Non-stationary conditional extremes

philip.jonathan@shell.com
www.lancs.ac.uk/∼jonathan

Lancaster Extremes Reading Group, October 2013

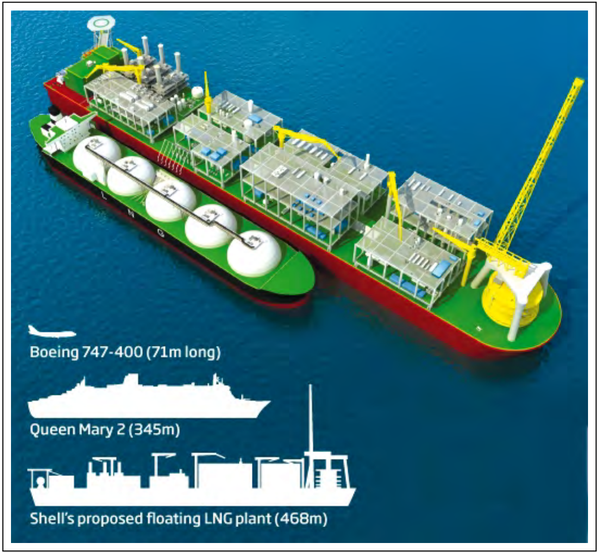
Thanks

Thanks for contributions by Shell colleagues:

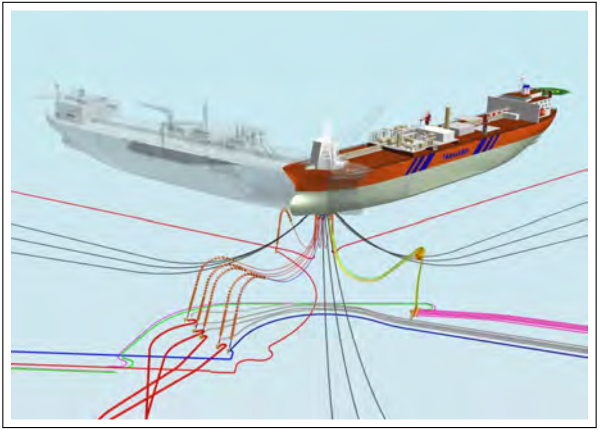
- Kevin Ewans, Graham Feld, David Randell, Yanyun Wu

... and Lancaster students:

- Kaylea Haynes, Emma Ross, Elena Zanini


Boeing 747-400 (71m long)
Queen Mary 2 (345m)
Shell's proposed floating LNG plant (468m)

Motivation: turret mooring

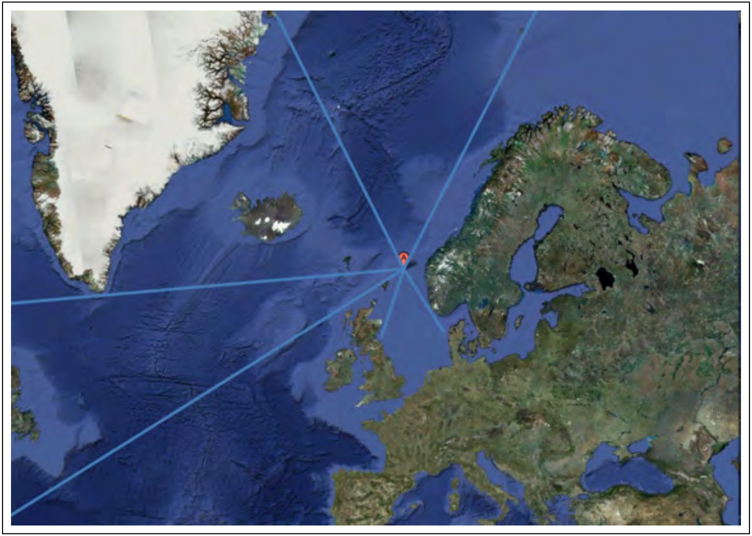


- Waves, winds, currents (all directional)

Extremes: multivariate challenges

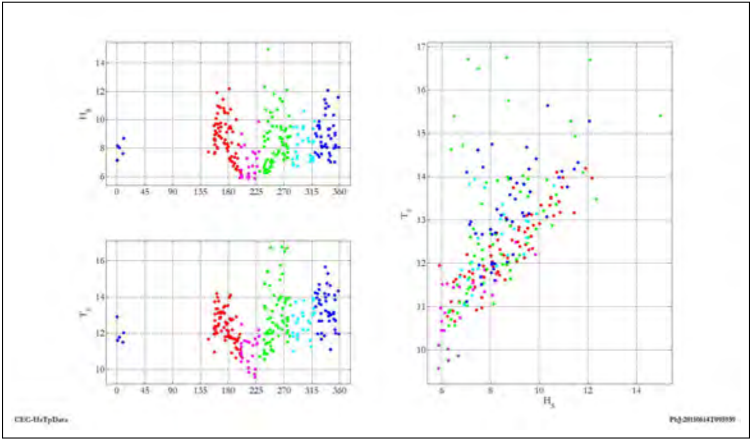
- **Spatial extremes using componentwise maxima**:
  - $\Leftrightarrow$ max-stability $\Leftrightarrow$ multivariate regular variation
  - Assumes all components extreme
  - $\Rightarrow$ Perfect independence or asymptotic dependence **only**
  - Composite likelihood for spatial extremes (Davison et al. 2012)
- **Extremal dependence**: (Ledford and Tawn 1997)
  - Assumes regular variation of joint survivor function
  - Gives more general forms of extremal dependence
  - $\Rightarrow$ Asymptotic dependence, asymptotic independence (with +ve, -ve association)
  - Hybrid spatial dependence model (Wadsworth and Tawn 2012)
- **Conditional extremes**: (Heffernan and Tawn 2004)
  - Assumes, given one variable being extreme, convergence of distribution of remaining variables
  - Allows some variables not to be extreme
  - Not equivalent to extremal dependence
- Application:
  - ... *a huge gap in the theory and practice of multivariate extremes* ... (Beirlant et al. 2004)

Application: location



- (Actually North West Shelf of Australia, South Atlantic Ocean)

Application: exploratory analysis



- Spread of T_P vs H_S different for different directions

Inference: objectives

Problem structure:

- Bivariate sample $\{\dot{x}_{ij}\}_{i=1,j=1}^{n,2}$ of random variables $\dot{X}_1, \dot{X}_2$
- Covariate values $\{\theta_{ij}\}_{i=1,j=1}^{n,2}$ associated with each individual

- For some choices of variables $\dot{X}$, e.g. $\dot{X}_1 = H_S$, $\dot{X}_2 = T_P$, $\theta_{i1} \triangleq \theta_{i2}$
- For other choices, e.g. $\dot{X}_1 = H_S$, $\dot{X}_2 =$WindSpeed, $\theta_{i1} \neq \theta_{i2}$ in general
- We will assume $\theta_{i1} = \theta_{i2} = \theta_i$

Objective:

- Objective: model the joint distribution of extremes of $\dot{X}_1$ and $\dot{X}_2$ as a function of θ

(Drop subscripts wherever possible for convenience)

Inference: outline

- Follows conditional extremes (Heffernan and Tawn 2004)
- Model $\dot{X}_1$ and $\dot{X}_2$ marginally as a function of θ
  - Quantile regression (QR) below threshold
  - Generalised Pareto (GP) above threshold
- Transform to standard Gumbel variates X_1 and X_2
- Model X_2 given large values of X_1 using non-stationary extension of conditional extremes model (incorporating θ)
- Simulate for long return periods
  - Generate samples of joint extremes on Gumbel scale
  - Transform to original scale
- Simulate structure variables $f(\dot{X}_1, \dot{X}_2)$ also

Penalised B-splines

- Physical considerations suggest model parameters vary smoothly with covariates θ
- A typical parameter η on (an index) set of covariates all take the form:

  $$\eta = B\beta_\eta$$

  for **B-spline** basis matrix B (defined on index set of covariate values) and some β_η to be estimated
- Multidimensional basis matrix B formulated using Kronecker products of marginal basis matrices:

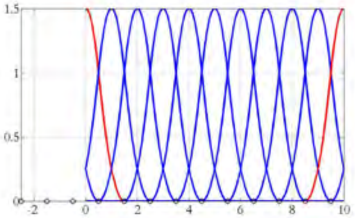
  $$B = B_\theta \otimes B_x \otimes B_y$$

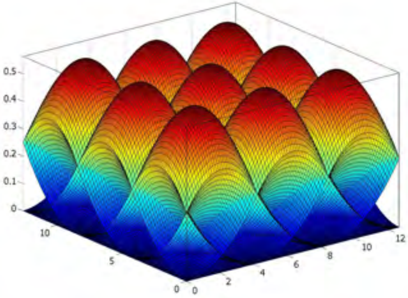
- Roughness R_η defined as:

  $$R_\eta = \beta_\eta' P \beta_\eta$$

  where effect of P is to difference neighbouring values of β_η

- **Wrapped** bases for periodic covariates (seasonal, direction)
- **Multidimensional** bases easily constructed. **Problem size** sometimes prohibitive
- Parameter **smoothness** controlled by roughness coefficient λ: cross validation chooses λ optimally
- We stick to a single dimension here





Non-stationary marginal threshold exceedances

For sufficiently large threshold, the $\dot{X}$s are marginally independently distributed according to:

$$Pr(\dot{X} > \dot{x}_i | \dot{X} > \phi_{\tau'}(\theta_i)) = (1 + \frac{\xi_i}{\zeta_i}(\dot{x}_i - \phi_{\tau'}(\theta_i)))^{-\frac{1}{\xi_i}}$$

where:

- $\phi_{\tau'}(\theta)$ is a quantile threshold with non-exceedance probability τ'
- $\xi_i = \xi(\theta_i)$ and $\zeta_i = \zeta(\theta_i)$
- ϕ, ξ and ζ are smooth functions

Use diagnostics to select an appropriate threshold level τ':

- Q-Q plot
- Stability of $\xi(\theta)$ with θ

Non-stationary marginal below threshold

The unconditional cumulative distribution function for threshold excesses is:

$$\begin{aligned} F(\dot{x}_i) &= Pr(\dot{X} \leq \dot{x}_i) \\ &= 1-(1-\tau^*)(1+\frac{\xi_i}{\zeta_i}(\dot{x}_i - \phi_{\tau'}(\theta_i)))^{-\frac{1}{\xi_i}} \quad \dot{x}_i > \phi_{\tau'}(\theta_i) \\ &= \tau_L + (\tau_H - \tau_L)\frac{(\dot{x}_i - \phi_{\tau_L}(\theta_i))}{(\phi_{\tau_H}(\theta_i) - \phi_{\tau_L}(\theta_i))} \quad \dot{x}_i \leq \phi_{\tau'}(\theta_i) \end{aligned}$$

where $\{\tau_d\}_{d=1}^{D}$ is a set of threshold probabilities for which quantile thresholds $\phi_{\tau_d}(\theta)$ have been estimated, and:

$$H = \arg\min_d \{\phi_{\tau_d}(\theta_i) \geq \dot{x}_i\}$$

with $L = H - 1$.

Typically we would have $\{\tau_d\}_{d=1}^{D} = 0.1, 0.2, ..., 0.9$ say, and evaluate quantile regressions for each. We would choice the smallest value for which GP gives good marginal fit, then use quantiles corresponding to smaller values to approximate the CDF

Threshold estimation: quantile regression

- Data $\{\theta_i, \dot{x}_i\}_{i=1}^n$
- In vector terms on the set $\{\theta_i\}_{i=1}^n$, τ^{th} conditional quantile function $\phi_\tau(\theta)$ is:

$$\phi_\tau = B\beta_{\phi\tau}$$

- Estimated by minimising criterion $\ell_{\phi\tau}$:

$$\ell_{\phi\tau} = \{\tau \sum_{r_i \geq 0}^{n} |r_i| + (1-\tau) \sum_{r_i < 0}^{n} |r_i|\}$$

in terms of residuals:

$$r_i = \dot{x}_i - \phi_\tau(\theta_i)$$

Roughness-penalised quantile regression

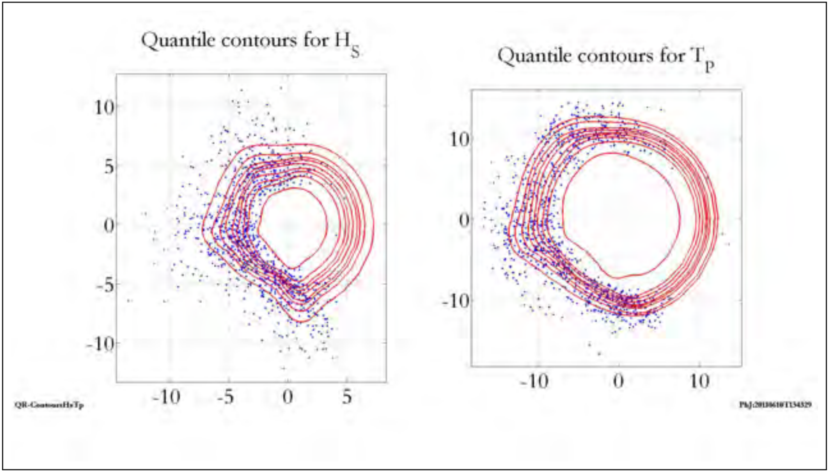
Use penalised criterion $\ell^*_{\phi\tau}$ instead of $\ell_{\phi\tau}$:

$$\ell^*_{\phi\tau} = \ell_{\phi\tau} + \lambda_{\phi\tau} R_{\phi\tau}$$

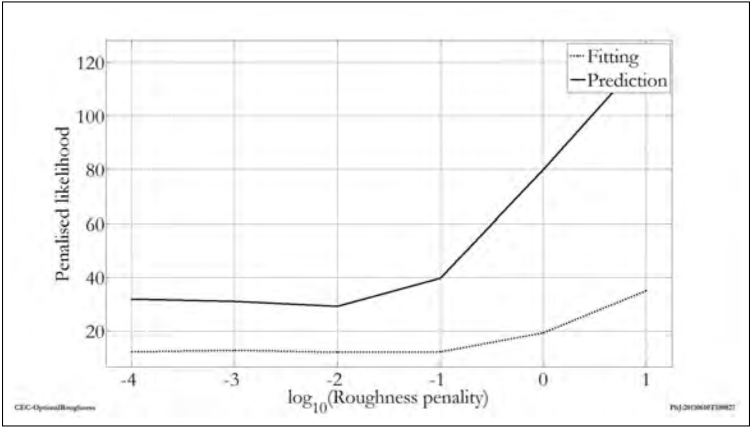
where $R_{\phi\tau}$ is the roughness of ϕ_τ, regulated using $\lambda_{\phi\tau}$, chosen using cross-validation or similar.

Solved simultaneously for set $\{\tau_d\}_{d=1}^{D}$ using linear programming.

Regression quantiles



- Transform directions to uniform prior using QR estimation
- Deciles to 80%



- Penalty of approximately 0.1 appropriate

Exceedance sizes: Generalised Pareto

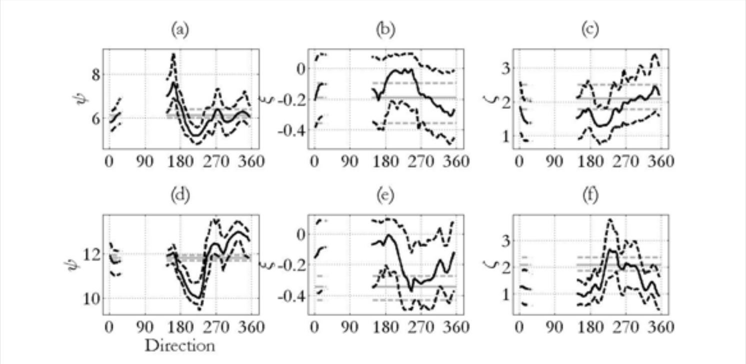
- Generalised Pareto model for size of threshold exceedance estimated by minimising roughness penalised log-likelihood:

$$\ell^*_{\xi,\zeta} = \ell_{\xi,\zeta} + \lambda_\xi R_\xi + \lambda_\zeta R_\zeta$$

- (Negative) conditional generalised Pareto log-likelihood:

$$\ell_{\xi,\zeta} = \sum_{i=1}^{n} \log \zeta_i + \frac{1}{\xi_i} \log(1 + \frac{\xi_i}{\zeta_i}(\dot{x}_i - \phi_{\tau'}(\theta_i)))$$

- Parameters: shape ξ, scale ζ
- Threshold $\phi_{\tau'}$ from quantile regression
- λ_ξ and λ_ζ estimated using cross validation. In practice set $\lambda_\xi = \kappa\lambda_\zeta$ for fixed κ



- Top line: H_S, bottom line: T_P
- Left to right: threshold, shape, scale
- Grey: stationary
- 95% bootstrap uncertainty bands also

Transformation to Gumbel scale

Transform sample $\{\dot{x}_i\}_{i=1}^{n}$ to sample $\{x_i\}_{i=1}^{n}$ on Gumbel scale using probability integral transform:

$$\exp(-\exp(-x_i)) = Pr(X \leq x_i) = Pr(\dot{X} \leq \dot{x}_i) \text{ from above}$$

Conditional extremes

On Gumbel scale, by analogy with Heffernan and Tawn [2004] we propose the following conditional extremes model:

$$(X_k | X_j = x_j, \theta) = \alpha_\theta x_j + x_j^{\beta_\theta}(\mu_\theta + \sigma_\theta Z) \text{ for } x_j > \phi^G_{j\tau'}(\theta)$$

where:

- $\phi^G_{j\tau'}(\theta)$ is a high directional quantile of X_j on Gumbel scale, above which the model fits well
- $\alpha_\theta \in [0, 1]$, $\beta_\theta \in (-\infty, 1]$, $\sigma_\theta \in [0, \infty)$
- Z is a random variable with unknown distribution G
- Z will be assumed to be approximately Normally distributed for the purposes of parameter estimation

Penalised likelihood

For sample $\{x_{ik}, x_{ij}, \theta_i\}_{i=1}^m$ corresponding to threshold exceedances $\{x_{ij}\}_{i=1}^m$ of $\phi^G_{j\tau'}$, negative log likelihood $\ell_{\alpha,\beta,\mu,\sigma}$ is given by:

$$\ell_{\alpha,\beta,\mu,\sigma} = \sum_{i=0}^{n} \log s_i + \frac{(x_{ik} - m_i)^2}{2s_i^2}$$

where:

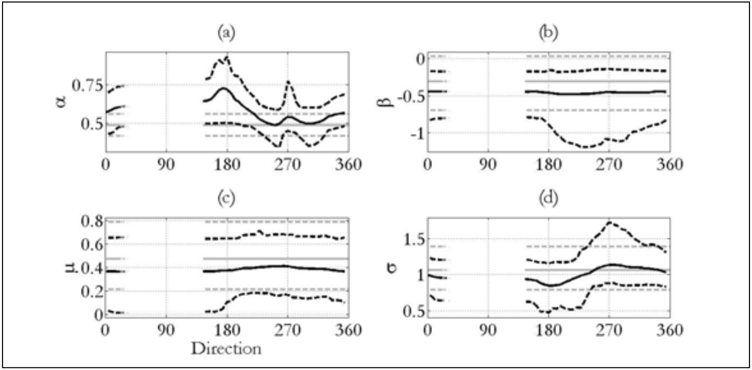
$$m_i = m_i(x_{ij}, \theta_i) = \alpha(\theta_i)x_{ij} + \mu(\theta_i)x_{ij}^{\beta(\theta_i)}$$
$$s_i = s_i(x_{ij}, \theta_i) = \sigma(\theta_i)x_{ij}^{\beta(\theta_i)}$$

Penalised negative log likelihood ℓ^* is given by

$$\ell^*_{\alpha,\beta,\mu,\sigma} = \ell_{\alpha,\beta,\mu,\sigma} + \lambda_\alpha R_\alpha + \lambda_\beta R_\beta + \lambda_\mu R_\mu + \lambda_\sigma R_\sigma$$

Imposing non-negativity: We choose to model $\sqrt{\alpha}$ and $\sqrt{\sigma}$ so that their squares are non-negative. Roughness penalty estimated using cross-validation.

Parameter estimates



Estimating T_P associated with extreme quantile of H_S

Given parameter estimates and sample of residuals:

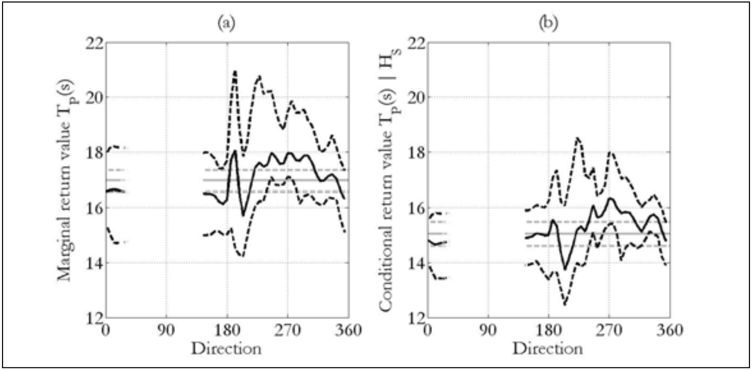
- Estimate quantiles of T_P given any quantile of H_S on Gumbel scale

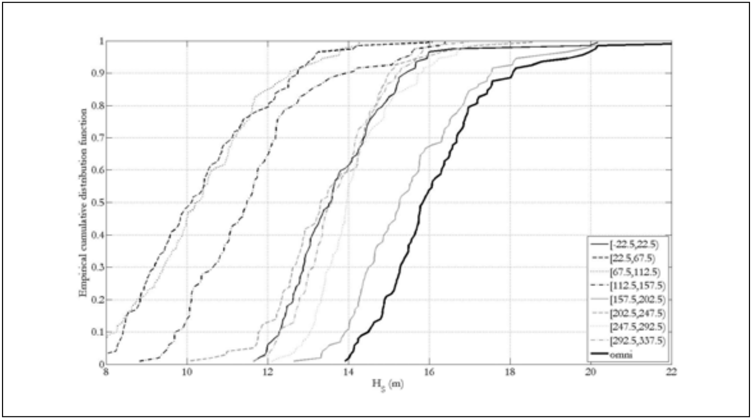
$$(T_P | H_S = h, \theta) = \hat{\alpha}_\theta h + h^{\hat{\beta}_\theta}(\hat{\mu}_\theta + \hat{\sigma}_\theta Z) \text{ for } h > \phi^G(\theta, \tau^G_{j*})$$

- Transform to original scale

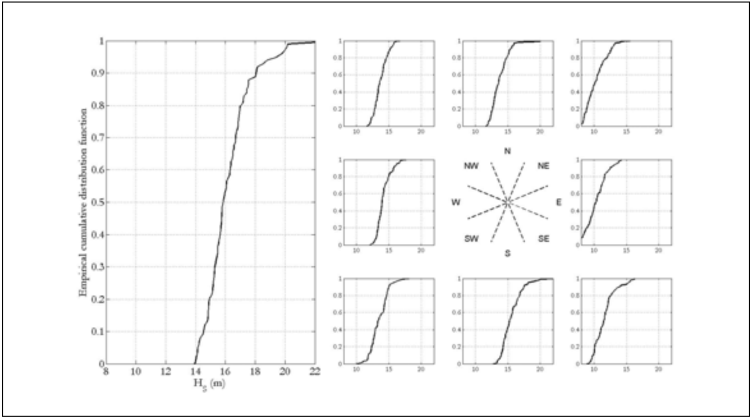
Compare with model ignoring covariate effects

Return values



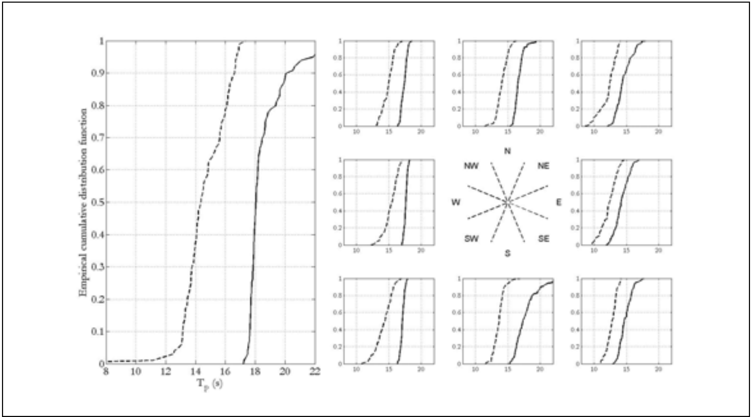


- Return values from simulation: maximum H_S per octant for 100-year return period.
- Octants centred on cardinal directions.
- Largest storms up North Sea from south.



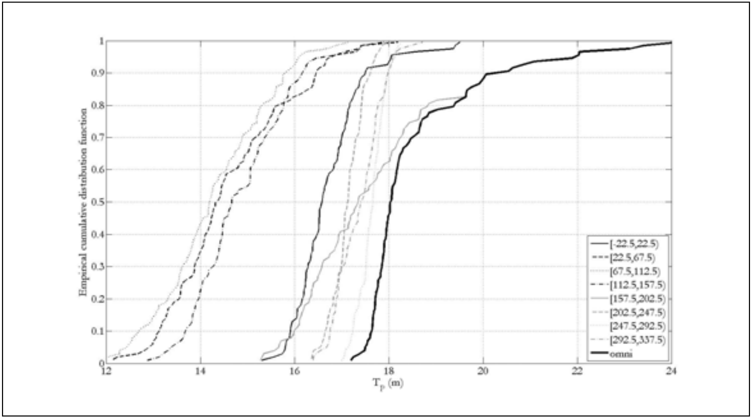
- Return values from simulation: marginal H_S for 100-years.
- Octants centred on cardinal directions.

Return values



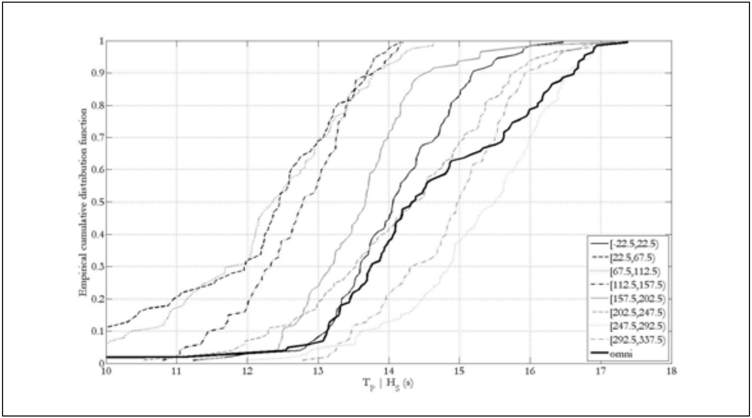
- Marginal maximum T_P for 100-year return period per octant (solid).
- Conditional T_P given 100-year H_S per octant (dashed).

Return values



- Marginal maximum T_P for 100-year return period per octant.
- Largest T_P from south and from Atlantic.

Return values



- Conditional T_P given 100-year H_S per octant.
- Largest conditional T_P from Atlantic.

Limit assumption

Extension of Heffernan and Tawn [2004]. The limit assumption required to justify the conditional model is:

$$Pr\left(\frac{x_j^{-\beta_\phi}(X_k - \alpha_\phi x_j) - \mu_\phi}{\sigma_\phi} \leq z | X_j = x_j, \theta\right) \rightarrow G(z) \text{ as } x_j \rightarrow \infty$$

Simulation study 1

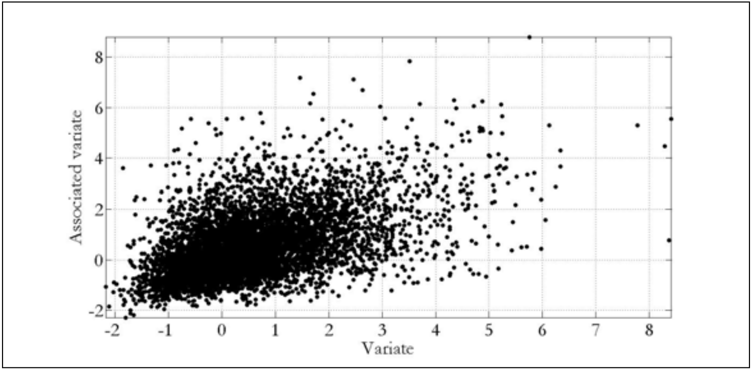
- Bivariate distribution with Normal dependence transformed marginally to standard Gumbel

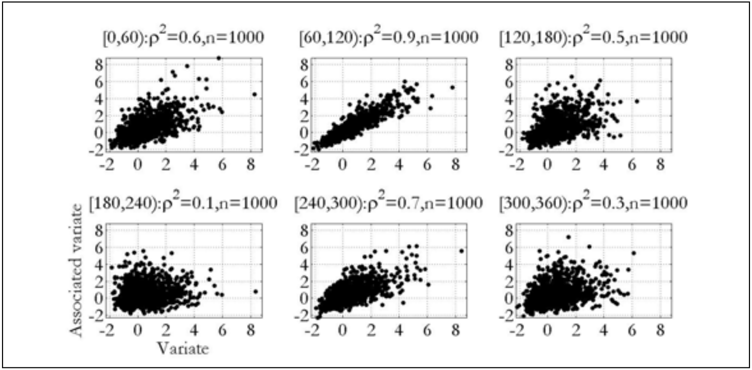
$$
\begin{aligned}
(X_1(\theta), X_2(\theta)) &= -\log(-\log(\Phi_{\Sigma(\theta)}(X_{1N}, X_{2N}))) \\
(X_2(\theta) | X_1(\theta) = x) &= \rho^2(\theta)x + x^{1/2}W(\theta) \text{ for large } x
\end{aligned}
$$

- 6 directional intervals: $\rho^2 = 0.6, 0.9, 0.5, 0.1, 0.7, 0.3$
- Sample size 1000×6

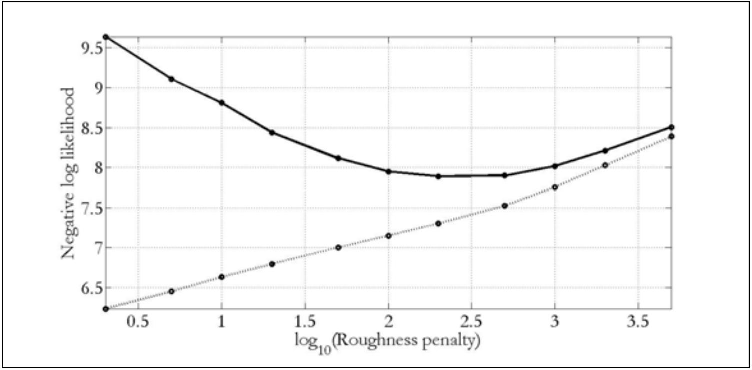
- Marginals assumed known
- Estimate conditional model only
- $\alpha = \rho^2$ and $\beta = 1/2$.

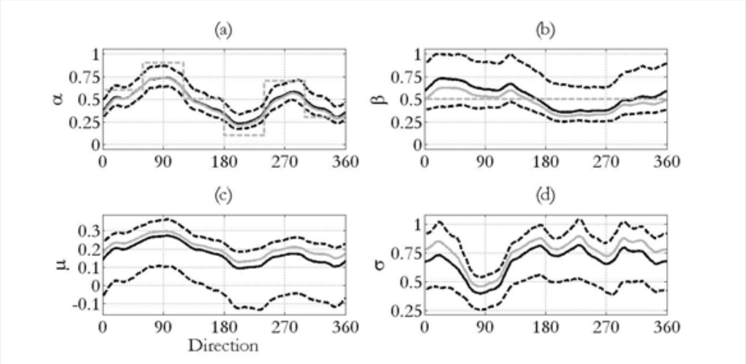
Study 1: sample

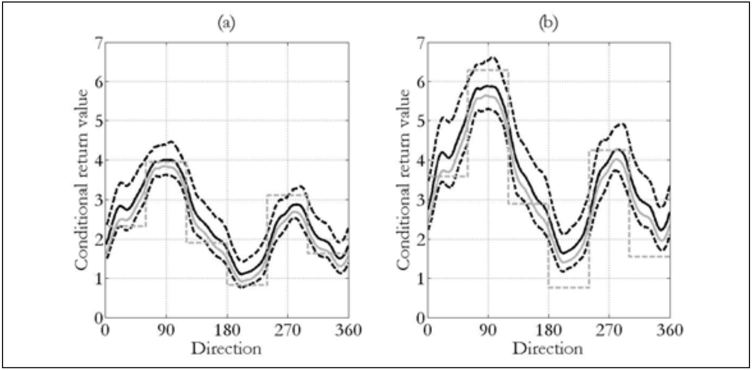



[0,60):ρ²=0.6,n=1000
[60,120):ρ²=0.9,n=1000
[120,180):ρ²=0.5,n=1000
[180,240):ρ²=0.1,n=1000
[240,300):ρ²=0.7,n=1000
[300,360):ρ²=0.3,n=1000
Associated variate
Variate

Study 1: roughness coefficient choice




(a)
1
0.75
0.5
0.25
0
α
0
90
180
270
360
(b)
β
(c)
0.3
0.2
0.1
0
-0.1
μ
Direction
(d)
σ


(a)
(b)
Conditional return value
Direction
0
90
180
270
360
0
1
2
3
4
5
6
7

Simulation study 2

- Mixture of bivariate distribution with Normal dependence transformed marginally to standard Gumbel, and bivariate extreme value distribution with exchangeable logistic dependency and Gumbel marginal distributions. Same intervals of θ.

For $\theta \in [0, 180)$:

- Dependence structure of study 1 with $\rho^2 = 0.8, 0.1, 0.8$.

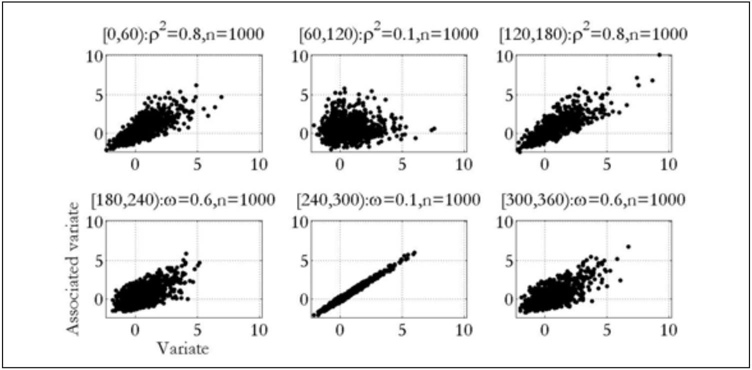
For $\theta \in [180, 360)$:

- Logistic dependence structure

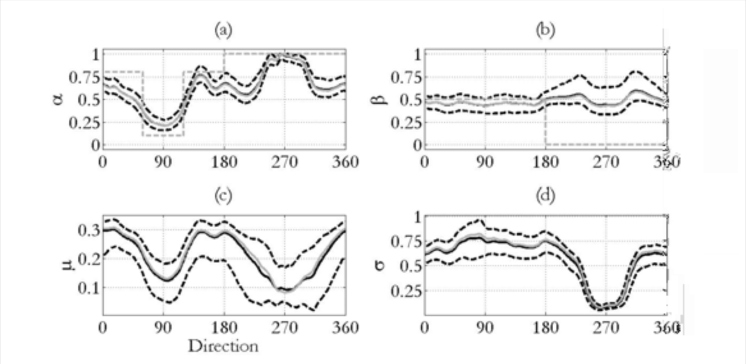
$$\begin{aligned} \Pr(X_1(\theta) \leq x_1, X_2(\theta) \leq x_2) &= \exp(-(\exp(x_1/\omega(\theta)) + \exp(x_2/\omega(\theta)))^{\omega(\theta)}) \\ (X_2(\theta)|X_1(\theta) = x) &= x + Z(\theta) \text{ for large } x \end{aligned}$$

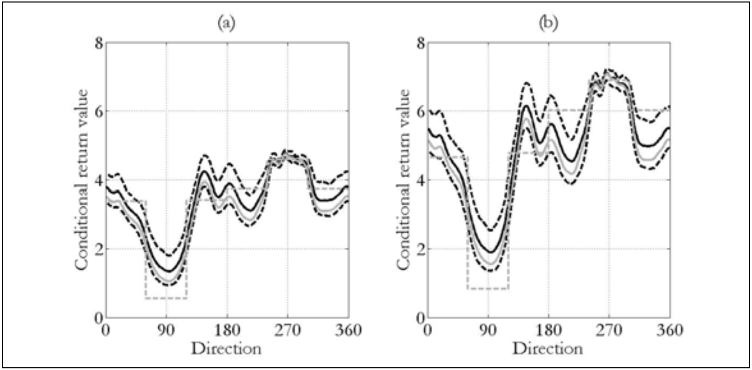
- $\omega = 0.6, 0.1, 0.6$
- Value of $\omega (< 1)$ has no effect on asymptotic conditional dependence structure.
- $\alpha = 1$ and $\beta = 0$.

Study 2: partitioned



Study 2: parameter estimates




(a)
(b)
Conditional return value
Conditional return value
0
2
4
6
8
0
90
180
270
360
Direction
Direction

Conclusions

- Non-stationary extension of conditional extremes method
- Requires efficient estimation of covariate effects (penalised B-splines here)
- Makes engineering application of conditional extremes model feasible, particularly for floating structures

References

- 2004: Heffernan and Tawn: A conditional approach for multivariate extreme values, J. R. Statist. Soc. B, v66, p497.
- 2005: Koenker: Quantile regression, Cambridge University Press.
- 2010: Eilers and Marx: Splines, knots and penalties, Computat. Statistics, v2, p637.
- 2013: Jonathan, Ewans and Randell, Non-stationary conditional extremes, Submitted to Environmetrics, draft at www.lancs.ac.uk/∼jonathan.

Thanks for listening

philip.jonathan@shell.com

References

A. C. Davison, S. A. Padoan, and M. Ribatet. Statistical modelling of spatial extremes. *Statistical Science*, 27:161–186, 2012.

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

A. W. Ledford and J. A. Tawn. Modelling dependence within joint tail regions. *J. R. Statist. Soc. B*, 59:475–499, 1997.

J.L. Wadsworth and J.A. Tawn. Dependence modelling for spatial extremes. *Biometrika*, 99:253–272, 2012.