covXtreme: open-source software for modelling extreme data sets

Ross Towe, Emma Ross, David Randell and Philip Jonathan

Guidance: DEFINITIONS & CAUTIONARY NOTE

The companies in which Shell plc directly and indirectly owns investments are separate legal entities. In this presentation"Shell", "Shell Group" and "Group" are sometimes used for convenience where references are made to Shell plc and its subsidiaries in general. Likewise, the words "we", "us" and "our" are also used to refer to Shell plc and its subsidiaries in general or to those who work for them. These terms are also used where no useful purpose is served by identifying the particular entity or entities. ''Subsidiaries'', "Shell subsidiaries" and "Shell companies" as used in this presentationrefer to entities over which Shell plc either directly or indirectly has control. Entities and unincorporated arrangements over which Shell has joint control are generally referred to as "joint ventures" and "joint operations", respectively. "Joint ventures" and "joint operations" are collectively referred to as "joint arrangements". Entities over which Shell has significant influence but neither control nor joint control are referred to as "associates". The term "Shell interest" is used for convenience to indicate the direct and/or indirect ownership interest held by Shell in an entity or unincorporated joint arrangement, after exclusion of all third-party interest.

Forward-Looking Statements This presentationcontains forward-looking statements (within the meaning of the U.S. Private Securities Litigation Reform Act of 1995) concerning the financial condition, results of operations and businesses of Shell. All statements other than statements of historical fact are, or may be deemed to be, forward-looking statements. Forward-looking statements are statements of future expectations that are based on management's current expectations and assumptions and involve known and unknown risks and uncertainties that could cause actual results, performance or events to differ materially from those expressed or implied in these statements. Forward-looking statements include, among other things, statements concerning the potential exposure of Shell to market risks and statements expressing management's expectations, beliefs, estimates, forecasts, projections and assumptions. These forward-looking statements are identified by their use of terms and phrases such as "aim", "ambition", ''anticipate'', ''believe'', ''could'', ''estimate'', ''expect'', ''goals'', ''intend'', ''may'', "milestones", ''objectives'', ''outlook'', ''plan'', ''probably'', ''project'', ''risks'', "schedule", ''seek'', ''should'', ''target'', ''will'' and similar terms and phrases. There are a number of factors that could affect the future operations of Shell and could cause those results to differ materially from those expressed in the forward-looking statements included in this presentation, including (without limitation): (a) price fluctuations in crude oil and natural gas; (b) changes in demand for Shell's products; (c) currency fluctuations; (d) drilling and production results; (e) reserves estimates; (f) loss of market share and industry competition; (g) environmental and physical risks; (h) risks associated with the identification of suitable potential acquisition properties and targets, and successful negotiation and completion of such transactions; (i) the risk of doing business in developing countries and countries subject to international sanctions; (j) legislative, judicial, fiscal and regulatory developments including regulatory measures addressing climate change; (k) economic and financial market conditions in various countries and regions; (l) political risks, including the risks of expropriation and renegotiation of the terms of contracts with governmental entities, delays or advancements in the approval of projects and delays in the reimbursement for shared costs; (m) risks associated with the impact of pandemics, such as the COVID-19 (coronavirus) outbreak; and (n) changes in trading conditions. No assurance is provided that future dividend payments will match or exceed previous dividend payments. All forward-looking statements contained in this presentationare expressly qualified in their entirety by the cautionary statements contained or referred to in this section. Readers should not place undue reliance on forward-looking statements. Additional risk factors that may affect future results are contained in Shell plc's Form 20-F for the year ended December 31, 2022 (available at www.shell.com/investor and www.sec.gov). These risk factors also expressly qualify all forward-looking statements contained in this presentationand should be considered by the reader. Each forward-looking statement speaks only as of the date of this presentation, September 2023. Neither Shell plc nor any of its subsidiaries undertake any obligation to publicly update or revise any forward-looking statement as a result of new information, future events or other information. In light of these risks, results could differ materially from those stated, implied or inferred from the forward-looking statements contained in this presentation.

Shell's net carbon intensity Also, in this presentationwe may refer to Shell's "Net Carbon Intensity", which includes Shell's carbon emissions from the production of our energy products, our suppliers' carbon emissions in supplying energy for that production and our customers' carbon emissions associated with their use of the energy products we sell. Shell only controls its own emissions. The use of the term Shell's "Net Carbon Intensity" is for convenience only and not intended to suggest these emissions are those of Shell plc or its subsidiaries.

Shell's net-Zero Emissions Target Shell's operating plan, outlook and budgets are forecasted for a ten-year period and are updated every year. They reflect the current economic environment and what we can reasonably expect to see over the next ten years. Accordingly, they reflect our Scope 1, Scope 2 and Net Carbon Intensity (NCI) targets over the next ten years. However, Shell's operating plans cannot reflect our 2050 net-zero emissions target and 2035 NCI target, as these targets are currently outside our planning period. In the future, as society moves towards net-zero emissions, we expect Shell's operating plans to reflect this movement. However, if society is not net zero in 2050, as of today, there would be significant risk that Shell may not meet this target.

Forward Looking Non-GAAP measures This presentationmay contain certain forward-looking non-GAAP measures such as cash capital expenditure and divestments. We are unable to provide a reconciliation of these forward-looking Non-GAAP measures to the most comparable GAAP financial measures because certain information needed to reconcile those Non-GAAP measures to the most comparable GAAP financial measures is dependent on future events some of which are outside the control of Shell, such as oil and gas prices, interest rates and exchange rates. Moreover, estimating such GAAP measures with the required precision necessary to provide a meaningful reconciliation is extremely difficult and could not be accomplished without unreasonable effort. Non-GAAP measures in respect of future periods which cannot be reconciled to the most comparable GAAP financial measure are calculated in a manner which is consistent with the accounting policies applied in Shell plc's consolidated financial statements. The contents of websites referred to in this presentationdo not form part of this presentation. We may have used certain terms, such as resources, in this presentationthat the United States Securities and Exchange Commission (SEC) strictly prohibits us from including in our filings with the SEC. Investors are urged to consider closely the disclosure in our Form 20-F, File No 1-32575, available on the SEC website www.sec.gov.

Motivation

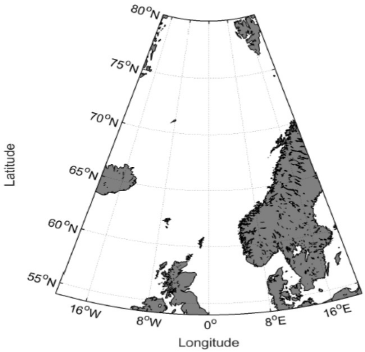
- In risk analysis, the **modelling of the extremes of natural phenomena** such as rainfall, temperature, winds and waves is needed
- Particularly interested in the largest events that we **might have seen or could possibly see**
- Statistical methods can be used to model these largest events and provide the basis for design criteria
- Want to be able to also **incorporate information about important covariates** e.g., direction or season that influence these natural phenomena
- Focus on oceanographic applications but covXtreme can be used more generally for **non-stationary multivariate extreme value analysis**

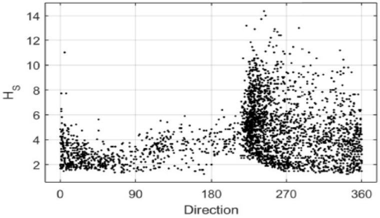
Motivation

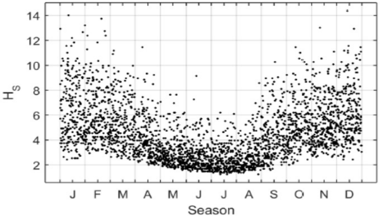
- Offshore operations **require the probability of failure** of manned structures and ships to be at the level of p=1e-4 per annum, corresponding to the so called **1 in 10 000 year criterion**
- This requires the understanding of the **extreme natural environment**:
  - Extreme behaviour of waves, winds and currents individually
  - Joint behaviour of waves, winds and currents
  - Impact of covariates such as direction and the time of year
- Want to be able to **propagate and quantify uncertainty** related to modelling extremes of oceanographic data

Motivation

Oceanographic data







Motivation

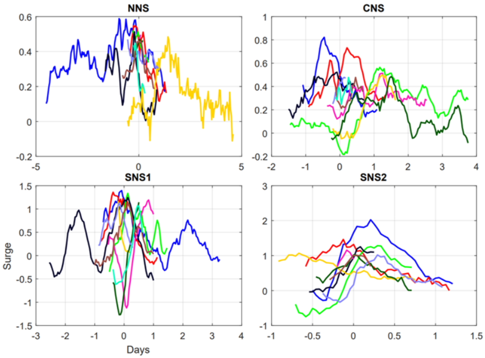
- Statistical tool should handle the following features:
  - Accurate estimation of the **tails of a data set**
  - Capture **covariate effects** such as direction and season
  - Account for the interaction between **multiple variables**
  - Careful handling of **uncertainty**

- As a result, we have developed covXtreme, a open source MATLAB software for the estimation of extreme conditions

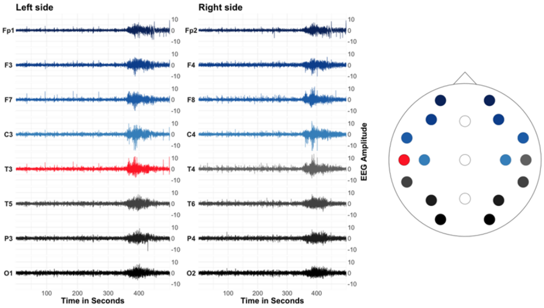
covXtreme

- covXtreme steps the user through a series of stages that result in the completion of a **full hazard risk analysis**
- MATLAB code is written in a **flexible modular way**
- Accompanied by an **user guide** that steps through two case studies
- A range of user settings can be specified - default settings are also provided
- Previous use of the code include Ross et al. [2018], Ross et al. [2020], Guerrero et al. [2021] and Barlow et al. [2023], example applications include surges, waves and neurology

Previous applications of covXtreme



Surge trajectories [Ross et al., 2018]



Brain signals [Guerrero et al., 2021]

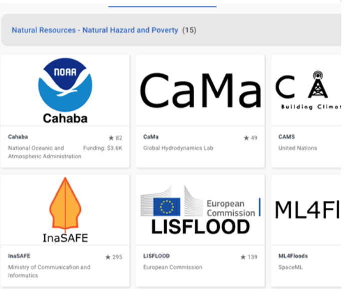
covXtreme

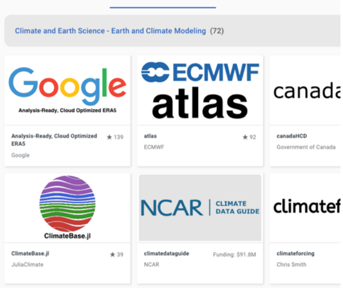
- **Stage 1:** selection of extreme events from an environmental data set or simulation of a data set: selection of independent events
- **Stage 2:** selection of covariate bins, for example wave height as a function of direction: capture covariates for upcoming marginal modelling
- **Stage 3:** estimation of marginal models with respect to covariates: non-stationary modelling as a function of covariate bin
- **Stage 4:** joint estimation of oceanographic variables, for example the behaviour of wind speed when wave height is large: account for interaction between multiple variables
- **Stage 5:** estimation of environmental contours for risk assessment: interpretable summary for design engineers

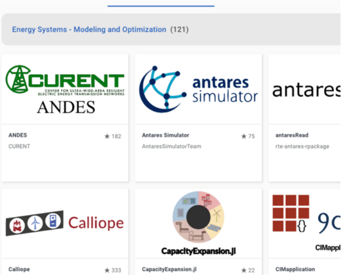
How covXtreme has been useful in Shell?

- Development of code bases such as covXtreme maintains **Shell's reputation as a leader in the sector**
- Application of covXtreme has yielded **cost savings and improved safety offshore**
- covXtreme is used for testing and scoping of improvements to Shell's proprietary risk analysis software
- **Established track record** of covXtreme through extensive testing using relevant case studies
- Ability to do an **efficient system analysis of offshore risk**
- Use in **upskilling and training** of new staff and working with academic and industry partners

How does covXtreme fit in LF Energy’s landscape?







How could this code be useful more generally?

- **Educational** tool for the offshore sector: users can be data scientists, statisticians and practitioners
- covXtreme **motivates and explains** how natural hazard analysis problems are solved
- **First openly available tool** that brings together important elements of an risk analysis:
  - modelling a single variable
  - modelling multiple variables
  - incorporating the effect of covariates
  - system based response estimates
- covXtreme has the **key functionality to solve a typical risk analysis problem in a pragmatic manner**

Public enagagement and future development

- covXtreme is **openly available through GitHub**: `https://github.com/sede-open/covXtreme`
- Journal article detailing methodology behind covXtreme is under review
- Presentations at statistics conferences e.g., RSS Conference (2023)
- Being **actively used by existing collaborators**
- External partners can build additional functionality into the code
- Opportunity to publicise with existing community and potentially interested communities, for example the flood risk sector

Potential future applications of covXtreme

- Pluvial and fluvial flooding
- Coastal risk of flooding
- Accounting for regime shifts in data sets e.g., financial or temperature data
- Operational maintenance of platforms, ships and wind turbines

BBC

Summary

- covXtreme enables **quick analysis of extreme data sets**
- **Computationally efficient and pragmatic** software for hazard risk analysis
- Non stationary **marginal and dependence modelling** with comprehensive uncertainty quantification
- **Improved quantification and communication** of risks associated with extreme events

References

A M Barlow, E Mackay, E Eastoe, and P Jonathan. A penalised piecewise-linear model for non-stationary extreme value analysis of peaks over threshold. *Ocean Eng.*, 2023.

Matheus B. Guerrero, Raphaël Huser, and Hernando Ombao. Conex-connect: Learning patterns in extremal brain connectivity from multi-channel eeg data, 2021. URL `https://arxiv.org/abs/2101.09352`.

J. E. Heffernan and J. A. Tawn. A conditional approach for multivariate extreme values. *J. R. Statist. Soc. B*, 66:497–546, 2004.

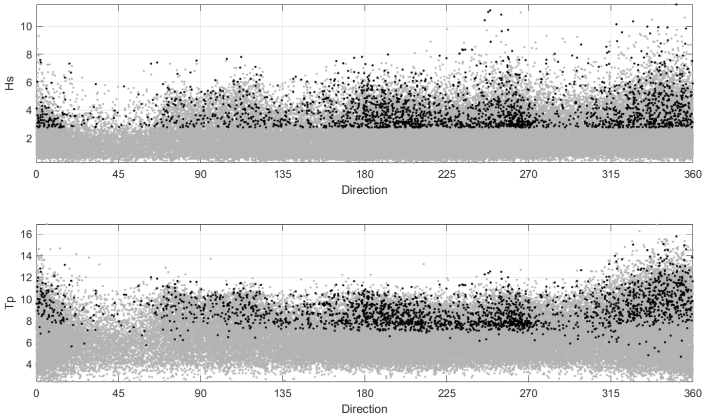
E Ross, S Sam, D Randell, G Feld, and P Jonathan. Estimating surge in extreme North Sea storms. *Ocean Eng.*, 154:430–444, 2018.

E. Ross, O. C. Astrup, E. Bitner-Gregersen, N. Bunn, G. Feld, B. Gouldby, A. Huseby, Y. Liu, D. Randell, E. Vanem, and P. Jonathan. On environmental contours for marine and coastal design. *Ocean Eng.*, 195:106194, 2020.

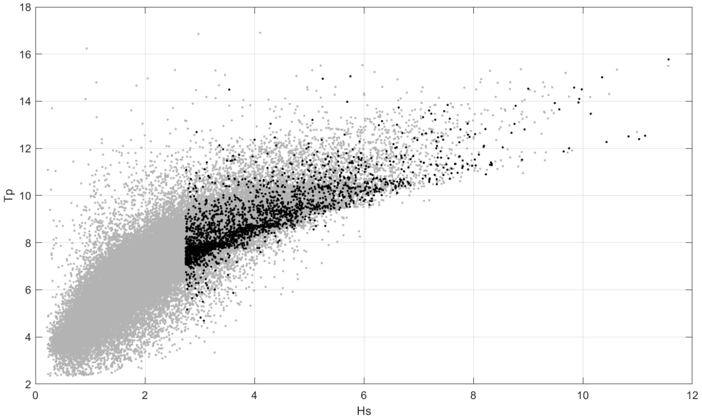
Case Study: Hs and Tp

Example application of covXtreme: modelling the relationship between significant wave height (Hs) and peak period (Tp)

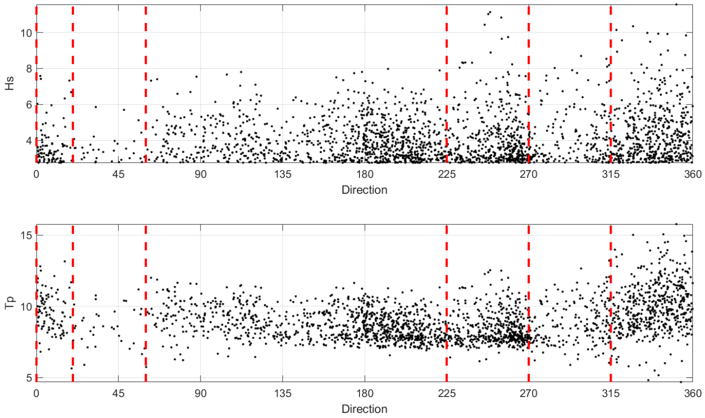
Stage 1: extraction of storm peaks



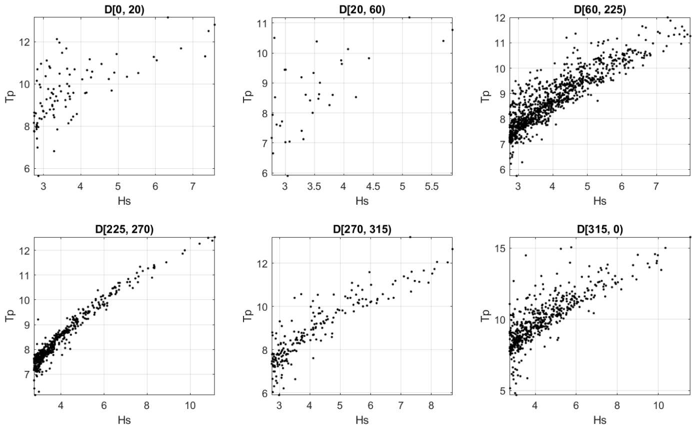
Stage 1: extraction of storm peaks



Stage 2: selection of bins



Stage 2: joint behaviour of Hs and Tp



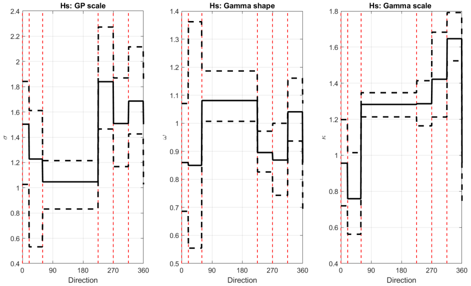
Stage 3: marginal model

- Set a bin dependent threshold ψ_b to define extreme events
- For data below the threshold fit a **Gamma distribution**
- For data above the threshold fit a **generalised Pareto (GP) distribution**:
  - Threshold ψ_b with scale ν_b and shape parameter ξ
- Likelihood above the threshold:

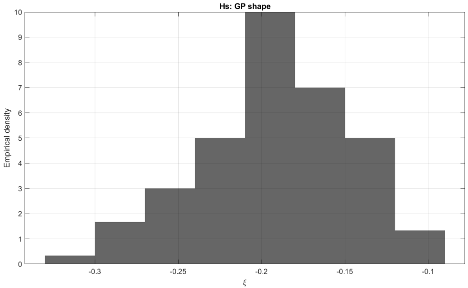
$$\ell(\dot{x}_i \mid \xi, \boldsymbol{\nu}, \boldsymbol{\psi}, \lambda) = \log \prod_{b=1}^{B} \prod_{\substack{i; A(i)=b; \\ \dot{x}_i > \psi_b}} f_{GP}(\dot{x}_i \mid \xi, \nu_b, \psi_b) + \lambda \left(\frac{1}{B} \sum_{b=1}^{B} \nu_b^2 - \left[\frac{1}{B} \sum_{b=1}^{B} \nu_b \right]^2 \right)$$

Stage 3: marginal model (Hs)

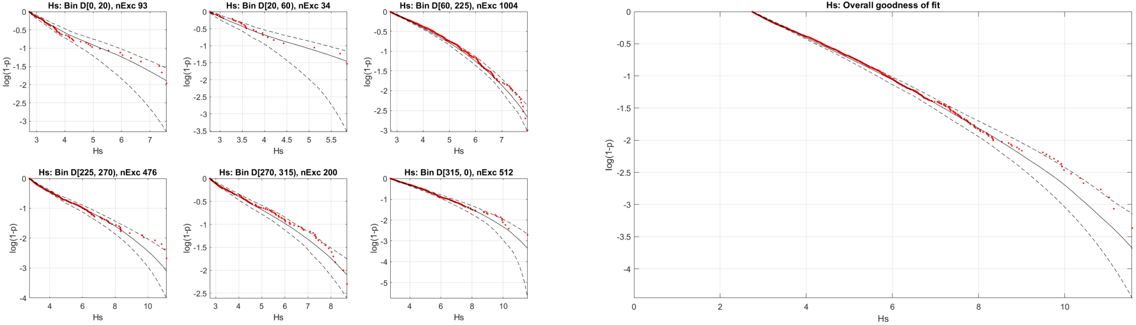
GP scale and Gamma parameters



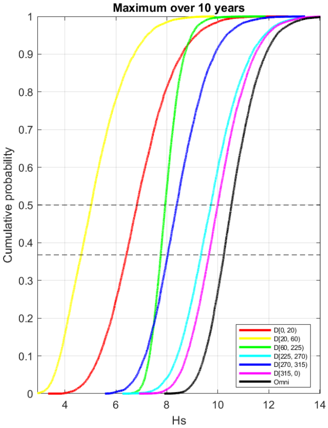
GP shape parameter

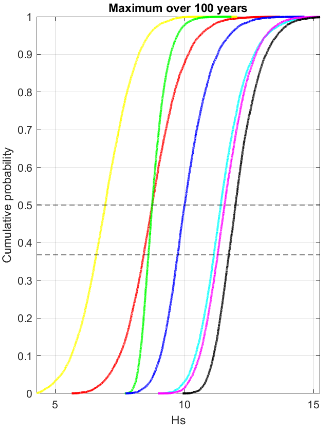


Stage 3: marginal model assessment (Hs)



Stage 3: marginal return values (Hs)



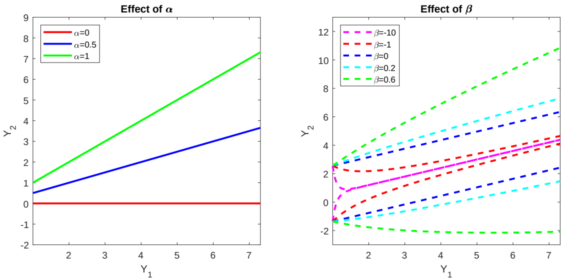


Stage 4: dependence model

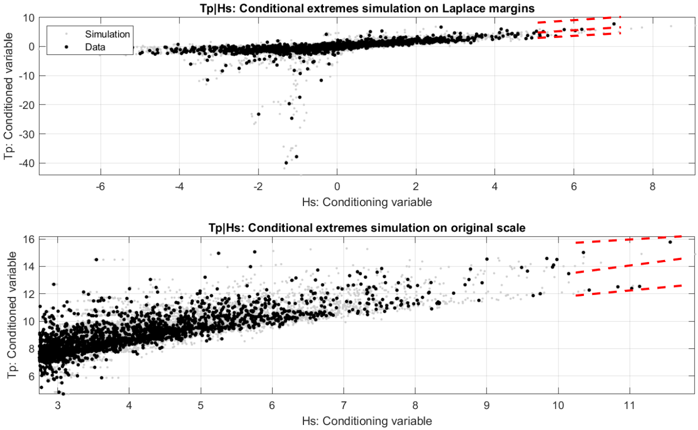
Conditional dependence model of Heffernan and Tawn [2004]:

$$(Y_2|Y_1 = y) = \alpha_b y + y^{\beta_b} W_b$$

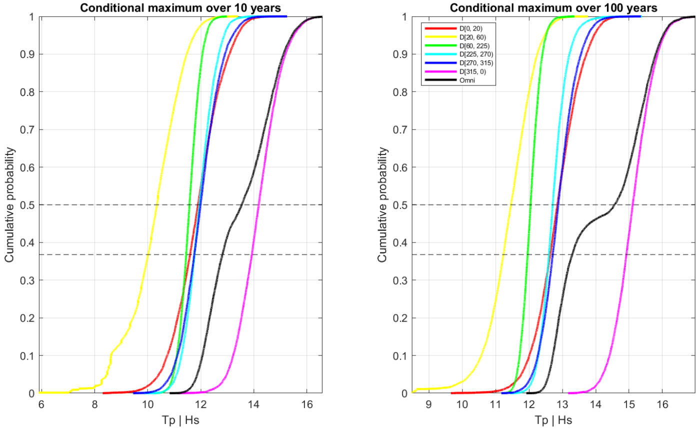
- $Y_2 = \text{Tp}$, $Y_1 = \text{Hs}$ *on Laplace scale*
- for $y >$ sufficiently large threshold ϕ
- $\alpha_b \in [-1, 1]$, $\beta_b \in (-\infty, 1]$
- $W_b \sim \text{DeltaLaplace}(\mu_b, \sigma_b, \delta)$



Stage 4 - simulations from the dependence model



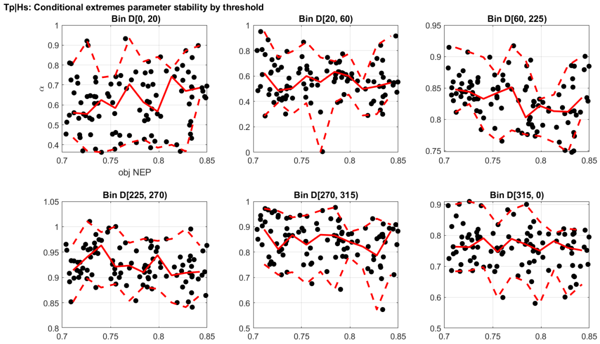
Stages 3 and 4: conditional return values (Tp|Hs)



Stages 3 and 4: dealing with uncertainty

Two sources of uncertainty:

- Bootstrap resampling
- Non exceedance probability threshold:
  $\phi \sim \text{Uniform}(\phi_{LB}, \phi_{UB})$



Dependence model threshold assessment

Stage 5: contour estimation

- Estimation of risk profiles
- Three different contour methods:
  - Exceedance (Exc)
  - Heffernan and Tawn (HTDns)
  - Huseby (Hus)
- Number of control factors