Characterising extreme ocean environments

David Randell [1], Philip Jonathan [2,3]

[1]Shell Global Solutions International BV, Amsterdam, NL.
[2]Shell Research Ltd., London, UK.
[3]Lancaster University, Department of Mathematics & Statistics, UK.

Statistics seminar, Exeter
(Slides at www.lancs.ac.uk/∼jonathan)

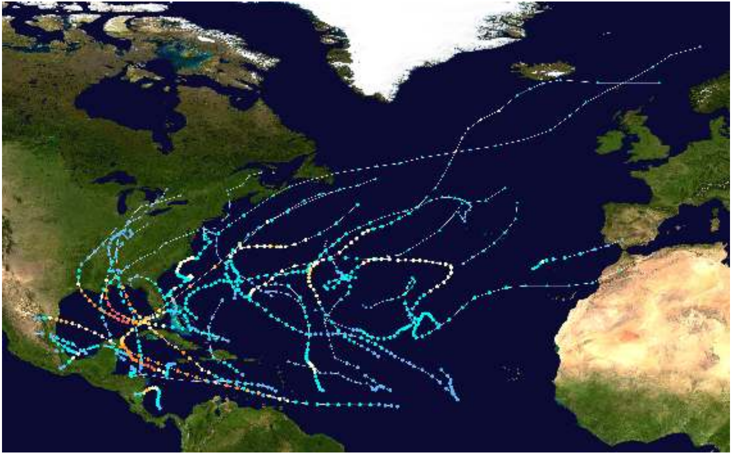
Acknowledgement

- Lancaster
- Shell
- Durham

Katrina

August 2015 (NOAA geostationary orbiting environmental satellite)

Hurricane tracks

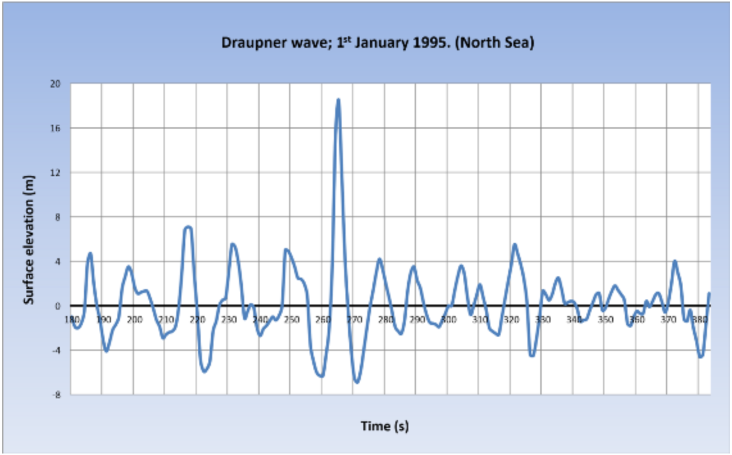


Summer 2005 (NASA, US National Hurricane Center)

Portugese coast

24m wave height, November 2017 (The Guardian)

Draupner



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

Roker Pier

Sunderland, every winter! (Daily Express)

Ship damage

Norwegian Dream, Atlantic, 2007 (gcaptain.com)

Wilstar, Agulhas current (Oceanography **18** 2005)

Structural damage

Ike, Gulf of Mexico, 2008 (Joe Richard)

North Sea, Winter 2015-16 (The Inertia)

Motivation

- Rational and consistent design and assessment of marine structures
  - Reduce bias and uncertainty in estimation of structural integrity
  - Quantify uncertainty as well as possible
- Non-stationary marginal, conditional, spatial and temporal extremes
  - Multiple locations, multiple variables, time-series
  - Multidimensional covariates
- Improved understanding and communication of risk
  - Incorporation within established engineering design practices
  - Knock-on effects of improved inference

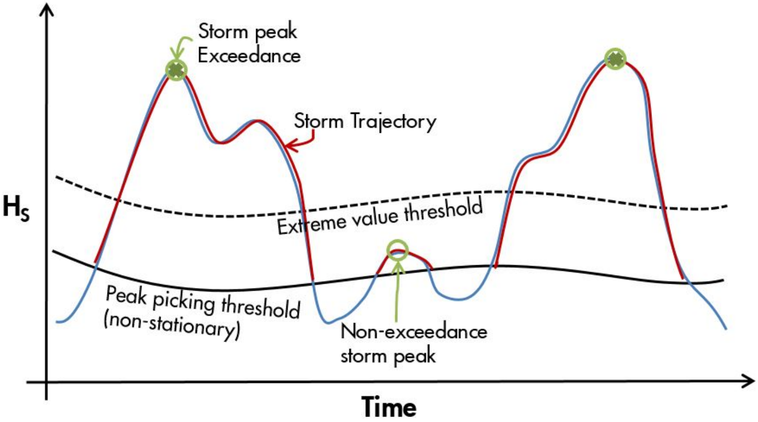
The ocean environment is an amazing thing to study ... especially if you like to combine beautiful physics, measurement and statistical modelling!

Fundamentals

- Environmental extremes vary smoothly with multidimensional covariates
  - Model parameters are non-stationary
- Environmental extremes exhibit spatial and temporal dependence
  - Characterise these appropriately
- Uncertainty quantification for whole inference
  - Data acquisition (simulator or measurement)
  - Data pre-processing (storm peak identification)
  - Hyper-parameters (extreme value threshold)
  - Model form (marginal measurement scale effect, spatial extremal dependence)
- Statistical and computational efficiency
  - Slick algorithms
  - Parallel computation
  - Bayesian inference

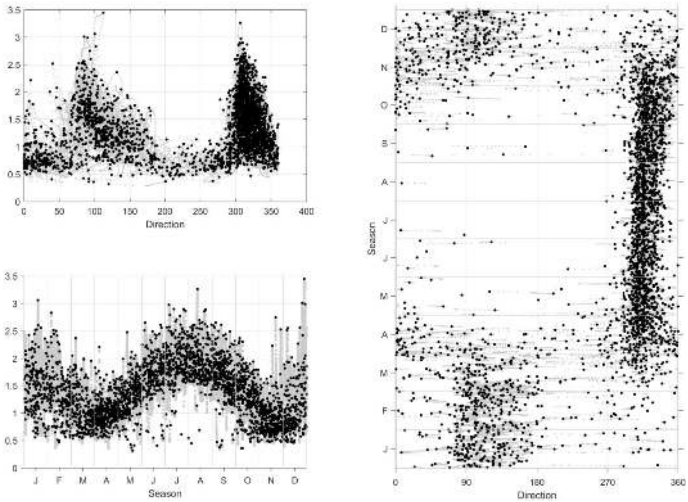
Storm model

$H_S \approx 4\times$ standard deviation of ocean surface time-series at a location corresponding to a time period (typically three hours)



A typical sample

Typical data for South China Sea location. Sea state (grey) and storm peak (black) H_S on season and direction



Outline

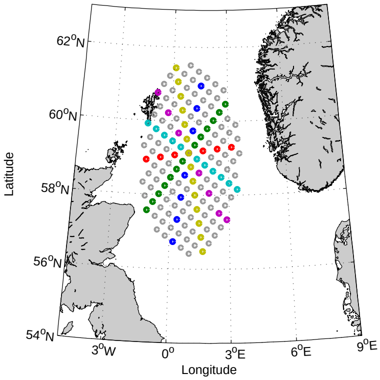
Covariate effects in:

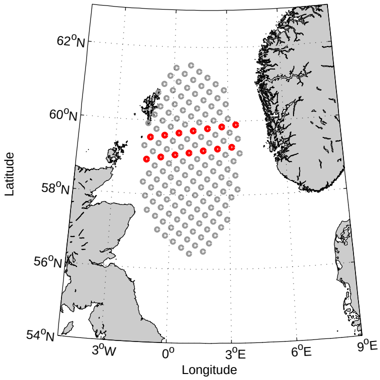
- Marginal extremes
  - Simple introductory example (directional model)
  - H_S^{sp} with 2D, 3D and 4D covariates
- Conditional extremes
  - Associated values of (e.g.) surge given extreme H_S^{sp}
- Temporal extremes
  - Conditional directional evolution of time-series of H_S
- Spatial extremes
  - Conditional spatial extremes of H_S^{sp}
  - Directional dependence in max-stable process parameters for H_S^{sp}

North Sea example as "connecting theme"; other examples to embellish

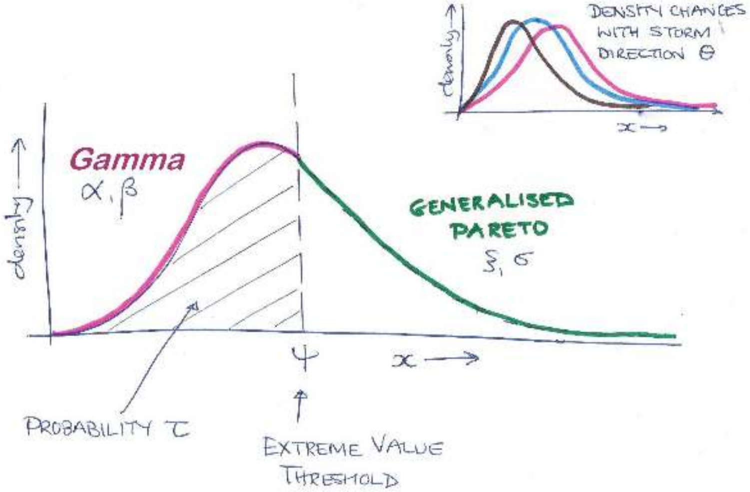
Outline: North Sea application

H_S^{sp} from gridded NEXTRA *winter* storm hindcast for North Sea locations; directional variability in storm severity; transects of locations with different orientations; central location for directional model





Simple gamma-GP model



Simple gamma-GP model

- Sample of peaks over threshold y, with covariates θ
  - θ is 1D in motivating example : directional
  - θ is nD later : e.g. $4D$ spatio-directional-seasonal
- Below threshold ψ
  - y follows truncated gamma with shape α, scale $1/\beta$
  - Hessian for gamma better behaved than Weibull
- Above ψ
  - y follows generalised Pareto with shape ξ, scale σ
- ξ, σ, α, β, ψ all functions of θ
- ψ for pre-specified threshold probability τ
  - Generalise later to estimation of τ

- Frigessi et al. [2002], Behrens et al. [2004], MacDonald et al. [2011]
- Randell et al. [2016]

Simple gamma-GP model

- Density is $f(y|\xi, \sigma, \alpha, \beta, \psi, \tau)$

$$= \begin{cases} \tau \times f_{TG}(y|\alpha, \beta, \psi) & \text{for } y \leq \psi \\ (1-\tau) \times f_{GP}(y|\xi, \sigma, \psi) & \text{for } y > \psi \end{cases}$$

- Likelihood is $\mathcal{L}(\xi, \sigma, \alpha, \beta, \psi, \tau | \{y_i\}_{i=1}^{n})$

$$= \prod_{i:y_i \leq \psi} f_{TG}(y_i|\alpha, \beta, \psi) \prod_{i:y_i > \psi} f_{GP}(y_i|\xi, \sigma, \psi)$$
$$\times \quad \tau^{n_B}(1-\tau)^{(1-n_B)} \text{ where } n_B = \sum_{i:y_i \leq \psi} 1.$$

Estimate all parameters as functions of θ

Rate of occurrence ρ

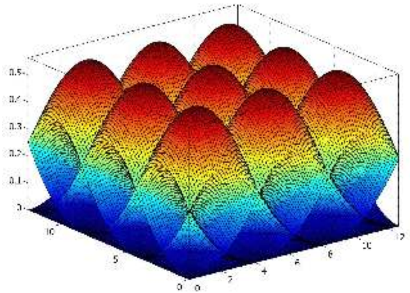
- Whole-sample rate of occurrence ρ modelled as Poisson process given counts c of numbers of occurrences per covariate bin
- Chavez-Demoulin and Davison [2005]

P-splines

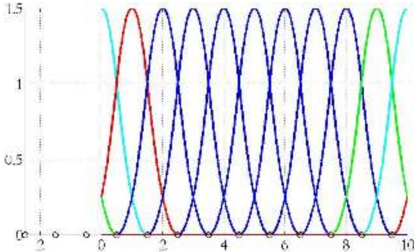
- Physical considerations suggest $\alpha, \beta, \rho, \xi, \sigma, \psi$ and τ vary smoothly with covariates θ
- Values of $\eta \in \{\alpha, \beta, \rho, \xi, \sigma, \psi, \tau\}$ on some index set of covariates take the form $\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\beta}_\eta$
  - For nD covariates, $\boldsymbol{B}$ takes the form of tensor product $\boldsymbol{B}_{\theta_n} \otimes ... \otimes \boldsymbol{B}_{\theta_\kappa} \otimes ... \otimes \boldsymbol{B}_{\theta_2} \otimes \boldsymbol{B}_{\theta_1}$
- Spline roughness with respect to each covariate dimension κ given by quadratic form $\lambda_{\eta\kappa}\beta'_{\eta\kappa}\boldsymbol{P}_{\eta\kappa}\boldsymbol{\beta}_{\eta\kappa}$
- $\boldsymbol{P}_{\eta\kappa}$ is a function of stochastic roughness penalties $\delta_{\eta\kappa}$

- Brezger and Lang [2006]

P-splines



Kronecker product



Periodic P-splines

Gibbs sampling on a page

POSTERIOR, LIKELIHOOD, PRIOR

$$p(\beta \mid y) = \frac{p(y \mid \beta)\; p(\beta)}{p(y)}$$

EVIDENCE

BAYES THEOREM

$$\propto p(y \mid \beta)\, p(\beta) \text{ when data } y \text{ is fixed.}$$

We start by guessing $p(\beta)$, and specifying $p(y \mid \beta)$. Then we can "learn" what β is when we've observed y.

$$p(\beta_1, \beta_2 \mid y) \propto p(y \mid \beta_1, \beta_2)\, p(\beta_1, \beta_2)$$

ITERATE

$$p(\beta_1 \mid y, \beta_2) \propto p(y \mid \beta_1, \beta_2)\, p(\beta_2)$$

$$p(\beta_2 \mid y, \beta_1) \propto p(y \mid \beta_1, \beta_2)\, p(\beta_1)$$

GIBBS SAMPLING

GIBBS SAMPLING allows us to learn about LOTS OF βs in a computationally efficient way.

Priors and conditional structure

Priors

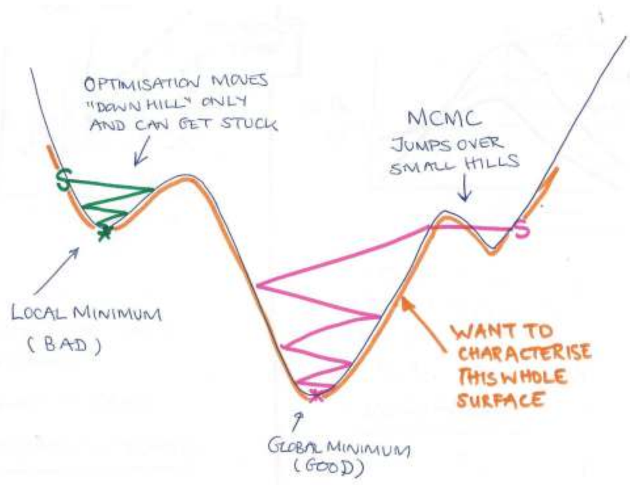
$$\begin{aligned} \text{density of } \boldsymbol{\beta}_{\eta\kappa} &\propto \exp\left(-\frac{1}{2}\lambda_{\eta\kappa}\boldsymbol{\beta}'_{\eta\kappa}\boldsymbol{P}_{\eta\kappa}\boldsymbol{\beta}_{\eta\kappa}\right) \\ \lambda_{\eta\kappa} &\sim \text{gamma} \\ (\text{ and } \tau &\sim \text{beta, when } \tau \text{ estimated }) \end{aligned}$$

Conditional structure

$$\begin{aligned} f(\tau|\boldsymbol{y}, \Omega \setminus \tau) &\propto f(\boldsymbol{y}|\tau, \Omega \setminus \tau) \times f(\tau) \\ f(\boldsymbol{\beta}_\eta|\boldsymbol{y}, \Omega \setminus \boldsymbol{\beta}_\eta) &\propto f(\boldsymbol{y}|\boldsymbol{\beta}_\eta, \Omega \setminus \boldsymbol{\beta}_\eta) \times f(\boldsymbol{\beta}_\eta|\boldsymbol{\delta}_\eta, \boldsymbol{\lambda}_\eta) \\ f(\boldsymbol{\lambda}_\eta|\boldsymbol{y}, \Omega \setminus \boldsymbol{\lambda}_\eta) &\propto f(\boldsymbol{\beta}_\eta|\boldsymbol{\delta}_\eta, \boldsymbol{\lambda}_\eta) \times f(\boldsymbol{\lambda}_\eta) \end{aligned}$$

$$\eta \in \Omega = \{\alpha, \beta, \rho, \xi, \sigma, \psi, \tau\}$$

Inference


OPTIMISATION MOVES "DOWN HILL" ONLY AND CAN GET STUCK
S
MCMC JUMPS OVER SMALL HILLS
S
LOCAL MINIMUM (BAD)
WANT TO CHARACTERISE THIS WHOLE SURFACE
GLOBAL MINIMUM (GOOD)

Inference

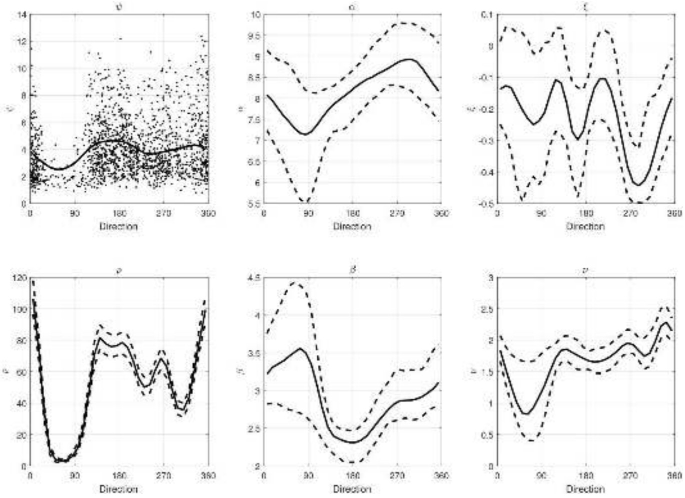
- Elements of $\boldsymbol{\beta}_\eta$ highly interdependent, correlated proposals essential for good mixing
- "Stochastic analogues" of IRLS and back-fitting algorithms for maximum likelihood optimisation used previously
- Estimation of different penalty coefficients for each covariate dimension

- Gibbs sampling when full conditionals available
- Otherwise Metropolis-Hastings (MH) within Gibbs, using suitable proposal mechanisms, mMALA where possible

- Roberts and Stramer [2002], Girolami and Calderhead [2011], Xifara et al. [2014]

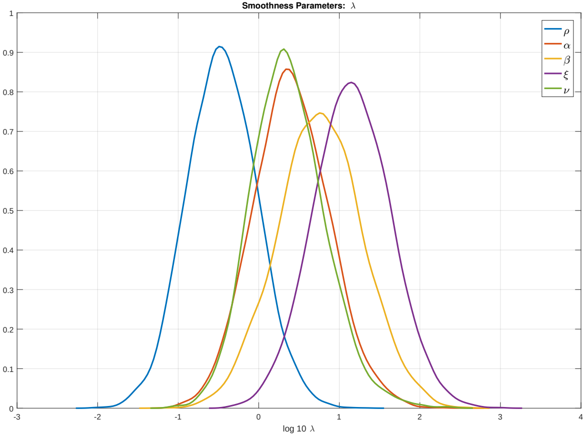
Posterior parameter estimates

Fetch characteristics obvious; land shadow of Norway at 60°



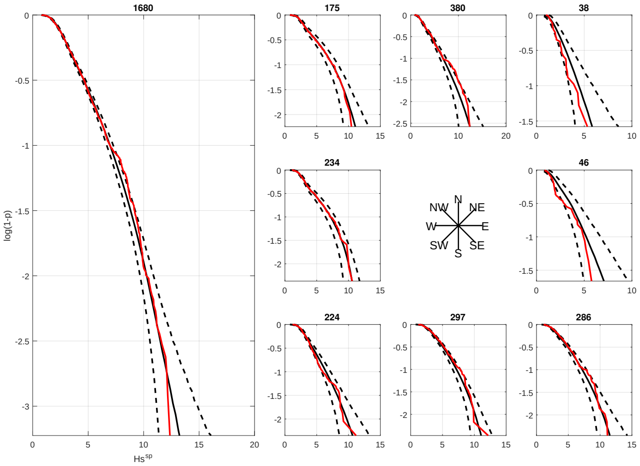
Posterior roughness penalty

Different scales so must be careful : rate is roughest, GP shape is smoothest



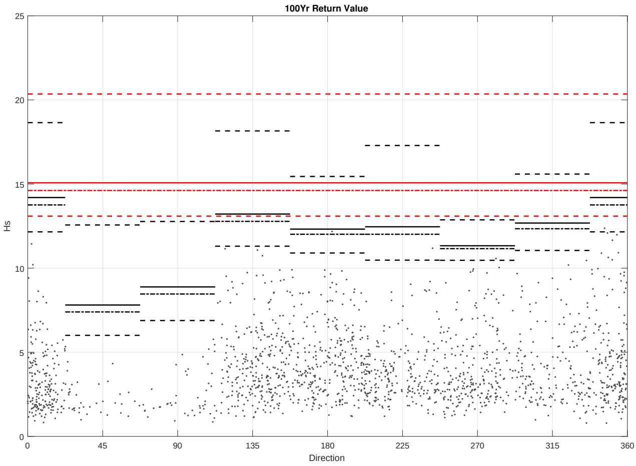
Validation

Compare sample with simulated values on partitioned covariate domain



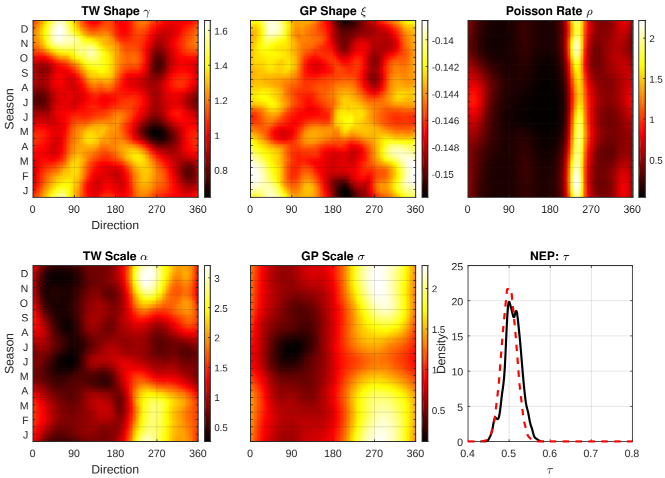
Return values

0.025, $\exp(-1)$, 0.5, 0.975 quantiles: omni (red), directional (black)



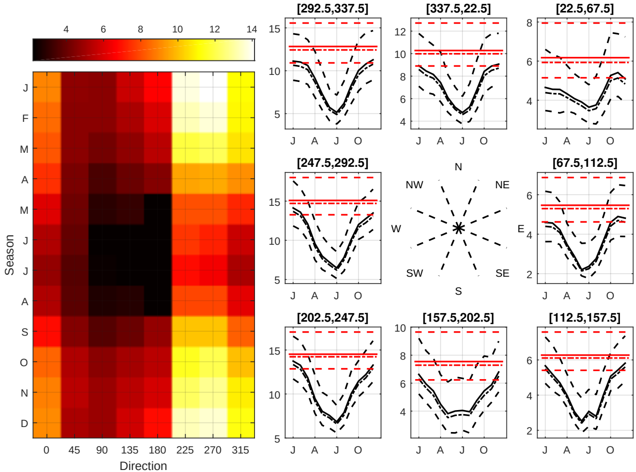
Extension to 2D

Directional-seasonal model; northern North Sea; τ estimated; land-shadow effect of Norway obvious; Randell et al. [2016]



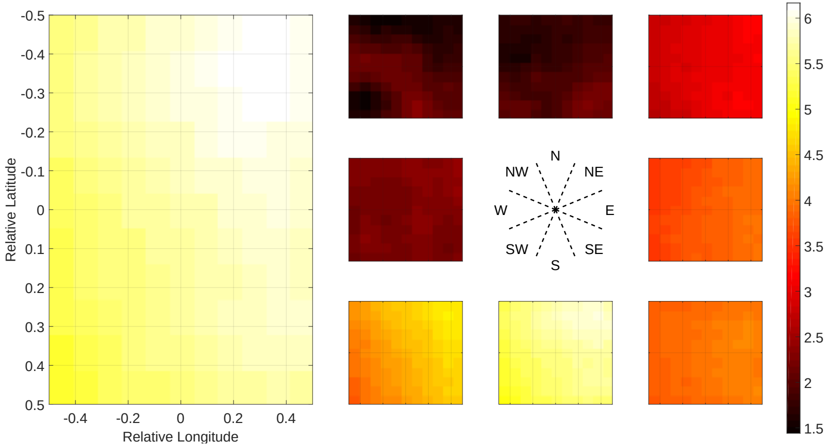
Extension to 2D

Summary statistics for return value distributions; seasonal campaigns can be optimised (offshore maintenance)



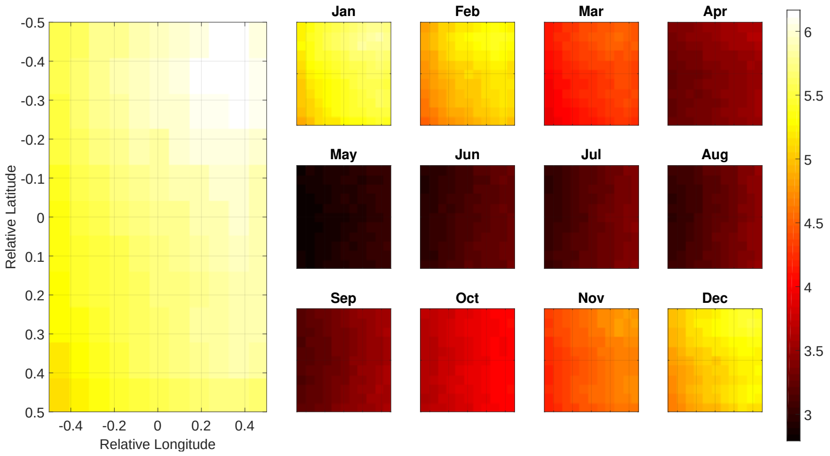
Extension to 4D

Spatio-directional-seasonal model for location in South China Sea; median estimate after integration over season; clear spatial and directional effects; Raghupathi et al. [2016] ML/CV/BS estimation

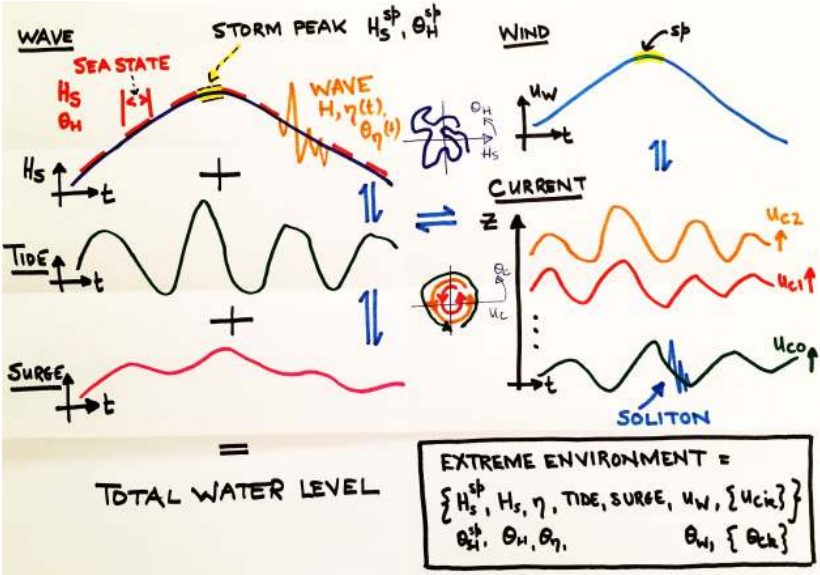


Extension to 4D

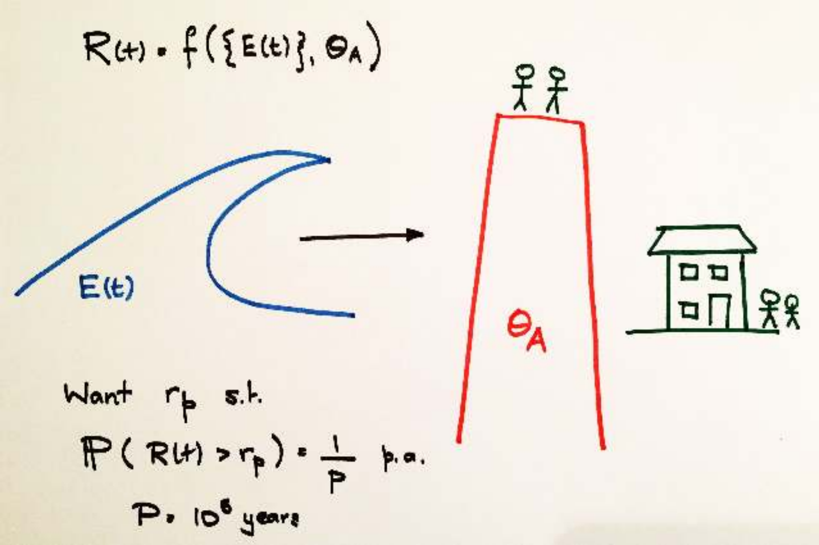
Median estimate after integration over direction; clear spatial and seasonal effects



An extreme environment



An extreme response



Motivating models for extremal dependence

Have (non-stationary) marginal model for dominant variable X_0^{sp} at storm peak. Need models for quantities conditional on X_0^{sp}

Conditional extremes

- Other "associated variables" at storm peak
  e.g. $T_P^{sp} \mid [H_S^{sp} > h, \theta_H^{sp}]$

Markov extremal process

- Evolution of variable around storm peak in time
  e.g $\{H_S(t_j), \theta_H(t_j)\}_j | [H_S^{sp} > h, \theta_H^{sp}]$

Max-stable processes and spatial conditional extremes

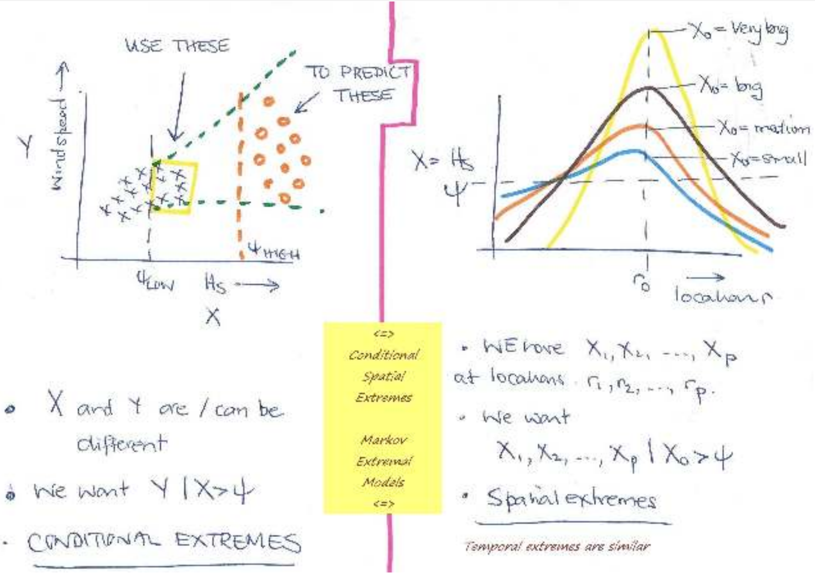
- Dependence of variable in space
  e.g. $\{H_{Sj}^{sp}, \theta_{Hj}^{sp}\}_j | [H_{S0}^{sp} > h, \theta_{H0}^{sp}]$

Hierarchical models for multivariate time-series of waves, crests, surge, tide, total water level, currents, winds. Characterise extreme safety-critical responses

Motivating models for extremal dependence

- Associated peak period: $T_P^{sp} \mid [H_S^{sp} > h, \theta_H^{sp}]$
  Jonathan et al. 2010, 2014
- Currents with depth: $\{u_{Cj}, \theta_{Cj}\}_j \mid [u_{C0} > u, \theta_{C0}]$
  Jonathan et al. 2012
- H_S given wind: $[H_S^{sp}, \theta_H^{sp}] \mid [u_W^{sp} > u, \theta_W^{sp}]$
  Towe et al. 2013
- Storm surge: $S^{sp} \mid [H_S^{sp} > h, \theta_H^{sp}]$
  Ross et al. 2018
- Spatial H_S (max-stable process): $\{H_{Sj}^{sp}\}_j | [H_{S0}^{sp} > x]$
  Ross et al. 2017
- Spatial H_S (conditional extremes): $\{H_{Sj}^{sp}\}_j | [H_{S0}^{sp} > x]$
  Shooter et al. 2018
- Temporal H_S: $\{H_S(t_k), \theta_H(t_k)\}_j | [H_S^{sp} > h, \theta_H^{sp}]$
  Tendijck et al. 2018

Conditional, spatial and temporal extremes



Simple (non-stationary) conditional extremes model

On standard **Laplace** scale, extend with covariates θ

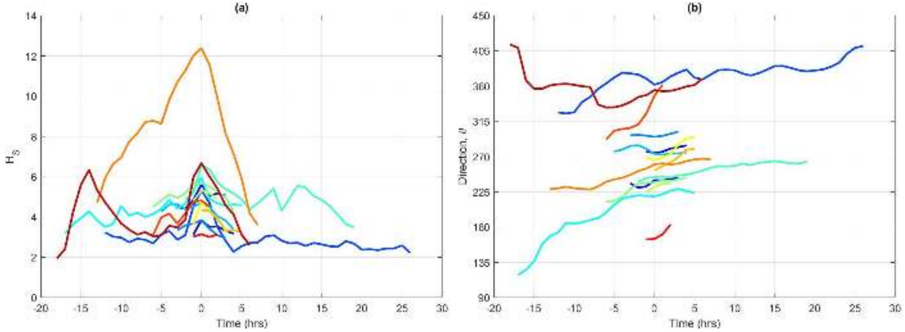
$$(X_2|X_1 = x, \theta) = \alpha x + x^{\beta}(\mu + \sigma Z) \text{ for } x > \psi_\tau$$

- ψ_τ is a high quantile of X_1, for non-exceedance probability τ, above which the model fits well
- $\alpha \in [-1, 1]$, $\beta \in (-\infty, 1]$, $\sigma \in [0, \infty)$
- Z is a random variable with **unknown** distribution G, assumed standard Gaussian for estimation
- $\eta \in \{\alpha, \beta, \mu, \sigma, \psi_\tau\}$ all functions of θ, written as $\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\beta}_\eta$ on index set of covariate values, for suitable covariate basis $\boldsymbol{B}$

- **Heffernan and Tawn [2004]** and derivatives
- Jonathan et al. [2013] for covariates

Motivating time-series extremes

Model for storm trajectories $\{X_t\}_{t \in I} | X_0 = x$ for $x > \psi_\tau$. Time evolution for the 15 typical storms (a) H_S in time, (b) θ in time.
Note change of notation: X_t is value of X at some location at time t



Evolution of X_t

For a "post-peak" portion $\{X_t\}_{t>0}$ of time-series following storm peak X_0, with covariate $\{\Theta_t\}_{t>0}$

On standard **Laplace** scale, for $x > \psi_\tau$

$$[X_{t+1}, X_{t+2}] \,|\, \{X_t = x\} = [\alpha_1, \alpha_2]\, x + x^{[\beta_1, \beta_2]} \left[\mu_1 + \sigma_1 Z_1, \mu_2 + \sigma_2 Z_2\right]$$

- High threshold ψ_τ with non-exceedance probability τ
- Parameters $\alpha_j \in [-1, 1]$, $\beta_j \in (-\infty, 1]$, $\sigma_j \in (0, \infty)$, $j = 1, 2$
- $[Z_1, Z_2]$ are dependent random variables, independent of X_t, with unknown joint distribution function $G_{1:2}$, assumed Gaussian for fitting, then estimated using KDE
- $\{\alpha_j\}$, $\{\beta_j\}$, $\{\mu_j\}$ and $\{\sigma_j\}$ are taken to be constant

- Winter and Tawn [2016, 2017]

Evolution of Θ_t

Given the directions Θ_t at time t relative to storm peak at $t = 0$, we model the rate of change of direction $\Delta_t = \dot{\Theta}_t$

Non-stationary AR(k) form is

$$(\Delta_t | X_t = x) \sim N\left(\sum_{j=1}^{k} \phi_j \Delta_{t-j}, \sigma^2(x)\right)$$

with auto-regressive parameters $\{\phi_j\}$, and variance $\sigma^2(x)$ where

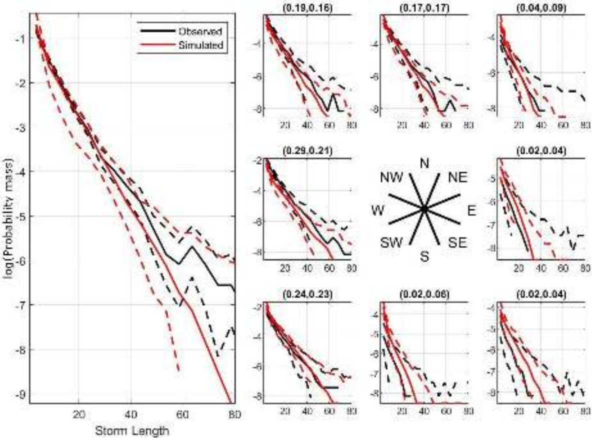
$$\sigma^2(x) = \lambda_1 \exp(-\lambda_2 x) + \lambda_3$$

and $\lambda_1, \lambda_2, \lambda_3 > 0$.

- Tendijck et al. [2018]

Illustrative validation: storm length

Directional comparison of logarithm of probability mass for storm lengths. The left hand panel shows the omni-directional comparison, and the smaller plots show comparisons for 8 directional octants centred on cardinal and inter-cardinal directions. Each panel shows original sample tail (black) and simulated tail (red) with 95% bootstrap uncertainty bands. Titles of smaller panels give the fraction of storm peak occurrences per directional octant, first from original sample and then from simulation



Conditional spatial extremes

Gaussian process representation for a pair of remote locations conditional on a reference location. Extendible to arbitrary number of locations

On Laplace scale

$$[X_{cj}, X_{cj'}] | \{X_{c0} = x\} \sim \text{MVN}\left(\mathcal{M}_{cjj'}, \mathcal{C}_{cjj'}\right), \quad x > \psi_\tau$$

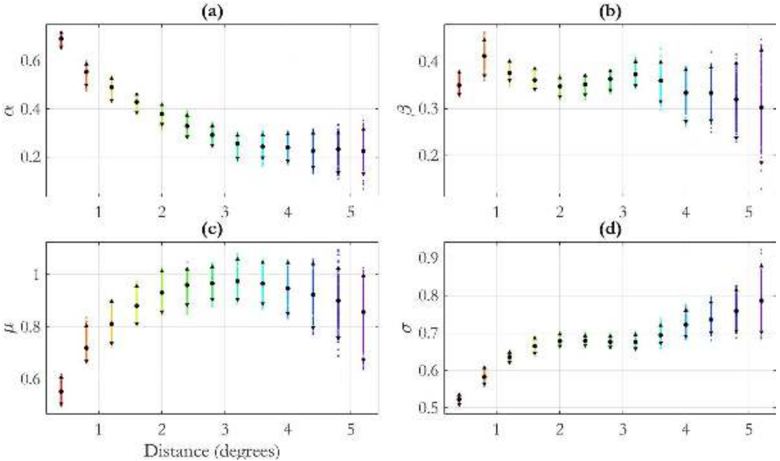
$$\mathcal{M}_{cjj'} = [\alpha(h_{c0j}), \alpha(h_{c0j'})] x_{c0} + [\mu(h_{c0j}), \mu(h_{c0j'})] x_{c0}^{[\beta(h_{c0j}), \beta(h_{c0j'})]}$$

$$\begin{aligned}\mathcal{C}_{cjj'} &= \begin{bmatrix} X_{c0}^{\beta(h_{c0j})} & 0 \\ 0 & X_{c0}^{\beta(h_{c0j'})} \end{bmatrix} \begin{bmatrix} \sigma(h_{c0j}) & 0 \\ 0 & \sigma(h_{c0j'}) \end{bmatrix} \begin{bmatrix} 1 & \rho^{h_{cjj'}} \\ \rho^{h_{cjj'}} & 1 \end{bmatrix} \\ &\times \begin{bmatrix} \sigma(h_{c0j}) & 0 \\ 0 & \sigma(h_{c0j'}) \end{bmatrix}^T \begin{bmatrix} X_{c0}^{\beta(h_{c0j})} & 0 \\ 0 & X_{c0}^{\beta(h_{c0j'})} \end{bmatrix}^T\end{aligned}$$

- Parameter set $\{\alpha_k\}$, $\{\beta_k\}$, $\{\mu_k\}$, $\{\sigma_k\}$, ρ with “gap” index k
- ρ is residual “gap” correlation parameter
- Wadsworth and Tawn [2018], Shooter et al. [2018]

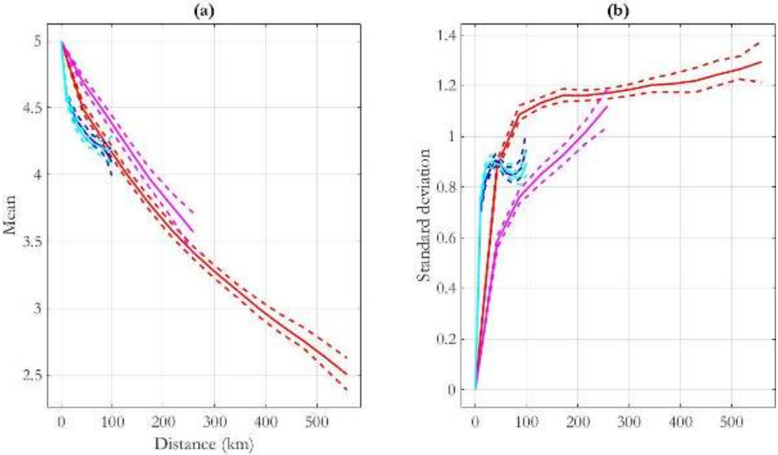
Parameter estimates

NNS:N-S transect, **free model**: (a) α, (b) β, (c) μ and (d) σ with distance h; posterior means (disk) and 95% credible intervals (solid triangles). $\rho \approx$ Gaussian, mean 0.73, 95% interval $(0.68, 0.77)$. **Suggests parametric possible**



Conditional profiles

Credible intervals for (a) conditional mean and (b) conditional standard deviation of fitted dependence model with distance for conditioning Laplace-scale value of 5. NNS:N-W (red), **NNS:E-W (magenta)**, CNS:N-S (blue), CNS:E-W (cyan). **c.f. MSP**



Max-stable processes

- **Max-stable process** (MSP) : a means of extending the GEV for modelling maxima at one location, to multivariate extreme value distributions for modelling of component-wise maxima observed on a lattice
- On unit Fréchet scale, only choices of F_Z exhibiting homogeneity are valid for spatial extreme value modelling
- **Exponent measure** V_Z

$$F_Z(z_1, z_2, ..., z_p) = \exp\{-V_Z(z_1, z_2, ..., z_p)\}$$

- **Extremal coefficient** θ_p

$$\begin{aligned} F_Z(z, z, ..., z) &= \exp\left(-V_Z(z, z, ..., z)\right) \\ &= \exp\left(-z^{-1} V_Z(1, 1, ..., 1)\right) \text{ for homogeneity} \\ &= \exp\left(-\theta_p / z\right) \end{aligned}$$

Exponent measures

- **Smith** : For two locations s_k, s_l in $\mathcal{S}$, V_{kl} for Smith process given by

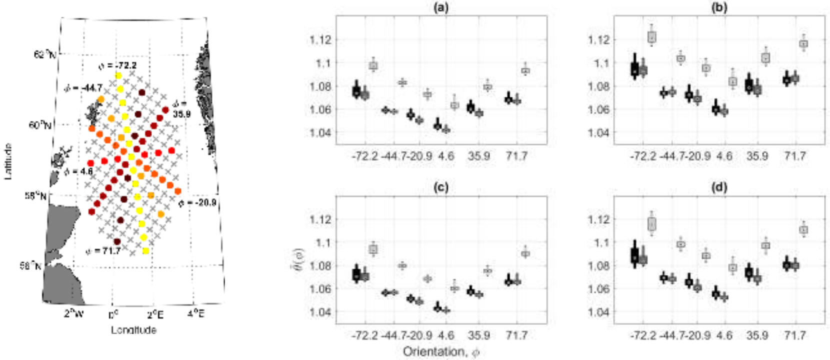
$$V_{kl}(z_k, z_l; h(\Sigma)) = \frac{1}{z_k}\Phi(\frac{m(h)}{2} + \frac{\log(z_l/z_k)}{m(h)}) + \frac{1}{z_l}\Phi(\frac{m(h)}{2} + \frac{\log(z_k/z_l)}{m(h)})$$

- $h = s_l - s_k$, $m(h)$ is Mahalanobis distance $(h'\Sigma^{-1}h)^{1/2}$ between s_k and s_l
- Σ **is** 2×2 **covariance matrix (2-D space) to be estimated**
- $V_{kl}(1, 1; h(\Sigma)) = 2\Phi(m(h)/2)$ by construction

- **Schlather** : similar likelihood, parameterised in terms of Σ only
- **Brown-Resnick** : identical likelihood, parameterised in terms of Σ and scalar Hurst parameter H (estimated up front)

Spatial : extremal coefficient $\hat{\theta}(\phi)$

Estimated extremal coefficient $\hat{\theta}(\phi)$ for all transects with a given orientation ϕ, estimated using 1-D Smith (black), Schlather (dark grey) and Brown-Resnick (light grey) processes. The first (second) row corresponds = marginal threshold with non-exceedance probability 0.5 (0.8). The first (second) column = censoring threshold with non-exceedance probability 0.5 (0.8)



$$F_{Z_1,Z_2}(z,z) = \exp[-\theta/z], \text{ for } \theta \in [1,2]$$

Summary

Today

- Covariate effects in marginal, conditional, spatial and temporal extremes of ocean storms

Also doing

- Bayesian uncertainty analysis (emulation and discrepancy)
- Alternative representations for covariate effects (e.g. tessellations)

Next

- More conditional spatial and (multivariate?) Markov extremal models
- “Measured” data (satellite altimeter, asymptotic independence?)
- Conditional profiles of extreme individual waves

Eventually

- Efficient whole-basin inference with $\approx$ 4D covariates

References

C N Behrens, H F Lopes, and D Gamerman. Bayesian analysis of extreme events with threshold estimation. *Stat. Modelling*, 4:227–244, 2004.

A. Brezger and S. Lang. Generalized structured additive regression based on Bayesian P-splines. *Comput. Statist. Data Anal.*, 50:967–991, 2006.

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.

I. D. Currie, M. Durban, and P. H. C. Eilers. Generalized linear array models with applications to multidimensional smoothing. *J. Roy. Statist. Soc. B*, 68:259–280, 2006.

P. H. C. Eilers and B. D. Marx. Splines, knots and penalties. *Wiley Interscience Reviews: Computational Statistics*, 2:637–653, 2010.

A. Frigessi, O. Haug, and H. Rue. A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes*, 5:219–235, 2002.

M. Girolami and B. Calderhead. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. Roy. Statist. Soc. B*, 73:123–214, 2011.

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

P. Jonathan, J. Flynn, and K. C. Ewans. Joint modelling of wave spectral parameters for extreme sea states. *Ocean Eng.*, 37:1070–1080, 2010.

P. Jonathan, K. C. Ewans, and J. Flynn. Joint modelling of vertical profiles of large ocean currents. *Ocean Eng.*, 42:195–204, 2012.

P. Jonathan, K. C. Ewans, and D. Randell. Joint modelling of environmental parameters for extreme sea states incorporating covariate effects. *Coastal Eng.*, 79:22–31, 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. MacDonald, C. J. Scarrott, D. Lee, B. Darlow, M. Reale, and G. Russell. A flexible extreme value mixture model. *Comput. Statist. Data Anal.*, 55:2137–2157, 2011.

L. Raghupathi, D. Randell, E. Ross, K. Ewans, and P. Jonathan. Multi-dimensional predictive analytics for risk estimation of extreme events. *(Big Data Foundations Workshop, IEEE High-Performance Computing, Data and Analytics Conference (HIPC2016), draft at www.lancs.ac.uk/∼jonathan)*, 2016.

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.

G. O. Roberts and O. Stramer. Langevin diffusions and Metropolis-Hastings algorithms. *Methodology and Computing in Applied Probability*, 4:337–358, 2002.

E Ross, M Kereszturi, M van Nee, D Randell, and P Jonathan. On the spatial dependence of extreme ocean storm seas. *Ocean Eng.*, 145:1–14, 2017.

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

R Shooter, E Ross, J A Tawn, and P Jonathan. On spatial conditional extremes for ocean storm severity. *In preparation for Environmetrics*, 2018.

S. Tendijck, E. Ross, D. Randell, and P. Jonathan. A non-stationary statistical model for the evolution of extreme storm events. *(Accepted by Environmetrics, draft at www.lancs.ac.uk/∼jonathan)*, 2018.

R. Towe, E. Eastoe, J. Tawn, Y. Wu, and P. Jonathan. Omae2013-10154: The extremal dependence of storm severity, wind speed and surface level pressure in the northern north sea. *Proc. 32nd Conf. Offshore Mech. Arct. Eng.*, 2013.

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.

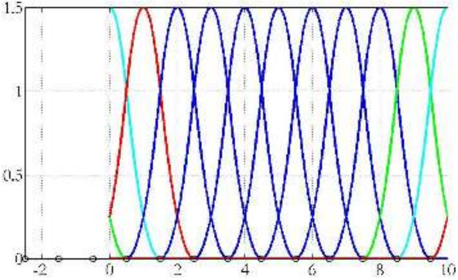
H. C. Winter and J. A. Tawn. kth-order Markov extremal models for assessing heatwave risks. *Extremes*, 20:393–415, 2017.

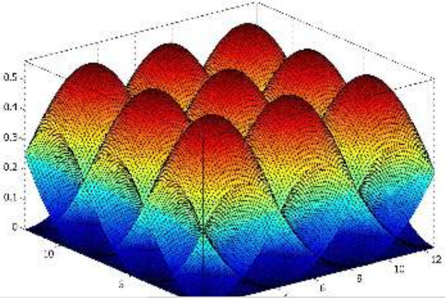
T. Xifara, C. Sherlock, S. Livingstone, S. Byrne, and M Girolami. Langevin diffusions and the Metropolis-adjusted Langevin algorithm. *Stat. Probabil. Lett.*, 91(2002):14–19, 2014.

Supporting material

Penalised B-splines

- **Wrapped** bases for periodic covariates (direction, season).
- **Multidimensional** bases easily constructed using tensor products, Eilers and Marx [2010].
- **GLAMs**, Currie et al. [2006] for efficient computation in high dimensions.





Gradient-based MCMC

- **HMC**: Hamiltonian Monte Carlo: uses first derivatives of parameters have momentum based on gradient. This approach can be unstable so several leapfrog steps are taken instead of single step.
- **Riemann manifold HMC**: uses second derivatives of parameters. Here 2 leapfrog steps are needs so this is computationally challenging
- **MALA** Metropolis adjusted Langevin algorithm: uses first derivatives steps. Proposal $\boldsymbol{\alpha}^* \sim N(\mu, \Sigma)$ where

$$\mu = \boldsymbol{\alpha} - \frac{\epsilon}{2}\frac{\partial}{\partial \boldsymbol{\alpha}}(L + L_{prior})$$
$$\Sigma = \epsilon I$$

and then implement standard MH based on this proposal.

mMALA

- Given a current state $\boldsymbol{\alpha}$ a proposal $\boldsymbol{\alpha}^*$ is sampled from $N(\mu(\boldsymbol{\alpha}), \Sigma)$, where

$$\mu(\boldsymbol{\alpha}) = \boldsymbol{\alpha} - \frac{\epsilon}{2} G^{-1}(\boldsymbol{\alpha}) \frac{\partial}{\partial \boldsymbol{\alpha}} (L + L_{prior})$$
$$\Sigma = \epsilon G^{-1}(\boldsymbol{\alpha})$$

  and then MH is carried through as before. As in MALA we again do not have symmetric proposals and so we must calculate the full acceptance probability.
- it is also interesting to notice the similarities between IWLS and mMALA. To see this compare

$$G(\boldsymbol{\alpha}_\xi)^{-1} = (B' \frac{\partial^2 L}{\partial \boldsymbol{\xi}^2} B + \lambda_\xi P)^{-1}$$
$$\hat{\alpha}_{t+1} = (B' \hat{W}_t B + \lambda D' D)^{-1} B' \hat{W}_t \hat{z}_t$$

Simple (non-stationary) conditional extremes model

On standard **Laplace** scale, extend with covariates θ

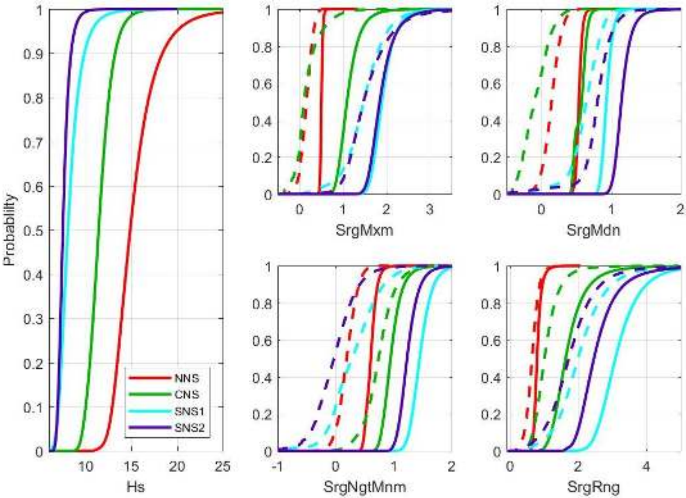
$$(X_2|X_1 = x, \theta) = \alpha x + x^{\beta}(\mu + \sigma Z) \text{ for } x > \psi_\tau$$

- ψ_τ is a high quantile of X_1, for non-exceedance probability τ, above which the model fits well
- $\alpha \in [-1, 1]$, $\beta \in (-\infty, 1]$, $\sigma \in [0, \infty)$
- Z is a random variable with **unknown** distribution G, assumed standard Gaussian for estimation
- $\eta \in \{\alpha, \beta, \mu, \sigma, \psi_\tau\}$ all functions of θ, written as $\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\beta}_\eta$ on index set of covariate values, for suitable covariate basis $\boldsymbol{B}$

- **Heffernan and Tawn [2004]** and derivatives
- Jonathan et al. [2013] for covariates

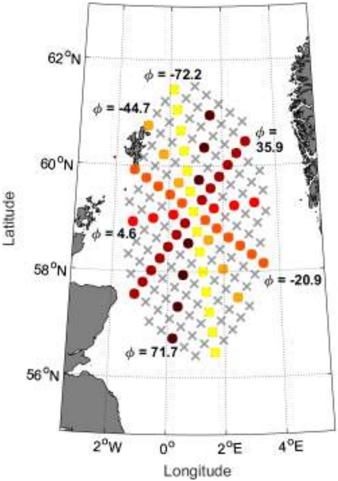
Example: Surge $|H_S^{sp}$

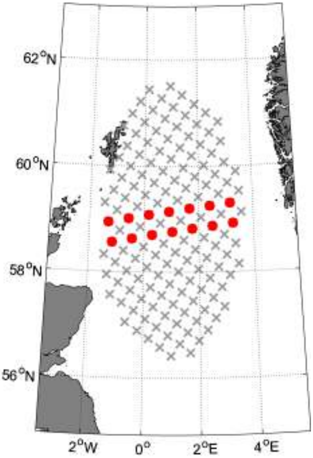
100-year H_S^{sp} together with marginal and conditional surge characteristics. SrgMxm: no associated surge for NNS and CNS



Spatial extremes

Storm peak H_S from gridded NEXTRA *winter* storm hindcast for North Sea locations; directional variability in storm severity; “strips” of locations with different orientations; central location for directional model





Motivation

- Improved inference for the characteristics of extremes at one location exploiting data from multiple locations in a spatial neighbourhood
- Improved estimation of risk for spatially-distributed structures (coastal defences, multiple installations) from spatially spread storm events

- Can we estimate spatial extremes models **usefully** from typical metocean hindcast data?
- Can we see evidence for **covariate effects** in extremal spatial dependence for ocean storm severity?

Spatial dependence

- Locations $j = 1, 2, ..., p$, continuous random variables $\{X_j\}$
- e.g. spatial distribution of H_S^{sp}

$$f(x_1, x_2, ..., x_p) = \left[f(x_1)f(x_2)...f(x_p)\right] \mathcal{C}(x_1, x_2, ..., x_p)$$

- $\{f(x_j)\}$ are marginal densities, $\mathcal{C}(x_1, x_2, ..., x_p)$ is dependence "copula"
- Interested in "the shape of an extreme storm"

$$f(x_1, x_2, ..., x_p | X_k = x_k > u_k) \text{ for large } u_k$$

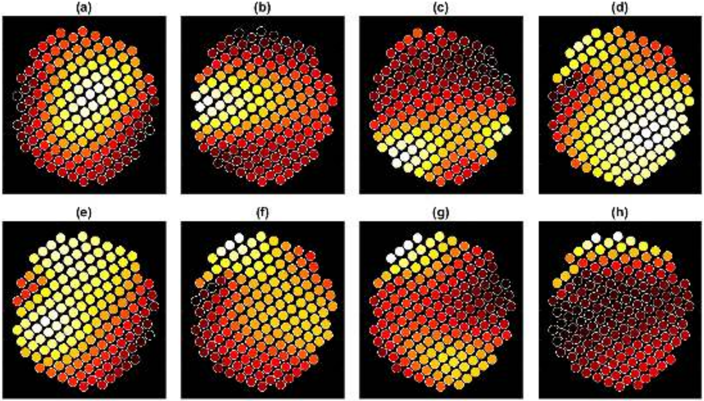
- We know how to estimate extremes marginally, but what about extremal dependence?
- $\Rightarrow$ **Sensible models for** $\mathcal{C}(x_1, x_2, ..., x_p)$

Inference procedures

- Sample of peaks $\{X_j\}$ from p locations, with covariates $\{\theta\}$
- Simple marginal gamma-GP model
- Sample transformed ("whitened") to standard Laplace or Fréchet scale per location
- Inference
    - Conditional spatial extremes
    - Spatial extremes ("max-stable process")
- Bayesian inference estimating joint distributions of parameters, uncertainties
    - Adaptive MCMC (Roberts and Rosenthal 2009) etc.

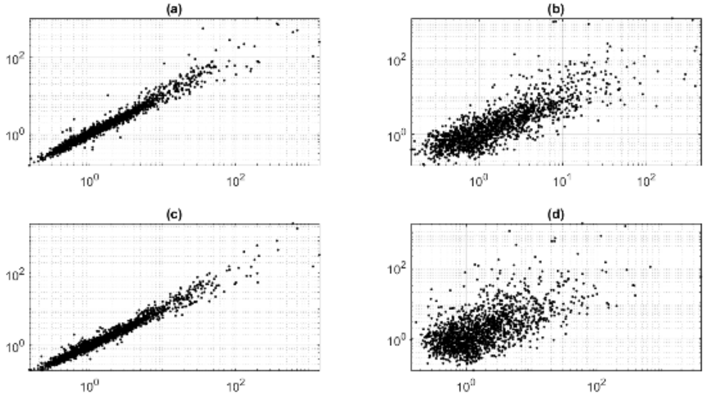
North Sea data

Standard scale observations of the spatial distribution of H_S^{sp} over the North Sea spatial grid for 8 typical events (a)-(h). The spatial maximum for each event is given as a white disc, and the spatial minimum as a black disc (with white outline). The white $\rightarrow$ yellow $\rightarrow$ red $\rightarrow$ black colour scheme indicates the spatial variation of relative magnitude of storm peak H_S^{sp}



North Sea data

Fréchet scale scatter plots of H_S^{sp} for different pairs of locations. Panel (a) for the central location and its nearest neighbour to the West along the approximate West-East transect with angle $\phi = 4.6$; panel (b) for the end locations of the same transect. Panel (c) for the central location and its nearest neighbour to the North along the approximate North-South transect with angle $\phi = -72.2$; panel (d) for the end locations of the same transect. **Higher dependence West-East (care with scale)**



Extremes basics : marginal

- Block maxima Y_k at location k have distribution F_{Y_k} which is max-stable in the sense that $F_{Y_k}^n(b'_{kn} + a'_{kn} y_k) = F_{Y_k}(y_k)$ for some sequences $\{a'_{kn} > 0\}$ and $\{b'_{kn}\}$

- Only possible limiting distribution for F_{Y_k} is generalised extreme value (GEV)

$$
\begin{aligned}
F_{Y_k}(y_k) &= \exp[-\exp\{(y_k - \eta)/\tau\}] \text{ for } \xi = 0 \\
&= \exp[-\{1 + \xi(y_k - \eta)/\tau\}_+^{-1/\xi}] \text{ otherwise}
\end{aligned}
$$

- For peaks over threshold, the equivalent asymptotic distribution is the generalised Pareto distribution.

Extremes basics : spatial

- Similarly, F_Y for block maxima Y at p locations "max-stable" when $F_Y^n(b'_{1n} + a'_{1n}y_1, b'_{2n} + a'_{2n}y_2, ..., b'_{pn} + a'_{pn}y_p) = F_Y(y_1, y_2, ..., y_p)$
- Transform to unit Fréchet $Z_k = \{1 + \xi(Y_k - \eta)/\tau\}^{1/\xi}$, $F_{Z_k}(z_k) = \exp(-1/z_k)$, for $z_k > 0$. Then

$$F_Z(z_1, z_2, ..., z_p) = F_Z(nz_1, nz_2, ..., nz_p)^n$$

- Only choices of F_Z exhibiting this homogeneity correspond to finite-dimensional distributions from max-stable processes (MSPs), and are hence valid for spatial extreme value modelling

Spatial : basic theory

- Max-stable process (MSP) : a means of extending the GEV for modelling maxima at one location, to multivariate extreme value distributions for modelling of component-wise maxima observed on a lattice
- On unit Fréchet scale, only choices of F_Z exhibiting homogeneity are valid for spatial extreme value modelling
- Terminology : exponent measure V_Z

$$F_Z(z_1, z_2, ..., z_p) = \exp\{-V_Z(z_1, z_2, ..., z_p)\}$$

- Terminology : extremal coefficient θ_p

$$\begin{aligned} F_Z(z, z, ..., z) &= \exp\left(-V_Z(z, z, ..., z)\right) \\ &= \exp\left(-z^{-1} V_Z(1, 1, ..., 1)\right) \text{ from homogeneity} \\ &= \exp\left(-\theta_p / z\right) \end{aligned}$$

Spatial : V_Z for Smith, Schlather and Brown-Resnick

- **Smith** : For two locations s_k, s_l in $\mathcal{S}$, V_{kl} for Smith process given by

$$V_{kl}(z_k, z_l; h(\Sigma)) = \frac{1}{z_k}\Phi(\frac{m(h)}{2} + \frac{\log(z_l/z_k)}{m(h)}) + \frac{1}{z_l}\Phi(\frac{m(h)}{2} + \frac{\log(z_k/z_l)}{m(h)})$$

- $h = s_l - s_k$, $m(h)$ is Mahalanobis distance $(h'\Sigma^{-1}h)^{1/2}$ between s_k and s_l
- Σ is 2×2 covariance matrix (2-D space) to be estimated. Σ scalar in 1-D
- $V_{kl}(1, 1; h(\Sigma)) = 2\Phi(m(h)/2)$ by construction

- **Schlather** : similar likelihood, parameterised in terms of Σ only
- **Brown-Resnick** : identical likelihood, parameterised in terms of Σ and scalar Hurst parameter H (estimated up front)

Spatial : constructive representation

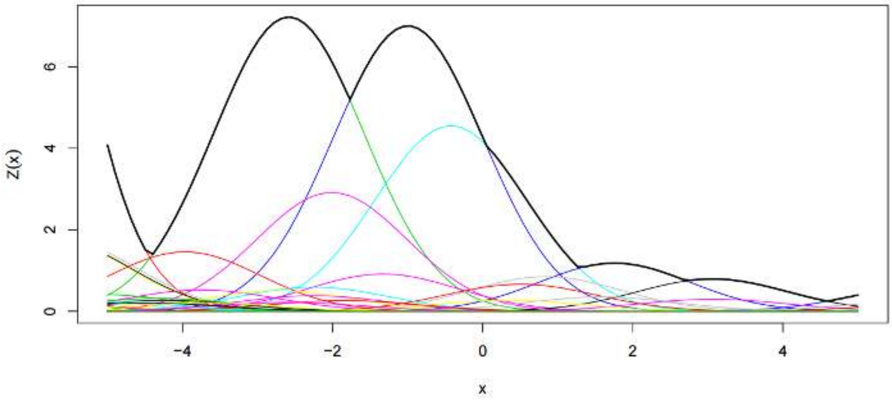
- MSP is maximum of multiple copies $\{W_i\}$ $(i \geq 1)$ of random function W
- Each W_i weighted using Poisson process $\{\rho_i\}$ $(i \geq 1)$

- The MSP $Z(s)$ for s in spatial domain $\mathcal{S}$ is $Z(s) = \mu^{-1} \max_i \{W_i^+(s)/\rho_i\}$

- $W_i^+ = \max\{W_i(s), 0\}$, $\mu = E(W^+(s)) = 1$ by construction typically
- $\rho_i = \epsilon_i$ for $(i = 1)$, $\rho_i = \epsilon_i + \rho_{i-1}$ for $(i > 1)$, and $\epsilon_i \sim \text{Exp}(1)$

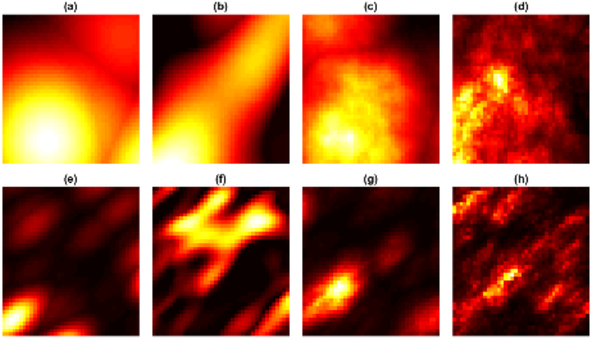
- Different choices of $W(s)$ give different MSPs
- Smith : $W_i(s; s_i, \Sigma) = \varphi(s - s_i; \Sigma)/f_S(s_i)$, with s_i sampled from density $f_S(s_i)$ on $\mathcal{S}$, with φ representing standard Gaussian density
- Schlather, Brown-Resnick : Similar

Spatial : constructive representation



Spatial : illustrations

Illustrative realisations of Smith (a,e), Schlather (b,f), and Brown-Resnick (c,d,g,h) processes for different parameter choices. The first row corresponds to parameter settings $(\Sigma_{11}, \Sigma_{22}, \Sigma_{12}) = (300, 300, 0)$ for all processes, and the second row to (30,20,15). For Brown-Resnick processes (c,g), Hurst parameter $H = 0.95$. For Brown-Resnick processes (d,h), $H = 0.65$. Each panel can be considered to show a possible spatial realisation of storm peak H_S, similar to those shown earlier



Spatial : estimation approximations

- Theory applies for (Fréchet scale) block maxima Z_Y, but we have (Fréchet scale) peaks over threshold Z_X. For $z_k, z_l > u$ for large u, approximate

$$\Pr\left[Z_{Xk} \leq z_k, Z_{Xl} \leq z_l\right] \approx \Pr\left[Z_{Yk} \leq z_k, Z_{Yl} \leq z_l\right]$$

- Theory gives us models for pairs of locations. Cannot write down full joint likelihood $\ell(\Sigma; \{z_j\})$. Approximate with composite likelihood $\ell_C(\Sigma; \{z_j\})$

$$\ell(\Sigma; \{z_j\}) \approx \ell_C(\Sigma; \{z_j\}) = \sum_{\{k,l\} \in \mathcal{N}} w_{kl} \log f_{kl}(z_k, z_l; h(\Sigma))$$

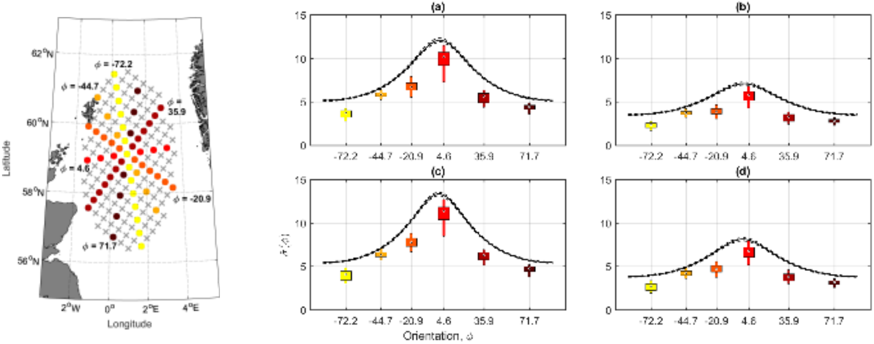
- Need $f_{kl}(z_k, z_l; h(\Sigma))$ for non-exceedances of u also, so make censored likelihood approximation

Spatial : estimation

- Estimate joint distribution of $\Omega = [\Sigma_{11}, \Sigma_{22}, \Sigma_{12}]$ (2-D space, or $\Omega = \Sigma$ in 1-D)
- MCMC using Metropolis-Hastings
  - Current state Ω_{r-1}, marginal posterior $f_M(\beta_M)$, original sample D of storm peak H_S.
  - Draw a set of marginal parameters β_{Mr} from f_M, independently per location.
  - Use β_{Mr} to transform D to standard Fréchet scale, independently per location, obtaining sample D_{Fr}.
  - Execute "adaptive" MCMC step from state Σ_{r-1} with sample D_{Fr} as input, obtain Σ_r.
- Adaptive MCMC candidates generated using $\Omega_r^c = \Omega_{r-1} + \gamma\epsilon_1 + (1-\gamma)\epsilon_2$
  - $\gamma \in [0,1]$, $\epsilon_1 \sim N(0, \delta_1^2 I_3/3)$, $\epsilon_2 \sim N(0, \delta_2^2 S_{\Omega_{r-1}}/3)$
  - $S_{\Omega_{r-1}}$ estimate of variance of Ω_{r-1} using samples to trajectory to date
  - Roberts and Rosenthal [2009]

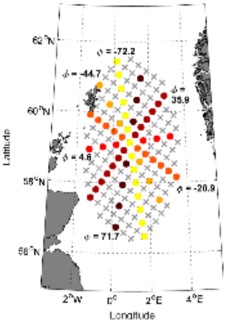
Spatial : $\hat{\sigma}(\phi)$ for Smith

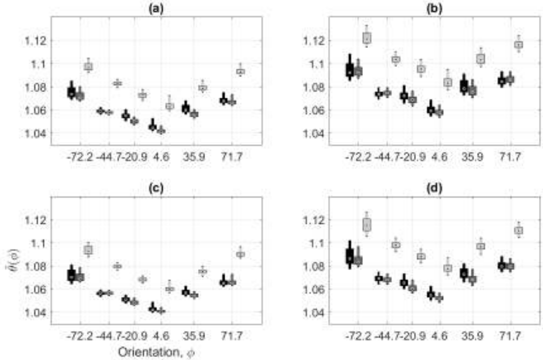
For all transects with a given orientation ϕ estimated using 1-D (box-whisker) and 2-D (black) Smith processes. ϕ is quantified as the transect angle anticlockwise from a line of constant latitude. The first (second) row: marginal threshold non-exceedance probability 0.5 (0.8). The first (second) column: censoring threshold non-exceedance probability 0.5 (0.8). For 1-D estimates with a given ϕ, box centres = median, box edges = 0.25 and 0.75 quantiles across all parallel transects; whisker edges = 0.025 and 0.975 quantiles. For 2-D estimates, the 0.025, 0.5 and 0.975 quantiles are shown as a function of ϕ. Note that the colour coding of box-whisker plots corresponds to that of transect orientation



Spatial : extremal coefficient $\hat{\theta}(\phi)$

Estimated extremal coefficient $\hat{\theta}(\phi)$ for all transects with a given orientation ϕ, estimated using 1-D Smith (black), Schlather (dark grey) and Brown-Resnick (light grey) processes. The first (second) row corresponds = marginal threshold with non-exceedance probability 0.5 (0.8). The first (second) column = censoring threshold with non-exceedance probability 0.5 (0.8)





Spatial : spatial dependence parameter $\hat{\sigma}(\phi, s)$ for individual transects

Smith process with marginal and censoring thresholds = non-exceedance probability of 0.8. (b)-(g): $\hat{\sigma}(\phi, s)$ for fixed orientation ϕ (given in the panel title) as a function of transect locator s. (a): transects with $s = 1$ for different orientations ϕ. (b)-(g): abscissa values for transect locators are scaled to physical perpendicular distances between parallel transects