Spatial Statistics - a one hour introduction

IGSSE Forum 2014
June 17, 2014

Benedikt Gräler
Institute for Geoinformatics
University of Münster
http://ifgi.uni-muenster.de/graeler

Spatial Data

From a purely statistical perspective, spatial data is multivariate data with special covariates: the coordinates.

Tobler's first law of Geography states [10]:

> *Everything is related to everything else, but near things are more related than distant things.*

Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space.

Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:

Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space.

Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:

  Degrees: 6.88

Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space.

Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:

  Degrees: 6.88
  Great Circle: 385 km

Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space.

Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:

  Degrees: 6.88
  Great Circle: 385 km
  Rate: 56 km/degree

Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space.

Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:

  Degrees: 6.88
  Great Circle: 385 km
  Rate: 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):

Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space.

Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:

  Degrees: 6.88
  Great Circle: 385 km
  Rate: 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):

  Degrees: 7.32

Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space.

Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:

  Degrees: 6.88
  Great Circle: 385 km
  Rate: 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):

  Degrees: 7.32
  Great Circle: 814 km

Coordinate Reference System

We model the earth, but think in maps: locations are projected from a curved surface in 3D to flat 2D space.

Be aware of geographic coordinates and different projections that maintain angles, certain distances or area.

Imagine the following distances between:

- the Fjord of Oslo (59.85 N 10.75 E) and Uppsala (59.85 N 17.63 E) that are at the same latitude:

  Degrees: 6.88
  Great Circle: 385 km
  Rate: 56 km/degree

- the intersections of the Congo river with the equator (0.00 N 18.21 E) and (0.00 N, 25.53 E):

  Degrees: 7.32
  Great Circle: 814 km
  Rate: 111 km/degree

CRS identifier

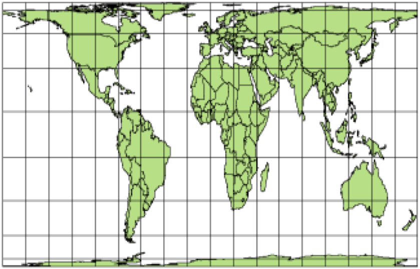
To distinguish different projections, a well prepared data set comes with its coordinate reference system (CRS) as metadata.

These are often encoded as

- EPSG-codes (by the European Petroleum Survey Group)
- proj4string

They define how the reference surface (sphere, ellipsoid) is fixed to the real world (called the datum) and how the projection (surface in 3D to 2D plane) is made.

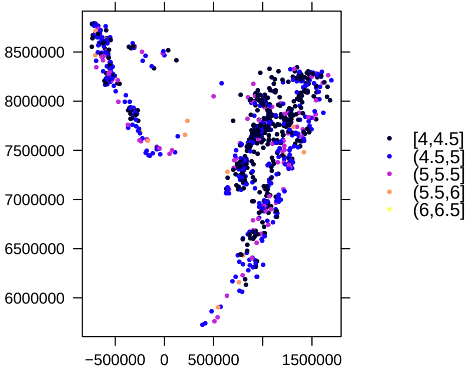
Projection



Point Patterns

Records of locations of events are referred to as *Point Patterns* (e.g. locations of earthquakes).

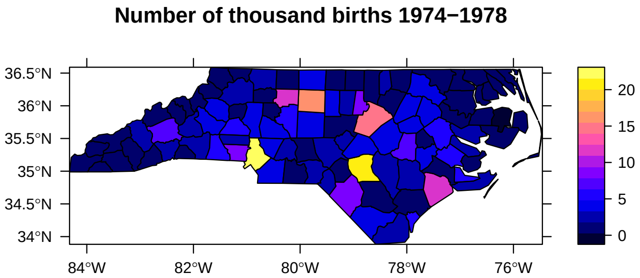
Adding an attribute (e.g. magnitude) to the raw locations generates a *marked point pattern*.



Lattice

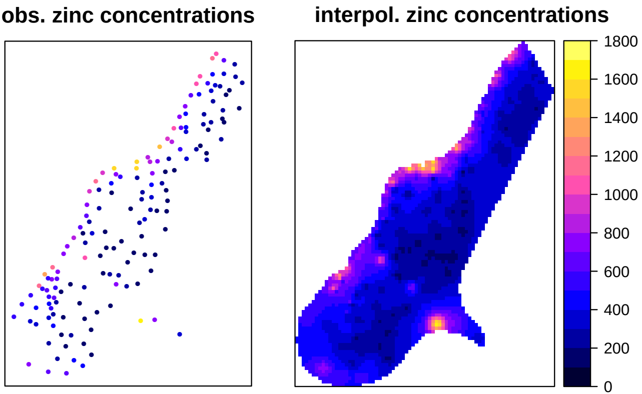
Data collected per region is referred to as *lattice* data (e.g. number of birth per county).

These values are true per region, but generally not observable at each point in that region.



Fields

Fields are understood as continuously spreading over space (e.g. temperature recordings) and typically observed at a set of distinct locations and illustrated as interpolated maps. Typically, fields are modelled as a realisation of a spatial random field.



Transitions between data types

- Points in a point pattern can be counted per pre-defined regions resulting in a lattice (earthquakes per region).

Transitions between data types

- Points in a point pattern can be counted per pre-defined regions resulting in a lattice (earthquakes per region).
- Smoothing techniques can generate a point density out of a point pattern resulting in a field (density of earthquakes).

Transitions between data types

- Points in a point pattern can be counted per pre-defined regions resulting in a lattice (earthquakes per region).
- Smoothing techniques can generate a point density out of a point pattern resulting in a field (density of earthquakes).
- Fields can be aggregated to regions resulting in a lattice (average temperature per federal state).

Motivation

- a better understanding of the observed phenomenon

Motivation

- a better understanding of the observed phenomenon
- prediction at unobserved locations

Motivation

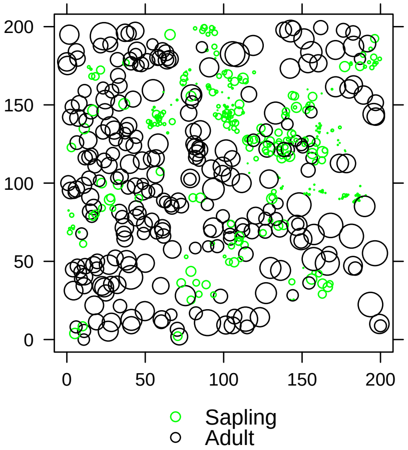
- a better understanding of the observed phenomenon
- prediction at unobserved locations
- prediction of the future

Motivation

- a better understanding of the observed phenomenon
- prediction at unobserved locations
- prediction of the future
- studying driving factors

Longleaf pine trees

The following examples follow those from the spatstat package [3]. More detailed explanations are found in [2].



Uniform intensity

How many trees are there per square unit on average?

$$\lambda = \frac{N}{area}$$

```
> summary(longleaf)
Marked planar point pattern:  584 points
Average intensity 0.0146 points per square metre

Coordinates are given to 1 decimal place
i.e. rounded to the nearest multiple of 0.1 metres

marks are  numeric,  of type ‘double’
Summary:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   2.00    9.10   26.15   26.84   42.12   75.90

Window: rectangle = [0, 200] x [0, 200] metres
Window area = 40000 square metres
Unit of length: 1 metre
```

Inhomogeneous intensity

How does the intensity change throughout the study area?

quadrat counting: the region is split into areas of equal size and a uniform density is estimated per area

ifgi
Institute for Geoinformatics
University of Münster

Inhomogeneous intensity

How does the intensity change throughout the study area?

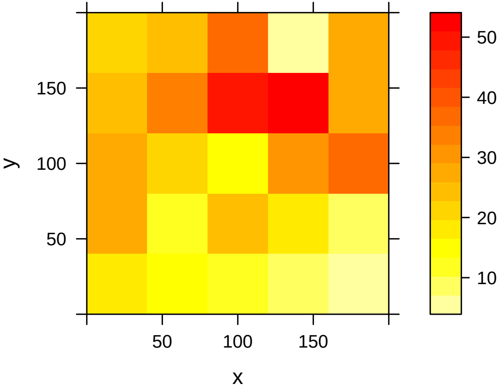
quadrat counting the region is split into areas of equal size and a uniform density is estimated per area

kernel smoothing the contribution of each point is spread across its neighbourhood based on some kernel density being properly rescaled

ifgi
Institute for Geoinformatics
University of Münster

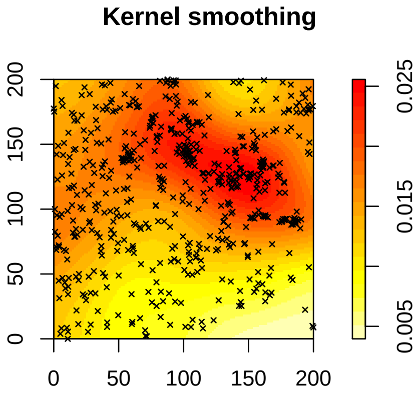
quadrat counting

```
> quadratcount(longleaf,nx=4,ny=4)
             x
y             [0,40] (40,80] (80,120] (120,160] (160,200]
  (160,200]       20      25       37         7        26
  (120,160]       25      34       50        51        27
  (80,120]        29      22       15        31        37
  (40,80]         26      12       24        19         8
  [0,40]          18      14       12         8         7
```



kernel smoothing

```
> density(longleaf)
```



Fitting Poisson processes I

ifgi
Institute for Geoinformatics
University of Münster

Following the algorithm by [4], we seek a model where the intensity λ is log-linear in the parameter θ:

$$\log\left(\lambda_\theta(x, y)\right) = \theta \cdot f(x, y)$$

A model fit with f being a simple linear model of the coordinates, is obtained through

```
> modelLl <- ppm(longleaf, ~x+y)
> modelLl
Nonstationary Poisson process

Trend formula: ~x + y
> AIC(modelLl)
6077
```

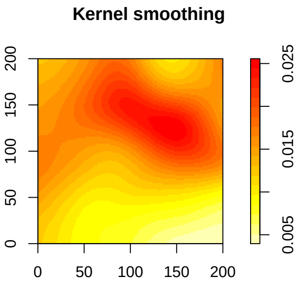
Fitting Poisson processes II

A model with f being a polynomial function can be fitted through:

```
> modelLl <- ppm(ppLl, ~polynom(x, y, 3))
> modelLl
Nonstationary Poisson process

Trend formula: ~polynom(x, y, 3)
> AIC(modelLl)
6027
```

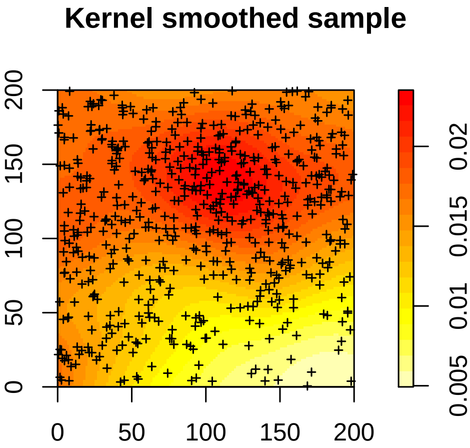
Fitting Poisson processes III



The fitted Poisson process

Now that the process is fitted, sampling can take place:

```
rpoispp(lambdaFun, lmax=0.03, win=owin(c(0,200),c(0,200)))
```



Additional options

- intensity based on covariates (e.g. elevation)

Additional options

- intensity based on covariates (e.g. elevation)
- intensity per category $\rightarrow$ multiple point process (e.g. sapling vs. adult)

Additional options

- intensity based on covariates (e.g. elevation)
- intensity per category $\rightarrow$ multiple point process (e.g. sapling vs. adult)
- interpoint interactions

ifgi
Institute for Geoinformatics
University of Münster

Additional options

- intensity based on covariates (e.g. elevation)
- intensity per category $\rightarrow$ multiple point process (e.g. sapling vs. adult)
- interpoint interactions
- modelling of marks: (L, M|L), (M, L|M), (L,M)

Additional options

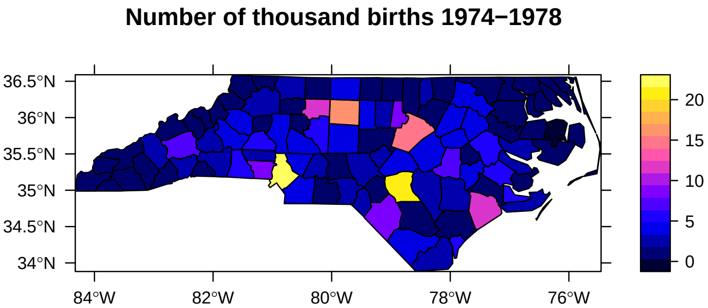
- intensity based on covariates (e.g. elevation)
- intensity per category $\rightarrow$ multiple point process (e.g. sapling vs. adult)
- interpoint interactions
- modelling of marks: (L, M|L), (M, L|M), (L,M)
- testing for independence

Additional options

- intensity based on covariates (e.g. elevation)
- intensity per category $\rightarrow$ multiple point process (e.g. sapling vs. adult)
- interpoint interactions
- modelling of marks: (L, M|L), (M, L|M), (L,M)
- testing for independence
- GOF testing

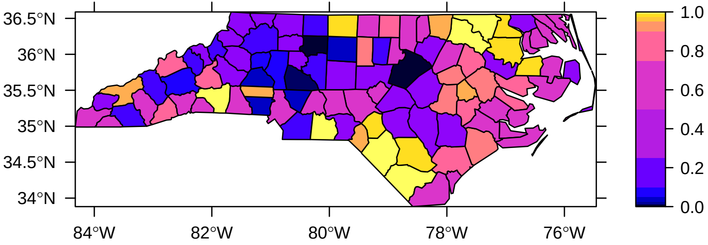
Sudden infant death syndrome - North Carolina

The following examples follow these of the spdep package and its vignettes [5].



Expected counts - probability map

Assuming Poisson distributions with mean values set to the expected number of cases ec per county based on the number of births $ec(county) = N_{Birth} \cdot rate$, the cumulated density of the observed values is derived.



CAR models I

The motivation of a *conditional autoregressive model* is given by the conditional distribution

$$
\begin{aligned}
& f\left(z(s_i)\middle|\{z(s_j) : j \neq i\}\right) \\
= \quad & \frac{1}{\tau_i\sqrt{2\pi}} \exp\left(\frac{-\left(z(s_i) - \theta_i(\{z(s_j) : j \neq i\})\right)^2}{2\tau_i^2}\right)
\end{aligned}
$$

with

$$
\theta_i(\{z(s_j) : j \neq i\}) = \mu_i + \sum_{j=1}^{n} c_{ij}\left(z(s_j) - \mu_j\right)
$$

while $(c_{ij})_{ij}$ with $c_{ij} = 0$ unless the locations s_i and s_j are pairwise dependent, $c_{ij}\tau_i^2 = c_{ji}\tau_j^2$ and $c_{ii} = 0$. τ_i^2.

CAR models II

The process Z can then be modelled as

$$Z \sim \mathrm{Gau}\big(\boldsymbol{\mu}, (\mathrm{I} - \mathrm{C})^{-1}\mathrm{M}\big)$$

$$Z(s_i) = \mu_i + \sum_{j=1}^{n} c_{ij}\big(Z(s_j) - \mu_j\big) + \nu_i$$

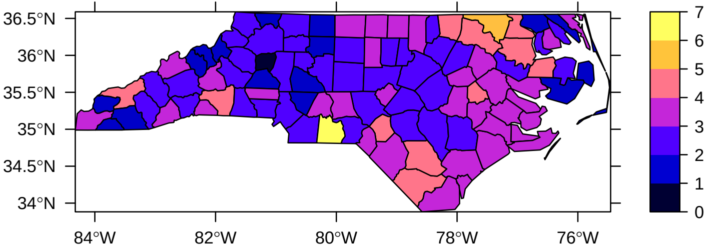
with $\boldsymbol{\nu} \sim \mathrm{Gau}(\mathbf{0}, \mathrm{M}\big(\mathrm{I} - \mathrm{C}^{\mathrm{t}})\big)$ while $M = \mathrm{diag}(\tau_1^2, \ldots, \tau_\mathrm{n}^2)$, $C = (c_{ij})_{ij}$.
Typically, only a small number of conditioning sites is used assuming the Markov property

$$f\big(z(s_i)|\{z(s_j) : j \neq i\}\big) = f\big(z(s_i)|\{z(s_j) : j \in N_i\}\big)$$

with N_i being the index set of selected neighbouring locations.

CAR Example: SIDS in North Carolina I

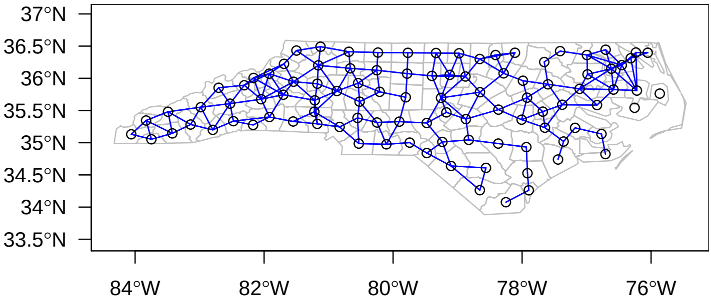
Freeman-Tukey transformed data



CAR Example: SIDS in North Carolina II

$$Z \sim \mathrm{Gau}(\boldsymbol{\mu}, (\mathrm{I} - \mathrm{C})^{-1}\mathrm{M})$$

We assume a constant mean $\boldsymbol{\mu} = m$. The neighbourhood sets of the Markov Random Field are based on the distance between centroids of the counties ($\leq$ 30 miles).



CAR Example: SIDS in North Carolina III

The spatial dependence matrix C is modelled as

$$c_{ij} := \begin{cases} \phi \cdot \frac{\min_{j \in N_i, i=1,\ldots,n}(d_{ij})}{d_{ij}} \cdot \sqrt{\frac{n_j}{n_i}}, & j \in N_i \\ 0, & j \notin N_i \end{cases}$$

and the conditional variance by

$$\tau_i^2 := \frac{\tau^2}{n_i}$$

such that $c_{ij}\tau_j^2 = c_{ji}\tau^2 i$ and $M = \mathrm{diag}(\tau_1^2, \ldots, \tau_\mathrm{n}^2)$.
The set of parameters is (m, ϕ, τ).

CAR Example: SIDS in North Carolina IV

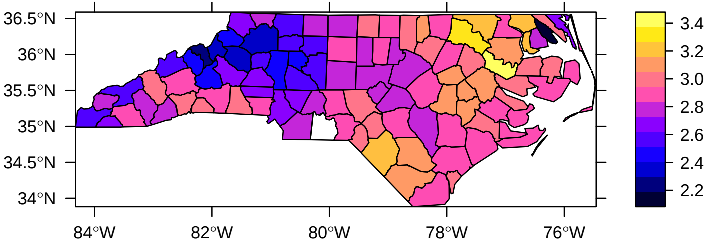
The call

```
car <- spautolm(ft.SID74 ~ 1, data=mdata.4,
                listw=sids.nhbr.listw.4,
                weights=BIR74, family="CAR")
summary(car)
```

yields estimates: $\hat{m} = 2.84$, $\hat{\phi} = 1.73$ and $\hat{\tau} = 36.16$

CAR Example: SIDS in North Carolina V

Predicted transformed data



SAR models

A *simultaneous autoregressive model* is given by

$$Z \sim \mathrm{Gau}\big(\boldsymbol{\mu}, (\mathrm{I} - \mathrm{B})^{-1} \Lambda (\mathrm{I} - \mathrm{B}^{\mathrm{t}})^{-1}\big)$$

$$Z(s_i) = \mu_i + \sum_{j=1}^{n} b_{ij} \big(Z(s_j) - \mu_j\big) + \epsilon_i$$

with $\boldsymbol{\epsilon} \sim \mathrm{Gau}(\mathbf{0}, \Lambda)$, $B = (b_{ij})_{ij}$ while $b_{ii} = 0$ and b_{ij} captures the dependence of location s_i on s_j. It is not necessarily $b_{ij} = b_{ji}$, but $(I - B)^{-1}$ is assumed to exist.

SAR Example: SIDS in North Carolina I

$$Z \sim \mathrm{Gau}\big(\boldsymbol{\mu}, (\mathrm{I}-\mathrm{B})^{-1}\Lambda(\mathrm{I}-\mathrm{B}^{\mathrm{t}})^{-1}\big)$$

We assume again a constant mean $\boldsymbol{\mu} = m$.
The spatial dependence matrix B is modelled as

$$b_{ij} := \phi \frac{1}{d_{ij} \cdot \sum_{j \in N_i} \frac{1}{d_{ij}}}$$

and the variance by

$$\sigma_i^2 := \frac{\sigma^2}{n_i}$$

such that $\Lambda = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_{\mathrm{n}}^2)$.
The set of parameters is (m, ϕ, σ).

SAR Example: SIDS in North Carolina II

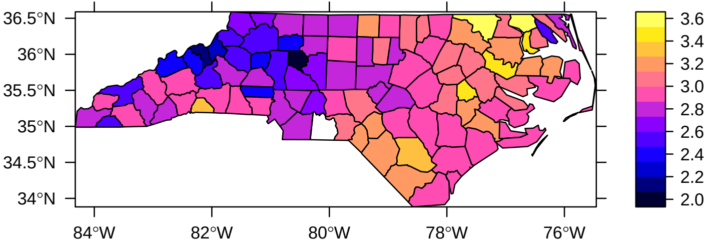
The call

```
sar <- spautolm(ft.SID74 ~ 1, data=mdata.4,
                listw=sids.nhbr.listw.4,
                weights=BIR74, family="SAR")
summary(sar)
```

yields estimates: $\hat{m} = 2.94$, $\hat{\phi} = 0.8683$ and $\hat{\sigma} = 35.55$

SAR Example: SIDS in North Carolina III

Predicted transformed data

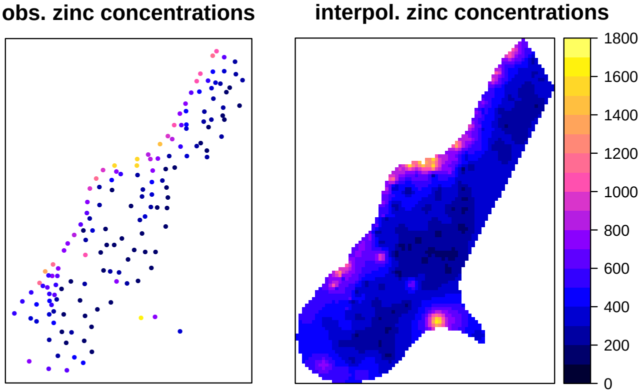


Random Fields

The following examples follow these of the gstat package [8] and the book [6].

Heavy metal concentrations in the soil along the Meuse riverbank have been sampled.

How do the concentrations on the full grid look like?



Stationarity and Isotropy

stationarity The process "looks" the same at each location (e.g. mean and variance do not change from east to west)

isotropy The dependence between locations is determined only by their separating distance neglecting the direction (e.g. locations 2 km apart along the north-south axis are as correlated as stations 2 km apart along the east-west axis)

Some tricks exist to weaken these assumptions (e.g. rotating and rescaling coordinates).

Variograms I

The dependence across space of a random field Z is assessed using a *variogram* γ:

$$\gamma(h) = \frac{1}{2}\mathrm{E}\big(\mathrm{Z}(\mathrm{s}) - \mathrm{Z}(\mathrm{s}+\mathrm{h})\big)^2$$

the empirical estimator looks like

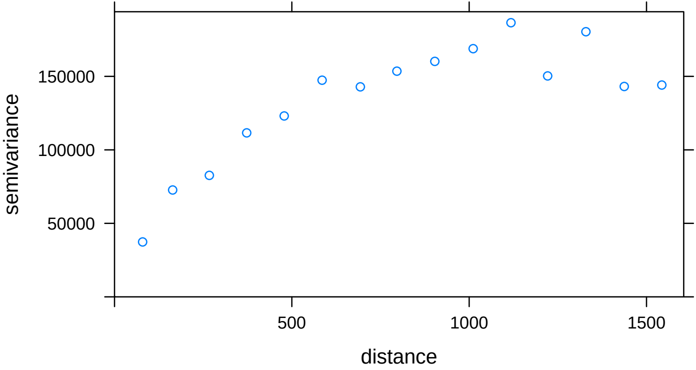
$$\hat{\gamma}(h) = \frac{1}{2|N_h|} \sum_{(i,j)\in N_h} \big(Z(s_i) - Z(s_j)\big)^2$$

while $N_h = \{(i,j) : h - \epsilon \leq ||s_i - s_j|| \leq h + \epsilon\}$

Variograms II

The *sample variogram* is obtained through

```
vgmMeuse <- variogram(zinc~1, meuse)
```



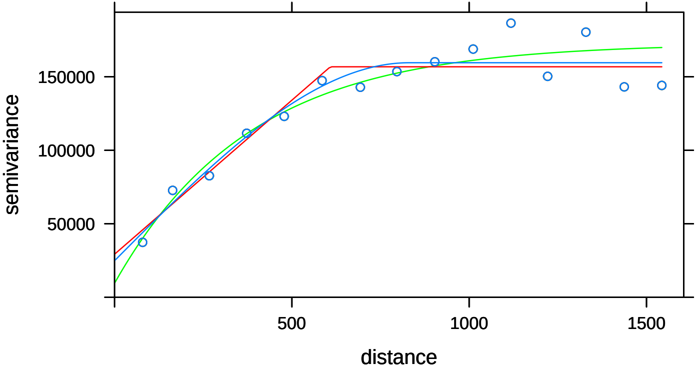
Variograms III

And a theoretical *variogram model* can be fitted

```
> head(vgm())
  short                          long
1   Nug                  Nug (nugget)
2   Exp             Exp (exponential)
3   Sph               Sph (spherical)
4   Gau                Gau (gaussian)
5   Exc Exclass (Exponential class)
6   Mat                  Mat (Matern)

> vgmModelMeuse <- fit.variogram(vgmMeuse,
                                  vgm(0.6, "Sph", 1000, 0.1))
vgmModelMeuse
  model     psill    range
1   Nug  24813.21   0.0000
2   Sph 134753.99 831.2953
```

Variograms IV



Kriging I

Certain variogram models can be used to parametrize a covariance matrix for a Gaussian random field over a finite set of locations $s_1, \dots, s_n$:

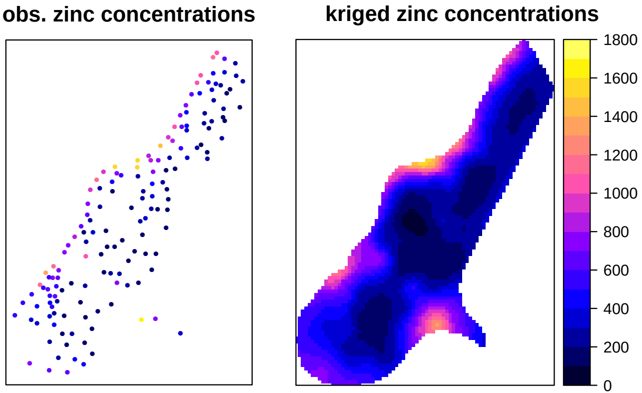
$$Z \sim \mathrm{Gau}\big(\boldsymbol{\mu}, \Sigma\big)$$

while $\Sigma = (\sigma_{ij}^2)_{ij}$ and $\sigma_{ij}^2 = \sigma^2 - \gamma(||s_i - s_j||)$, $1 \leq i, j \leq n$ with $\sigma^2 = \mathrm{Var}\big(\mathrm{Z(s)}\big)$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)$.

Predictions can be made using matrix inversion and matrix multiplications.

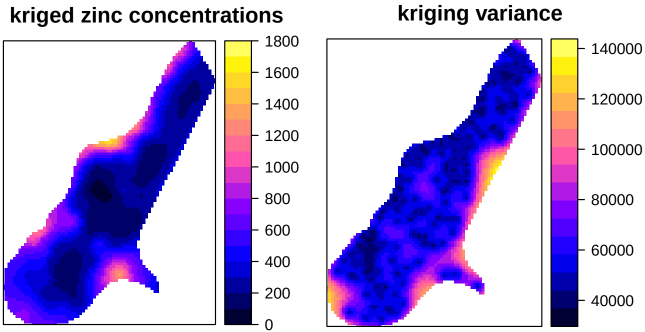
Kriging II

```
krige(zinc~1, meuse, meuse.grid, model=vgmModelMeuse)
```



Kriging III

The model quantifies how *uncertain* it is about the estimates through the kriging variance:



Overview of kriging types

simple kriging the mean value is known

ifgi
Institute for Geoinformatics
University of Münster

Overview of kriging types

simple kriging the mean value is known

ordinary kriging prediction based on coordinates

Overview of kriging types

simple kriging the mean value is known

ordinary kriging prediction based on coordinates

universal kriging prediction based on coordinates and additional regressors (distance to the river)

ifgi
Institute for Geoinformatics
University of Münster

Overview of kriging types

simple kriging: the mean value is known

ordinary kriging: prediction based on coordinates

universal kriging: prediction based on coordinates and additional regressors (distance to the river)

co-kriging: the cross-variogram between two variables is as well exploit (zinc and lead)

Spatio-Temporal Data

$S \times T$ works as a data structure, but modelling needs to consider special properties of the product of space and time.

direction Today's values influence tomorrow, but will not take effect on yesterday's values.

anisotropy What is the equivalent in terms of dependence of 1 m separation in seconds or minutes?

The easiest way to think of spatio-temporal data is as time slices - but this neglects the temporal dependence.

After modelling temporal trend or periodicities, the residuals might be modelled as a spatio-temporal random field.

The spatio-temporal variogram

Extending the variogram to a twoplace function for spatio-temporal random fields $Z(s,t)$:

$$\gamma(h,u) = \mathrm{E}\big(\mathrm{Z}(\mathrm{s},\mathrm{t}) - \mathrm{Z}(\mathrm{s}+\mathrm{h},\mathrm{t}+\mathrm{u})\big)^2$$

at any location (s,t). And empirical version

$$\hat{\gamma}(h,u) = \frac{1}{2|N_{h,u}|} \sum_{(i,j) \in N_{h,u}} \big(Z(s_i,t_i) - Z(s_j,t_j)\big)^2$$

while $N_{h,u} = \left\{ (i,j) \;\middle|\; \begin{array}{l} h - \epsilon_s \leq ||s_i - s_j|| \leq h + \epsilon_s \\ u - \epsilon_t \leq t_i - t_j \leq u + \epsilon_t \end{array} \right\}$

Spatio-Temporal Variogram Models I

Modelling dependence over space differently from the dependence over time yields another degree of freedom.

metric the temporal component is adjusted using only an anisotropy correction

$$\gamma_m(h,u) = \gamma_j(\sqrt{h^2 + (\kappa \cdot u)^2})$$

separable spatial and temporal component are added up neglecting interactions between space and time

$$\begin{aligned}\gamma_{sep}(h,u) &= \mathrm{nug} \cdot \mathbf{1}_{h>0,u>0} \\ &+ \mathrm{sill} \cdot \big(\gamma_s(h) + \gamma_t(u) - \gamma_s(h)\gamma_t(u)\big)\end{aligned}$$

Spatio-Temporal Variogram Models II

product-sum a product interaction term is added to the separable model

$$\begin{aligned}\gamma_{ps}(h,u) &= \mathrm{nug}\cdot \mathbf{1}_{h>0,u>0} \\ &+ \gamma_s(h)+\gamma_t(u)-k\gamma_s(h)\gamma_t(u)\end{aligned}$$

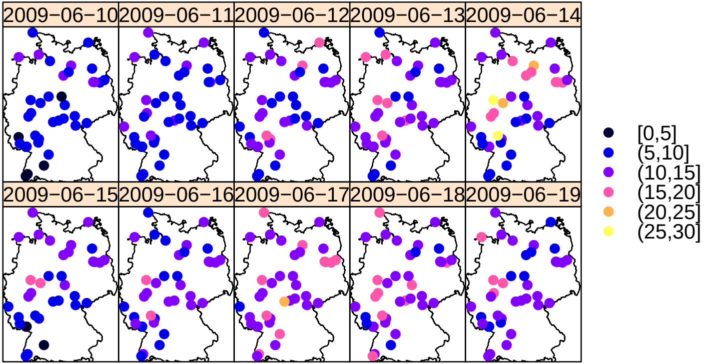
The parameter k needs to fulfil $0 < k \leq 1/(\max(\mathrm{sill}_s, \mathrm{sill}_t))$.

sum-metric the covariance structure is composed out of spatial and temporal components added with a metric model capturing interactions

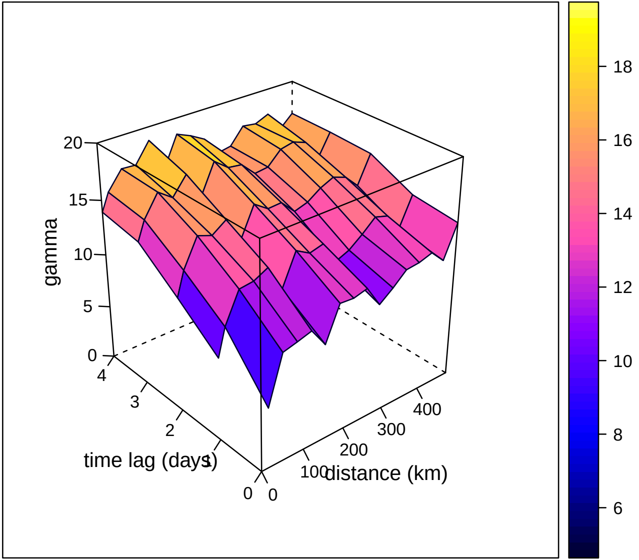
$$\gamma_{sm}(h,u) = \gamma_s(h)+\gamma_t(u)+\gamma_j(\sqrt{h^2+(\kappa\cdot u)^2})$$

Example

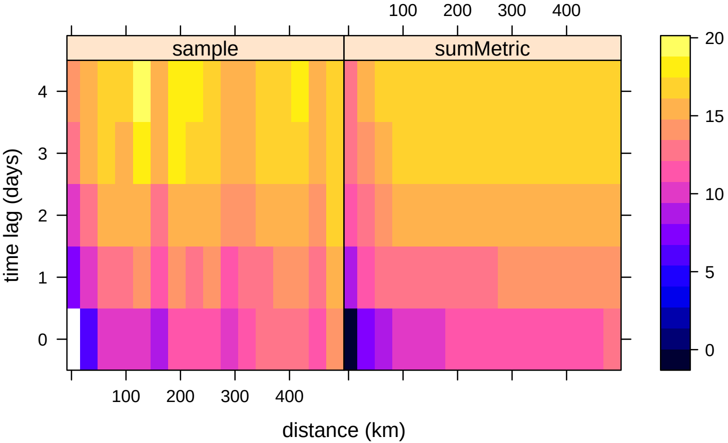
A set of spatially spread time series of daily measurements: what does the random field look like all over Germany?



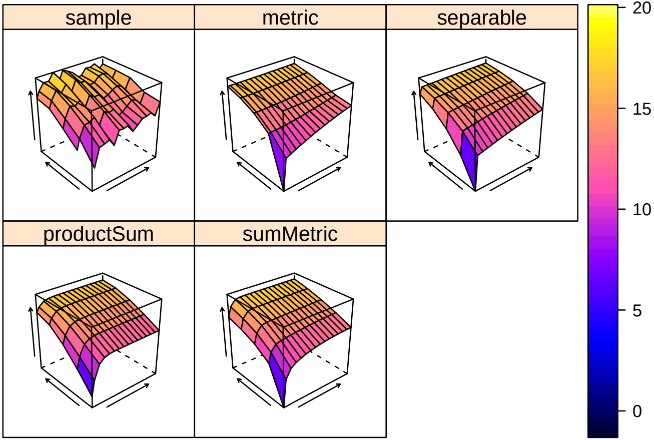
Spatio-Temporal Sample Variogram



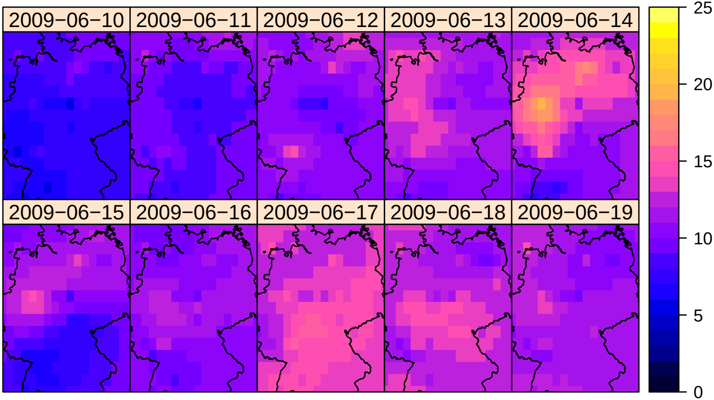
variogram of the sum-metric model



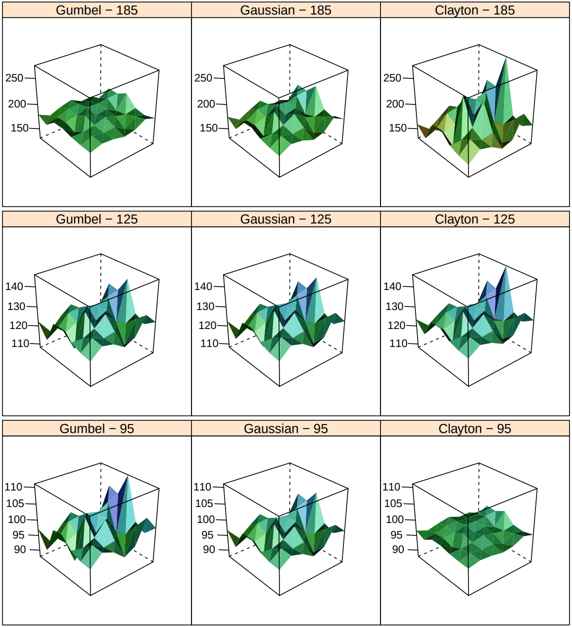
variogram of all spatio-temporal models



kriged map for 10 days



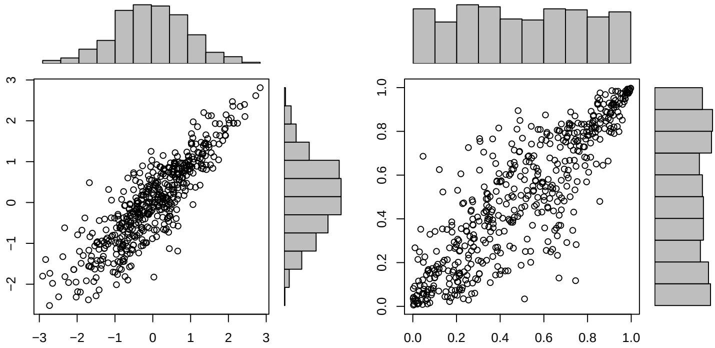
What if the world happens to be non-Gaussian?



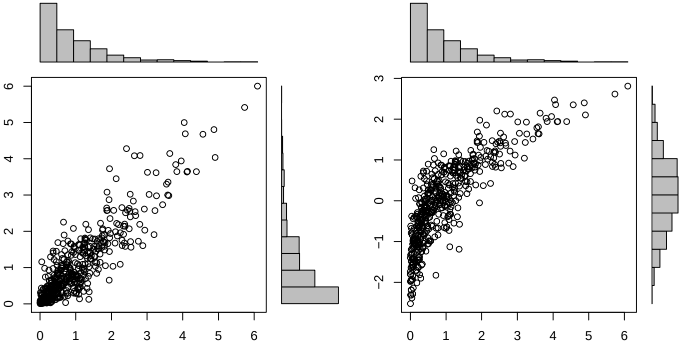
Bivariate Copulas I

Copulas allow to model dependencies much more detailed than a typical correlation value.

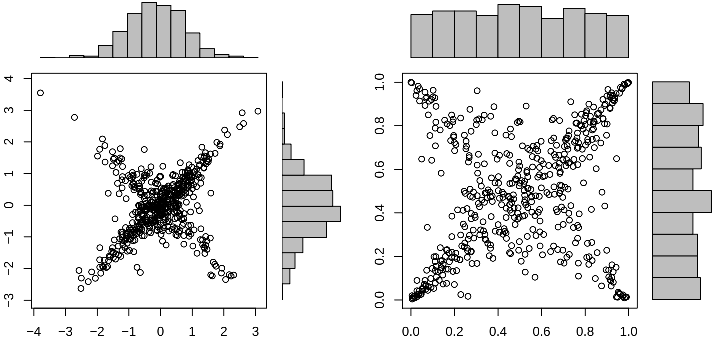
Instead of a single value, a full distribution is fitted describing dependence.



Bivariate Copulas II



Bivariate Copulas III



See the copulatheque for further interactive examples.

Sklar's Theorem

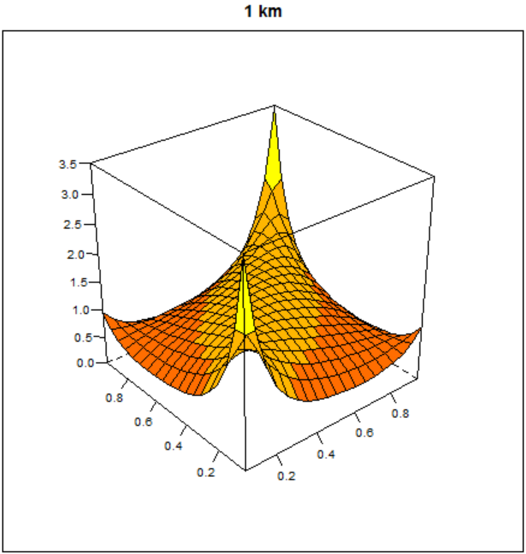
The power of copulas originates from Sklar's Theorem [9]:

Any multivariate distribution H can be decomposed into its marginal distribution functions F_1, ..., F_d and its copula C:

$$H(x_1, \ldots, x_d) = C\big(F_1(x_1), \ldots, F_d(x_d)\big)$$

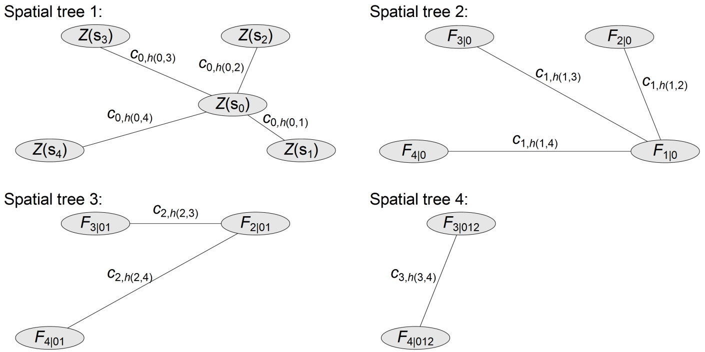
Strength and Shape changes with distance

The *Bivariate Spatial Copula* is a convex combination of bivariate copula families parametrised by distance.

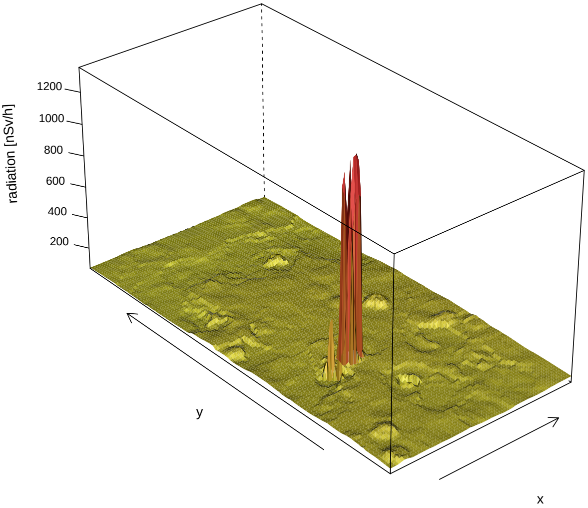


Spatial Vine Copulas

A *Vine Copula* connects bivariate copulas to multivariate copulas [1, 7].



Spatial Vine Copula Prediction



References I

[1] Kjersti Aas, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44:182–198, 2009.

[2] Adrian Baddeley. *Analysing spatial point patterns in R*, 2010. Version 4.1.

[3] Adrian Baddeley and Rolf Turner. spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software*, 12(6):1 – 42, 2005.

[4] Mark Berman and T. Rolf Turner. Approximating point process likelihoods with GLIM. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 41(1):pp. 31–38, 1992.

[5] Roger Bivand. *spdep: Spatial dependence: weighting schemes, statistics and models*, 2014. R package version 0.5-71.

[6] Roger S. Bivand, Edzer Pebesma, and Virgilio Gómez-Rubio. *Applied Spatial Data Analysis with R*. Use R! Springer, 2nd edition, 2013.

[7] Roger M.. Cooke, Harry Joe, and Kjersti Aas. Vines arise. In Dorota Kurowicka and Harry Joe, editors, *Dependence Modelling*, chapter 3, pages 37–72. World Scientific Publishing Co, 2011.

[8] Edzer J. Pebesma. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, 30:683–691, 2004.

[9] A. Sklar. Fonctions de répartition á n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229 – 231, 1959.

[10] W. R. Tobler. A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240, 1970.