# The Extremal Dependence of Storm Severity, Wind Speed and Surface Level Pressure in the Northern North Sea

Ross Towe,
Emma Eastoe, Jonathan Tawn, Philip Jonathan and Yanyun Wu

r.towe@lancaster.ac.uk

# Return levels and Pooling

- Return levels are used in the design criteria of an offshore structure to ensure the structure withstands a particularly extreme event
- Can calculate return levels for Hs for a given site in the North sea

# Return levels and Pooling

- Return levels are used in the design criteria of an offshore structure to ensure the structure withstands a particularly extreme event
- Can calculate return levels for Hs for a given site in the North sea
- A model is fitted to estimate the return levels of $H_s$, as there is not enough data to estimate the return levels empirically
- To increase our confidence in the estimates of the return levels, we can pool data across sites

# Pooling

What happens if there is dependence between sites?

| Data set | SE(50yr) | SE(100yr) |
|---|---|---|
| One site | 0.65 | 1.00 |
| Pooled (perfectly dependent) sites | 0.62 | 0.94 |
| Pooled (independent) sites | 0.22 | 0.30 |

Standard errors (uncertainty) of the 50 and 100 year return levels

# Pooling

What happens if there is dependence between sites?

| Data set | SE(50yr) | SE(100yr) |
|---|---|---|
| One site | 0.65 | 1.00 |
| Pooled (perfectly dependent) sites | 0.62 | 0.94 |
| Pooled (independent) sites | 0.22 | 0.30 |

Standard errors (uncertainty) of the 50 and 100 year return levels

- Pooling dependent data produces false estimates of the standard errors
- The effective sample size is smaller for pooled perfectly dependent data compared to pooled independent data

# Extremal dependence

- We can also consider the extremal dependence **Wind Speed** (**$WS$**) and **Significant Wave Height** (**$H_s$**) at a given site
- A storm is characterised by a number of different factors, such as wind and waves
- A severe storm can be a result of high waves but also a combination of extreme wind and waves

# Extremal dependence

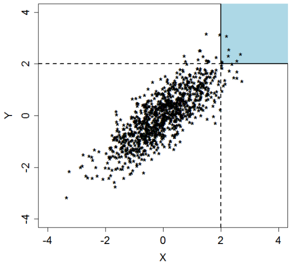
- We can also consider the extremal dependence **Wind Speed (*WS*)** and **Significant Wave Height ($H_s$)** at a given site
- A storm is characterised by a number of different factors, such as wind and waves
- A severe storm can be a result of high waves but also a combination of extreme wind and waves
- **Should we model *WS* and $H_s$ jointly or independently?**
- **We are interested in the joint behaviour *WS* and $H_s$ during a storm**

# Extremal dependence

## Bivariate Normal distribution

**The dependence in the body of the data is different to that of the tail**

Consider the bivariate Normal distribution with $\rho = 0.8$




Calculate,

$P(X > z \mid Y > z)$ for $z \to \infty$,

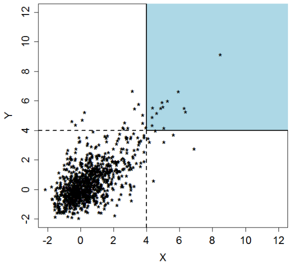
| $z$ | $P(X > z \mid Y > z)$ |
|---|---|
| 1.00 | 0.60 |
| 1.50 | 0.49 |
| 2.00 | 0.33 |
| 2.50 | 0.00 |
| 3.00 | 0.00 |

# Extremal dependence

## Bivariate extreme value distribution

**The dependence in the body of the data is different to that of the tail**

Consider the bivariate extreme value distribution with $\rho = 0.8$



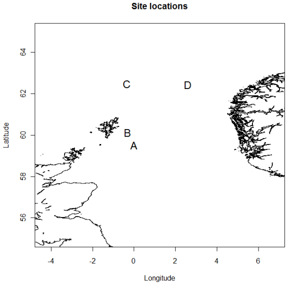

Calculate

$P(X > z \mid Y > z)$ for $z \to \infty$,

| $z$ | $P(X > z \mid Y > z)$ |
|---|---|
| 2.00 | 0.72 |
| 4.00 | 0.79 |
| 6.00 | 1.00 |
| 8.00 | 1.00 |
| 10.00 | 1.00 |

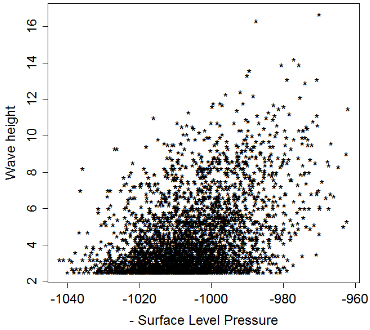
# Data

- Storm peak events from a 50 year data set:
  - Significant wave height [metres]
  - Wind speed [miles per hour]
  - Surface level pressure [hectopascal]
  - Storm direction [degrees]
- Sites have different local characteristics

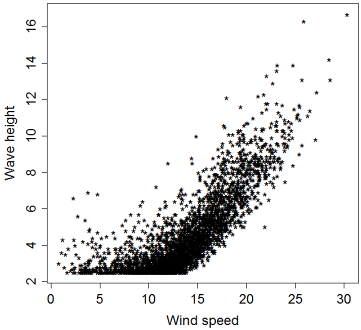


# Original margins

$(NSLP \text{ vs } H_s)$



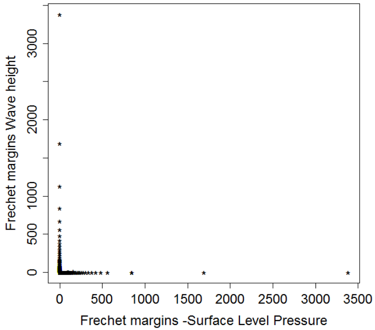

$(WS \text{ vs } H_s)$



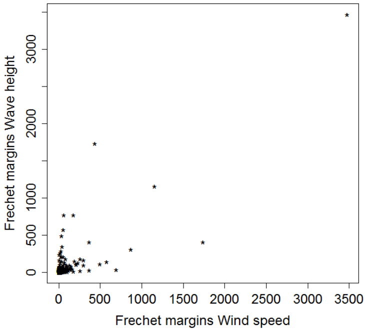
# Fréchet margins

Consider the data on Fréchet margins, $F(z) = \exp(-1/z)$

$(NSLP \text{ vs } H_s)$




$(WS \text{ vs } H_s)$



# Definition

Consider a pair of random variables *(X,Y)* with identical margins,

**Asymptotic independence:**

$$P(X > z\ ,\ Y > z)\ =\ 0,\ \text{as } z \to \infty$$

*Two extreme events are unlikely to occur simultaneously*
*Consider the measure:* $\bar{\chi}$

**Asymptotic dependence:**

$$P(X > z\ ,\ Y > z)\ >\ 0,\ \text{as } z \to \infty$$

*Two extreme events are likely to occur simultaneously*
*Consider the measure:* $\chi$

# Extremal dependence measures

**Asymptotic independence:**

- $\bar{\chi}$ determines the strength of asymptotic independence
- $\bar{\chi} \in (-1, 1)$.
- If $\bar{\chi} = 1 \Rightarrow$ asymptotic dependence

# Extremal dependence measures

**Asymptotic independence:**

- $\bar{\chi}$ determines the strength of asymptotic independence
- $\bar{\chi} \in (-1, 1)$.
- If $\bar{\chi} = 1 \Rightarrow$ asymptotic dependence

**Asymptotic dependence:**

- $\chi$ determines the strength of asymptotic dependence

$$\chi = P(X > z \mid Y > z)$$

- $\chi \in (0, 1)$.
- When $\chi = 0 \Rightarrow$ asymptotic independence

# Extremal dependence measures

**Asymptotic independence:**

- $\bar{\chi}$ determines the strength of asymptotic independence
- $\bar{\chi} \in (-1, 1)$.
- If $\bar{\chi} = 1 \Rightarrow$ asymptotic dependence

**Asymptotic dependence:**

- $\chi$ determines the strength of asymptotic dependence

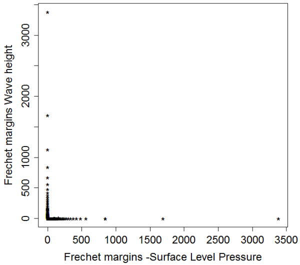
$$\chi = P(X > z \mid Y > z)$$

- $\chi \in (0, 1)$.
- When $\chi = 0 \Rightarrow$ asymptotic independence

**$\bar{\chi}$ and $\chi$ can be estimated empirically or by using a model**

# Asymptotic independence case

**(NSLP v $\mathbf{H_s}$)**

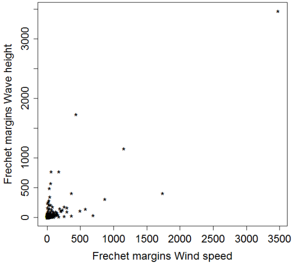



| Site | $\hat{\bar{\chi}}$ (95% *CI*) |
|---|---|
| A | 0.61 (0.50,0.72) |
| B | 0.61 (0.50,0.72) |
| C | 0.59 (0.47,0.70) |
| D | 0.65 (0.53,0.77) |

Estimates of $\bar{\chi}$ for the 4 sites with the threshold set at the respective 80% quantile estimation threshold

# Asymptotic dependence case

## (WS v $H_s$)

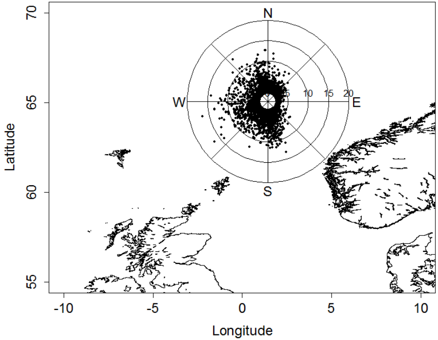



| Site | $\hat{\bar{\chi}}$ (95% *CI*) | $\hat{\chi}$ (95% *CI*) |
|---|---|---|
| A | 0.95(0.81,1.00) | 0.73(0.69,0.79) |
| B | 0.86(0.73,0.98) | 0.74(0.69,0.78) |
| C | 0.85(0.71,0.99) | 0.73(0.68,0.78) |
| D | 0.84(0.71,0.97) | 0.72(0.67,0.77) |

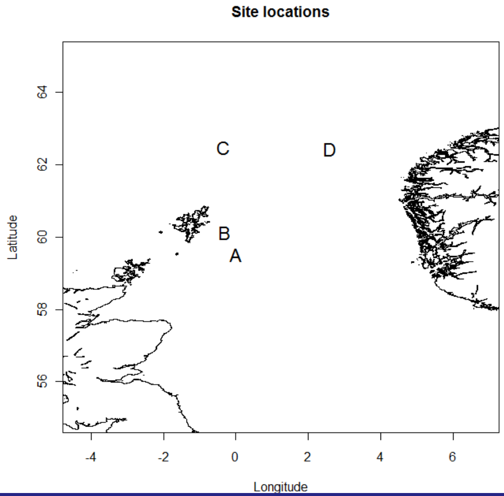
Estimates of ($\bar{\chi}$ and $\chi$) for the four sites with the threshold set at the respective 80% quantile estimation threshold

# Possible covariates

- The marginal behaviour of the height of waves and the speed of winds is influenced by:
  - Land shadow
  - Storm direction
- Can consider storm direction to be a covariate
- Is the dependence structure, therefore influenced by the same covariates?



# Site locations



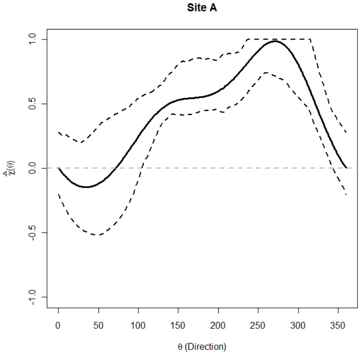
# Asymptotic independence case

Introduce storm direction ($\theta$) as a covariate by using a Fourier series:

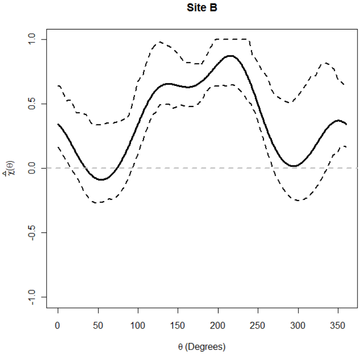
$$\begin{aligned}
\bar{\chi}(\theta) &= 2\min\{\eta_+(\theta), 1\} - 1, \text{ where} \\
\log[\eta_+(\theta)] &= \alpha_0 + \sum_{i=1}^{d} \alpha_{1i} \cos(\alpha_{2i} + i\theta), \quad \theta \in [0, 360^\circ).
\end{aligned}$$

# $(NSLP\ vs\ H_s)$ : $\bar{\chi}(\theta)$

$\bar{\chi}(\theta)$ for Site A

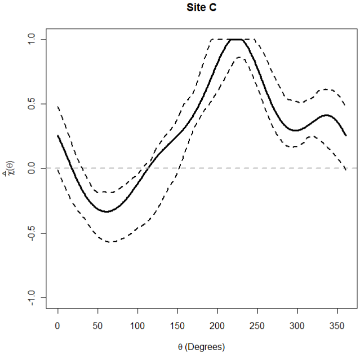



$\bar{\chi}(\theta)$ for Site B

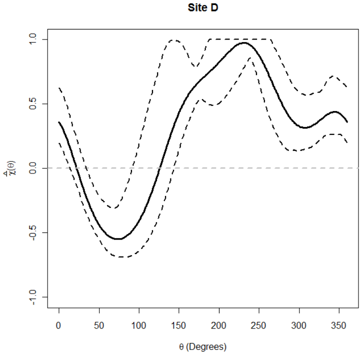


# $(NSLP\ vs\ H_s)$ : $\bar{\chi}(\theta)$

$\bar{\chi}(\theta)$ for Site C




$\bar{\chi}(\theta)$ for Site D



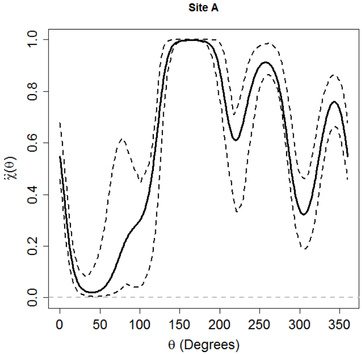
# Asymptotic dependence case

Introduce storm direction ($\theta$) as a covariate by using a Fourier series:

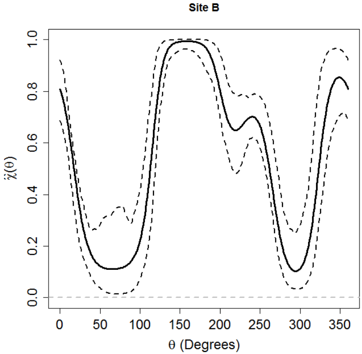
$$\text{logit}[\chi(\theta)] = \beta_0 + \sum_{i=1}^{d} \beta_{1i} \cos[\beta_{2i} + i\theta], \quad \text{for } \theta \in [0, 360^\circ)$$

# $(WS \text{ vs } H_s)$ : $\chi(\theta)$

$\chi(\theta)$ for Site A

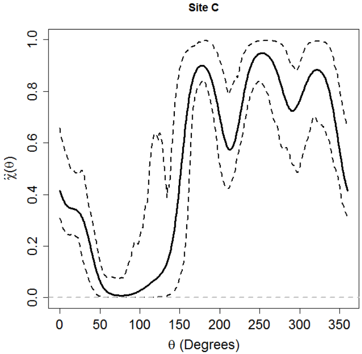



$\chi(\theta)$ for Site B

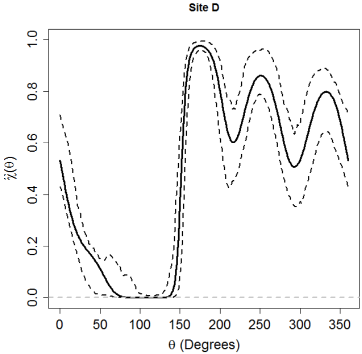


# $(WS\ vs\ H_s)$ : $\chi(\theta)$

$\chi(\theta)$ for Site C




$\chi(\theta)$ for Site D



# Concluding Remarks

- Need to think carefully about pooling of dependent samples, and dependence between extreme values. Intuition can be misleading
- Wind speed and wave height are asymptotically dependent
- NSLP and wave height are asymptotically independent
- Extremal dependence varies with storm direction are vital for reliable estimation of joint design values for winds, waves and currents
- Empirical modelling of $\chi$ and $\bar{\chi}$ may guide the choice of covariates for joint extreme value modelling