Obtaining extremes from models of future climate

Peter Challenor [1]

[1]National Oceanography Centre

Climate models

If we are to make inferences about future climate we cannot rely on observations. We must use models.

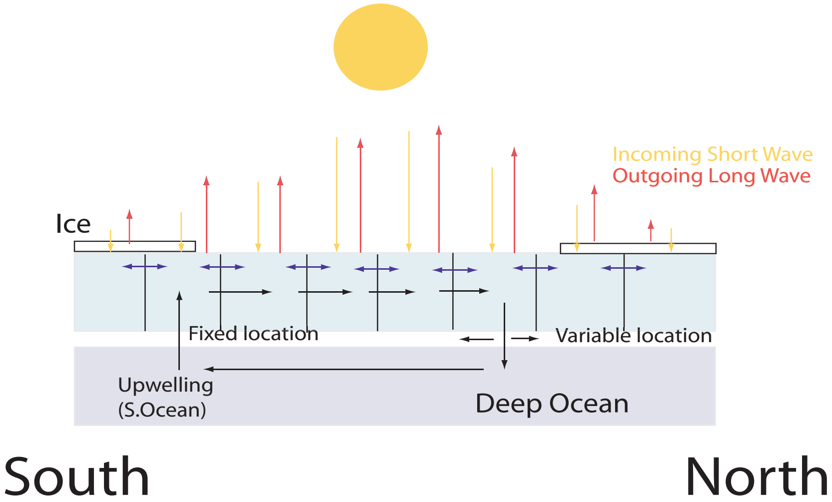
These vary in complexity from simple energy balance models to large scale AOGCM's

Climate models/simulators

- Energy Balance Models (EBM)
- Earth Models of Intermediate Complexity (EMIC)
- General Circulation Models (GCM)
- Earth System Models

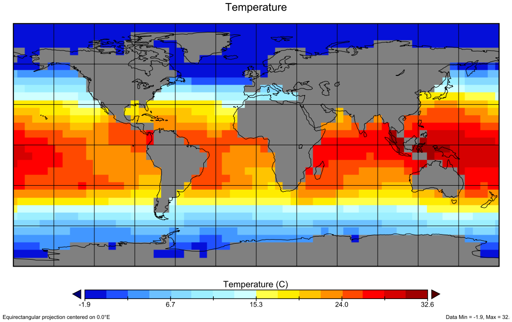
An Energy Balance Model

SurfEBM

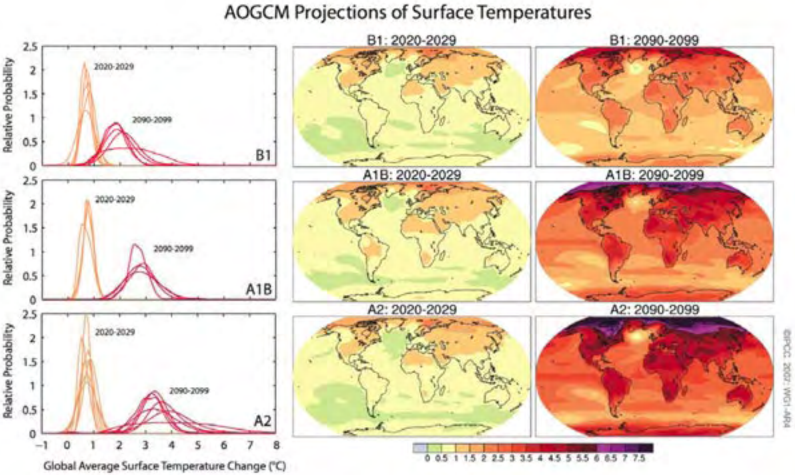


An EMIC

GENIE



AOGCM



AOGCM Projections of Surface Temperatures
B1: 2020-2029
B1: 2090-2099
A1B: 2020-2029
A1B: 2090-2099
A2: 2020-2029
A2: 2090-2099
Relative Probability
2020-2029
2090-2099
B1
A1B
A2
Global Average Surface Temperature Change (°C)
©IPCC 2007: WG1-AR4

Extremes from Climate Models

- Extreme Climates
- Extreme weather within predicted climates

Uncertainty in Climate Modelling

Our models are in general deterministic but uncertainty comes from a number of sources

- Parameter Uncertainty
- Initial Condition Uncertainty
- Boundary Condition Uncertainty
- Structural Uncertainty

Uncertainty in Forecasting

Lorentz defined two types of forecasting uncertainty

Type 1:

- Uncertainty comes from Initial conditions (Weather forecasting)

Type 2:

- Uncertainty comes from parameter/ boundary condition uncertainty (Climate prediction)

Extreme Climates

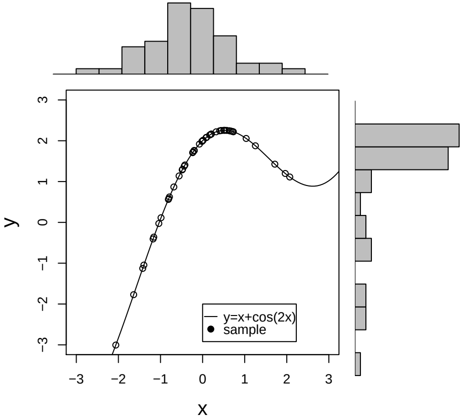
Climate models give predictions of the *expected* climate

However because of the uncertainties the model projections are not inconsistent with other more extreme climates.

To guide policy we would like to know the chance of getting such extreme (and unpleasant) climates

Could we cross a *tipping point*?

Monte Carlo Estimate



MC variance is proportional to $\frac{1}{n}$

Emulators

To enable us to do Monte Carlo (or similar) calculations we need to be able to run the model many times

But climate models are computationally very expensive

We therefore build emulators - fast approximations to the climate model (or simulator)

Emulators

- We have a simulator
$$y = f(x)$$
- Treat $f(.)$ as an unknown random function and use Bayesian methods to estimate it
- (Note it is possible to redo all this work in a frequentist way see book by *Santner et al.*)

Gaussian Processes

- We model the simulator with a Gaussian Process (GP)
- This has mean

  $$y = f(x) = h(x)^T \beta$$

  h(.) is a known vector of regressor (or basis) functions
  e.g. $h(x)^T = (1, x, x^2)$
- Variance σ^2
- And correlation function $c(x, x')$ Usually, but not invariably, we take

  $$c(x, x') = exp(-(x - x')^T C(x - x'))$$

  $$C^{-1} = diag(\delta)$$
- so we have a set of GP parameters (β, σ, δ)

The Prior

If we have prior information on the parameters we can include this but often we use a non-informative prior

$$\pi(\beta, \sigma^2) \propto \sigma^{-2}$$

The Posterior

$\eta(x) \sim t_{n-q}$

$$E(\eta(x)) = h(x)^T \beta' + t(x)^T A^{-1} (y - H\beta')$$

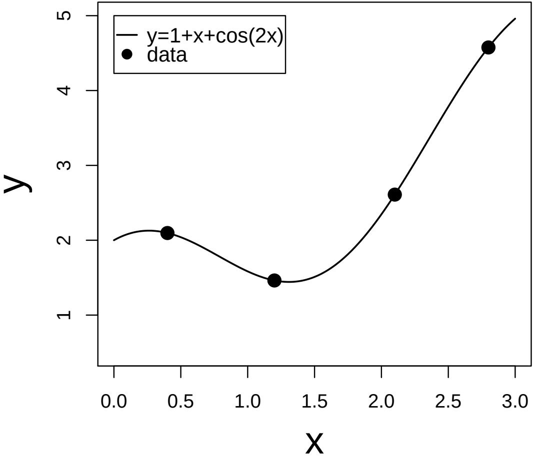
$$\beta' = \left(H^T A^{-1} H\right)^{-1} H^T A^{-1} y$$

H is the matrix $\{h(x_1), \cdots, h(x_n)\}^T$
These are the regression terms at x

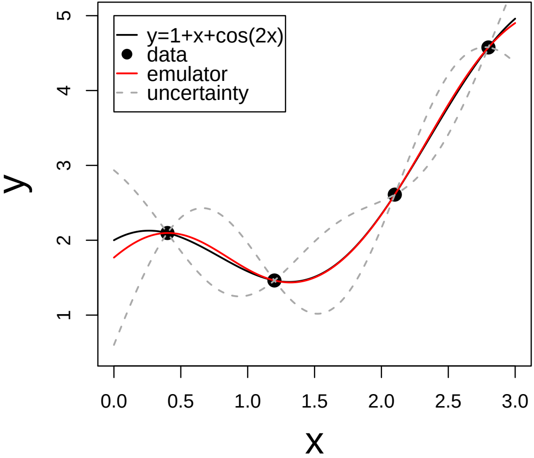
$$t(x) = \{c(x, x_1), \cdots, c(x, x_n)\}$$

This is the correlation of x with the data, x_i
A is the matrix $\{c(x_i, x_j)\}$
This is the correlation matrix of the data with itself
And there are similar, but more complex, expressions for the variance

Example



Example


y=1+x+cos(2x)
data
emulator
uncertainty
y
x
0.0
0.5
1.0
1.5
2.0
2.5
3.0
1
2
3
4
5

Smoothness

- δ is the smoothness or scale of the GP
- δ is not included in our posterior because it isn't included in the Bayesian solution
- Use maximum posterior (likelihood) or cross validation to estimate δ
- Can use MCMC to include uncertainty on δ in our emulator

Smoothness

- δ is the smoothness or scale of the GP
- δ is not included in our posterior because it isn't included in the Bayesian solution
- Use maximum posterior (likelihood) or cross validation to estimate δ
- Can use MCMC to include uncertainty on δ in our emulator

Smoothness

- δ is the smoothness or scale of the GP
- δ is not included in our posterior because it isn't included in the Bayesian solution
- Use maximum posterior (likelihood) or cross validation to estimate δ
- Can use MCMC to include uncertainty on δ in our emulator

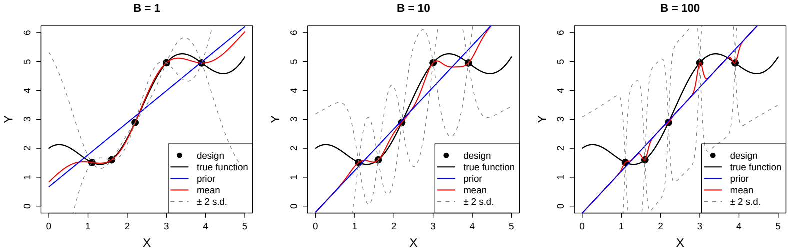
Smoothness

- δ is the smoothness or scale of the GP
- δ is not included in our posterior because it isn't included in the Bayesian solution
- Use maximum posterior (likelihood) or cross validation to estimate δ
- Can use MCMC to include uncertainty on δ in our emulator

Smoothness

- δ is the smoothness or scale of the GP
- δ is not included in our posterior because it isn't included in the Bayesian solution
- Use maximum posterior (likelihood) or cross validation to estimate δ
- Can use MCMC to include uncertainty on δ in our emulator

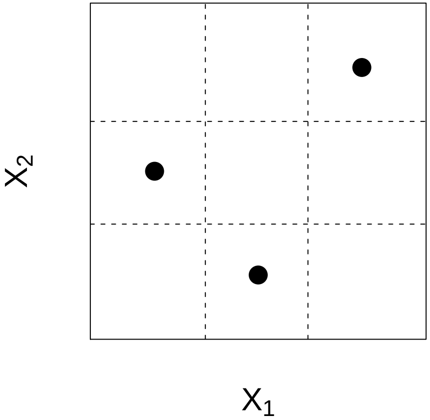
Example



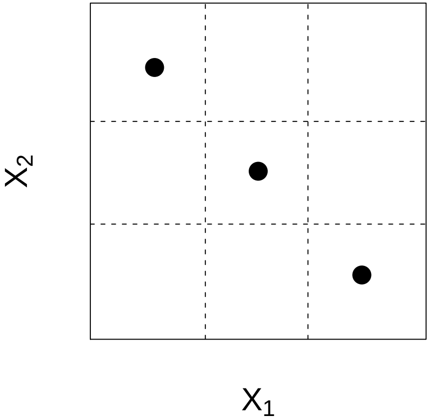
Design

- We need to design our initial experiment to make the training set
- We want a design that spans input parameter space
- Popular designs are
  - Latin Hypercubes
  - Sobol' Sequences
- Sequential Designs

The Latin hypercube



Not all Latin hypercubes are equal



Further Information on Emulators

See

www.mucm.ac.uk

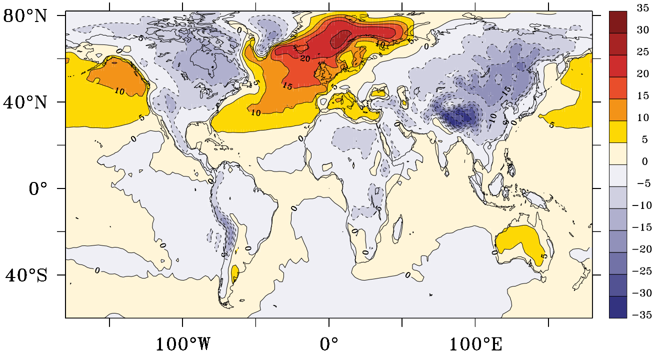
Look under toolkit

An example

The Collapse of the Thermohaline Circulation

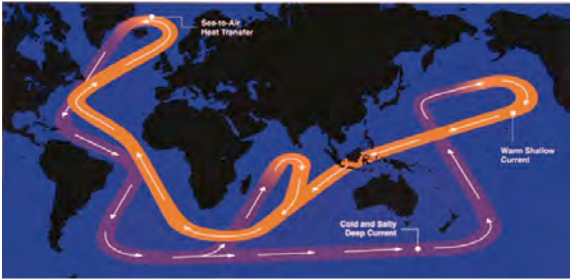
The thermohaline circulation

North West Europe is warm compared to similar latitudes



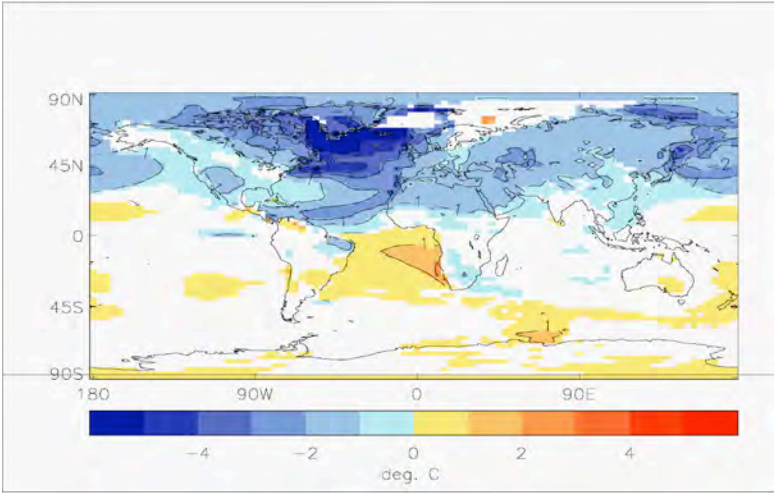
The thermohaline circulation

- This is because heat is transported N in the Atlantic
- This heat comes, not from the Gulf Stream but from the Thermohaline Circulation
- Cold salty water sinks in the North and flows south at depth
- Warm, fresh water is brought north



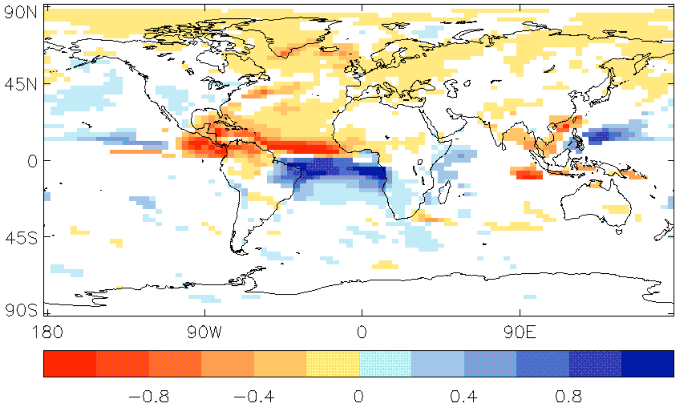
What if the thermohaline circulation collapsed today?

Surface air temperature change 20-30 years after THC shutdown. THC recovers after 120 years (Vellinga and Wood, 2002)



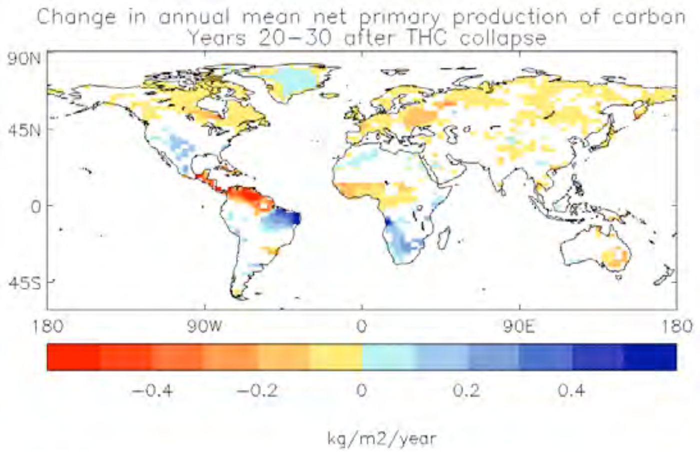
What if the thermohaline circulation collapsed today?

And precipitation



What if the thermohaline circulation collapsed today?

And primary production



The Big Question

- What is the probability that the Meridional Overturning Circulation in the North Atlantic (MOC) will collapse by 2100?
- What is the probability that the Meridional Overturning Circulation in the North Atlantic (MOC) in a ensemble of different models will collapse by 2100?
- What is the probability that the Meridional Overturning Circulation in the North Atlantic (MOC) in a particular model will collapse by 2100?

The Big Question

- What is the probability that the Meridional Overturning Circulation in the North Atlantic (MOC) will collapse by 2100?
- What is the probability that the Meridional Overturning Circulation in the North Atlantic (MOC) in a ensemble of different models will collapse by 2100?
- What is the probability that the Meridional Overturning Circulation in the North Atlantic (MOC) in a particular model will collapse by 2100?

The Big Question

- What is the probability that the Meridional Overturning Circulation in the North Atlantic (MOC) will collapse by 2100?
- What is the probability that the Meridional Overturning Circulation in the North Atlantic (MOC) in a ensemble of different models will collapse by 2100?
- What is the probability that the Meridional Overturning Circulation in the North Atlantic (MOC) in a particular model will collapse by 2100?

GENIE-1

- GENIE has *about* 15 (+2) unknown input parameters

Inputs

Parameters

- Ocean viscosity
- Moisture transport
- Climate Sensitivity ...

Forcings

- Carbon dioxide
- Greenland Melting ...

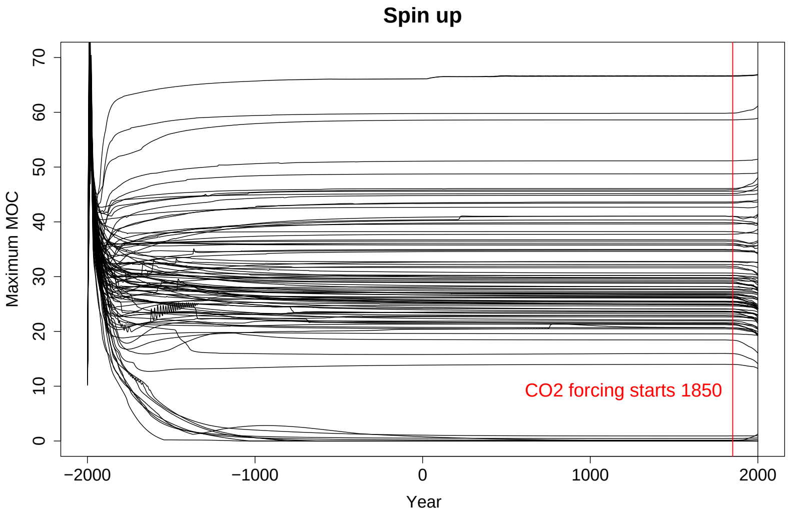
Outputs

- Ocean/air temperature
- Rainfall
- Salinity
- Heat/moisture fluxes
- Ocean currents
- Max. Atlantic overturning circulation

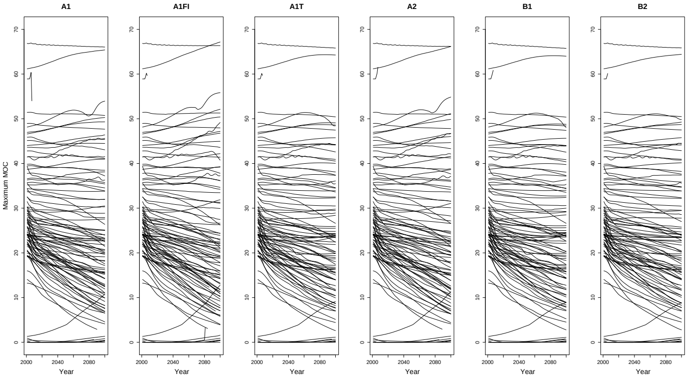
The Training Experiment

- To generate the training set we run the model in a designed experiment
- This ensemble is not designed to give a realistic climate but to span parameter space
- In our experiments we have one hundred member ensembles for training
- The design is a maxi-min Latin hypercube

Spin-up of GENIE-1



GENIE projects the Future



Calibrating the model

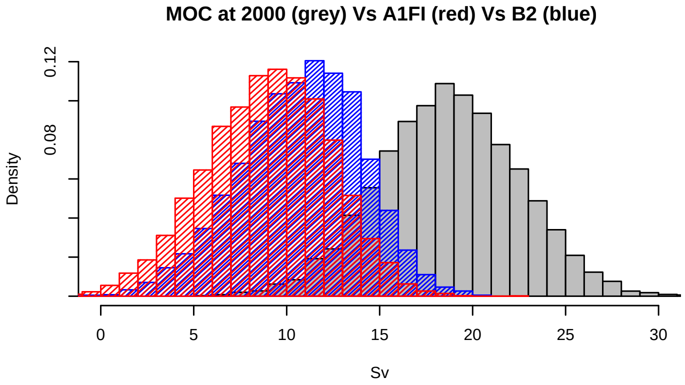
- So far we have not included any data in our analysis
- Only 5 estimates of the strength of the MOC available (1957, 1981, 1989, 1998, 2004)
- Rejection Sampling

Calibrating the model

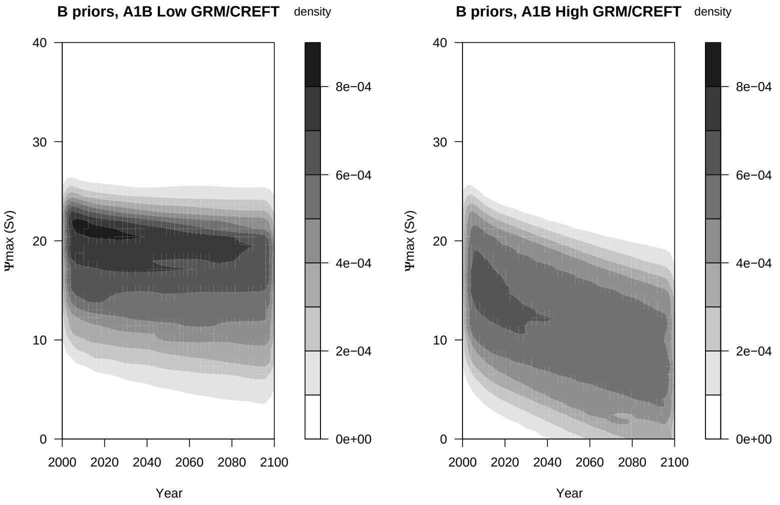
- So far we have not included any data in our analysis
- Only 5 estimates of the strength of the MOC available (1957, 1981, 1989, 1998, 2004)
- Rejection Sampling

Calibrating the model

- So far we have not included any data in our analysis
- Only 5 estimates of the strength of the MOC available (1957, 1981, 1989, 1998, 2004)
- Rejection Sampling


MOC at 2000 (grey) Vs A1FI (red) Vs B2 (blue)
Density
0.08
0.12
0
5
10
15
20
25
30
Sv

Our pdf max MOC through the 21st century



Probability of the MOC dropping below a threshold of 10 Sv during the 21st century, for three indicative scenarios.

Scenario	Low melt rate	High melt rate
A1B	0.09	0.35
A2	0.10	0.33
B1	0.10	0.30

Weather extremes from climate models

- Downscaling
  - Dynamic downscaling (regional climate model)
  - Statistical downscaling
- Stationarity
  - Assume stationarity
  - Model stationarity statistically
  - Generate a stationary series

Generating a stationary series

- Run the model with fixed forcing for 20-30 years
- Run a ’weather’ generator’

The resulting series can be analysed by standard EV theory methods

Uncertainty in Extreme Weather

- These estimates don't have the simulator uncertainty included
- One way to do this would be to build an emulator
- Can we emulate extremes directly across an ensemble of climate models?
- It is unrealistic to use GP's for this as we know that extremes have very non-Gaussian distributions

Max stable processes

- In a similar way to Gaussian processes we can define max stable processes whose marginals are the extreme value distributions
- These are much more complex than Gaussian processes
- Brown-Resnick processes may be suitable for building extremal emulators

$$\eta(t) = \bigvee_{i=1}^{\infty} \left\{ U_i + W_i(t) - \frac{\sigma^2(t)}{2} \right\}$$

An Alternative

- The Generalised Extreme Value (GEV) distribution has three parameters
- Extremes do not have Gaussian distributions
- But the parameters of the GEV can be assumed to do so
- Therefore we can jointly emulate the three parameters
- Since these three jointly define the extremes we have, in effect, a GP emulator for extremes
- The max stable process defined this way is less adaptable than the Brown-Resnick process