Extreme Value Analysis Part II: Multivariate and Spatial Extremes

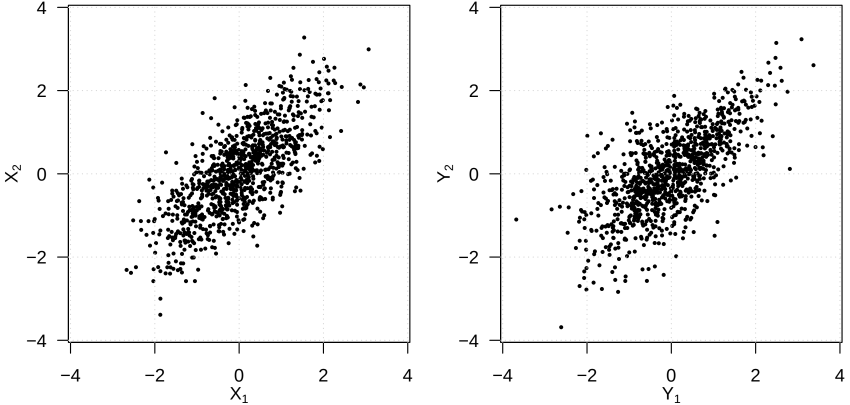
Whitney Huang and Brook Russell

May 24, 2022

Today's Meeting Agenda

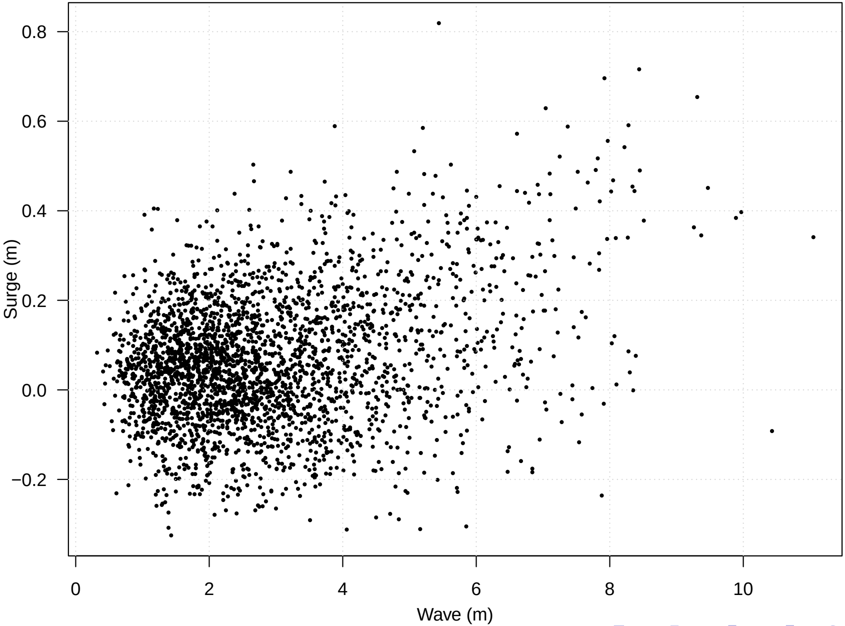
- Multivariate extremes
  - Describing tail dependence
  - Block Maxima Models
  - Threshold Exceedance Models
  - Conditional Extreme Value Models
- Spatial extremes
  - Weather vs. Climate
  - Bayesian Hierarchical Approaches
  - Max-Stable Process Models

Motivating Example

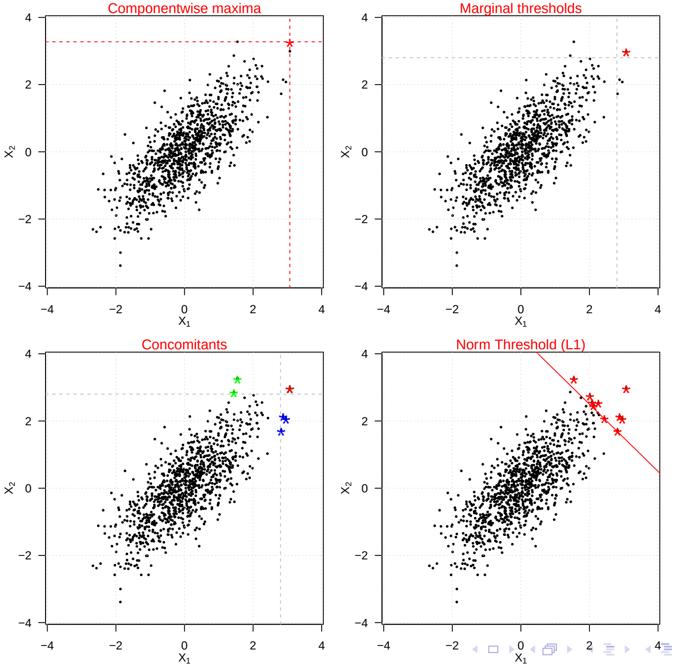


A central aim of multivariate extremes is to describe tail dependence

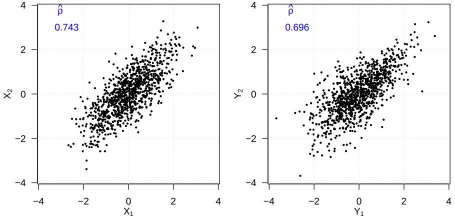
Example: Wave–Surge Data at Newlyn, UK



What is a Multivariate Extreme?



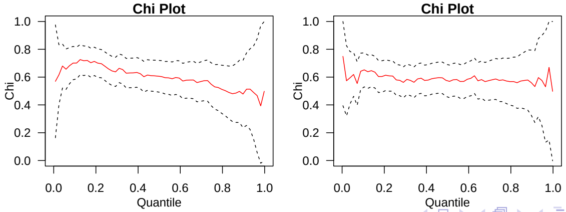
Measuring Tail Dependence



ρ measures "spread from center", does not focus on extremes

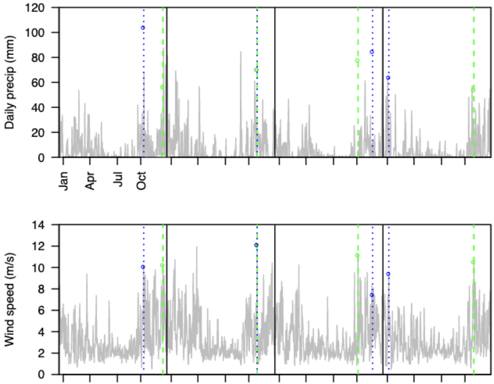
Upper tail dependence parameter:

$$\chi = \lim_{u \to 1} \chi(u) = \mathbb{P}\left(F_{X_1}(X_1) > u | F_{X_2}(X_2) > u\right)$$



Componentwise Maxima

Let $\{\boldsymbol{X}_i = (X_{1,i}, \cdots, X_{d,i})\} \in \mathbb{R}^d$ are i.i.d. random vectors. Let's first consider **componentwise maxima**:
$\boldsymbol{M}_n = (M_{1,n}, \cdots, M_{d,n})^T$, where $M_{j,n} = \max_{i=1}^n X_{j,i}$



- Univariate GEV still apply in each margin
- Need to model the interdependence across $\{M_{j,n}\}_{j=1}^d$
- $\boldsymbol{M}_n$ is not necessarily one the the original observations

Multivariate Extreme Value Theorem

Let $\boldsymbol{X}_1, \cdots, \boldsymbol{X_n} \overset{\text{iid}}{\sim} \boldsymbol{F}$ where $\boldsymbol{X}_i \in \mathbb{R}^d$, if

$$\frac{\max_{i \leq i \leq n} \mathbf{X}_i - \mathbf{b}_n}{\mathbf{a}_n} = \left(\frac{\bigvee_{i=1}^n X_{1,i} - b_{1,n}}{a_{1,n}}, \cdots, \frac{\bigvee_{i=1}^n X_{d,i} - b_{d,n}}{a_{d,n}} \right) \xrightarrow{d} \mathbf{G}$$

then $\mathbf{G}$ must be a multivariate extreme value distribution

- Each marginal (approximately) follows a GEV distribution (i.e., $M_j \approx \mathrm{GEV}(\mu_j, \sigma_j, \xi_j)$)
- No limiting parametric family exists for describing extremal dependence, i.e., the interdependence across $\{M_{j,n}\}_{j=1}^d$
- Marginals and dependence are typically handled separately $\Rightarrow$ "Copula-like" approach

Multivariate GEV

- Transform each marginal to a common marginal distribution, e.g., unit Fréchet $(\mathrm{GEV}(1,1,1))$

$$\tilde{M}_j = \left[1 + \xi_j \left(\frac{M_j - \mu_j}{\sigma_j}\right)\right]^{\frac{1}{\xi_j}}$$

with $\tilde{G}(m) = \exp(m^{-1})$

- Then

$$\mathbb{P}(\tilde{M}_1 \leq \tilde{m}_1, \cdots, \tilde{M}_d \leq \tilde{m}_d) = \exp\left(-V\left(\tilde{m}_1, \cdots, \tilde{m}_d\right)\right),$$

where $V : \mathbb{R}_+^d \mapsto \mathbb{R}_+$ is the exponent measure that characterizes the extremal dependence

- In practice, modeling usually involves fitting a parametric family of V

Exponent Measure for Bivariate Extremes ($d = 2$)

$$\mathbb{P}(M_1 \leq m_1, M_2 \leq m_2) = \mathbb{P}(\tilde{M}_1 \leq \tilde{m}_1, \tilde{M}_2 \leq \tilde{m}_2) = \exp(-V(\tilde{m}_1, \tilde{m}_2)),$$

where

$$V(m_1, m_2) = 2 \int_0^1 \max\left(\frac{\omega}{m_1}, \frac{1-\omega}{m_2}\right) dH(\omega)$$

and H is a distribution on $[0, 1]$ satisfying the mean constraint:

$$\int_0^1 \omega \, dH(\omega) = \frac{1}{2}$$

Example: If H places mass 0.5 on $\omega = 0$ and 1, then we get

$$\mathbb{P}(\tilde{M}_1 \leq m_1, \tilde{M}_2 \leq m_2) = \exp\left\{-\left(m_1^{-1} + m_2^{-1}\right)\right\},$$

corresponding to independent of M_1 and M_2

The Logistic Model

$$G(m_1, m_2) = \exp\left\{-\left(m_1^{-\frac{1}{\alpha}} + m_2^{-\frac{1}{\alpha}}\right)^{\alpha}\right\},$$

where m_1 and $m_2 > 0$, $\alpha \in (0, 1)$

- $\alpha \to 1$ corresponds to independent
- $\alpha \to 0$ corresponds to perfectly dependent
- This model is symmetric $\Rightarrow$ the variables are exchangeable, that is, M_1 depends on M_2 to exactly the same degree that M_2 depends on M_1

The Bilogistic Model

$$G(m_1, m_2) = \exp\left\{m_1\gamma^{1-\alpha} + m_2(1-\gamma)^{1-\beta}\right\},$$

where $0 < \alpha < 1$, $0 < \beta < 1$, and γ is the solution of:

$$(1-\alpha)m_1(1-\gamma)^{\beta} = (1-\beta)m_2\gamma^{\alpha}$$

- Independence is obtained when $\alpha = \beta \to 1$ and when one of α or β is fixed and the other approaches 1
- When $\alpha = \beta$ the model reduces to the logistic model
- The value of $\alpha - \beta$ determines the extent of asymmetry in the dependence structure

Parameter Estimation

Models (e.g., Logistic or Bilogistic) can be fitted by maximum likelihood estimation, either:

- two steps: marginal GEV components followed by dependence function, or
- one step that estimates marginals and dependence structure simultaneously

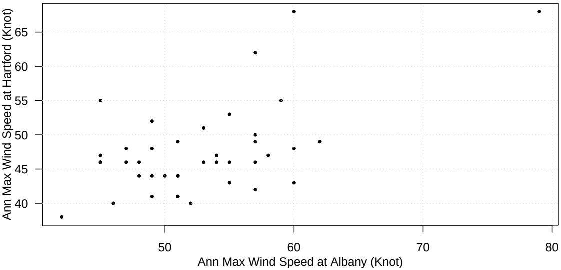
The probability density function is

$$g(m_1, m_2) = \{V_{m_1} V_{m_2} - V_{m_1 m_2}\} \exp(-V(m_1, m_2)),$$

where V_{m_1}, V_{m_2}, and $V_{m_1 m_2}$ denote partial derivatives of V.

With the (transformed) data $(\tilde{m}_{1,i}, \tilde{m}_{2,i})_{i=1}^{m}$, one can maximze the log-likelihood $\ell(\theta) = \sum_{i=1}^{m} \log g(\tilde{m}_{1,i}, \tilde{m}_{2,i})$ to obtain maximum likelihood estimates and standard errors

Annual Maximum Wind Speeds at Albany and Hartford



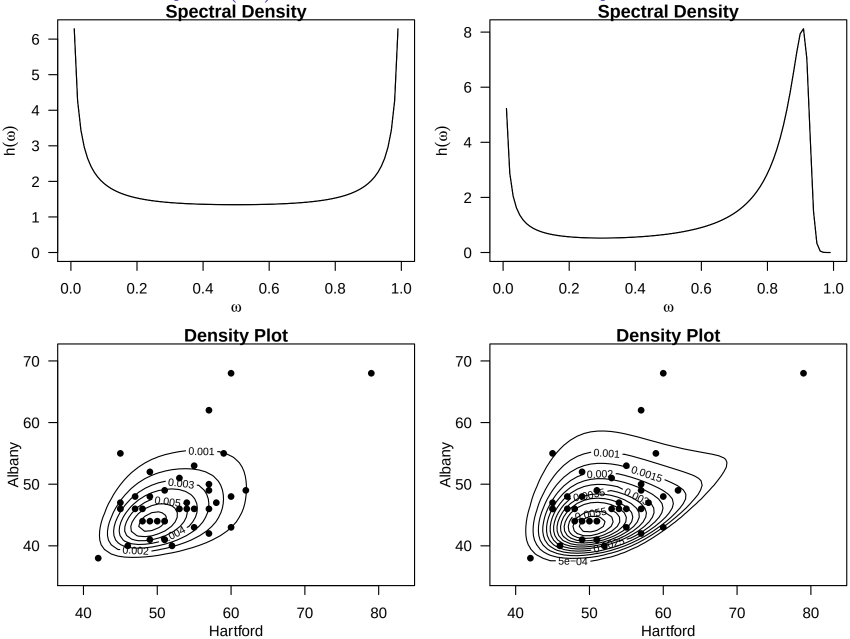
Marginal Parameter Estimates

Model	μ_1	σ_1	ξ_1	μ_2	σ_2	ξ_2
Logistic	49.97	5.04	0.01	44.58	4.34	0.08
Bilogisitc	50.08	5.07	-0.04	44.64	4.40	0.10

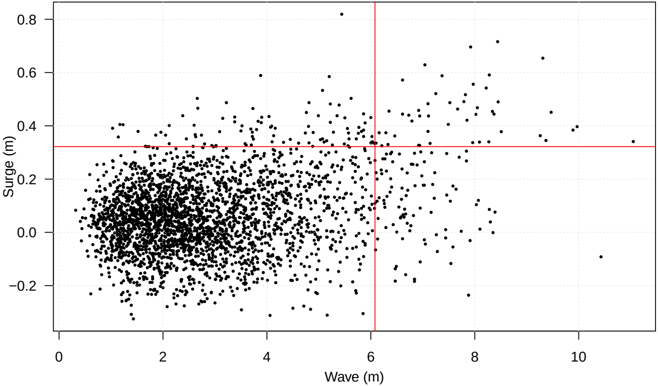
Dependence Parameter Estimates

Model	log-lik.	α	β
Logistic	-246.07	0.71	
Bilogisitc	-244.89	0.10	0.90

Spectral Density $h(\omega)$ and Bivariate Density Contour



Bivariate Threshold Excess Model

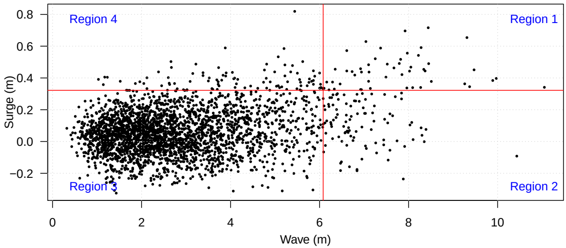


Main idea:

$$F(x_1, x_2) \approx G(x_1, x_2) = \exp(-V(\tilde{x}_1, \tilde{x}_2)),\ x_1 > u_{x_1}, x_2 > u_{x_2},$$

where marginal transformation (from GPD to unit Fréchet) can be achieved by probability integral transformation

Likelihood Calculations



$$g(x_1, x_2) = \begin{cases} \dfrac{\partial^2 \boldsymbol{G}}{\partial x_1 \partial x_2}\Big|_{(\tilde{x}_1, \tilde{x}_2)} & \text{if } (x_1, x_2) \in \text{Region 1;} \\ \dfrac{\partial \boldsymbol{G}}{\partial x_1}\Big|_{(\tilde{x}_1, \tilde{u}_{x_2})} & \text{if } (x_1, x_2) \in \text{Region 2;} \\ \dfrac{\partial \boldsymbol{G}}{\partial x_2}\Big|_{(\tilde{u}_{x_1}, \tilde{x}_2)} & \text{if } (x_1, x_2) \in \text{Region 3;} \\ \boldsymbol{G}(\tilde{u}_{x_1}, \tilde{u}_{x_2}) & \text{if } (x_1, x_2) \in \text{Region 4.} \end{cases}$$

Wave–Surge Analysis at Newlyn

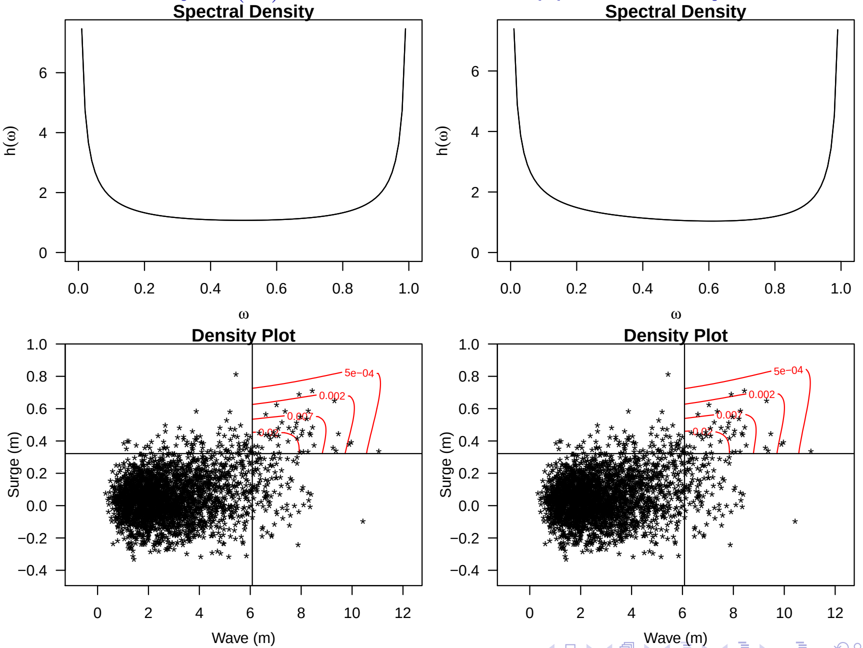
Marginal Parameter Estimates

Model	σ_{u_1}	ξ_1	σ_{u_2}	ξ_2
Logistic	1.26	-0.13	0.09	0.01
Bilogisitc	1.27	-0.14	0.09	0.01

Dependence Parameter Estimates

Model	log-lik.	α	β
Logistic	-1018.04	0.76	
Bilogisitc	-1017.90	0.79	0.73

Spectral Density $h(\omega)$ and Bivariate Upper Density Contour



Asymptotic Dependence and Independence

One key problem with using limit distributions for multivariate extremes is that they force one of two possibilities:

- Asymptotic Dependence (AD): extremes occur with a dependence structure which conforms to an extreme value distribution
  $\rightarrow \chi = \lim_{u \rightarrow 1} \mathbb{P}\left(F_{X_1}(X_1) > u | F_{X_2}(X_2) > u\right) > 0$
- Asymptotic Independence (AI): extremes occur independently in the different margins
  $\rightarrow \chi = \lim_{u \rightarrow 1} \mathbb{P}\left(F_{X_1}(X_1) > u | F_{X_2}(X_2) > u\right) = 0$

Often AI is suggested by the data, and yet quite strong dependence is still present at high levels (e.g., Gaussian copula with $\rho \neq 1$)

Conditional Extreme Value (CEV) Models [Heffernan & Tawn, 04]

Models the conditional distribution by assuming a **parametric** location-scale form after marginal transformation

- **Marginal modeling**:
  1. Estimate marginal distributions of Y and X
  2. Transform $(Y, X)^T$ to Laplace marginals $(\tilde{Y}, \tilde{X})^T$

- **Dependence modeling**:

  Assume for large u,

$$\left[\frac{\tilde{Y} - a\left(\tilde{X}\right)}{b\left(\tilde{X}\right)} \leq z | \tilde{X} > u\right] \sim G(z),$$

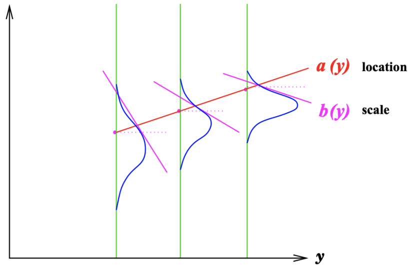
where $a(x) = \alpha x$ and $b(x) = x^{\beta}$, $\alpha \in [-1, 1]$, $\beta \in (-\infty, 1)$

A cartoon illustration of the CEV dependence modeling

Assume for large u,

$$\left[\frac{\tilde{Y} - a\left(\tilde{X}\right)}{b\left(\tilde{X}\right)} \leq z | \tilde{X} > u\right] \sim G(z),$$

where $a(x) = \alpha x$ and $b(x) = x^{\beta}$, $\alpha \in [-1, 1]$, $\beta \in (-\infty, 1)$

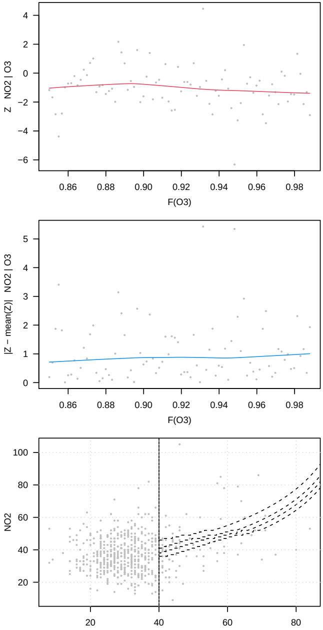


Source: Heffernan's slides given at the Interface 2008 Symposium

- $\tilde{Y} = \alpha\tilde{X} + \tilde{X}^{\beta} Z$,
  $\Rightarrow Z = \frac{\tilde{Y} - \alpha\tilde{X}}{\tilde{X}^{\beta}} \sim G$
- α and β are estimated by making a parametric assumption of $\tilde{Y}$
- G estimated nonparametrically

Modeling "Summer" O_3 and NO_2 [Heffernan & Tawn, 2004]

- Apr.-July, 1994-1998 daily maximum air pollution data from Leeds, U.K.
- Need to choice several thresholds: one for each marginal and a threshold for the dependence structure
- $\tilde{Y} = \alpha \tilde{X} + \tilde{X}^{\beta} Z,$
  $\Rightarrow Z = \frac{\tilde{Y} - \alpha \tilde{X}}{\tilde{X}^{\beta}} \sim G$
  $\hat{\alpha} = 0.77,\ \hat{\beta} = 0.188$

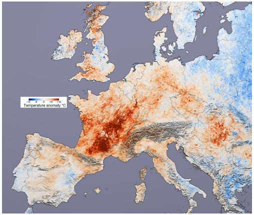


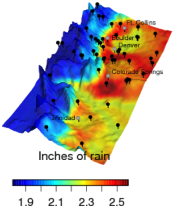
Summary of Multivariate Extreme Value Analysis

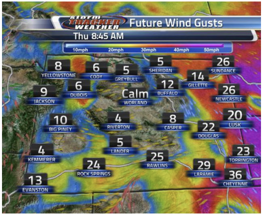
- Definition of a multivariate extreme is not obvious
- Tail dependence NOT summarized with correlations, we looked at the extremal coefficient
- Methodologies exist for both block maxima and threshold exceedance approaches
- Conditional extreme value models for asymptotic independence cases

Spatial Extremes: Why Spatial Matters?

- Quantity of interest often has a (continuous) spatial domain







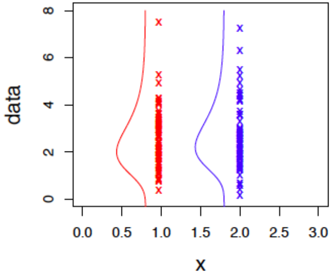
- "Everything is related to everything else, but near things are more related than distant things." – Waldo Tobler

- Spatial interpolation of:
  - **Climate (marginal)**: model how tail distributions vary in space, e.g., return level maps
  - **Weather (dependence)**: fill in extreme values in space

Main Objectives

- Marginal modeling: model the site–wise behavior of extremes by using extreme value distribution (i.e. GEV, GPD)

- Dependence modeling: model the spatial dependence of the extreme values

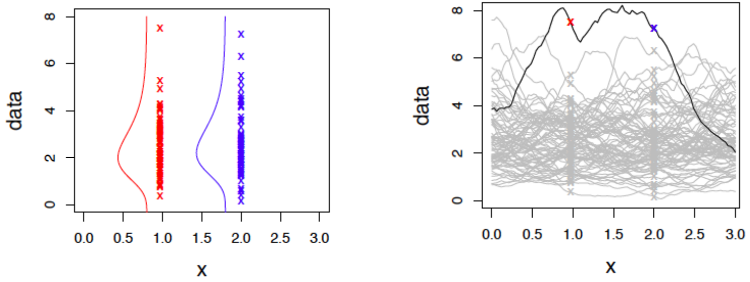


Figures courtesy of Jenny Wadsworth

Main Objectives

- Marginal modeling: model the site–wise behavior of extremes by using extreme value distribution (i.e. GEV, GPD)

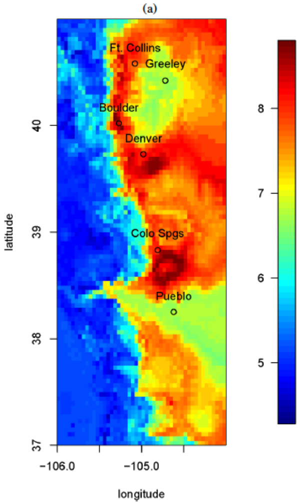
- Dependence modeling: model the spatial dependence of the extreme values



Figures courtesy of Jenny Wadsworth

Estimating Return Level Maps

- $M(\mathbf{s}) \approx \mathrm{GEV}(\mu(\mathbf{s}), \sigma(\mathbf{s}), \xi(\mathbf{s}))$, $\mathbf{s} \in \mathcal{S}$ where $M(\mathbf{s}) = \max_{i=1}^{n} X_i(\mathbf{s})$. E.g. $X_i(\mathbf{s})$: daily cumulative precipitation at site $\mathbf{s}$
- Have observations at $\mathbf{s}_1, \cdots, \mathbf{s}_k$, could fit a GEV to each location But ...
- Need spatial models for $\mu(\mathbf{s}), \sigma(\mathbf{s}), \xi(\mathbf{s})$, $\mathbf{s} \in \mathcal{S}$
  - spatial interpolation to produce the return level map
  - borrowing strength over space



Source: Fig. 8 (a) of Cooley et al. 2007

Bayesian Hierarchical Approach [Cooley et al. 2007]

Model the spatial dependence of parameters of extreme value distribution

$$\left\{ \underbrace{\pi(\boldsymbol{\theta_1}(\mathbf{s}), \boldsymbol{\theta_2} | m(\mathbf{s}))}_{\text{posterior}} \propto \underbrace{\pi(m(\mathbf{s}) | \theta_1(\mathbf{s}))}_{\text{data}} \underbrace{\pi(\boldsymbol{\theta_1}(\mathbf{s}) | \boldsymbol{\theta_2})}_{\text{process}} \underbrace{\pi(\boldsymbol{\theta_2})}_{\text{prior}} \right\}_{\mathbf{s} \in \mathcal{D}}$$

- **Process level**: model each parameter across the spatial region as a Gaussian process to induce the spatial dependence
- **Data level**: marginals follow independent extreme value distribution given the parameters at the process level

> The main purpose is to generate the return level map, e.g. 100 year site-wise precipitation event

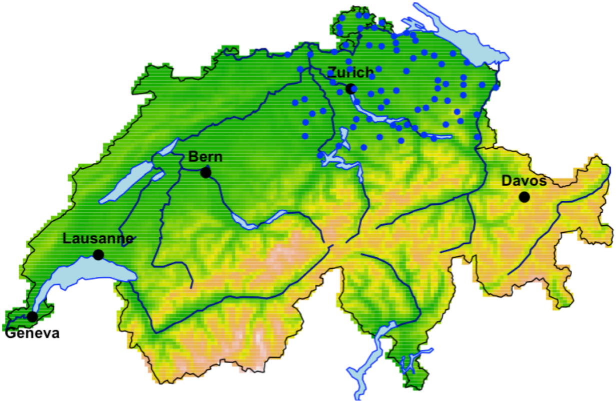
Hierarchical Approach Cont'd

- Data: $Y(\mathbf{s}_i) | (\mu(\mathbf{s}), \sigma(\mathbf{s}), \xi(\mathbf{s})) \overset{i.i.d.}{\sim} \mathrm{GEV}(\mu(\mathbf{s}_i), \sigma(\mathbf{s}_i), \xi(\mathbf{s}_i))$, $\mathbf{s}_i \in \mathcal{S}$
- Latent processes: $(\mu(\mathbf{s}), \phi(\boldsymbol{s}) = \log(\sigma(\mathbf{s})), \xi(\mathbf{s}))$ are Gaussian Processes $\mathrm{GP}(m(\cdot), K_\theta(\cdot,\cdot))$
- Prior: need to specify the prior distributions of the parameters for mean and covariance functions to carry out Bayesian inference $\Rightarrow$ can be difficult to specify

Estimating Switzerland Rainfall Return Level Map

[Davison, Padoan & Ribatet, 2012]

Summer (June–Aug.) annual maxima precipitation in Switzerland over the years 1962–2008



Let's give a try on using Bayesian hierarchical approach to produce, say, 20-year return level

R Implementation in SpatialExtremes

```
## Prior specification

```{r}
hyper <- list()
hyper$betaMeans <- list(loc = rep(0, 3), scale = rep(0, 3), shape = 0)
hyper$betaIcov <- list(loc = diag(rep(1/1000, 3)), scale = diag(rep(1 / 1000, 3)),
shape = 1 / 10)
hyper$sills <- list(loc = c(1, 12), scale = c(1, 1), shape = c(1, 0.04))
hyper$ranges <- list(loc = c(5, 3), scale = c(5, 3), shape = c(5, 3))
hyper$smooths <- list(loc = c(1, 1), scale = c(1, 1), shape = c(1, 1))
prop <- list(gev = c(3, 0.1, 0.3), ranges = c(1, 0.8, 1.2), smooths = rep(0, 3))
start <- list(sills = c(10, 10, 0.5), ranges = c(20, 10, 10), smooths = c(1, 1, 1),
beta = list(loc = c(25, 0, 0), scale = c(33, 0, 0), shape = 0.001))
```

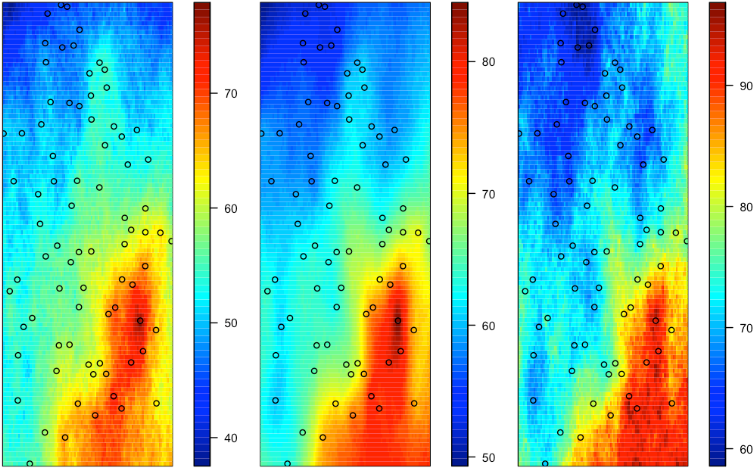
Running the Gibbs sampler

```{r}
loc.form <- scale.form <- y ~ lon + lat; shape.form <- y ~ 1
chain <- latent(rain, coord[, 1:2], "powexp", loc.form, scale.form, shape.form, hyper =
hyper,
                prop = prop,
start = start, n = 100, burn.in = 500, thin = 5)
chain
```
```

Summary of Bayesian Hierarchical Approach

- +: The latent variable approach is very flexible and can helps in improving "weak" trend surfaces
- +: Easily extends to threshold exceedances
- -: It is however time consuming and no spatial dependence (on the data layer) is taken into account
- -: Quite difficult to define prior distributions and identifiability problems—typically sills and ranges.

Inference for 20-Year Return Level



Extremes of Stochastic Processes

So far the spatial dependence in extremes has been ignored, it is time to take it into account using max-stable processes

- Motivating example:
  $1 - \mathbb{P}\left(X_t(\boldsymbol{s}) \leq h(\boldsymbol{s}) \quad \forall \boldsymbol{s} \in \mathcal{S}, \quad \forall t \in [a, b]\right)$

  e.g. $X_t(\boldsymbol{s})$ be the sea level at location $\boldsymbol{s}$ at time t and $h(\boldsymbol{s})$ be the height of seawall at $\boldsymbol{s}$

- Want: a limiting stochastic process $\{M(\boldsymbol{s})\}_{\boldsymbol{s} \in \mathcal{S}}$ s.t.

$$\left\{\frac{\max_{t=1}^{n} X_t(\boldsymbol{s}) - b_n(\boldsymbol{s})}{a_n(\boldsymbol{s})}\right\}_{\boldsymbol{s} \in \mathcal{S}} \overset{fdd}{\to} \{M(\boldsymbol{s})\}_{\boldsymbol{s} \in \mathcal{S}}$$

It turns out that if such a $\{M(\boldsymbol{s})\}_{\boldsymbol{s} \in \mathcal{S}}$ exists, it must be a max-stable process

Max-stable processes

A stochastic process $\{M(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}}$ is max-stable if there exist functions $\{a_n(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}} > 0$ and $\{b_n(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}}$ such that

$$\left\{\frac{\max_{i=1}^{n} M_i - b_n(\boldsymbol{s})}{a_n(\boldsymbol{s})}\right\}_{\boldsymbol{s}\in\mathcal{S}} \stackrel{fdd}{=} \{M(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}},\ n \geq 1$$

where $M_1, \cdots, M_n$ be the independent copies of $\{M(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}}$

- $\{M(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathcal{S}}$: pointwise maxima over the spatial domain
- Any max-stable process, after marginal transformation, has its spectral representation (see next slide) $\Rightarrow$ it is easier to specify/construct a max-stable process via the spectral representation

Spectral Representations for Max-Stable Processes [de Haan, 1984; Schlather, 2002]

Let $\{\zeta_i\}_{i\geq 1}$ be the points of Poisson process on $(0,\infty]$ with intensity $d\Lambda(\zeta) = \zeta^{-2}\, d\zeta$ and $W_1, W_2, \cdots$ be independent copies of a non-negative stochastic process $\{W(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathbb{R}^k}$ such that $\mathbb{E}[W(\boldsymbol{s})] = 1\ \forall \boldsymbol{s} \in \mathbb{R}^k$. Suppose $\{\zeta_i\}_{i\geq 1}$ independent of $W_i's$. Then

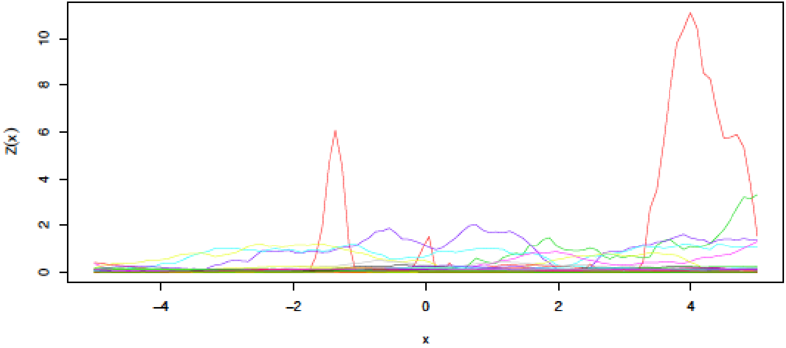
$$\left\{ \bigvee_{i\geq 1} \zeta_i W_i(\boldsymbol{s}) \right\}_{\boldsymbol{s}\in\mathbb{R}^k}$$

is max-stable process with unit Fréchet marginals

e.g. Brown–Resnick model: $W(\boldsymbol{s}) = \exp\left(\varepsilon(\boldsymbol{s}) - \gamma(\boldsymbol{s})\right)$, $\boldsymbol{s} \in \mathbb{R}^k$, where $\{\varepsilon(\boldsymbol{s})\}$ is a Gaussian process with stationary increments and semivariogram γ

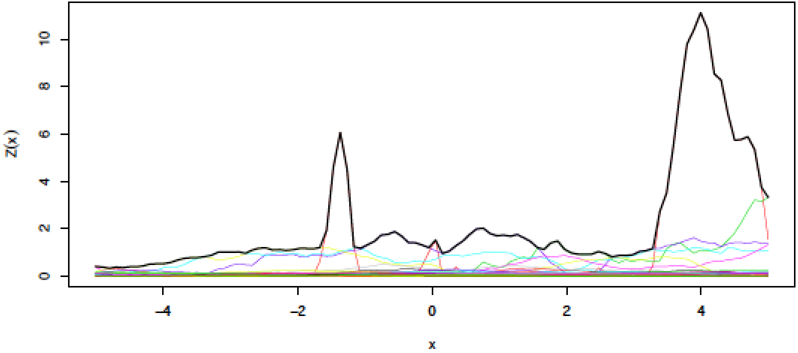
Schlather's Spectral Representation

$$\{\zeta_i W_i(\boldsymbol{s})\}_{\boldsymbol{s}\in\mathbb{R}^k}$$



Schlather's Spectral Representation Cont'd

$$\left\{ \bigvee_{i \geq 1} \zeta_i W_i(\boldsymbol{s}) \right\}_{\boldsymbol{s} \in \mathbb{R}^k}$$



Likelihood for a Max-Stable Process

Recall

$$\mathbb{P}(M_1 \le m_1, \cdots .M_d \le m_d) = \exp(-\underbrace{\mathbb{E}\left[\max\left\{\frac{W(\boldsymbol{s}_1)}{m_1}, \cdots, \frac{W(\boldsymbol{s}_d)}{m_d}\right\}\right]}_{V(m_1,\cdots,m_d)\equiv V})$$

joint density

$$\frac{\partial}{\partial m_1 \cdots \partial m_d} e^{-V} = e^{-V} \sum_{\pi \in \mathcal{P}_d} (-1)^{|\tau|} \prod_{j=1}^{|\tau|} \frac{\partial^{|\tau_j|}}{\partial m_{\tau_j}} V$$

where the cardinality of $\mathcal{P}_d = B_d$, the Bell number of order d

$$B_3 = 5$$
$$B_{10} = 115,975$$
$$B_{25} = 4,638,590,332,229,999,353!$$

Current Modeling Practice on Spatial Extremes

- Composite likelihood (Lindsay, 1988): pairwise likelihoods (Padoan et al. (2010)), Triplewise likelihoods (Genton et al. (2011), Huser and Davison (2013))

- Bayesian hierarchical approach for max-stable processes: Reich and Shaby (2012)

- Full likelihood by exploiting "temporal" information: Wadsworth and Tawn (2014), Thibaud et al. (2015)

Extremal Coefficient Function

The extremal coefficient is

$$\theta = \frac{\log \mathbb{P}(\max_{j=1}^{d} M(\boldsymbol{s}_j) \leq m)}{\log \mathbb{P}(M(\boldsymbol{s}) \leq m)} = V(1, \cdots, 1).$$

$$1\,(\text{dependence}) \leq \theta \leq d\,(\text{independence})$$

- In a spatial context it is often more convenient to consider bivariate extremal coefficients and link to their pairwise spatial distances
- One can use F-madogram (Cooley, Naveau, & Poncet, 2006)

$$\nu_F(h) = \frac{1}{2}\mathbb{E}\left[|F(M(\boldsymbol{s})) - F(M(\boldsymbol{s}'))|\right]$$

to estimate

$$\theta(h) = \frac{1 - 2\nu_F(h)}{1 - 2\nu_F(h)}$$

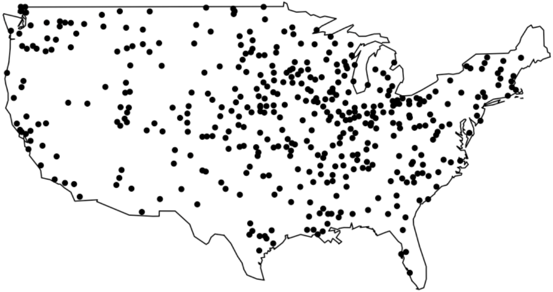
Fitting Max-Stable Porcesses using Pairwise Likelihood

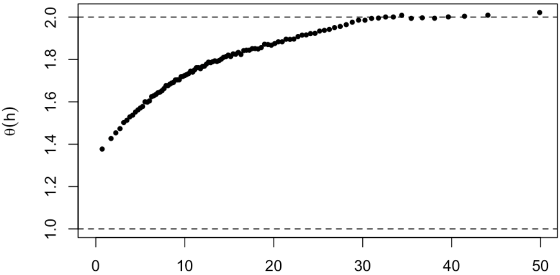
Since the (full) likelihood is too expensive to compute, we consider pairwise log-likelihood

$$\ell_p(\theta; \boldsymbol{m}) = \sum_{t=1}^{T} \sum_{j=1}^{d-1} \sum_{k=j+1}^{d} w_{jk} \log g(m_t^{(j)}, m_t^{(k)}; \theta)$$

- g is the density of the associated bivariate GEV where spatial covariates can be added to define trend surfaces
- Maximize this gives the maximum composite likelihood estimator (MCLE)
- Inference can be performed under the composite likelihood framework

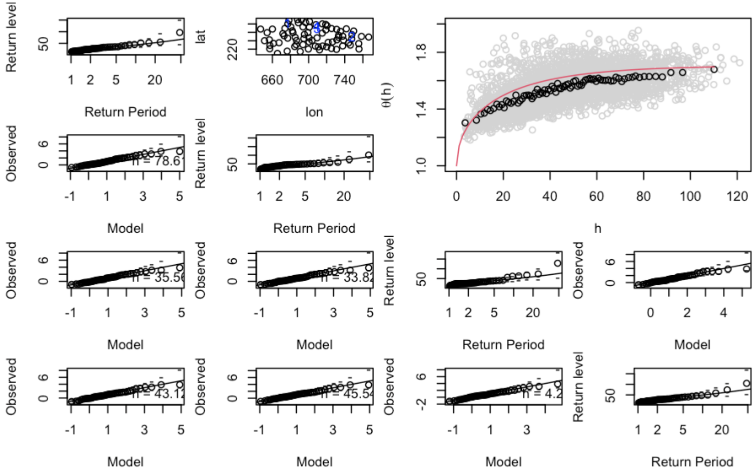
Summer Annual Maxima Temperature in Continental US, 1911-2010





Fitting Max-Stable Process Model!

```{r}
loc.form <- y ~ lon + lat + alt
scale.form <- y ~ lon + lat
shape.form <- y ~ 1
(fit <- fitmaxstab(rain, coord, "powexp", nugget=0, loc.form, scale.form, shape.form,
marg.cov = cbind(alt = alt)))
plot(fit)
```



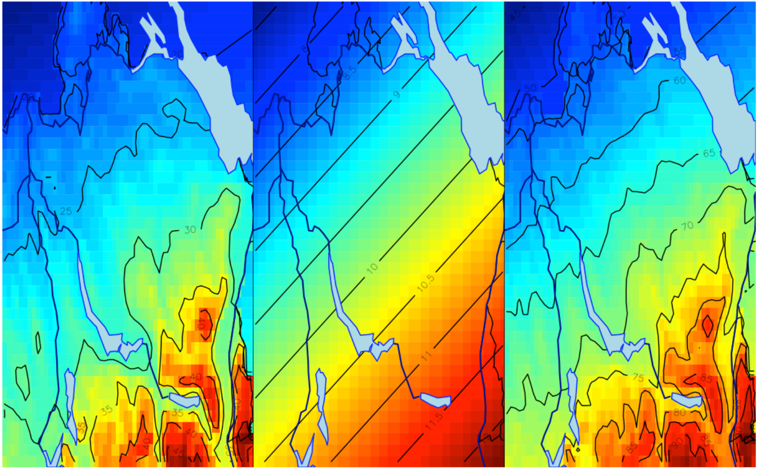
Model Selection

```
> M0 <- fitmaxstab(rain, coord, "powexp", nugget = 0, locCoeff4 = 0, loc.form, scale.form, shape.for
m, marg.cov = cbind(alt = alt))
Computing appropriate starting values
Starting values are defined
Starting values are:
      range       smooth    locCoeff1    locCoeff2    locCoeff3    locCoeff4
27.257344723  0.974059614 29.764826111  0.037985247 -0.135302350  0.008372902
 scaleCoeff1  scaleCoeff2  scaleCoeff3  shapeCoeff1
10.529521199  0.013058818 -0.040111724  0.156381541
> M1 <- fitmaxstab(rain, coord, "powexp", nugget = 0,
+ loc.form, scale.form, shape.form, marg.cov = cbind(alt = alt))
Computing appropriate starting values
Starting values are defined
Starting values are:
      range       smooth    locCoeff1    locCoeff2    locCoeff3    locCoeff4
27.257344723  0.974059614 29.764826111  0.037985247 -0.135302350  0.008372902
 scaleCoeff1  scaleCoeff2  scaleCoeff3  shapeCoeff1
10.529521199  0.013058818 -0.040111724  0.156381541
> anova(M0, M1)
Eigenvalue(s): 161.12

Analysis of Variance Table
   MDf Deviance Df Chisq Pr(> sum lambda Chisq)
M0   9  2246053
M1  10  2232232  1 13821              < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Prediction

Spatial maps for μ, σ, and 20-year return level



Summary of Max-Stable Process Models

- $+$: Justified by extreme value theory
- $+$: Able to describe asymptotic dependence
- -: Describe everything that is asymptotically independent as exactly independent