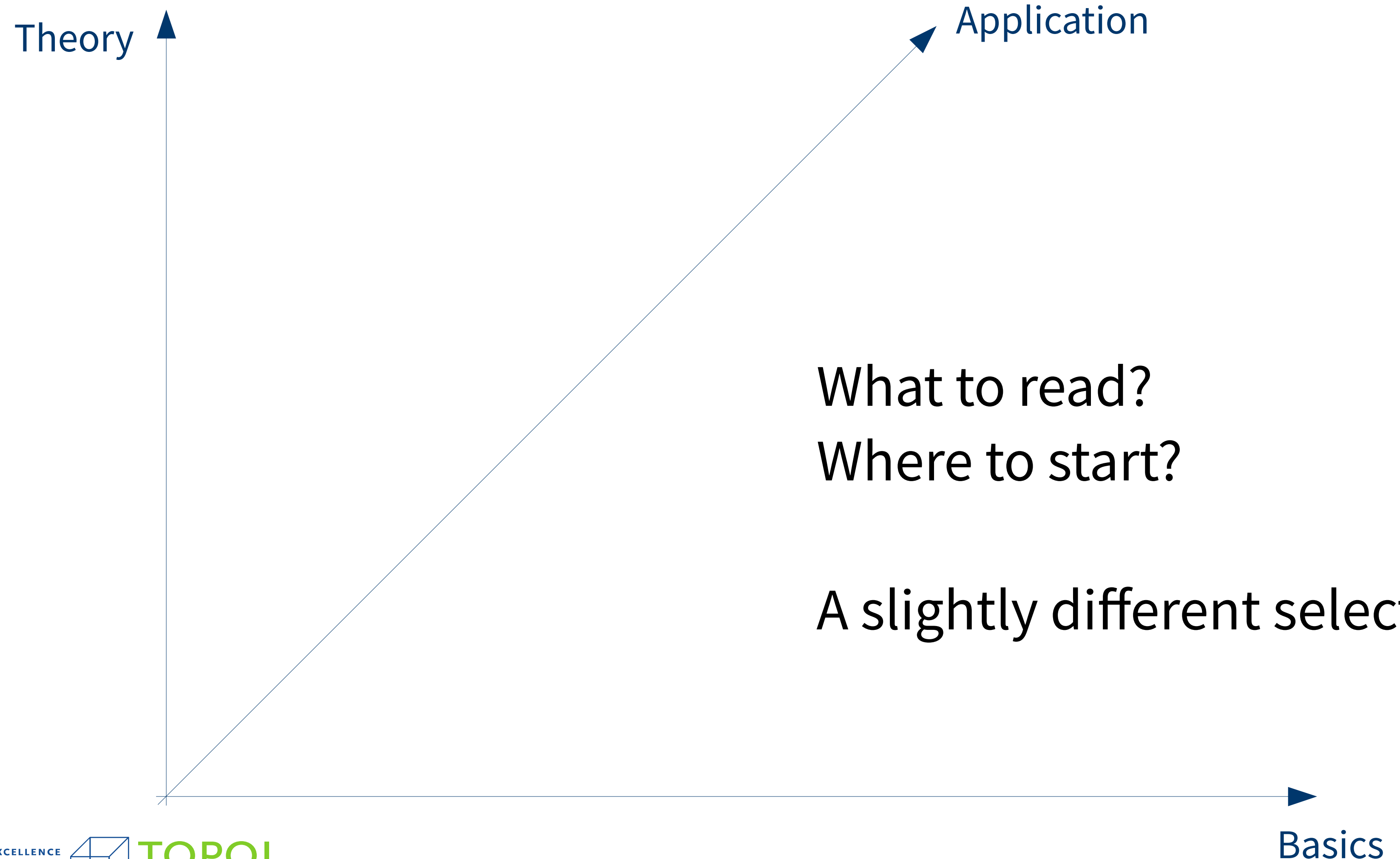
Modeling in (Landscape) Archaeology

Modeling in Landscape Archaeology

What to read?
Where to start?

A slightly different selection...

Modeling in Landscape Archaeology



What to read?
Where to start?

A slightly different selection...


Theory
Application
Basics
Statistical Analysis and Modelling of Spatial Point Patterns
JANINE ILLIAN, ANTTI PENTTINEN HELGA STOYAN AND DIETRICH STOYAN
WILEY
STATISTICS IN PRACTICE
A Practical Guide to Geostatistical Mapping
Tomislav Hengl
Chapman & Hall/CRC Interdisciplinary Statistics Series
Spatial Point Patterns
Methodology and Applications with R
GEOGRAPHIC INFORMATION ANALYSIS
second edition
DAVID O'SULLIVAN
DAVID J. UNWIN
Use R!
Roger S. Bivand
Edzer Pebesma
Virgilio Gómez-Rubio
Applied Spatial Data Analysis with R
Second Edition
Springer
Quantitative Archaeology and Archaeological Modelling
Oliver Nakoinz
Daniel Knitter
Modelling Human Behaviour in Landscapes
Basic Concepts and Modelling Elements
Springer

Modeling in Landscape Archaeology

- Spatial phenomena are the result of processes
- Modeling = (re-)construction of these processes

Modeling in Landscape Archaeology

- Spatial phenomena are the result of processes
- Modeling = (re-)construction of these processes

Artifacts
(pottery, metals, bricks, …)

Spatial **process** →

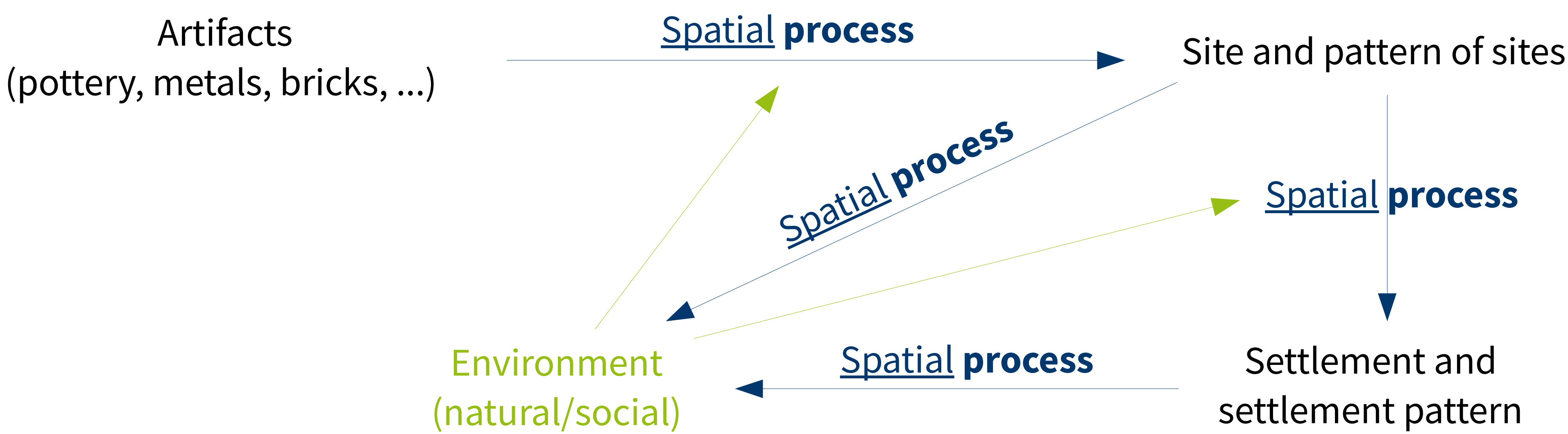
Site and pattern of sites

Spatial **process** ↓

Settlement and
settlement pattern

Modeling in Landscape Archaeology

- Spatial phenomena are the result of processes
- Modeling = (re-)construction of these processes

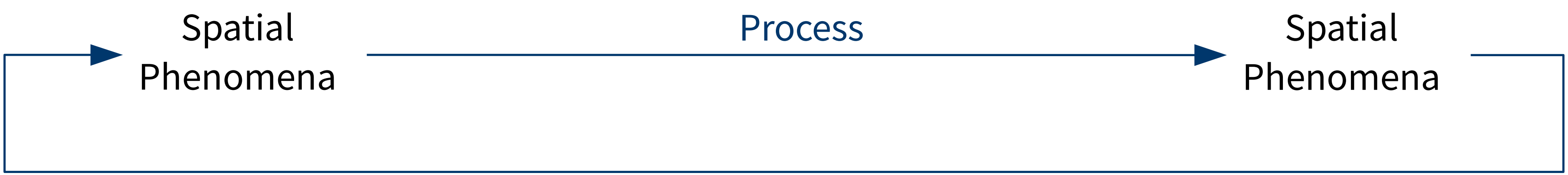


Modeling in Landscape Archaeology

Spatial Phenomena $\xrightarrow{\text{Process}}$ Spatial Phenomena

Modeling in Landscape Archaeology

Modeling as a recursive activity.



Content

- Modeling and processes
- **Why space matters: Potentials and Pitfalls of Spatial Data**
- Point Pattern Analyses: What processes might have caused the distribution of findings?
  - First order processes (density based approaches)
  - Second order processes (distance based approaches)
- From single attributes to continuous surfaces: geostatistics
- Geomorphometry: What is the character of the landscape?
- Outlook

Spatial data are special

Pitfalls of spatial data:

→ spatial autocorrelation
→ modifiable areal unit problem
→ scale
→ ecological fallacy
→ non-uniformity of space
→ edge effects

Spatial autocorrelation

The Earth is not an isotropic plate

Spatial data are not random

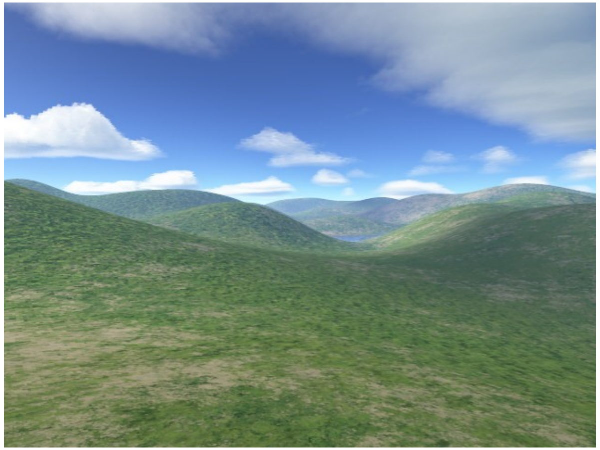
Data from locations near to each other are usually more similar than data far away from each other

Space and location are important because of Spatial autocorrelation.

Without spatial autocorrelation, spatial analyses would be pointless

Spatial autocorrelation

Low



High

Maximum

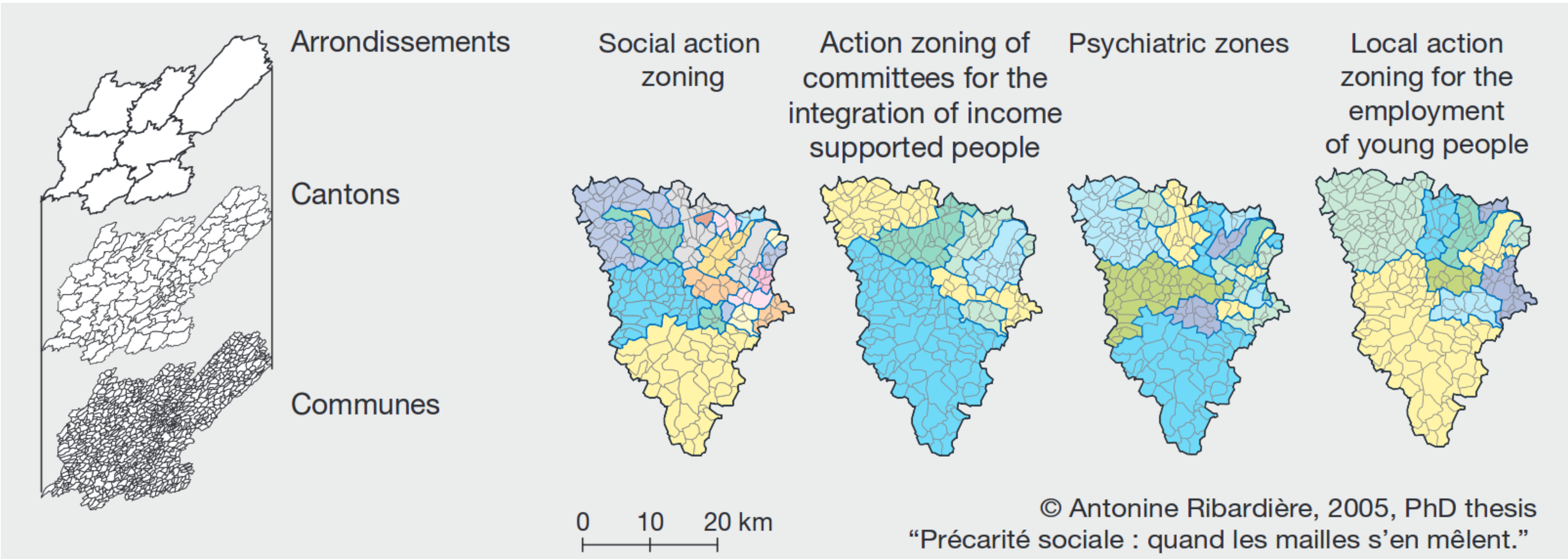
www.geog.ucsb.edu/~kclarke/G176B/Lecture12.ppt

Modifiable Areal Unit Problem (MAUP)

- aggregation units used are arbitrary with respect to the phenomena under investigation
- If spatial units are specified differently, one might get very different patterns

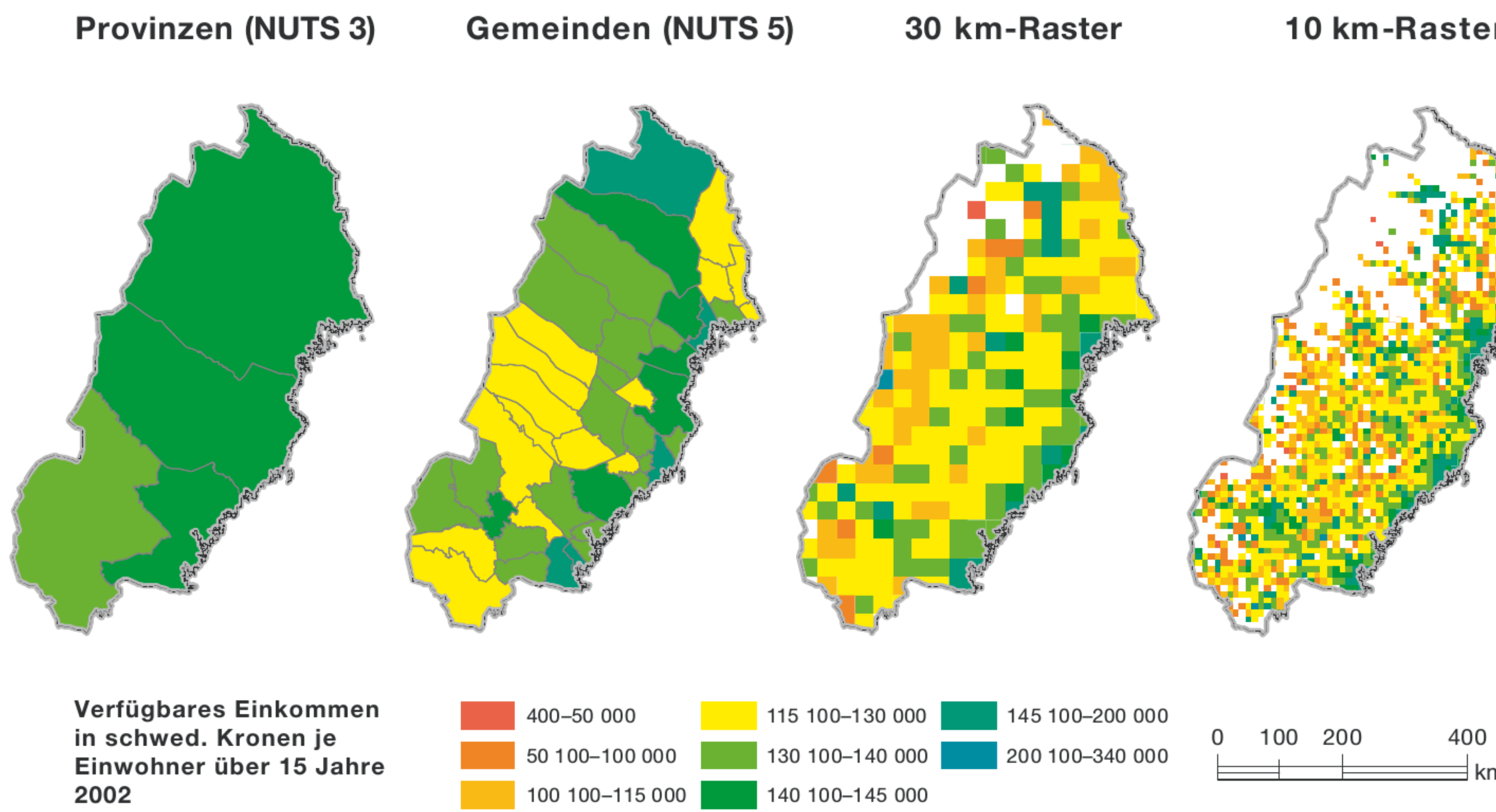
→ Openshaw, Stan (1984): The modifiable areal unit problem

Modifiable Areal Unit Problem (MAUP)



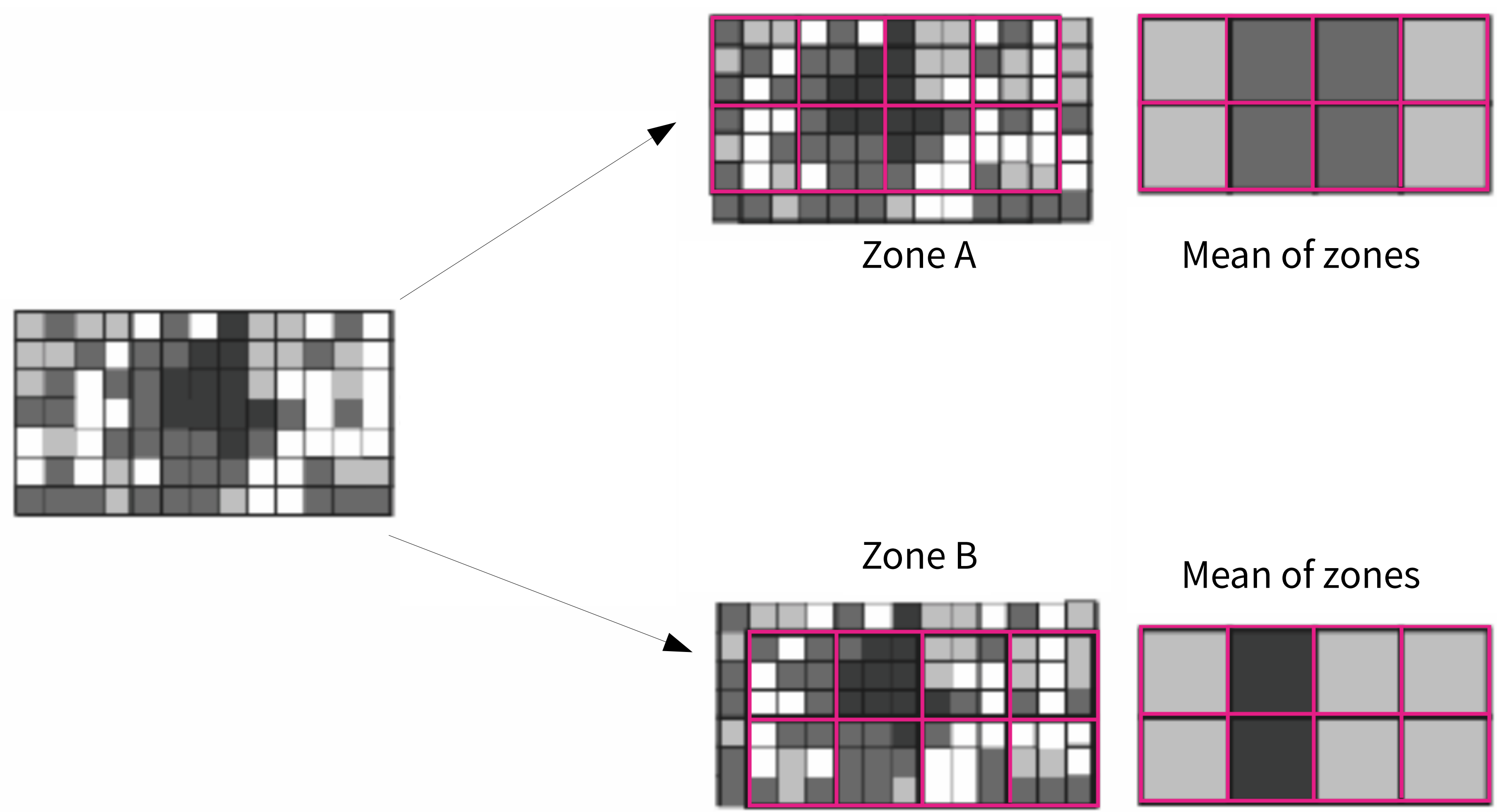
http://www.bbsr.bund.de/BBSR/DE/Veroeffentlichungen/IzR/2009/10_11/Inhalt/DL_Madelin_ua.pdf?__blob=publicationFile&v=2

Modifiable Areal Unit Problem (MAUP)



http://www.bbsr.bund.de/BBSR/DE/Veroeffentlichungen/IzR/2009/10_11/Inhalt/DL_Madelin_ua.pdf?__blob=publicationFile&v=2

Modifiable Areal Unit Problem (MAUP)



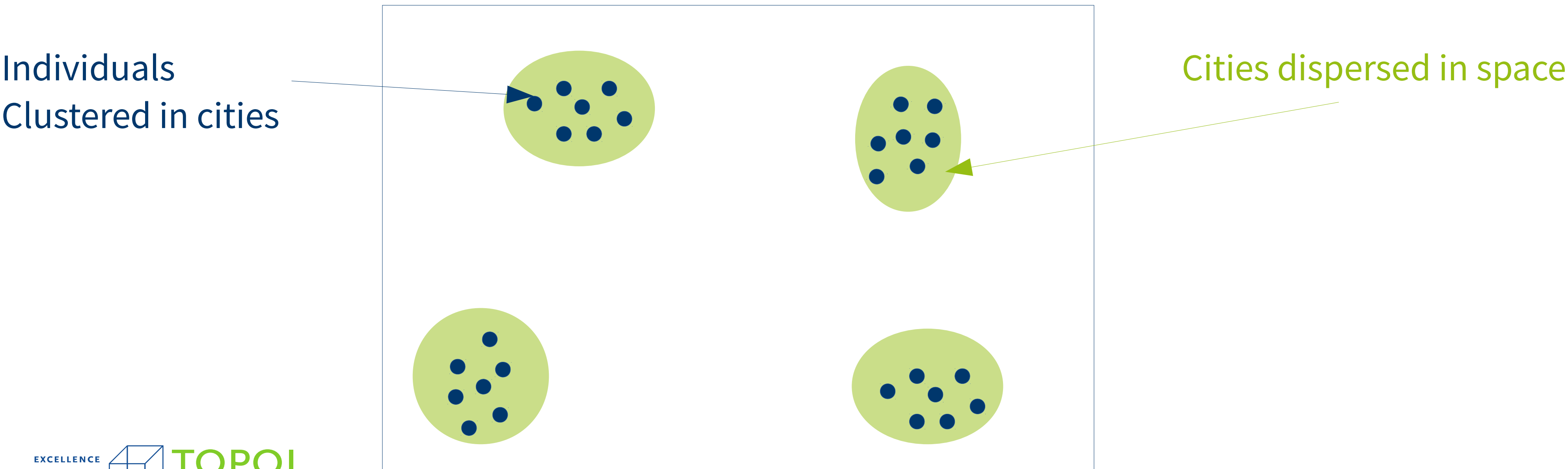
http://www.bbsr.bund.de/BBSR/DE/Veroeffentlichungen/IzR/2009/10_11/Inhalt/DL_Madelin_ua.pdf?__blob=publicationFile&v=2

Scale effects

- Scale affects results and representativity of data

- e.g. cities may be represented as points or polygons

- results depend on the scale of analysis, e.g. province or county → MAUP

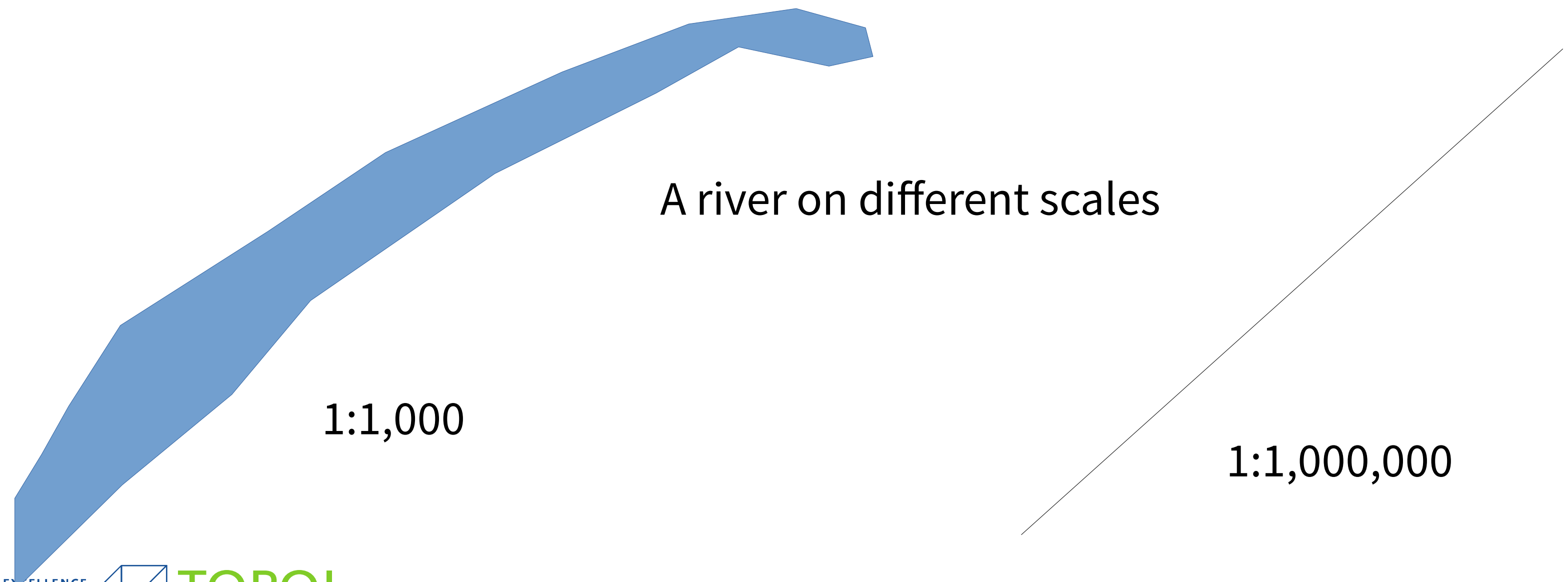


Scale effects

- Scale affects results and representativity of data

- e.g. cities may be represented as points or polygons

- results depend on the scale of analysis, e.g. province or county → MAUP



A river on different scales

Ecological fallacy

Arises when:
- Statistical relationship at one level of aggregation…
- …is assumed to hold true because it holds at a more detailed level

correlation IS NOT causation

For example, we might observe a strong relationship between income and crime at the county level, with lower-income counties being associated with higher crime rates. If from this we conclude that lower-income individuals are more likely to commit a crime, then we are falling for the ecological fallacy. In fact, it is only valid to say exactly what the data say: that lower-income counties tend to experience higher crime rates. What causes the observed effect may be something entirely different

(O'Sullivan & Unwin, 2010)

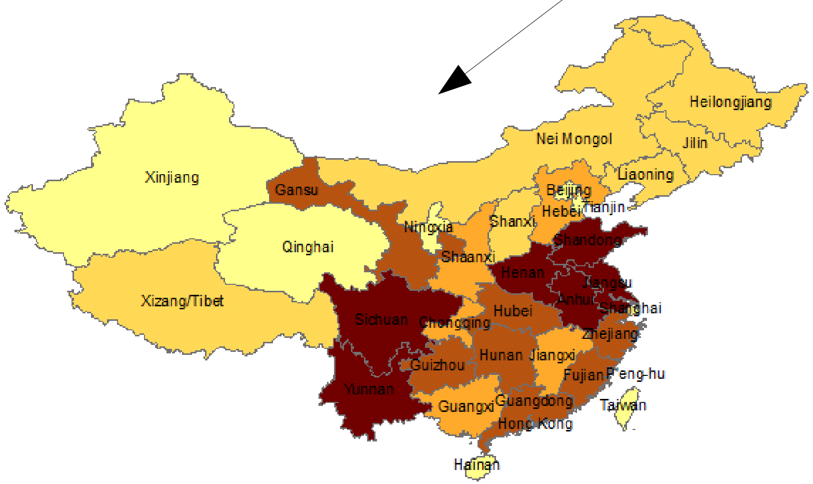
Non-uniformity of space

Bank robberies are clustered
→ because banks are clustered

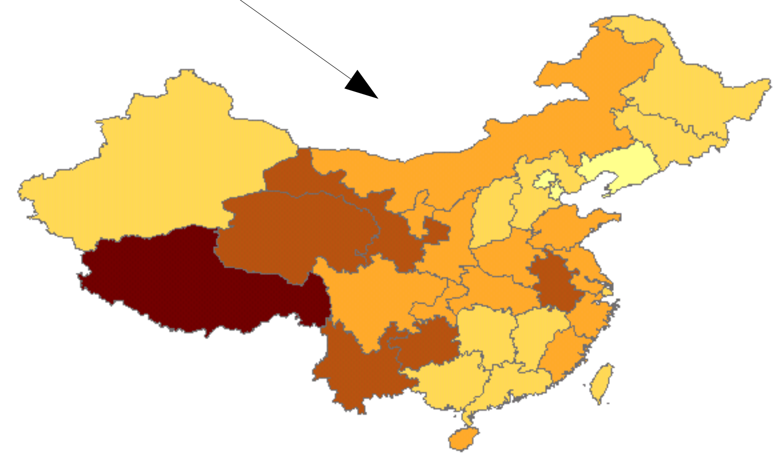
Diseases due to bad air are clustered
→ because factories are clustered

Illiteracy in China

absolute



as ratio of population



www.utdallas.edu/~briggs/henan/2spatanal.ppt

Edge effects

- unless you study the entire world: Every study region has a boundary

- You do not have data outside your study region

- But: the “outside” data may/do affect the data of your study region (if there is spatial autocorrelation...)

Spatial data are special

Pitfalls of spatial data:

→ spatial autocorrelation
→ modifiable areal unit problem
→ scale
→ ecological fallacy
→ non-uniformity of space
→ edge effects

Potentials of spatial data:

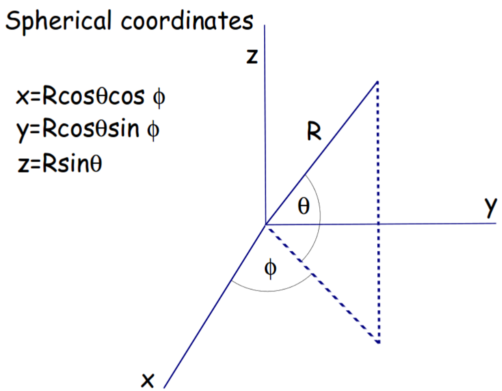
→ distance
→ adjacency
→ interaction
→ neighborhood

Distance

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

Euclidean distance as calculated using *Pythagoras'* formula

→ for *projected* data on the local/regional scale



R = radius of earth, ca. 6378 km

Spherical distance via spherical coordinates and a 3D version of *Pythagoras'* formula

$$d^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2$$

→ for *unprojected* data on a global scale

https://www.math.ksu.edu/~dbski/writings/haversine.pdf

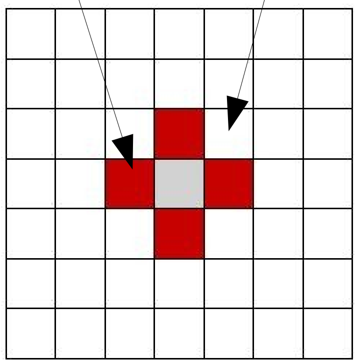
Distance – Adjacency – Interaction - Neighborhood

$$\mathbf{D} = \begin{bmatrix} 0 & 66 & 68 & 68 & 24 & 41 \\ 66 & 0 & 51 & 110 & 99 & 101 \\ 68 & 51 & 0 & 67 & 91 & 116 \\ 68 & 110 & 67 & 0 & 60 & 108 \\ 24 & 99 & 91 & 60 & 0 & 45 \\ 41 & 101 & 116 & 108 & 45 & 0 \end{bmatrix}$$

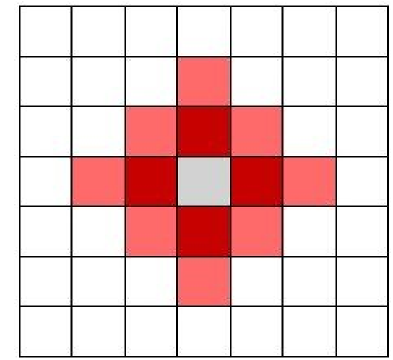
(O'Sullivan & Unwin, 2010)

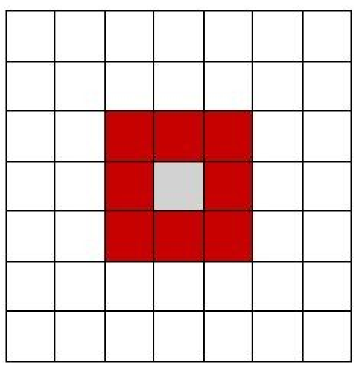
Adjacency

Binary
→ things are adjacent or not

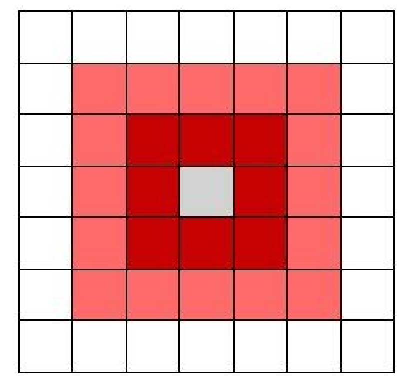


Rook case
→ sharing a boundary





Queen case
→ sharing a boundary or point



1st order neighbor

2nd order neighbor

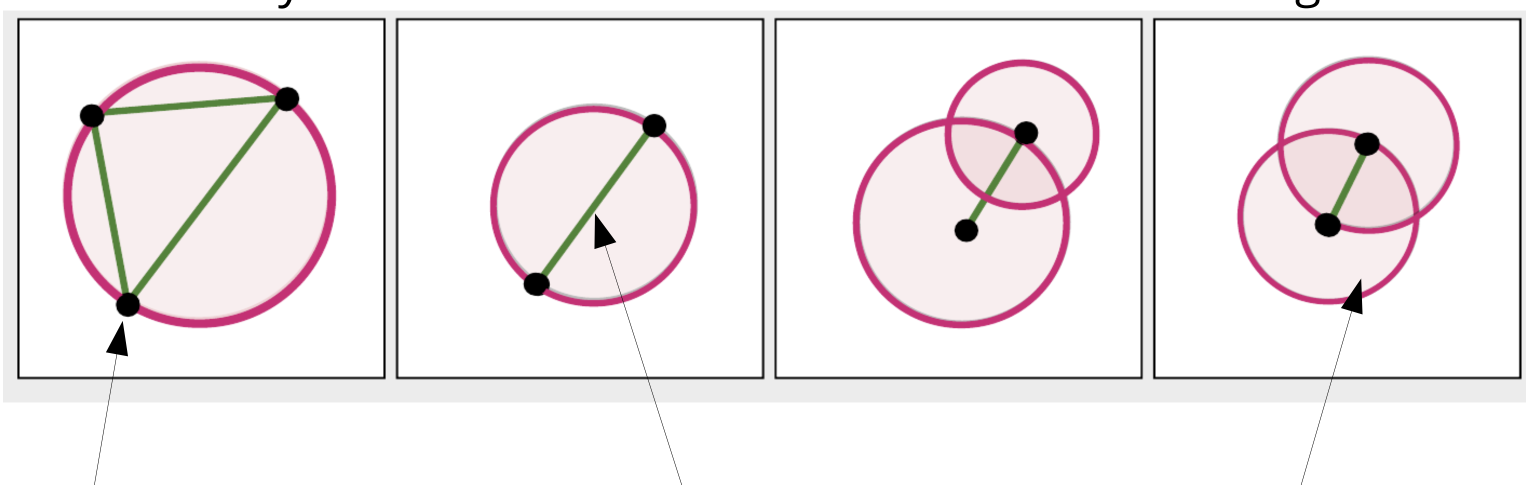
and also…distance based neighbors…

www.utdallas.edu/~briggs/henan/2spatanal.ppt

Adjacency

...or graph based neighbors.

Delaunay | Gabriel | Sphere of influence | Relative neighbor



Neighbor

Edge between neighbor

Space without points

(Nakoinz & Knitter 2016)

Distance – Adjacency – Interaction - Neighborhood

$$\mathbf{A}_{k=3} = \begin{bmatrix} * & 1 & 0 & 0 & 1 & 1 \\ 1 & * & 1 & 0 & 1 & 0 \\ 1 & 1 & * & 1 & 0 & 0 \\ 1 & 0 & 1 & * & 1 & 0 \\ 1 & 0 & 0 & 1 & * & 1 \\ 1 & 1 & 0 & 0 & 1 & * \end{bmatrix}$$

$$\mathbf{A}_{d \leq 50} = \begin{bmatrix} * & 0 & 0 & 0 & 1 & 1 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ 1 & 0 & 0 & 0 & * & 1 \\ 1 & 0 & 0 & 0 & 1 & * \end{bmatrix}$$

(O'Sullivan & Unwin, 2010)

Interaction

- A combination of distance and adjacency
- related to Tobler's first law of geography:

"(...) everything is related to everything else, but near things are more related than distant things" (Tobler 1970, 236)

interaction *weight* between i and *j*

rate of decline

$$w_{ij} \propto \frac{1}{d^k}$$

distance between *i* and *j*

...strength of interaction between i and j is proportional to their inverse distance

Distance – Adjacency – Interaction - Neighborhood

$$\mathbf{W} = \begin{bmatrix} \infty & 0.0152 & 0.0147 & 0.0147 & 0.0417 & 0.0244 \\ 0.0152 & \infty & 0.0196 & 0.0091 & 0.0101 & 0.0099 \\ 0.0147 & 0.0196 & \infty & 0.0149 & 0.0110 & 0.0086 \\ 0.0147 & 0.0091 & 0.0149 & \infty & 0.0167 & 0.0093 \\ 0.0417 & 0.0101 & 0.0110 & 0.0167 & \infty & 0.0222 \\ 0.0244 & 0.0099 & 0.0086 & 0.0093 & 0.0222 & \infty \end{bmatrix} \quad \begin{array}{c} \text{row totals:} \\ 0.1106 \\ 0.0639 \\ 0.0688 \\ 0.0646 \\ 0.1016 \\ 0.0744 \end{array}$$

W[1,2] = W[1,] / sum(W[1,])

...

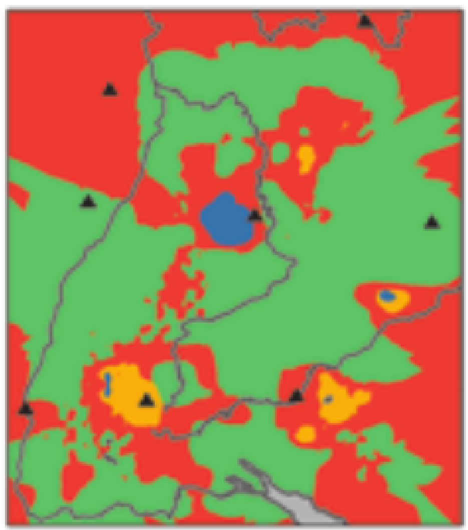
$$\begin{array}{rl} \mathbf{W} = & \begin{bmatrix} \infty & 0.1370 & 0.1329 & 0.1329 & 0.3767 & 0.2205 \\ 0.2373 & \infty & 0.3071 & 0.1424 & 0.1582 & 0.1551 \\ 0.2136 & 0.2848 & \infty & 0.2168 & 0.1596 & 0.1252 \\ 0.2275 & 0.1406 & 0.2309 & \infty & 0.2578 & 0.1432 \\ 0.4099 & 0.0994 & 0.1081 & 0.1640 & \infty & 0.2186 \\ 0.3279 & 0.1331 & 0.1159 & 0.1245 & 0.2987 & \infty \end{bmatrix} \\ \text{column totals:} & \begin{matrix} 1.4161 & 0.7949 & 0.8949 & 0.7805 & 1.2510 & 0.8626 \end{matrix} \end{array}$$

Influence of element A on all other elements

(O'Sullivan & Unwin, 2010)

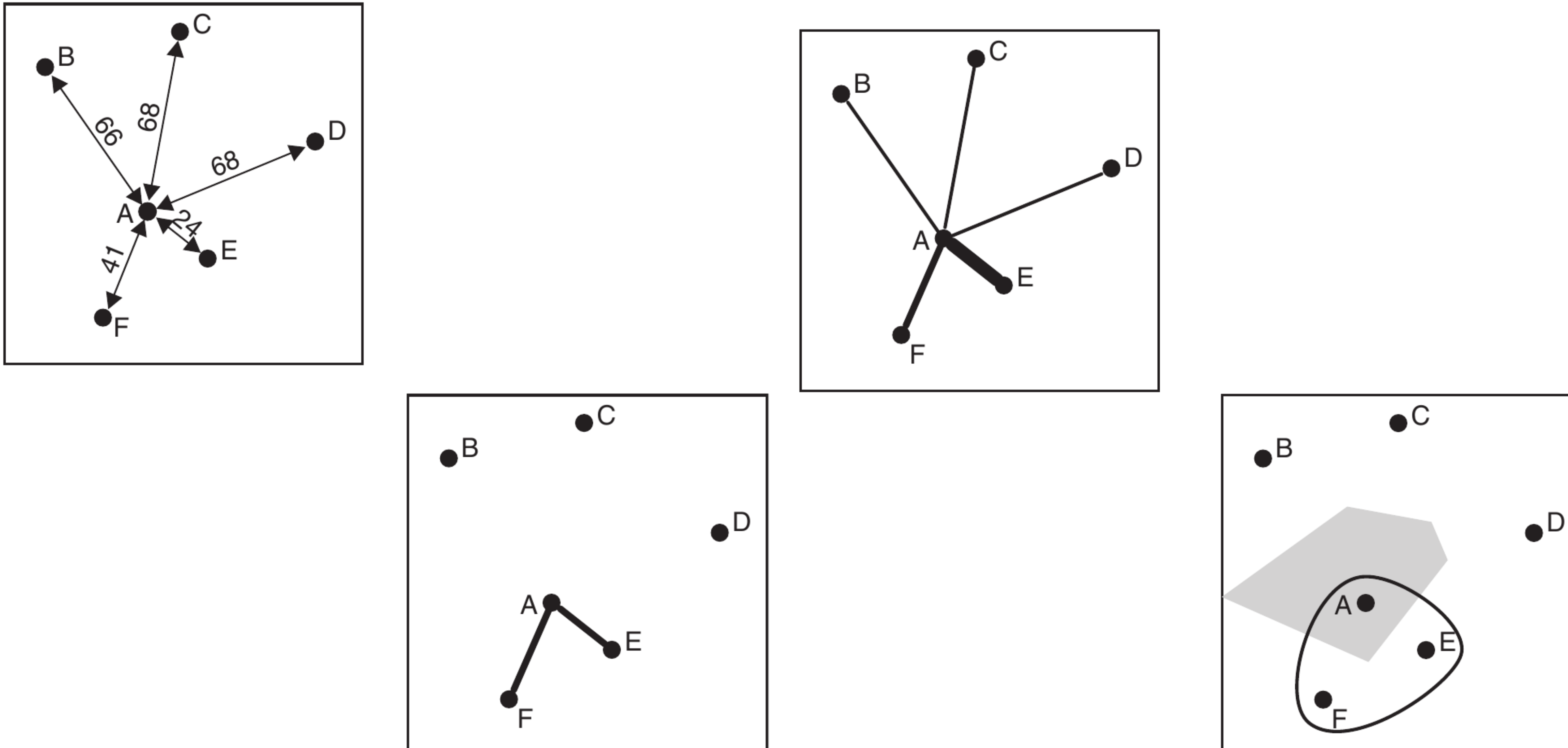
Neighborhood

Neighborhood of X is:
- the set of adjacent cells/areas
- the set of points/cells/areas within a certain distance
- the set of points/cells/areas that share some specific attribute or that are alike



(Nakoinz 2013)

Distance – Adjacency – Interaction - Neighborhood



EXCELLENCE CLUSTER TOPOI

(O'Sullivan & Unwin, 2010)

Spatial data are special

Pitfalls of spatial data:

→ spatial autocorrelation
→ modifiable areal unit problem
→ scale
→ ecological fallacy
→ non-uniformity of space
→ edge effects

Potentials of spatial data:

→ distance
→ adjacency
→ interaction
→ neighborhood

...it's all about processes

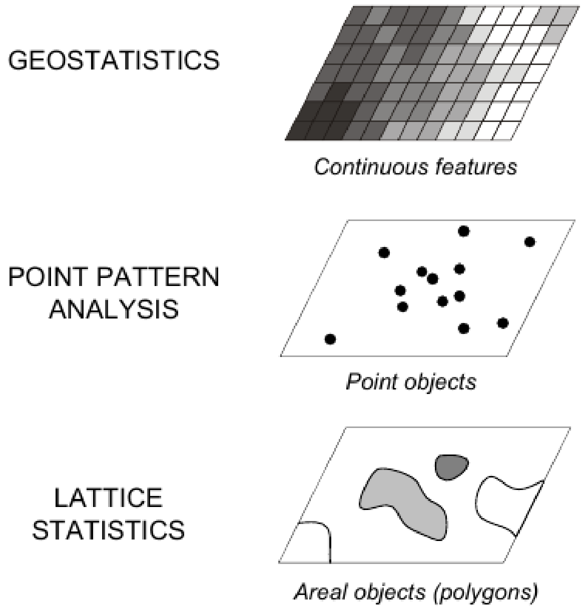


Fig. 1.2: Spatial statistics and its three major subfields after Cressie (1993).

However you call it and focus on...

everything is about processes...

Because they cause spatial patterning, autocorrelation, ...

EXCELLENCE CLUSTER TOPOI

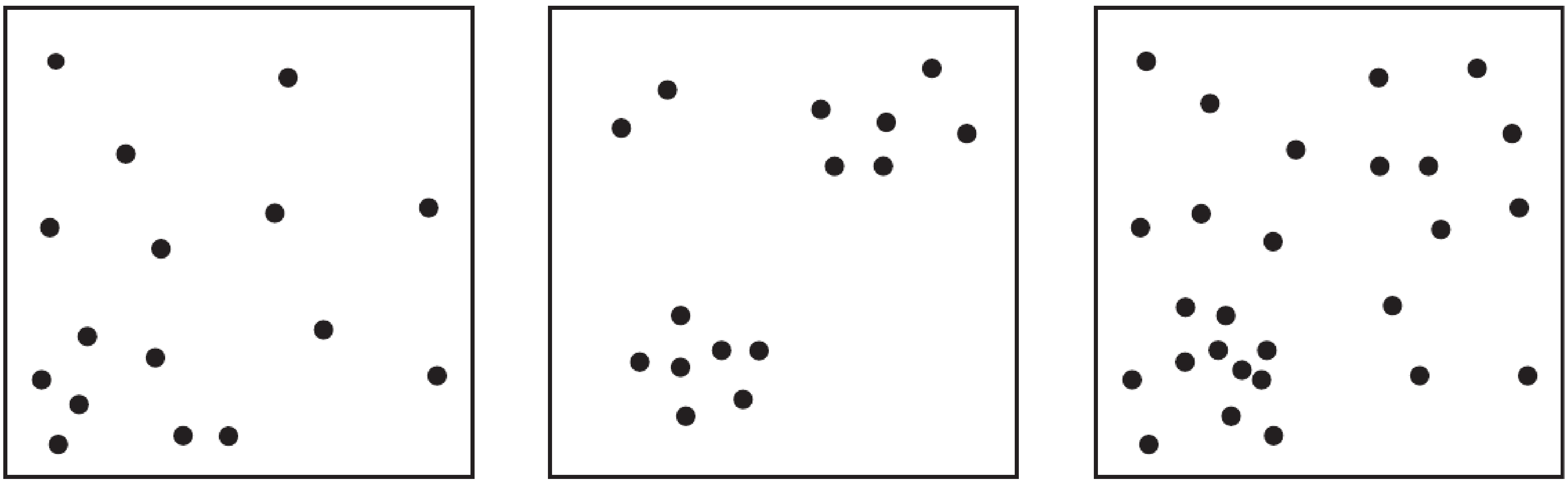
(Hengl, 2009)

Content

- Modeling and processes
- Why space matters: Potentials and Pitfalls of Spatial Data
- **Point Pattern Analyses: What processes might have caused the distribution of findings?**
  - **First order processes (density based approaches)**
  - **Second order processes (distance based approaches)**
- From single attributes to continuous surfaces: geostatistics
- Geomorphometry: What is the character of the landscape?
- Outlook

Point Pattern

What caused these patterns? First- or Second order effects?

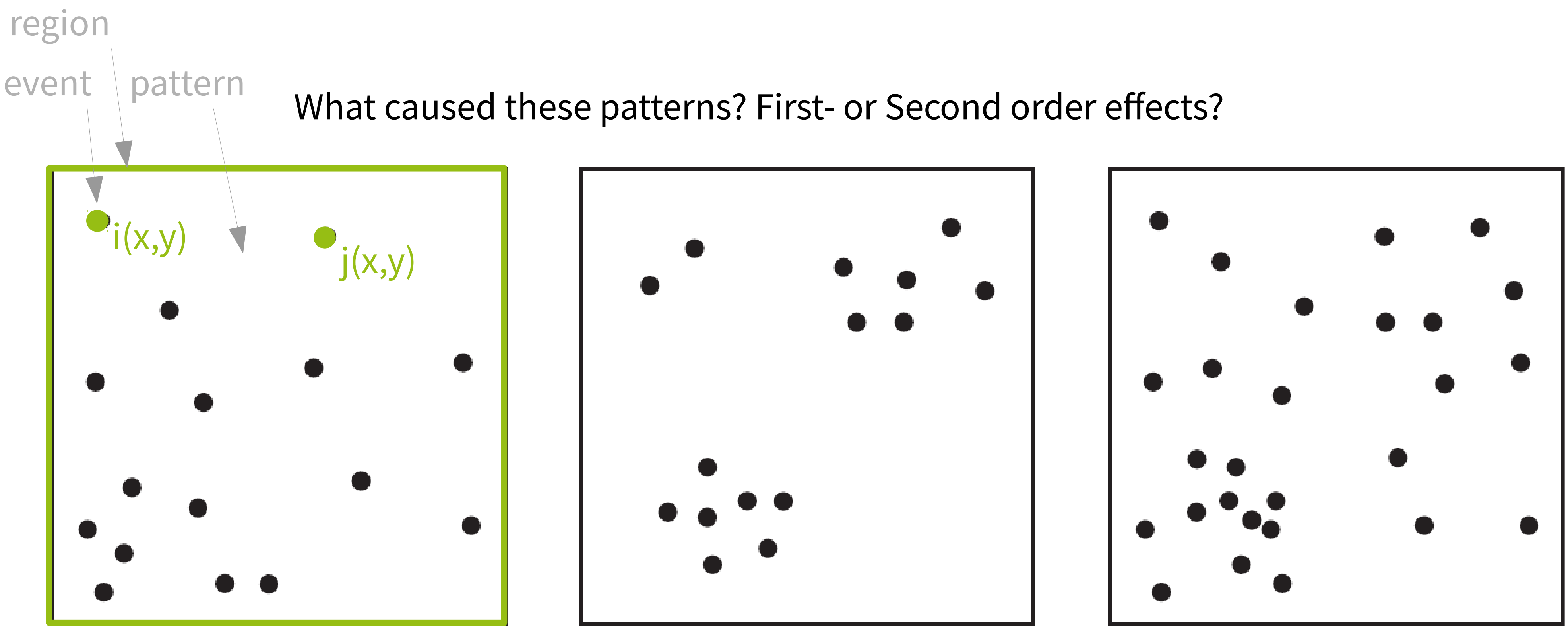


Point pattern analysis aims to distinguish between both effects
– helping you to understand the genuine process creating the pattern –

(O'Sullivan & Unwin, 2010)

Point Pattern

What caused these patterns? First- or Second order effects?



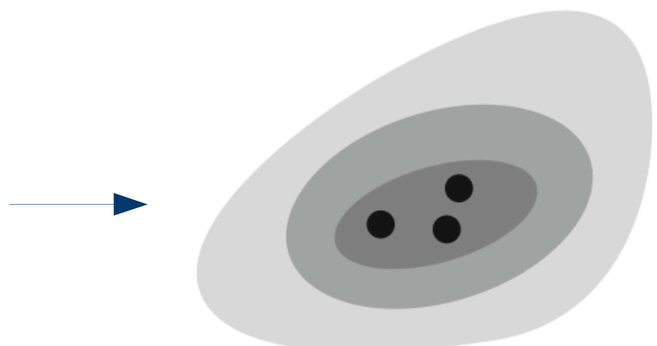
Point pattern analysis aims to distinguish between both effects
– helping you to understand the genuine process creating the pattern –

(O’Sullivan & Unwin, 2010)

Point Pattern

What caused these patterns? First- or Second order effects?

e.g. mining sites and resources



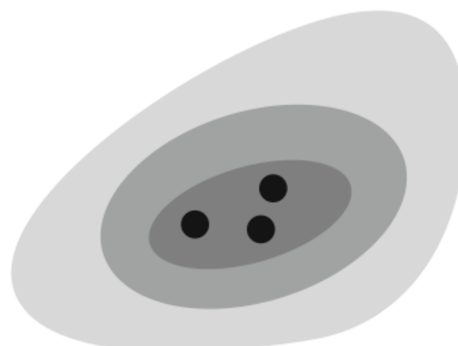
First order property: location depends on an spatial parameter

(Nakoinz & Knitter 2016)

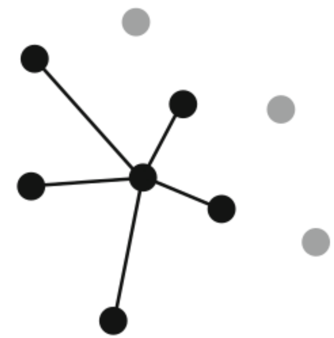
Point Pattern

What caused these patterns? First- or Second order effects?

e.g. mining sites and resources



First order property: location depends on an spatial parameter



Second order property: location depends on the relationship to other points

e.g. graveyard and settlement

(Nakoinz & Knitter 2016)

Point Pattern

Simple measures:

- mean center

- standard distance

- intensity of a pattern

$$\bar{\mathbf{s}} = (\mu_x, \mu_y) = \left(\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n} y_i}{n} \right)$$

$$d = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu_x)^2 + (y_i - \mu_y)^2}{n}}$$

$$\hat{\lambda} = \frac{n}{a} = \frac{\#(S \in A)}{a}$$

Number of events in the pattern found in study region

Area of study region

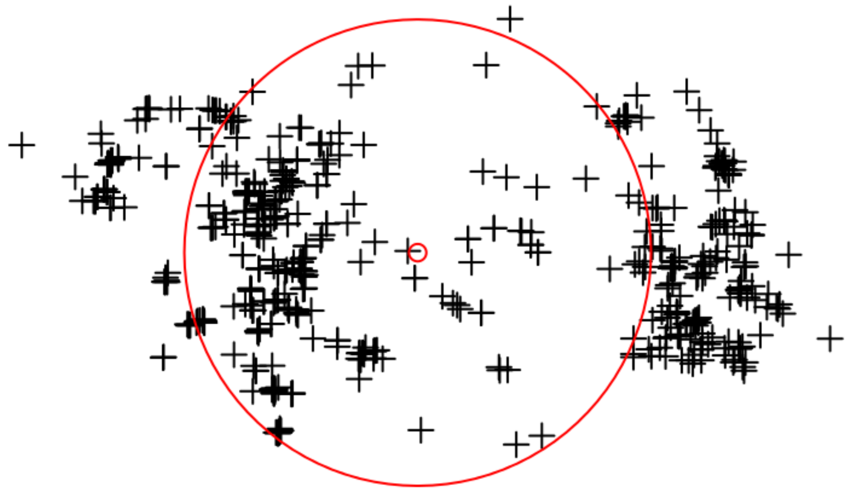
Point Pattern

Simple measures:

- mean center

- standard distance

- intensity of a pattern



0.015 sites/sqkm

Sensitive...what is the area?

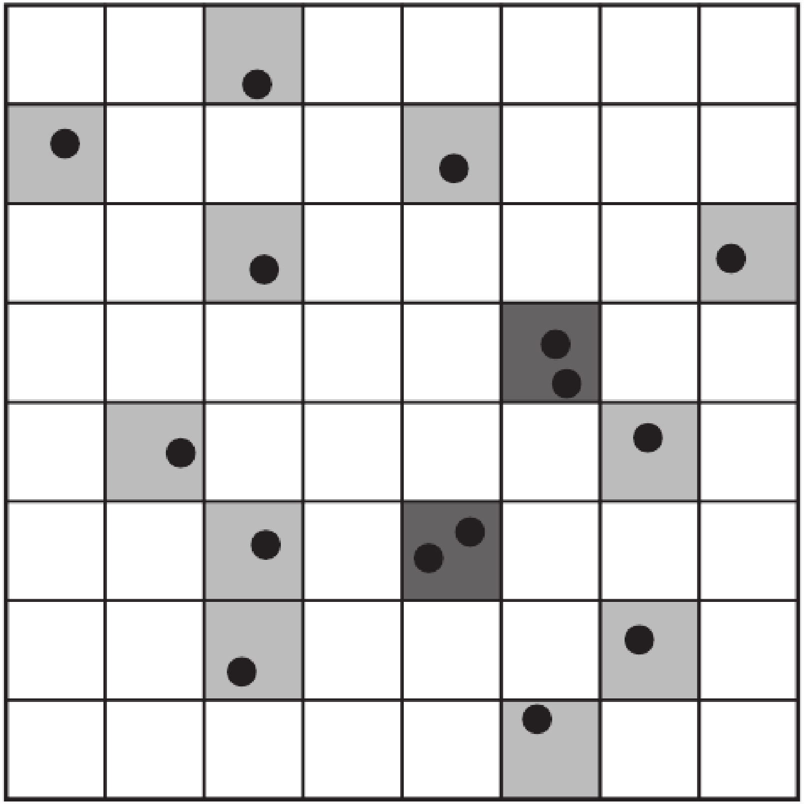
Point Pattern – first order effects

Intensity – more advanced

Independent Random Process (IRP)
= Complete Spatial Randomness (CSR)

→ any event has equal probability to be located in a quadrant
→ the occurrence of points is independent of the Positioning of other events

P (event in a quadrant) = 1/64



(O'Sullivan & Unwin, 2010)

Point Pattern – first order effects

Intensity – more advanced

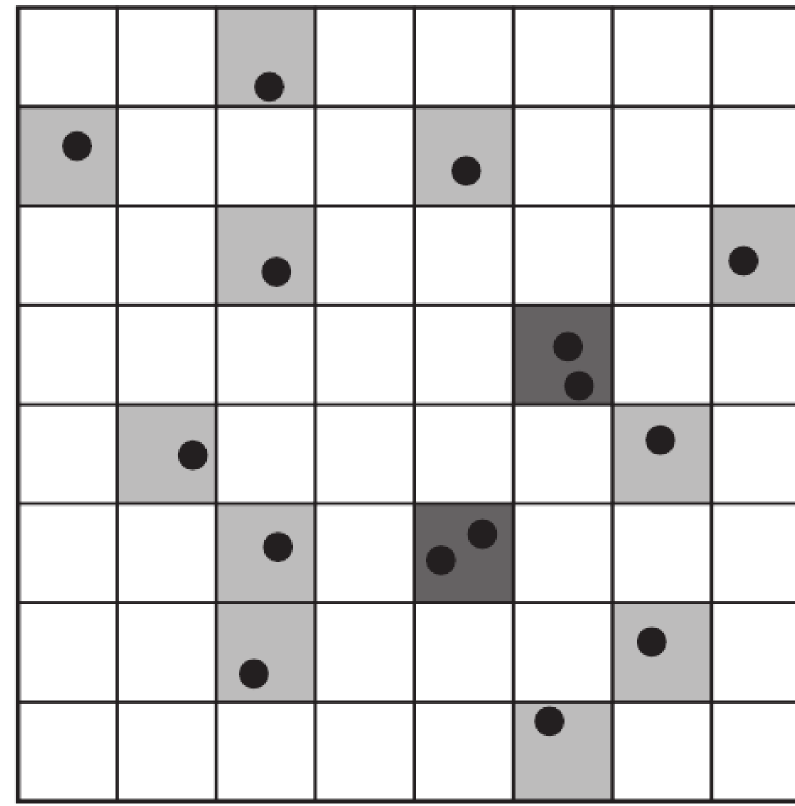
Independent Random Process (IRP)
= Complete Spatial Randomness (CSR)

Probability of *k* events in a quadrant is calculated with Poisson distribution

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

intensity

Base of natural logarithm (2.78...)

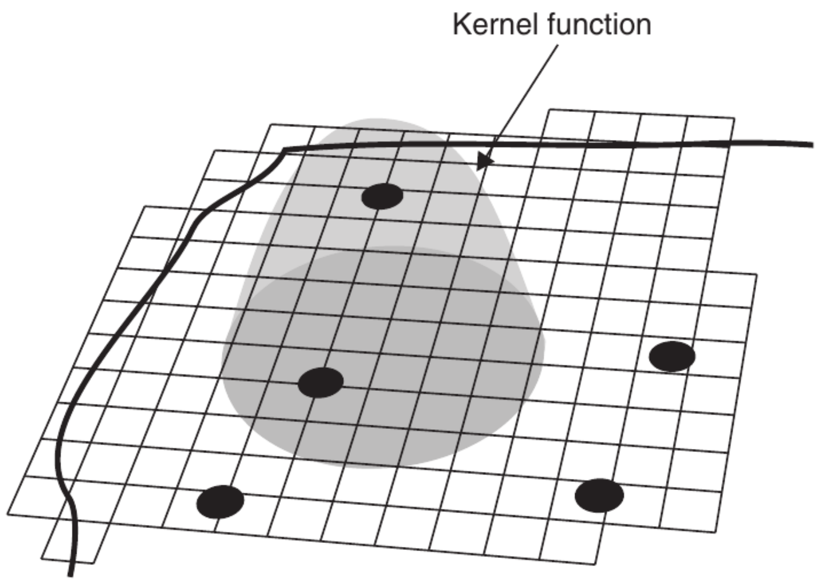


(O'Sullivan & Unwin, 2010)

Point Pattern – first order effects

Intensity – more advanced

Kernel density estimation (KDE)



Edge correction

kernel location

$$\hat{\lambda}_p = e(p) \sum_i k(x_i - p)$$

(Gaussian) Kernel function

mean

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

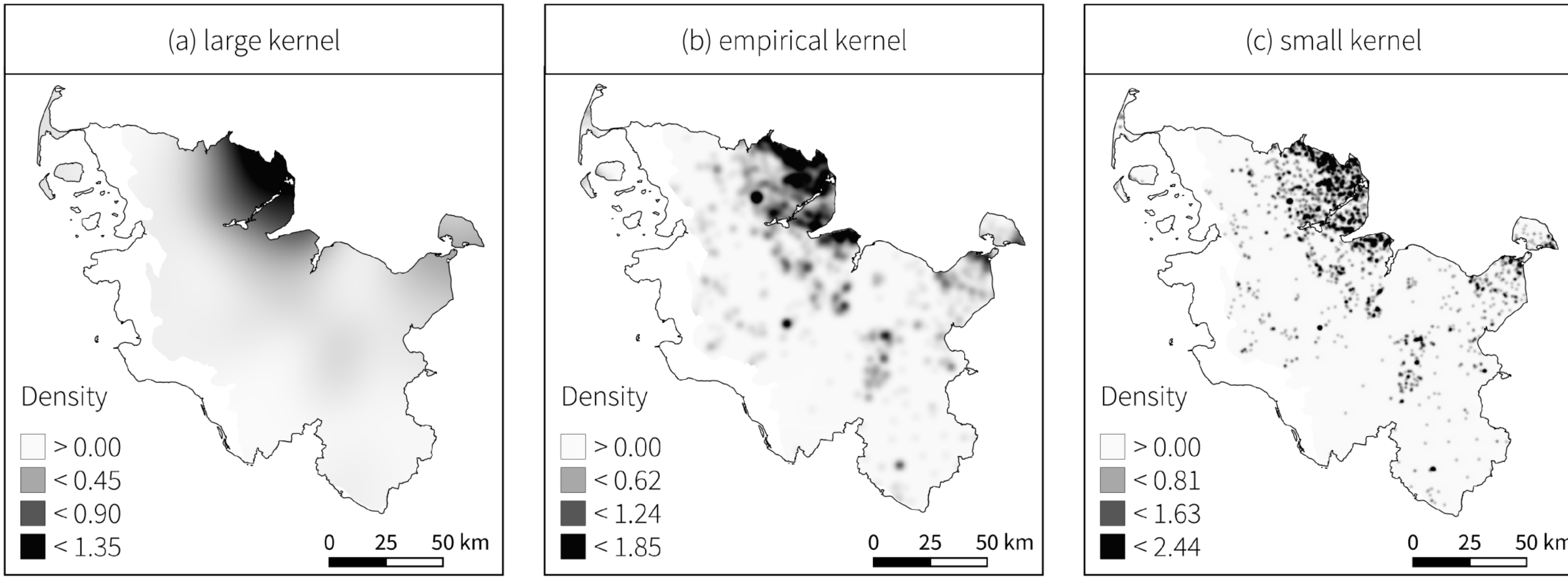
Standard deviation

(Knitter & Nakoinz (in press); (O'Sullivan & Unwin, 2010)

Point Pattern – first order effects

Sigma large

Sigma small



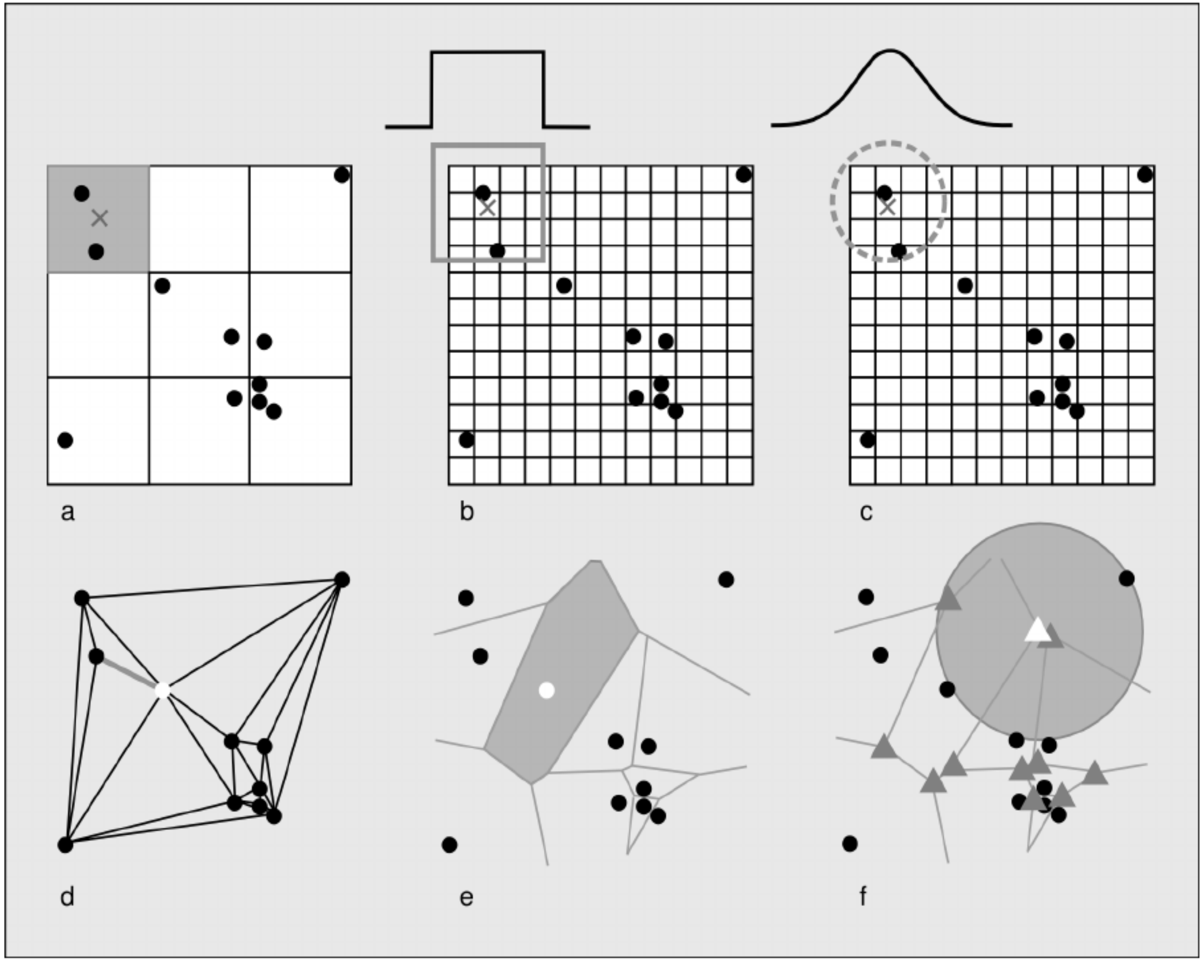
General trend

Local trend

Knitter & Nakoinz (in press)

Point Pattern – first order effects

Intensity – possibilities to calculate

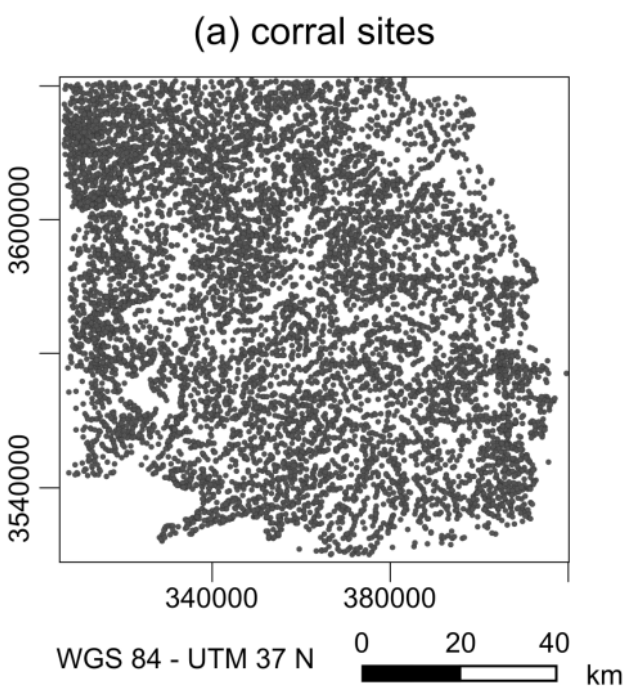


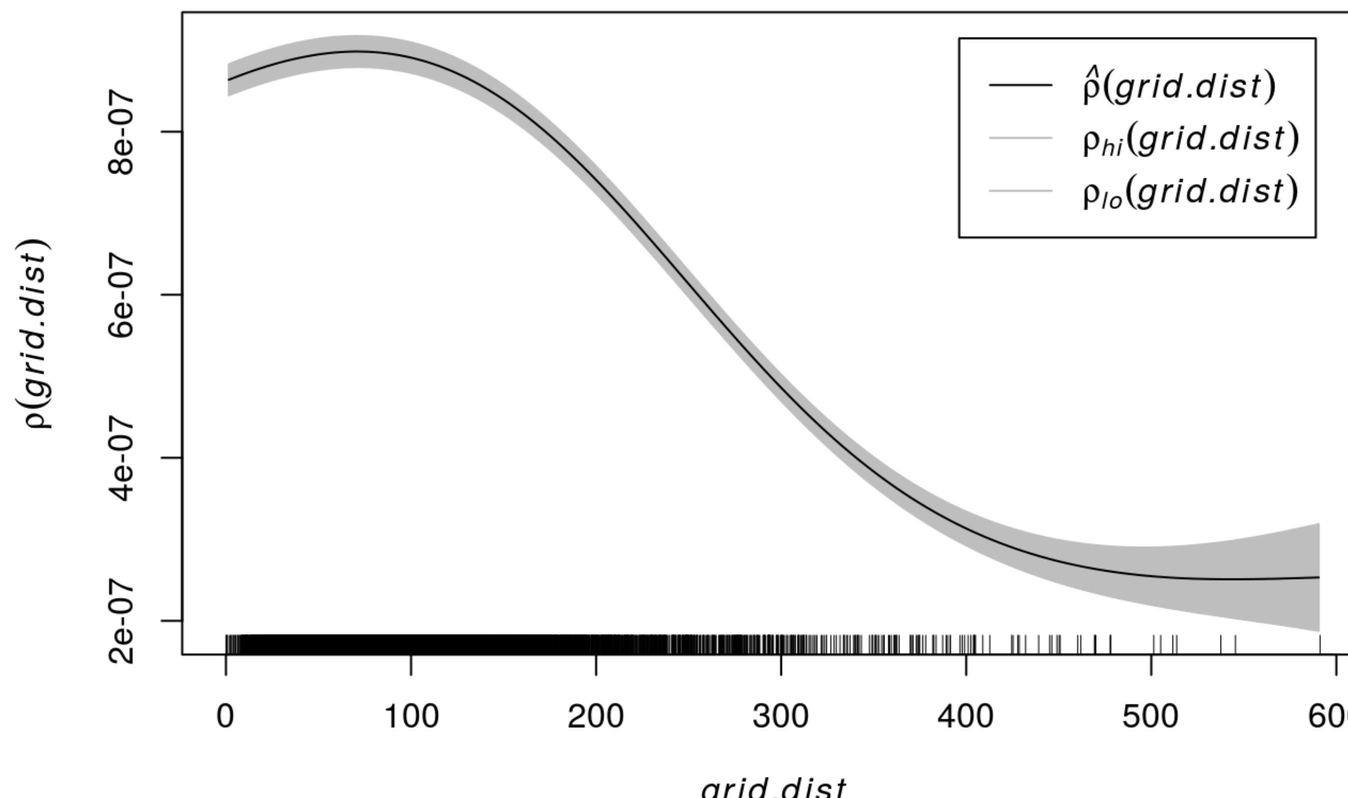
EXCELLENCE CLUSTER TOPOI

(Nakoinz & Knitter 2016)

Point Pattern – first order effects

Intensity as tool to investigate dependence on covariates

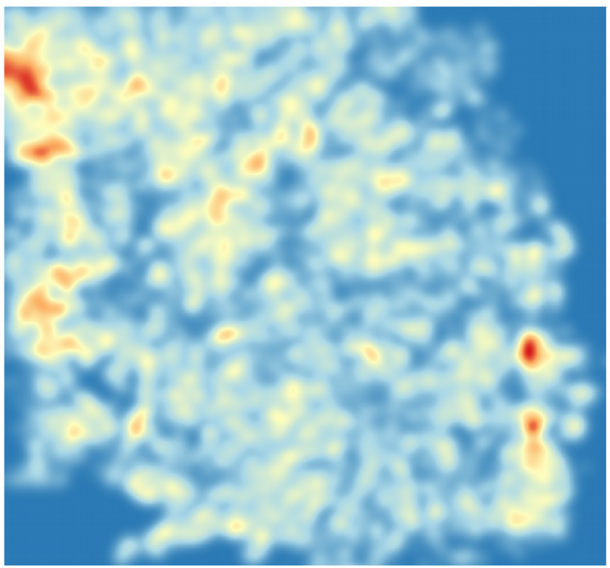


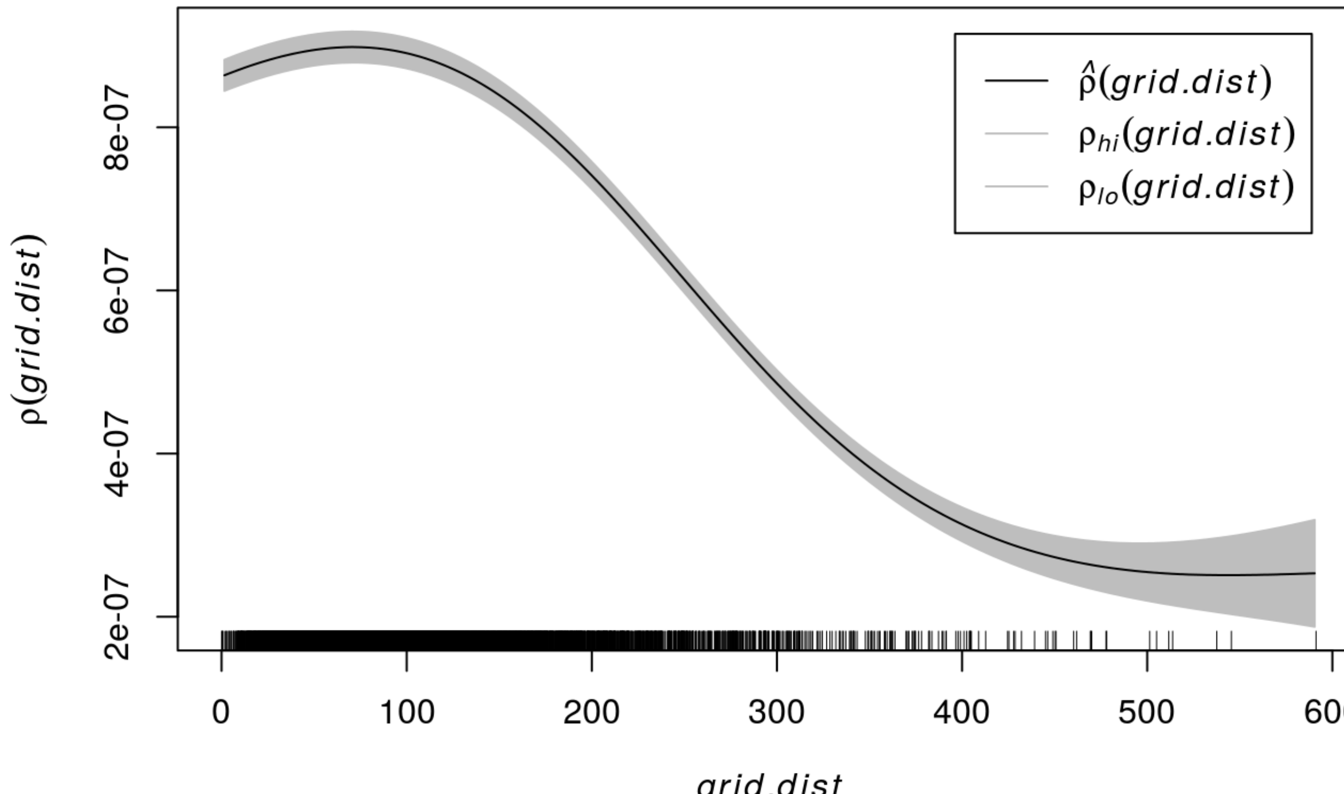


(Meister et al. forthcoming)

Point Pattern – first order effects

Intensity as tool to investigate dependence on covariates



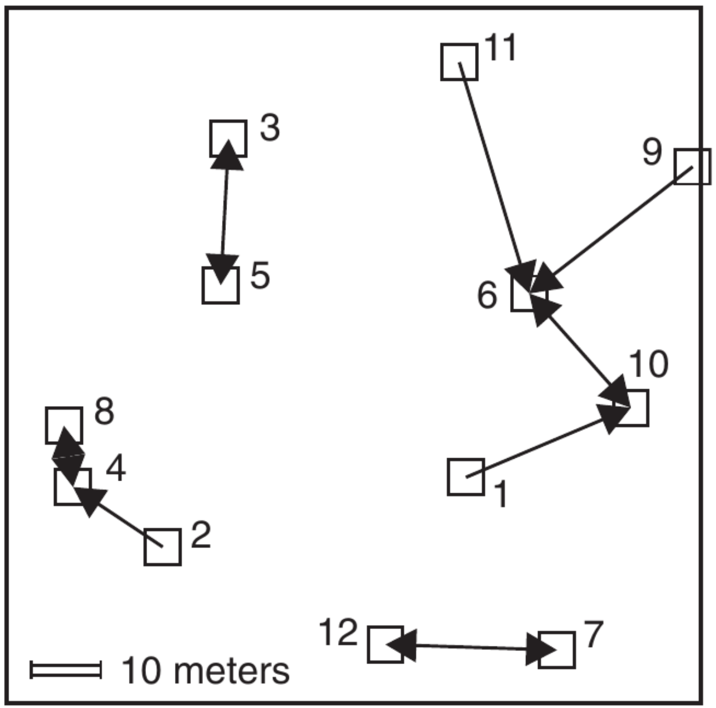


(Meister et al. forthcoming)

Point Pattern – second order effects

Point Pattern – second order effects

Point patterns described by distance measures
- nearest-neighbor distance → euclidean distance



mean nearest-neighbor distance

$$\bar{d}_{\min} = \frac{\sum_{i=1}^{n} d_{\min}(\mathbf{s}_i)}{n}$$

Are the points clustered or dispersed?
Use Clark and Evan's *R* statistic of nearest neighbor distances

$$R = \bar{d}_{\min} \Big/ \frac{1}{(2\sqrt{\lambda})}$$

expected nearest-neighbor distance

R < 1: more clustered
R > 1: more evenly spaced

(O'Sullivan & Unwin, 2010)

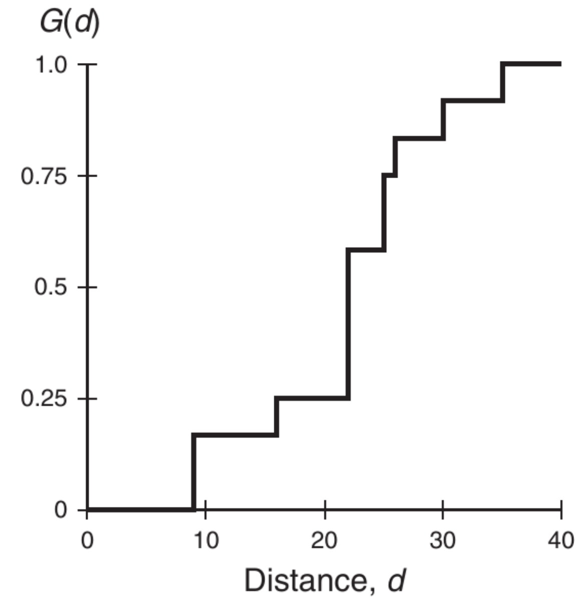
EXCELLENCE CLUSTER TOPOI

Point Pattern – second order effects

Point patterns described by distance measures
→ cumulative frequency distribution of the nearest-neighbor distances = G(d)

$$G(d) = \frac{\#(d_{\min}(\mathbf{s}_i) < d)}{n}$$

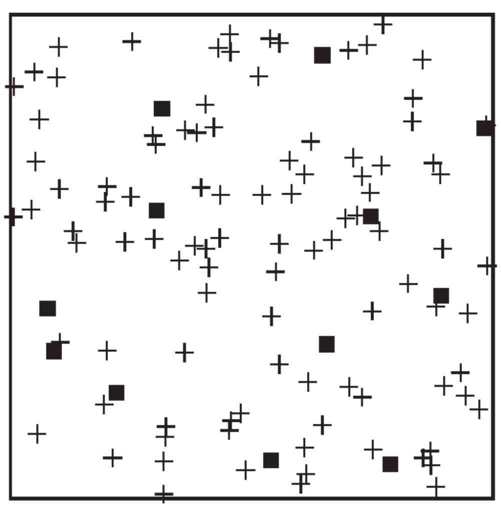
→ G(d) tells us what fraction of all n-n distances is less than d

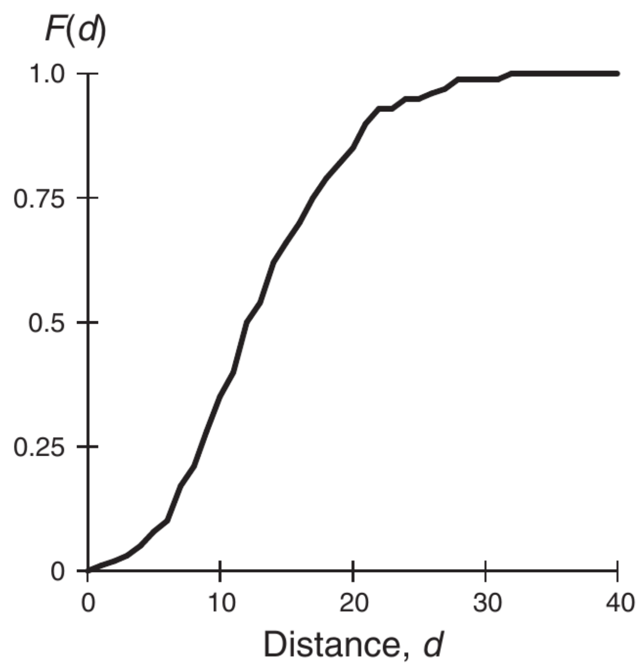


(O'Sullivan & Unwin, 2010)

Point Pattern – second order effects

Point patterns described by distance measures
→ cumulative frequency distribution of the nearest-neighbor distances of *arbitrary events* to known events = F(d)





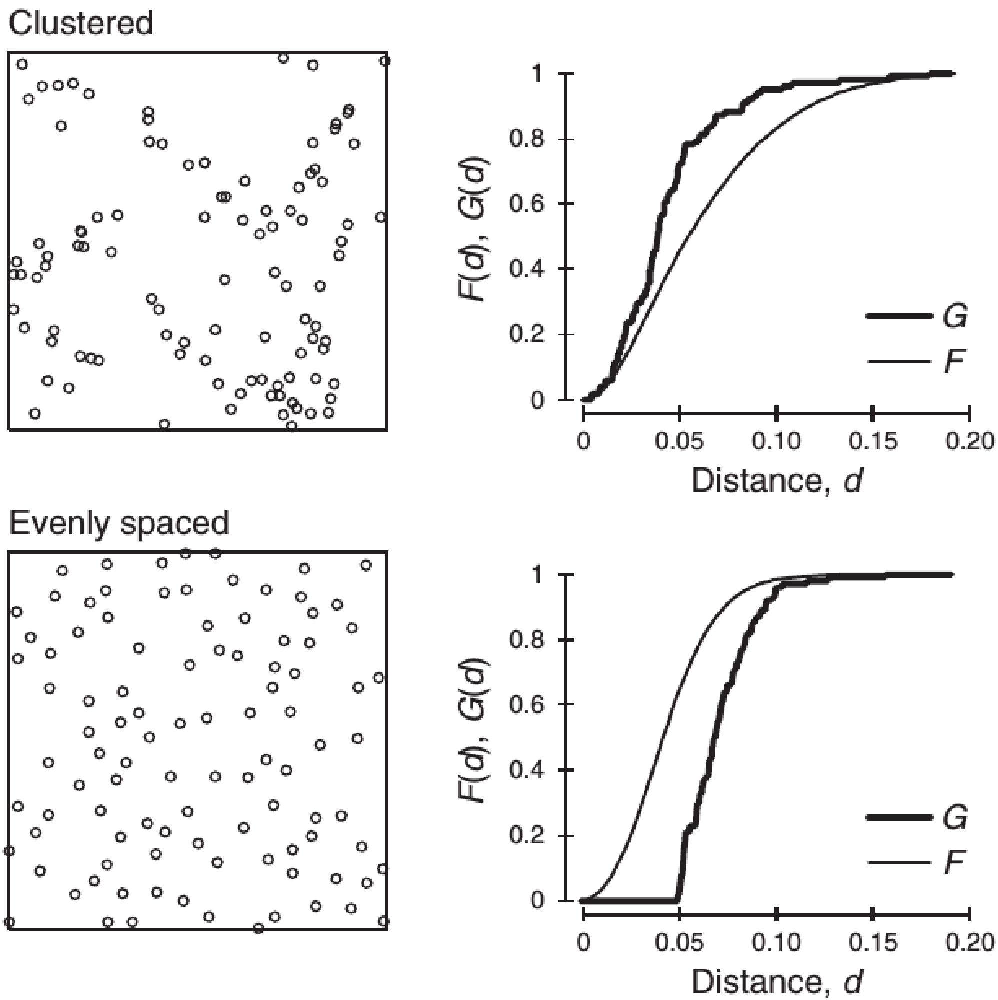
minimum distance from random point pi to an event

$$F(d) = \frac{\#[d_{\min}(\mathbf{p}_i, S) < d]}{m}$$

set of randomly selected locations

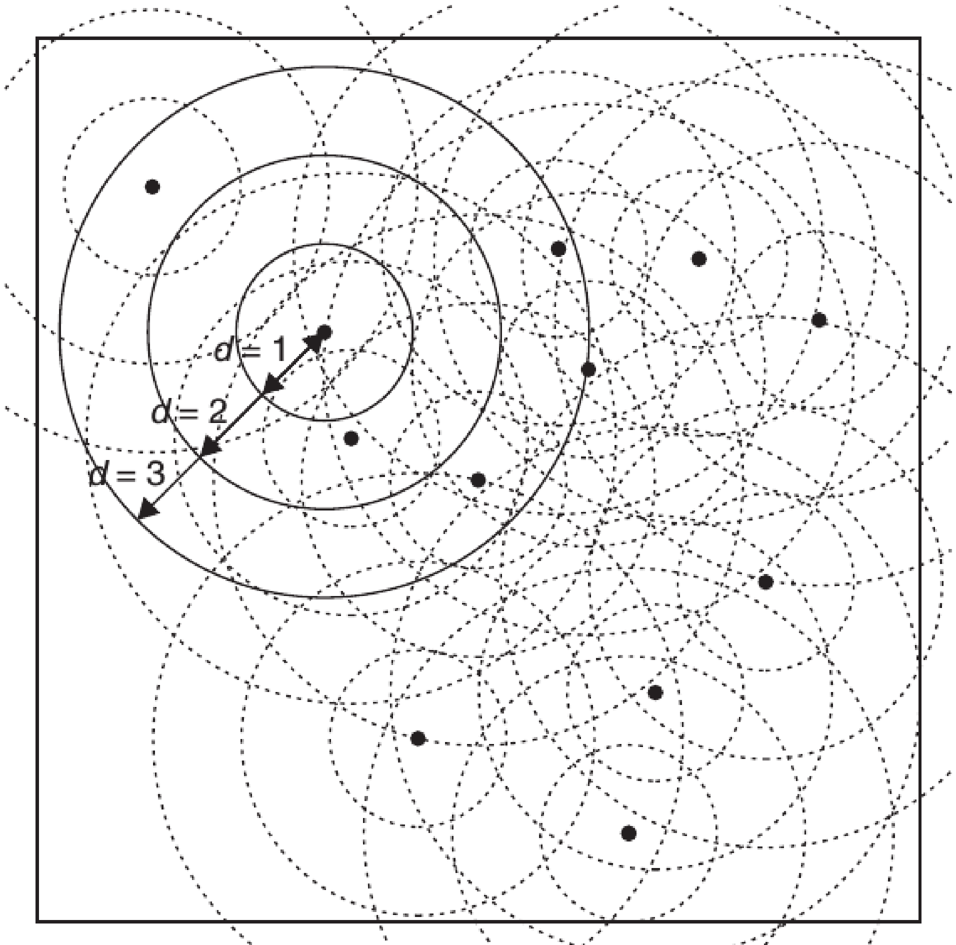
(O'Sullivan & Unwin, 2010)

Point Pattern – second order effects



(O'Sullivan & Unwin, 2010)

Point Pattern – second order effects



To get rid of the nearest-neighbor limitation: use the K Function

Events in circle radius d centered at s

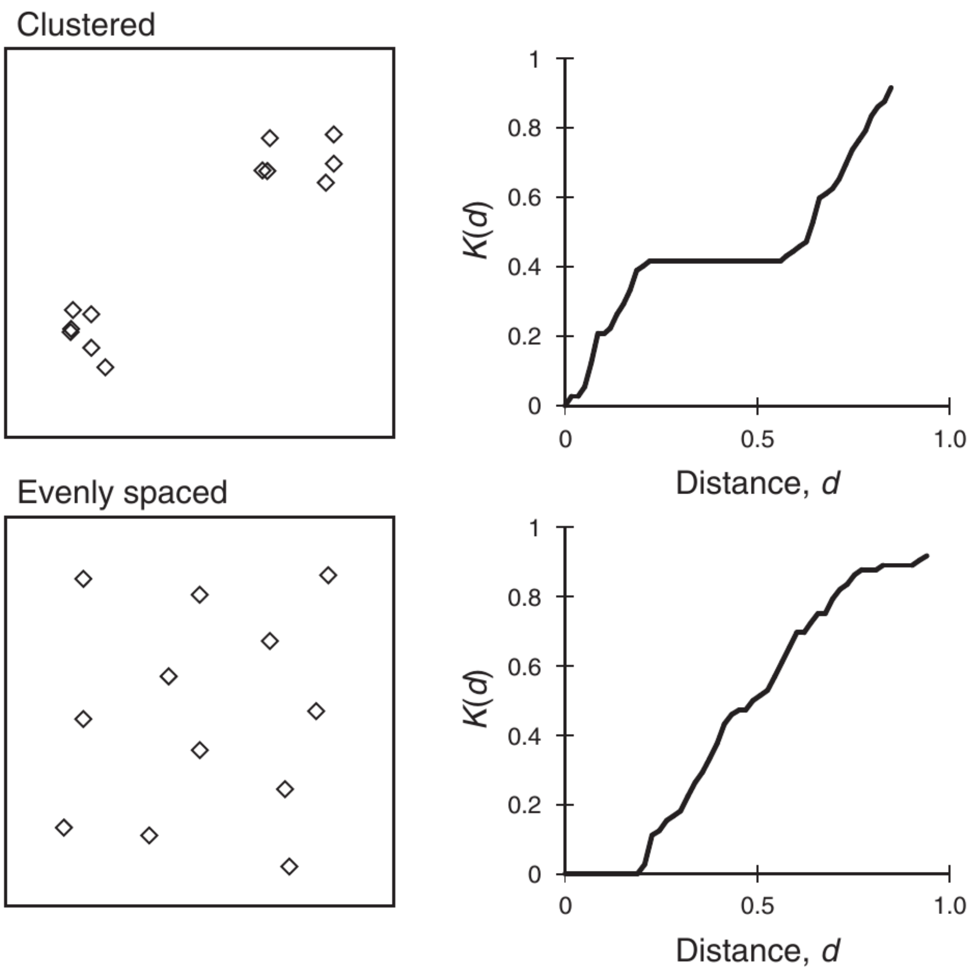
$$K(d) = \frac{\sum_{i=1}^{n} \#[S \in C(\mathbf{s}_i, d)]}{n\lambda}$$

Event density in the study area/region

EXCELLENCE CLUSTER TOPOI

(O'Sullivan & Unwin, 2010)

Point Pattern – second order effects



To get rid of the nearest-neighbor limitation: use the K Function

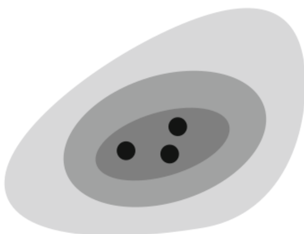
Events in circle radius d centered at s

$$K(d) = \frac{\sum_{i=1}^{n} \#[S \in C(\mathbf{s}_i, d)]}{n\lambda}$$

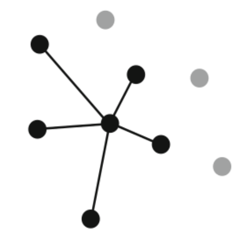
Event density in the study area/region

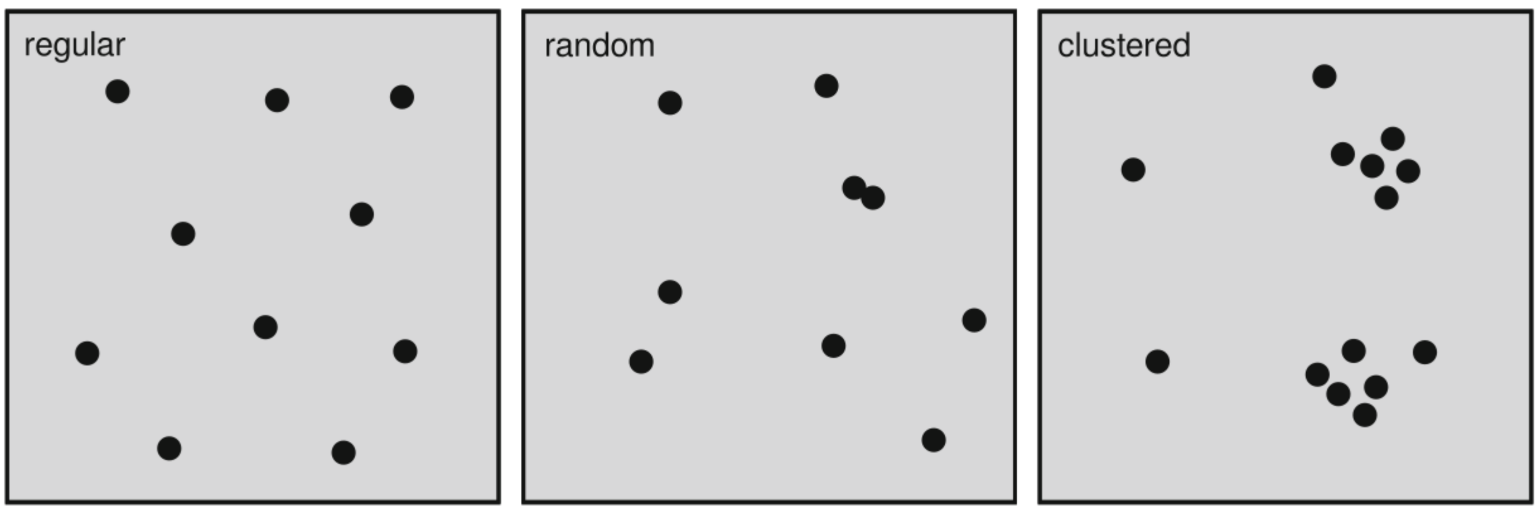
(O'Sullivan & Unwin, 2010)

Point Pattern Analyses



Help to distinguish between first and second order effects that caused your spatial pattern at hand
→ you gain insights in the underlying processes





Content

- Modeling and processes
- Why space matters: Potentials and Pitfalls of Spatial Data
- Point Pattern Analyses: What processes might have caused the distribution of findings?
  - First order processes (density based approaches)
  - Second order processes (distance based approaches)
- **From single attributes to continuous surfaces: geostatistics**
- Geomorphometry: What is the character of the landscape?
- Outlook

Interpolation

We start with points

But wait…

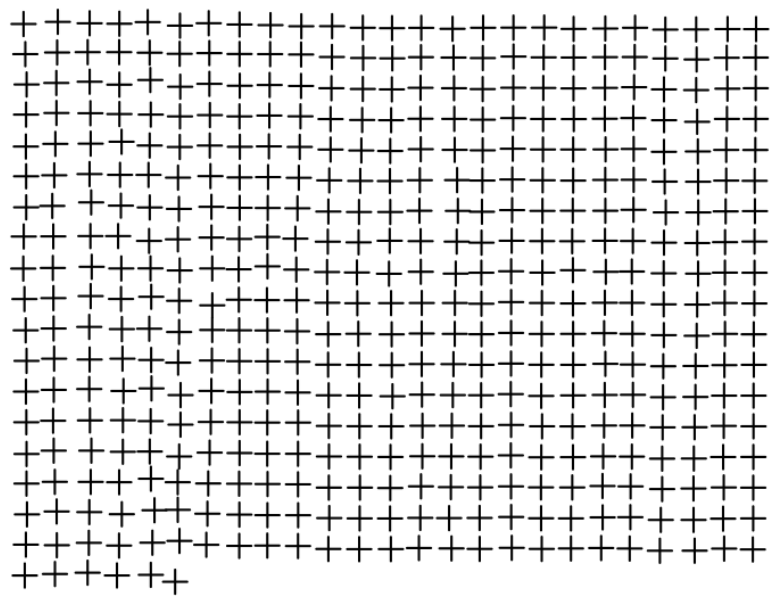
Why are the methods of point patterns not used??

Because *point data* is not necessarily a *point pattern*

e.g. sampling locations of soil samples to analyze phosphorous (=points)
are *artificial* and irrelevant to the study of phosphorous distribution
→ such points do *not* reflect a process.

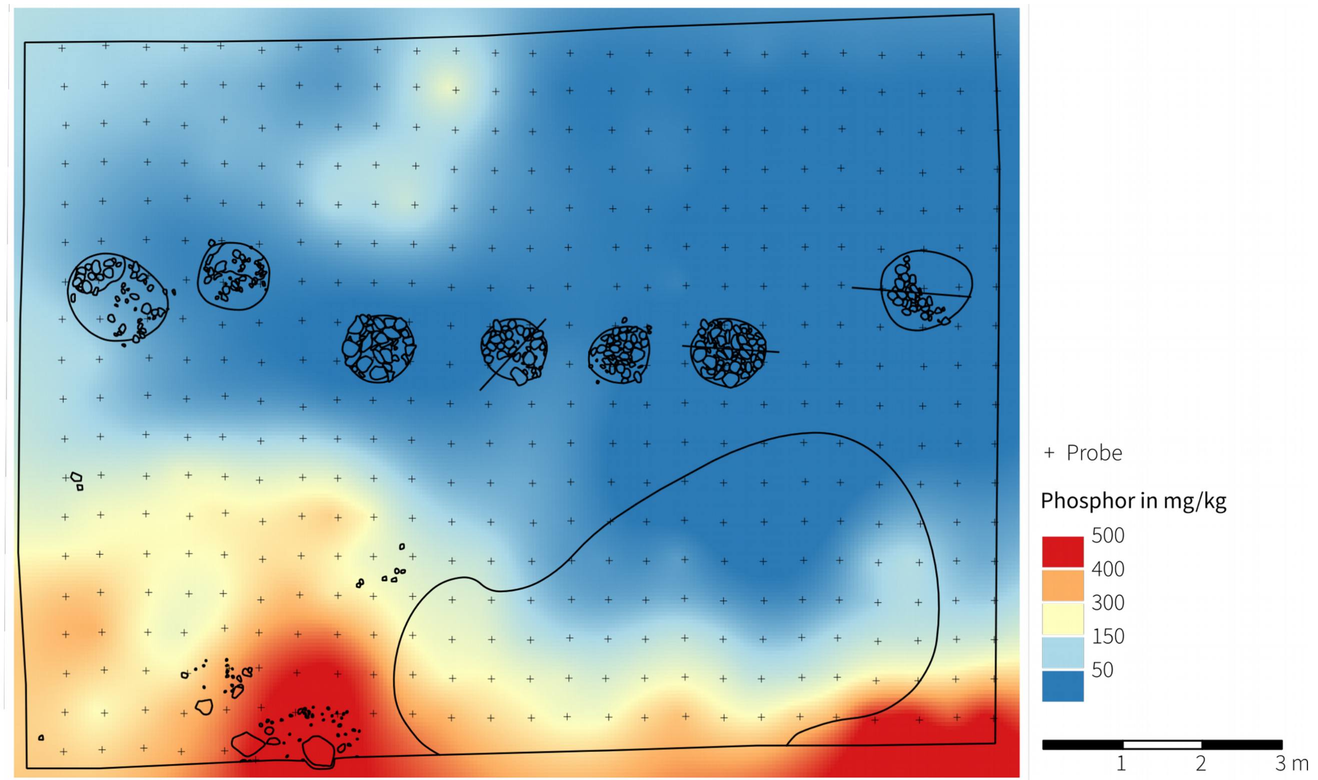
Interpolation

Aim: get a continuous representation of attribute information.



Interpolation

Aim: get a continuous representation of attribute information.

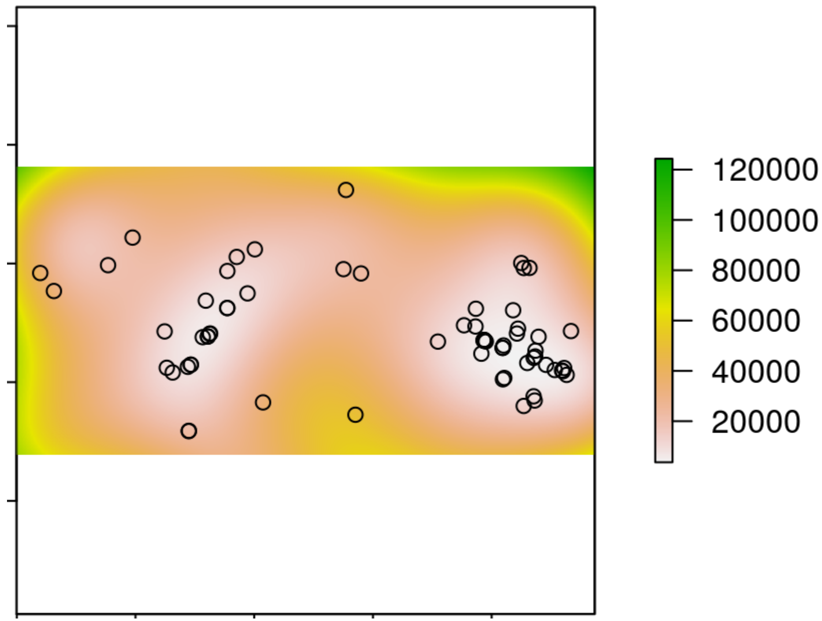


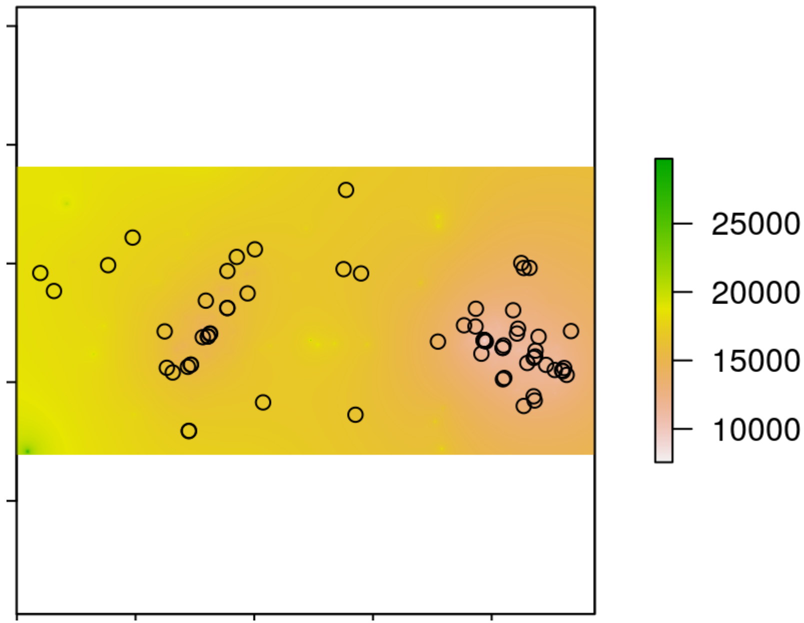
Hoelzmann & Knitter forthcoming

Interpolation

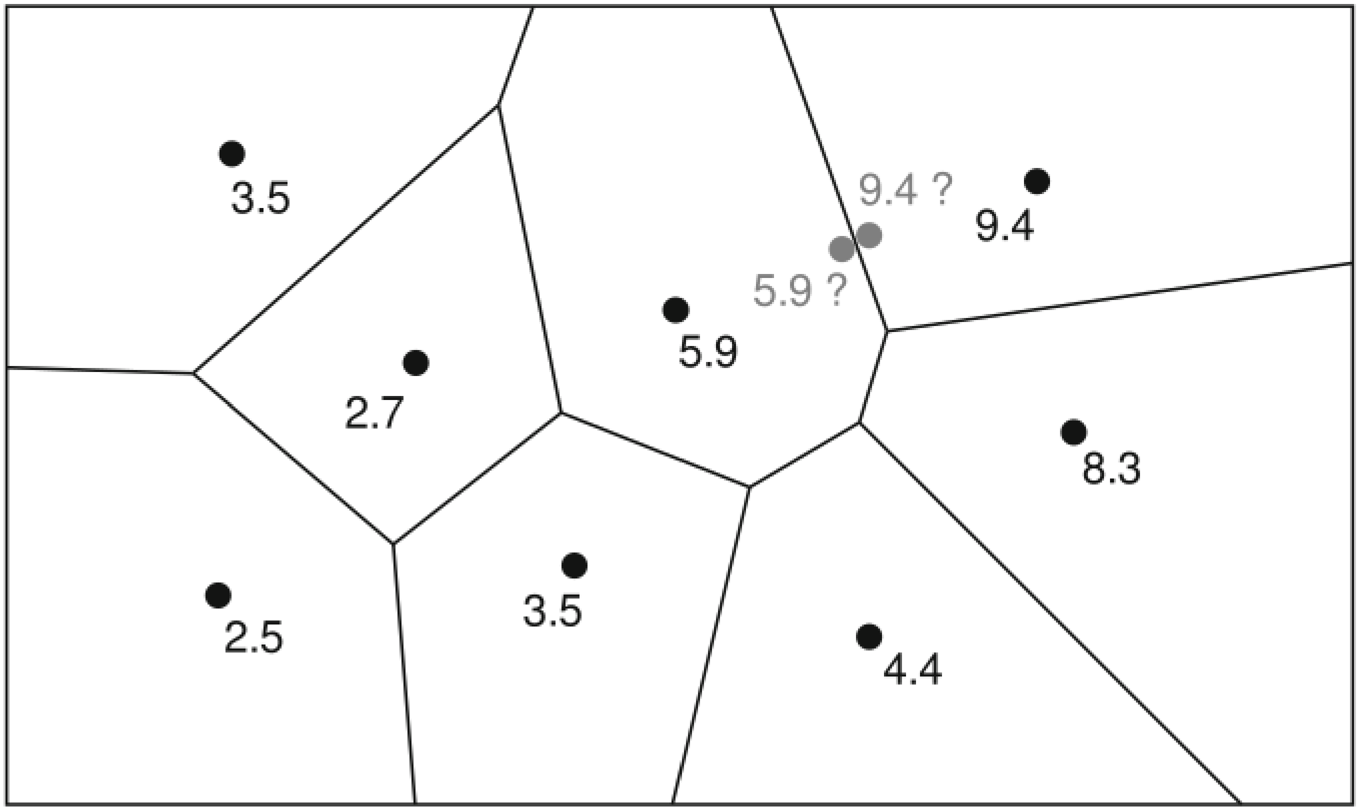
Aim: get a continuous representation of settlement density.

Mission accomplished…!?





Interpolation



Interpolation

Inverse Distance Weighting

Control points *i* in neighborhood *m*

$$\hat{z}_j = \sum_{i=1}^{m} w_{ij} z_i$$

Estimate at *j*

$$w_{ij} \propto \frac{1}{d_{ij}^k}$$

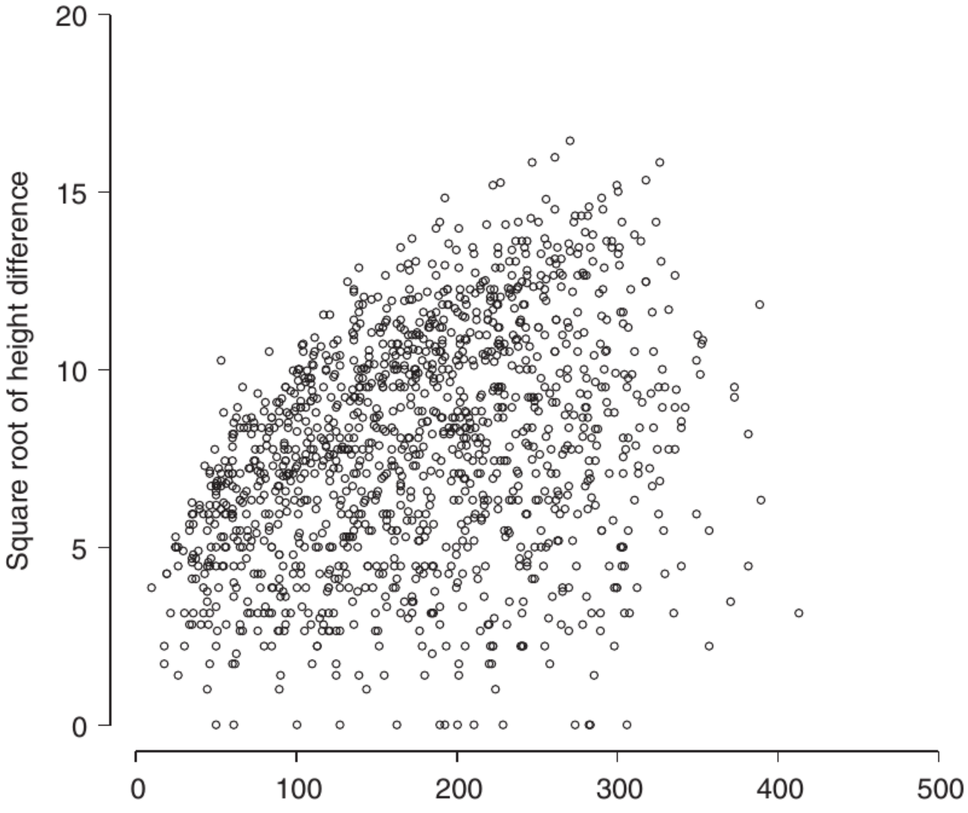
(O'Sullivan & Unwin, 2010)

Interpolation

Kriging

Square root differences cloud

1. Describing spatial variation



(O'Sullivan & Unwin, 2010)

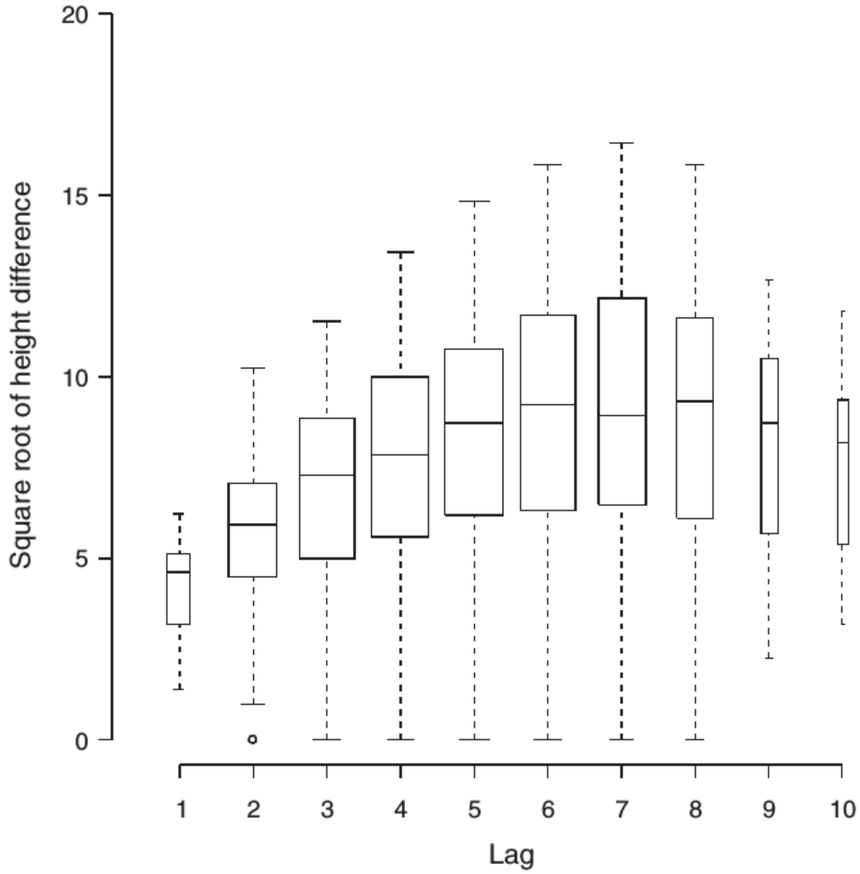
Interpolation

Kriging

Square root differences cloud

Distance classes → lags
and summary statistics

1. Describing spatial variation



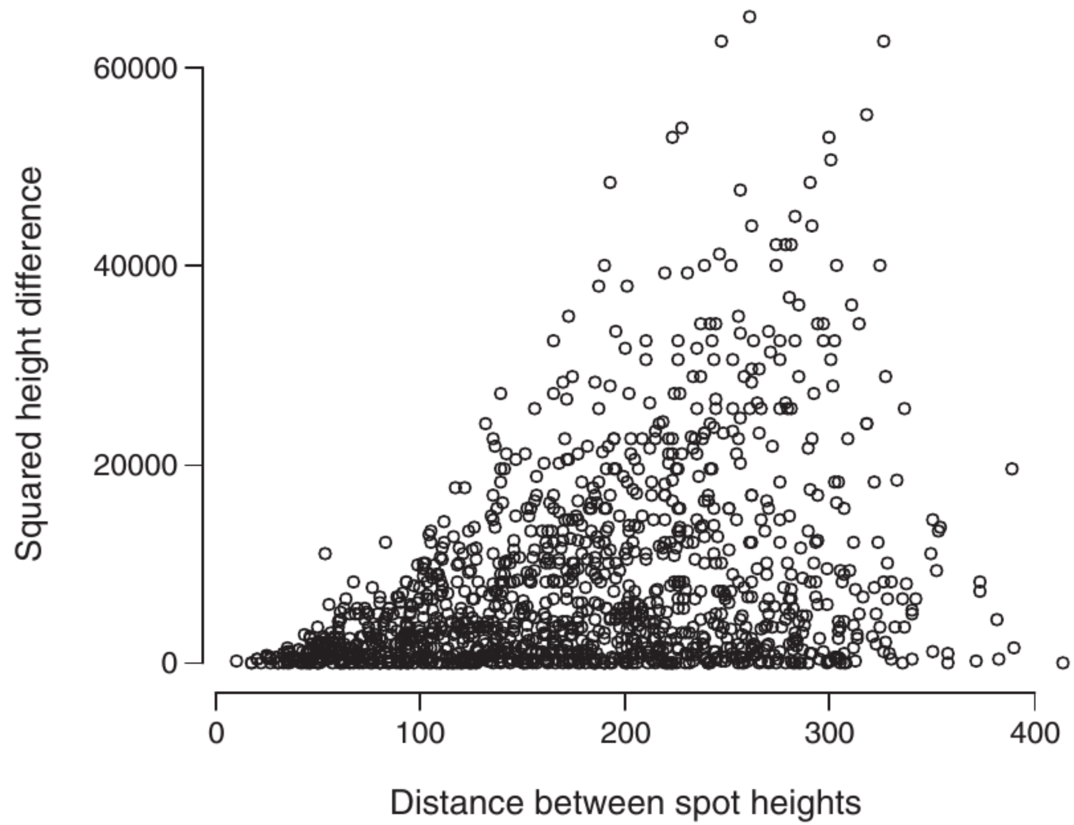
(O'Sullivan & Unwin, 2010)

Interpolation

Kriging

1. Describing spatial variation

Square differences, i.e. semivariogram cloud



(O'Sullivan & Unwin, 2010)

Interpolation

Kriging

Square differences, i.e. semivariogram cloud

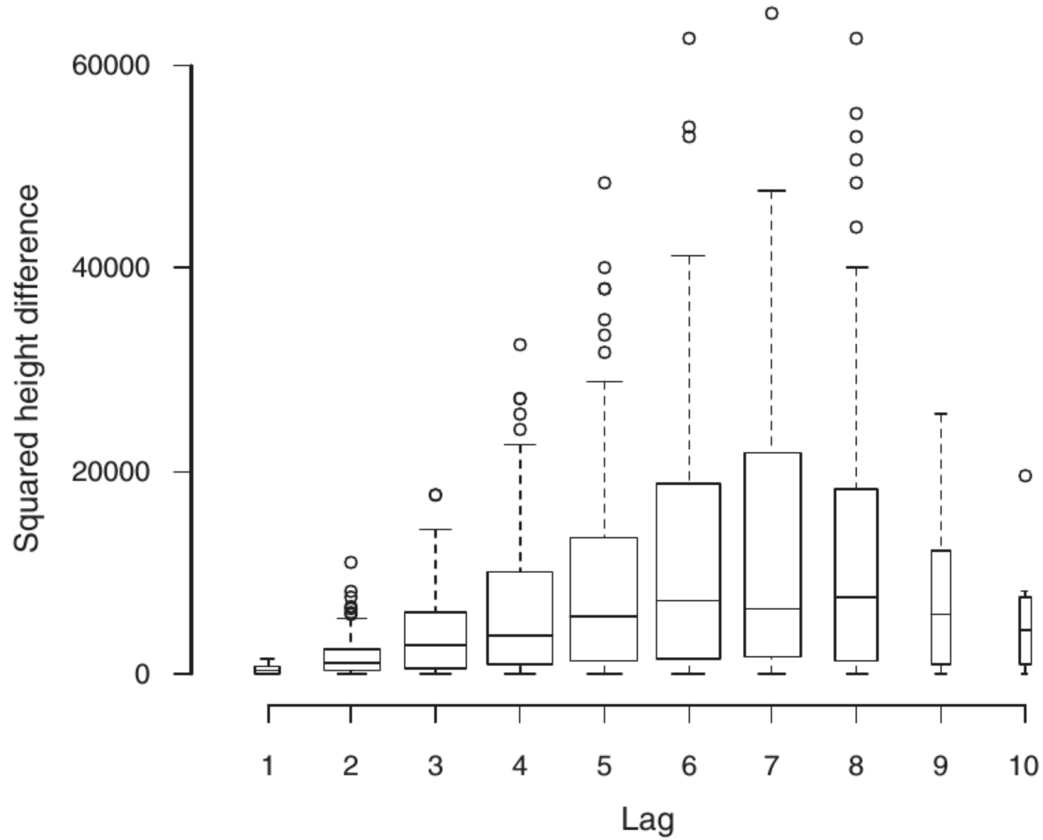
Distance classes → lags
and summary statistics
→ *experimental semivariogram*
[or short variogram]

Semivariance

$$2\hat{\gamma}(d) = \frac{1}{n(d)} \sum_{d_{ii}=d} \left(z_i - z_j\right)^2$$

= mean of sum of squared differences at distance d

1. Describing spatial variation



(O'Sullivan & Unwin, 2010)

Interpolation

Kriging

Square differences, i.e. semivariogram cloud

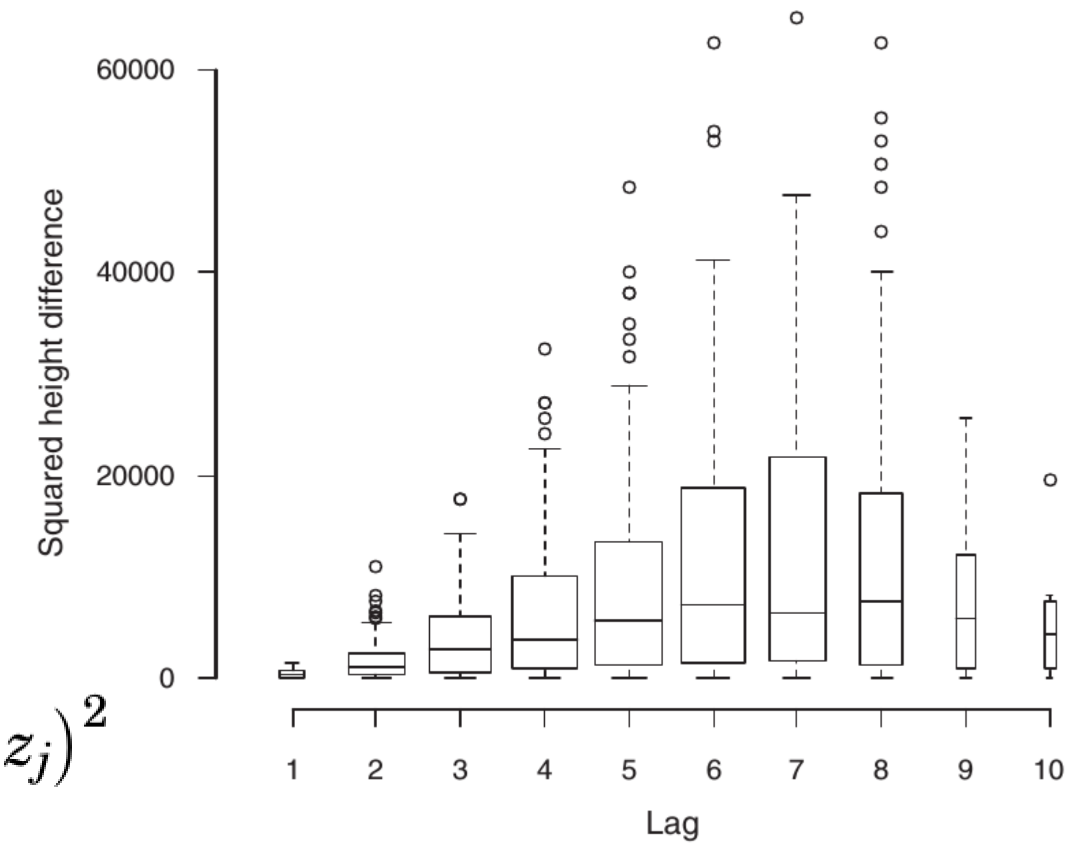
Distance classes → lags
and summary statistics
→ *experimental semivariogram*
[or short variogram]

Semivariance

$$2\hat{\gamma}(d) = \frac{1}{n(d \pm \Delta/2)} \sum_{d \pm \Delta/2} \left(z_i - z_j\right)^2$$

= mean of sum of squared differences at distance lags

1. Describing spatial variation



(O'Sullivan & Unwin, 2010)

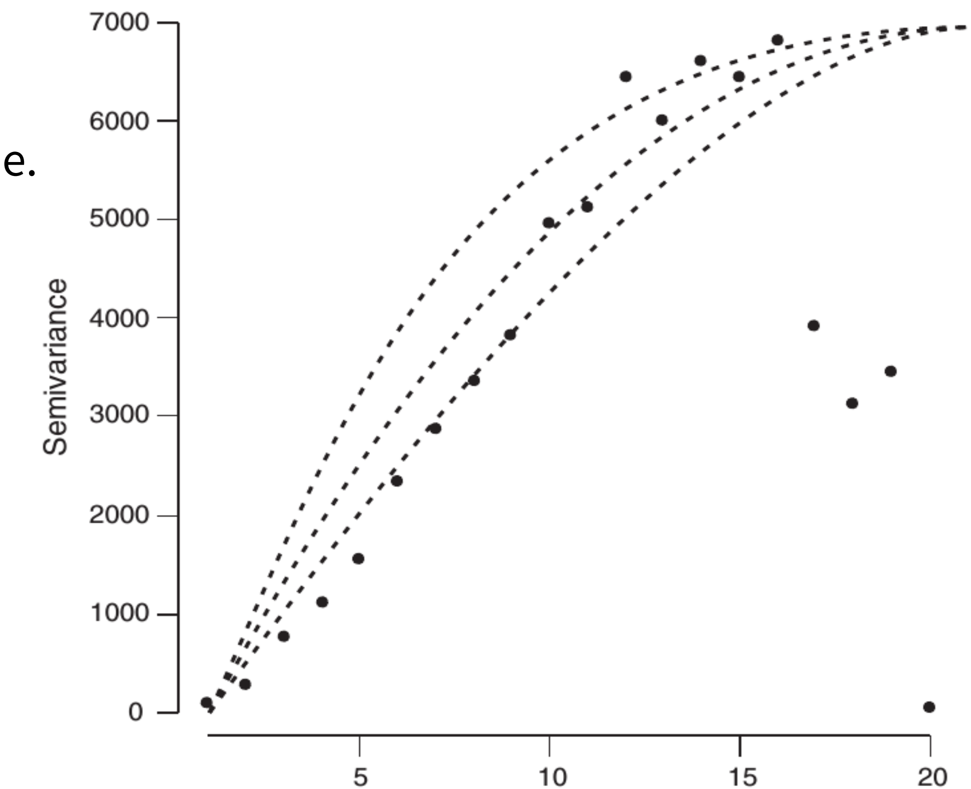
Interpolation

Kriging

Fitting a function to the experimental estimates from the semivariogram

Any function? No. Function has to be authorized, i.e.
→ function can only be positive
→ intercept at zero

2. Summarizing spatial variation



(O'Sullivan & Unwin, 2010)

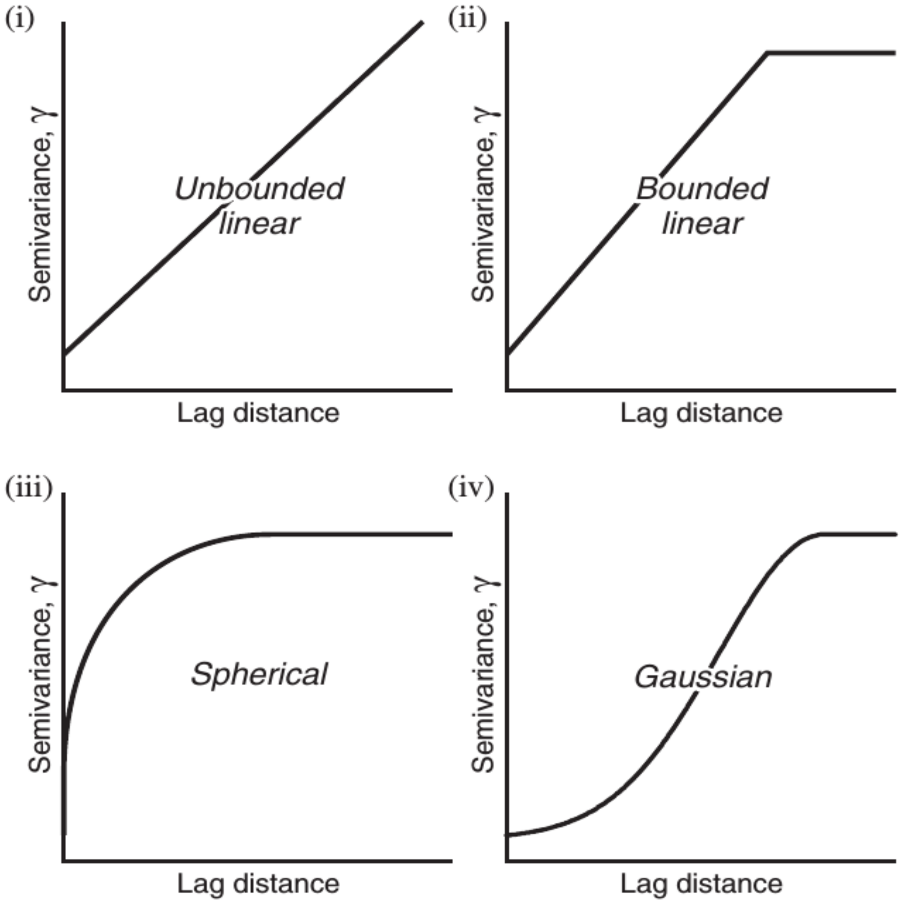
Kriging

Interpolation

Fitting a function to the experimental estimates from the semivariogram

Exemplary functions to fit

2. Summarizing spatial variation

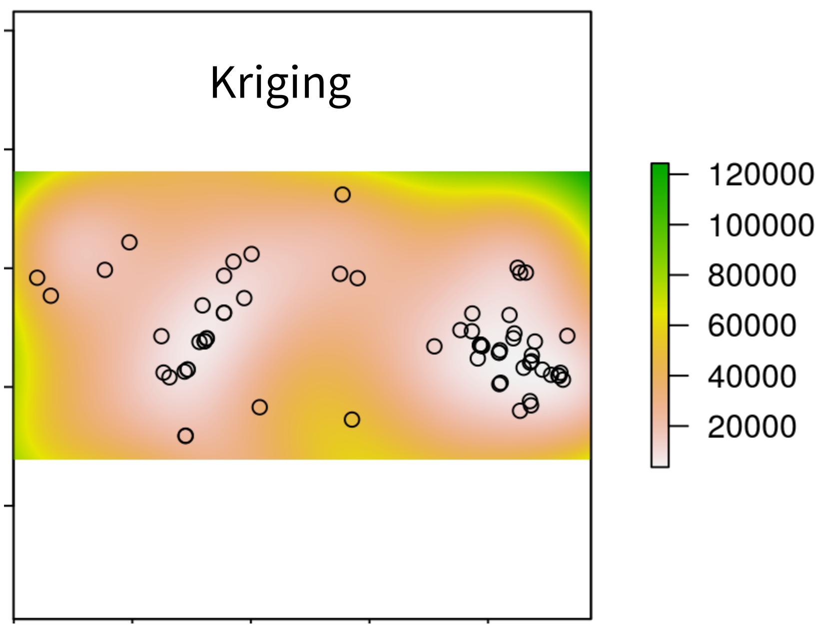


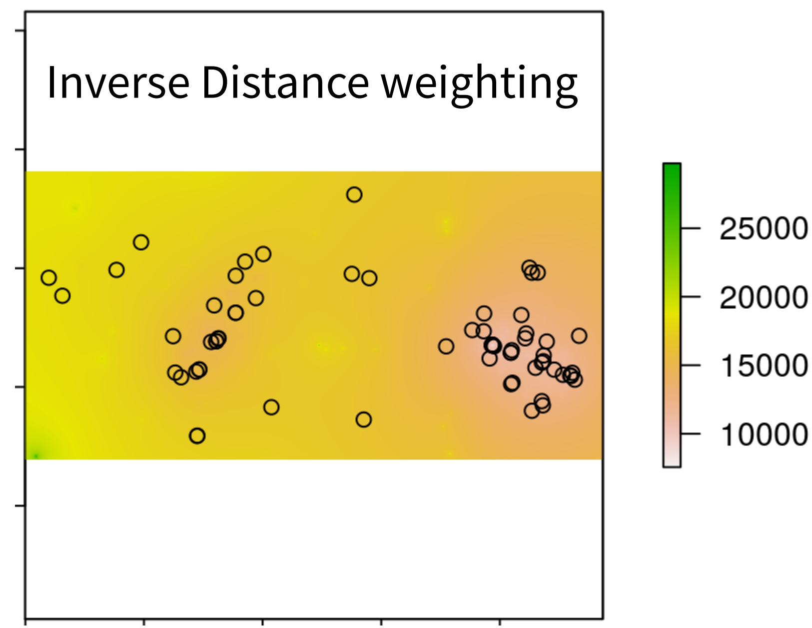
(O'Sullivan & Unwin, 2010)

Interpolation

Aim: get a continuous representation of settlement density.

Mission accomplished…!?





Content

- Modeling and processes
- Why space matters: Potentials and Pitfalls of Spatial Data
- Point Pattern Analyses: What processes might have caused the distribution of findings?
  - First order processes (density based approaches)
  - Second order processes (distance based approaches)
- From single attributes to continuous surfaces: geostatistics
- **Geomorphometry: What is the character of the landscape?**
- Outlook

Terrain analyses

Geomorphometry: the science of the quantitative representation of the Earth's surface

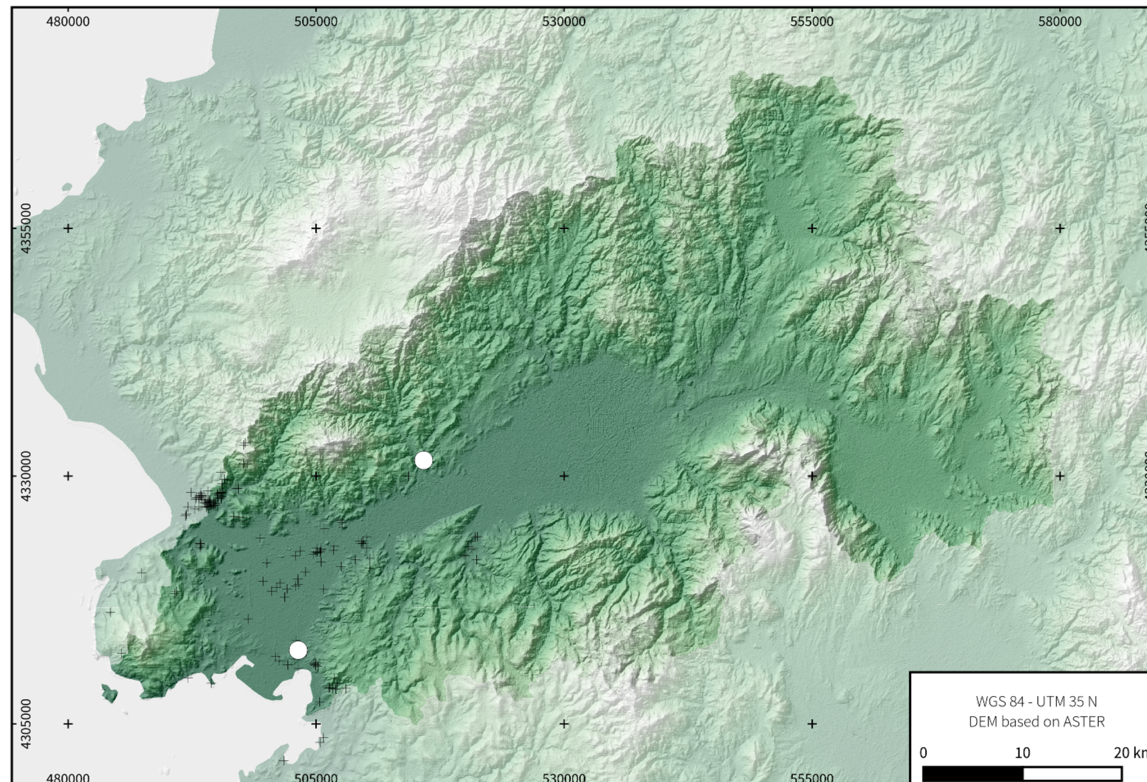
A great starting point:

Tomislav Hengl and Hannes I. Reuter (eds):
Geomorphometry Concepts, Software, Applications.
Developments in Soil Science, Volume 33, Pages 1-765 (2009)

Terrain analyses

Geomorphometry: the science of the quantitative representation of the Earth's surface

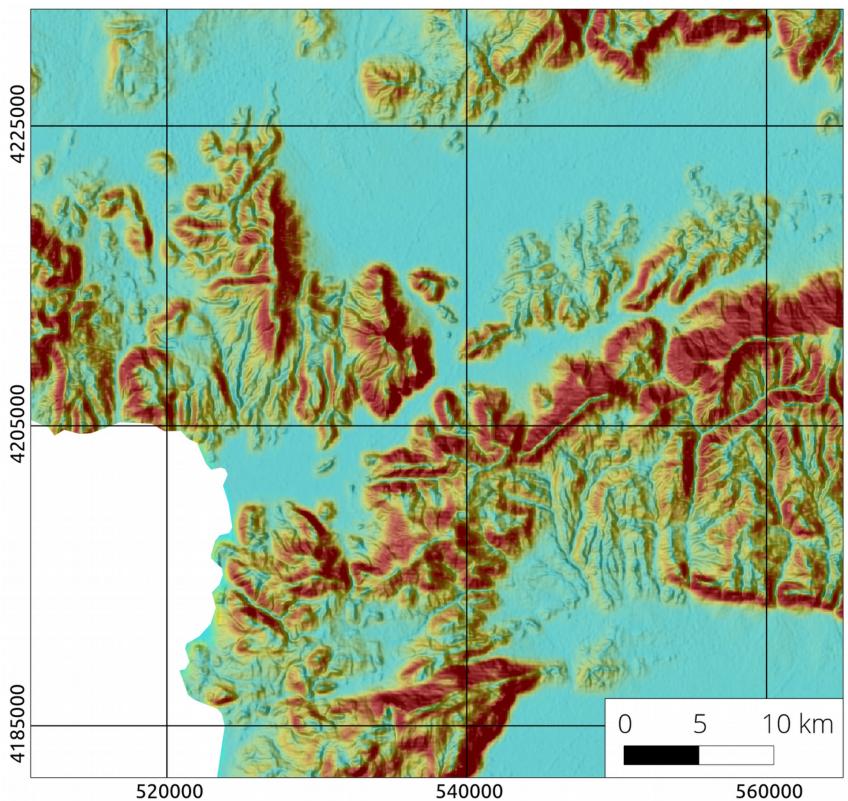
Everything starts with a DEM…

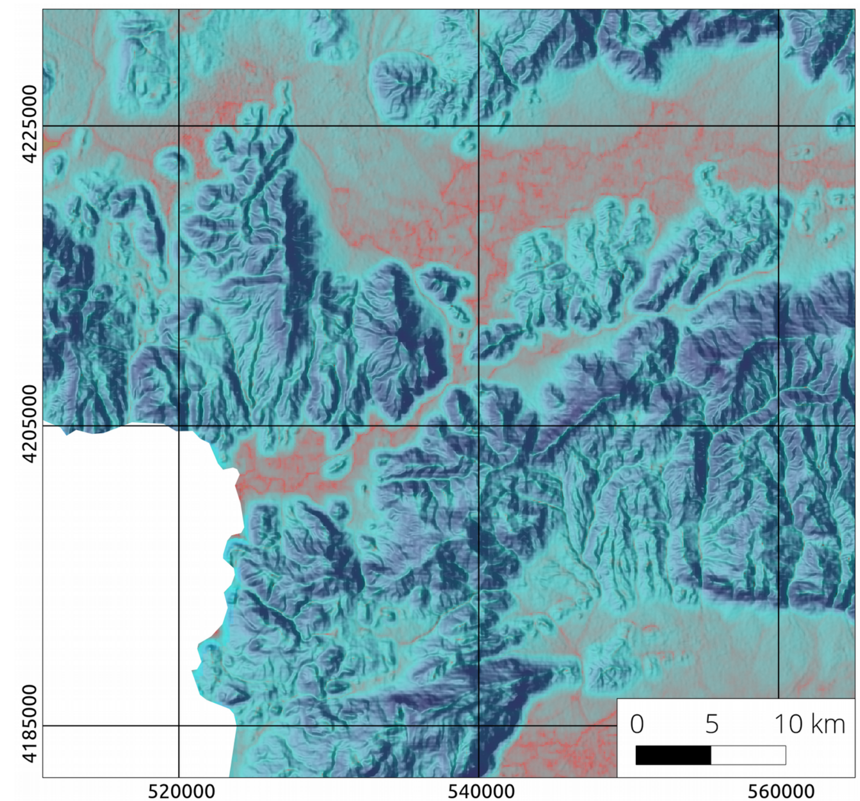


Terrain analyses

Slope = difference length/difference height

$$MTI = ln\left(\frac{A^n}{tan\beta}\right)$$





Terrain analyses – what was perceivable?

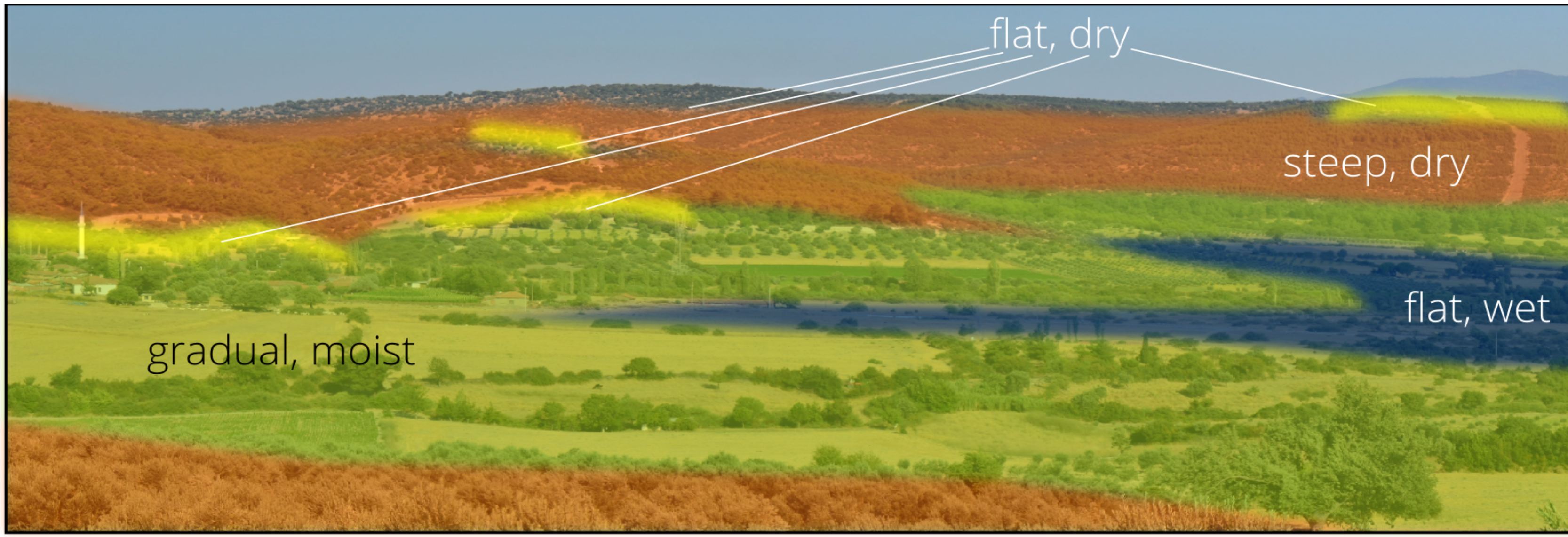
Terrain analyses – what was perceivable?

Terrain analyses – what was perceivable?

Terrain analyses – what was perceivable?

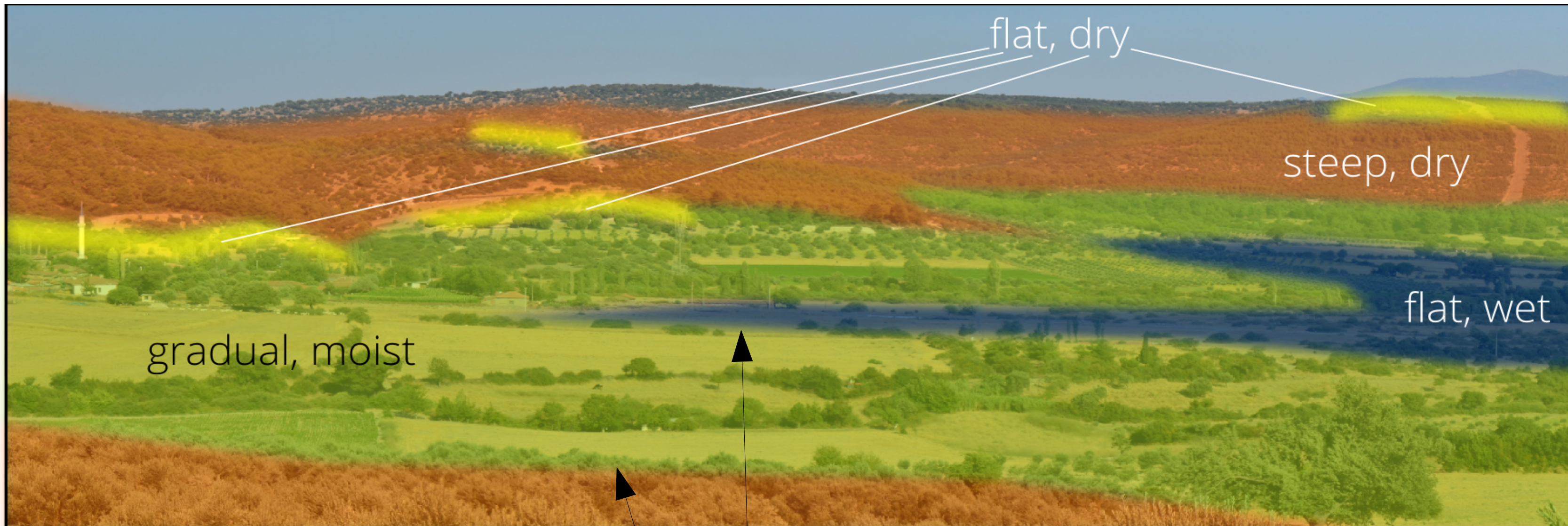
Terrain analyses – what was perceivable?

Terrain analyses – what was perceivable?



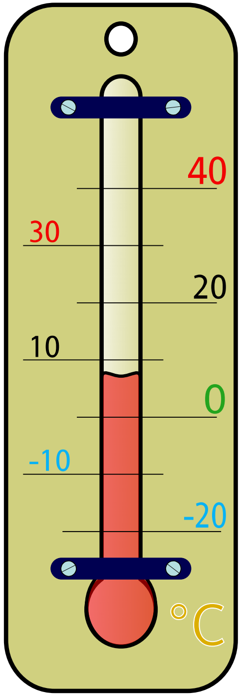
Popa & Knitter 2015

Terrain analyses – what was perceivable?



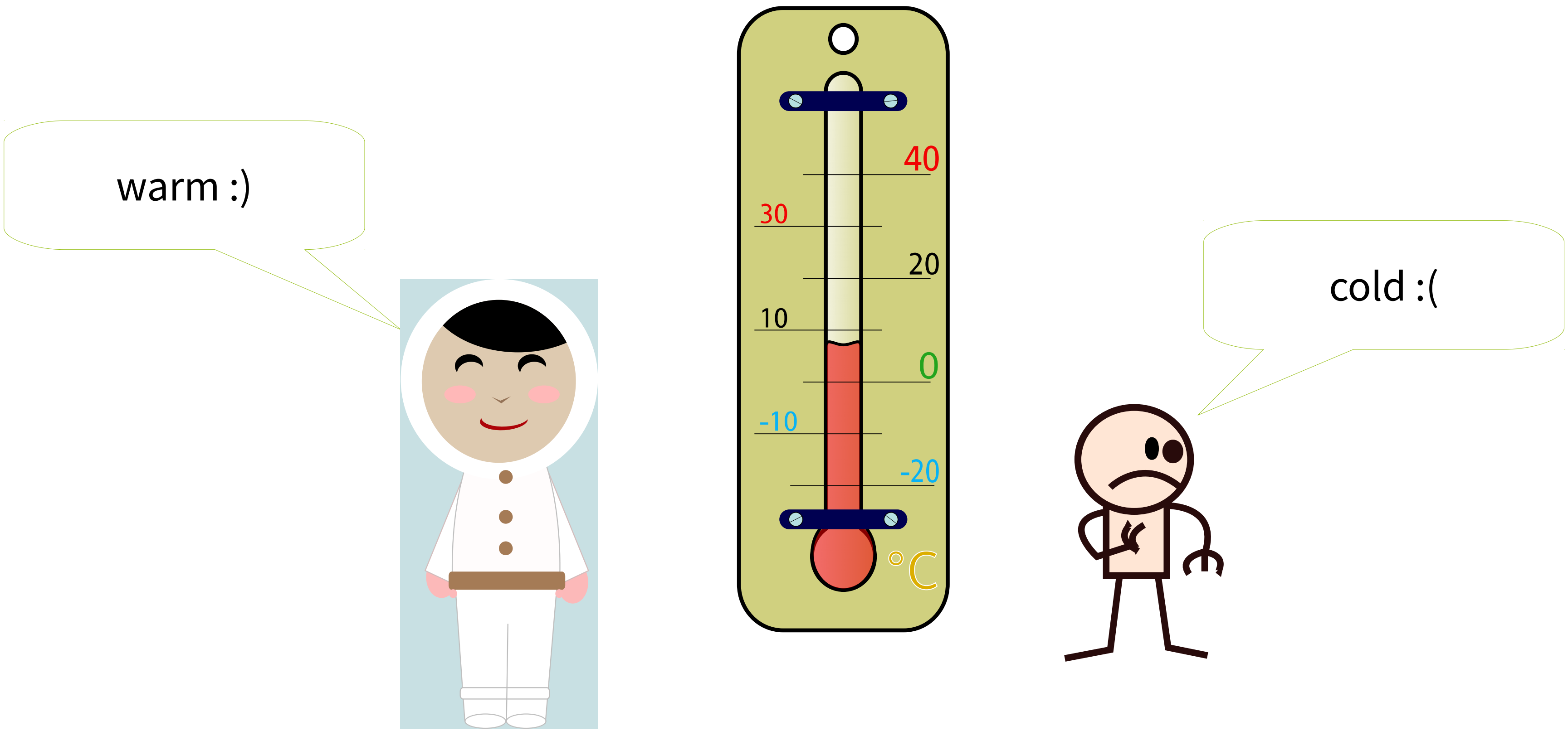
Where is the transition?

Crisp vs. fuzzy boundaries



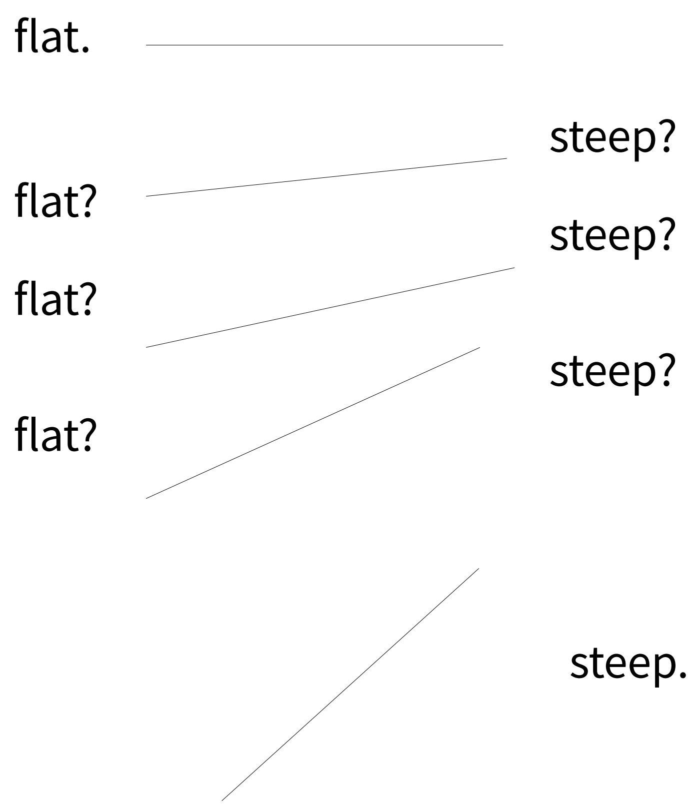
openclipart.org

Crisp vs. fuzzy boundaries



openclipart.org

Crisp vs. fuzzy boundaries



Crisp vs. fuzzy boundaries

Flatness = 100% — Steepness = 0%

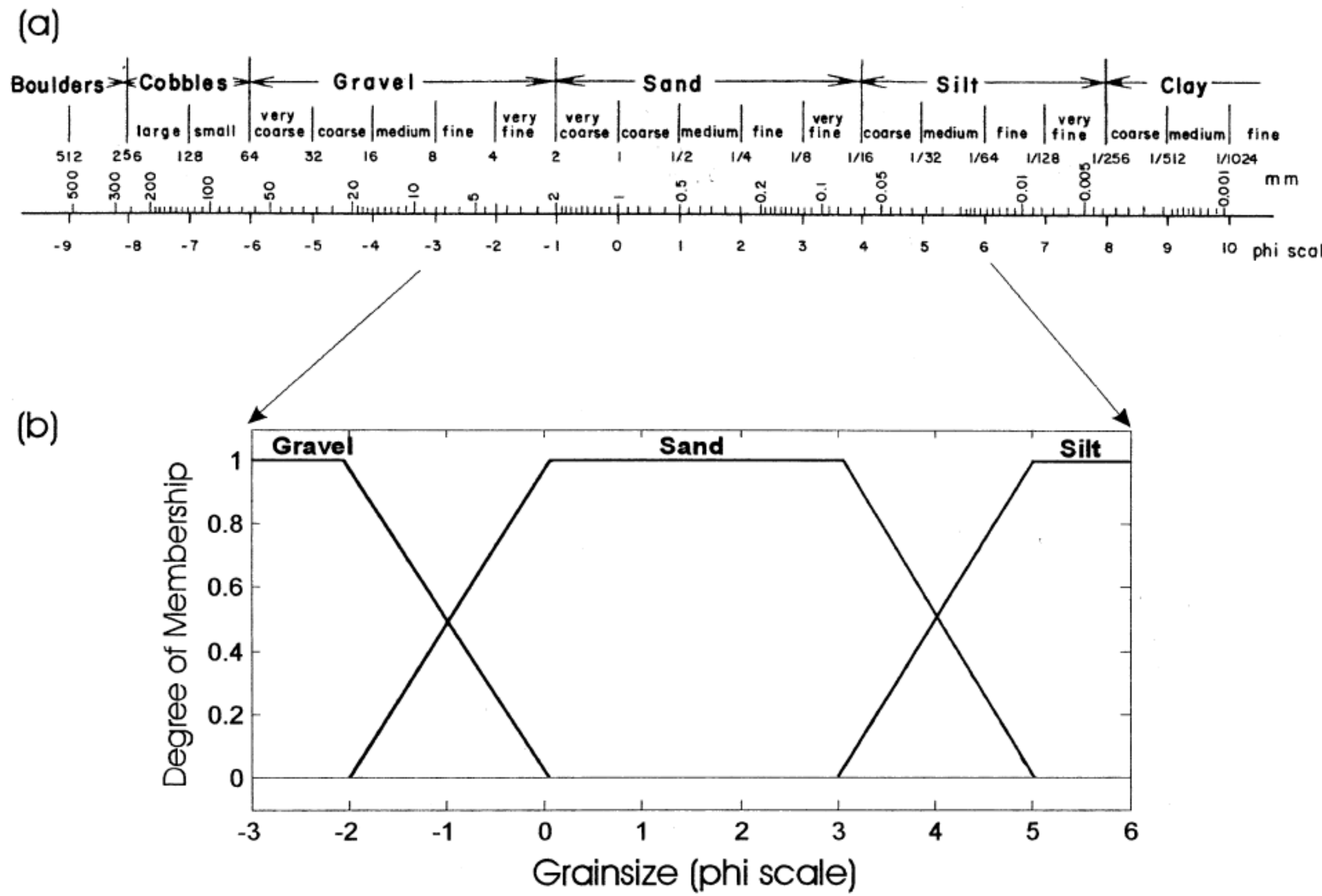
Flatness = 80% — Steepness = 10%

Flatness = 20% — Steepness = 40%

Flatness = 10% — Steepness = 80%

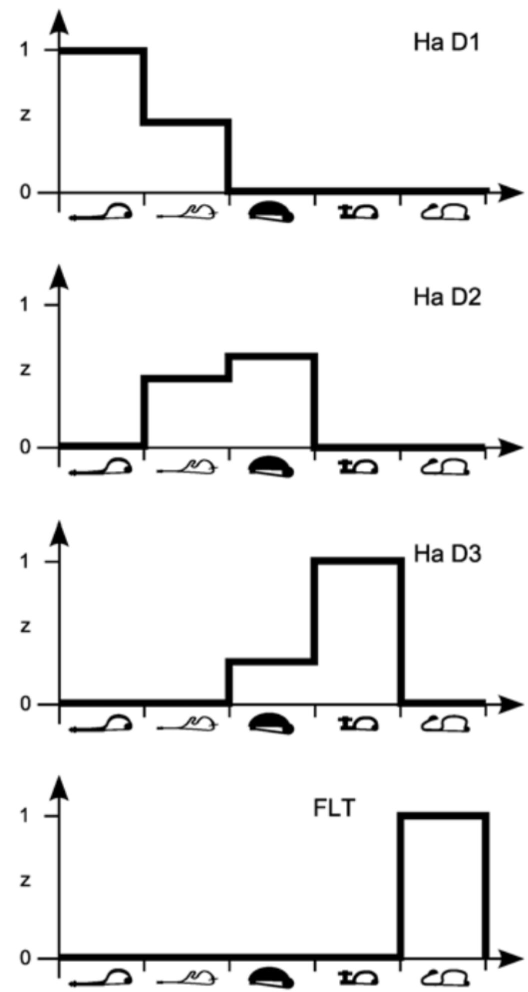
Flatness = 0% — Steepness = 100%

Crisp vs. fuzzy boundaries



Demicco 2004

Crisp vs. fuzzy boundaries

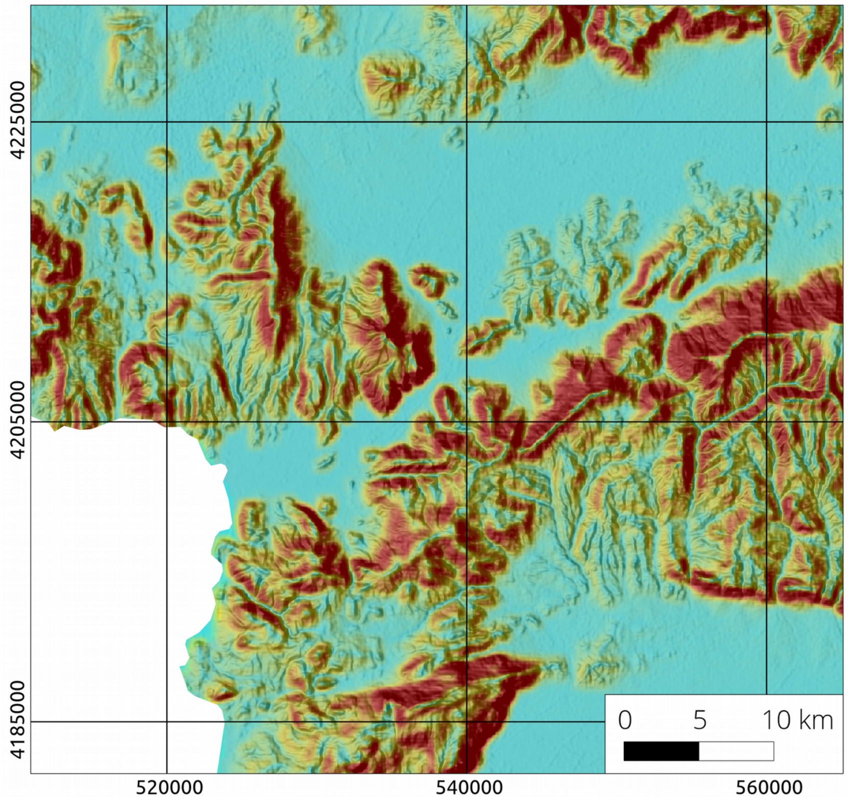


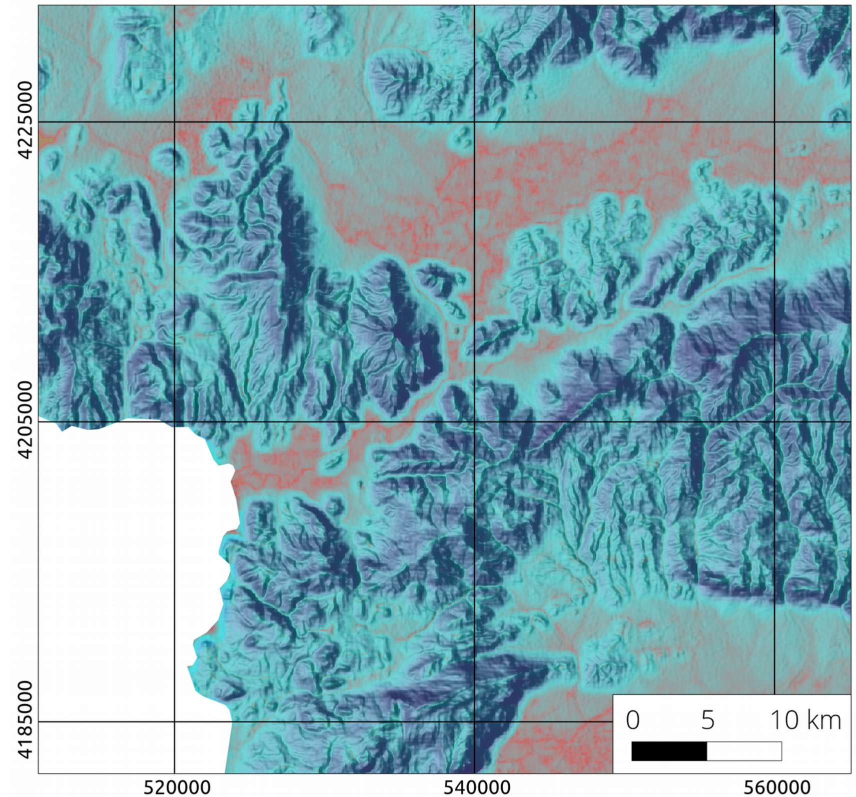
Nakoinz 2012

Terrain analyses – what was perceivable?

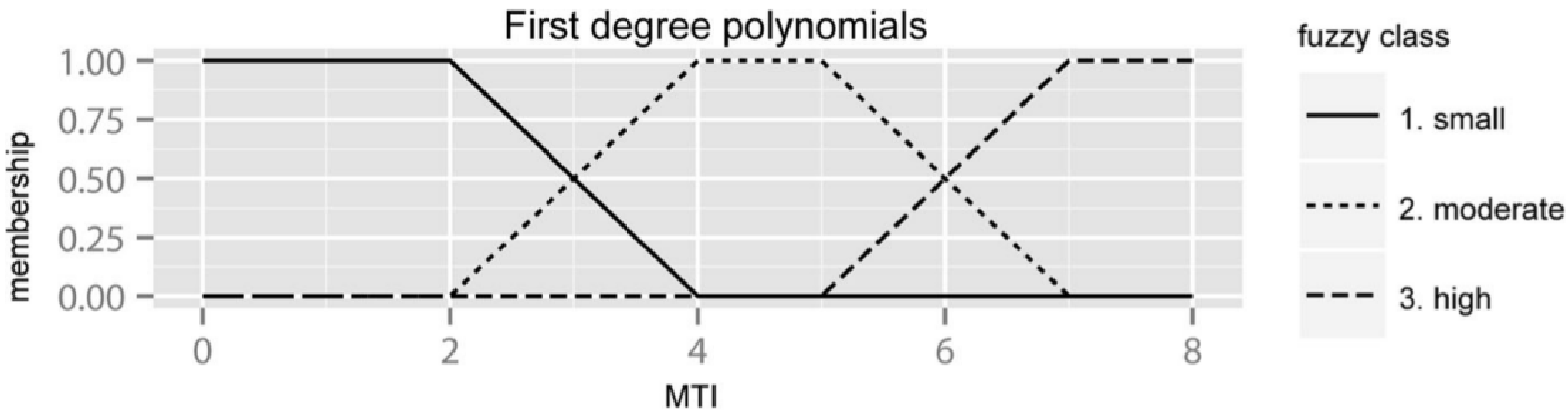
Slope = difference length/difference height

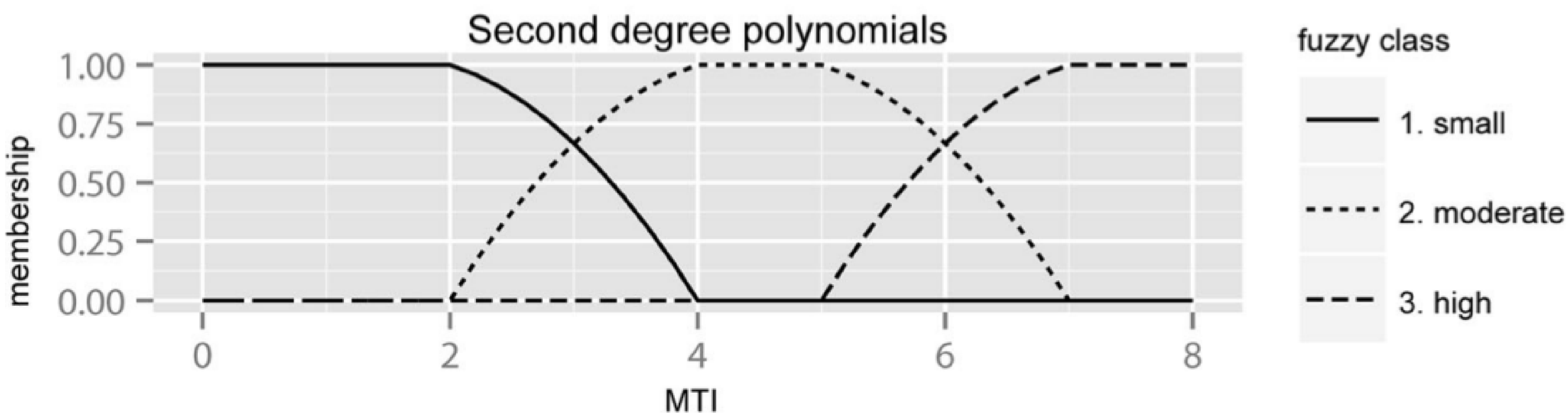
$$MTI = ln\left(\frac{A^n}{tan\beta}\right)$$





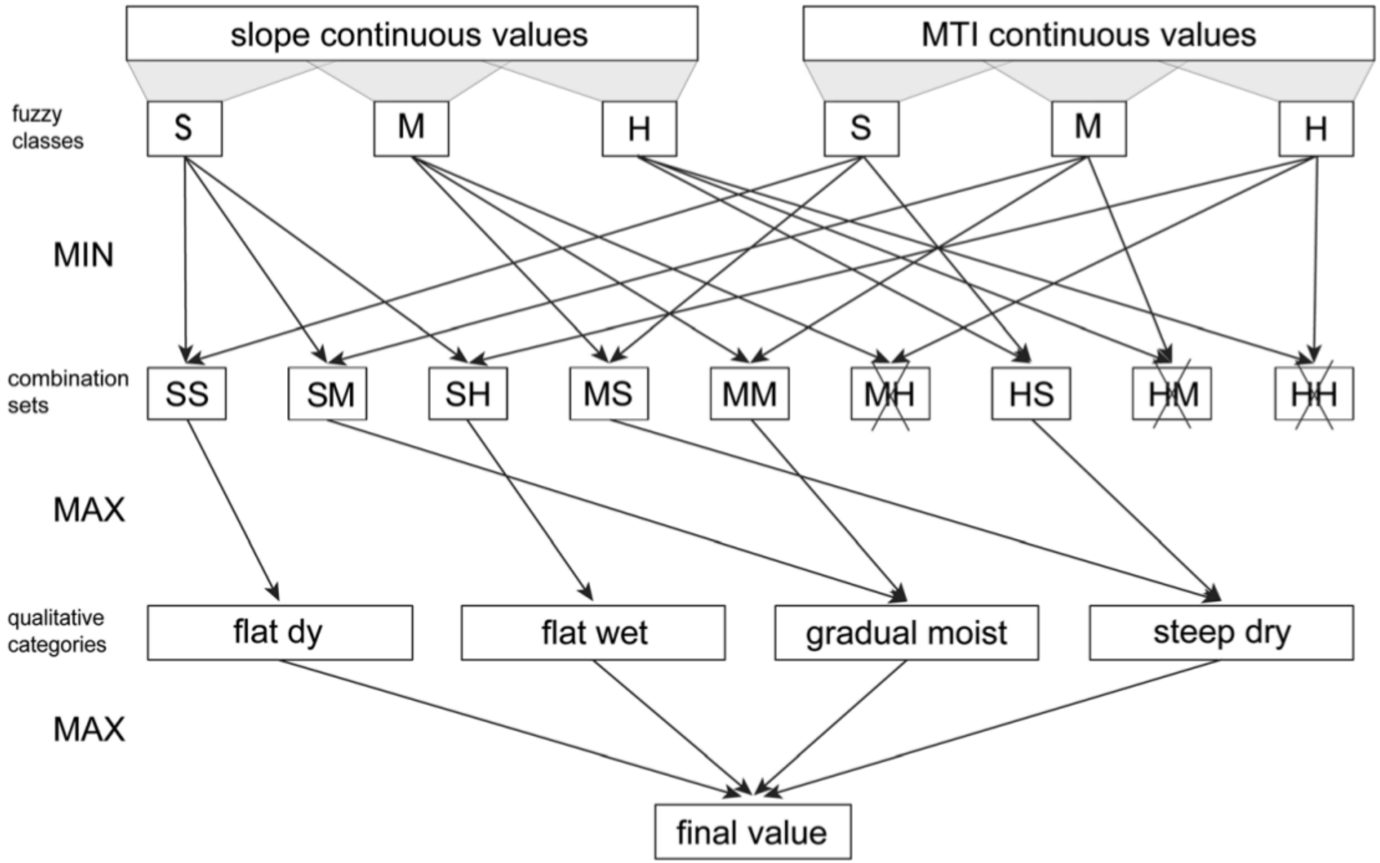
Terrain analyses – what was perceivable?





Popa & Knitter 2015

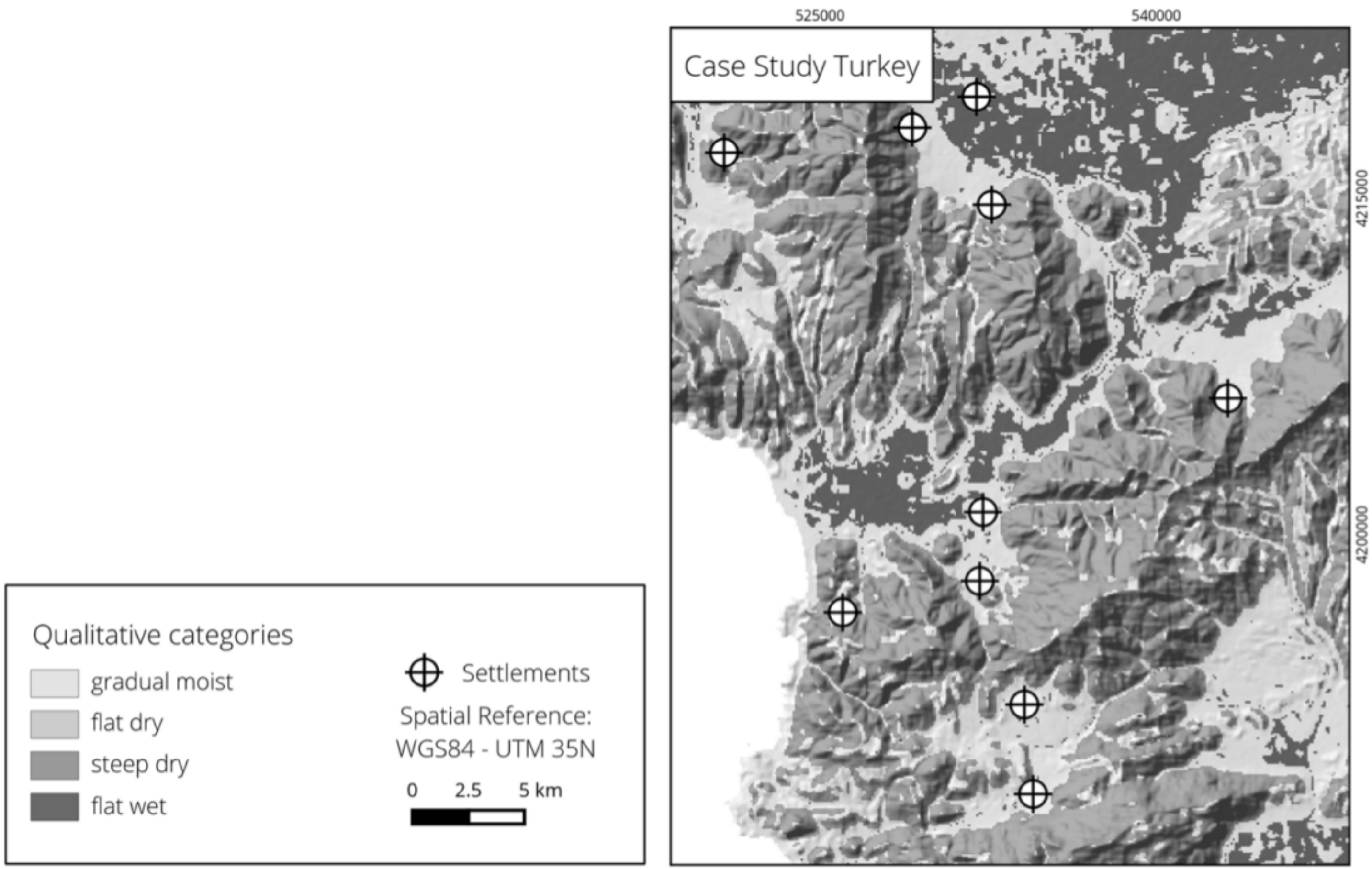
Terrain analyses – what was perceivable?



EXCELLENCE CLUSTER TOPOI

Popa & Knitter 2015

Terrain analyses – what was perceivable?

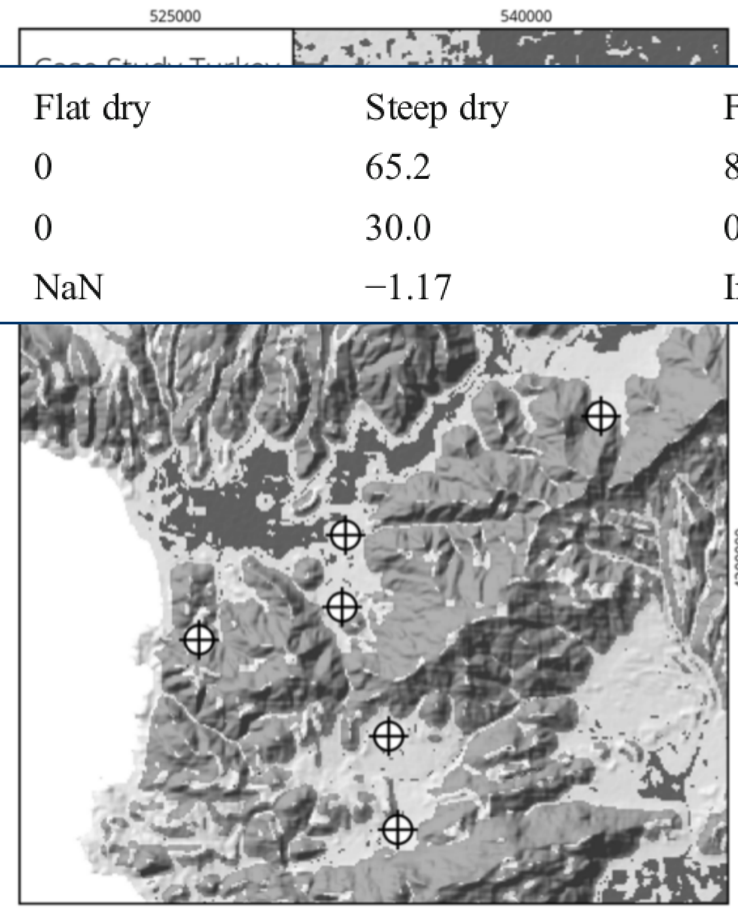


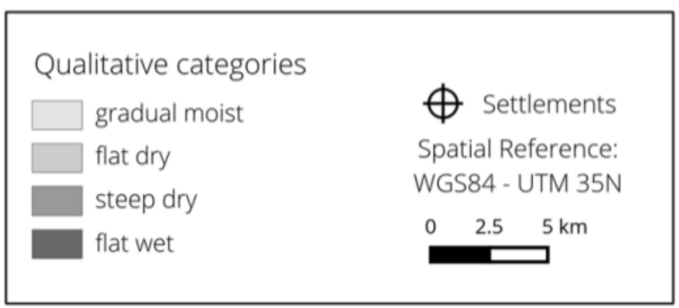
EXCELLENCE CLUSTER TOPOI

Popa & Knitter 2015

Terrain analyses – what was perceivable?

Turkey	Gradual moist	Flat dry	Steep dry	Flat wet
Area (%)	25.8	0	65.2	8.99
Sites (%)	70.0	0	30.0	0
Gain	0.63	NaN	−1.17	Inf





Terrain analyses – what was perceivable?

If “they” preferred certain areas, how would a resulting interaction network look like?

Terrain analyses – what was perceivable?

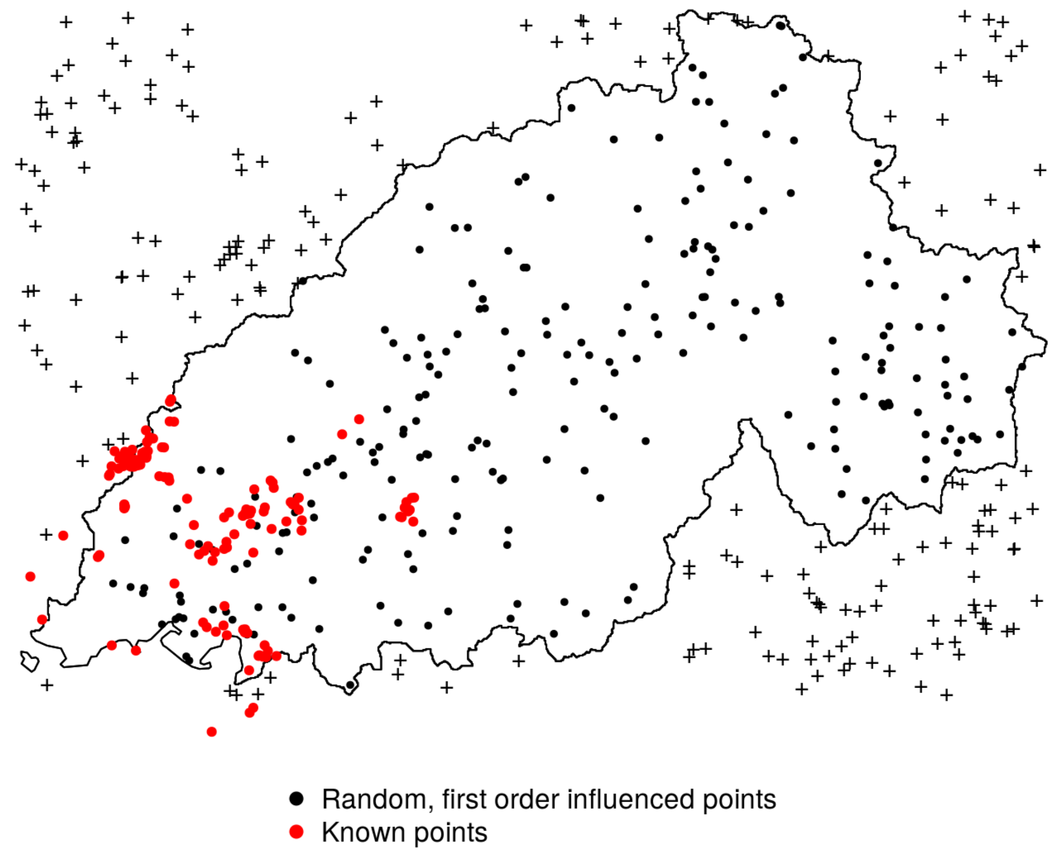
If “they” preferred certain areas,
how would a resulting interaction
network look like?

Workflow:
- fuzzy terrain

Terrain analyses – what was perceivable?

If "they" preferred certain areas, how would a resulting interaction network look like?

Workflow:
- fuzzy terrain
- point pattern analyses



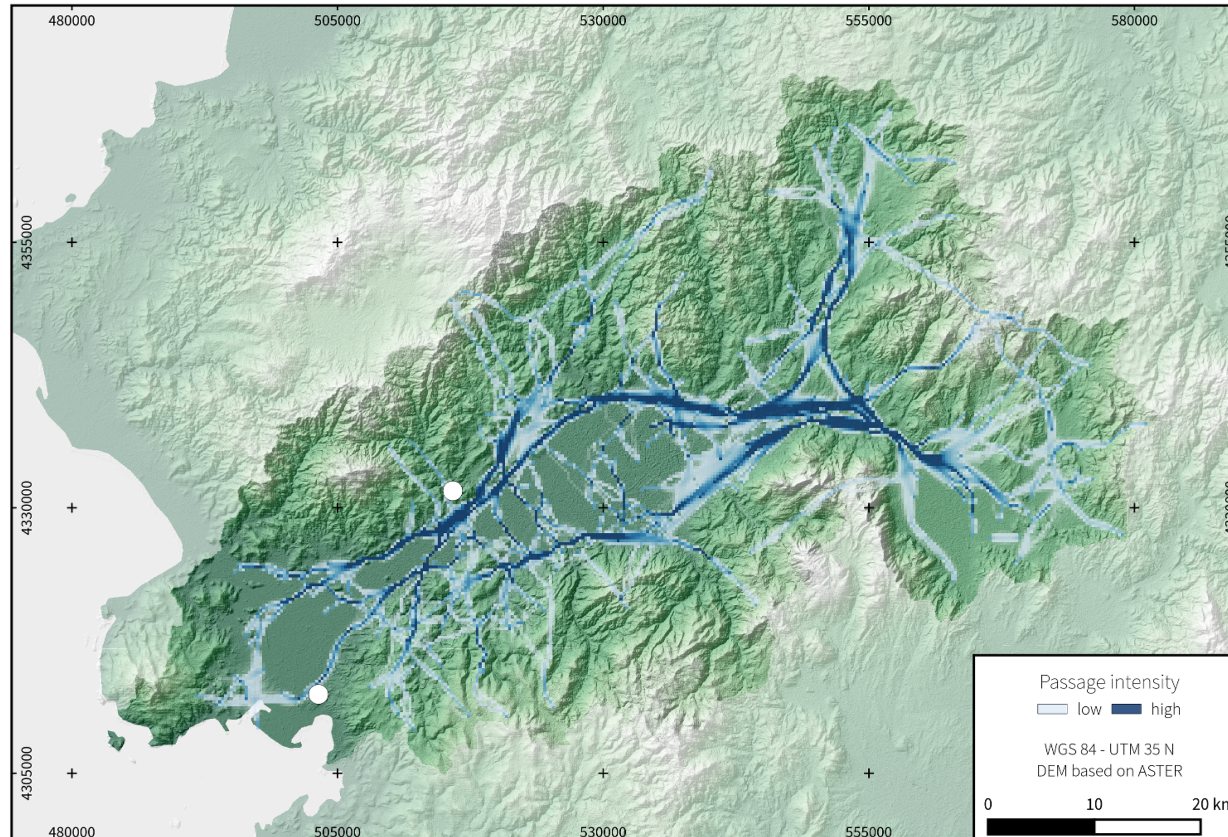
EXCELLENCE CLUSTER TOPOI

Terrain analyses – what was perceivable?

If “they” preferred certain areas, how would a resulting interaction network look like?

Workflow:
- fuzzy terrain
- point pattern analyses
- cost-surface creation
- pseudo-random walk

  (the smaller the cost, the higher the probability)
- sum of all random walks



EXCELLENCE CLUSTER TOPOI

Content

- Modeling and processes
- Why space matters: Potentials and Pitfalls of Spatial Data
- Point Pattern Analyses: What processes might have caused the distribution of findings?
  - First order processes (density based approaches)
  - Second order processes (distance based approaches)
- From single attributes to continuous surfaces: geostatistics
- Geomorphometry: What is the character of the landscape?
- **Outlook**

Outlook: personal advices

use open-source software, i.e. free software published under terms of GNU GENERAL PUBLIC LICENSE

Why?

- They grow. Fast. So you can work at the bleeding edge
- YOU can contribute
- others can contribute so YOU can profit from their methods and tools
- you can investigate how methods are implemented.

Publish open access. www.sci-hub.io is/was no alternative and science needs to be free.

Share your code. It is easy: github.com or gitlab.com

Thank you very much for your attention.

Entity-Attribute Spatial Data Types

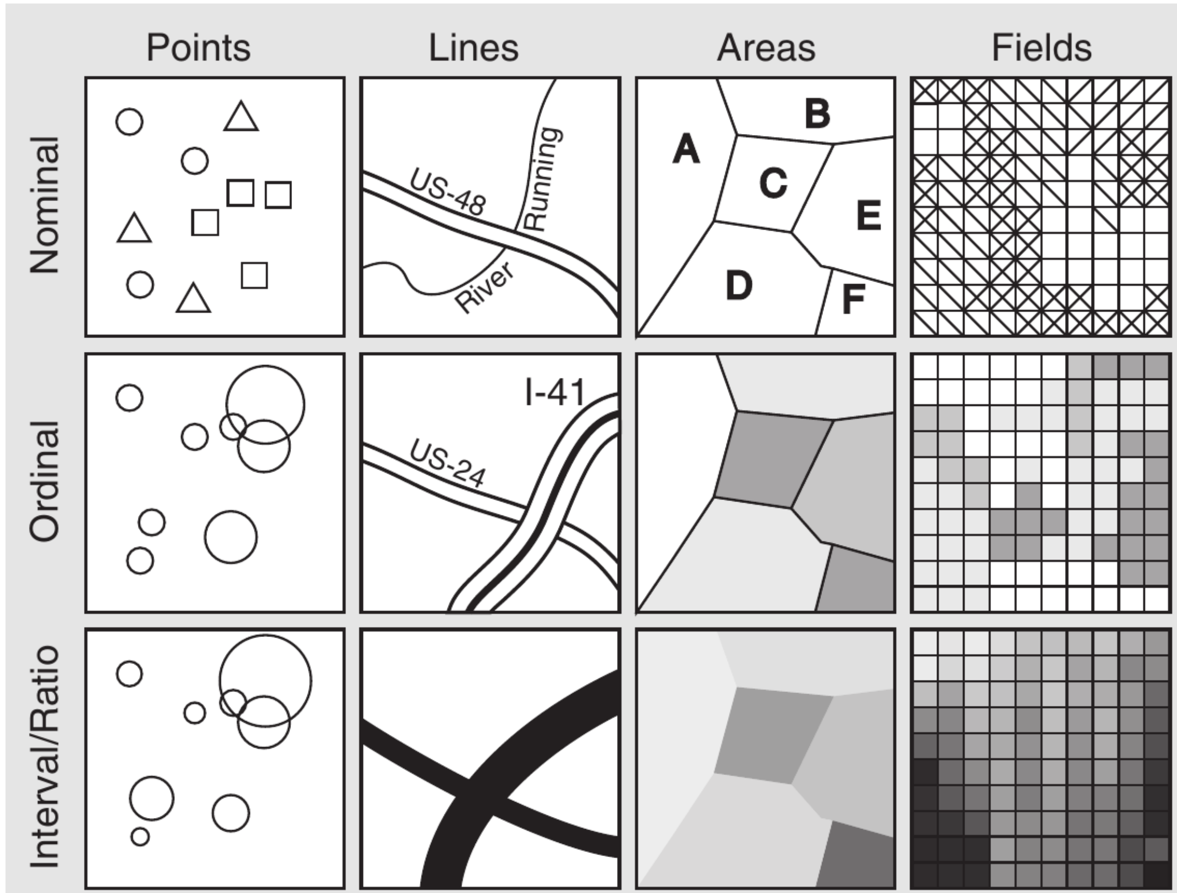
Processes caused the spatial occurrence of different entities

Ceramic types

Site types

Hierarchy

Site size



Territories

Population density

(O'Sullivan & Unwin, 2010)

Entity-Attribute Spatial Data Types

Processes define the transformation of spatial data

FROM	TO: Point, L^0	Line, L^1	Area, L^2	Field, L^3
Point, L^0	Mean center	Network graphs	territories Proximity polygons TIN, point buffer	Interpolation. Kernel density estimation Distance surfaces Population density
Line, L^1	Intersection junction	Shortest distance path	Line buffer	Distance to nearest line object surface
Area, L^2	Centroid City center	Graph of area skeleton	Area buffer, Polygon overlay	Pycnophylatic interpolation and other surface models
Field, L^3	Surface specific points VIPs	Surface network	Watershed delineation, Hill masses territories	Equivalent vector field

EXCELLENCE CLUSTER TOPOI

(O'Sullivan & Unwin, 2010)

Spatial Autocorrelation – Moran's I

Weight from spatial weight matrix

Covariance term

$$I = \left[\frac{n}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \right] \times \left[\frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}} \right]$$

Division by total data variance

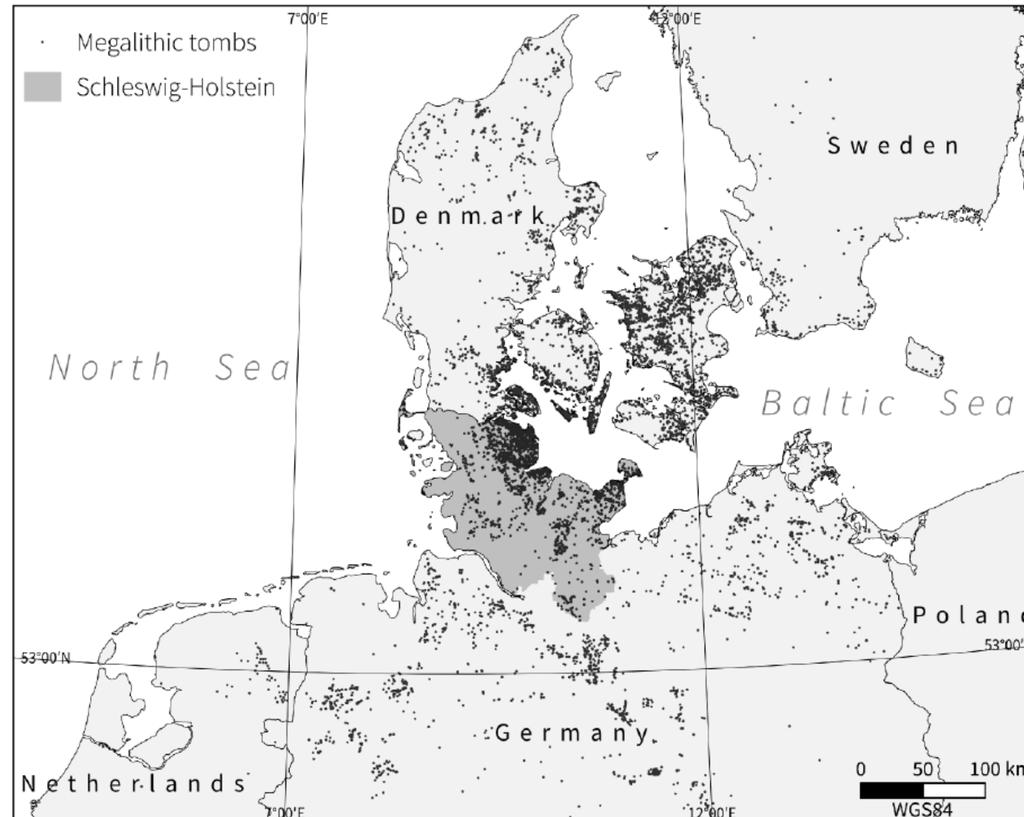
Normalize to total spatial weights

(O'Sullivan & Unwin, 2010)

Point Pattern – second order effects

Shows the point pattern clustering/dispersion; does it deviate from CSR?

→ Simulation approach based on Monte Carlo simulations



Knitter & Nakoinz (in press)

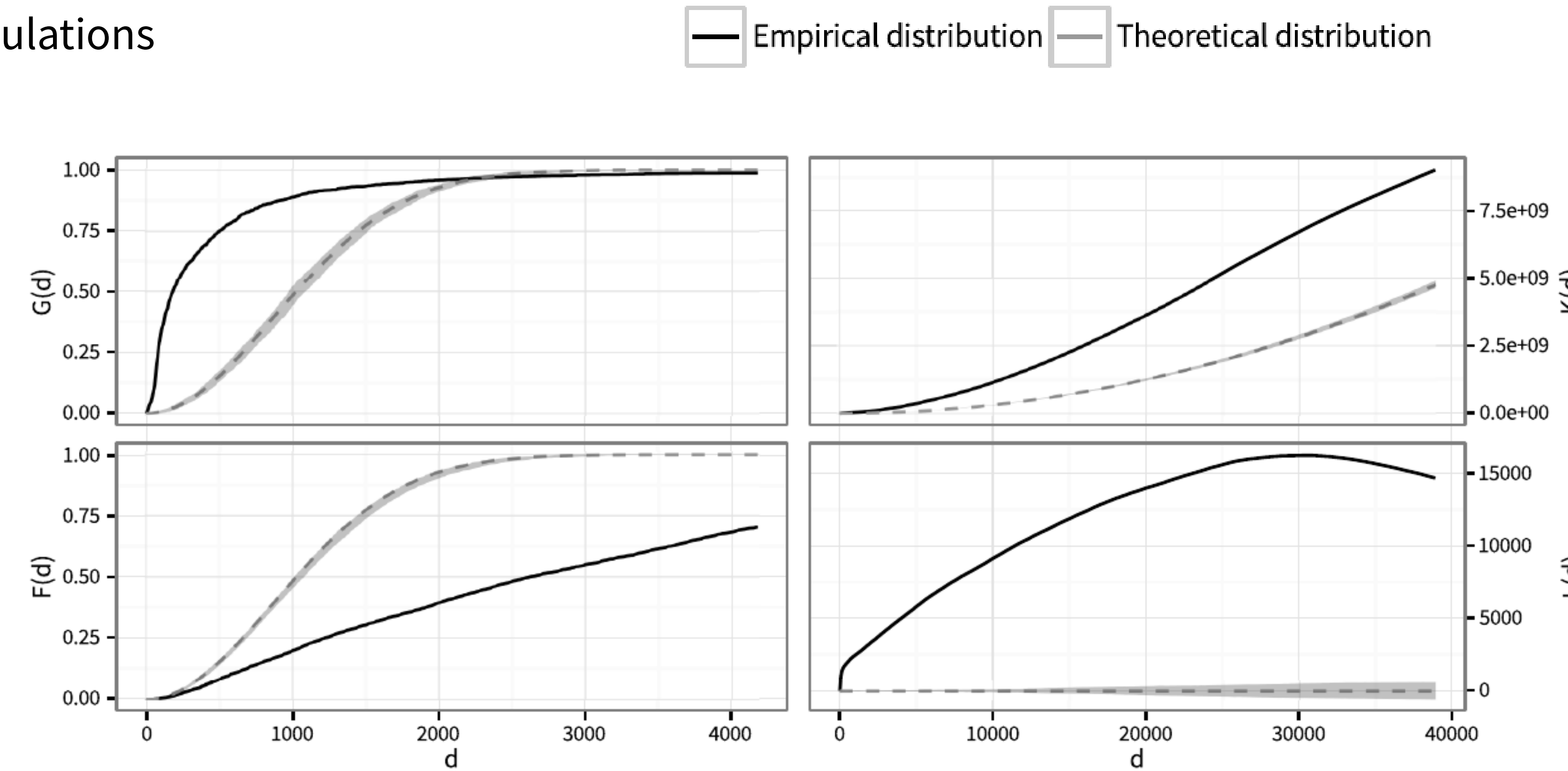
Point Pattern – second order effects

Shows the point pattern clustering/dispersion; does it deviate from CSR?

→ Simulation approach based on
Monte Carlo simulations

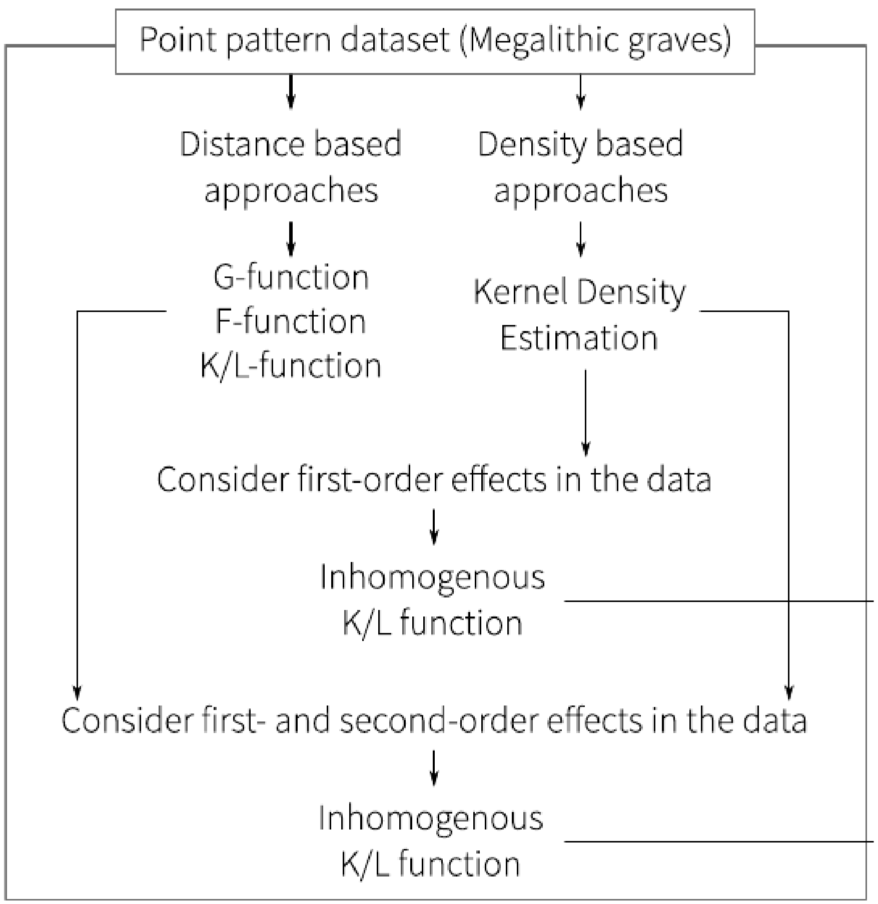
$$L(d) = \sqrt{\frac{K(d)}{\pi}}$$

Just a square root transformation of K(d)



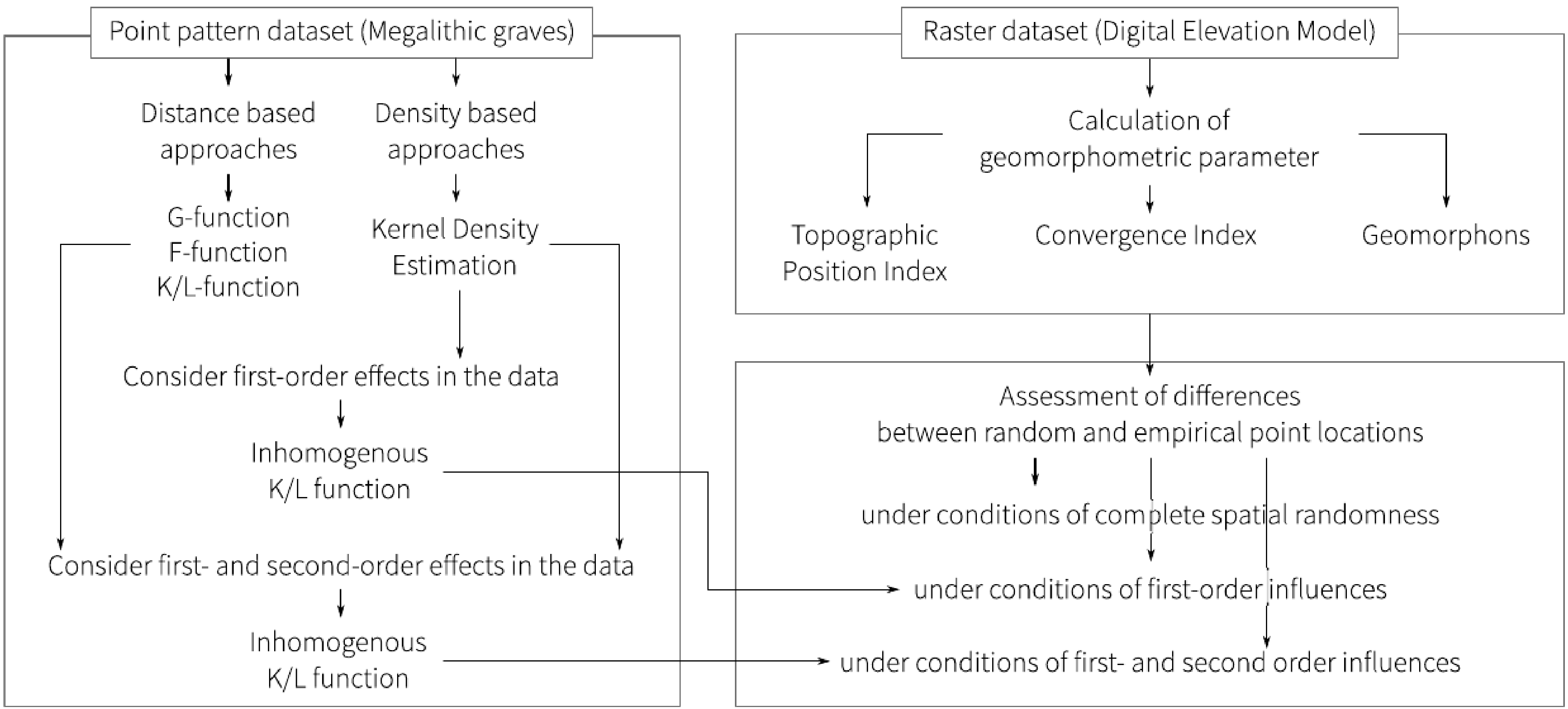
Knitter & Nakoinz (in press)

Point Pattern – a look ahead



Knitter & Nakoinz (in press)

Point Pattern – a look ahead



Knitter & Nakoinz (in press)

Interpolation

Kriging

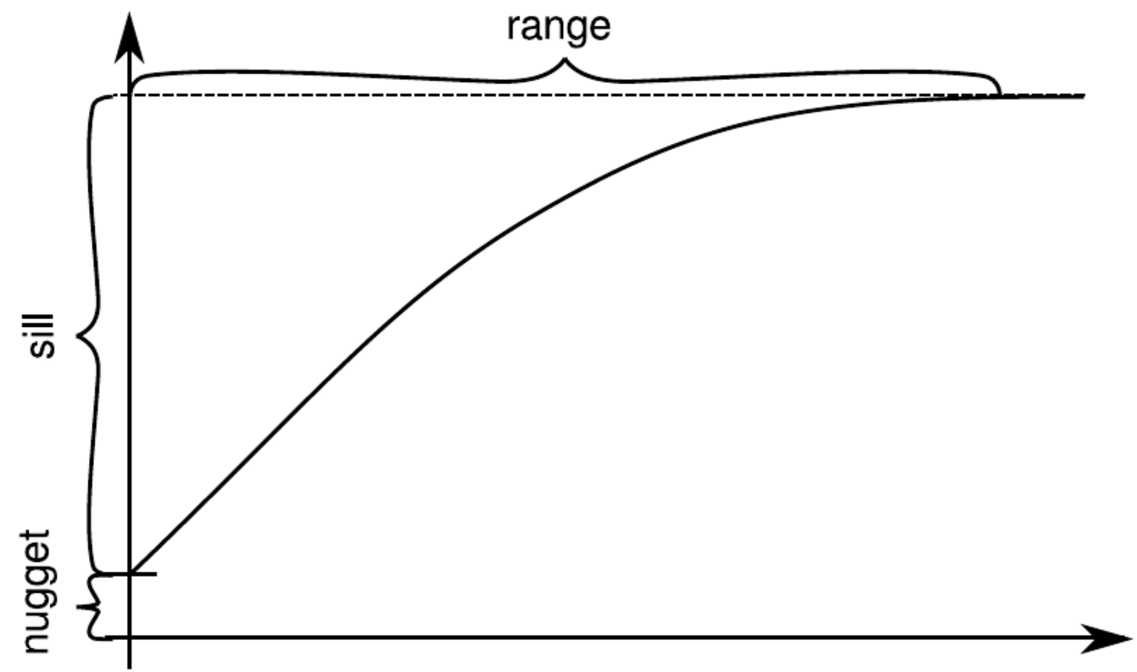
2. Summarizing spatial variation

"Elements" of a semivariogram

Nugget
→ measurement error
→ variation below shortest sampling interval

Sill
→ range of semivariance

Range
→ distance at which semivariance reaches maximal value



(Nakoinz & Knitter 2016)