Extremes for metocean: Covariates

Philip Jonathan

Lancaster University

IOGP, London, September 2026
(slides at *www.lancs.ac.uk/∼jonathan*)

Overview of masterclass

- A: Fundamentals
- B: Covariates
- C: Multivariate
- D: covXtreme
- E: Recent developments & take-aways

Overview of B: Covariates

- Motivating thought experiment
- Motivating application
- Overview of analysis procedure
  - Frequentist: maximum penalised likelihood, cross-validation, bootstrapping
  - Bayesian: priors, Markov Chain Monte Carlo
- Typical results
- Key issues

Sanity check

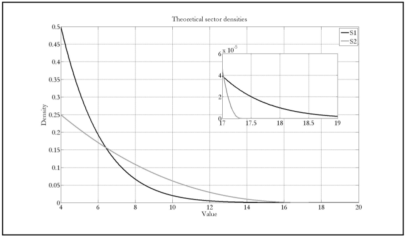
All models are wrong, some models are useful

Consistent balance of engineering and statistical insights

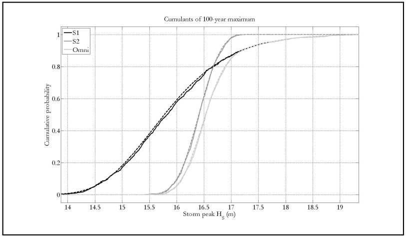
A motivating “toy” thought experiment

- A directional extremes problem
- Two homogeneous 180° directional sectors S_1 and S_2, extremes are GP-distributed in each
- ξ, σ and u values all potentially different between sectors
- Random samples of size 1250 from each sector corresponding to 25 years
- Test
  - H_0: $\xi_1 = \xi_2, \sigma_1 = \sigma_2$
  - H_A: $\xi_1 \neq \xi_2, \sigma_1 \neq \sigma_2$
- We will consider case (ξ_1,σ_1,u_1)=(-0.1,2,4) and (ξ_2,σ_2,u_2)=(-0.3,4,4)
  - See Jonathan et al. [2008] for more

Tail characteristics



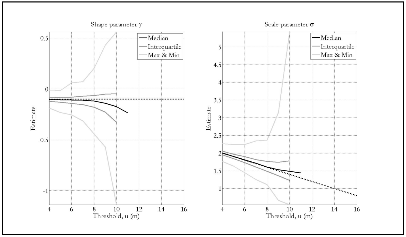
Sector densities (theory)



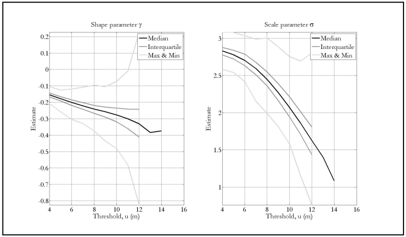
100-year maximum

- S1 (black) dominates far tail, but S2 (grey) prevalent for values $\in [7, 14]$
- Typical sample will have more examples from S2 than S1, causing issues
- 100-year maximum dominated by S2 below 17, and by S1 above

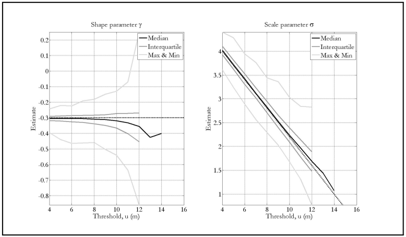
Parameter estimates with threshold



H_A: S_1 estimates with threshold



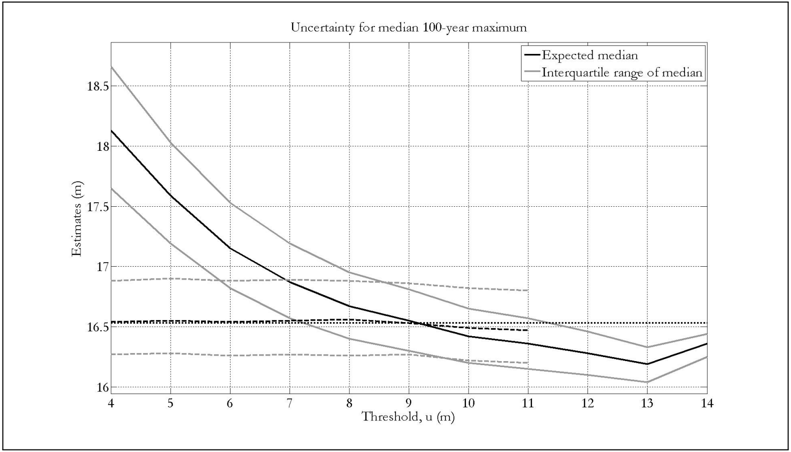
H_0: estimates with threshold



H_A: S_2 estimates with threshold

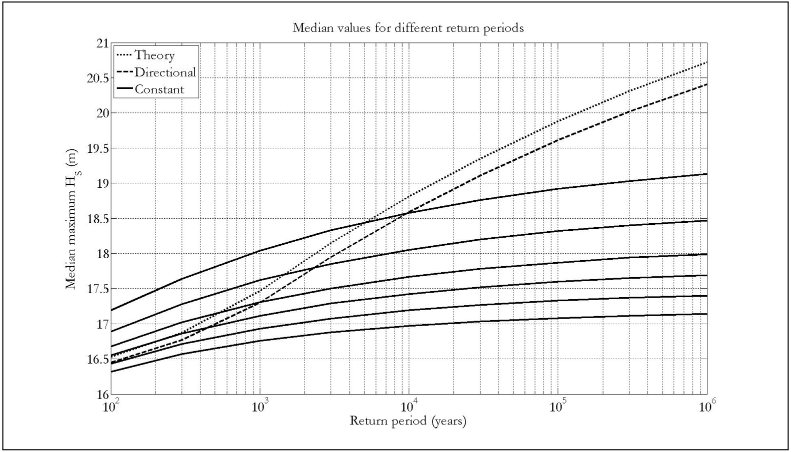
- H_A looks OK in S_1 and S_2
  - $\hat{\xi}$ constant with u, $\hat{\sigma}$ linear
- H_0 suspicious!
  - $\hat{\xi}$ NOT constant with u

Estimated median 100-year maximum



Median 100-year event: true (dot), H_A (dash), H_0 (solid)

Implications for long return periods



Median P-year event: true (dot), H_A (dash),

H_0 (solid, for thresholds 6,7,...,11m)

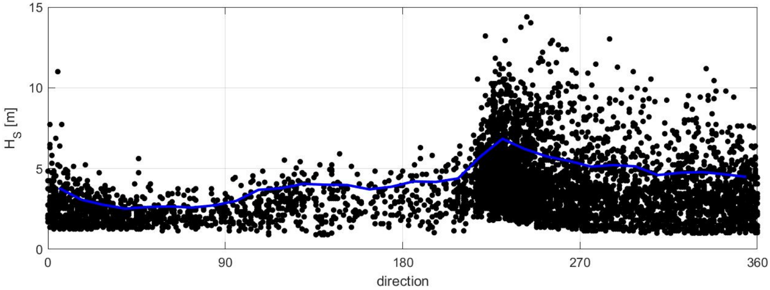
Take-aways from thought experiment

- If we misspecify the extreme value model, results are unpredictable
  - We honestly don't know whether return value estimates will be biased high, low or be unbiased
  - Threshold choice becomes a bigger problem
  - Ignore covariates at your peril in general
- If we are confident that the far tail is dominated by one direction, it makes sense to look at that sector in isolation to estimate the omni-directional return value
  - Provided we have enough data from that direction
- If we want consistent estimates of return values across every directional sector, we must fit an appropriately-specified directional extreme value model

- Similar reasoning for seasonal extremes
  - We need directional-seasonal covariate models for quantities like H_S
  - Similar logic to include other important covariates

Motivating application

- Typical data for northern North Sea
  - Storm peak H_S ($\triangleq H_S^{sp}$) on direction
  - Rate and size of occurrence varies with direction



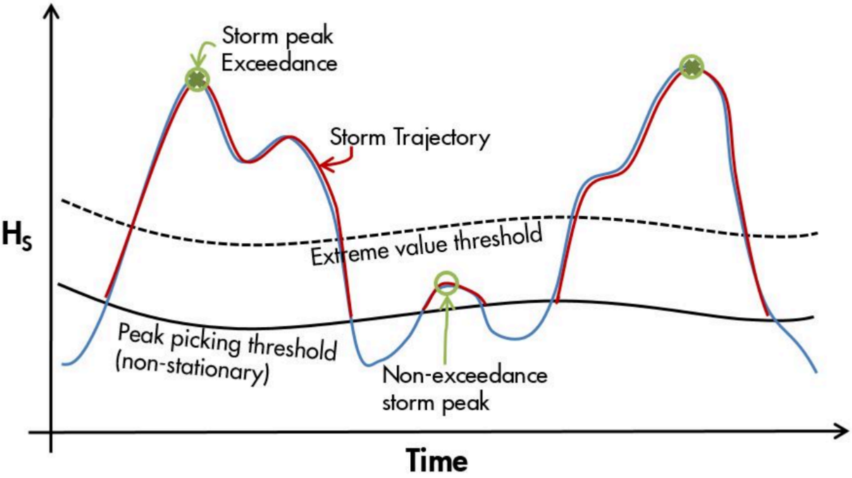
Storm peak H_S^{sp} on direction, with $\tau = 0.8$ extreme value threshold

- Environmental extremes vary smoothly with multidimensional covariates
  - We will use only direction for simplicity here

Full analysis procedure

- Input
  - Time-series of sea state H_S and associated variables (e.g. T_P, surge)
- Preprocessing ("peak picking")
  - Isolate storms
  - Isolate storm peak quantities (H_S^{sp} and associated)
  - Isolate accompanying multivariate storm trajectories
- Key statistical models (all directional)
  - Model for distributional quantiles, used as extreme value thresholds u
  - Model for rate of occurrence of exceedances of threshold u per annum
  - Model for size of occurrence of exceedance of threshold u
- Metocean inferences
  - Use fitted models to estimate return values for H_S^{sp} and associated storm peak quantities
  - Combine simulated storm peak quantities with storm trajectories to estimate return values for sea-state quantities and individual waves

Key quantities



Key variables describing a storm. Peak-picking threshold (solid black), extreme value threshold (dashed black), storm peak (green circle), storm peak over threshold (green star), storm trajectory (red)

Managing computational complexity of parameter variation with direction

- Periodic domain $\mathcal{D}$ ($= [0, 360)^\circ$ say) for direction θ
- Partition directional domain $\mathcal{D}$ into m bins of equal width $d > 0$ and bin centres $\boldsymbol{\theta} = \{\theta_j\}_{j=1}^m$
- Index set $\mathcal{I}$ maps each storm peak observation i onto one of the bins j; so $\theta_{\mathcal{I}(i)}$ is the bin centre direction for observation i
- Any model parameters $\eta(\theta)$ for $\theta \in \mathcal{D}$ are also evaluated at bin centres only, and can be written as a vector $\boldsymbol{\eta} = \{\eta_j\}_{j=1}^m = \{\eta(\theta_j)\}_{j=1}^m$
- To keep things simple, we might set $m = 360$ so that each direction is specified to the nearest degree
    - But this might not always be the wisest choice, for physical and especially computational reasons

A model for extreme value threshold, $u(\theta)$

- Gamma model for full distribution of storm peaks X with direction $\theta \in \mathcal{D}$
- Density for a single storm peak observation x

$$f_{\text{Gmm}}(x|\alpha, \kappa) = \frac{\kappa^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\kappa x)$$

  with parameters shape $\alpha > 0$ and scale (or rate) $\kappa > 0$
- Sample likelihood for storm peak values $\boldsymbol{x} = \{x_i\}_{i=1}^{n}$ with associated directional bins $\{\theta_{\mathcal{I}(i)}\}_{i=1}^{n}$ is

$$\mathcal{L}_{\text{Gmm}}(\boldsymbol{\alpha}, \boldsymbol{\kappa}|\boldsymbol{x}) = \prod_{i=1}^{n} f_{\text{Gmm}}(x_i; \alpha_{\mathcal{I}(i)}, \kappa_{\mathcal{I}(i)})$$

  for parameter vectors $\boldsymbol{\alpha} = \{\alpha_j\}_{j=1}^{m}$ and $\boldsymbol{\kappa} = \{\kappa_j\}_{j=1}^{m}$
- α and κ vary smoothly with direction on $\mathcal{D}$

- Quantiles of the gamma distribution define extreme value thresholds u. We consider model averages over multiple thresholds
- Obvious alternative approach is quantile regression

A model for expected annual rate $\rho(\theta)$ of threshold exceedance

- Poisson model for annual rate $\rho(\theta)$ of threshold exceedance for $\theta \in \mathcal{D}$
- Count the number C of storm peak threshold exceedance counts $\boldsymbol{c} = \{c_j\}_{j=1}^m$ in the directional bins
- Density for a single directional bin j

$$f_{\text{Pss}}(c_j|\rho_j) = \frac{(d\rho_j)^{c_j}}{c_j!} \exp\{-d\rho_j\}$$

  with rate $\rho \geq 0$, where $d\rho_j$ is the expected number of threshold exceedances in the bin
- Sample likelihood

$$\mathcal{L}_{\text{Pss}}(\boldsymbol{\rho}|\boldsymbol{c}) = \prod_{j=1}^{m} f_{\text{Pss}}(c_j|\rho_j)$$

  for parameter vector $\boldsymbol{\rho} = \{\rho_j\}_{j=1}^m$
- ρ varies smoothly with direction on $\mathcal{D}$

A model for the size X' of threshold exceedances

- GP model for storm peaks X' over threshold $u(\theta)$ with direction $\theta \in \mathcal{D}$
- Density for a single storm peak threshold exceedance x'

$$f_{\mathrm{GP}}(x'|\xi,\nu,u) = \frac{1}{\sigma}\left(1+\frac{\xi}{\sigma}\left(x'-u\right)\right)_{+}^{-1/\xi-1} \text{ with } \nu = \sigma(1+\xi)$$

  for $x' > u$ with parameters shape $\xi > -1$ and adjusted scale $\nu > 0$
- Sample likelihood for storm peak threshold exceedances $\boldsymbol{x}' = \{x'_{i'}\}_{i'=1}^{n'}$ with associated directional bins $\{\theta_{\mathcal{I}(i')}\}_{i'=1}^{n'}$ is

$$\mathcal{L}_{\mathrm{GP}}(\boldsymbol{\xi},\boldsymbol{\nu}|\boldsymbol{x}',\boldsymbol{u}) = \prod_{i=1}^{n} f_{\mathrm{GP}}(\boldsymbol{x}';\xi_{\mathcal{I}(i')},\nu_{\mathcal{I}(i')},,u_{\mathcal{I}(i')})$$

- ξ and ν (and u) vary smoothly with direction on $\mathcal{D}$

Smooth covariate representations in 1-D

Covariate representation

- Index set $\mathcal{I}$ of m directional bins on $\mathcal{D}$; each observation belongs to exactly one bin
- On $\mathcal{I}$ assume $\eta_j = \sum_{k=1}^{p} B_{jk}\beta_k$, $j = 1, 2, \ldots, m$, or

$$\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\beta}, \quad \eta \in (\xi, \nu, \rho, u)$$

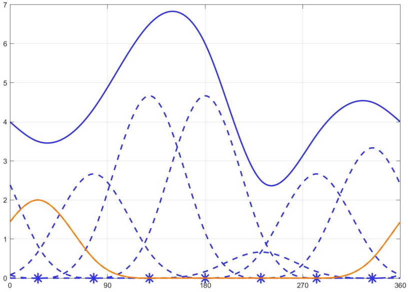
- $\boldsymbol{B} = \{B_{jk}\}_{j=1;k=1}^{m;p}$ set of p basis vectors for $\mathcal{D}$
- $\boldsymbol{\beta} = \{\beta_k\}_{k=1}^{p}$ basis coefficients
- Inference reduces to estimating $\boldsymbol{B}$ and $\boldsymbol{\beta}$ for each η

Roughness penalty

- $\boldsymbol{D}$ is a $p \times p$ (wrapped) differencing matrix
- $\boldsymbol{D}\boldsymbol{\beta}$ is a vector of differences between consecutive values of $\boldsymbol{\beta}$
- $\boldsymbol{\beta}'\boldsymbol{D}'\boldsymbol{D}\boldsymbol{\beta}$ is the scalar sum of squares difference, a roughness penalty!

P-splines

- p regularly-spaced knots on $\mathcal{D}$
- $\boldsymbol{B}$ consists of p B-spline bases
  - Order d
  - Each using $d + 1$ consecutive knot locations
  - **Local support**
  - Wrapped on $\mathcal{D}$
  - Cox - de Boor recursion formula



Periodic P-splines

- p is fixed and "over-specified"
- Knot locations fixed
- Roughness $\boldsymbol{\beta}'\boldsymbol{D}'\boldsymbol{D}\boldsymbol{\beta}$ penalised

- Alternatives: Fourier, Voronoi partition, adaptive splines, ...

Maximised penalised likelihood

- Choose $\boldsymbol{\eta}$ to maximise

$$\mathcal{L}^*(\boldsymbol{\eta}) = \mathcal{L}(\boldsymbol{\eta}) - \lambda \boldsymbol{\beta}' \boldsymbol{D}' \boldsymbol{D} \boldsymbol{\beta}$$

  for given roughness coefficient $\lambda > 0$
- Use cross-validation to estimate the value of λ which maximises predictive performance
- Use bootstrapping to estimate uncertainties on everything of interest

- General procedure: works for all of u, ρ, ξ and ν
- Using a simple piecewise constant covariate representation, this is exactly the idea behind covXtreme

- This is a "frequentist" approach; alternative is Bayesian inference

Bayesian inference

Priors

$$\begin{aligned} \text{prior density of } \boldsymbol{\beta} &\propto \exp\left(-\lambda \boldsymbol{\beta}' \boldsymbol{P} \boldsymbol{\beta}\right) \\ \text{prior density of } \lambda &\sim \text{gamma} \end{aligned}$$

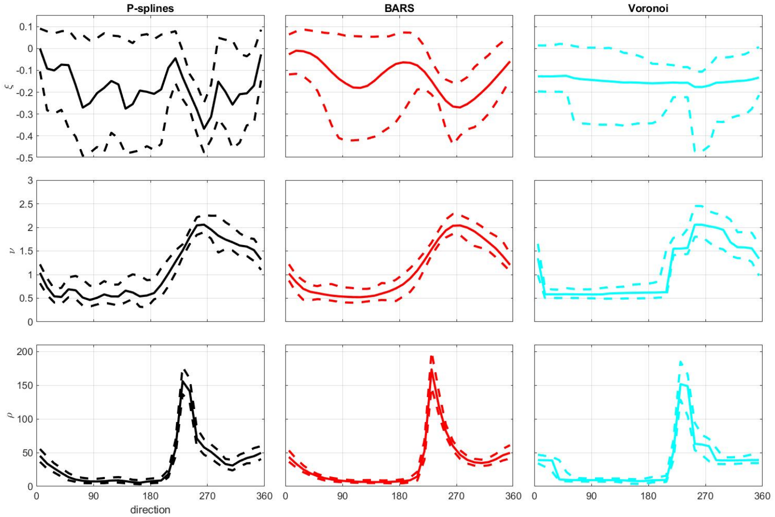
Inference for GP

- Objective: sample from posterior density $f(\boldsymbol{\beta}_\xi, \lambda_\xi, \boldsymbol{\beta}_\nu, \lambda_\nu | \boldsymbol{x'})$
- Parameter set $\Omega = \{\boldsymbol{\beta}_\xi, \lambda_\xi, \boldsymbol{\beta}_\nu, \lambda_\nu\}$
- Generic conditional structure for Metropolis Hastings within Gibbs sampling from full conditionals

$$\begin{aligned} f(\boldsymbol{\beta}_\xi | \boldsymbol{x}', \Omega \setminus \boldsymbol{\beta}_\xi) &\propto f(\boldsymbol{x}' | \boldsymbol{\beta}_\xi, \Omega \setminus \boldsymbol{\beta}_\xi) \times f(\boldsymbol{\beta}_\xi | \lambda_\xi) \\ f(\boldsymbol{\beta}_\nu | \boldsymbol{x}', \Omega \setminus \boldsymbol{\beta}_\nu) &\propto f(\boldsymbol{x}' | \boldsymbol{\beta}_\nu, \Omega \setminus \boldsymbol{\beta}_\nu) \times f(\boldsymbol{\beta}_\nu | \lambda_\nu) \\ f(\lambda_\xi | \boldsymbol{x}', \Omega \setminus \lambda_\xi) &\propto f(\boldsymbol{\beta}_\xi | \lambda_\xi) \times f(\lambda_\xi) \\ f(\lambda_\nu | \boldsymbol{x}', \Omega \setminus \lambda_\nu) &\propto f(\boldsymbol{\beta}_\nu | \lambda_\nu) \times f(\lambda_\nu) \end{aligned}$$

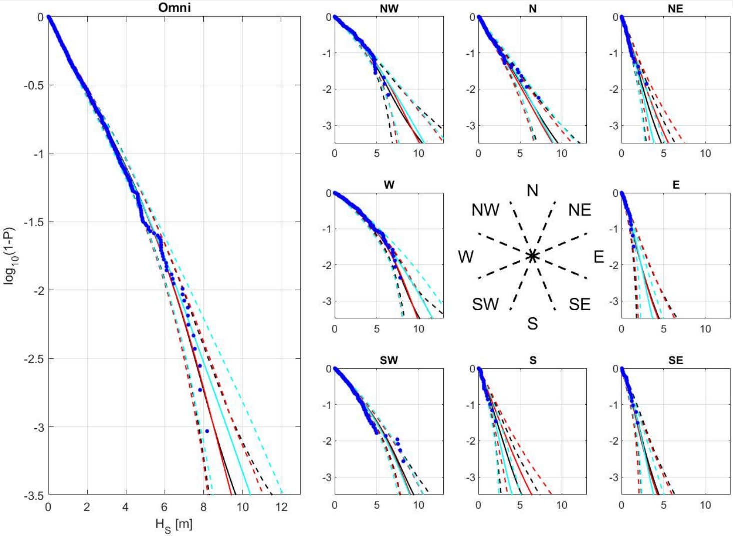
- Inference for ρ, u similar

Posterior parameter estimates for northern North Sea



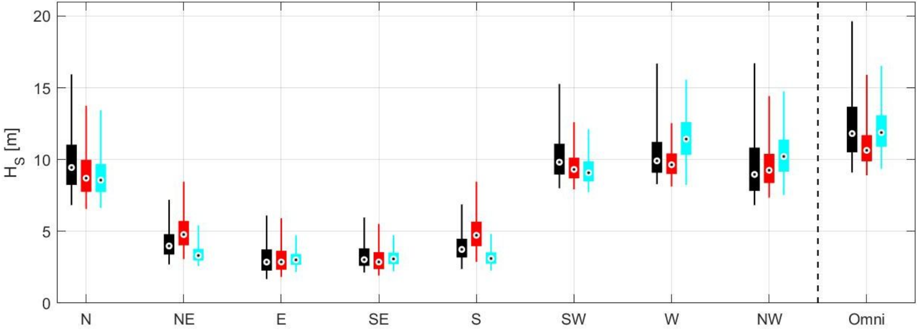
Comparing Bayesian inference using P-splines, adaptive regression splines and Voronoi partition; Zanini et al. [2020]

Fit diagnostic



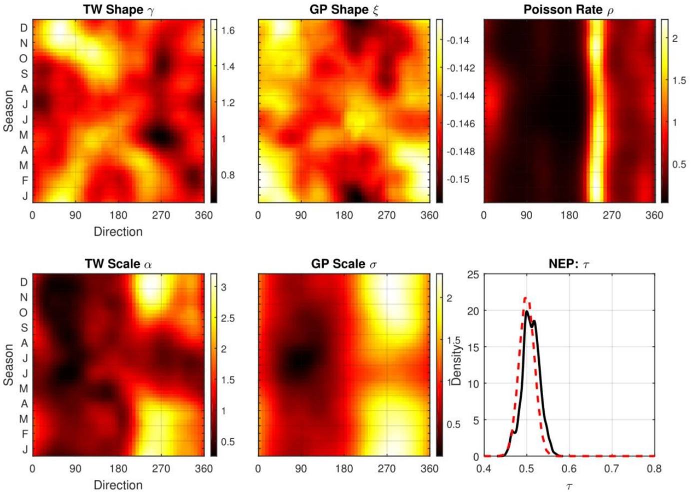
Empirical tail (blue); posterior means and 95% credible intervals for quantile levels from different models

Directional posterior predictive distribution of 1000-year maximum



- Box-whiskers with 2.5%, 25%, 50%, 75% and 97.5% percentiles

Extension to 2D : directional-seasonal



- 2-D tensor product P-spline bases for same northern North Sea location
- Marginal posterior median estimates (plus posterior density for τ)
- Uses hybrid truncated Weibull-GP model

Covariate modelling: what can go wrong?

- Small sample size
- Selection of appropriate non-stationary threshold(s)
  - Use diagnostics e.g. $\hat{\xi}|(u,\theta)$
  - Model averaging probably better when choice not obvious
- Complex tail structures requiring multiple covariates or very high u
- Estimation of optimal roughness coefficients λ using cross-validation in frequentist analysis can be problematic; Bayesian inference tends to behave more reasonably

Implications

- Propagate and quantify uncertainty carefully
  - No excuses!
- Model averaging in place of hyper-parameter selection
- Use cross-validation and bootstrapping in frequentist analysis; otherwise use Bayesian inference

All models are wrong, some models are useful

If you want to read more

- Key papers for spline applications: Marx and Eilers [1998], Chavez-Demoulin and Davison [2005]
- Basic ideas around covariate modelling: Jonathan et al. [2008], Jones et al. [2016], Feld et al. [2019a], Feld et al. [2019b]
- Lancaster / Shell: Randell et al. [2016] (North Sea), Ross et al. [2017] (South China Sea), Zanini et al. [2020] (North Sea)
- Lots of related work at `www.lancaster.ac.uk/~jonathan`

References

V. Chavez-Demoulin and A. C. Davison. Generalized additive modelling of sample extremes. *J. Roy. Statist. Soc. Series C: Appl. Stat.*, 54:207–222, 2005.

G. Feld, D. Randell, and P. Jonathan. On the estimation and application of directional design criteria. *Proc. 38th Int. Conf. on Ocean, Offshore & Arctic Engineering, Scotland*, 2019a.

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, 2019b.

P. Jonathan, K. C. Ewans, and G. Z. Forristall. Statistical estimation of extreme ocean environments: The requirement for modelling directionality and other covariate effects. *Ocean Eng.*, 35:1211–1225, 2008.

M. J. Jones, D. Randell, K. Ewans, and P. Jonathan. Statistics of extreme ocean environments: non-stationary inference for directionality and other covariate effects. *Ocean Eng.*, 119:30–46, 2016.

B. D. Marx and P. H. C. Eilers. Direct generalised additive modelling with penalised likelihood. *Comput. Statist. Data Anal.*, 28:193–209, 1998.

D. Randell, K. Turnbull, K. Ewans, and P. Jonathan. Bayesian inference for non-stationary marginal extremes. *Environmetrics*, 27:439–450, 2016.

E. Ross, D. Randell, K. Ewans, G. Feld, and P. Jonathan. Efficient estimation of return value distributions from non-stationary marginal extreme value models using Bayesian inference. *Ocean Eng.*, 142:315–328, 2017.

E. Zanini, E. Eastoe, M. J. Jones, D. Randell, and P. Jonathan. Flexible covariate representations for extremes. *Environmetrics*, 31:e2624, 2020.