Spatial Conditional Extremes for Ocean Storm Severity

Rob Shooter[1], Jonathan Tawn[1], Jenny Wadsworth[1], Phil Jonathan[2], Emma Ross[2]

[1] Lancaster University, [2] Shell Research

1. Introduction

- Offshore structures rely on Extreme Value Analysis to aid in the design of their construction.
- Spatial extremes are of particular use in this instance; a key consideration when doing this involves characterising the nature of spatial dependence.
- We are working on a model which allows a flexible dependence structure, using a spatial conditional extremes framework.

2. Conditional Extremes Modelling

- The basis of our spatial model is the conditional extremes approach of Heffernan and Tawn (2004).
- For data (X, Y) with Laplace marginal distributions, the bivariate dependence model for $Y \mid (X > u)$, for some high threshold u, is

$$Y = \alpha X + X^{\beta} Z,$$

where $\alpha \in (-1, 1)$, $\beta \in (0, 1)$ and Z is some residual process - for example this may be a Normal distribution. Fitting this model amounts to non-linear regression.

- For higher-dimensional data, a multivariate extension of this can be obtained by allowing Z to be a multivariate Normal distribution with marginal distributions $N(\mu_j, \gamma_j^2)$.
- Then for data $(X_1, \ldots, X_n)$, the marginal models are as follows:

$$X_j \mid \{X_1 = x\} \sim N\left(\alpha_j + \mu_j x^{\beta_j}, \gamma_j^2 x^{2\beta_j}\right);$$

we use this as the basis of our spatial model.

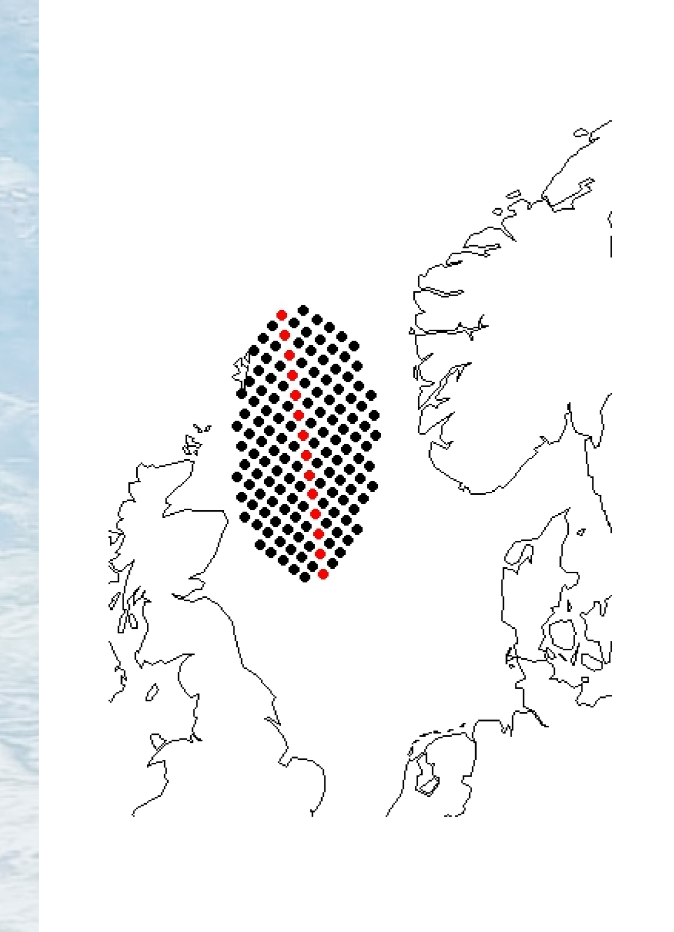
3. Spatial Conditional Extremes

- Suppose that $X(\cdot)$, the process of interest, is stationary and isotropic and has Laplace marginal distributions.
- Also suppose that we have sampling locations $s, s_0 \in \mathcal{S}$, then for $h = |s - s_0|$, the distance/lag between two sites, then we have

$$X(s) \mid \{X(s_o) > u\} = \alpha(h) X(s_0) + X(s_0)^{\beta(h)} W(s - s_0). \quad (1)$$

- We assume W is a Gaussian process with some correlation structure to be estimated, incorporating $\sigma(h)$, and with mean function $\mu(h)$.
- Different combinations of parameter values correspond to different types of spatial dependence:
  1. Asymptotic dependence at all distances h is seen when $\alpha(h) = 1$ and $\beta(h) = 0$ for all $h \geq 0$.
  2. A mixture of dependence types is observed if $(\alpha(h), \beta(h)) = (1, 0)$ for $h \leq h_{AD}$ but also $\alpha(h) < 1$ for $h > h_{AD}$. The process exhibits asymptotic dependence up to distance h_{AD} and asymptotic independence after this point.
- Thus, we have a model able to flexibly model extremal dependence.

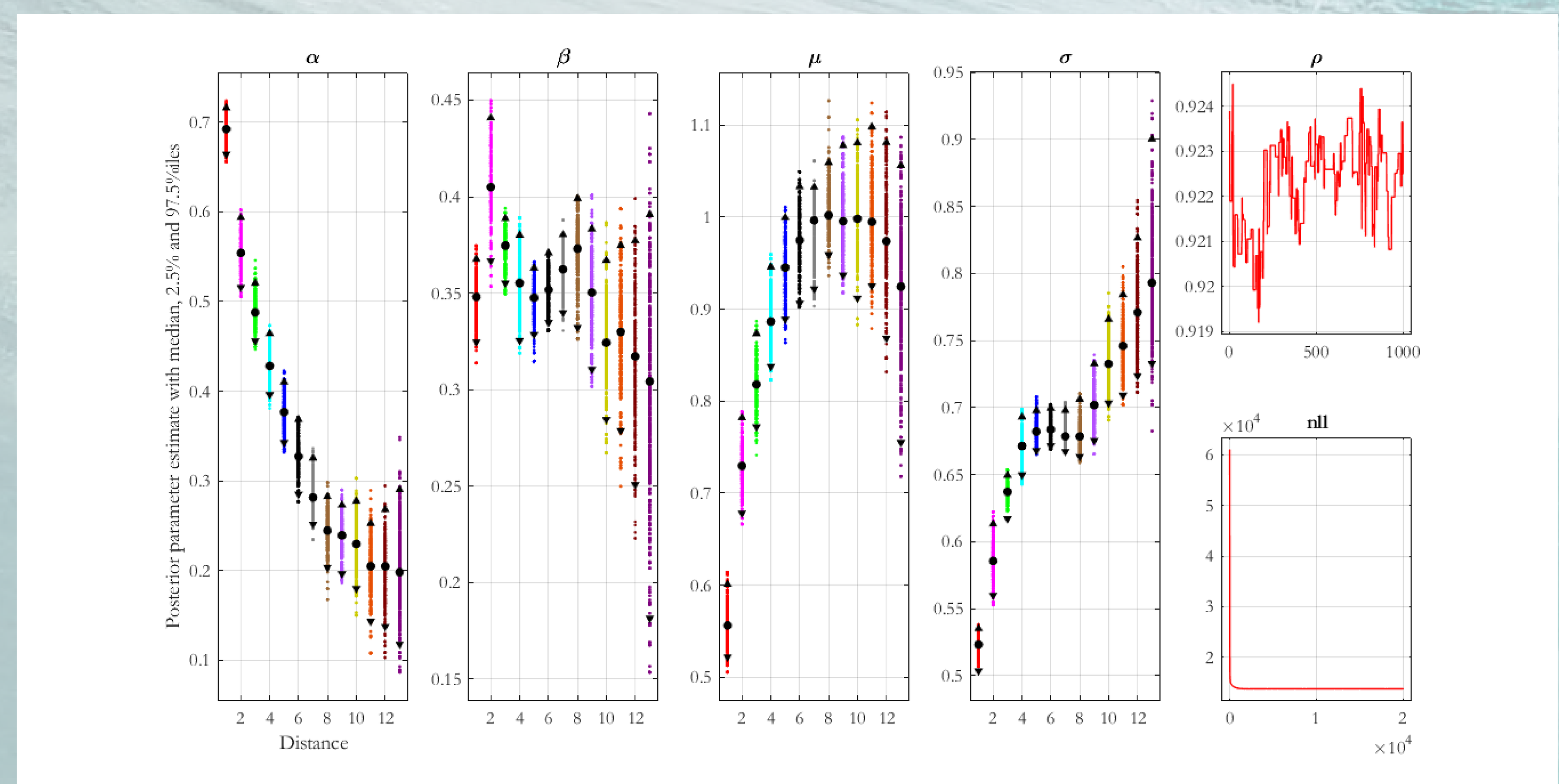
4. Northern North Sea (NNS) Data



- We use whitened data from a North Sea hindcast model, which comprises of storm peak significant wave heights during 1680 storm events at 150 sites; these sites are shown in the map to the left.
- As there is directionality within the data, we make use of directional transects, which allow simpler fitting on an approximately 1-dimensional line.
- For this analysis, a North-South transect is used; this is highlighted in red on the map.

5. Inference Using MCMC

- To obtain inference for the spatial Heffernan-Tawn model parameters, we utilise MCMC methods on model (1).
- In particular, we use an adaptive Metropolis-within-Gibbs algorithm, which proposes the parameters $\{\alpha(i), \beta(i), \mu(i), \sigma(i)\}_{i=1}^{13}$ jointly, so that parameters of a given lag are estimated jointly.
- The correlation parameter ρ of the Gaussian process $W(\cdot)$ is estimated separately.
- Constraints on $\alpha(h)$ and $\beta(h)$, such as those suggested by Keef et al. (2013), are also implemented within the algorithm.
- We pool data corresponding to particular lags (thus assuming the spatial grid is stationary and isotropic) to give us more information and should improve inference.
- Posterior estimates of these parameters are shown below.

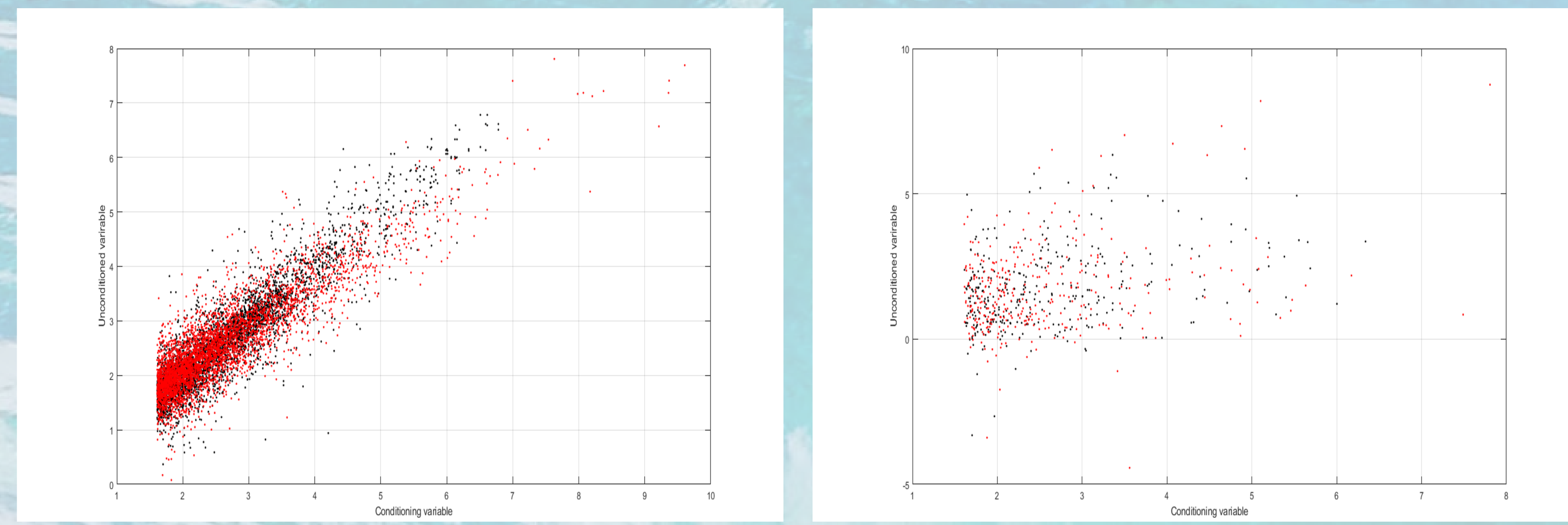


6. Inference using MCMC (continued)

- Results appear to be physically plausible; posterior variance increases with lag (so less information available). It appears that $\alpha(h)$ decays exponentially with distance h, whilst $\mu(h)$ and $\sigma(h)$ increase as h increases.
- Trace plots suggest that mixing of the MCMC chains is reasonable, with the greatest issues caused by confounding of $\alpha(h)$ and $\mu(h)$ and the trade-off required in $\alpha(h)$ and $\beta(h)$ by the imposed constraints.

7. Simulation from Fitted Model

- To further assess model fit, we simulate from the fitted model and compare the simulated data to the original data; examples are shown below.



- From this, it can be seen that the simulated data at both the shortest (left) and longest (right) lags are able to capture the behaviour in the original data quite well - suggesting the model fit is reasonably good.

8. Further Work

- Improving the MCMC procedure by utilising, for example, MMALA techniques, plus implementation of a parameterised form for the parameters $\{\alpha_i, \beta_i, \mu_i, \sigma_i\}$, for all lags i.
- Extension of the model to 2 dimensions, so as to reflect the spatial grid more accurately, as well as to including covariate effects, such as storm direction, and spatial effects such as the effect of the conditioning location.
- Application of the model to different ocean basins, which may have different physical characteristics, and hence different behaviour of spatial dependence of the extremes.

References

- Heffernan, J. E. and Tawn, J. A. (2004). Journal of the Royal Statistical Society: Series B 66:497-546. A conditional approach for multivariate extreme values.
- Keef, C., Papastathopoulos, I., Tawn, J. A.(2013). Journal of Multivariate Analysis. Estimation of the conditional distribution of a vector variable given that one of its components is large: additional constraints for the Heffernan and Tawn model.
- Tawn, J. A., Shooter, R., Towe, R. P., Lamb, R. (2018). Spatial Statistics. Modelling Spatial Extreme Events with Environmental Applications.

http://www.lancs.ac.uk/pg/shooter

r.shooter@lancaster.ac.uk