

Extreme ocean environments

Philip Jonathan

Lancaster University, Department of Mathematics and Statistics

Wales Mathematics Colloquium
(Slides at www.lancs.ac.uk/~jonathan)



Acknowledgement and overview

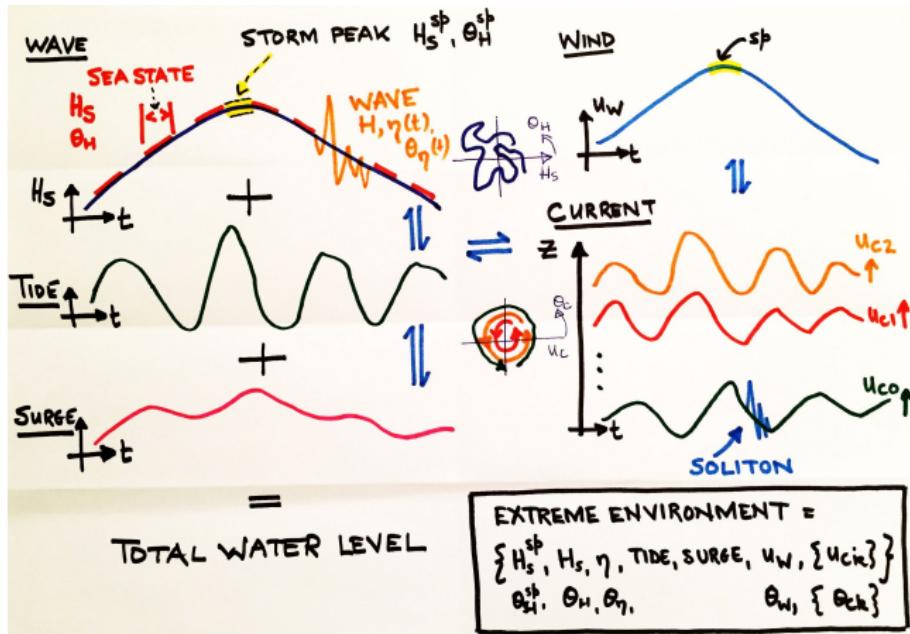
Thanks

- Lancaster : Emma Eastoe, Jon Tawn, Stan Tendijck, Elena Zanini
- Metocean Research Limited (NZ) : Kevin Ewans
- Shell : Graham Feld, Matthew Jones, David Randell, Emma Ross, Ross Towe
- UK Metoffice : Rob Shooter

Overview

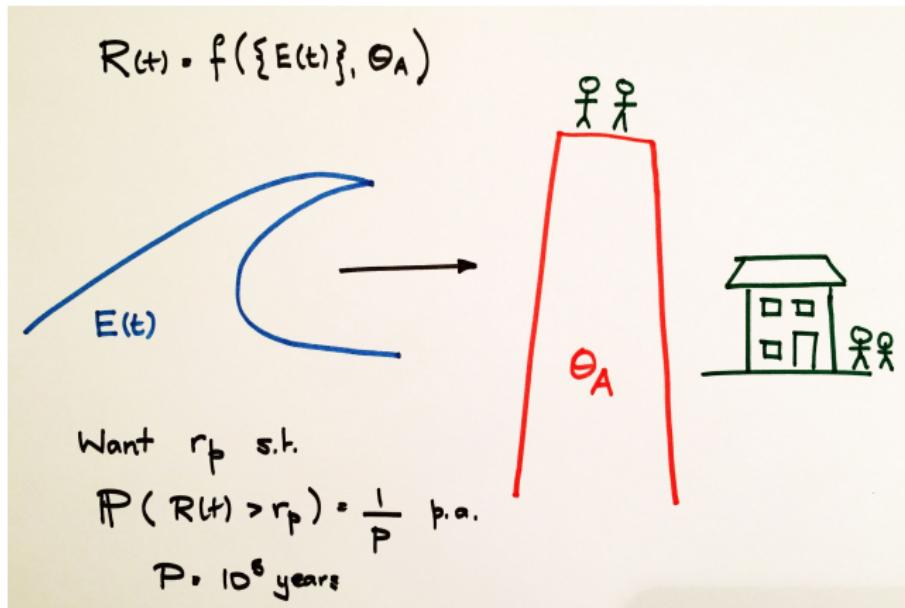
- Motivation
- Marginal extremes
- Multivariate conditional extremes

Modelling ocean storm environment



- Multiple coupled physical processes
- Rare, extreme events

Modelling structural risk



- o Ocean environment is harsh
- o Marine structures at risk of failure
- o Reliability standards must be met

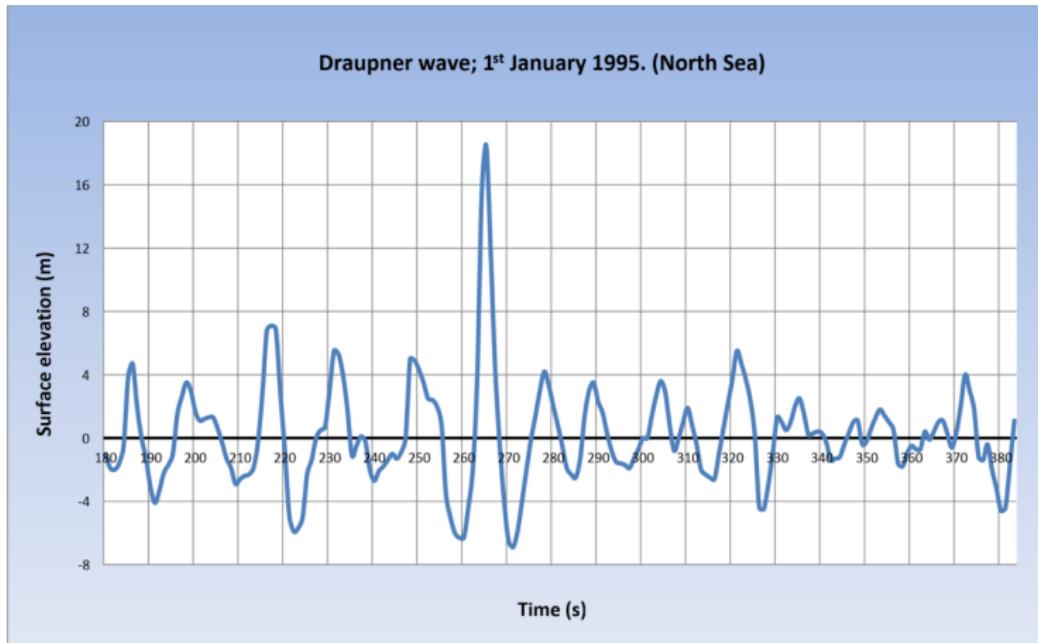
Spectacular scale



Offshore Portugal, 24m wave height, November 2017 (The Guardian)

- Nazaré is a great source of huge coastal waves

Spectacular scale



Laser readings, 1 January 1995. Wave 25.6m, crest 18.5m (Statoil / Equinor)

- Maximum recorded wave height > 30m (multiple events, various sources)
- Maximum recorded significant wave height : 19.0m (buoy, North Atlantic, 4 Feb 2013, WMO)

Wave impact damage



Norwegian Dream, Atlantic, 2007
(gcaptain.com)



Ike, Gulf of Mexico, 2008
(Joe Richard)

Optimal design

Set-up

- A marine system with “strength” specifications \mathcal{S}
- An ocean environment X dependent on covariates Θ
- A structural “loading” Y as a result of environment X and covariates Θ
- System utility (or risk) $U(Y|\mathcal{S})$ for loading Y and specification \mathcal{S}
- Desired U typically specified in terms of annual probability of failure
- $Y|X, \Theta$ and $X|\Theta$ (and $U?$) subject to uncertainty Z
- Z, Θ, X, Y are multidimensional random variables

Optimal design

- Estimate a model $f_{X|\Theta, Z}$ for the environment
- Estimate a model $f_{Y|X, \Theta, Z}$ for environment-structure interaction
- Estimate a model $f_{\Theta|Z}$ for the covariates

$$\mathbb{E}[U|\mathcal{S}] = \int_z \int_y \int_x \int_\theta U(y|\mathcal{S}, Z) f_{Y|X, \Theta, Z}(y|x, \theta, z) f_{X|\Theta, Z}(x|\theta, z) f_{\Theta|Z}(\theta|z) f_Z(z) d\theta dx dy dz$$

⇒ solve for \mathcal{S} to achieve required (safety) utility

Return values : conventional engineering practice

- Estimating $\mathbb{E}[U|\mathcal{S}]$ is difficult
- Design to extreme quantile of marginal **annual** distribution of single X instead

$$F_A(x) = \int_{\mathbf{Z}} \int_{\boldsymbol{\theta}} \int_k F_{X|\boldsymbol{\Theta},\mathbf{Z}}^k(x|\boldsymbol{\theta}, \mathbf{Z}) f_{C|\boldsymbol{\Theta},\mathbf{Z}}(k|\boldsymbol{\theta}, z) f_{\boldsymbol{\Theta}|\mathbf{Z}}(\boldsymbol{\theta}|z) f_{\mathbf{Z}}(z) dk d\boldsymbol{\theta} dz$$

where $f_{C|\boldsymbol{\Theta},\mathbf{Z}}$ is the density of annual rate of events given covariate $\boldsymbol{\Theta}$.

- Set the **return value** x_T (for $T = 1000$ years say) such that

$$F_A(x_T) = 1 - \frac{1}{T}$$

- Specify **conditional** return values for other X s given $X = x_T$
- Potentially as a function of covariates
- Ambiguous ordering of expectation operators ... a can of worms!

A model for the (non-stationary multivariate extreme) environment

- Expected utility and return values are dominated by **extreme** environments
- Have to estimate **tails** of distributions well
- Focus on a simple **Z-free** 2-D environment with stationary dependence

$$F_{X|\Theta,Z}(x|\theta, z) = \textcolor{red}{C}\left(F_{X_1|\Theta}(x_1|\theta), F_{X_2|\Theta}(x_2|\theta)\right) \text{ for simplicity, so}$$

$$\begin{aligned} f_{X|\Theta,Z}(x|\theta, z) &= f_{X_1, X_2|\Theta}(x|\theta) \\ &= f_{X_1|\Theta}(x_1|\theta) f_{X_2|\Theta}(x_2|\theta) \times \textcolor{red}{c}\left(F_{X_1|\Theta}(x_1|\theta), F_{X_2|\Theta}(x_2|\theta)\right) \text{ typically} \end{aligned}$$

- Marginal models (**non-stationary**, extreme) $f_{X_1|\Theta}(x_1|\theta), f_{X_2|\Theta}(x_2|\theta)$
- Multivariate model on **standard** marginal scale (**stationary**, “extreme”) $c(u_1, u_2)$

Marginal extremes

- Theory : Beirlant et al. [2004]
- Method : Dey and Yan [2016]

Generalised extreme value distribution

- F_X^n is the distribution of the maximum of n independent draws of X
- If F_X^n “looks like” $F_X^{n'}$, we say F_X is **max-stable**
- More formally, F_X is max-stable if there exist sequences of constants $a_n > 0, b_n$, and **non-degenerate** G_ξ such that

$$\lim_{n \rightarrow \infty} F_X^n(a_n x + b_n) = G_\xi(x)$$

- We say $F_X \in D(G_\xi)$ or that F_X lies in the **max-domain of attraction** of G_ξ
- The Fisher–Tippett–Gnedenko theorem states that G_ξ is the generalised extreme value distribution with parameter ξ

$$G_\xi(y) = \exp\left(-(1 + \xi y)^{-1/\xi}\right), \quad \xi \in \mathbb{R}$$

- For sufficiently large n , it makes sense to model **block maxima** of n independent identically-distributed draws of X using G_ξ (with $(x - \mu)/\sigma$ in place of y above)

Generalised Pareto distribution

- Now suppose we have an exceedance X of high threshold $\psi \in \mathbb{R}$
- The Pickands-Balkema-De Haan theorem states

$$\begin{aligned}\lim_{\psi \rightarrow \infty} \mathbb{P}[X \leq x | X > \psi] &= \lim_{\psi \rightarrow \infty} \frac{F_X(x)}{1 - F_X(\psi)} \\ &= \text{GP}(x | \xi, \sigma, \psi) \\ &= 1 - \left(1 + \frac{\xi}{\sigma}(x - \psi)\right)_+^{-1/\xi}, \quad \sigma > 0, \quad \xi \in \mathbb{R}\end{aligned}$$

Theory

- Derived from **max-stability** of F_X
- Threshold-stability property
- "Poisson \times GP = GEV"

Practicalities

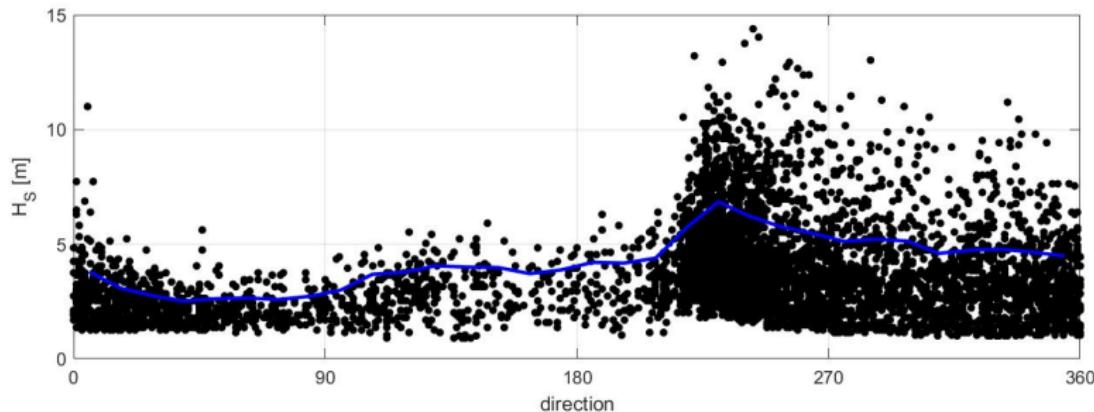
- How to isolate independent threshold exceedances from observed time-series?
- How to specify extreme threshold ψ ?
- ξ, σ, ψ functions of covariates
- Davison and Smith [1990]

Marginal extremes in practice

- Motivation : Chavez-Demoulin and Davison [2005]
- Practicalities : Jonathan and Ewans [2013], Feld et al. [2019]
- Semi-parametric : Randell et al. [2016], Zanini et al. [2020]
- ... lots more
- Non-stationary marginal extremes

Motivation

- Environmental extremes vary smoothly with multidimensional covariates
- Generic modelling framework for different covariate representations
- Statistical and computational efficiency for **n-D** covariates
- Full (Bayesian or **Bayes-Price?**) uncertainty quantification



Typical data for northern North Sea. Storm peak H_S on direction, with $\tau = 0.8$ extreme value threshold.
Rate and size of occurrence varies with direction.

Model for size of occurrence

- Sample of **storm peaks** Y over threshold $\psi_\theta \in \mathbb{R}$, with **1-D covariate** $\theta \in \mathcal{D}_\theta$
- Extreme value threshold ψ_θ **assumed known**
- Y assumed to follow generalised Pareto distribution with shape ξ_θ , (modified) scale ν_θ

$$f_{\text{GP}}(y|\xi_\theta, \nu_\theta) = \frac{1}{\sigma_\theta} \left(1 + \frac{\xi_\theta}{\sigma_\theta} (y - \psi_\theta) \right)_+^{-1/\xi_\theta - 1} \quad \text{with } \nu_\theta = \sigma_\theta(1 + \xi_\theta)$$

- Shape parameter $\xi_\theta \in \mathbb{R}$ and scale parameter $\nu_\theta > 0$
- Non-stationary Poisson model for rate of occurrence, with rate $\rho_\theta \geq 0$

Covariate representations

- Index set $\mathcal{I}_\theta = \{\theta_s\}_{s=1}^m$ on **periodic** covariate domain \mathcal{D}_θ
- Each observation belongs to exactly one θ_s

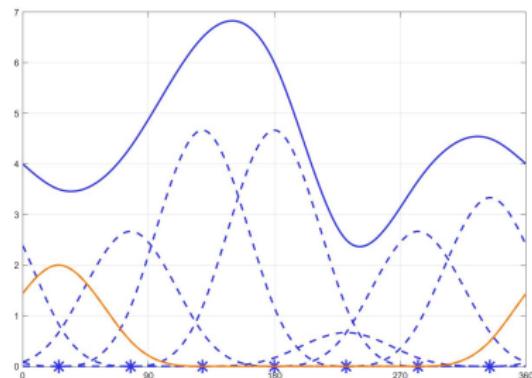
- On \mathcal{I}_θ , assume

$$\begin{aligned}\eta_s &= \sum_{k=1}^n B_{sk} \beta_k, s = 1, 2, \dots, m, \text{ or} \\ \boldsymbol{\eta} &= \mathbf{B} \boldsymbol{\beta}\end{aligned}$$

- $\eta \in (\xi, \nu)$ (and similar for ρ)
- $\mathbf{B} = \{B_{sk}\}_{s=1;k=1}^{m;n}$ basis for \mathcal{D}_θ
- $\boldsymbol{\beta} = \{\beta_k\}_{k=1}^n$ basis coefficients
- Inference reduces to estimating $n_\xi, n_\nu, \mathbf{B}_\xi, \mathbf{B}_\nu, \boldsymbol{\beta}_\xi, \boldsymbol{\beta}_\nu$ (and roughnesses λ_ξ, λ_ν)
- P-splines, BARS and Voronoi are different forms of \mathbf{B}

P-splines

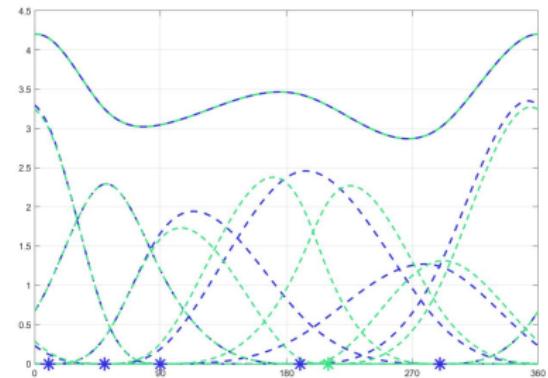
- n regularly-spaced knots on \mathcal{D}_θ
- B consists of n B-spline bases
 - Order d
 - Each using $d + 1$ consecutive knot locations
 - **Local support**
 - Wrapped on \mathcal{D}_θ
 - Cox - de Boor recursion formula
- n is fixed and “over-specified”
- Knot locations $\{r_k\}_{k=1}^n$ fixed
- Local roughness λ of β penalised



Periodic P-splines

BARS basis

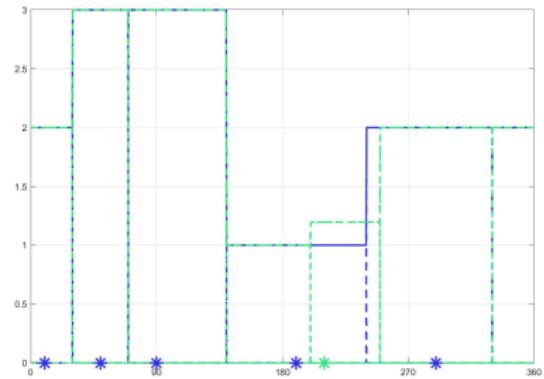
- n irregularly-spaced knots on \mathcal{D}_θ
- B consists of n B-spline bases
- Knot locations $\{r_k\}_{k=1}^n$ can change
- Number of knots n can change



Periodic BARS knot birth and death

Voronoi partition

- n irregularly-spaced centroids on \mathcal{D}_θ
 - Define n neighbourhoods or “cells”
- B consists of n basis functions
 - Piecewise constant on \mathcal{D}_θ
 - = 1 “within cell”, = 0 “outside”
- Centroid locations $\{r_k\}_{k=1}^n$ can change
- Number of centroids n can change
- Trivial extension to n-D



Periodic Voronoi centroid birth and death

Prior for β (all representations)

$$\text{prior density of } \beta \propto \exp\left(-\frac{1}{2}\beta' P \beta\right)$$

- $P = \lambda D'D$, D is a $n \times n$ (wrapped) differencing matrix
- P-splines: D represents first-difference; prior equivalent to local roughness penalty
- BARS and Voronoi: D is I_n ; prior is “ridge-type” for Bayesian regression

Prior for λ (all representations)

$$\lambda \sim \text{Gamma}$$

Prior for n (BARS and Voronoi)

$$n \sim \text{Poisson}$$

Prior for $r_k, k = 1, 2, \dots, n$ (BARS and Voronoi)

$$r_k \sim \text{Uniform}$$

Inference for GP

Parameter set Ω

- P-splines: $\Omega = \{\beta_\xi, \lambda_\xi, \beta_\nu, \lambda_\nu\}$ with n_ξ, r_ξ, n_ν and r_ν pre-specified
- BARS and Voronoi: $\Omega = \{n_\xi, r_\xi, \beta_\xi, \lambda_\xi, n_\nu, r_\nu, \beta_\nu, \lambda_\nu\}$
- $\mathbf{r} = \{r_k\}_{k=1}^n, \boldsymbol{\beta} = \{\beta_k\}_{k=1}^n$

Inference

- Gibbs sampling when full conditionals available
- Otherwise Metropolis-Hastings (MH) within Gibbs, using suitable proposal mechanisms, mMALA where possible
- Reversible-jump for n, r (satisfy dimension-jumping **detailed balance**)

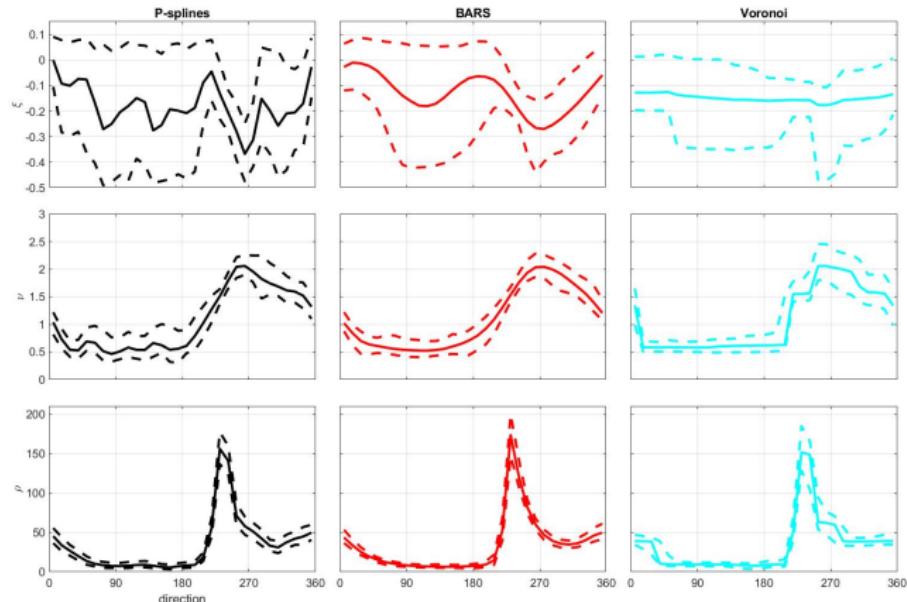
Basic conditional structure for non-dimension-jumping

$$\begin{aligned} f(\boldsymbol{\beta}_\eta | \mathbf{y}, \Omega \setminus \boldsymbol{\beta}_\eta) &\propto f(\mathbf{y} | \boldsymbol{\beta}_\eta, \Omega \setminus \boldsymbol{\beta}_\eta) \times f(\boldsymbol{\beta}_\eta | \lambda_\eta) \\ f(\lambda_\eta | \mathbf{y}, \Omega \setminus \lambda_\eta) &\propto f(\boldsymbol{\beta}_\eta | \lambda_\eta) \times f(\lambda_\eta) \\ f(\mathbf{r}_\eta | \mathbf{y}, \Omega \setminus \mathbf{r}_\eta) &\propto f(\mathbf{y} | \mathbf{r}_\eta, \Omega \setminus \mathbf{r}_\eta) \times f(\mathbf{r}_\eta), \end{aligned}$$

- $\eta \in (\xi, \nu)$ (and ρ)

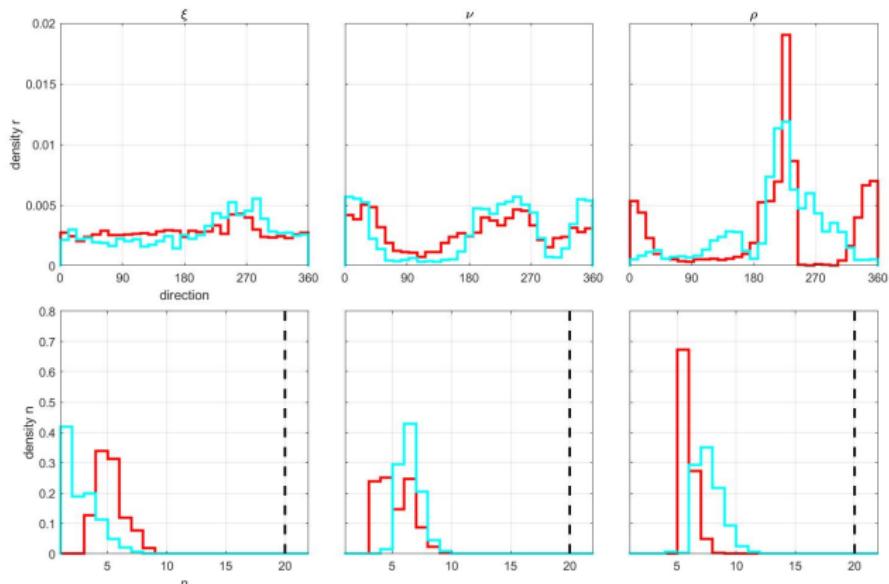
Posterior parameter estimates for ξ , ν and ρ for northern North Sea

- Note colour scheme
- Rate ρ and ν very similar
- Voronoi gives almost constant ξ
- Voronoi piecewise constant
- Land shadow effects
- General agreement

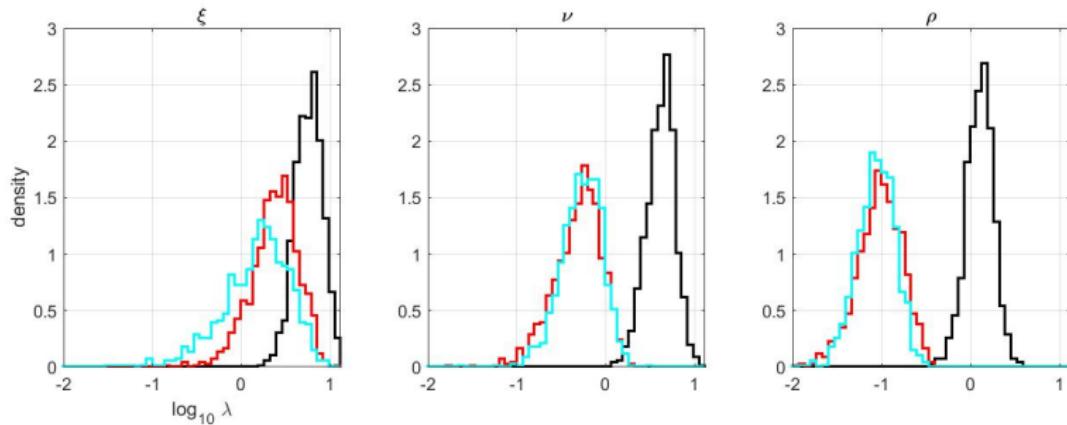


Posterior densities for locations r and numbers n

- Prior uniform knot placement for r
- Knot placement uniform for ξ , clear effect for ρ
- n close to 1 for Voronoi ξ
- General agreement
- Effect of different priors on n checked



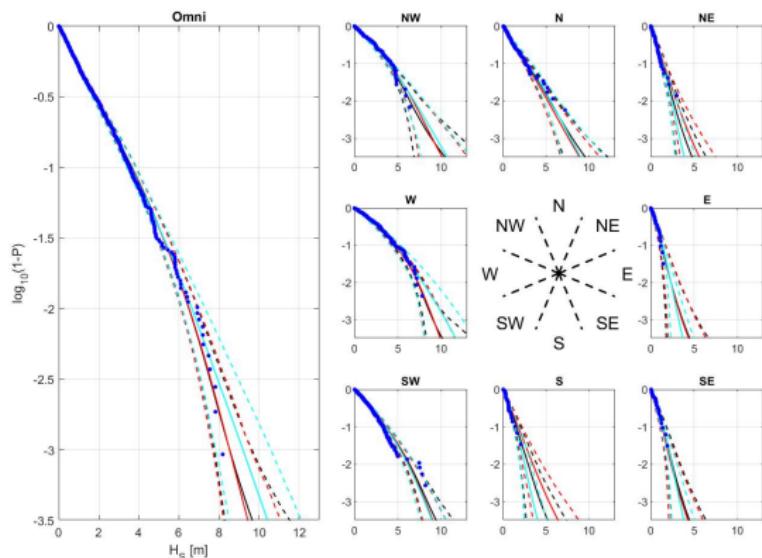
Posterior densities for penalty coefficients λ



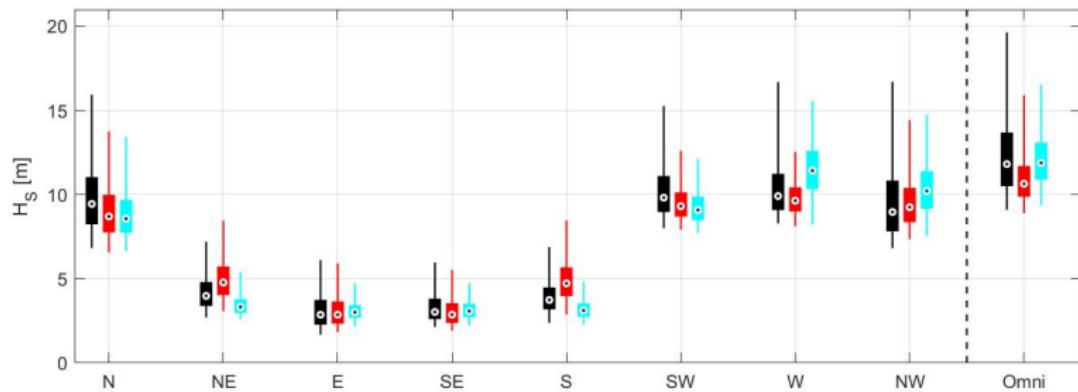
- Prior density is $\text{Gamma}(1,1)$ ($f(x) \propto \exp(-x)$, $x \geq 0$)
- Ridge penalties for BARS and Voronoi, but roughness for P-splines
- λ somewhat lower for Voronoi, but also this has smaller n
- General consistency

Fit diagnostic

- Empirical tail (blue)
- Posterior means and 95% credible intervals for quantile levels from different models
- General consistency



Directional posterior predictive distribution of $T = 1000$ -year maximum



- Box-whiskers with 2.5%, 25%, 50%, 75% and 97.5% percentiles
- Uncertainties larger for P-splines?
- General consistency
- This is more-or-less what the engineer needs to design a “compliant” structure

Multivariate extremes

- Theory : Beirlant et al. [2004]
- Copulas : Joe [2014]
- Method : Dey and Yan [2016]

Multivariate extreme value distribution, MEVD

- $\mathbf{X}_i = (X_{i1}, \dots, X_{ij}, \dots, X_{ip})$, $i = 1, \dots, n$ iid p -vectors, distribution F
- $M_{n,j} = \max_i X_{ij}$, component-wise maximum
- Then for $Z_{n,j} = (M_{n,j} - b_{n,j})/a_{n,j}$, normalised with scaling constants:

$$\mathbb{P}(\mathbf{Z} \leq \mathbf{z}) = F^n(a_n \mathbf{z} + \mathbf{b}_n) \rightarrow G(\mathbf{z}) \quad \text{as } n \rightarrow \infty$$

- Non-degenerate $G(\mathbf{z})$ must be max-stable, so $\forall k \in \mathbb{N}, \exists \boldsymbol{\alpha}_k > \mathbf{0}, \beta_k$ s.t.

$$G^k(\boldsymbol{\alpha}_k \mathbf{z} + \beta_k) = G(\mathbf{z})$$

- We say $F \in D(G)$
- Margins G_1, \dots, G_p are unique GEV, but $G(\mathbf{z})$ is not unique
- The component-wise maximum is not “observed” (especially as $n \rightarrow \infty$)

MEVD on common margins

- On uniform margins, we have **extreme value copula**: $C(\mathbf{u}) = C^k(\mathbf{u}^{1/k})$
- On standard Fréchet margins ($G_j(z) = \exp(-z^{-1})$), with pseudo-polars (r, \mathbf{w})

$$\begin{aligned} G(\mathbf{z}) &= \exp(-V(\mathbf{z})), \quad \text{for } \text{exponent measure } V \\ \text{with } V(\mathbf{z}) &= \int_{\Delta} \max_j \left\{ \frac{w_j}{z_j} \right\} S(d\mathbf{w}), \quad \text{on } \Delta = \{\mathbf{w} \in \mathbb{R}^p : \|\mathbf{w}\| = 1\} \\ \text{and } 1 &= \int_{\Delta} w_j S(d\mathbf{w}), \quad \forall j, \text{ for } \text{angular measure } S \end{aligned}$$

- Max-stability : $V(r\mathbf{z}) = r^{-1}V(\mathbf{z})$, homogeneity order -1
- Rich spatial extensions to **max-stable processes**, MSPs
- Multivariate generalised Pareto distribution, MGPD
- Condition of **multivariate regular variation**, MRV

$$\frac{1 - F(t\mathbf{x})}{1 - F(t\mathbf{1})} \rightarrow \lambda(\mathbf{x}) \text{ as } t \rightarrow \infty, \mathbf{x} \in \mathbb{R}^p$$

useful to prove that $F \in D(G)$ for some MEVD G

Extremal dependence (2D, uniform margins)

$$\chi(u) = \frac{\mathbb{P}(U > u, V > u)}{\mathbb{P}(U > u)} = \frac{\bar{C}(u, u)}{1 - u} \rightarrow \chi \text{ as } u \rightarrow 1$$

- $\chi = 1$ perfect dependence
- $\chi \in (0, 1)$ asymptotic dependence, AD
- $\chi = 0$ perfect independence

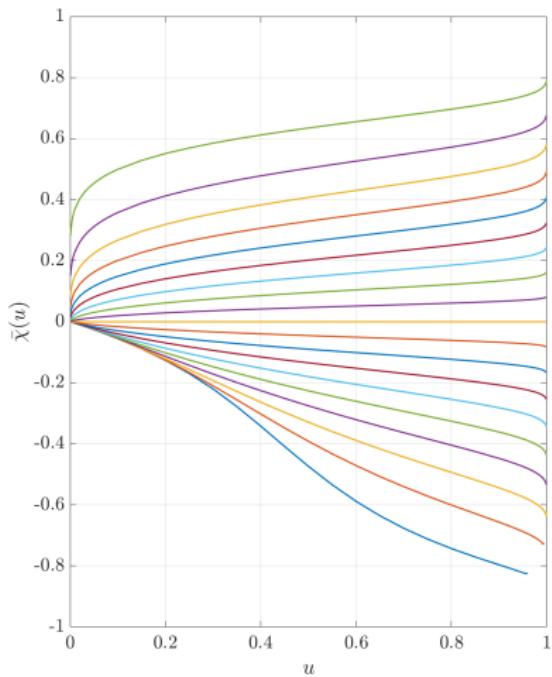
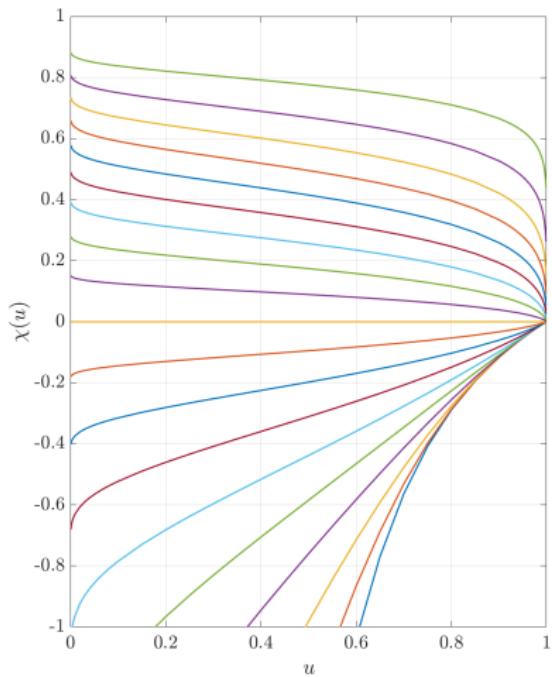
$$\bar{\chi}(u) = 2 \frac{\log \mathbb{P}(U > u)}{\log \mathbb{P}(U > u, V > u)} - 1 = 2 \frac{\log(1 - u)}{\log \bar{C}(u, u)} - 1 \rightarrow \bar{\chi} \text{ as } u \rightarrow 1$$

- $\bar{\chi} = 1$ perfect dependence and AD
- $\bar{\chi} \in (0, 1)$ asymptotic independence, AI
- $\bar{\chi} = 0$ perfect independence
- See η for motivation

$$\theta(u) = \frac{\log \mathbb{P}(U \leq u, V \leq u)}{\log \mathbb{P}(U \leq u)} = \frac{\log C(u, u)}{\log u} \rightarrow \theta \text{ as } u \rightarrow 1$$

- $\theta = 2 - \chi$
- MEVDs do not admit asymptotic independence

Extremal dependence (bivariate Gaussian)



$\chi(u)$ and $\bar{\chi}(u)$ for bivariate Gaussian ($\Rightarrow \chi = 0, \bar{\chi} = \rho$)
 Colours are correlations ρ on $-0.9, -0.8, \dots, 0.9$
 (Recreated from Coles et al. 1999)

Beyond component-wise maxima

- Many (almost all?) environmental extremes problems involve asymptotic independence, at least in part
- Need to move away from MEVDs
- On Fréchet margins ($F(z) = \exp(-z^{-1})$), assume

$$\frac{\mathbb{P}(Z_1 > z, Z_2 > z)}{(\mathbb{P}(Z_1 > z))^{1/\eta}} = \mathcal{L}(z)$$

where \mathcal{L} is slowly varying : $\mathcal{L}(xz)/\mathcal{L}(z) \rightarrow 1$ as $z \rightarrow \infty$

- $\bar{\chi} = 2\eta - 1$
- Ledford and Tawn [1996], Ledford and Tawn [1997]
- e.g. use non-extreme value copulas or inverted EV copulas
- $\mathbb{P}(Z_2 > z | Z_1 > z) \approx Cz^{1-1/\eta}$ from above
- Idea: assume a max-stable-like normalisation for **conditional extremes**

Conditional extremes

- $\mathbf{X} = (X_1, \dots, X_j, \dots, X_p)$
- Each X and Y have standard Laplace margins ($f(x) = \exp(-|x|)/2, x \in \mathbb{R}$)
- Seek a model for $\mathbf{X}|(Y = y)$ for $y > u$
- Assume we can find p -dimensional scaling $\mathbf{a}, \mathbf{b} > \mathbf{0}$ such that

$$\begin{aligned}\mathbb{P}(\mathbf{Z} \leq \mathbf{z} | Y = y) &\rightarrow G(\mathbf{z}) \quad \text{as} \quad u \rightarrow \infty \\ \text{for } \mathbf{Z} &= \frac{\mathbf{X} - \mathbf{a}(y)}{\mathbf{b}(y)}\end{aligned}$$

- Non-degenerate G is unknown, and estimated empirically
- Typical scaling is $\mathbf{a} = \alpha y$ and $\mathbf{b} = y^\beta$, $\alpha \in [-1, 1]^p$, $\beta \in (-\infty, 1]^p$
- So simply fit regression model

$$\mathbf{X}|(Y = y) = \alpha y + y^\beta \mathbf{Z}$$

- $\alpha = 1, \beta = 0$: perfect dependence and AD, and $\alpha \in (0, 1)$: AI
- Heffernan and Tawn [2004] find choices for α and β for popular bivariate cases
- Bivariate Gaussian : $\alpha = \rho^2$, $\beta = 1/2$

Conditional extremes in practice

- Non-stationary : Jonathan et al. [2014]
- Time-series : Winter and Tawn [2016], Tendijck et al. [2019]
- Mixture model : Tendijck et al. [2021]
- Spatial : Shooter et al. [2021b], Shooter et al. [2021a]
- ... lots more
- **Multivariate spatial** : Shooter et al. [2022]

Multivariate spatial conditional extremes (MSCE)

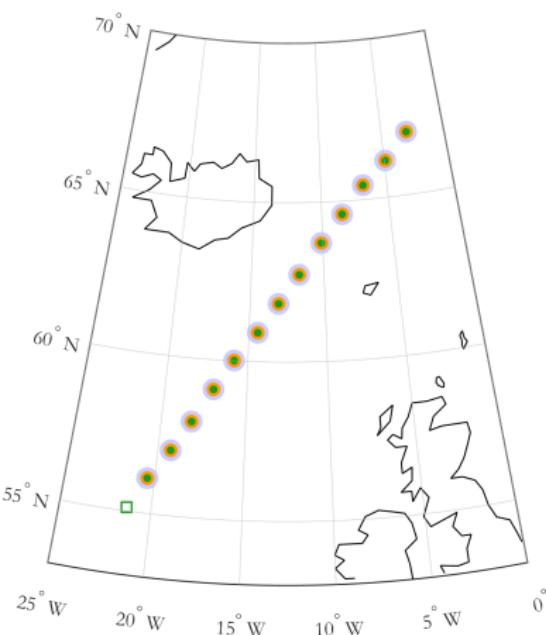
Motivation

- How useful are satellite observations of ocean waves and winds?
- Could they become the primary data source for decisions soon?
- What are the **spatial characteristics of extremes** from satellite observations?

Overview

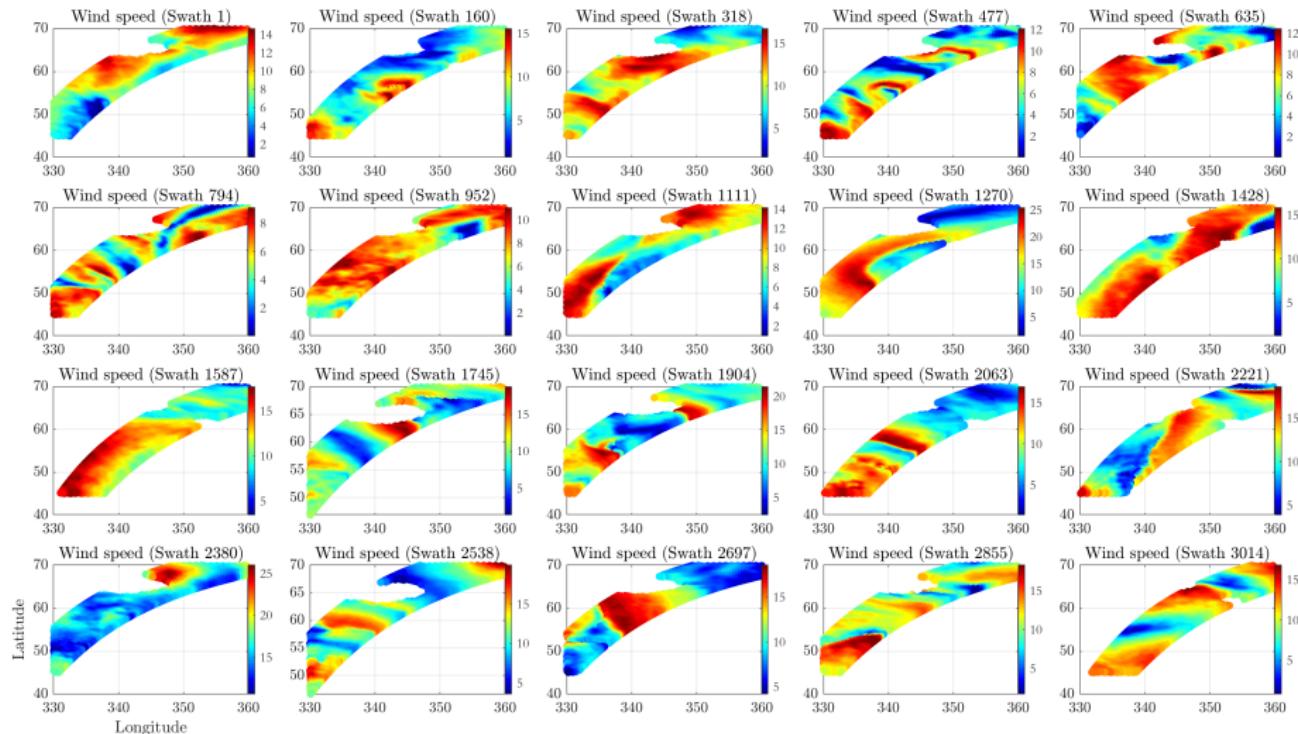
- A look at the data
- Brief overview of methodology
- Results for joint spatial structure of extreme scatterometer wind speed, hindcast wind speed and hindcast significant wave height in the North Atlantic
- Implications for future practical applications

In a nut-shell



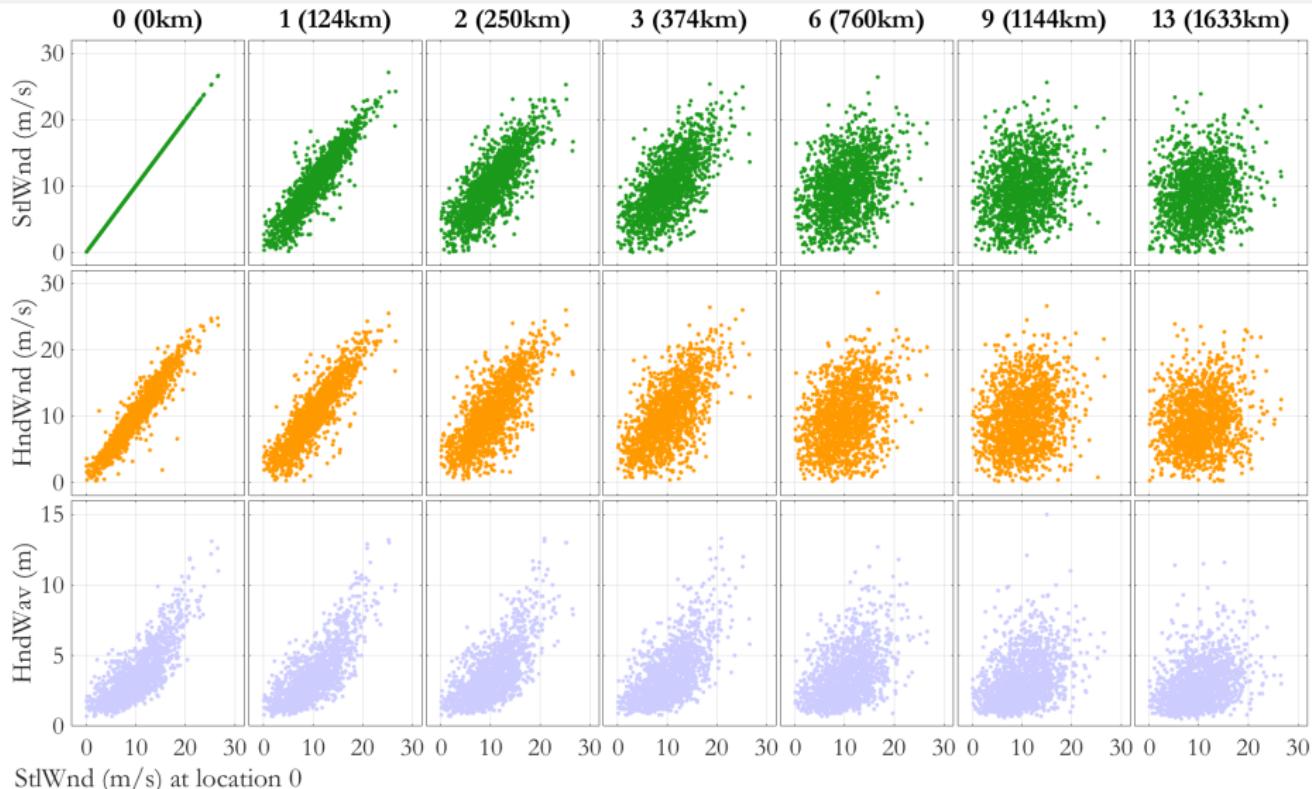
- Condition on **large value** x of first quantity X_{01} at **one location** $j = 0$ (**green square**)
 - Estimate “conditional spatial profiles” for $m > 1$ quantities $\{X_{jk}\}_{j=1,k=1}^{p,m}$ at $p > 0$ **other locations** (**green**, **orange** and **blue** circles)
- $$X_{jk} \sim Lpl$$
- $$x > u$$
- $$X| \{X_{01} = x\} = \alpha x + x^\beta Z$$
- $$Z \sim DL(\mu, \sigma^2, \delta; \Sigma(\lambda, \rho, \kappa))$$
- MCMC to estimate $\alpha, \beta, \mu, \sigma, \delta$ and ρ, κ, λ
 - $\alpha, \beta, \mu, \sigma, \delta$ spatially smooth for each quantity
 - Residual correlation Σ for conditional Gaussian field, powered-exponential decay with distance

Swath wind speeds



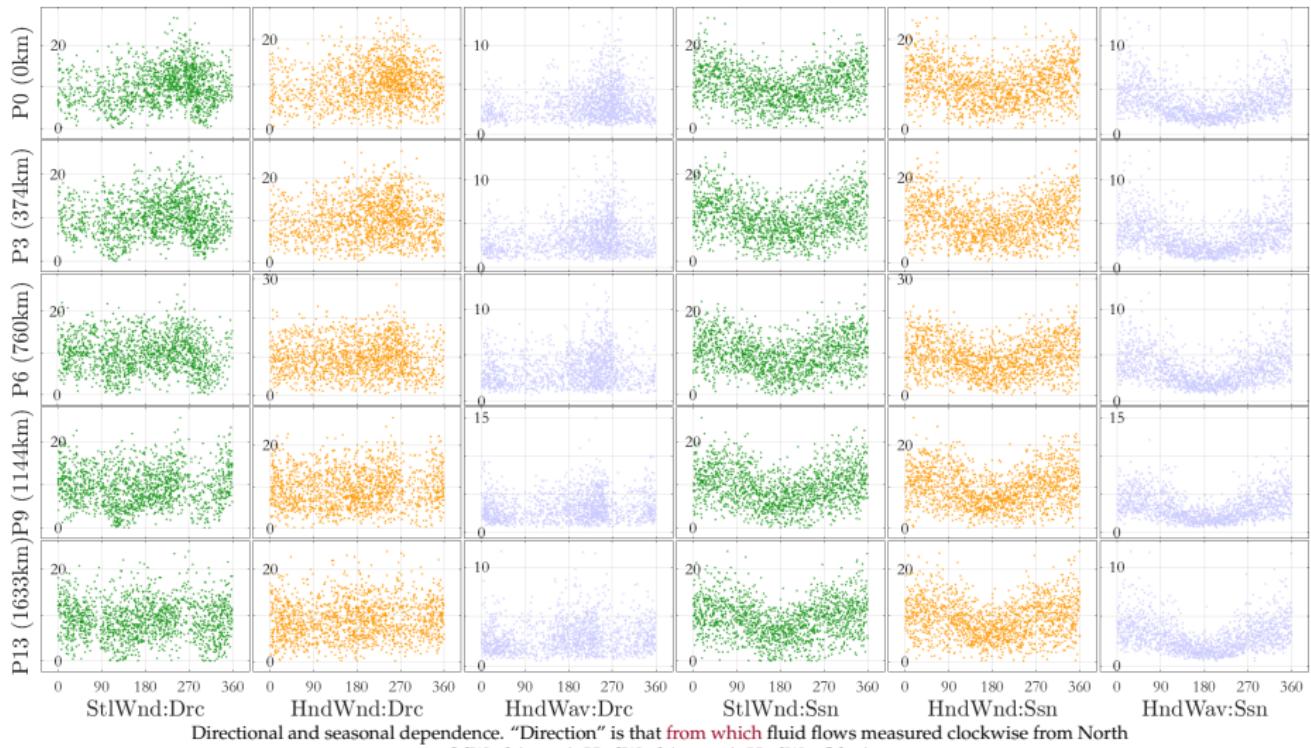
Daily descending METOP swaths. Satellite swath location changes over time. Spatial structure evident

Scatter plots on physical scale



Scatter plots of registered data : StlWnd (green), HndWnd (orange), HndWav(blue)

Covariate dependence

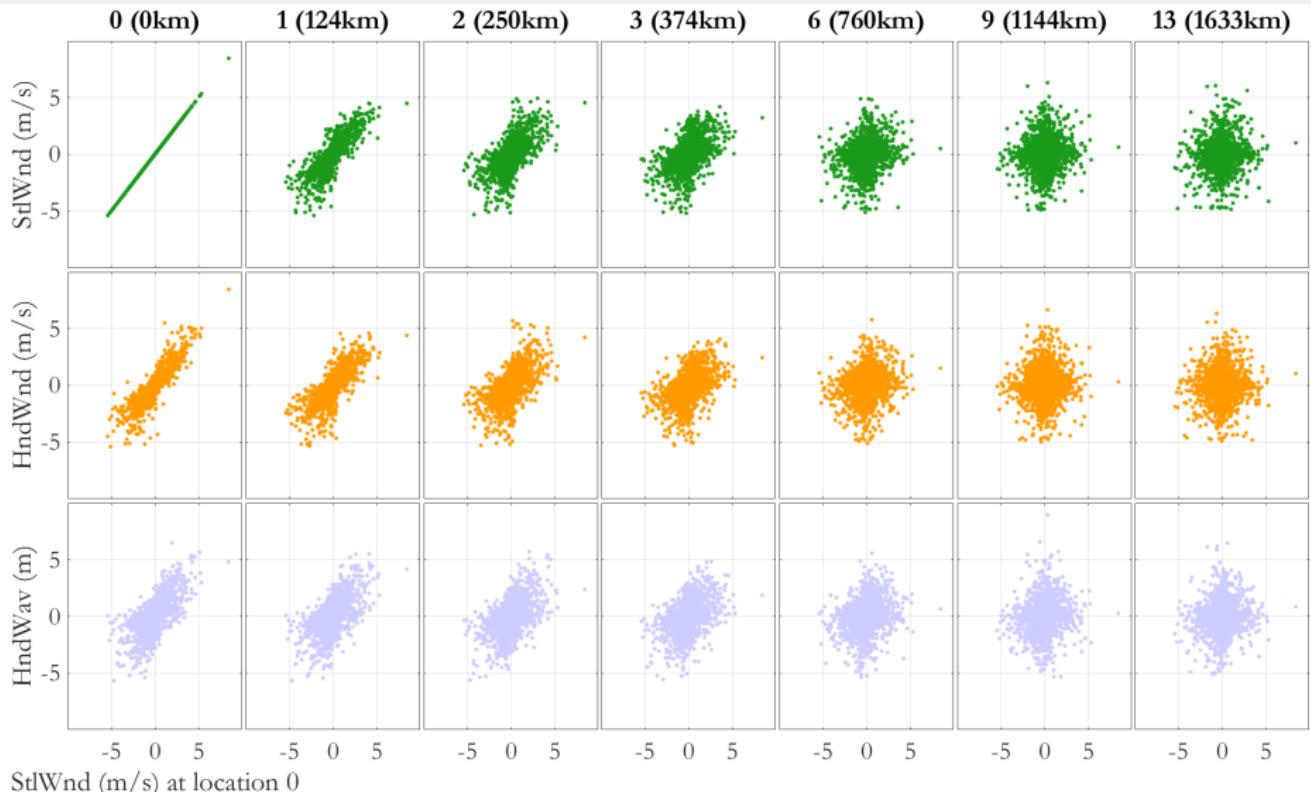


Marginal transformation to standard Laplace scale

Procedure

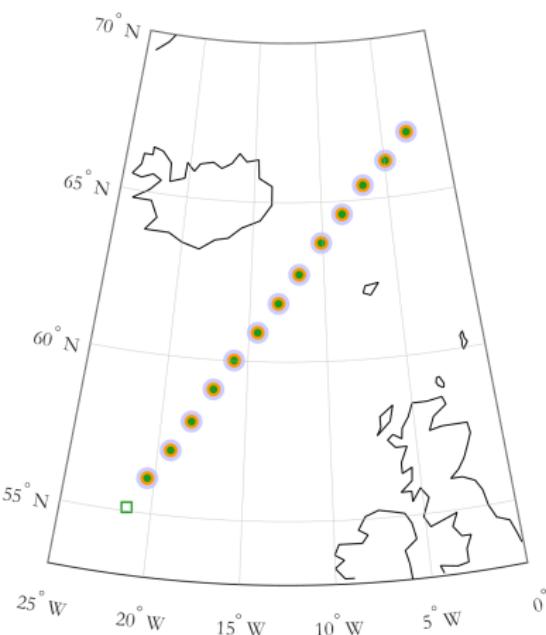
- Non-stationary piecewise constant **directional-seasonal marginal extreme value model**
- Pre-specified 8 directional bins (“octants”) of equal width centred on cardinal and semi-cardinal directions
- Pre-specified “summer” and “winter” seasonal bins
- Generalised Pareto model for peaks over threshold
- Model parameters vary smoothly between bins, optimal roughness found using cross-validation
- Multiple extreme value thresholds with non-exceedance probabilities between 0.7 and 0.9 considered
- Bootstrapping for uncertainties
- **Uncertainty in marginal model not propagated**
- Independent marginal models for pair of variable (**St1Wnd**, **HndWnd**, **HndWav**) and location (0,1,...,13)
- Software : github.com/ECSADES/ecsades-matlab

Scatter plots on Laplace scale



Registered data on Laplace scale: StlWnd (green), HndWnd (orange), HndWav(blue)

In a nut-shell



- Condition on **large value** x of first quantity X_{01} at **one location** $j = 0$ (**green square**)
 - Estimate “conditional spatial profiles” for $m > 1$ quantities $\{X_{jk}\}_{j=1,k=1}^{p,m}$ at $p > 0$ **other locations** (**green**, **orange** and **blue** circles)
- $$X_{jk} \sim Lpl$$
- $$x > u$$
- $$X| \{X_{01} = x\} = \alpha x + x^\beta Z$$
- $$Z \sim DL(\mu, \sigma^2, \delta; \Sigma(\lambda, \rho, \kappa))$$
- MCMC to estimate $\alpha, \beta, \mu, \sigma, \delta$ and ρ, κ, λ
 - $\alpha, \beta, \mu, \sigma, \delta$ spatially smooth for each quantity
 - Residual correlation Σ for conditional Gaussian field, powered-exponential decay with distance

Inference

- Delta-Laplace residual margins

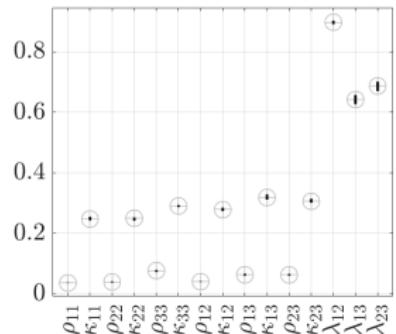
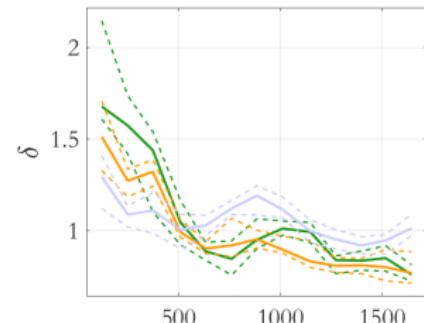
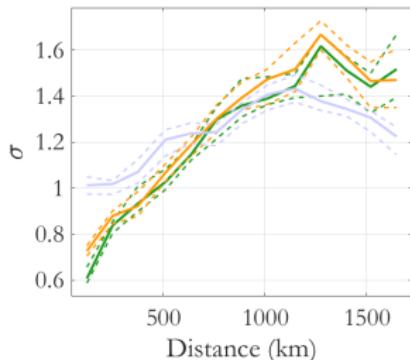
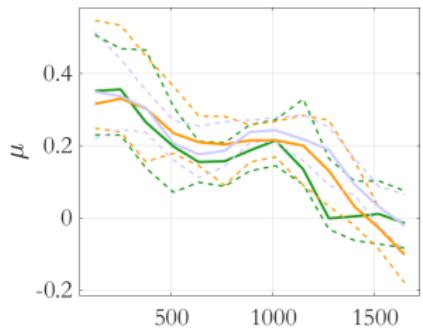
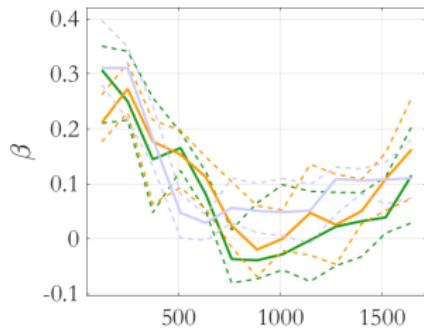
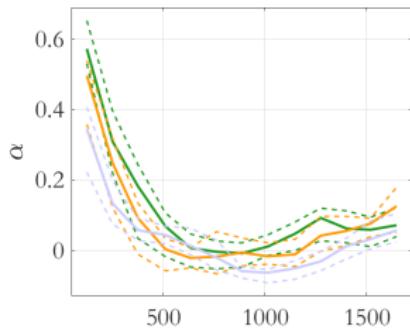
$$f_{Z_{j,k}}(z_{j,k}) = \frac{\delta_{j,k}}{2\kappa_{j,k}\sigma_{j,k}\Gamma\left(\frac{1}{\delta_{j,k}}\right)} \exp\left\{-\left|\frac{z - \mu_{j,k}}{\kappa_{j,k}\sigma_{j,k}}\right|^{\delta_{j,k}}\right\}, \quad \kappa_{j,k}^2 = \Gamma\left(1/\delta_{j,k}\right)/\Gamma\left(3/\delta_{j,k}\right)$$

- Gaussian residual dependence

$$\Sigma_{\mathcal{A}^*(j,k)\mathcal{A}^*(j',k')}^* = \lambda_{k,k'}^{|k-k'|} \exp\left(-\left(\frac{\text{dist}(r_j, r_{j'})}{\rho_{k,k'}}\right)^{\kappa_{k,k'}}\right)$$

- Piecewise linear forms for all parameters with distance using n_{Nod} spatial nodes
- Adaptive MCMC, Roberts and Rosenthal [2009]
- Total of $m(5n_{\text{Nod}} + (3m + 1)/2)$ parameters
- Rapid convergence, 10k iterations sufficient

Parameter estimates



Estimates for α , β , μ , σ and δ with distance, and residual process estimates ρ , κ and λ . Model fitted with $\tau = 0.75$

StlWnd (green), HndWnd (orange), HndWav(blue)

Residual Gaussian field : ρ =scale (need to $\times 100$ km), κ =exponent (need to $\times 5$), λ =cross-correlation

Summary

Why?

- Careful quantification of “rare-event” risk
- Characterise tails of (multivariate) distributions
- Limited observations
- Immediate real-world consequences

The next 10 years?

- Univariate : fuller covariate descriptions, exploit measurement scale / sub-asymptotics, UQ, provide real-world decision-support
- Multivariate : theoretical development, computational tractability, expansion in scope (time-series, spatial), serious real-world applications

An interesting field for research?

- Environmental extremes is a nice area if you like a mix of statistical theory, method, computation and serious physical science-based application

References

- J. Beirlant, Y. Goegebeur, J. Segers, and J. Teugels. *Statistics of extremes: theory and applications*. Wiley, Chichester, UK, 2004.
- V. Chavez-Demoulin and A.C. Davison. Generalized additive modelling of sample extremes. *J. Roy. Statist. Soc. Series C: Applied Statistics*, 54:207–222, 2005.
- S Coles, J Heffernan, and J Tawn. Dependence measures for extreme value analyses. *Extremes*, 2:339–365, 1999.
- A.C. Davison and R. L. Smith. Models for exceedances over high thresholds. *J. R. Statist. Soc. B*, 52:393, 1990.
- D. Dey and J. Yan, editors. *Extreme value modeling and risk analysis: methods and applications*. CRC Press, Boca Raton, USA, 2016.
- G. Feld, D. Randell, E. Ross, and P. Jonathan. Design conditions for waves and water levels using extreme value analysis with covariates. *Ocean Eng.*, 173: 851–866, 2019.
- J. E. Heffernan and J. A. Tawn. A conditional approach for multivariate extreme values. *J. R. Statist. Soc. B*, 66:497–546, 2004.
- H. Joe. *Dependence modelling with copulas*. CRC Press, 2014.
- P. Jonathan and K. C. Ewans. Statistical modelling of extreme ocean environments with implications for marine design : a review. *Ocean Eng.*, 62:91–109, 2013.
- P. Jonathan, K. C. Ewans, and D. Randell. Non-stationary conditional extremes of northern North Sea storm characteristics. *Environmetrics*, 25:172–188, 2014.
- A. W. Ledford and J. A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83:169–187, 1996.
- A. W. Ledford and J. A. Tawn. Modelling dependence within joint tail regions. *J. R. Statist. Soc. B*, 59:475–499, 1997.
- D. Randell, K. Turnbull, K. Ewans, and P. Jonathan. Bayesian inference for non-stationary marginal extremes. *Environmetrics*, 27:439–450, 2016.
- G. O. Roberts and J. S. Rosenthal. Examples of adaptive MCMC. *J. Comp. Graph. Stat.*, 18:349–367, 2009.
- R. Shooter, E. Ross, A. Ribal, I. R. Young, and P. Jonathan. Spatial conditional extremes for significant wave height from satellite altimetry. *Environmetrics*, 32: e2674, 2021a.
- R. Shooter, J. A. Tawn, E. Ross, and P. Jonathan. Basin-wide spatial conditional extremes for severe ocean storms. *Extremes*, 24:241–265, 2021b.
- R. Shooter, E. Ross, A. Ribal, I. R. Young, and P. Jonathan. Multivariate spatial conditional extremes for extreme ocean environments. *Ocean Eng.*, 247:110647, 2022.
- S. Tendijck, E. Ross, D. Randell, and P. Jonathan. A non-stationary statistical model for the evolution of extreme storm events. *Environmetrics*, 30:e2541, 2019.
- S. Tendijck, E. Eastoe, J. Tawn, D. Randell, and P. Jonathan. Modeling the extremes of bivariate mixture distributions with application to oceanographic data. *J. Am. Statist. Soc.*, 2021. doi: 10.1080/01621459.2021.1996379.
- H. C. Winter and J. A. Tawn. Modelling heatwaves in central France: a case-study in extremal dependence. *J. Roy. Statist. Soc. C*, 65:345–365, 2016.
- E. Zanini, E. Eastoe, M. Jones, D. Randell, and P. Jonathan. Covariate representations for non-stationary extremes. *Environmetrics*, 31:e2624, 2020.

Thanks for listening / Diolch am wrando!

