Estimating extremes from global ocean and climate models: A Bayesian hierarchical model approach

E. C. J. Oliver[1,2], Simon J. Wotherspoon[1], and Neil. J Holbrook[1,2]

[1] Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Australia

[2] Australian Research Council Centre of Excellence for Climate System Science

1 Introduction

- **Global climate and ocean models** are indispensible tools for understanding of the ocean-atmosphere-climate system
- Models usually simulate well the basic oceanic and atmospheric climate state (mean, variance, etc.) but **often fail to properly represent extreme events**
- It has been shown that extreme events can be well predicted given only knowledge of the **central statistics (mean, variance, etc.)**
- On this principle we have developed a **Bayesian hierarchical model** which **improves the estimates of extremes** from global climate and ocean models

2 Data

Ocean model

The **Ocean Forecasting Australia Model (OFAM)** was used to model the marine climate. The horizontal resolution is 1/10^o in latitude and longitude (eddy-resolving) around Australia and coarser elsewhere. We used two OFAM model runs (Chamberlain et al, 2012):

CTRL: Forced by normal-year ERA-Interim surface fluxes, representing the 1990s

A1B: Dynamically downscaled CSIRO Mk3.5 GCM projections under the A1B emissions scenario, representing the 2060s

Climate statistics were calculated including the mean ($\boldsymbol{\mu}$), variance ($\boldsymbol{\sigma^2}$), third central moment ($\boldsymbol{m_3}$), and eddy kinetic energy ($\boldsymbol{K}$) over a domain consisting of *J* locations. Various combinations are assembled into the covariate matrices $\boldsymbol{X}$, e.g,

$$\boldsymbol{X} = [\mathbf{1} \mid \boldsymbol{\mu} \mid \boldsymbol{\sigma}^2 \mid \boldsymbol{K}]$$

Observations

Daily fields of remote-sensed **observed sea surface temperatures (SSTs)** from the AVHRR were obtained for the period 1/1/1982 to 31/12/2009 (28 years) and are defined on a 4 km grid. Data were taken at the same *J* locations as in the ocean model.

3 Extreme Value Theory

Extreme events are those producing climate anomalies which are **rare** and whose **magnitudes deviate significantly** from the expected value

Consider the sequence $\{x_t | t = 1, 2, \dots\}$ (e.g., SST time series)

Define *y* to be the maximum over a block of length *n* = 1 year (a **"block maxima" approach**; Coles, 2001)

$$y = \max(x_1, x_2, \dots, x_n)$$

The **annual maxima** (*y*) can be modeled using an Extreme Value Distribution (EVD), and we will use the Type I, or the Gumbel, distribution:

$$F(y|a,b) = \exp\left[-\exp\left(-\frac{x-a}{b}\right)\right]$$

The return period for a specified extreme value represents the expected frequency with which that extreme value will repeat. For example, if the return period for a 30^{o}C extreme value, denoted T_{30}, is 50 years then there is a 1 in 50 chance of a 30^{o}C event occurring in any given year. Conversely, the return level is the extreme value associated with a particular return period. In the previous example, the 50-year return level, denoted z_{50}, is 30^{o}C.

Return levels z_T and **return periods** T_z are defined as:

$$z_T(y) = a - b\log\left[-\log F_{\mathrm{I}}(y|a,b)\right]$$
$$T_z(y) = \left[1 - F_{\mathrm{I}}(y|a,b)\right]^{-1}$$

Given a vector of annual maxima ($\boldsymbol{y}$), the parameters $\boldsymbol{\theta}$=(*a*,*b*) can be estimated by minimizing the likelihood (maximum likelihood estimation):

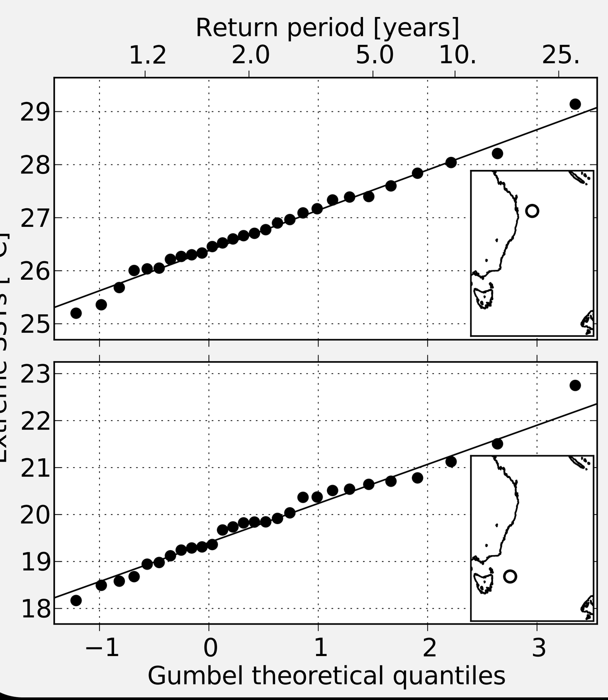
$$L(\boldsymbol{\theta}|\boldsymbol{y}) = p(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{i=1}^{N} f(y_i|\boldsymbol{\theta})$$

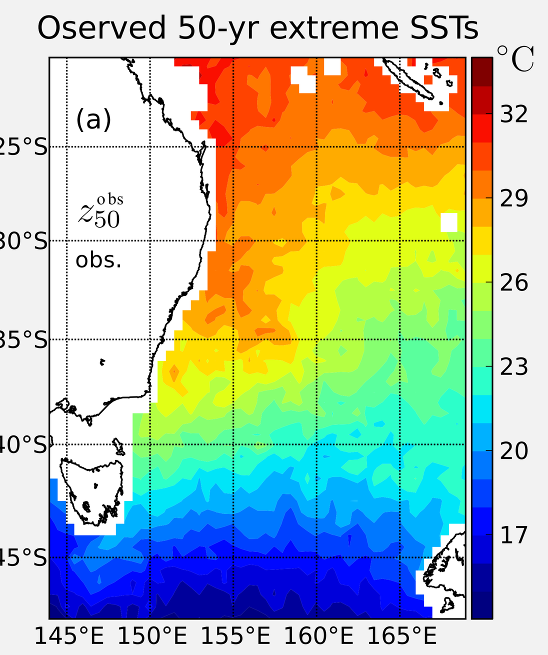
Gumbel pdf

or by **Bayesian estimation**:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\theta}) p(\boldsymbol{\theta})$$

posterior distribution, likelihood, prior distribution





4 Bayesian Hierarchical Extremes Model

Define annual maxima at all *J* locations as a list of vectors: $\boldsymbol{Y} = \{\boldsymbol{y}_t | j = 1, 2, \dots J\}$

We model the extremes using a **bayesian hierarchical model (BHM)**: a model with several nested layers

$$p(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{X}) \propto p(\boldsymbol{Y}|\boldsymbol{\theta}_2)\, p(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1, \boldsymbol{X})\, p(\boldsymbol{\theta}_1)$$

posterior distribution, data layer, climate process layer, prior distribution

i. Data layer

Model the annual maxima $\boldsymbol{Y}$ using the Gumbel dist.:

$$p(\boldsymbol{Y}|\boldsymbol{\theta}_2) = \prod_{j=1}^{J} p(\boldsymbol{y}_j|a_j, \phi_j) = \prod_{j=1}^{J}\prod_{i=1}^{N} f(y_{ji}|a_j, \phi_j)$$

where $\phi = \log(b)$, $\boldsymbol{\theta}_2 = (\boldsymbol{a}, \boldsymbol{\phi})$, $\boldsymbol{a} = \{a_j | j = 1, 2, \dots, J\}$, and $\boldsymbol{\phi} = \{\phi_j | j = 1, 2, \dots, J\}$.

ii. Climate process layer

The parameters of the Gumbel dist. are modeled as a function of the ocean model marine climate $\boldsymbol{X}$:

$$\boldsymbol{a} = \boldsymbol{X}\beta_a + \epsilon_a \quad \rightarrow \quad p(\boldsymbol{a}|\beta_a, \tau_a, \boldsymbol{X}) = \mathcal{N}_J(\boldsymbol{X}\beta_a, \tau_a^{-1}\mathbf{I})$$
$$\boldsymbol{\phi} = \boldsymbol{X}\beta_\phi + \epsilon_\phi \quad \rightarrow \quad p(\boldsymbol{\phi}|\beta_\phi, \tau_a, \boldsymbol{X}) = \mathcal{N}_J(\boldsymbol{X}\beta_\phi, \tau_\phi^{-1}\mathbf{I})$$

where $\mathcal{N}_J(\boldsymbol{X}\beta, \tau^{-1}\mathbf{I})$ is a *J*-dim Normal distribution with mean $\boldsymbol{X\beta}$ and covariance $\tau^{-1}\mathbf{I}$ and $\theta_1 = (\beta_a, \beta_\phi, \tau_a, \tau_\phi)$.

Since the models for $\boldsymbol{a}$ and $\boldsymbol{b}$ are independent, we can factor the climate process layer as

$$p(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1, \boldsymbol{X}) = p(\boldsymbol{a}|\beta_a, \tau_a, \boldsymbol{X})\; p(\boldsymbol{\phi}|\beta_\phi, \tau_\phi, \boldsymbol{X})$$

iii. Priors

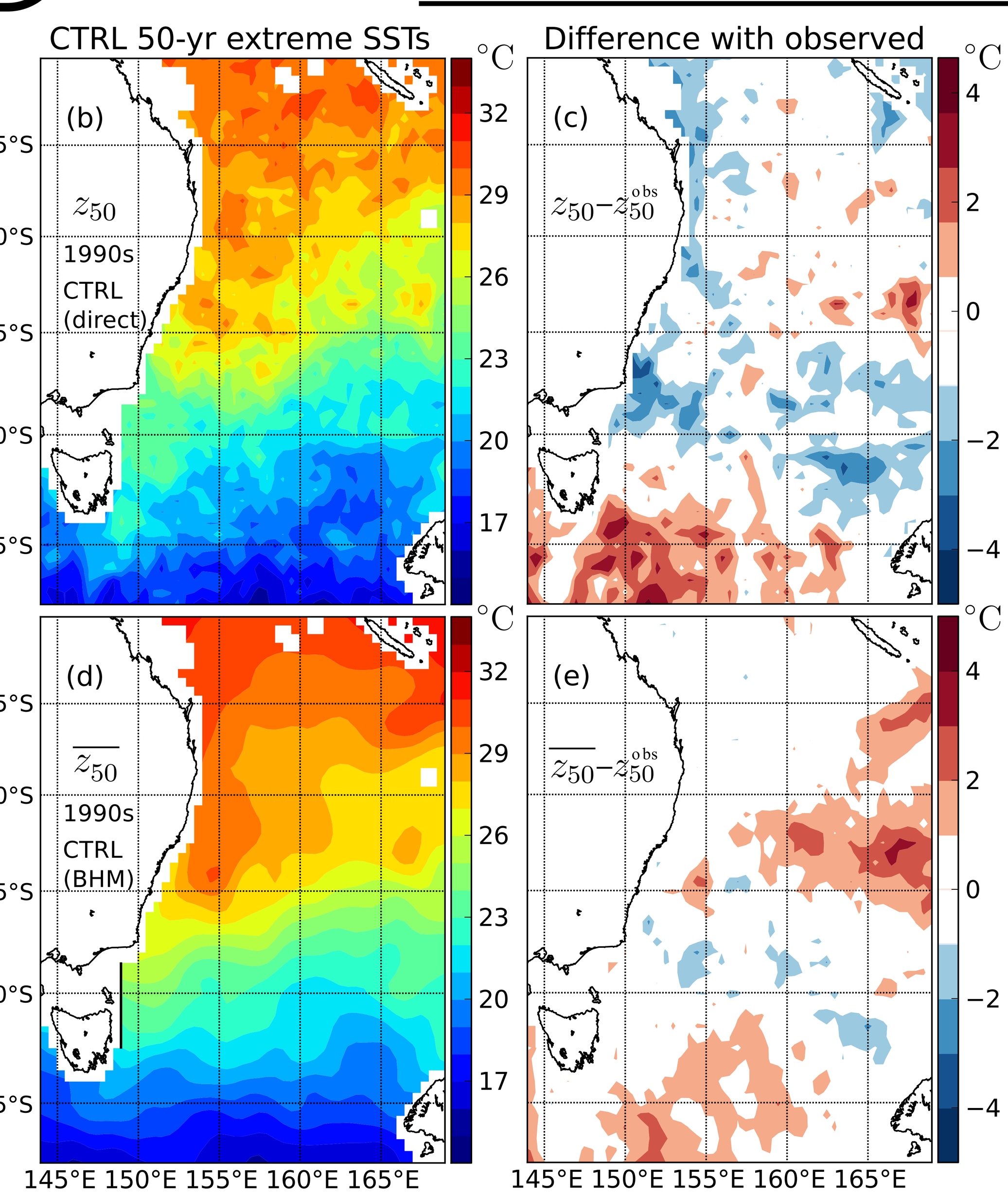
Assume that the parameters $\boldsymbol{\theta_1}$ are independent

$$p(\boldsymbol{\theta}_1) = p(\beta_a)p(\beta_\phi)p(\tau_a)p(\tau_\phi)$$

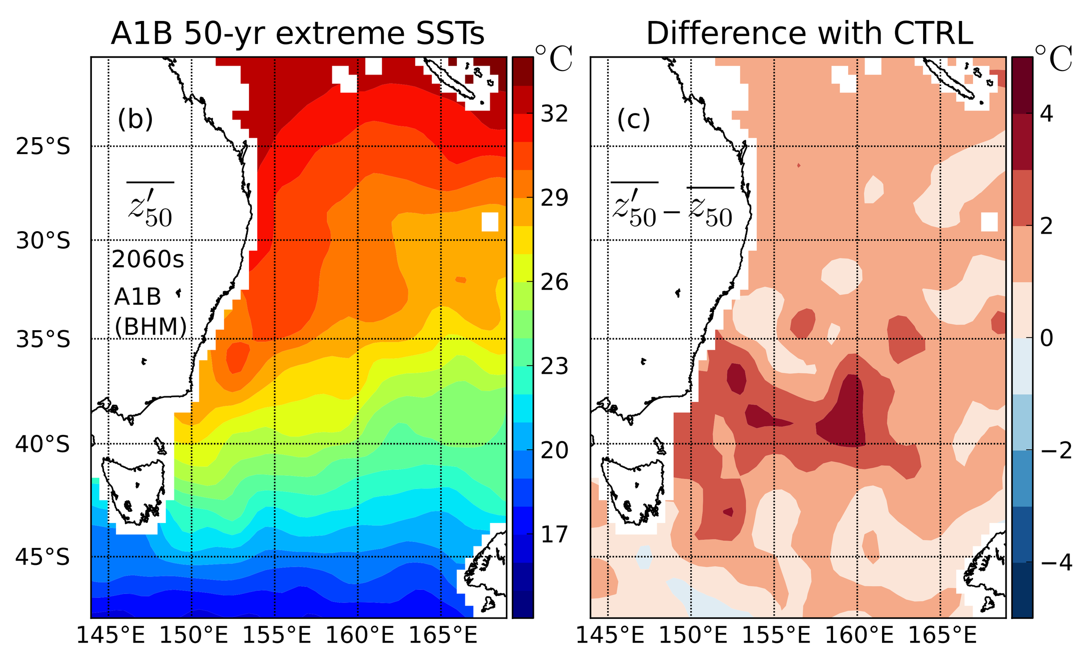
and with no prior knowledge regarding how the Gumbel parameters are related to the climate variables we choose diffuse non-informative priors.

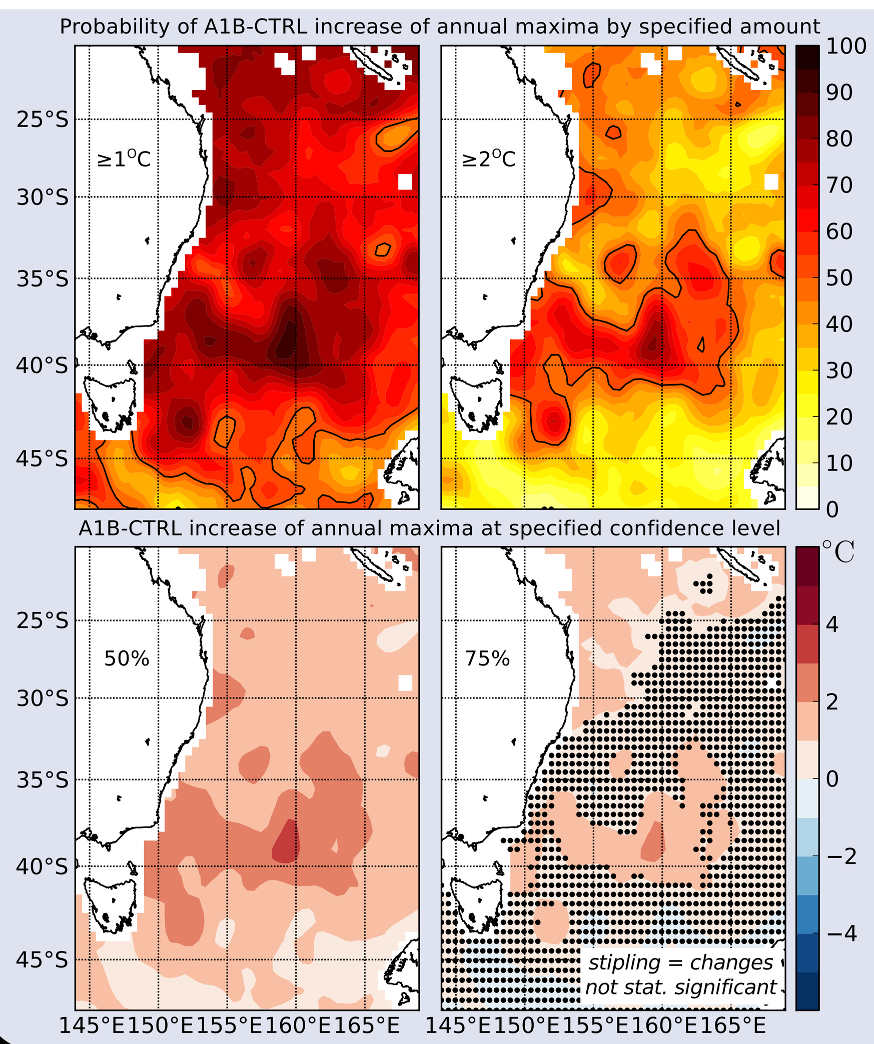
Model parameters are estimated numerically using Marlov chain Monte Carlo, Metropolis rule and Gibbs sampling methods

5 Model Estimates of Extreme SSTs



- The best **combination of climate variables** ($\boldsymbol{X}$) was determined (using the Deviance Information Criteria) to be the **mean, variance, third moment and EKE**
- Samples from the posterior are used to estimate $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{z_T}$...
- Estimates from the BHM **improve** on those from OFAM
- **Assuming stationarity** of the BHM parameters ($\boldsymbol{\theta_2}$), we construct a marine climate matrix from the 2060s ($\boldsymbol{X'}$) ocean simulation, then we can sample from $p(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{X}')$ and obtain **estimates of projected future extreme SSTs**



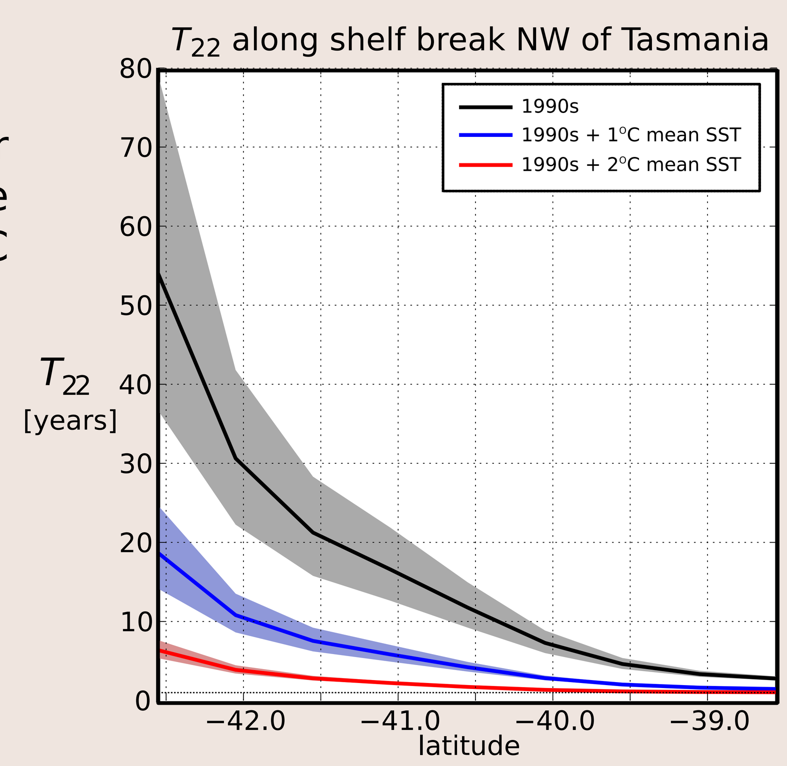


By drawing independent samples of $\boldsymbol{Y}$ from the posterior for the 1990s we can test if the projected change in the 2060s is **statistically significant**, when compared to the change possible in an **unchanged 1990s climate (due to random variations)**

Toy model

We can specify a particular climate ($\boldsymbol{X}_{\mathrm{spec}}$), such as the 1990s climate plus a 2^{o}C warming of the mean SST

Then drawing from the posterior distribution given the specified climate **we can test what the response of the extremes are to large-scale changes in the climate**



6 Conclusions

- The BHM extreme model provides a technique to **improve estimates of extremes** from global ocean and climate models
- The basic approach is to **model observed extremes as a function of the historical climate statistics**, assume stationarity, and then use climate projections or specified climates to estimate extremes for other climate scenarios
- The method allows for an **estimation of statistical significance** (by comparing against randomness in an unchanged climate) and can also be used as a **toy model** to test the response of the extremes to prescribed climate changes

Oliver, E. C. J., S. J. Wotherspoon and N. J. Holbrook, Estimating extremes from global ocean and climate models: A Bayesian hierarchical model approach, *Progress in Oceanography*, doi: 10.1016/j.pocean.2013.12.004 (in press, available online)

Oliver, E. C. J., S. J. Wotherspoon, M. A. Chamberlain and N. J. Holbrook, Projected Tasman Sea extremes in sea surface temperature through the 21st century, *Journal of Climate*, doi: 10.1175/JCLI-D-13-00259.1 (in press, available online)