Conditional Modeling of Extreme Wind Gusts by Bivariate Brown-Resnick Processes

Marco Oesting

INRA / AgroParisTech

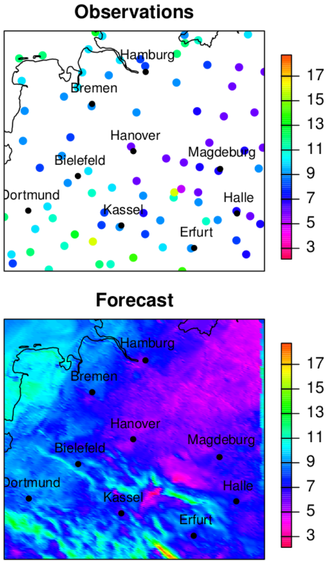
Joint Work with Martin Schlather (University of Mannheim)
and Petra Friederichs (University of Bonn)

Workshop on Stochastic Weather Generators
September 17, 2014, Avignon

This work has been supported by Volkswagen Stiftung within the project 'WEX-MOP'.

Extreme Wind Gusts

- wind gusts are strongly varying in space
- high uncertainty in forecasts, particularly for extreme wind gusts



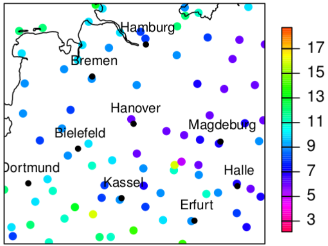
Extreme Wind Gusts

- wind gusts are strongly varying in space
- high uncertainty in forecasts, particularly for extreme wind gusts

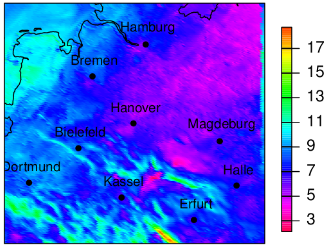
Goal:

Model for the (observed) wind gusts $V_{\max}^{\text{obs}}$ conditional on the forecast $V_{\max}^{\text{pred}}$

Observations



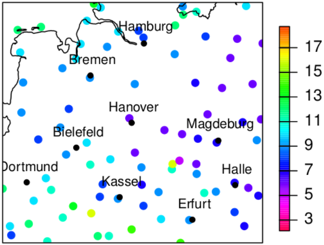
Forecast



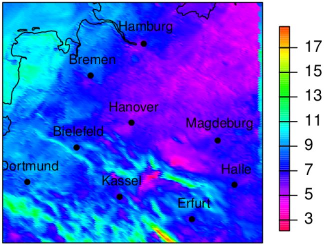
Extreme Wind Gusts

- wind gusts are strongly varying in space
- high uncertainty in forecasts, particularly for extreme wind gusts

Observations



Forecast



Goal:

Model for the (observed) wind gusts $V_{\max}^{\mathrm{obs}}$ conditional on the forecast $V_{\max}^{\mathrm{pred}}$

Two Modelling Steps:

1. model for marginal distributions (at single location) of $V_{\max}^{\mathrm{obs}}$ & $V_{\max}^{\mathrm{pred}}$
2. dependence model
   $\rightsquigarrow$ bivariate stochastic process

Outline

Outline

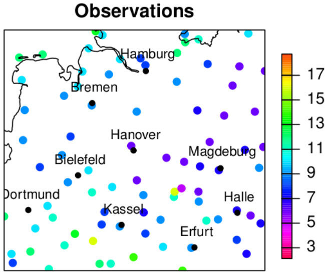
1. Marginal Model

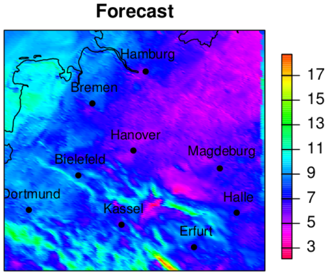
2. Dependence Model
   - Brown-Resnick Processes
   - Application to Data

The Data

Observation data:
for the maximal wind speed $\rightsquigarrow$ $V_{\text{max}}^{\text{obs}}$

- at 110 DWD stations in Northern Germany
- for 358 days (03/2011 – 02/2012)





The Data

Observation data:

for the maximal wind speed $\rightsquigarrow V_{\text{max}}^{\text{obs}}$

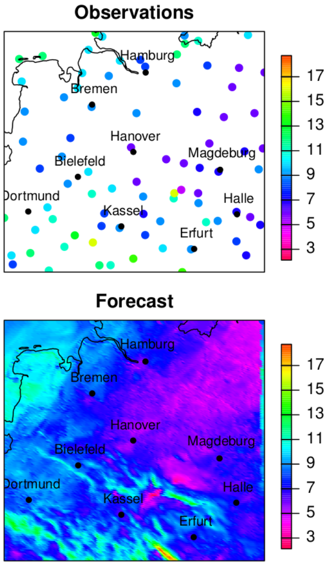
- at 110 DWD stations in Northern Germany
- for 358 days (03/2011 – 02/2012)

Forecast data:

from COSMO-DE EPS

- on a grid with mesh size 2.8 km covering Germany
- 20 ensemble members

1. for the maximal wind speed $\rightsquigarrow V_{\text{max}}^{\text{pred}}$
2. for the mean wind speed



Marginal Model for Extreme Wind Gusts

(single) **wind speed** $V(l, d)$ at location l and day d:

$$V(l, d) =_d \sqrt{\mathrm{Var}\{V(l, d)\}}\, V_0 + \mathbb{E}\{V(l, d)\}$$

with V_0 following some standardized distribution

Marginal Model for Extreme Wind Gusts

(single) **wind speed** $V(l,d)$ at location l and day d:

$$V(l,d) =_d \sqrt{\mathrm{Var}\{V(l,d)\}}\, V_0 + \mathbb{E}\{V(l,d)\}$$

with V_0 following some standardized distribution

$\mathbb{E}\{V(l,d)\}$ and $\mathrm{Var}\{V(l,d)\}$:

- general weather conditions
- estimated from forecast

Marginal Model for Extreme Wind Gusts

(single) **wind speed** $V(l, d)$ at location l and day d:

$$V(l,d) =_d \sqrt{\mathrm{Var}\{V(l,d)\}}\,V_0 + \mathbb{E}\{V(l,d)\}$$

with V_0 following some standardized distribution

maximal wind speed $V_{\max}(l, d)$ at location l and day d:

$$\mathbb{P}\left(\frac{V_{\max}(l,d) - \mathbb{E}\{V(l,d)\}}{\sqrt{\mathrm{Var}\{V(l,d)\}}} \leq x\right) \approx \exp\left(-\left(1 + \xi\frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right)$$

$\mathbb{E}\{V(l,d)\}$ and $\mathrm{Var}\{V(l,d)\}$:

- general weather conditions
- estimated from forecast

Marginal Model for Extreme Wind Gusts

(single) **wind speed** $V(l,d)$ at location l and day d:

$$V(l,d) =_d \sqrt{\mathrm{Var}\{V(l,d)\}}V_0 + \mathbb{E}\{V(l,d)\}$$

with V_0 following some standardized distribution

maximal wind speed $V_{\max}(l,d)$ at location l and day d:

$$\mathbb{P}\left(\frac{V_{\max}(l,d) - \mathbb{E}\{V(l,d)\}}{\sqrt{\mathrm{Var}\{V(l,d)\}}} \leq x\right) \approx \exp\left(-\left(1+\xi\frac{x-\mu}{\sigma}\right)_+^{-1/\xi}\right)$$

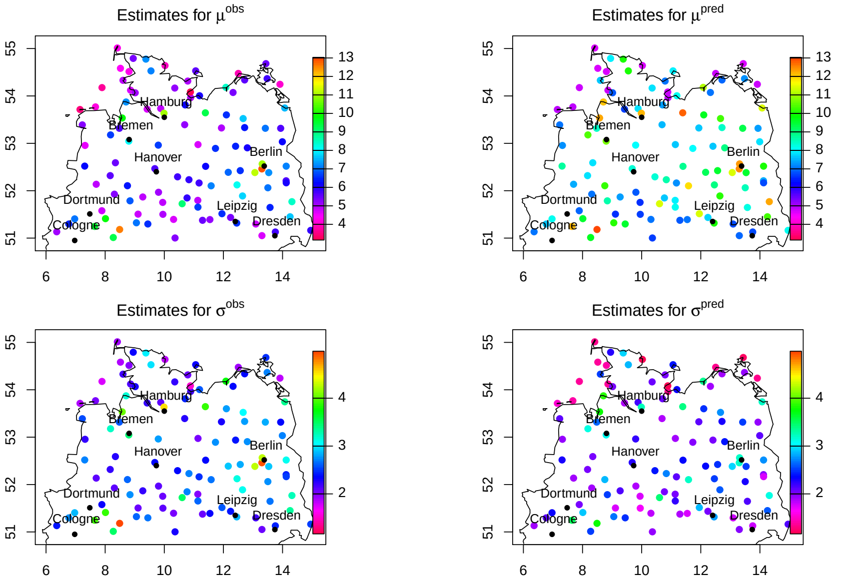
$\mathbb{E}\{V(l,d)\}$ and $\mathrm{Var}\{V(l,d)\}$:

- general weather conditions
- estimated from forecast

GEV parameters (ξ, μ, σ):

- ξ constant in space in time
- μ and σ vary spatially
- estimated via ML

GEV Parameters



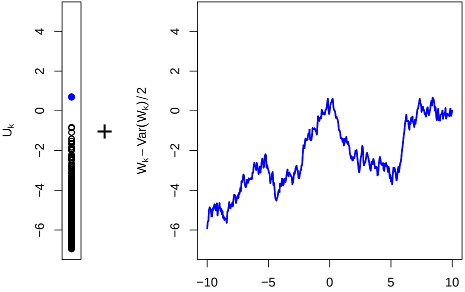
Outline

Spectral Representation

(Brown & Resnick 1977, Kabluchko, Schlather & de Haan 2009)

- $\{U_k\}_{k\in\mathbb{N}}$: PPP on $\mathbb{R}$ with intensity measure $e^{-u}\,\mathrm{d}u$ (magnitude)
- $W(\cdot)$: centered Gaussian process with stationary increments

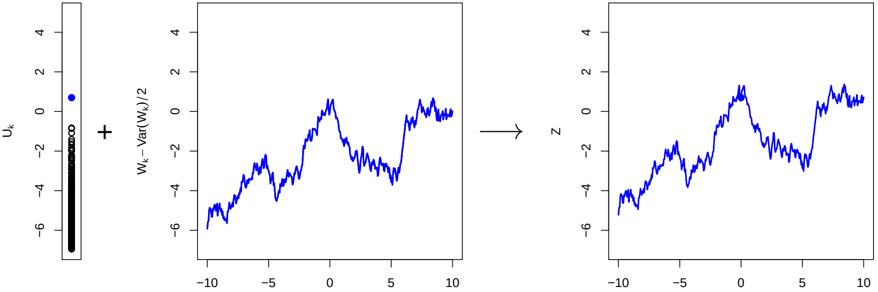
$$\left(W_k(\cdot) - \frac{\mathrm{Var}(W_k(\cdot))}{2}\right) \sim_{iid} \left(W(\cdot) - \frac{\mathrm{Var}(W(\cdot))}{2}\right) \quad \text{(spatial profile)}$$



Spectral Representation

(Brown & Resnick 1977, Kabluchko, Schlather & de Haan 2009)

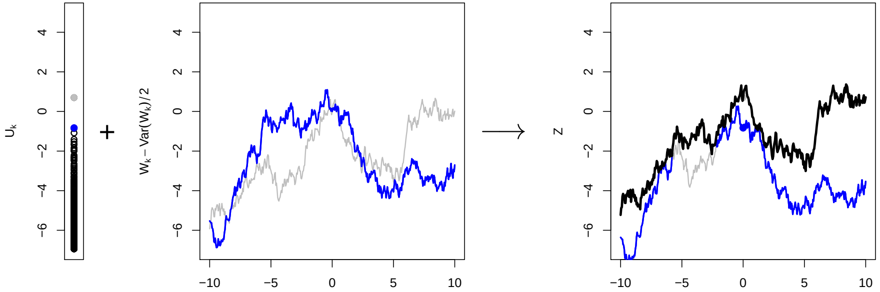
$$Z(x) = \max_{k \in \mathbb{N}} \left(U_k + W_k(x) - \frac{\mathrm{Var}(W_k(x))}{2} \right), \quad x \in \mathbb{R}^d$$



Spectral Representation

(Brown & Resnick 1977, Kabluchko, Schlather & de Haan 2009)

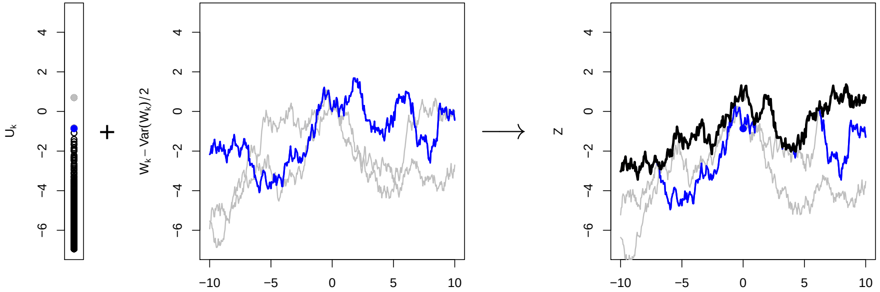
$$Z(x) = \max_{k \in \mathbb{N}} \left(U_k + W_k(x) - \frac{\mathrm{Var}(W_k(x))}{2} \right), \quad x \in \mathbb{R}^d$$



Spectral Representation

(Brown & Resnick 1977, Kabluchko, Schlather & de Haan 2009)

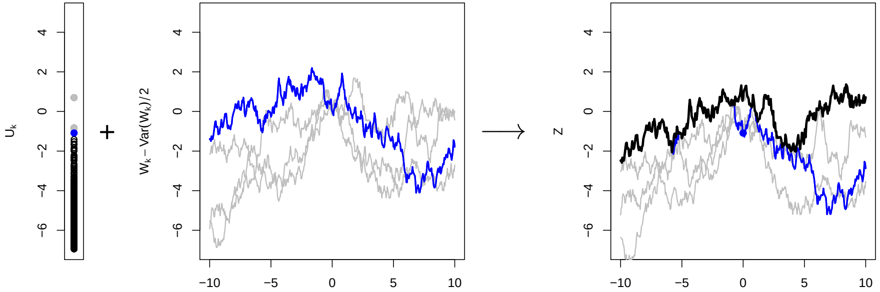
$$Z(x) = \max_{k \in \mathbb{N}} \left(U_k + W_k(x) - \frac{\operatorname{Var}(W_k(x))}{2} \right), \quad x \in \mathbb{R}^d$$



Spectral Representation

(Brown & Resnick 1977, Kabluchko, Schlather & de Haan 2009)

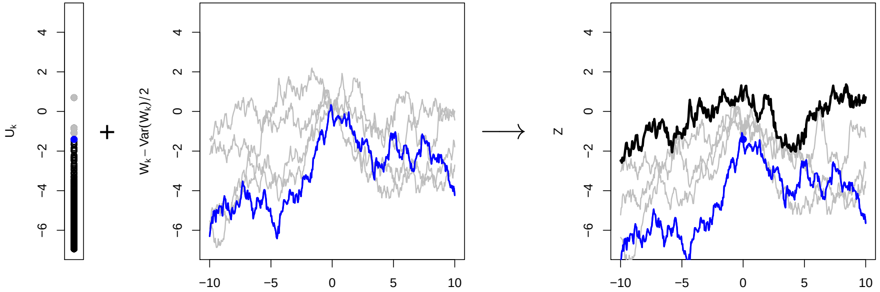
$$Z(x) = \max_{k \in \mathbb{N}} \left(U_k + W_k(x) - \frac{\mathrm{Var}(W_k(x))}{2} \right), \quad x \in \mathbb{R}^d$$



Spectral Representation

(Brown & Resnick 1977, Kabluchko, Schlather & de Haan 2009)

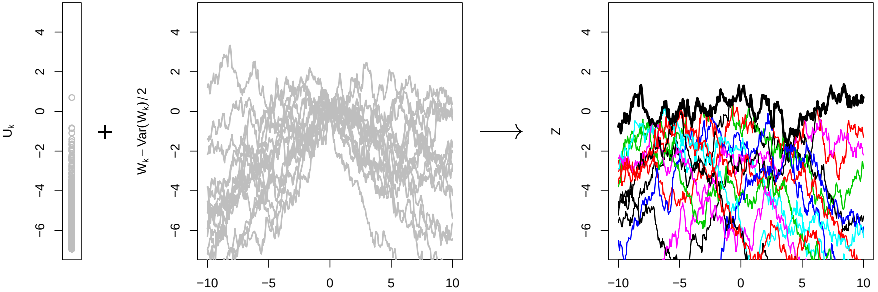
$$Z(x) = \max_{k \in \mathbb{N}} \left(U_k + W_k(x) - \frac{\mathrm{Var}(W_k(x))}{2} \right), \quad x \in \mathbb{R}^d$$



Spectral Representation

(Brown & Resnick 1977, Kabluchko, Schlather & de Haan 2009)

$$Z(x) = \max_{k \in \mathbb{N}} \left(U_k + W_k(x) - \frac{\mathrm{Var}(W_k(x))}{2} \right), \quad x \in \mathbb{R}^d$$



Z is max-stable and stationary!

Multivariate Generalization

(cf. Stucki & Molchanov, 2013, O. et al., 2013)

- $\{U_k\}_{k\in\mathbb{N}}$: Poisson point process with intensity $e^{-u}\,\mathrm{d}u$
- $W(\cdot)$: centered Gaussian process s.t.

variogram

$$\gamma(x_1, x_2) = (\mathrm{Var}(W \quad (x_1) - W \quad (x_2)))$$

depends on $x_1 - x_2 \in \mathbb{R}^d$ only

- $\{W_k(\cdot)\}_{k\in\mathbb{N}}$: independent copies of W

$$Z \quad (x) = \max_{k\in\mathbb{N}} \left(U_k + W_k \quad (x) - \frac{\mathrm{Var}(W_k \quad (x))}{2} \right), \quad x \in \mathbb{R}^d,$$

Multivariate Generalization

(cf. Stucki & Molchanov, 2013, O. et al., 2013)

- $\{U_k\}_{k\in\mathbb{N}}$: Poisson point process with intensity $e^{-u}\,\mathrm{d}u$
- $W(\cdot) = \left(W^{(1)}(\cdot), W^{(2)}(\cdot)\right)$: centered Gaussian process s.t.

pseudo-variogram

$$\gamma(x_1, x_2) = (\mathrm{Var}(W^{(i)}(x_1) - W^{(j)}(x_2)))_{i,j=1,2}$$

depends on $x_1 - x_2 \in \mathbb{R}^d$ only

- $\{W_k(\cdot)\}_{k\in\mathbb{N}}$: independent copies of W

$$Z^{(i)}(x) = \max_{k\in\mathbb{N}} \left(U_k + W_k^{(i)}(x) - \frac{\mathrm{Var}(W_k^{(i)}(x))}{2} \right), \quad x \in \mathbb{R}^d, \; i = 1, 2.$$

Multivariate Generalization

(cf. Stucki & Molchanov, 2013, O. et al., 2013)

- $\{U_k\}_{k\in\mathbb{N}}$: Poisson point process with intensity $e^{-u}\,\mathrm{d}u$
- $W(\cdot) = \left(W^{(1)}(\cdot), W^{(2)}(\cdot)\right)$: centered Gaussian process s.t.

pseudo-variogram

$$\gamma(x_1, x_2) = (\mathrm{Var}(W^{(i)}(x_1) - W^{(j)}(x_2)))_{i,j=1,2}$$

depends on $x_1 - x_2 \in \mathbb{R}^d$ only

- $\{W_k(\cdot)\}_{k\in\mathbb{N}}$: independent copies of W

$$Z^{(i)}(x) = \max_{k\in\mathbb{N}} \left(U_k + W_k^{(i)}(x) - \frac{\mathrm{Var}(W_k^{(i)}(x))}{2} \right), \quad x \in \mathbb{R}^d, \ i = 1, 2.$$

Then,

- Z is max-stable and stationary (as bivariate process)
- law of Z depends on γ only

How to construct a valid pseudo-variogram?

Necessary condition:
(essentially) same behaviour of all components $\gamma_{ij}(h)$ as $\|h\| \to \infty$

Construction principle:

- $Y(\cdot)$: univariate Gaussian process with stationary increments and variogram γ^*
- $V(\cdot) = (V^{(1)}(\cdot), V^{(2)}(\cdot))$: bivariate stationary Gaussian process with covariance function

$$C(h) = \begin{pmatrix} C_{11}(h) & C_{12}(h) \\ C_{21}(h) & C_{22}(h) \end{pmatrix}$$

$W(\cdot) = (Y(\cdot) + V^{(1)}(\cdot), Y(\cdot) + V^{(2)}(\cdot))$ has a pseudo-variogram

$$\gamma(h) = \left(\gamma^*(h) + C_{ii}(0) + C_{jj}(0) - 2C_{ij}(h)\right)_{i,j=1,2}.$$

Outline

Full model

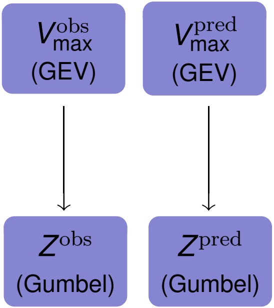
- **Reminder:**
  marginals of $V_{\text{max}}^{\text{obs}}$ and $V_{\text{max}}^{\text{pred}}$ are modelled by GEVs (parameters estimated via ML)

$V_{\text{max}}^{\text{obs}}$ (GEV)

$V_{\text{max}}^{\text{pred}}$ (GEV)

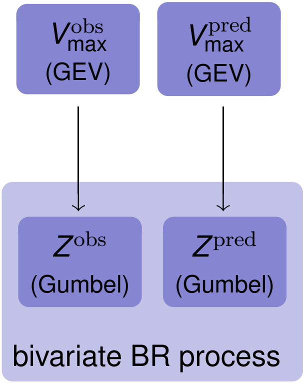
Full model

- **Reminder:**
  marginals of $V_{\text{max}}^{\text{obs}}$ and $V_{\text{max}}^{\text{pred}}$ are modelled by GEVs (parameters estimated via ML)
- data are transformed to standard Gumbel margins ($\rightsquigarrow Z^{\text{obs}}, Z^{\text{pred}}$)



Full model

- **Reminder:**
  marginals of $V_{\text{max}}^{\text{obs}}$ and $V_{\text{max}}^{\text{pred}}$ are modelled by GEVs (parameters estimated via ML)
- data are transformed to standard Gumbel margins ($\rightsquigarrow Z^{\text{obs}}, Z^{\text{pred}}$)
- standardized observation and forecast are jointly modelled by bivariate BR process
  - dependence in space
  - dependence between observations and forecast

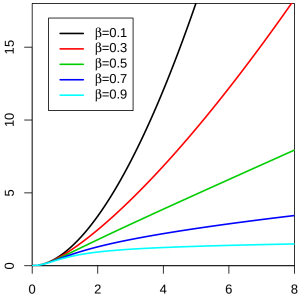


Bivariate Pseudo-Variogram Model

$$\gamma(h) = \left(\gamma^*(h) + C_{ii}(0) + C_{jj}(0) - 2C_{ij}(h)\right)_{i,j=1,2}.$$

- γ^*: variogram of power law type

$$\gamma^*(h) = \frac{\|h\|^2}{(1+\|h\|^2)^\beta}, \quad \beta \in (0,1)$$



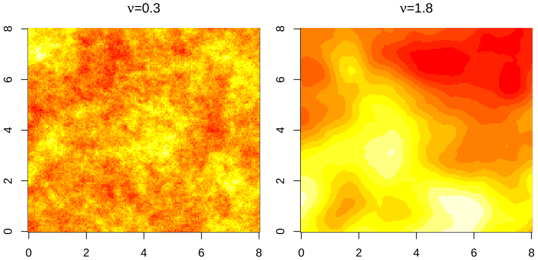
Bivariate Pseudo-Variogram Model

$$\gamma(h) = \left(\gamma^*(h) + C_{ii}(0) + C_{jj}(0) - 2C_{ij}(h)\right)_{i,j=1,2}.$$

- C: bivariate Matérn model (Gneiting et. al., 2010)

$$C_{ij}(h) = \rho_{ij}\sigma_i\sigma_j \frac{2^{1-\nu_{ij}}}{\Gamma(\nu_{ij})}(a\|h\|)^{\nu_{ij}} K_{\nu_{ij}}(a\|h\|)$$

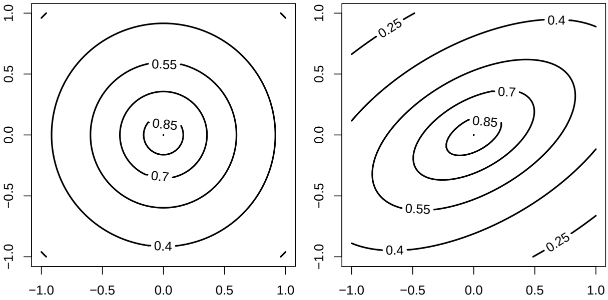
for $\sigma_1, \sigma_2 \geq 0$, $a, \nu_{11}, \nu_{22} > 0$, $\nu_{12} = (\nu_{11} + \nu_{22})/2$ and suitable ρ_{ij}



Bivariate Pseudo-Variogram Model

$$\gamma(h) = \left(\gamma^*(h) + C_{ii}(0) + C_{jj}(0) - 2C_{ij}(h)\right)_{i,j=1,2}.$$

- γ^*: $\gamma^*(h) = \|Ah\|^2 \Big/ (1 + \|Ah\|^2)^\beta, \quad \beta \in (0,1)$
- C: $C_{ij}(h) = \rho_{ij}\sigma_i\sigma_j 2^{1-\nu_{ij}}\Gamma(\nu_{ij})^{-1}\left(a\|Ah\|\right)^{\nu_{ij}}\mathcal{K}_{\nu_{ij}}\left(a\|Ah\|\right)$
- introduce anisotropy matrix A (dilation/rotation)

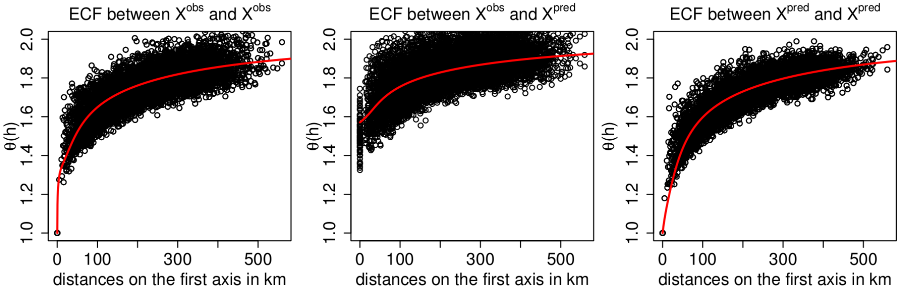


Extremal Coefficient Function

(cf. Schlather and Tawn 2003, Cooley, Naveau, Poncet 2009)

ECF between Z^{obs} and Z^{obs}

$$\mathbb{P}(Z^{\mathrm{obs}}(h) \leq z, Z^{\mathrm{obs}}(0) \leq z) = \mathbb{P}(Z^{\mathrm{obs}}(0) \leq z)^{\theta(h)}$$

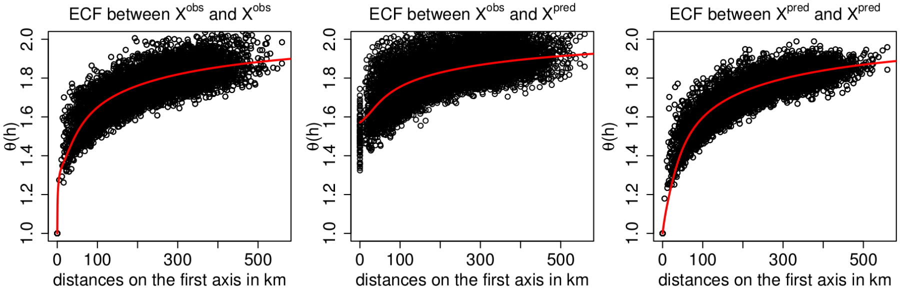


Extremal Coefficient Function

(cf. Schlather and Tawn 2003, Cooley, Naveau, Poncet 2009)

ECF between Z^{obs} and Z^{pred}

$$\mathbb{P}(Z^{\mathrm{obs}}(h) \leq z, Z^{\mathrm{pred}}(0) \leq z) = \mathbb{P}(Z^{\mathrm{pred}}(0) \leq z)^{\theta(h)}$$

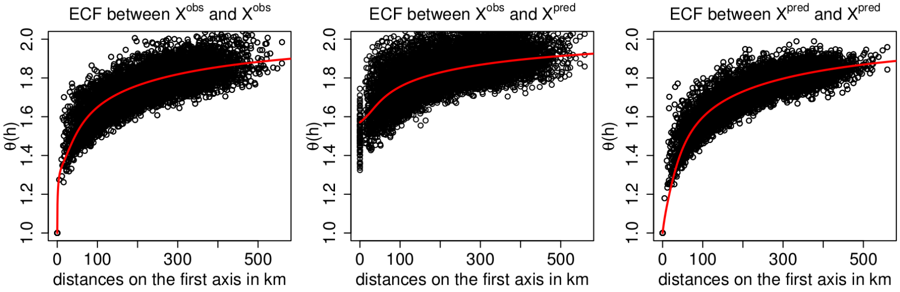


Extremal Coefficient Function

(cf. Schlather and Tawn 2003, Cooley, Naveau, Poncet 2009)

ECF between Z^{pred} and Z^{pred}

$$\mathbb{P}(Z^{\mathrm{pred}}(h) \leq z,\, Z^{\mathrm{pred}}(0) \leq z) = \mathbb{P}(Z^{\mathrm{pred}}(0) \leq z)^{\theta(h)}$$

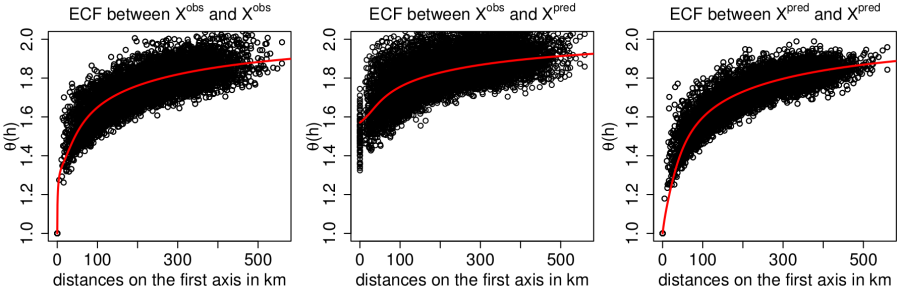


Extremal Coefficient Function

(cf. Schlather and Tawn 2003, Cooley, Naveau, Poncet 2009)

ECF between Z^{pred} and Z^{pred}

$$\mathbb{P}(Z^{\mathrm{pred}}(h) \leq z,\, Z^{\mathrm{pred}}(0) \leq z) = \mathbb{P}(Z^{\mathrm{pred}}(0) \leq z)^{\theta(h)}$$

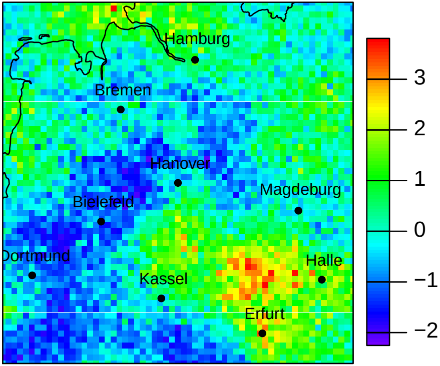


ECF for bivariate BR processes: $\theta_{ij}(h) = 2\Phi\left(\sqrt{\gamma_{ij}(h)}/2\right)$

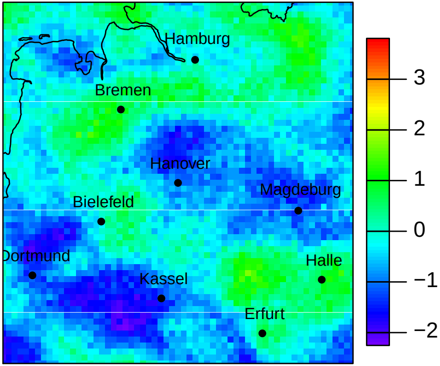
$\rightsquigarrow$ Fit ECF of bivariate BR process to estimated ECF!

Unconditional simulation of the Brown-Resnick process:

Realisation of X^{obs}



Realisation of X^{pred}



Post-Processing of the Forecast

Given Data

- historical observation & forecast data
- forecast for today

Post-Processing of the Forecast

Given Data

- historical observation & forecast data
  $\rightsquigarrow$ parameter estimation
- forecast for today

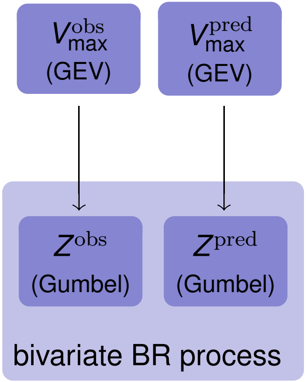
Post-Processing of the Forecast

Given Data

- historical observation & forecast data
  $\rightsquigarrow$ parameter estimation
- forecast for today

Post-Processing:

1. standardized forecast $V_{\max}^{\mathrm{pred}}$ to standard Gumbel margins $\rightsquigarrow Z^{\mathrm{pred}}$
2. simulate realizations of $Z^{\mathrm{obs}} \mid Z^{\mathrm{pred}}$
3. transform Z^{obs} from Gumbel to GEV margins $\rightsquigarrow V_{\max}^{\mathrm{obs}}$



Summary:

- bivariate Brown-Resnick processes model
  - dependence in space
  - dependence between forecast and observation
- flexible modelling via pseudo-variogram
- application to data shows promising results

Outlook/Open Questions:

- simulation of observations conditional on forecast
  $\rightsquigarrow$ conditional simulation of bivariate Brown-Resnick processes (Dombry, Eyi-Minko & Ribatet, 2013)
- downscaling (how to choose GEV parameters?)

B. M. Brown and S. I. Resnick.
Extreme values of independent stochastic processes.
J. Appl. Probab., 14(4):732–739, 1977.

D. Cooley, P. Naveau, and P. Poncet.
Variograms for spatial max-stable random fields.
In *Dependence in Probability and Statistics*, pages 373–390. Springer, New York, 2006.

L. de Haan.
A spectral representation for max-stable processes.
Ann. Probab., 12(4):1194–1204, 1984.

C. Dombry, F. Éyi-Minko, and M. Ribatet.
Conditional simulation of max-stable processes.
Biometrika, 100(1):111–124, 2013.

T. Gneiting, W. Kleiber, and M. Schlather.
Matérn cross-covariance functions for multivariate random fields.
J. Amer. Statist. Assoc., 105(491):1167–1177, 2010.

T. Gneiting and A. E. Raftery.
Strictly proper scoring rules, prediction, and estimation.
J. Amer. Statist. Assoc., 102(477):359–378, 2007.

Z. Kabluchko, M. Schlather, and L. de Haan.
Stationary max-stable fields associated to negative definite functions.
Ann. Probab., 37(5):2042–2065, 2009.

I. Molchanov and K. Stucki.
Stationarity of multivariate particle systems.
Stochastic Process. Appl., 123(6):2272–2285, 2013.

M. Oesting, M. Schlather, and P. Friederichs.
Conditional modelling of extreme wind gusts by bivariate brown-resnick processes.
Available from `http://arxiv.org/abs/1312.4584.`, 2013.

M. D. Penrose.
Semi-min-stable processes.
Ann. Probab., 20(3):1450, 1992.

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.

Verification

For every station, compare

“deviation” of COSMO-DE-EPS forecast from observation

and

“deviation” of post-processed forecast* from observation

*conditioned on forecast for the same station only

“Deviation” measure: Continuous Ranked Probability Score (CRPS)

Verification

For every station, compare

“deviation” of COSMO-DE-EPS forecast from observation

and

“deviation” of post-processed forecast* from observation

*conditioned on forecast for the same station only

“Deviation” measure: Continuous Ranked Probability Score (CRPS)

CRPS (cf. Gneiting & Raftery, 2007)

F: probability distribution (forecast) x: real value (observation)

$$\mathrm{CRPS}(F, x) = \int (F(y) - \mathbf{1}_{\{y \geq x\}})^2 \, \mathrm{d}y = \mathbb{E}|X - x| - \frac{1}{2}\mathbb{E}|X - X'|$$

for $X, X' \sim_{iid} F$ with $\mathbb{E}|X| < \infty$.

Verification

For every station, compare

“deviation” of COSMO-DE-EPS forecast from observation	and	“deviation” of post-processed forecast* from observation

*conditioned on forecast for the same station only

“Deviation” measure: Continuous Ranked Probability Score (CRPS)

Results:

For each of the 110 stations:

CRPS for post-processed forecast < CRPS for COSMO-DE-EPS,

on average by 13%.