# Multivariate Extremes Value Methods for Spatial Flood Risk Assessment

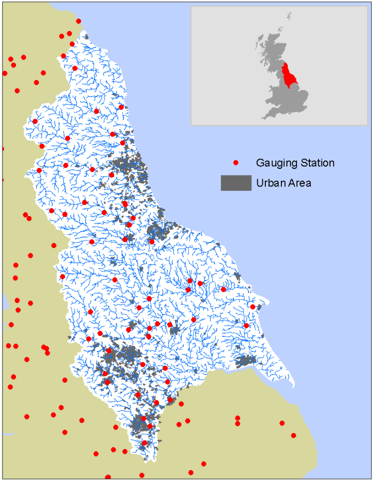
**Jonathan Tawn (Lancaster University)**

**with**

**Caroline Keef (Yorkshire Water), Rob Lamb (JBA)**

**Based on papers:**
**JH+JT (2004) JRSS B**
**CK+JT+ RL (2013) Environmetrics**

# Problem: What is the distribution of annual financial loss from river flows for the North East of England?

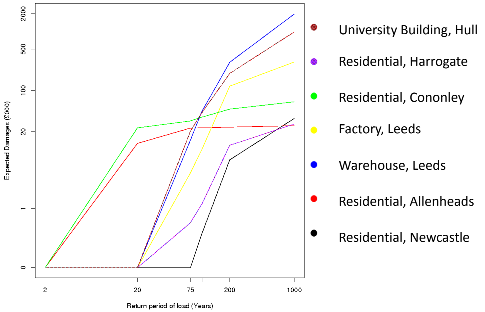



**Daily mean flow measurements at 145 gauging stations in and around the region**

## Link between River Flow and Loss at a Site

- At each location the river flow determines the financial loss
- This loss is tabulated as a function of flow return period

Individual property damage data



## Why Not Univariate Approach on Extreme Losses?

Fails to incorporate:

- knowledge of non-linear structure of loss functions
- which locations were hit and missed by past flood events
- changes in portfolio of insured properties
- changes in flood defences, inflation

Does not give information about nature of extreme events for planning and scenario assessment

## Strategy

- model joint distribution of flows at 145 sites
- simulation extreme events from this distribution
- interpolate return levels along rivers
- evaluate loss over network using known loss functions
- evaluate distribution of losses in event by Monte Carlo methods
- model distribution of number of events per year
- evaluate distribution of annual losses by Monte Carlo methods

# Multivariate Extreme Values: Copulas

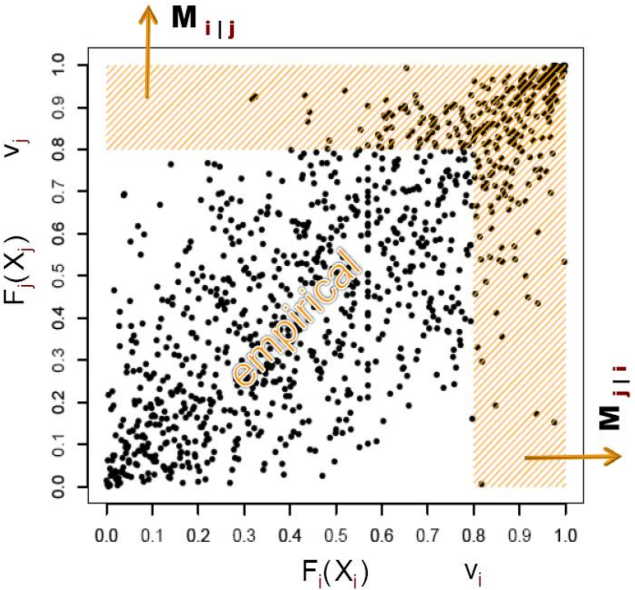
**Model joint distribution function $F_{\mathbf{X}}$ of $\mathbf{X} = (X_1, \ldots, X_m)$**

$$F_{\mathbf{X}}(x_1, \ldots, x_m) = C\{F_1(x_1), \ldots, F_m(x_m)\}$$

**where**

- $F_i$ **is the marginal distribution function for** $X_i$
- $C$ **is the copula with uniform margins**

# Focus on Joint Tail Modelling



## Marginal Extremes

- **Univariate variable** $X$
- **Marginal distribution function** $\mathrm{Pr}(X < x)$
- **Upper end point** $x_F$
- **Assume that there exists** $\sigma_u > 0$ **such that limit is non-degenerate for** $x > 0$**, then limit must be RHS**

$$\lim_{u \to x_F} \mathrm{Pr}\left(\frac{X-u}{\sigma_u} > x \,\middle|\, X > u\right) = \begin{cases} [1+\xi x]_+^{-1/\xi} & \xi \neq 0 \\ \exp(-x) & \xi = 0 \end{cases}$$

**where** $\xi$ **is a shape parameter,** $y_+ = \max(y, 0)$

## Generalised Pareto distribution (GPD)

- For $u$ close to $x_F$, motivates the asymptotic approximation for $x > 0$

$$\Pr(X - u > x \,|\, X > u) = \left[1 + \xi \frac{x}{\sigma_u}\right]_+^{-1/\xi}$$

- For large $u$

$$\Pr(X > x) = p_u \left[1 + \xi \frac{x - u}{\sigma_u}\right]_+^{-1/\xi} \qquad x > u$$

where $p_u = \Pr(X > u)$

- GPD tail for $X$

## GPD Extrapolation

**For large** $u$ **and** $x > 0$

$$\Pr(X > x + u) = \left(1 + \xi \frac{x}{\sigma_u}\right)_+^{-1/\xi} \Pr(X > u)$$

**We estimate** $\Pr(X > u)$ **empirically and use the formula for extrapolation**

## Copulas with Laplace margins

- By suitable transformation $\mathbf{X} \to \mathbf{S}$, $C$ could have any marginal
- We take $\mathbf{S} = (S_1, \ldots, S_m)$ to have Laplace marginals
- View Laplace scale as log return period
- Laplace density is

$$f(s) = \frac{1}{2}\exp(-|s|), \quad -\infty < s < \infty$$

- For Laplace distribution $(\sigma_u = 1, \xi = 0)$ with $x > 0$ and $u > 0$

$$\Pr(S > x + u \mid S > u) = \exp(-x)$$

## Extremal Dependence

**Pair** $(S_1, S_j)$

$$\chi_j = \lim_{v \to \infty} \Pr(S_j > v \mid S_1 > v)$$

- **Asymptotic dependence** $\chi_j > 0$
- **Asymptotic independence** $\chi_j = 0$

## Asymptotic Dependence: a conditional viewpoint

**If all variables are asymptotically dependent on $S_1$ then for**
$\mathbf{S} = (S_1, S_2, \ldots S_m) = (S_1, \mathbf{S}_{-1})$

$$\lim_{v\to\infty} \Pr\left(S_1 - v > s, \mathbf{S}_{-1} - S_1 < \mathbf{z} | S_1 > v\right) = \exp(-s) H_{m-1}(\mathbf{z})$$

**with $H_{m-1}$ non-degenerate in each margin and $s > 0$**
**Note: limiting conditional independence**

**If all components of $\mathbf{S}_{-1}$ are asymptotic independent of $S_1$ then $H_{m-1}$ puts all mass at $-\infty$ for each component**

## Conditional Asymptotics:

**Look for functions $\mathbf{a}$ and $\mathbf{b}$: $(\mathbb{R} \to \mathbb{R}^{m-1})$**

$$\lim_{v \to \infty} \Pr\left(S_1 - v > s, \frac{\mathbf{S}_{-1} - \mathbf{a}(S_1)}{\mathbf{b}(S_1)} \leq \mathbf{z} \mid S_1 > v\right) = \exp(-s) G_{m-1}(\mathbf{z})$$

$G_{m-1}$ **is non-degenerate in each margin and** $s > 0$

**Applies for asymptotic dependence and asymptotic independence**

**Simple forms for $\mathbf{a}(s) = \alpha s$ and $\mathbf{b}(s) = s^{\beta}$ are sufficient in all theoretical examples**

## Conditional Method: Heffernan and T. (2004, JRSS B)

**Multivariate thresholded regression:**
**Given** $S_1 = s > u$

$$\mathbf{S}_{-1} = \boldsymbol{\alpha} s + s^{\boldsymbol{\beta}} \boldsymbol{\mu} + \boldsymbol{\sigma} s^{\boldsymbol{\beta}} \mathbf{Z}$$

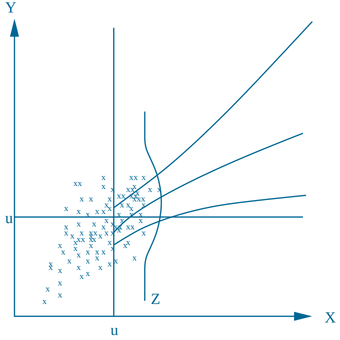
- $m-1$**-dimensional parameters** $-\mathbf{1} \leq \boldsymbol{\alpha} \leq \mathbf{1}$, $\boldsymbol{\beta} < \mathbf{1}$, $\boldsymbol{\sigma} > \mathbf{0}$
- $E(\mathbf{Z}) = \mathbf{0}$ **and** $Var(Z_j) = 1 \quad \forall j$
- $\mathbf{Z}$ **is independent of** $S_1$
- **Estimate** $\mathbf{Z} \sim G_{m-1}$ **the distribution of multivariate residuals empirically**

# Conditional Method: Heffernan and T. (2004, JRSS B)

## Multivariate thresholded regression:

**Given** $S_1 = s > u$

$$\mathbf{S}_{-1} = \left\{\alpha s + s^{\beta}\boldsymbol{\mu}\right\} + \sigma s^{\beta}\mathbf{Z}$$



## Special Cases

$$\mathbf{S}_{-1} = \boldsymbol{\alpha} S_1 + S_1^{\boldsymbol{\beta}}(\boldsymbol{\mu} + \boldsymbol{\sigma}\mathbf{Z})$$

### Asymptotic Dependence

$$\boldsymbol{\alpha} = \mathbf{1} \text{ and } \boldsymbol{\beta} = \mathbf{0}$$

### Asymptotic Independence with $S_j$ (independence)

$$\alpha_j < 1 \qquad (\alpha_j = 0, \beta_j = 0)$$

### Positive (negative) extremal dependence with $S_j$

$$0 < \alpha_j < 1 \qquad (-1 < \alpha_j < 0)$$

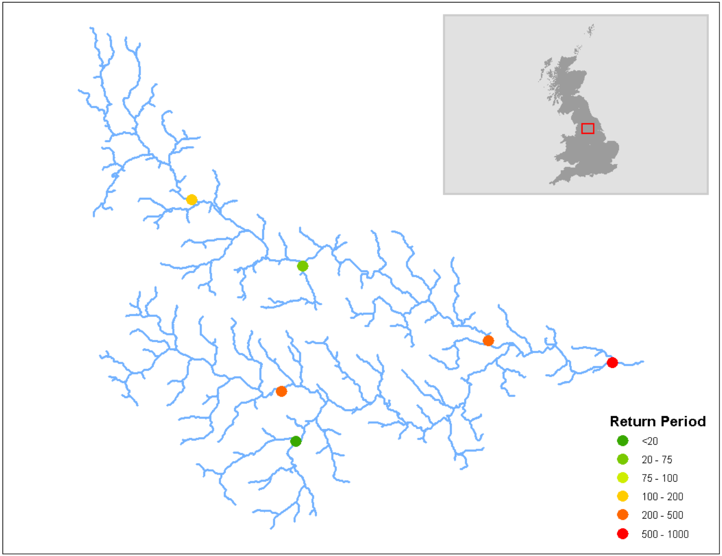
### Multivariate Normal Copula

$$\alpha_j = \mathbf{sign}(\rho_{1j})\rho_{1j}^2 \text{ and } \beta_j = \frac{1}{2} \text{ for } j = 2, \ldots, m$$

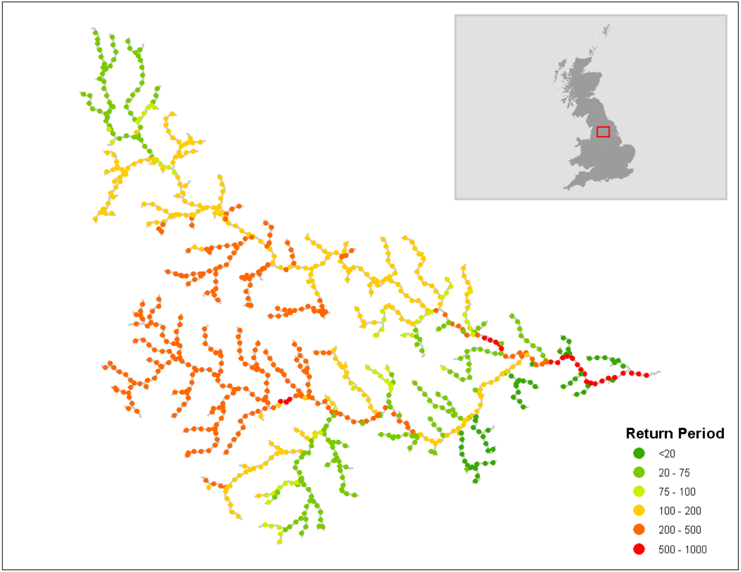
## Conditional Method: Assessment of fit Videos

- Pairwise assessment
- Higher order assessment

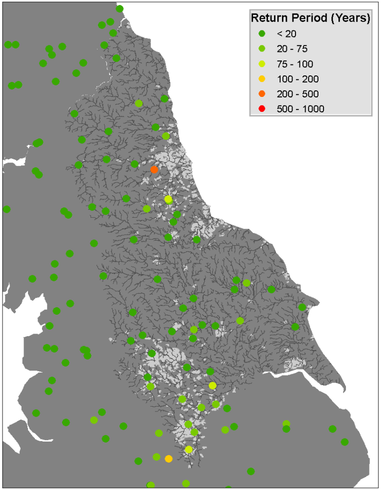
# Simulated Event

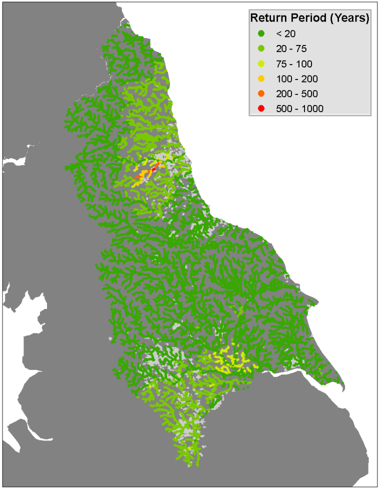


# Interpolation

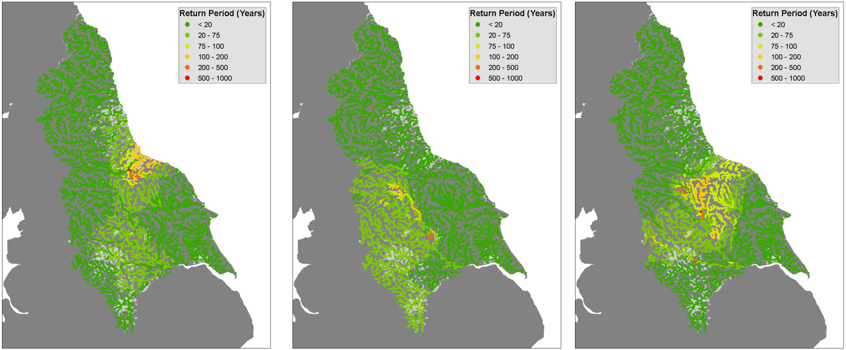


# Worst Observed Loss Event: Oct 2000

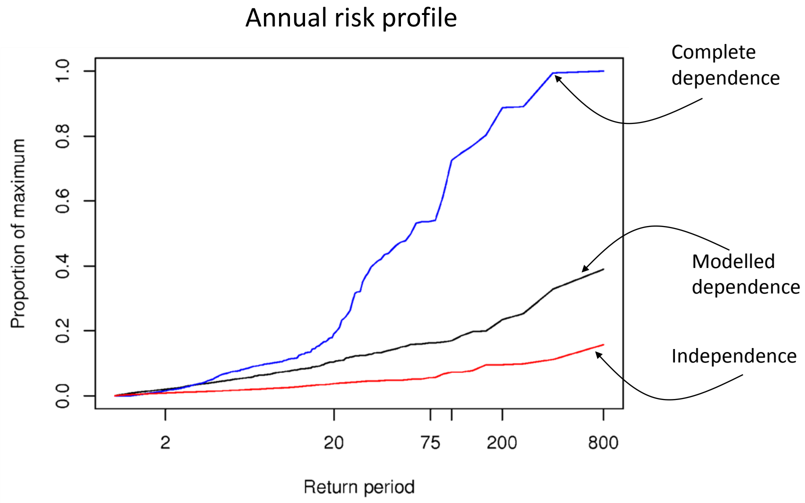


# Different Simulated 100 Year Loss Events



# Estimated Annual Loss Distribution



# Ongoing work: with Ross Towe KTP with JBA

**Inclusion of rainfall data**

- 15 min data at gauges
- daily gridded data

**Added value of rainfall data**

- conditionally independent of flows given close neighbouring flows
- small catchments with rapid response
- catchments in headwaters
- areas without flow gauges
- areas with limited overlapping flow gauges
- joint pluvial and fluvial flooding