

# Stochastic Modelling of Spatial Data

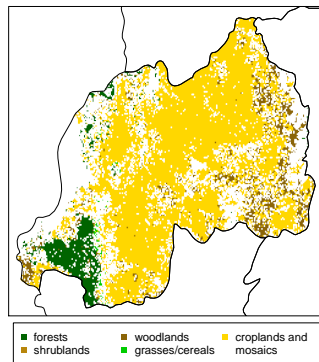
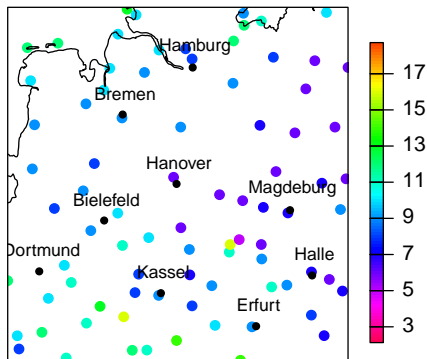
Marco Oesting

Universität Stuttgart

KCDS Summer School 2023  
September 19, 2023, Karlsruhe



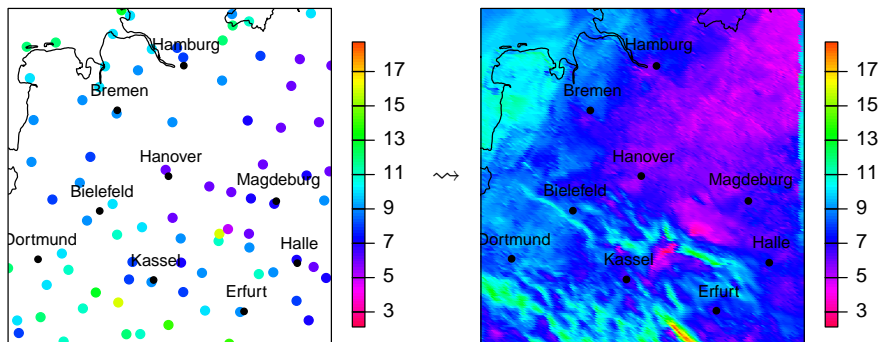
# Spatial Data in Environmental Science



## Two Important Types:

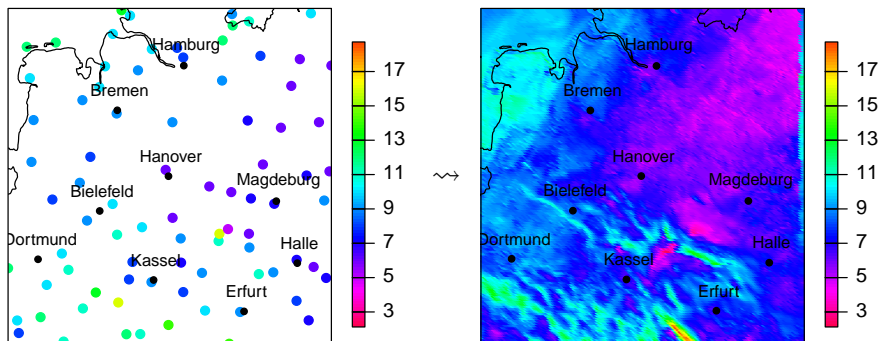
- data at scattered sites  
(in-situ measurements)
- gridded data  
(satellite data, output of numerical models)

# Spatial Modelling: Aims



- 1 prediction at ungauged sites, higher resolution (interpolation, “downscaling”)
- 2 simulation of “artificial data”, unobserved scenarios

# Spatial Modelling: Aims



- 1 prediction at ungauged sites, higher resolution (interpolation, “downscaling”)
- 2 simulation of “artificial data”, unobserved scenarios

↪ probabilistic models

# Building Blocks of Probabilistic Models

- 1 probability distribution at a single site

**Assumption:** same type of distribution in whole region of interest;  
exact distribution might depend on covariates (LON, LAT, ALT, ...)

# Building Blocks of Probabilistic Models

- 1 probability distribution at a single site

**Assumption:** same type of distribution in whole region of interest; exact distribution might depend on covariates (LON, LAT, ALT, ...)

- 2 model for spatial dependence structure

## Tobler's First Law of Geography

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

**Assumption:** strength of dependence between data at two sites is function of distance and direction

# Outline

## 1 “Classical” Gaussian Modelling

- Gaussian Modelling at a Single Site
- Gaussian Modelling in Space
- Simulation & Interpolation

## 2 Modelling of Extreme Events

- Modelling an Extreme Event at a Single Site
- Modelling an Extreme Event in Space
- Simulation & Interpolation

# Outline

## 1 “Classical” Gaussian Modelling

- Gaussian Modelling at a Single Site
- Gaussian Modelling in Space
- Simulation & Interpolation

## 2 Modelling of Extreme Events

- Modelling an Extreme Event at a Single Site
- Modelling an Extreme Event in Space
- Simulation & Interpolation



# Outline

## 1 “Classical” Gaussian Modelling

- Gaussian Modelling at a Single Site
- Gaussian Modelling in Space
- Simulation & Interpolation

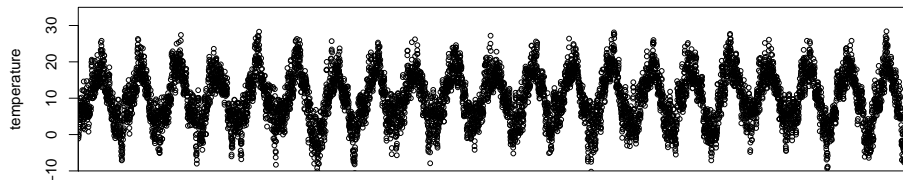
## 2 Modelling of Extreme Events

- Modelling an Extreme Event at a Single Site
- Modelling an Extreme Event in Space
- Simulation & Interpolation

# Gaussian Data

**Aim:** model environmental variable  $X$  at site  $s$  (Notation:  $X(s)$ )

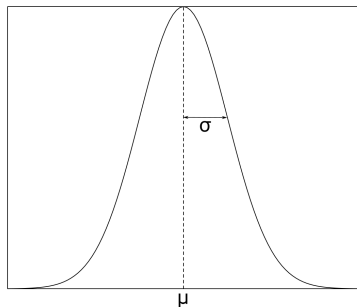
**Example:** Daily temperature data at Münster Airport 1990–2010



For continuous variables (e.g. temperature, pressure, air pollution):  
 $X(s)$  can often be assumed to follow **Gaussian distribution**

$$X(s) \sim \mathcal{N}(\mu, \sigma^2)$$

# Gaussian Distribution



$$X(s) \sim \mathcal{N}(\mu, \sigma^2)$$

$\mu$ : mean (location parameter)

$\sigma$ : standard deviation  
(scale parameter)

## Gaussian PDF

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ - \left( \frac{x - \mu}{\sigma} \right)^2 \right\}, \quad x \in \mathbb{R}$$

## Gaussian Distribution (cont'd)

$$X(\mathbf{s}) \sim \mathcal{N}(\mu, \sigma^2)$$

## Gaussian Distribution (cont'd)

mean and variance may differ from site to site due to conditions . . .

$$X(\mathbf{s}) \sim \mathcal{N}(\mu(\mathbf{s}), \sigma^2(\mathbf{s}))$$

# Gaussian Distribution (cont'd)

mean and variance may differ from site to site due to conditions . . .

$$X(\mathbf{s}) \sim \mathcal{N}(\mu(\mathbf{s}), \sigma^2(\mathbf{s}))$$

**First Approach:** estimate  $\mu$  and  $\sigma^2$  from data separately for each site, e.g. by empirical counterparts

$$\begin{aligned}\hat{\mu}(\mathbf{s}) &= \frac{1}{n} \sum_{i=1}^n X_i(\mathbf{s}) \\ \hat{\sigma}^2(\mathbf{s}) &= \frac{1}{n-1} \sum_{i=1}^n (X_i(\mathbf{s}) - \hat{\mu}(\mathbf{s}))^2\end{aligned}$$

# Gaussian Distribution (cont'd)

mean and variance may differ from site to site due to conditions . . .

$$X(\mathbf{s}) \sim \mathcal{N}(\mu(\mathbf{s}), \sigma^2(\mathbf{s}))$$

**First Approach:** estimate  $\mu$  and  $\sigma^2$  from data separately for each site, e.g. by empirical counterparts

$$\begin{aligned}\hat{\mu}(\mathbf{s}) &= \frac{1}{n} \sum_{i=1}^n X_i(\mathbf{s}) \\ \hat{\sigma}^2(\mathbf{s}) &= \frac{1}{n-1} \sum_{i=1}^n (X_i(\mathbf{s}) - \hat{\mu}(\mathbf{s}))^2\end{aligned}$$

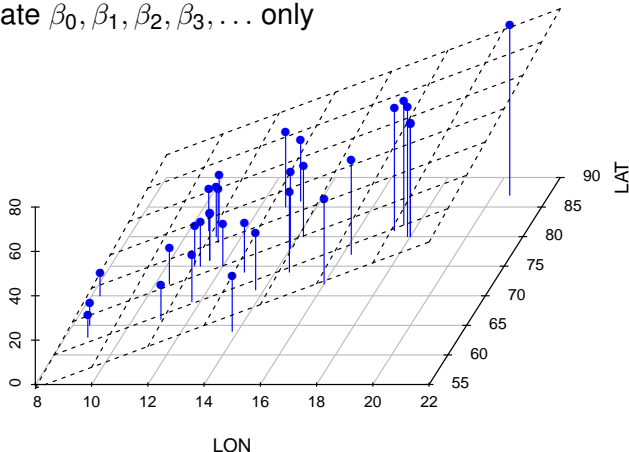
**Problem:** how to model variable at  **ungauged**  sites?

## Gaussian Distribution (cont'd)

**Second Approach:** assume further spatial structure of  $\mu$  and  $\sigma^2$ , e.g.

$$\mu(s) = \beta_0 + \beta_1 \cdot \text{LON}(s) + \beta_2 \cdot \text{LAT}(s) + \beta_3 \cdot \text{ALT}(s) + \dots$$

and estimate  $\beta_0, \beta_1, \beta_2, \beta_3, \dots$  only



~> model for **any site** within region



# How to estimate the $\beta$ 's?

## 1. Least-Squares-Fit of the Raw Estimates

Minimize

$$\sum_s (\hat{\mu}(s) - \beta_0 - \beta_1 \cdot \text{LON}(s) - \beta_2 \cdot \text{LAT}(s) - \beta_3 \cdot \text{ALT}(s) - \dots)^2$$

# How to estimate the $\beta$ 's?

## 1. Least-Squares-Fit of the Raw Estimates

Minimize

$$\sum_s (\hat{\mu}(s) - \beta_0 - \beta_1 \cdot \text{LON}(s) - \beta_2 \cdot \text{LAT}(s) - \beta_3 \cdot \text{ALT}(s) - \dots)^2$$

## 2. Fit of Linear Regression Model for the Observations

If the variance  $\sigma^2(s) \equiv \sigma^2$  is nearly constant and the data are assumed to be independent (also across sites), we can write

$$X_i(s) = \beta_0 + \beta_1 \cdot \text{LON}(s) + \beta_2 \cdot \text{LAT}(s) + \beta_3 \cdot \text{ALT}(s) + \dots + \varepsilon_i(s)$$

where  $\varepsilon_i(s) \sim_{i.i.d.} \mathcal{N}(0, \sigma^2)$ .

# How to estimate the $\beta$ 's?

## 1. Least-Squares-Fit of the Raw Estimates

Minimize

$$\sum_s (\hat{\mu}(s) - \beta_0 - \beta_1 \cdot \text{LON}(s) - \beta_2 \cdot \text{LAT}(s) - \beta_3 \cdot \text{ALT}(s) - \dots)^2$$

## 2. Fit of Linear Regression Model for the Observations

If the variance  $\sigma^2(s) \equiv \sigma^2$  is nearly constant and the data are assumed to be independent (also across sites), we can write

$$X_i(s) = \beta_0 + \beta_1 \cdot \text{LON}(s) + \beta_2 \cdot \text{LAT}(s) + \beta_3 \cdot \text{ALT}(s) + \dots + \varepsilon_i(s)$$

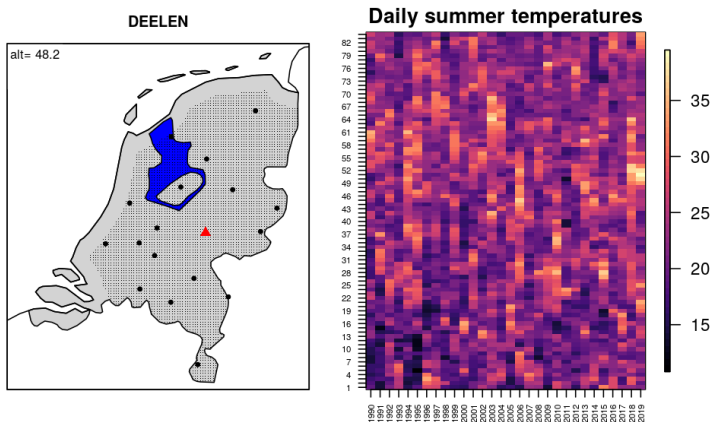
where  $\varepsilon_i(s) \sim_{i.i.d.} \mathcal{N}(0, \sigma^2)$ .

Then, minimizing

$$\sum_s \sum_{i=1}^n (X_i(s) - \beta_0 - \beta_1 \cdot \text{LON}(s) - \beta_2 \cdot \text{LAT}(s) - \beta_3 \cdot \text{ALT}(s) - \dots)^2$$

is equivalent to the (independent) maximum likelihood estimator.

# Data Example in Julia: Dutch Summer Temperatures



daily mean temperature data for...

- 18 Dutch weather stations
- 29 summers (JJA 1991–2019)

}  $\Rightarrow$  14-day block means  
(6 per year)

# Julia Packages

`using Statistics`

`using DataFrames`

`using GLM`

`using Plots`

`using GaussianProcesses`

`using GaussianRandomFields`

`using Extremes`

`using Optim`

`using JLD2`

# Outline

## 1 “Classical” Gaussian Modelling

- Gaussian Modelling at a Single Site
- **Gaussian Modelling in Space**
- Simulation & Interpolation

## 2 Modelling of Extreme Events

- Modelling an Extreme Event at a Single Site
- Modelling an Extreme Event in Space
- Simulation & Interpolation

# Gaussian Processes

**Main Assumption:** Data belong to a Gaussian process  $\{X(s)\}_{s \in S}$ !

## Ingredients:

- mean  $\mu(s) = \mathbb{E}\{X(s)\}$  ✓
- variance  $\sigma^2(s) = \text{Var}\{X(s)\}$  ✓

# Gaussian Processes

**Main Assumption:** Data belong to a Gaussian process  $\{X(s)\}_{s \in S}$ !

## Ingredients:

- mean  $\mu(s) = \mathbb{E}\{X(s)\}$  ✓
- variance  $\sigma^2(s) = \text{Var}\{X(s)\}$  ✓
- correlation  $\rho(s_1, s_2) = \text{Corr}\{X(s_1), X(s_2)\}$  ?

## Tobler's First Law of Geography

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

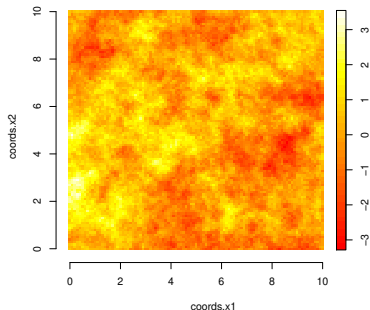
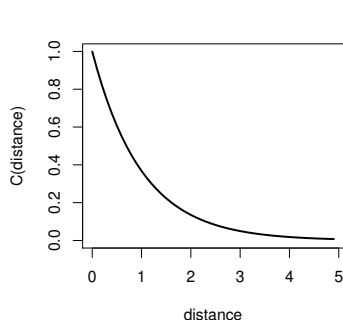
typical model for correlation:

- ▶ close to 1 (strong correlation) as  $s_1$  and  $s_2$  are close to each other
- ▶ decreasing as distance increases
- ▶ eventually tends to 0 (independence)



# Example: Exponential Correlation

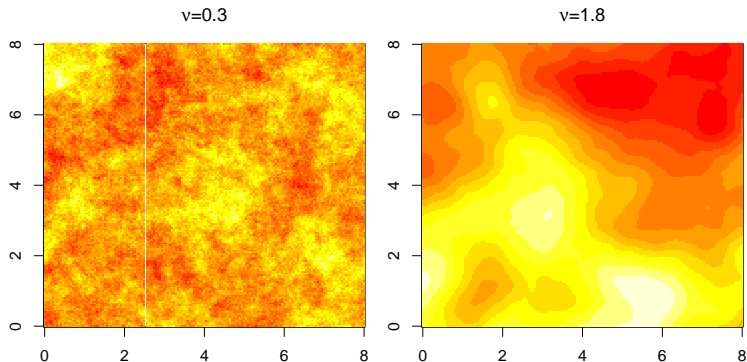
$$\rho(s_1, s_2) = \exp(-a\|s_1 - s_2\|), \quad a > 0$$



**Note:** Not every function is an admissible correlation function, only **positive definite** ones.

## Example: Whittle–Matérn Model

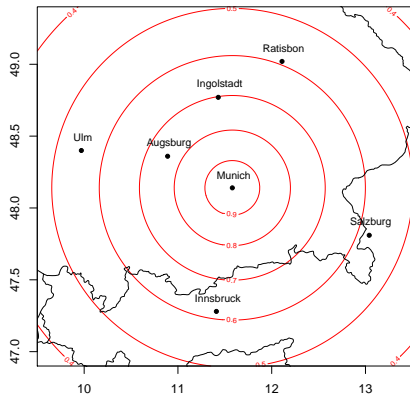
$$\rho(s_1, s_2) = \frac{2^{\nu-1}}{\Gamma(\nu)} (a\|s_1 - s_2\|)^{\nu} K_{\nu}(a\|s_1 - s_2\|), \quad a > 0, \nu > 0$$



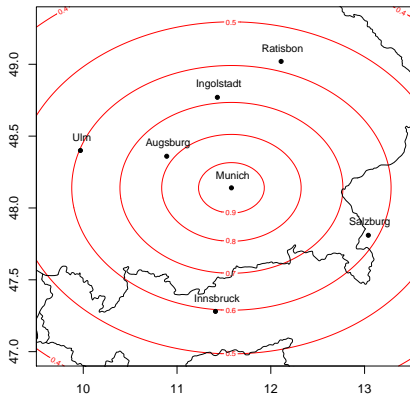
**Note:** parameter  $\nu$  allows for different degrees of smoothness

# Anisotropy: The direction matters

isotropic:



anisotropic:



# Model Fit

## 1 build parametric models for

- ▶ mean  $\mu(s)$
- ▶ variance  $\sigma^2(s)$
- ▶ correlation  $\rho(s_1, s_2)$

## 2 estimate parameters from data

- ▶ least squares fit to empirical estimates (see above)
- ▶ joint **maximum likelihood** (including dependence structure)
- ▶ ...

~> full probabilistic model for variable within whole region

# Outline

## 1 “Classical” Gaussian Modelling

- Gaussian Modelling at a Single Site
- Gaussian Modelling in Space
- **Simulation & Interpolation**

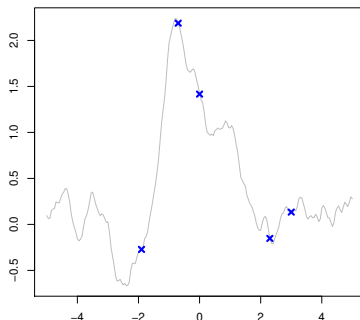
## 2 Modelling of Extreme Events

- Modelling an Extreme Event at a Single Site
- Modelling an Extreme Event in Space
- Simulation & Interpolation

# Interpolation & Prediction

**Given:** data at sites  $s_1, \dots, s_n$  belonging to Gaussian process

**Question:** What does the process look like in between?



# Interpolation & Prediction

**Given:** data at sites  $s_1, \dots, s_n$  belonging to Gaussian process

**Question:** What does the process look like in between?

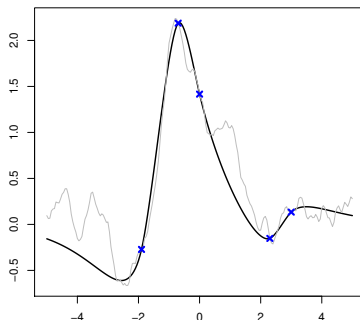
**(At least) Two Options:**

- pointwise prediction

↪ Kriging /

Gaussian Process Regression

pointwise “optimal” prediction,  
smooth interpolation by  
conditional mean



# Interpolation & Prediction

**Given:** data at sites  $s_1, \dots, s_n$  belonging to Gaussian process

**Question:** What does the process look like in between?

**(At least) Two Options:**

- pointwise prediction

↪ **Kriging /**

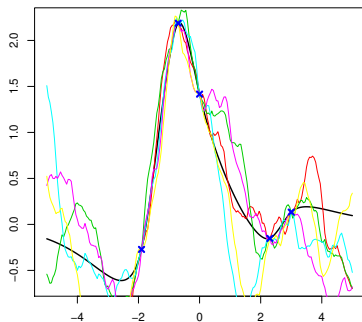
**Gaussian Process Regression**

pointwise “optimal” prediction,  
smooth interpolation by  
conditional mean

- structure of the sample path

↪ **conditional simulation**

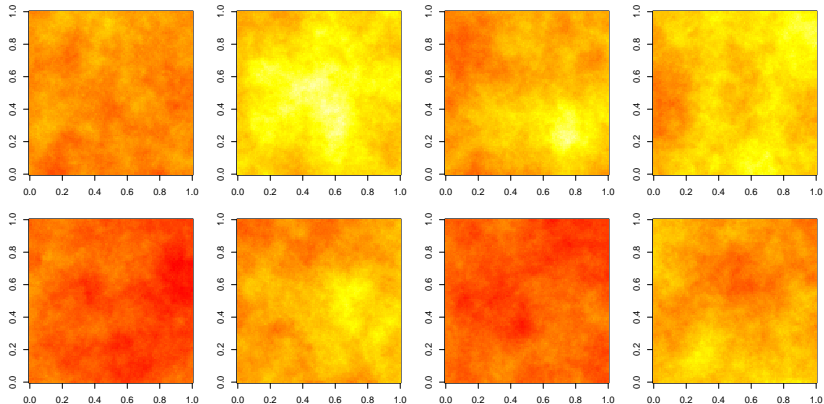
recovers all potential sample  
paths matching the data





# Simulation of Gaussian Processes

**Aim:** generate realizations of process



# Simulation of Gaussian Processes

How to simulate  $N$ -dimensional vector

$$\mathbf{Y} \sim \mathcal{N}(\mathbf{0}, \Sigma) \quad ?$$

## 1. Direct Simulation:

- ▶ based on observation  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \Rightarrow \mathbf{AX} \sim \mathcal{N}(\mathbf{0}, \mathbf{AA}^\top)$
- ▶ find decomposition (e.g. Cholesky)  $\Sigma = \mathbf{AA}^\top$  and simulate  $\mathbf{AX}$
- ▶ computational costs typically of order  $\mathcal{O}(N^3)$

## 2. Circulant Embedding:

- based on FFT
- works for GPs with isotropic covariance function on regular grid
- computational costs of order  $\mathcal{O}(N \log N)$

...

# Outline

## 1 “Classical” Gaussian Modelling

- Gaussian Modelling at a Single Site
- Gaussian Modelling in Space
- Simulation & Interpolation

## 2 Modelling of Extreme Events

- Modelling an Extreme Event at a Single Site
- Modelling an Extreme Event in Space
- Simulation & Interpolation

# Extreme Events ...

... have high impact

## Examples of extreme events in environmental science:

- floods
- storms
- avalanches
- heat waves
- drought
- ...



New Orleans after Hurricane Katrina (Photo: K. Niemi)

# Extreme Events ...

... are rare

## Example: Dutch Delta Project

building dikes which are supposed to fail once per 10 000 years

⇒ **Challenge:** assess probability of unobserved events



New Orleans after Hurricane Katrina (Photo: K. Niemi)

# Extreme Events ...

... are rare

## **Example: Dutch Delta Project**

building dikes which are supposed to fail once per 10 000 years

⇒ **Challenge:** assess probability of unobserved events

... often have a spatial extent

⇒ **Challenge:** accurate modelling of spatial dependence structure



New Orleans after Hurricane Katrina (Photo: K. Niemi)

# Outline

## 1 “Classical” Gaussian Modelling

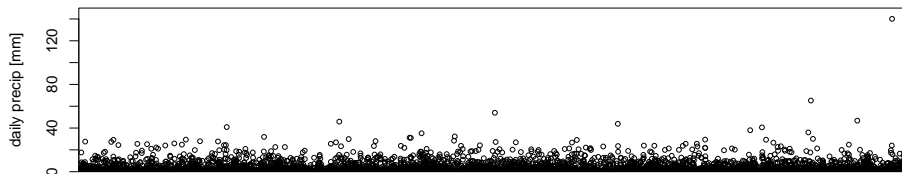
- Gaussian Modelling at a Single Site
- Gaussian Modelling in Space
- Simulation & Interpolation

## 2 Modelling of Extreme Events

- **Modelling an Extreme Event at a Single Site**
- Modelling an Extreme Event in Space
- Simulation & Interpolation

# Block Maxima Approach

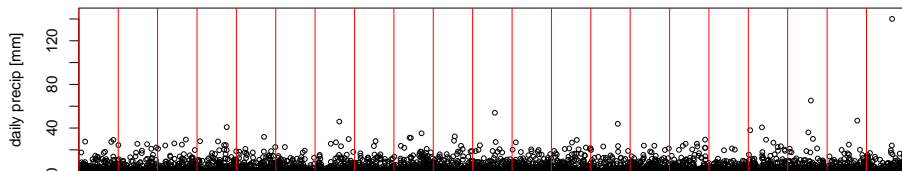
**Example:** Daily precipitation at Münster Airport 1990–2010





# Block Maxima Approach

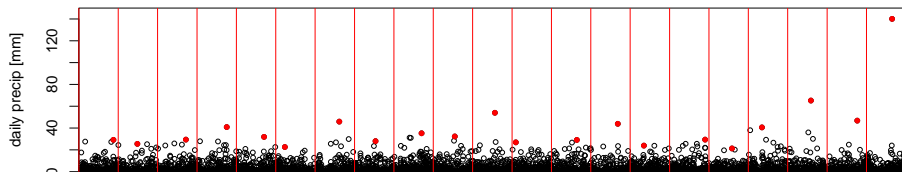
**Example:** Daily precipitation at Münster Airport 1990–2010



**Idea:** Divide time series into blocks of size 365 (or 366)

# Block Maxima Approach

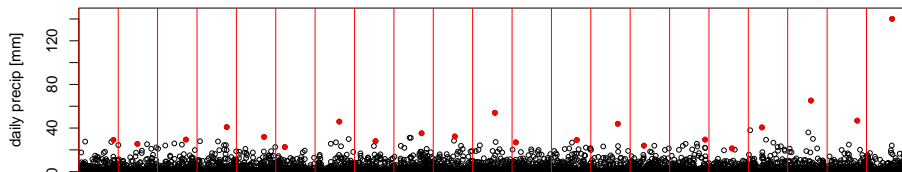
**Example:** Daily precipitation at Münster Airport 1990–2010



**Idea:** Divide time series into blocks of size 365 (or 366)  
and take maximum within each block  $\rightsquigarrow$  21 annual maxima

# Block Maxima Approach

**Example:** Daily precipitation at Münster Airport 1990–2010



**Idea:** Divide time series into blocks of size 365 (or 366)  
and take maximum within each block  $\rightsquigarrow$  21 annual maxima

**Theoretical result:** For large blocks, block maximum  $Z$  approximately  
follows **Generalized Extreme Value (GEV)** distribution.

$$\mathbb{P}(Z \leq z) \approx \exp \left( - \left( 1 + \xi \frac{z - \mu}{\sigma} \right)^{-1/\xi} \right)$$

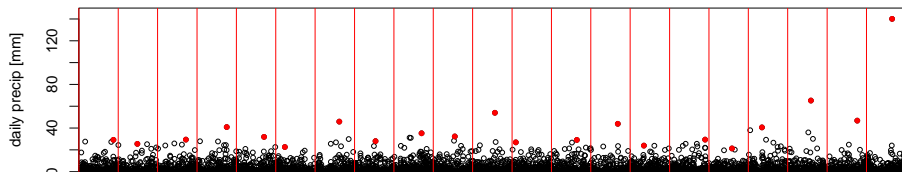
$\mu$ : location parameter

$\sigma$ : scale parameter

$\xi$ : shape parameter

# Block Maxima Approach

**Example:** Daily precipitation at Münster Airport 1990–2010



**Idea:** Divide time series into blocks of size 365 (or 366)  
and take maximum within each block  $\rightsquigarrow$  21 annual maxima

**Theoretical result:** For large blocks, block maximum  $Z$  approximately  
follows **Generalized Extreme Value (GEV)** distribution.

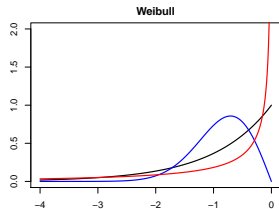
$$\mathbb{P}(Z \leq z) \approx \exp \left( - \left( 1 + \xi \frac{z - \mu}{\sigma} \right)^{-1/\xi} \right)$$

$\mu$ : location parameter

$\sigma$ : scale parameter

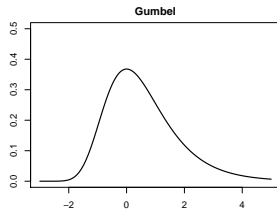
$\xi$ : **shape parameter**

# GEV: Flexible Modelling of Tails



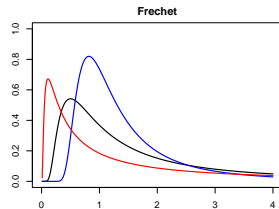
$$\xi < 0$$

upper end point



$$\xi = 0$$

light tail



$$\xi > 0$$

heavy tail

# GEV Distributions in Space

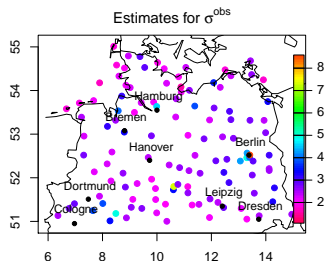
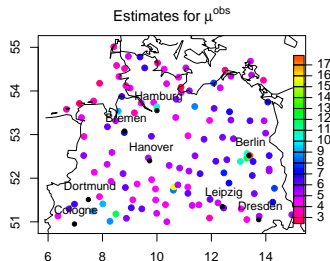
Block maximum  $Z$  at site  $s$ :

$$Z(s) \sim \text{GEV}(\mu(\mathbf{s}), \sigma(\mathbf{s}), \xi(\mathbf{s}))$$

## First Approach:

Estimate  $\mu$ ,  $\sigma$  and  $\xi$  separately for each site, e.g. via individual **maximum likelihood** (derivative of c.d.f. can be calculated easily)

⇒ how to model at ungauged stations?



Source: O., Schlather & Friedrichs 2017

# GEV Distributions in Space (cont'd)

Block maximum  $Z$  at site  $s$ :

$$Z(s) \sim \text{GEV}(\mu(\mathbf{s}), \sigma(\mathbf{s}), \xi(\mathbf{s}))$$

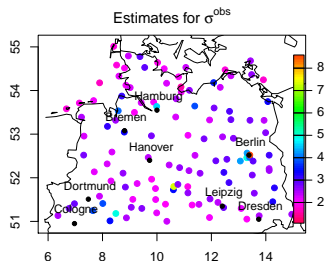
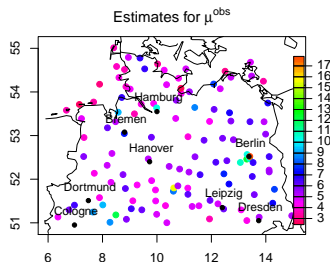
## Second Approach:

Assume some spatial structure of  $\mu$ ,  $\sigma$  and  $\xi$ , e.g.

$$\mu(\mathbf{s}) = \beta_0 + \beta_1 \cdot \text{LON} + \beta_2 \cdot \text{LAT} + \beta_3 \cdot \text{ALT}$$

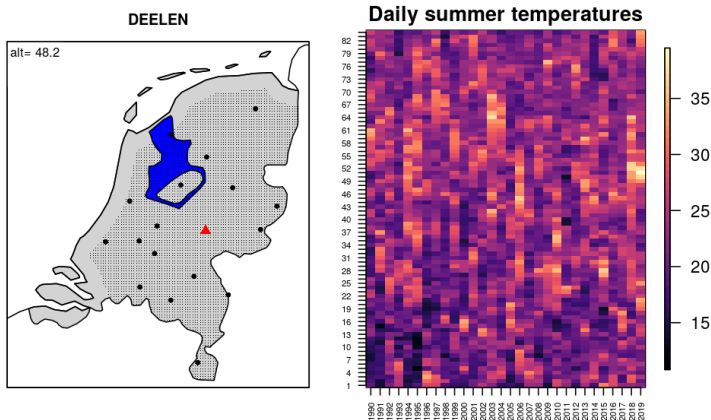
and estimate coefficients  $\beta_0, \beta_1, \beta_2, \beta_3$  (e.g. by joint **maximum likelihood** assuming independence across sites)

$\rightsquigarrow$  model for **any site** within region



Source: O., Schlather & Friedrichs 2017

# Data Example in Julia: Dutch Summer Temperatures



daily **maxima** temperature data for...

- 18 Dutch weather stations
- 29 summers (JJA 1991–2019)

}  $\Rightarrow$  14-day block **maxima**  
(6 per year)



# Outline

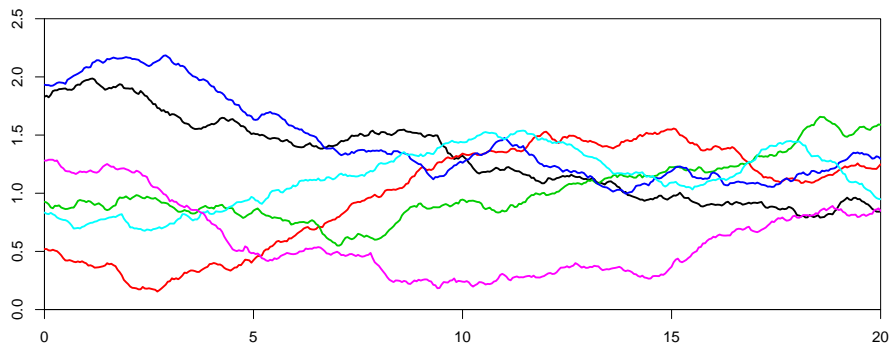
## 1 “Classical” Gaussian Modelling

- Gaussian Modelling at a Single Site
- Gaussian Modelling in Space
- Simulation & Interpolation

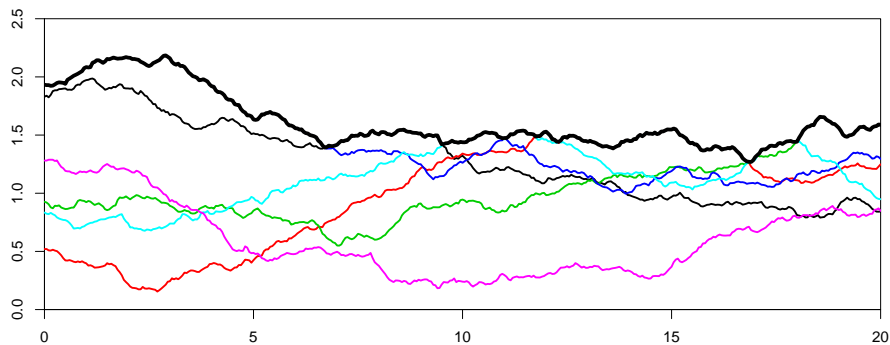
## 2 Modelling of Extreme Events

- Modelling an Extreme Event at a Single Site
- **Modelling an Extreme Event in Space**
- Simulation & Interpolation

# Block Maxima in Space

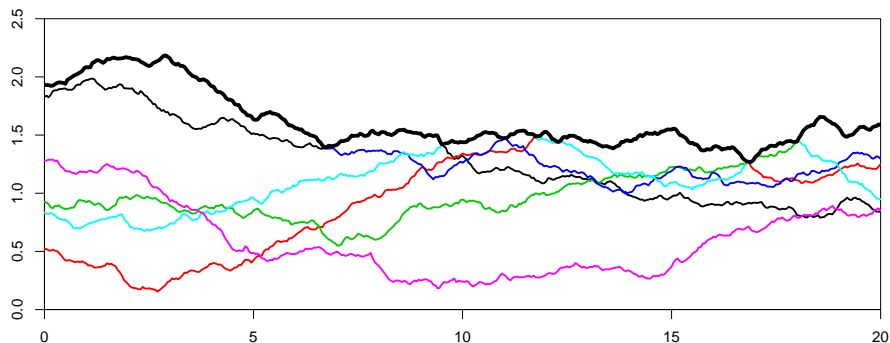


# Block Maxima in Space



Take pointwise maxima!

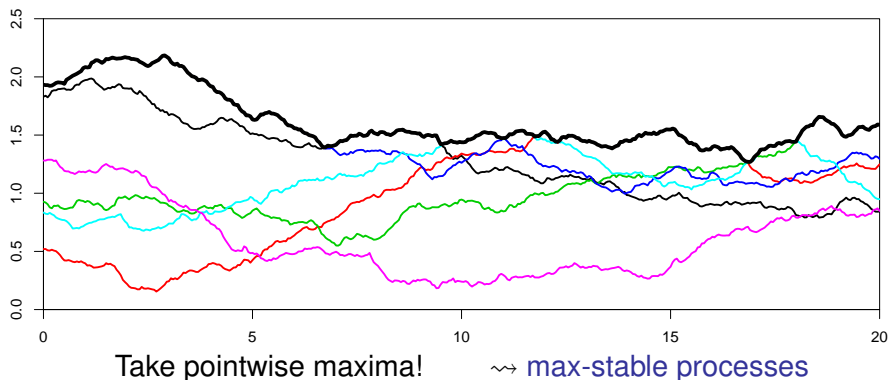
# Block Maxima in Space



Take pointwise maxima!

$\rightsquigarrow$  max-stable processes

# Block Maxima in Space



**Example:** Mixed Moving Maxima Processes (Smith, 1990)

~> interpretation as models for storms

Components:

- magnitude of storm
- storm center
- shape of storm

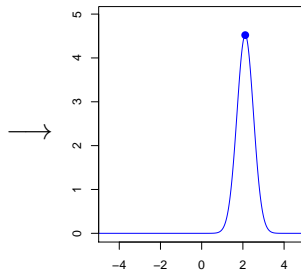
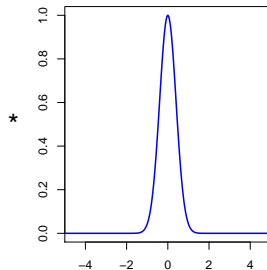
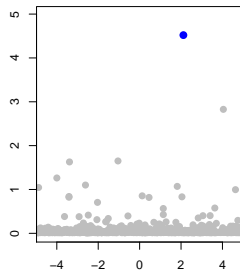
# Construction of Mixed Moving Maxima Processes

•  $U_i$ : magnitudes

•  $S_i$ : centers

•  $F_i$ : shapes

$$U_i \cdot F_i(x - S_i)$$



center / magnitude

shape

storm

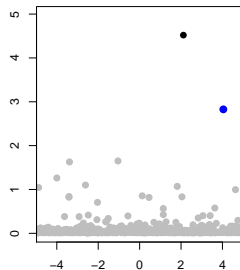
# Construction of Mixed Moving Maxima Processes

•  $U_i$ : magnitudes

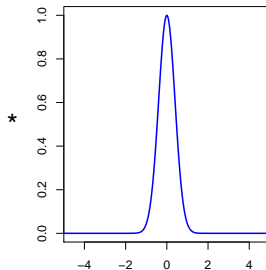
•  $S_i$ : centers

•  $F_i$ : shapes

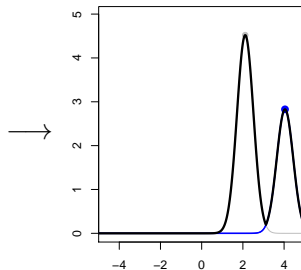
$$Z(x) = \max_{i=1,2,\dots} U_i \cdot F_i(x - S_i)$$



center / magnitude



shape



storm

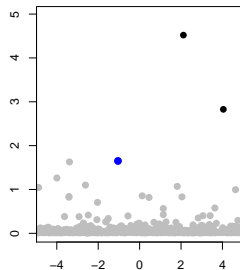
# Construction of Mixed Moving Maxima Processes

•  $U_i$ : magnitudes

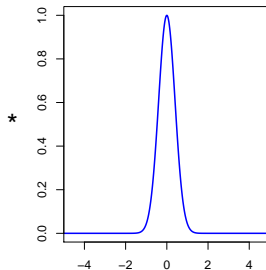
•  $S_i$ : centers

•  $F_i$ : shapes

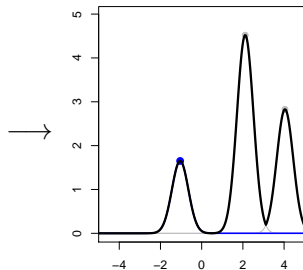
$$Z(x) = \max_{i=1,2,\dots} U_i \cdot F_i(x - S_i)$$



center / magnitude



shape



storm



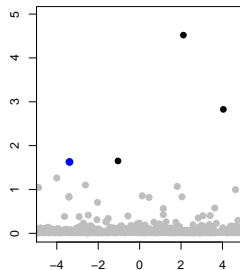
# Construction of Mixed Moving Maxima Processes

•  $U_i$ : magnitudes

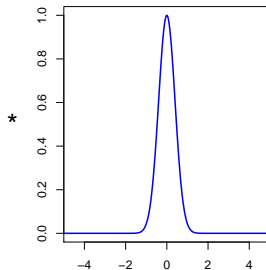
•  $S_i$ : centers

•  $F_i$ : shapes

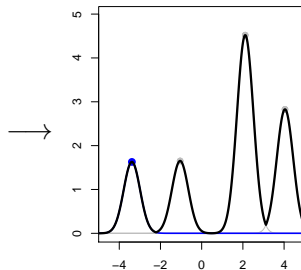
$$Z(x) = \max_{i=1,2,\dots} U_i \cdot F_i(x - S_i)$$



center / magnitude



shape



storm

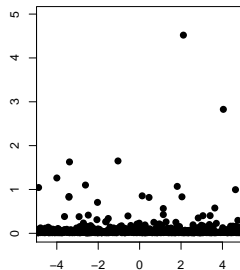
# Construction of Mixed Moving Maxima Processes

•  $U_i$ : magnitudes

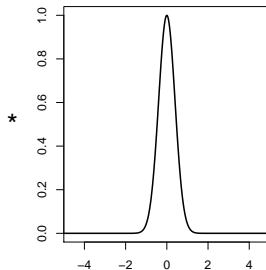
•  $S_i$ : centers

•  $F_i$ : shapes

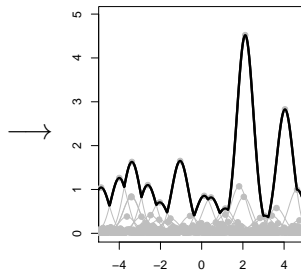
$$Z(x) = \max_{i=1,2,\dots} U_i \cdot F_i(x - S_i)$$



center / magnitude



shape



storm

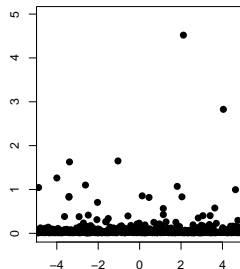
# Construction of Mixed Moving Maxima Processes

•  $U_i$ : magnitudes

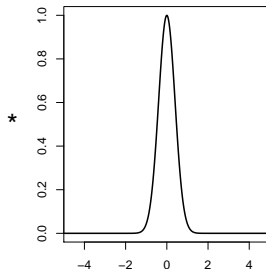
•  $S_i$ : centers

•  $F_i$ : shapes

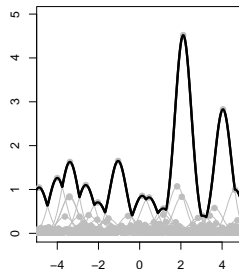
$$Z(x) = \max_{i=1,2,\dots} U_i \cdot F_i(x - S_i)$$



center / magnitude



shape



storm

~> storm shape determines the **spatial dependence** structure!

# Popular Max-Stable Models

Several Models use Gaussian processes as building blocks (storm shapes) . . .

- extremal Gaussian processes  
(Schlather 2002)
- Brown–Resnick processes  
(Kabluchko, Schlather & de Haan 2009)
- extremal- $t$  processes  
(Opitz 2009)

# Popular Max-Stable Models

Several Models use Gaussian processes as building blocks (storm shapes) . . .

- extremal Gaussian processes  
(Schlather 2002)
- Brown–Resnick processes  
(Kabluchko, Schlather & de Haan 2009)
- extremal- $t$  processes  
(Opitz 2009)

Some dependence properties inherited by Gaussian process, e.g.

- smoothness of the process
- anisotropy

# How to Measure Spatial Dependence in Extremes

**Problem:** Correlation not always defined!

Extremal Coefficient (Smith 1990, Schlather & Tawn 2003)

$$\mathbb{P}(Z(s_2) \text{ large} \mid Z(s_1) \text{ large})$$

# How to Measure Spatial Dependence in Extremes

**Problem:** Correlation not always defined!

Extremal Coefficient (Smith 1990, Schlather & Tawn 2003)

$$\theta(s_1, s_2) \approx 2 - \mathbb{P}(Z(s_2) \text{ large} \mid Z(s_1) \text{ large})$$

# How to Measure Spatial Dependence in Extremes

**Problem:** Correlation not always defined!

Extremal Coefficient (Smith 1990, Schlather & Tawn 2003)

$$\theta(s_1, s_2) = 2 - \lim_{u \rightarrow \infty} \mathbb{P}(Z(s_2) > u \mid Z(s_1) > u)$$



# How to Measure Spatial Dependence in Extremes

**Problem:** Correlation not always defined!

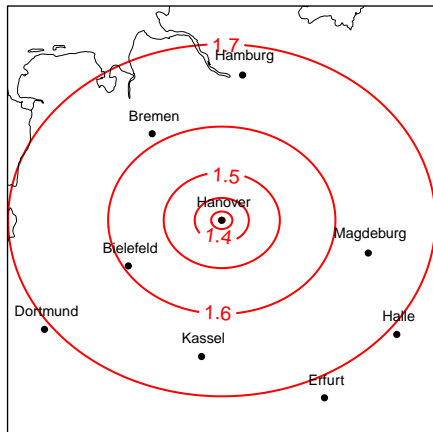
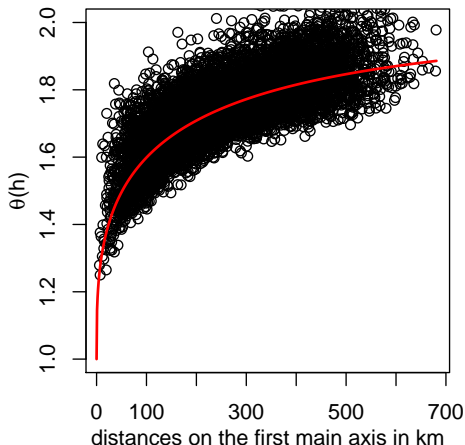
Extremal Coefficient (Smith 1990, Schlather & Tawn 2003)

$$\theta(s_1, s_2) = 2 - \lim_{u \rightarrow \infty} \mathbb{P}(Z(s_2) > u \mid Z(s_1) > u)$$

**Properties:**

- $1 \leq \theta(s_1, s_2) \leq 2$
- $\theta(s_1, s_2) = 1$  :  $Z(s_1)$  and  $Z(s_2)$  fully dependent
- $\theta(s_1, s_2) = 2$  :  $Z(s_1)$  and  $Z(s_2)$  (asymptotically) independent
- typically:  $\theta(s_1, s_2)$  increases as distance  $\|s_1 - s_2\|$  increases

# Example: Extreme Wind Gusts in Northern Germany



Source: O., Schlather & Friederichs 2017

# Model Fit

- 1 Parametric models for ...
  - ▶ marginal parameters  $\mu(s)$ ,  $\sigma(s)$  and  $\xi(s)$
  - ▶ parameters of underlying Gaussian process

# Model Fit

## 1 Parametric models for ...

- ▶ marginal parameters  $\mu(s)$ ,  $\sigma(s)$  and  $\xi(s)$
- ▶ parameters of underlying Gaussian process

## 2 Estimation of parameters:

- ▶ first estimation marginal parameters, then least squares fit to empirical extremal coefficients
- ▶ **maximum likelihood** more complicated, but feasible, cf. Huser, Dombry, Ribatet & Genton 2019, Dombry, Engelke & O. 2017
- ▶ ...

↪ full probabilistic model for variable within whole region

# Outline

## 1 “Classical” Gaussian Modelling

- Gaussian Modelling at a Single Site
- Gaussian Modelling in Space
- Simulation & Interpolation

## 2 Modelling of Extreme Events

- Modelling an Extreme Event at a Single Site
- Modelling an Extreme Event in Space
- **Simulation & Interpolation**

# Simulation of Spatial Extremes

**Aim:** create large amount of “artificial data” from fitted model  
(including rare/unobserved events) by simulations

~> assess risk of certain “extreme” scenarios

**Remind:** max-stable process = maximum over infinitely many “storms”

~> simulation difficult

# Simulation of Spatial Extremes

**Aim:** create large amount of “artificial data” from fitted model  
(including rare/unobserved events) by simulations

~> assess risk of certain “extreme” scenarios

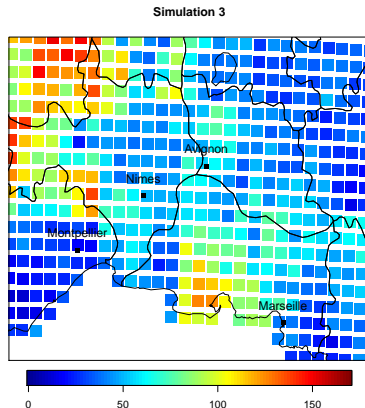
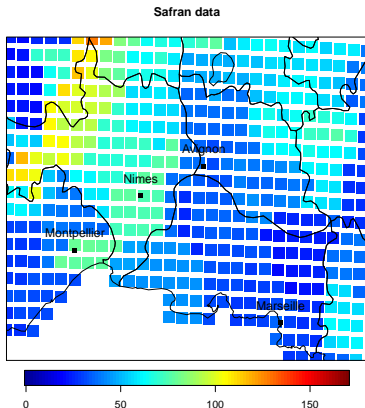
**Remind:** max-stable process = maximum over infinitely many “storms”

~> simulation difficult

## Algorithms for exact simulation:

- Dieker & Mikosch 2015
- Dombry, Engelke & O. 2016
- O., Schlather & Zhou 2018

# Example: Autumnal Rainfall in Southern France

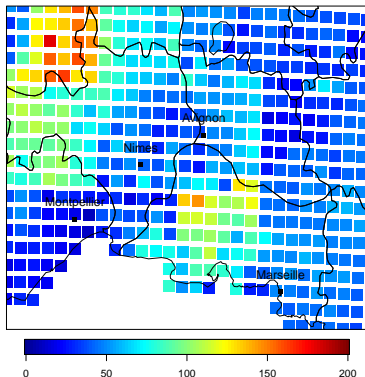


Source: O., Bel & Lantuéjoul 2018

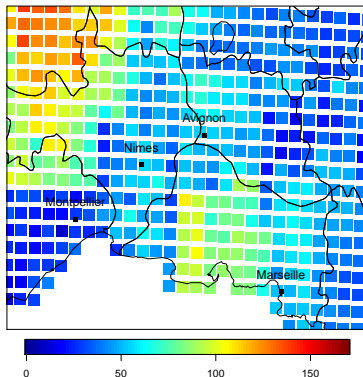


# Example: Autumnal Rainfall in Southern France

Simulation 20



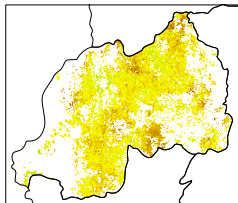
Simulation 42



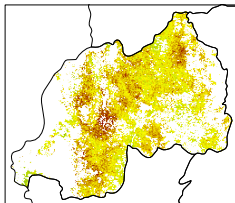
Source: O., Bel & Lantuéjoul 2018

# Example: Drought in Rwanda

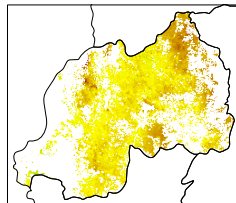
**Simulated NDVI at Season 1**



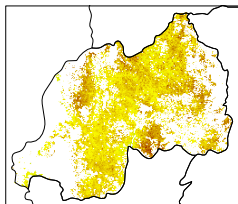
**Simulated NDVI at Season 2**



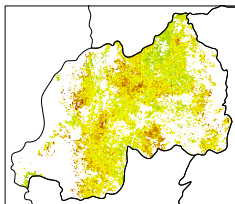
**Simulated NDVI at Season 3**



**Simulated NDVI at Season 4**



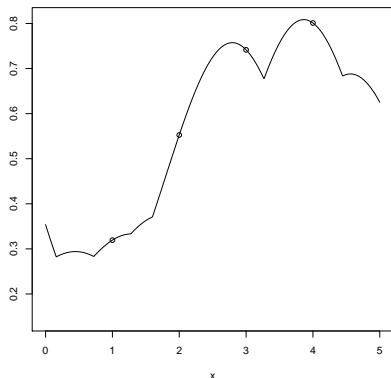
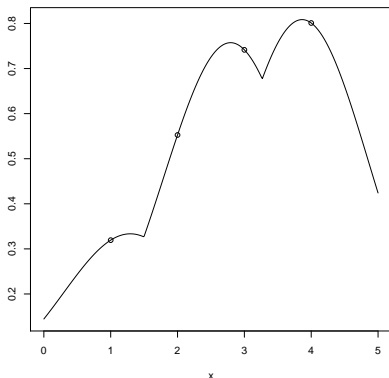
**Simulated NDVI at Season 5**



Source: O. & Stein 2018

# Prediction & Interpolation

- no analogue to Kriging for max-stable processes
- conditional simulations more difficult, but feasible for small number of conditions (Dombry, Eyi-Minko & Ribatet 2013)



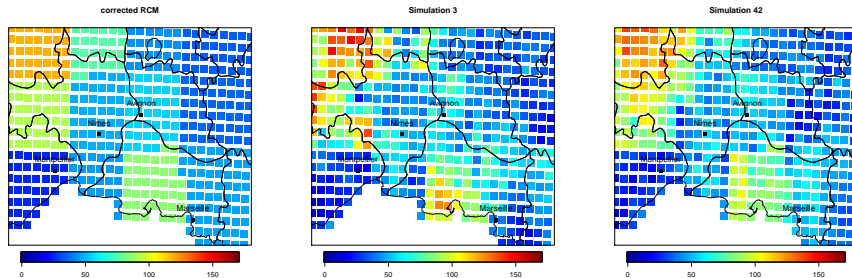
Source: Dombry, O. & Ribatet 2016

# Conditional Simulation & Downscaling

also other conditions possible, e.g. average over large grid cell

→ “downscaling”

## Example: Rainfall in Southern France



Source: O., Bel & Lantuéjoul 2018

# Conclusion

- probabilistic spatial models include distributions for each site (with spatially varying parameters) and spatial dependence structure
- type of model depends on variable of interest and situation, e.g. Gaussian processes for “normal” data and max-stable processes for “extreme” data
- flexible models for a wide range of applications are available
- simulations enable us to go beyond the data
- conditional simulations allow for interpolation and downscaling

# Conclusion

- probabilistic spatial models include distributions for each site (with spatially varying parameters) and spatial dependence structure
- type of model depends on variable of interest and situation, e.g. Gaussian processes for “normal” data and max-stable processes for “extreme” data
- flexible models for a wide range of applications are available
- simulations enable us to go beyond the data
- conditional simulations allow for interpolation and downscaling

## Caution:

- Risk of model misspecification!
- High uncertainties in extreme events!

# References I



A. B. Dieker and T. Mikosch.

Exact simulation of Brown-Resnick random fields at a finite number of locations.  
*Extremes*, 18(2):301–314, 2015.



C. Dombry, S. Engelke, and M. Oesting.

Exact simulation of max-stable processes.  
*Biometrika*, 103(2):303–317, 2016.



C. Dombry, S. Engelke, and M. Oesting.

Bayesian inference for multivariate extreme value distributions.  
*Electron. J. Stat.*, 11(2):4813–4844, 2017.



C. Dombry, F. Éyi-Minko, and M. Ribatet.

Conditional simulation of max-stable processes.  
*Biometrika*, 100(1):111–124, 2013.



C. Dombry, M. Oesting, and M. Ribatet.

Conditional simulation of max-stable processes.  
In D. K. Dey and J. Yan, editors, *Extreme Value Modeling and Risk Analysis*, pages 215–238. CRC Press, 2016.



R. Huser, C. Dombry, M. Ribatet, and M. G. Genton.

Full likelihood inference for max-stable data.  
*Stat*, 8(1):e218, 2019.



Z. Kabluchko, M. Schlather, and L. de Haan.

Stationary max-stable fields associated to negative definite functions.  
*Ann. Probab.*, 37(5):2042–2065, 2009.



M. Oesting, L. Bel, and C. Lantuéjoul.

Sampling from a max-stable process conditional on a homogeneous functional with an appl. for downscaling climate data.  
*Scand. J. Stat.*, 45(2):382–404, 2018.

# References II



M. Oesting, M. Schlather, and P. Friederichs.

Statistical post-processing of forecasts for extremes using bivariate Brown–Resnick processes with an appl. to wind gusts.  
*Extremes*, 20(2):309–332, 2017.



M. Oesting, M. Schlather, and C. Zhou.

Exact and fast simulation of max-stable processes on a compact set using the normalized spectral representation.  
*Bernoulli*, 24(2):1497–1530, 2018.



M. Oesting and A. Stein.

Spatial modeling of drought events using max-stable processes.  
*Stoch. Env. Res. Risk A.*, 32(1):63–81, 2018.



T. Opitz.

Extremal t processes: Elliptical domain of attraction and a spectral representation.  
*J. Multivar. Anal.*, 122:409–413, 2013.



M. Schlather.

Models for stationary max-stable random fields.  
*Extremes*, 5(1):33–44, 2002.



M. Schlather and J. A. Tawn.

A dependence measure for multivariate and spatial extreme values: Properties and inference.  
*Biometrika*, 90(1):139–156, 2003.



R. L. Smith.

Max–stable processes and spatial extremes.  
Unpublished Manuscript, 1990.