Spatial conditional extremes via the Gibbs sampler

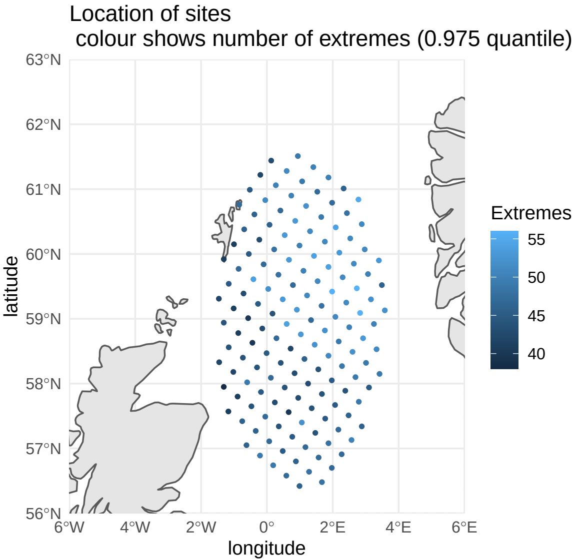
Adrian Casey
School of Mathematics
University of Edinburgh

April 28, 2020

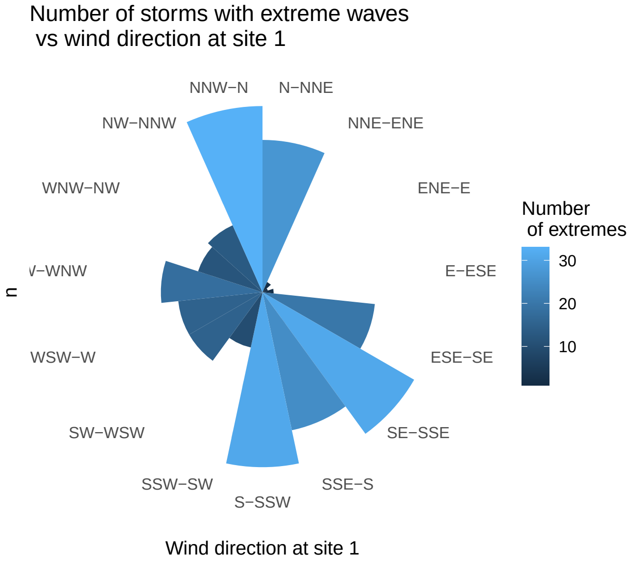
Data

- The wave height data consists of 1680 hindcast observations of significant wave height during storms measured at 150 locations.
- Sites arranged on an approximate grid in the North Sea, with axes running NE-SW and SE-NW.
- The data includes the wind direction at each site for each storm. Extremes are dependent upon this variable.

Data


Location of sites
colour shows number of extremes (0.975 quantile)
63°N
62°N
61°N
60°N
59°N
58°N
57°N
56°N
latitude
6°W
4°W
2°W
0°
2°E
4°E
6°E
longitude
Extremes
55
50
45
40

Wind direction



Heffernan & Tawn

Suppose that $\boldsymbol{X}$ is a random vector with exponential marginal distributions. We condition on the random variable X_k being above a high threshold u. Let $\boldsymbol{X}_{-k}$ be the other components of the vector. The conditional model rests on the relatively weak assumption that, given $X_i > u$ there exist vector functions $\boldsymbol{a}^{\mathbf{k}} : \mathbb{R} \rightarrow \mathbb{R}^{d-1}$ and $\boldsymbol{b}^{\mathbf{k}} : \mathbb{R} \rightarrow \mathbb{R}^{d-1}$ such that,

$$\Pr\left\{X_k - u > y, \frac{\mathbf{X}_{-k} - \mathbf{a}^{\mathbf{k}}(X_k)}{\mathbf{b}^{\mathbf{k}}(X_k)} \leq \mathbf{z} | X_k > u\right\} \rightarrow exp(-y)G^{\mathbf{k}}(\boldsymbol{z}), \quad u \rightarrow \infty. \quad (1)$$

Spatial extremes

Questions that might be asked:

- How many sites are likely to experience extreme values simultaneously?
- Given an extreme value at one location, what is the distribution of values at other sites ?

Particular issues:

- Asymptotics give no general form for the distribution $G^{\mathbf{k}}(\boldsymbol{z})$. In spatial statistics this distribution will be high-dimensional and so difficult to model outside the Gaussian framework.
- An alternative approach is the graphical extremes methods outlined in the recent paper by Engelke & Hitz. This has the disadvantage of assuming asymptotic dependence, which is generally not the case in spatial statistical problems.

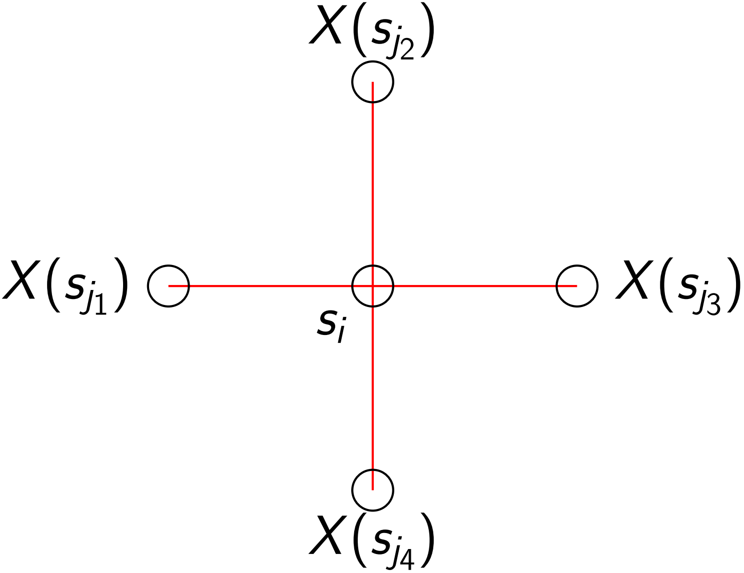
Gaussian Markov random fields

Let $\{s_i\}$ be a set of sites with associated random variables $\{X(s_i)\}$. Let Δ_i be the neighbourhood (the Markov blanket) of $X(s_i)$. For a Gaussian distibution $\mathcal{N}(0, Q^{-1})$ and precision matrix $Q_{ij=}Q(s_i, s_j)$,

$$E[X(s_i)] = -\sum_{s_j \in \Delta_i} \frac{Q_{ij}}{Q_{ii}} X(s_j), \tag{2}$$

$$Var[X(s_i)] = \frac{1}{Q_{ii}}. \tag{3}$$

Δ_i for first order GMRF



A local extreme value model

Assume that, for a range of (positively dependent) MRFs with exponential margins, there exists a scalar function acting on the neighbourhood of site i $a_\Delta(X_{\Delta_i})$ such that ,

$$Z = \frac{X(s_i) - \alpha a_\Delta(X_{\Delta_i})}{(a_\Delta(X_{\Delta_i}))^\beta} \sim G, \tag{4}$$

$$a_\Delta(X_{\Delta_i}) > u \tag{5}$$

where G is a non-degenerate univariate distribution function, with convergence above some high threshold value u.

The function a_Δ

We propose a form for the function a_Δ which is homogeneous in each of the components in the neighbourhood. This function arises in the analysis of asymptotically independent k^{th} order Markov chains.

$$a_\Delta(X_{\Delta_i}) = \left\{ \sum_{j \in \Delta_i} \gamma_j (\gamma_j X(s_j))^\delta \right\}^{1/\delta} \tag{6}$$

$$\sum_{j \in \Delta_i} \gamma_j = 1, \quad \delta > 0 \tag{7}$$

Example:GMRF

Consider the GMRF in Slide 6, drop the i suffix and and set $\beta_j = \frac{Q(s_i,s_j)}{Q(s_i,s_i)}$, $\sigma = \frac{1}{Q(s_i,s_i)}$, then it can be shown that,

$$a_\Delta(X_\Delta) = \left(\sum_{j \in \Delta} \gamma_j (\gamma_j X_j)^{1/2} \right)^2, \tag{8}$$

$$\gamma_j = \frac{\beta_j^{2/3}}{\sum_k \beta_k^{2/3}} \tag{9}$$

$$\beta = 0.5 \tag{10}$$

$$\alpha = \sum_k \beta_k^{2/3} \tag{11}$$

$$Z \sim N(0, \sqrt{2}\sigma) \tag{12}$$

The Gibbs sampler

The Gibbs sampling method is a method of sampling from the full joint distribution by sequential sampling from a series of conditional distributions. So the method follows the following pattern,

Draw x_1^* from $f(X(s_1)|X(s_2) = x_2 \dots X_(s_n) = x_n)$

Draw x_2^* from $f(X(s_2)|X(s_1) = x_1^*, X(s_3) = x_3, \dots X(s_n) = x_n)$

$\vdots$

- This method is often applied to GMRFs because the conditioning set is reduced to the neighbourhood at each point. The existence of this local model suggests a path to sampling from the full joint distribution of $\{X(s_1), \dots X(s_n)\}$, given that the vector at some site, say $X(s_1)$, is extreme.

The Gibbs sampler algorithm

Algorithm 1: A Gibbs sampler for a local extreme model

Data: Regular lattice data

Result: A distribution that allows effective extrapolation to higher quantiles of conditional extremes

Initialization; Set the lattice values to a random sample from the exponential distribution . One site is set to be above some threshold

while *samples* $< N$ **do**

    while $i < n$ **do**

        read current lattice values $\{X(s_1), \ldots X(s_n)\}$;

        if $i = 1$ **then**

            sample from $u + Exp(1)$;

        else if *The neighbourhood function* $a_\Delta(X_{\Delta_i}) > u$ **then**

            Simulate a value from the fitted local extremes model;

            Replace $X(s_i)$ with that value;

        else

            simulate a new value from some model of the bulk density $f(X(s_i) \mid X(s_j) : j \neq i)$;

            Replace $X(s_i)$ with that value;

    After one sweep of the lattice, save as sample;

Working through the lattice

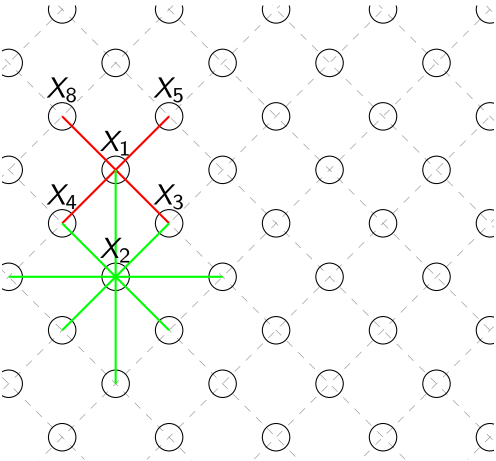


Figure 1: $a_\Delta(X_{\Delta_1}) > u$, so use local model sample$a_\Delta(X_{\Delta_2}) < u$ so use a bulk model to sample X_2, here dependent on 8 neighbours.

Fitting a local extremes model to data

Asuming stationarity, we seek a model for the dependence of a site $X(s_i)$ on the four nearest lattice points.

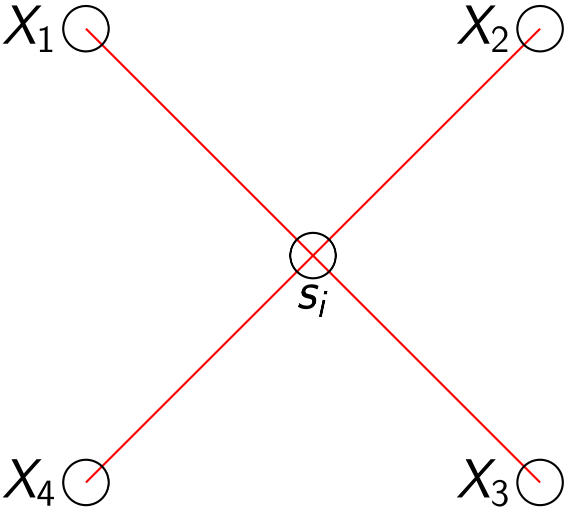
$$X(s_i)|\Delta_i = \alpha a_\Delta(X_{\Delta_i}) + a_\Delta(X_{\Delta_i})^\beta Z, \quad a_\Delta(X_{\Delta_i}) > u, \tag{13}$$

$$a_\Delta(X_{\Delta_i}) = \left\{ \sum_{j \in \Delta_i} \gamma_j (\gamma_j X(s_j))^\delta \right\}^{1/\delta} \tag{14}$$

$$\sum_{j \in \Delta_i} \gamma_j = 1, \quad \delta > 0 \tag{15}$$

$$i = 1 \ldots n. \tag{16}$$

Fitting a local extremes model to data



Fitting a model for the bulk distribution

We need a model for the conditional distribution applicable in the non-extreme region,

$$p(X(s_i)|X(s_j) : j \neq i). \tag{17}$$

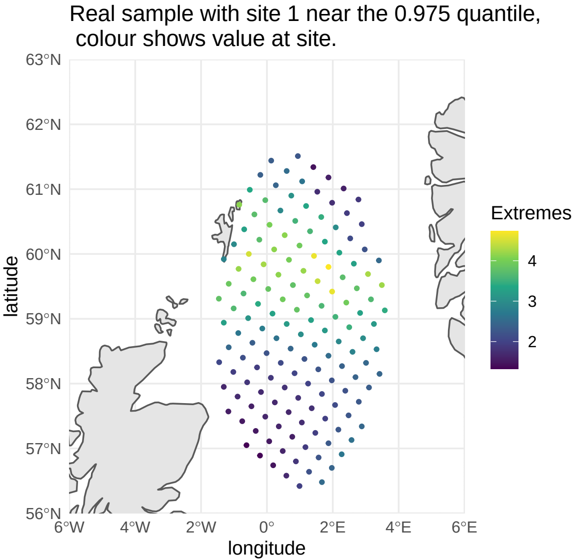
The modelling strategy chosen for this step is to first transform the data to normal scale and then to fit a linear model to each full conditional in turn. In order to avoid over-fitting and to manage the number of parameters in this model, a Lasso is applied.

Parameter stability

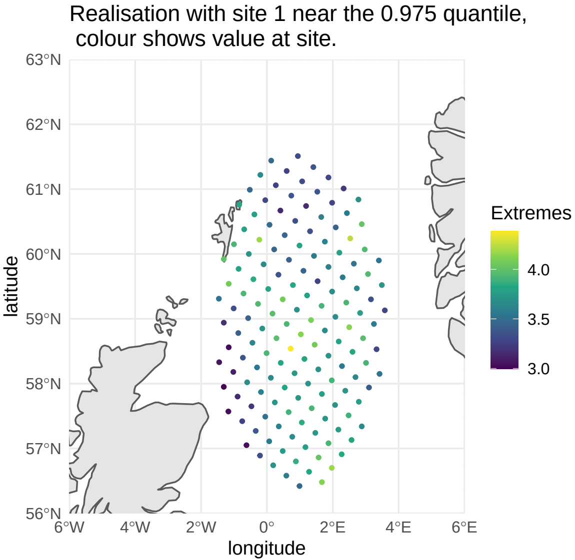
Quantile	γ_1	γ_2	γ_3	γ_4	β	δ	α
0.92	0.27	0.23	0.28	0.22	0.36	0.86	4.00
0.95	0.27	0.23	0.28	0.22	0.16	0.91	3.98
0.975	0.29	0.22	0.28	0.20	0.01	0.75	3.96
0.99	0.25	0.26	0.26	0.24	0.02	1.43	4.12

Table 1: Table showing parameter stability in Θ_1

Sample from data



Realisation from the Gibbs sampler



Expected extreme waves in the same storm

- Simulation based on 3000 Gibbs sampling results with a 1000 sample burn-in.
- Results

Extreme quantile	Data	Simulation
0.95	102.80	108.80
0.975	90.72	92.64
0.99	86.85	88.65
0.995	78.80	68.00
0.999	NA	28.80

Table 2: Expected number of sites with a wave in the extreme quantile, given a similarly extreme wave at site 1.

Conclusions and next steps

- Generally encouraging results for this method of conditional spatial extremes.
- Significant problems remain.
  - Boundary conditions.
  - Non-stationarity and wind dependence.
  - The method relies on some knowledge of the distribution of a_Δ.
  - Analysis of the distribution $\{X(s_1), \ldots, X(s_n)\} \mid \max\{X(s_1), \ldots, X(s_n)\} > u$ is computationally demanding.

References

S. Engelke and A. S. Hitz. Graphical models for extremes. *arXiv preprint arXiv:1812.01734*, 2018.

J. E. Heffernan and J. A. Tawn. A conditional approach for multivariate extreme values (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3):497–546, 2004.

I. Papastathopoulos and J. A. Tawn. Extreme events of higher-order markov chains: hidden tail chains and extremal yule-walker equations. *arXiv preprint arXiv:1903.04059*, 2019.

M. Reistad, Ø. Breivik, H. Haakenstad, O. J. Aarnes, B. R. Furevik, and J.-R. Bidlot. A high-resolution hindcast of wind and waves for the north sea, the norwegian sea, and the barents sea. *Journal of Geophysical Research: Oceans*, 116(C5), 2011.