# Extremes for metocean: covXtreme

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

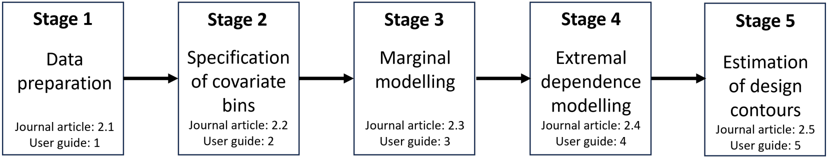
# Overview of D: covXtreme

- What is covXtreme?
- Walk-through of typical analysis in 5 stages
    - What is the analysis doing?
    - How does covXtreme report the analysis
    - Notes and issues
- Reflections: covXtreme in context

# What is covXtreme?

- Entry-level open-source MATLAB software for estimation of
  - non-stationary marginal extreme value models
  - non-stationary conditional extreme value models
  - associated values and environmental design contours
- Models used
  - Simplified gamma-GP model used for marginal distributions
  - Conditional extremes model for dependence
  - Different contour formulations
- Covariate representation
  - Piecewise constant on covariate domain
  - Pre-specified covariate "bins" within which stationary assumed
  - GP $\xi$ assumed constant on covariate domain
  - CE $\beta$, $\mu$ and $\sigma$ assumed constant on covariate domain
  - Roughness of GP $\nu$ and CE $\alpha$ penalised for optimal predictive performance
- Model fitting and uncertainty quantification
  - Cross-validation for optimal parameter roughness
  - Bootstrapping for parameter uncertainty
  - Model averaging over multiple marginal and dependence thresholds

# Analysis procedure




- Journal article: Towe et al. [2024]
- User guide and software: `https://github.com/covXtreme`

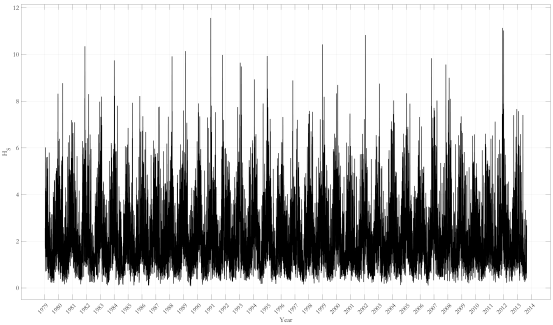
**Notes**

- Not pushing covXtreme for applications (but would use as benchmark)
- Using covXtreme here as a computationally fast teaching tool, incorporating key analysis steps in statistically sound manner

# Case study

- "Case study 1" from Towe et al. [2024]
- Software and data to reproduce analysis available on GitHub

- Objective: characterise the joint distribution of (storm peak) $H_S$ and (associated) $T_P$ for a northern North Sea location
- Approximately 35 years of data for sea state $H_S$, $T_P$ and wave direction $\theta$

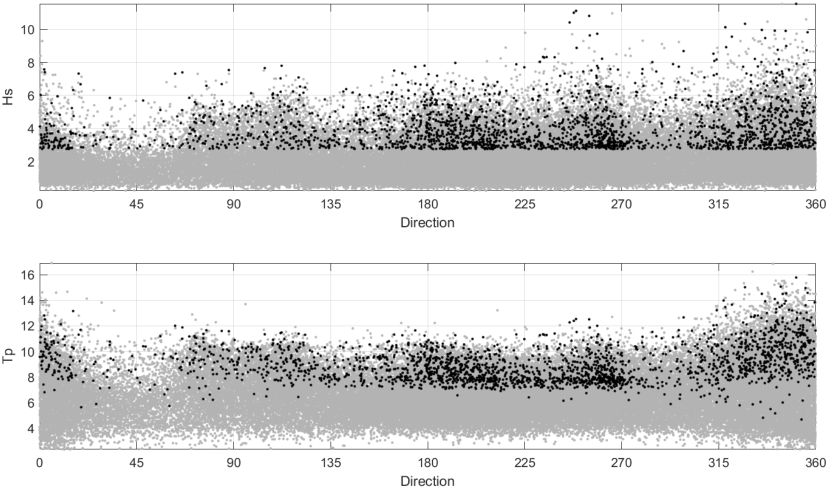


# Stage 1: Data preparation

**Objective**: isolate storm peak events, assumed to be statistically independent given covariates, which can used fairly in an extreme value analysis

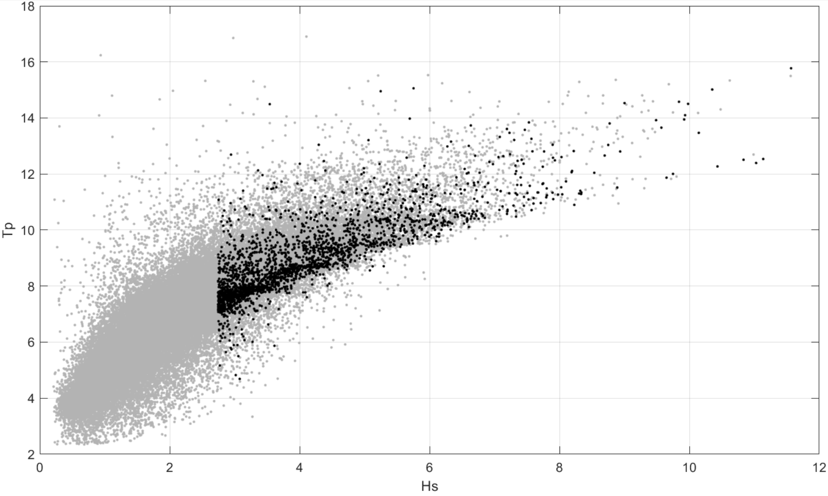
- Define "peak picking" threshold
  - An empirical quantile $q$ of the sea state $H_S$ data
  - Corresponds to $\mathbb{P}(H_S > q) = 0.7$ for the current application
- Identify contiguous intervals of $H_S$ in time above $q$
  - These are called "storms"
- Identify the times of storm peak $H_S$ within each storm
  - These are the "storm peak times"
- Extract the values of all variables of interest at the storm peak times
  - Storm peak significant wave height $H_S^{sp}$
  - Associated peak wave period $T_P^{ass}$
  - Storm peak direction $\theta^{sp}$
  - ...

# Stage 1: Data preparation: $H_S^{sp}$ and $T_P^{ass}$ on direction




- Sea state $H_S$ and $T_P$ variables (grey)
- Storm peak $H_S$ and associated $T_P$ (black)

# Stage 1: Data preparation: $T_P^{ass}$ on $H_S^{sp}$




- Sea state $H_S$ and $T_P$ variables (grey)
- Storm peak $H_S$ and associated $T_P$ (black)

# Stage 1: Data preparation: Comments

**Notes**

- Directional effects in $H_S^{sp}$ and $T_P^{ass}$
- Positive dependence between $H_S^{sp}$ and $T_P^{ass}$
- Both these effects should be incorporated in the statistical model

**Potential issues**

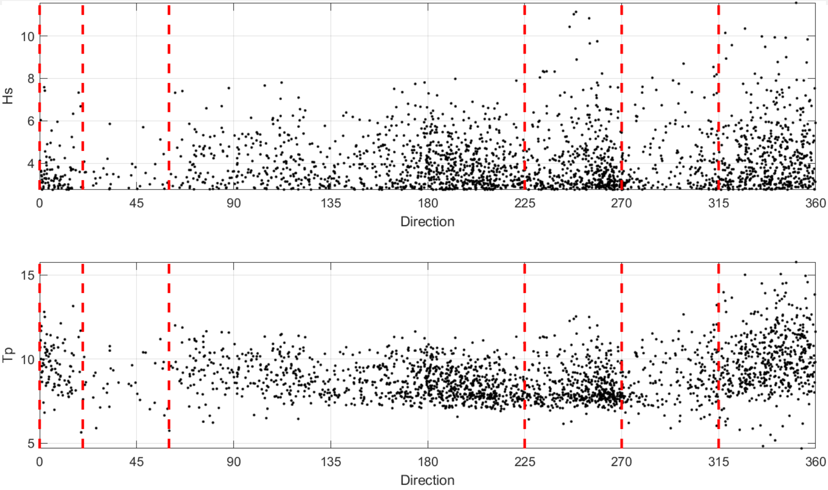
- Peak picking ignores covariate effects in covXtreme
- Sensitivity to peak picking threshold choice not assessed in covXtreme
- Both these issues could be addressed, at increased computational cost

# Stage 2: Covariate partition

**Objective**: partition the covariate domain into intervals, within which the marginal and dependence characteristics of $H_S^{sp}$ and $T_P^{ass}$ do not depend on the covariate(s)

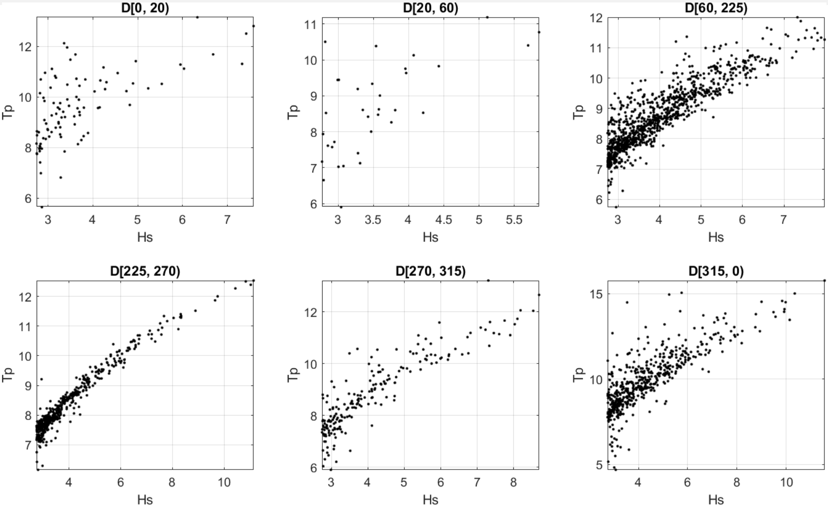
- Partition the domain $[0, 360)$ for storm peak direction $\theta^{sp}$ into intervals
- Within each interval, there is no obvious influence of $\theta$ on the marginal and dependence characteristics of $H_S^{sp}$ and $T_P^{ass}$
- Partition is performed manually

# Stage 2: Covariate partition: Bin edge specification




- Bin edges (red) located manually using physical intuition and data
- Partition separates Atlantic, Norwegian Sea, North Sea and Norwegian landshadow

# Stage 2: Covariate partition: Within-bin characteristics




- Positive dependence between $H_S^{sp}$ and $T_P^{ass}$ in all bins
- Detail of dependence varies between bins?

# Stage 2: Covariate partition: Comments

**Notes**

- Directional "bins" look reasonable
  - Marginal and dependence characteristics are different between bins
  - Marginal and dependence characteristics are approximately stationary within bin
- This means we can perform stationary marginal and dependence modelling within each bin

**Potential issues**

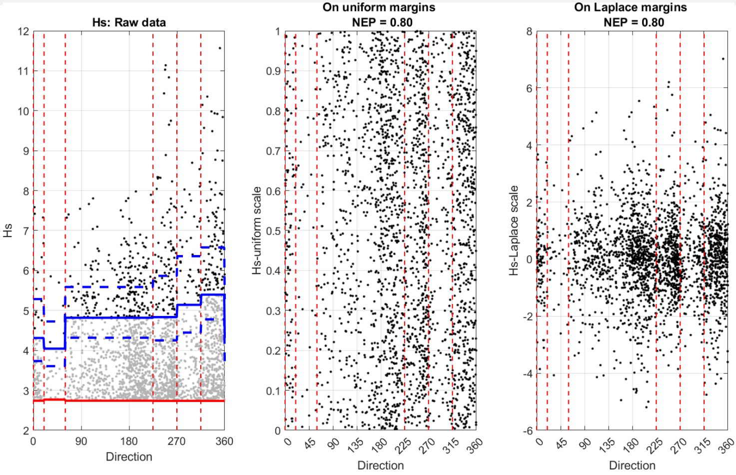
- It is likely that characteristics of $H_S^{sp}$ and $T_P^{ass}$ vary smoothly with $\theta^{sp}$
  - But piecewise constant representation provides good performance in practice (e.g. Zanini et al. 2020)
  - Resolving this issue would incur considerable computational cost (and see e.g. Jones et al. 2016, Randell et al. 2016)
- Covariate bins are defined by hand
  - Resolving this issue is possible, but again at considerable computational cost (e.g. Jonathan 2021)

# Stage 3: Marginal modelling

**Objective**: characterise the marginal tails of $H_S^{sp}$ (and $T_P^{ass}$ independently). Exploit the fact that neighbouring covariate bins are likely to have similar marginal characteristics. Propagate uncertainty using plausible random marginal thresholds and bootstrapping.

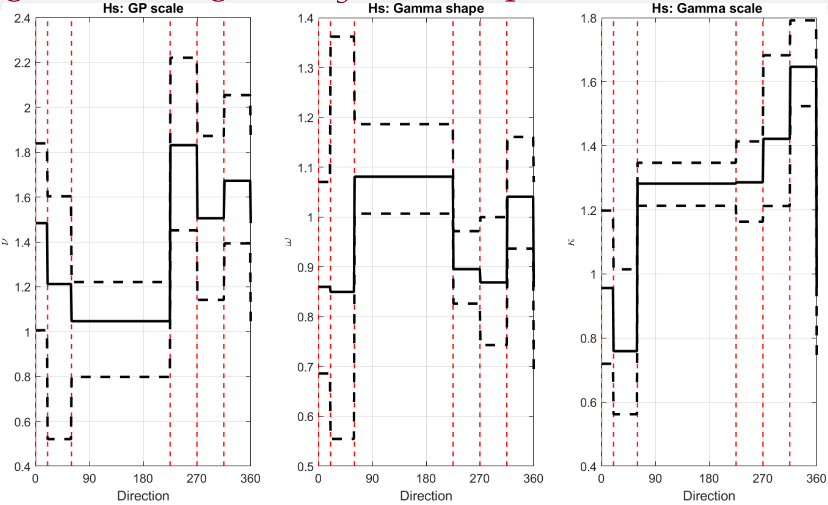
- Extreme value threshold estimation
  - Fit the gamma distribution to $H_S^{sp}$ on each covariate bin independently
  - Use a high quantile $q$ of the distribution as the extreme value threshold for that bin
  - Use lots of different plausible quantile levels to propagate uncertainty in threshold choice
  - Use threshold stability diagnostic for estimated GP shape
- Tail fitting
  - Fit the GP distribution to threshold exceedances of $q$ per bin
  - Penalise the roughness of GP scale parameter across bins
  - Select the roughness coefficient which optimises predictive performance across all bins

# Stage 3(a): Marginal $H_S^{sp}$: Extreme value threshold



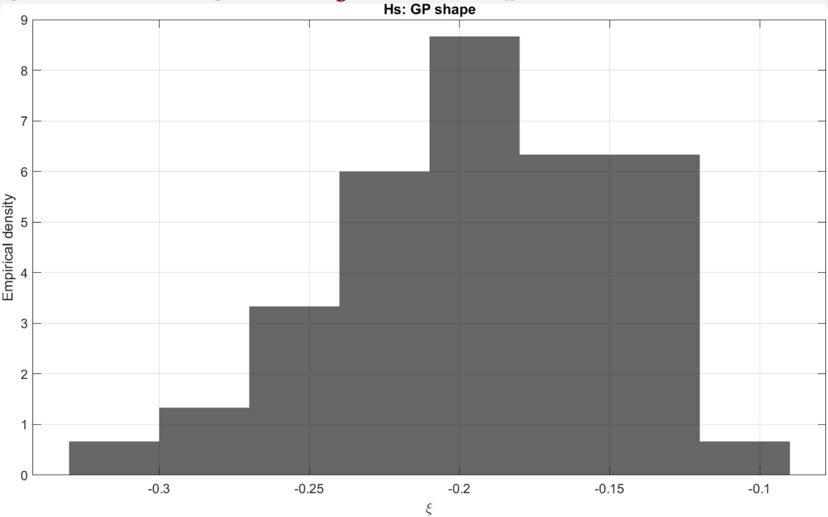

- Median and 95% spread of extreme value threshold (left: blue)
- Peak picking threshold (left: red)
- Inspection of data on different marginal scales (centre and right)

# Stage 3(a): Marginal $H_S^{sp}$: Model parameters



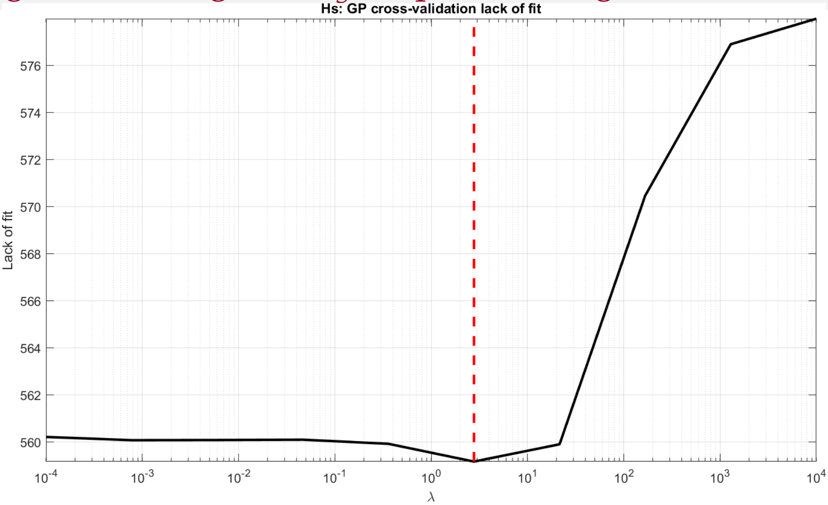

- Median and 95% intervals for GP shape, gamma shape and gamma scale
- Uncertainty from EV threshold and bootstrap resampling
- Effect of direction obvious

# Stage 3(a): Marginal $H_S^{sp}$: GP shape



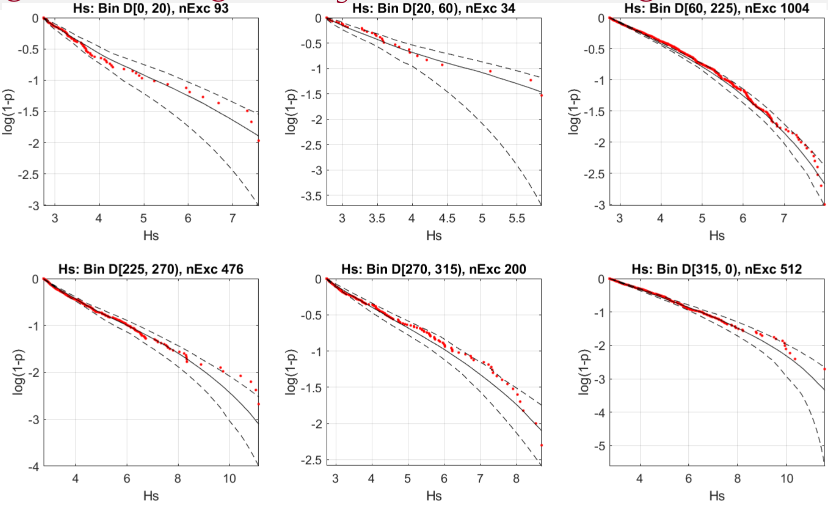

- GP shape $< 0$, and centred around -0.2, typical for northern North Sea
- Uncertainty from EV threshold and bootstrap resampling

# Stage 3(a): Marginal $H_S^{sp}$: Optimal roughness $\lambda$



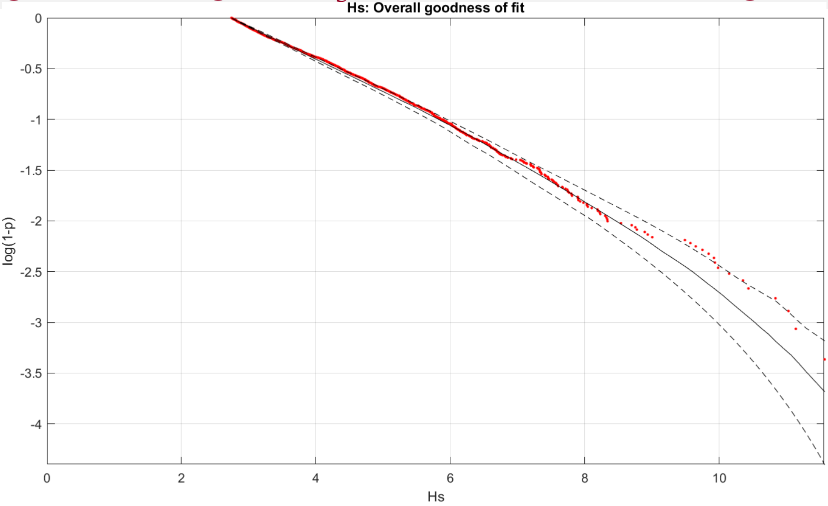

- Large $\lambda \Rightarrow$ low variability of GP scale with direction
- Small $\lambda \Rightarrow$ high variability of GP scale with direction
- Optimal $\lambda$ ensures best out-of-sample predictive performance

# Stage 3(a): Marginal $H_S^{sp}$: Directional diagnostics



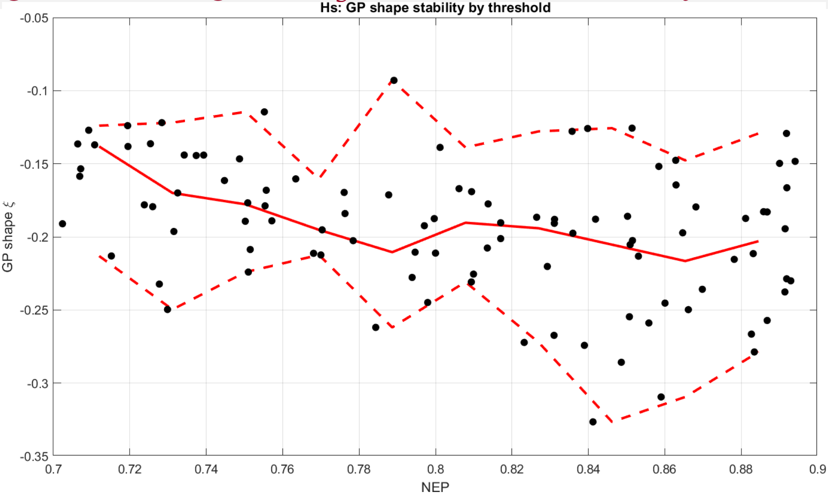

- Sample data (red)
- Median and 95% uncertainty bands from statistical model (black)
- Uncertainty from EV threshold and bootstrap resampling

# Stage 3(a): Marginal $H_S^{sp}$: Omni-directional diagnostic



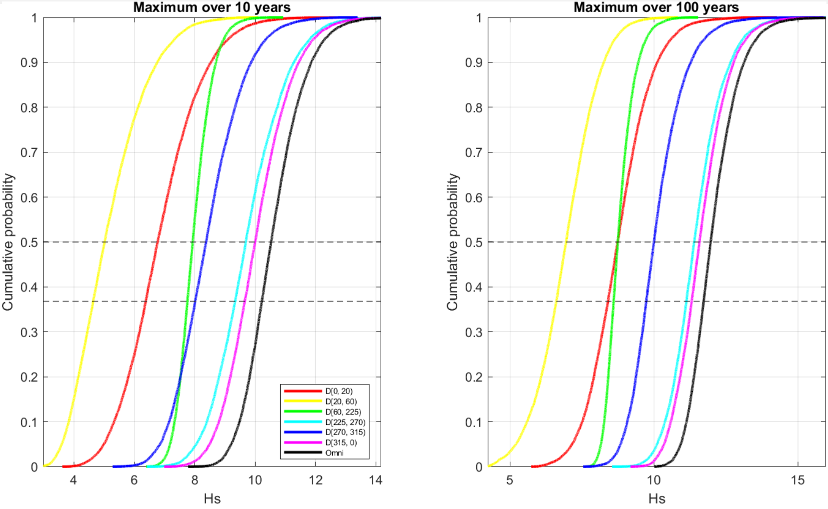

- Sample data (red)
- Median and 95% uncertainty bands from statistical model (black)
- Uncertainty from EV threshold and bootstrap resampling

# Stage 3(a): Marginal $H_S^{sp}$: Threshold stability




- Individual estimates of GP shape (from analysis using particular EV threshold and sample bootstrap resample; black)
- Median and 95% uncertainty band with threshold NEP (red)

# Stage 3(a): Marginal $H_S^{sp}$: $P$-year maxima

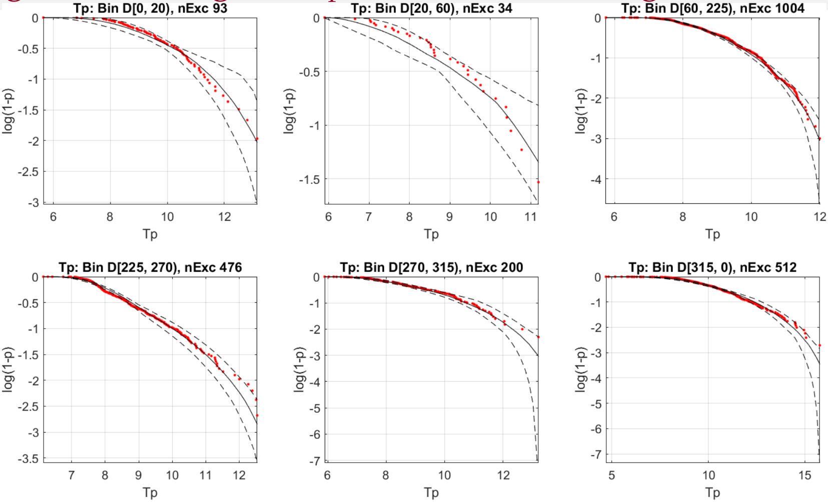



- Distribution of 10-year and 100-year maxima per directional bin
- Estimated using simulation (or efficiently using importance sampling)

# Stage 3(b): Marginal modelling: $T_P^{ass}$

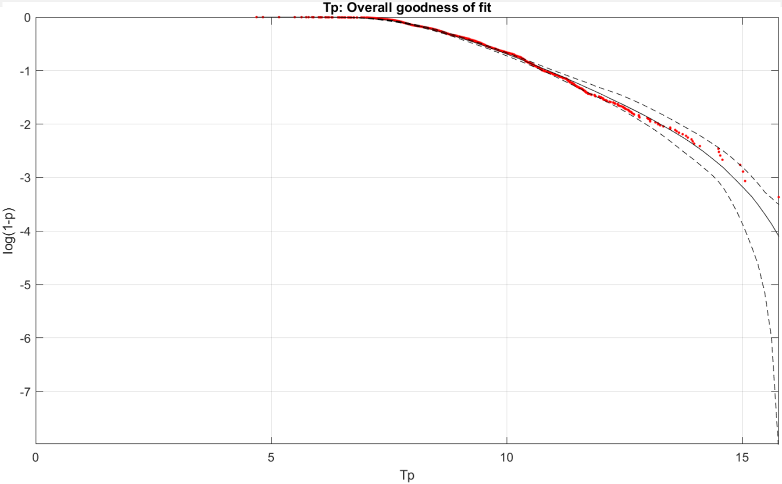
- Analogous analysis
- Diagnostics and $P$-year maxima shown here for illustration

# Stage 3(b): Marginal $T_P^{ass}$: Directional diagnostics



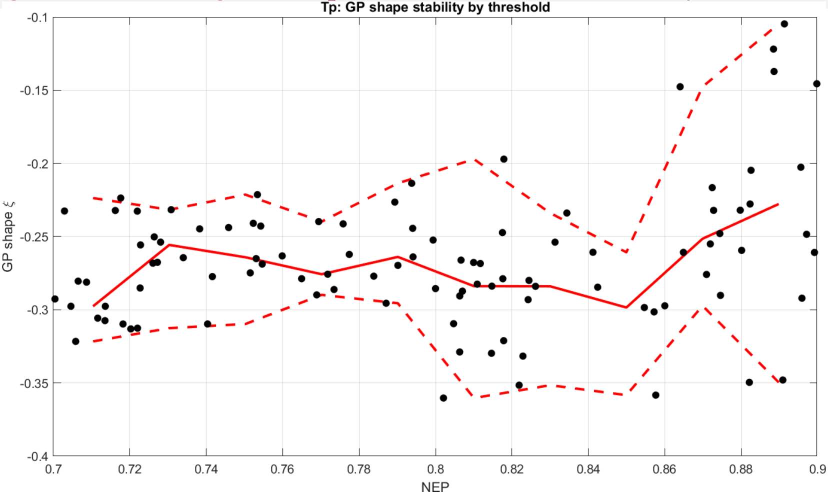

- Sample data (red)
- Median and 95% uncertainty bands from statistical model (black)
- Uncertainty from EV threshold and bootstrap resampling

# Stage 3(b): Marginal $T_P^{ass}$: Omni-directional diagnostic



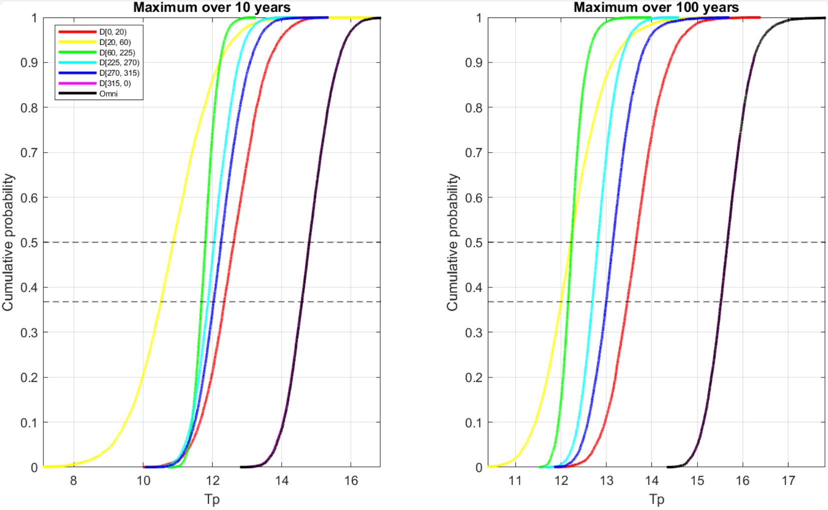

- Sample data (red)
- Median and 95% uncertainty bands from statistical model (black)
- Uncertainty from EV threshold and bootstrap resampling

# Stage 3(b): Marginal $T_P^{ass}$: Threshold stability




- Individual estimates of GP shape (from analysis using particular EV threshold and sample bootstrap resample; black)
- Median and 95% uncertainty band with threshold NEP (red)

# Stage 3(b): Marginal $T_P^{ass}$: $P$-year maxima




- Distribution of 10-year and 100-year maxima per directional bin
- Estimated using simulation (or efficiently using importance sampling)

# Stage 3: Marginal modelling: Comments

**Notes**

- Diagnostics indicate reasonable tail fits per bin and omni-directionally
- Diagnostics indicate reasonable threshold stability
- Nice optimal roughness coefficient $\lambda$ for GP scale
- GP shape estimates consistent with experience

**Potential issues**

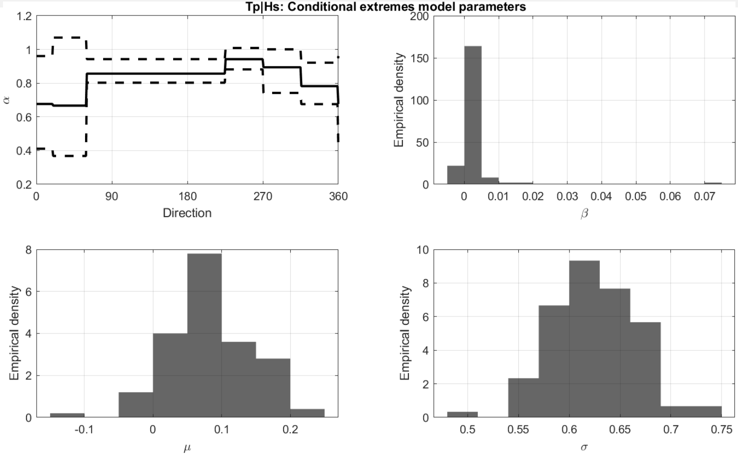
- Likely that characteristics of $H_S^{sp}$ and $T_P^{ass}$ vary smoothly with $\theta^{sp}$
- Optimal $\lambda$ chosen using original sample only for computational efficiency
- GP shape held fixed with direction

# Stage 4: Dependence modelling

**Objective**: characterise the extremal dependence between $T_P^{ass}$ and $H_S^{sp}$ using the conditional extremes (CE) model. Exploit the fact that neighbouring covariate bins are likely to have similar dependence characteristics. Propagate uncertainty using plausible random marginal thresholds, dependence thresholds and bootstrapping.

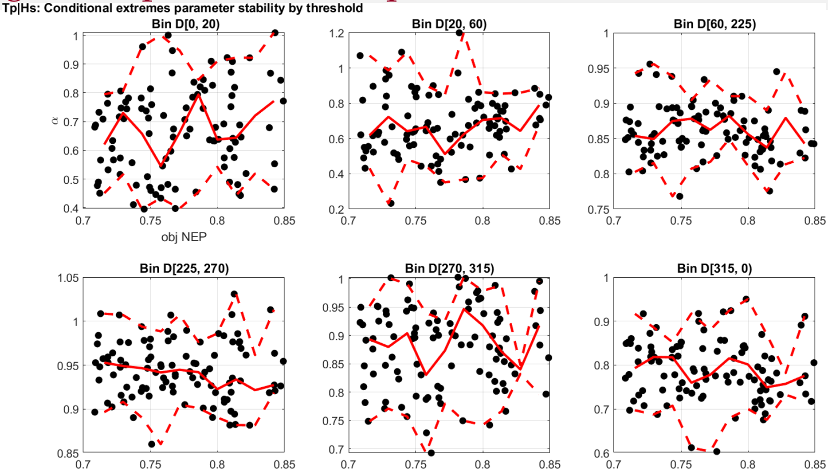
- Estimate an interval of plausible dependence thresholds $q^*$ for $H_S^{sp}$
  - Use threshold stability diagnostic for CE slope "$\alpha$" parameter
  - Use lots of plausible $q^*$ values to propagate uncertainty
- Fit the CE model for $T_P^{ass}|(H_S^{sp} > q^*)$
  - Fit CE model to pairs of observations of $H_S^{sp}$ and $T_P^{ass}$ with $H_S^{sp} > q^*$ per bin
  - Penalise the roughness of estimated "$\alpha$" parameter across bins
  - Select the roughness coefficient which optimises predictive performance across all bins

# Stage 4: Dependence: CE parameter estimates



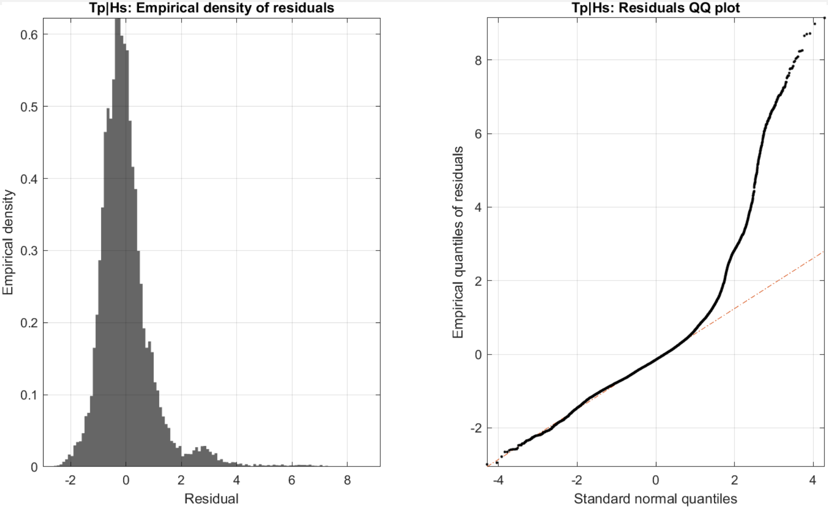

- Median and 95% uncertainty for estimated non-stationary CE slope $\alpha$
- Empirical densities for estimated $\beta$, $\mu$ and $\sigma$ (all assumed stationary)
- Uncertainty from marginal and dependence thresholds and bootstrap resampling

# Stage 4: Dependence: CE parameter estimates



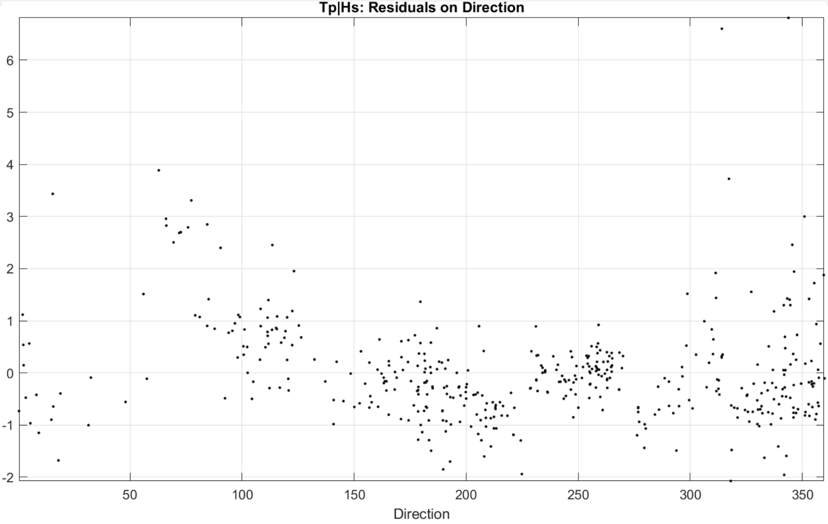

- Individual estimates of CE $\alpha$ with dependence threshold NEP (black)
- Median and 95% uncertainty band with dependence threshold NEP (red)
- Uncertainty from marginal and dependence thresholds and bootstrap resampling

# Stage 4: Dependence: CE residuals



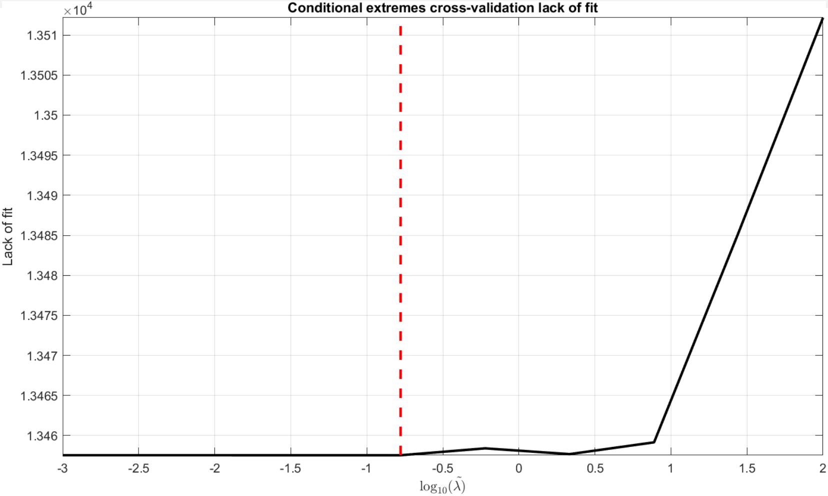

- Left: empirical density of CE residuals (usual long tail)
- Right: QQ plot of residuals (usual long tail)

# Stage 4: Dependence: CE residuals



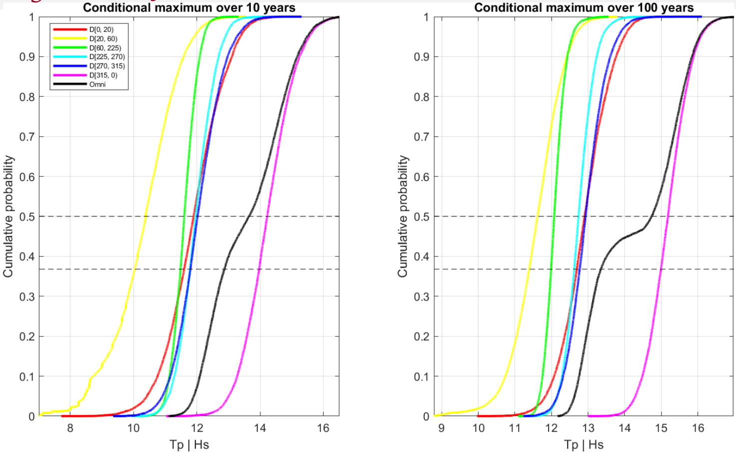

- No obvious relation between residual value and $\theta^{sp}$
  - Maybe some concern around 60°

# Stage 4: Dependence: Optimal roughness $\lambda^*$




- Large $\lambda^* \Rightarrow$ low variability of GP scale with direction
- Optimal $\lambda^*$ ensures best out-of-sample predictive performance
- No increase in lack of fit with decreasing $\lambda^*$ is not unusual

# Stage 4: Dependence: Conditional distribution of $T_P^{ass}|(H_S^{sp} = P\text{-year maximum})$




- Estimated using simulation (or efficiently using importance sampling)
- Reduction in 1 to 2 seconds in median value omni-directionally with respect to unconditional $P$-year maxima for $T_P^{ass}$

# Stage 4: Dependence modelling: Comments

**Notes**

- Diagnostics indicate reasonable threshold stability
- Diagnostics indicate independence of residual and $\theta^{sp}$
- Reasonable optimal roughness coefficient $\lambda^*$ for CE slope $\alpha$
- Clear variation of $\alpha$ with $\theta^{sp}$
- Reduction in most probable and median conditional value with respect to $P$-year equivalent

**Potential issues**

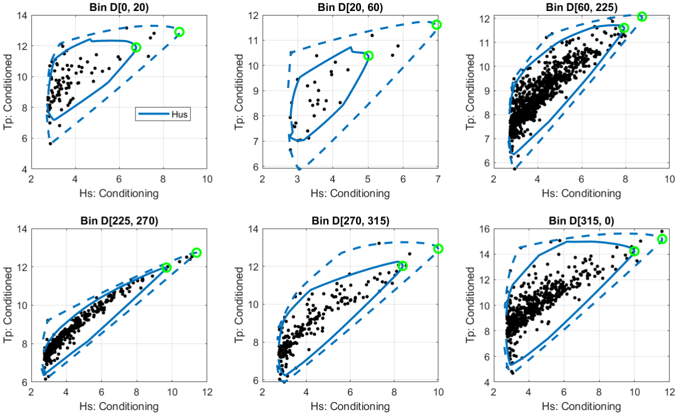
- Likely that $\alpha$ varies smoothly with $\theta^{sp}$
- Optimal $\lambda^*$ chosen using original sample only for computational efficiency
- CE parameters $\beta$, $\mu$ and $\sigma$ assumed stationary
- See Jonathan et al. [2014] for discussion

# Stage 5: Environmental contour estimation

**Objective**: simulate under fitted marginal and dependence models to estimate environmental contours for the space of $H_S^{sp}$ and $T_P^{ass}$. Use Huseby "direct sampling" contours as illustration. Propagate uncertainty using plausible random marginal thresholds, dependence thresholds and bootstrapping

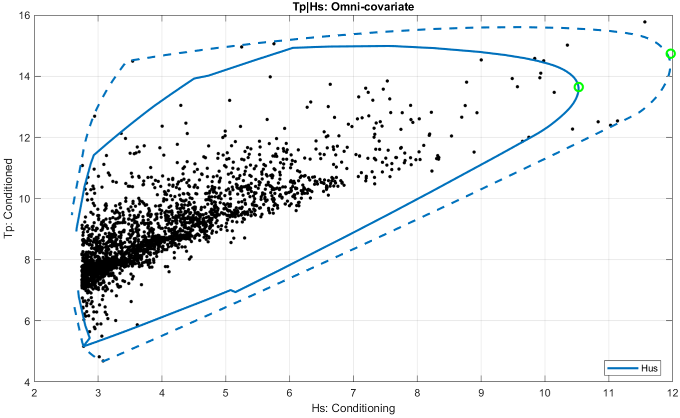
- Simulate a large sample, sufficient to characterise the return period of interest well
- Estimate the "direct sampling" contours of Huseby et al. [2015]
- Incorporate uncertainty from plausible random marginal thresholds, dependence thresholds and bootstrapping

# Stage 5: Contour estimation: Directional Huseby contours




- Directional direct sampling contours corresponding to the 10-year and 100-year $H_S^{sp}$ event

# Stage 5: Contour estimation: Omni-directional Huseby contours




- Omni-directional direct sampling contours corresponding to the 10-year and 100-year $H_S^{sp}$ event

# Stage 5: Environmental contours: Comments

- Environmental contours have their place, but are they really the best we can do?
  - Lots of different contour types; never clear which is best
- Full probabilistic risk analysis and optimal decision making better
  - Couple metocean and structures analysis fully
  - See Speers et al. [2024]

# covXtreme: Reflections

**A simple entry-level benchmarking tool**

- Statistically and computationally well-founded
- Open-source: you can download code, modify and share improvements
  - A PYTHON version would be nice for accessibility
- Computationally fast
  - Analysis (although limited) complete in seconds
  - Good tool to explore ideas: marginal and multivariate extremes, covariate effects, uncertainty propagation

**Not a panacea**

- Lots of considerably more complex tools targetting metocean specifically
  - DNV
  - Gibson
  - Hansen: JEVA
  - Norwegian metoffice: EnvEx
  - Shell: CEVA
  - ...
- Tools in R from applied statistics literature: threshr, texmex, evgam, ...

All models are wrong, some models are useful

# References


A Huseby, E Vanem, and B Natvig. Alternative environmental contours for structural reliability analysis. *Struct. Saf.*, 54:32–45, 2015.

P. Jonathan. A non-parametric representation for multidimensional covariates in an extreme value model. *Gwerddon*, 33:68–84, 2021.

P. Jonathan, K. C. Ewans, and D. Randell. Non-stationary conditional extremes of northern North Sea storm characteristics. *Environmetrics*, 25:172–188, 2014.

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.

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

M. Speers, D. Randell, J. A. Tawn, and P. Jonathan. Estimating metocean environments associated with extreme structural response to demonstrate the dangers of environmental contour methods. *Ocean Eng.*, 311:118754, 2024.

R. Towe, E. Ross, D. Randell, and P. Jonathan. covXtreme: MATLAB software for non-stationary penalised piecewise constant marginal and conditional extreme value models. *Environ. Model. Softw.*, 177:106035, 2024.

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