Improving estimates of extremes from global ocean and climate models and assigning confidence to future projections

Eric C. J. Oliver[1,2], Simon J. Wotherspoon[1], Matthew A. Chamberlain[3], and Neil J. Holbrook[1,2]

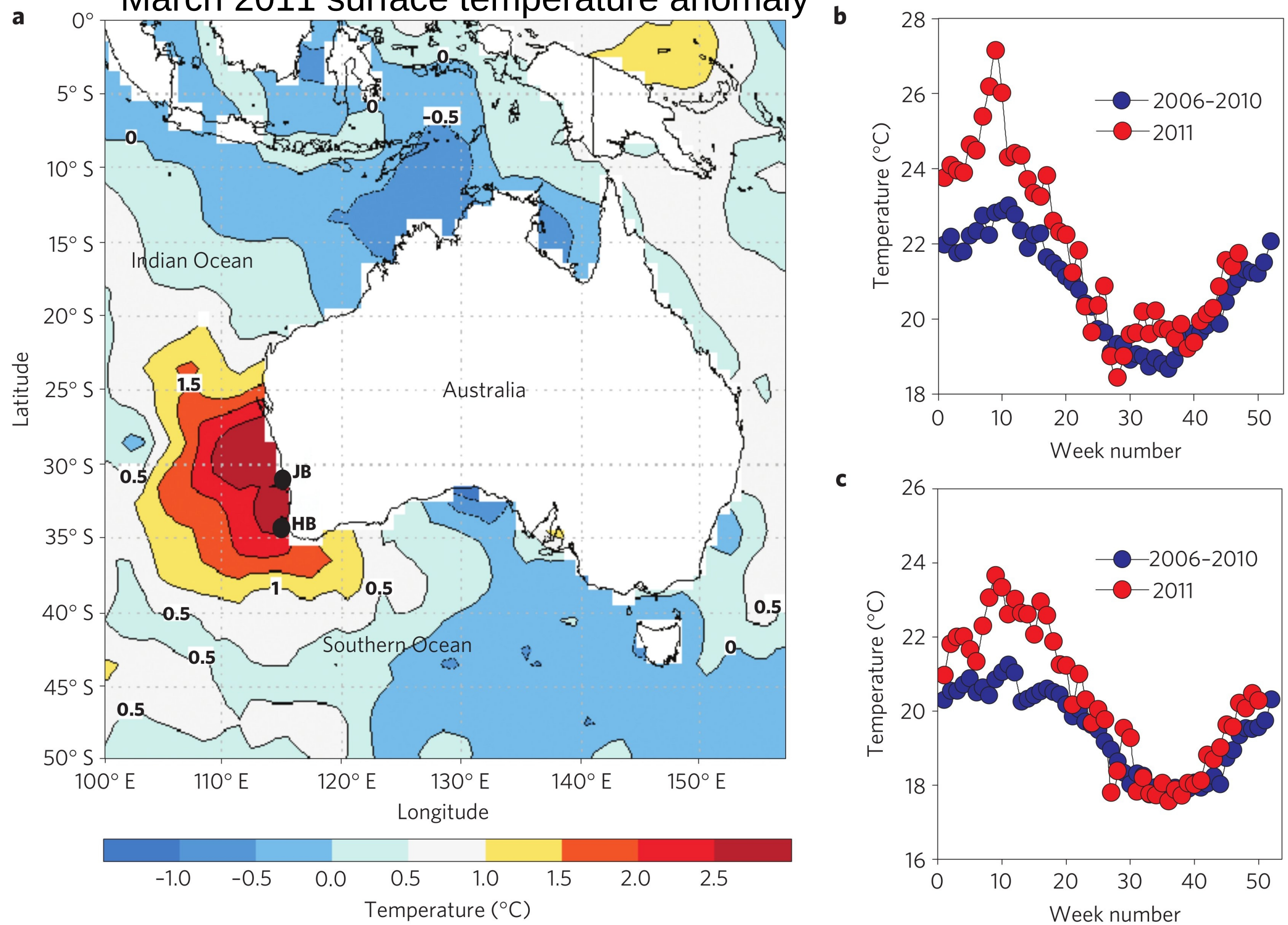
[1] Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, AUS
[2] Australian Research Council Centre of Excellence for Climate System Science
[2] CSIRO Marine and Atmospheric Research (CMAR), Hobart, AUS

translatingnatureintoknowledge

Marine Heat Waves

- In **2011**, a "**marine heat wave**" off of **Western Australia** was documented (Pearce and Feng, 2013; Feng et al., 2013)

March 2011 surface temperature anomaly



Wernberg et al (2013)

- Some species experienced range extensions during the marine heat wave which persisted after the heat wave dissipated (Wernberg et al. 2013)

Extreme Value Theory

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

Extreme Value Theory

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

- Consider a series ***y*** of **annual maxima** so that each element of the series is the maximum value of a variable that occurred within a unique year
- The annual maxima can be modeled using an **Extreme Value Distribution (EVD)**, e.g., the Type I or Gumbel distribution:

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

Extreme Value Theory

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

- Consider a series ***y*** of **annual maxima** so that each element of the series is the maximum value of a variable that occurred within a unique year
- The annual maxima can be modeled using an **Extreme Value Distribution (EVD)**, e.g., the Type I or Gumbel distribution:

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

- Given a vector of annual maxima (***y***), the parameters ***θ*** = (*a*,*b*) can be estimated by **Bayesian estimation**:

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

posterior distribution

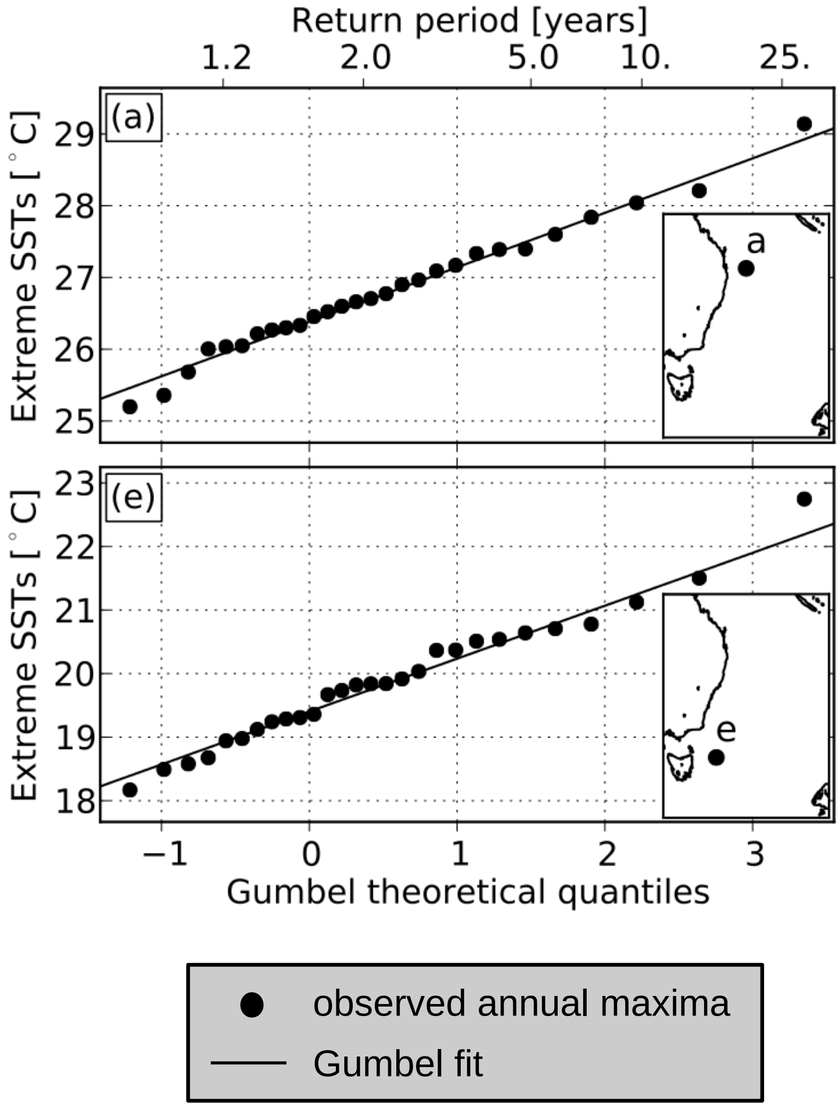
likelihood

prior distribution

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

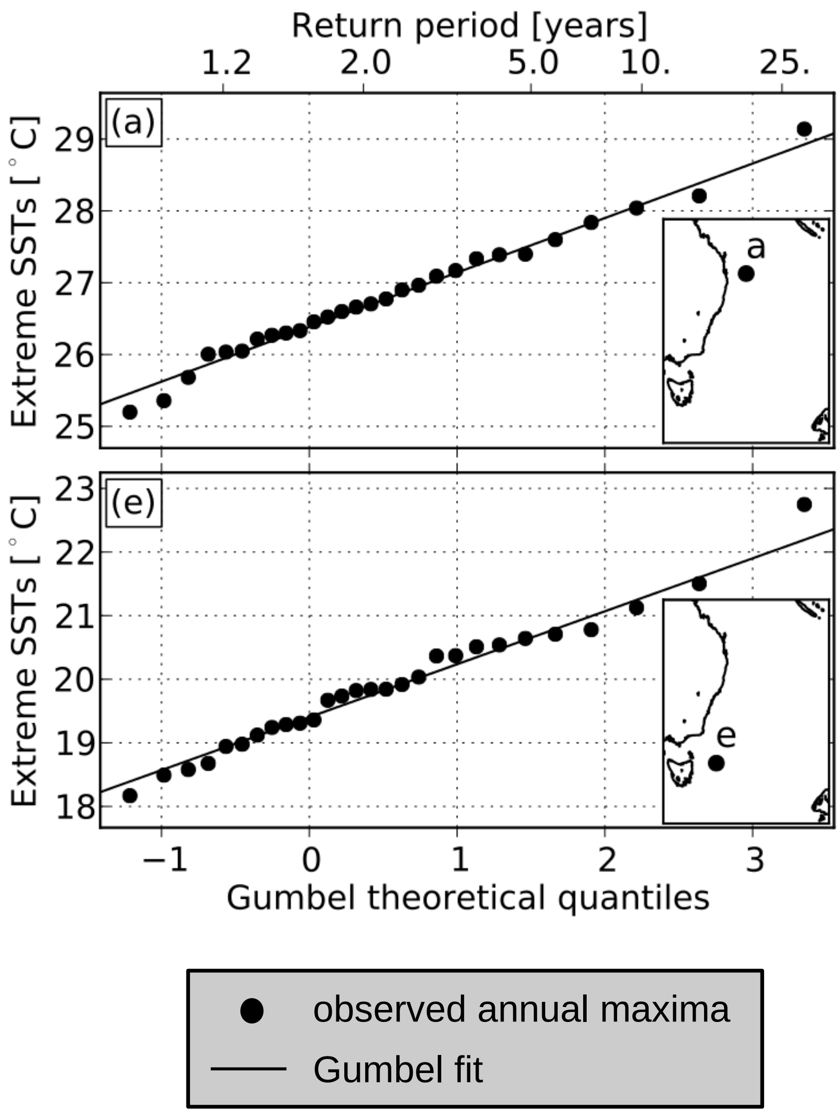
Extreme Value Theory

- The **fit of the Gumbel distribution** can be compared with the annual maxima using a return level plot:



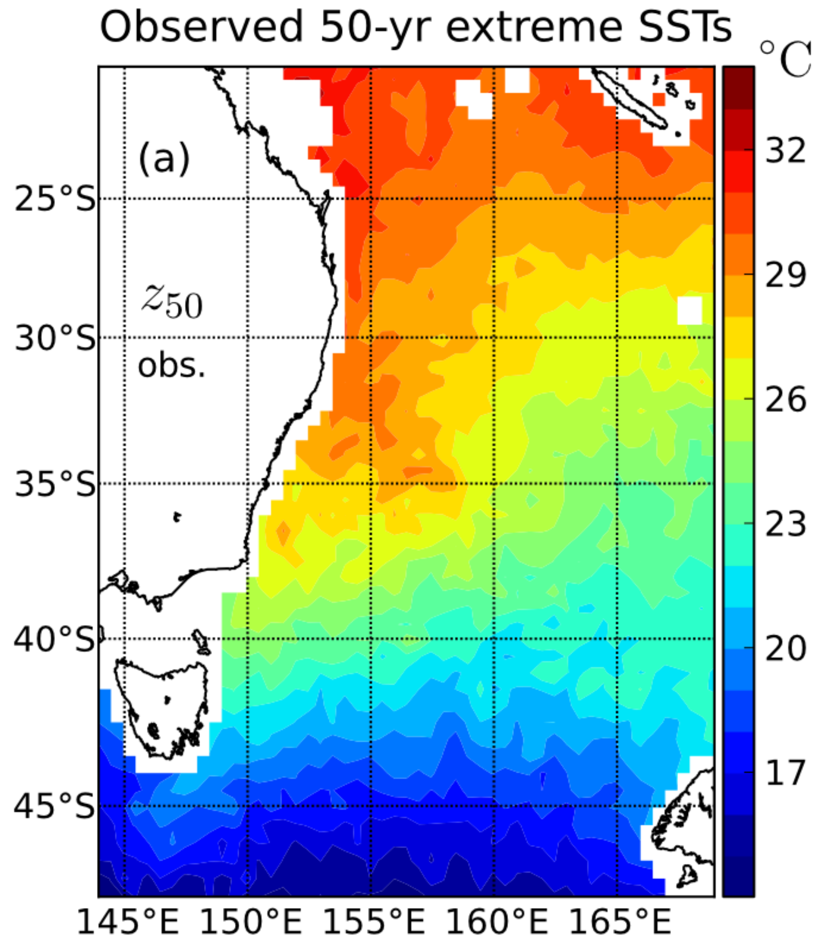
Extreme Value Theory

- The **fit of the Gumbel distribution** can be compared with the annual maxima using a return level plot:



- **Return levels** z_T and **return periods** T_z are defined using

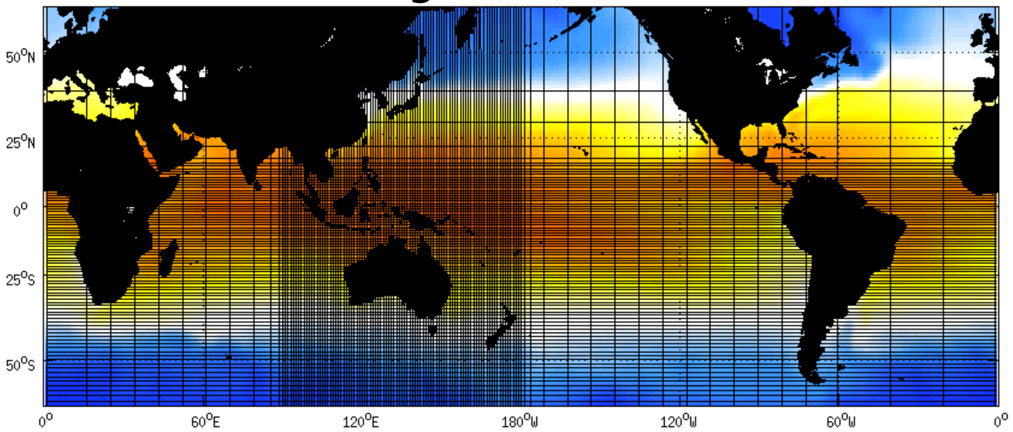
$$
\begin{aligned}
z_T(y) &= a - b\log\left[-\log F_{\mathrm{I}}(y|a,b)\right] \\
T_z(y) &= \left[1 - F_{\mathrm{I}}(y|a,b)\right]^{-1}
\end{aligned}
$$



Model Projections

- **Eddy-resolving dynamical downscaling in Australia region performed by Chamberlain et al. (2010):**
- Two ocean model runs using Ocean Forecasting Australia Model (**OFAM**; $70^{O}S–70^{O}N$ domain, $1/10^{O}$ resolution around Australasia)

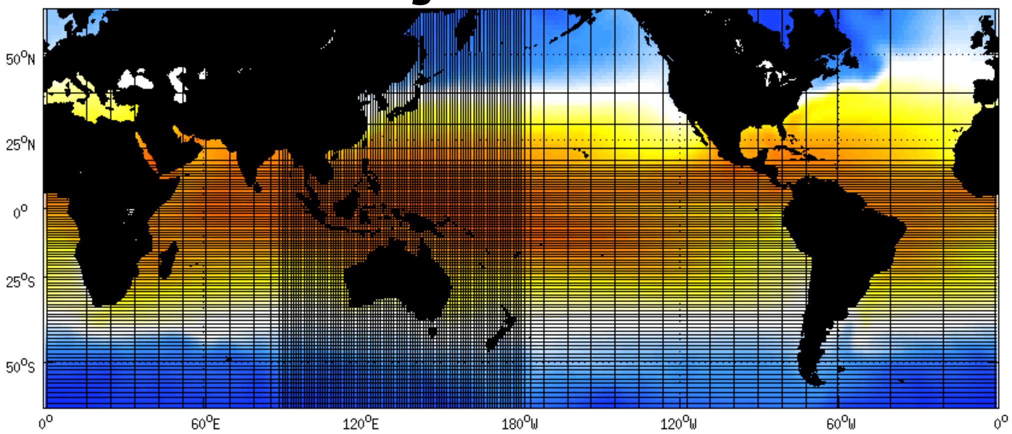
OFAM grid with mean SST



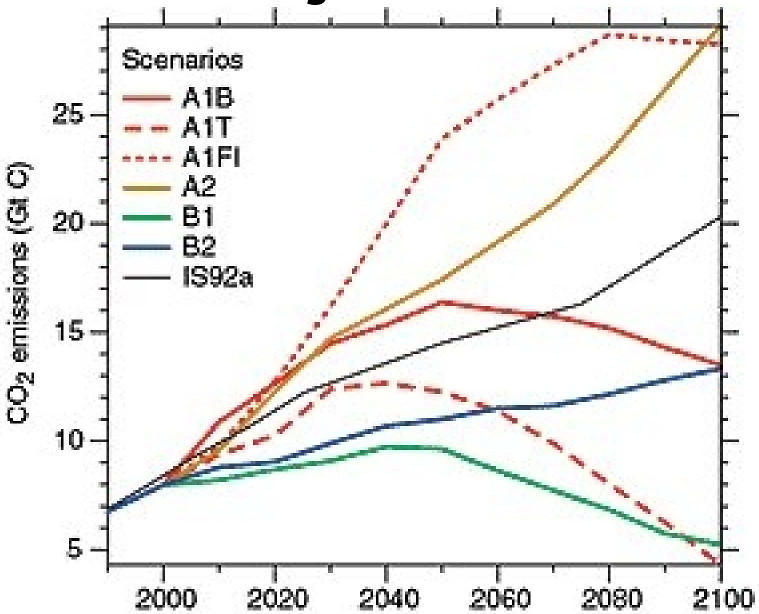
Model Projections

- **Eddy-resolving dynamical downscaling in Australia region performed by Chamberlain et al. (2010)**:
- Two ocean model runs using Ocean Forecasting Australia Model (**OFAM**; 70°S–70°N domain, 1/10° resolution around Australasia)
- Forcings representative of:
  - **1990s (CTRL run)**, and
  - **2060s (A1B run)**
- Control run forced by **historical reanalysis**
- **Climate change scenario** provided by CSIRO Mk3.5 GCM with an A1B emissions scenario

OFAM grid with mean SST



Climate change emissions scenarios



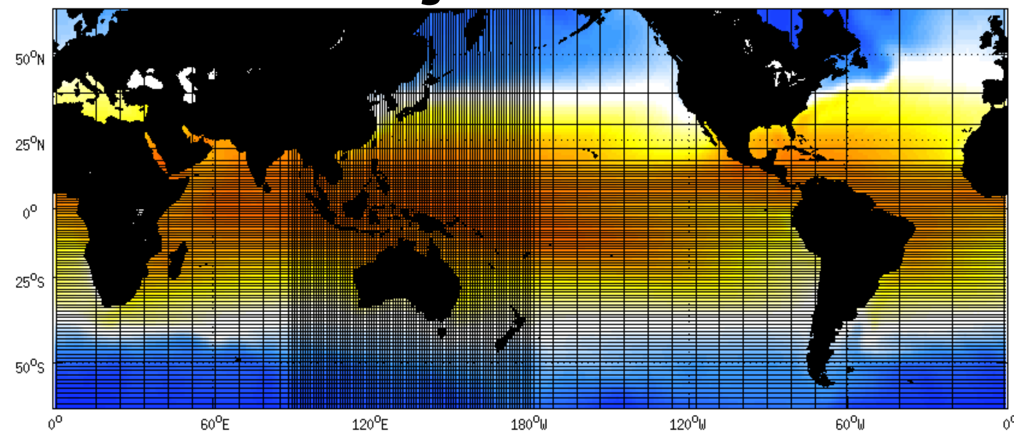
IMAS INSTITUTE FOR MARINE AND ANTARCTIC STUDIES

Model Projections

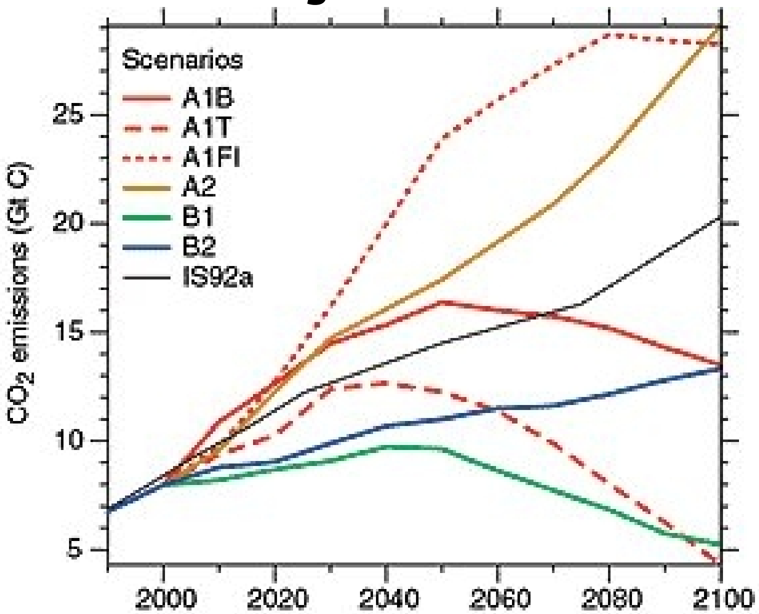
UTAS

- **Eddy-resolving dynamical downscaling in Australia region performed by Chamberlain et al. (2010):**
- Two <u>ocean model runs</u> using Ocean Forecasting Australia Model (**OFAM**; 70°S–70°N domain, 1/10° resolution around Australasia)
- Forcings representative of:
  - **1990s (CTRL run)**, and
  - **2060s (A1B run)**
- Control run forced by **historical reanalysis**
- **Climate change scenario** provided by CSIRO Mk3.5 GCM with an A1B emissions scenario
- Models represent well general circulation and temperature distribution around Australia, including seasonality **[Sun et al, 2012; Matear et al., 2013]**

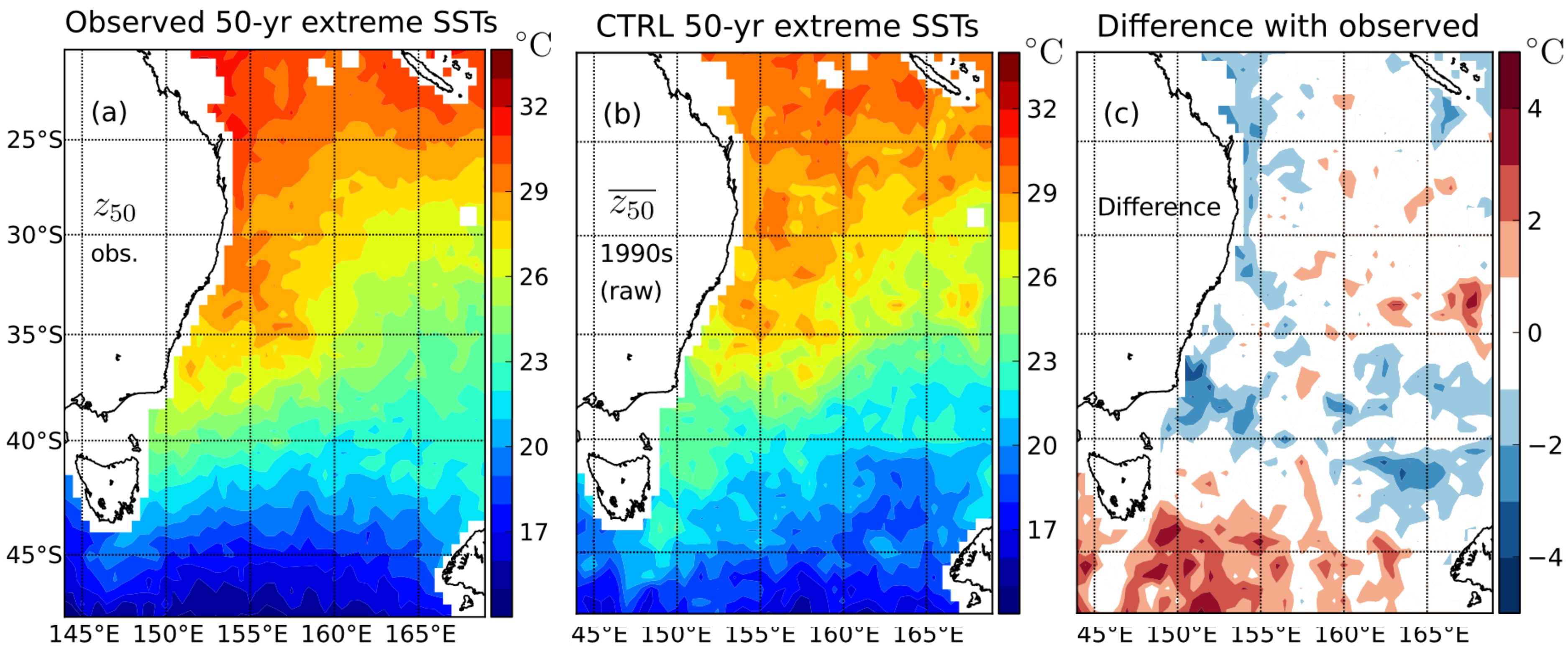
OFAM grid with mean SST



Climate change emissions scenarios



Model Projected Extremes



- The ocean model runs **do not** fully represent the extremes
- The ocean model runs **do represent** well the overall climate
- Extremes can be represented using the "climate" alone, e.g.:
  - Griffiths et al. (2005), Ballester et al. (2010), Simolo et al. (2011), de Vries at al. (2012)
- So, can we model **"observed extremes" = f("simulated climate")** **???**

Bayesian Hierarchical Model

- Define **<u>observed</u> annual maxima** at all J locations as a list of vectors $\boldsymbol{Y} = \{\boldsymbol{y}_t | j = 1, 2, \ldots J\}$
- We model the extremes using a Bayesian hierarchical model (**BHM**): a model with several nested layers:

$$\underbrace{p(\boldsymbol{\theta}|\boldsymbol{Y},\boldsymbol{X})}_{\text{posterior distribution}} \propto \underbrace{p(\boldsymbol{Y}|\boldsymbol{\theta}_2)}_{\text{data layer}}\ \underbrace{p(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1,\boldsymbol{X})}_{\text{climate process layer}}\ \underbrace{p(\boldsymbol{\theta}_1)}_{\text{prior distribution}}$$

Bayesian Hierarchical Model

- Define **<u>observed</u> annual maxima** at all J locations as a list of vectors $\boldsymbol{Y} = \{\boldsymbol{y}_t | j = 1, 2, \ldots J\}$
- 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 observed annual maxima ***Y*** using the Gumbel distribution:

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

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

Bayesian Hierarchical Model

- Define **<u>observed</u> annual maxima** at all J locations as a list of vectors $\boldsymbol{Y} = \{\boldsymbol{y}_t | j = 1, 2, \ldots J\}$
- We model the extremes using a bayesian hierarchical model (**BHM**): a model with several nested layers:

$$\underbrace{p(\boldsymbol{\theta}|\boldsymbol{Y}, \boldsymbol{X})}_{\text{posterior distribution}} \propto \underbrace{p(\boldsymbol{Y}|\boldsymbol{\theta}_2)}_{\text{data layer}} \; \underbrace{p(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1, \boldsymbol{X})}_{\text{climate process layer}} \; \underbrace{p(\boldsymbol{\theta}_1)}_{\text{prior distribution}}$$

- **i. Data layer.** Model the observed annual maxima ***Y*** using the Gumbel distribution:

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

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

- **ii. Climate process layer.** The parameters of the Gumbel distribution are modeled as a function of the ocean model marine climate ***X***:

$$\boldsymbol{a} = \boldsymbol{X}\boldsymbol{\beta}_a + \boldsymbol{\epsilon}_a \qquad \longrightarrow \qquad p(\boldsymbol{a}|\boldsymbol{\beta}_a, \tau_a, \boldsymbol{X}) = \mathcal{N}_J(\boldsymbol{X}\boldsymbol{\beta}_a, \tau_a^{-1}\mathbf{I})$$

$$\boldsymbol{\phi} = \boldsymbol{X}\boldsymbol{\beta}_\phi + \boldsymbol{\epsilon}_\phi \qquad \qquad p(\boldsymbol{\phi}|\boldsymbol{\beta}_\phi, \tau_a, \boldsymbol{X}) = \mathcal{N}_J(\boldsymbol{X}\boldsymbol{\beta}_\phi, \tau_\phi^{-1}\mathbf{I})$$

  where **β**s and τs are parameters of the regression model

  – Since the models for ***a*** and ***b*** are independent, we can factor the climate process layer as

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

Bayesian Hierarchical Model

- Define **observed annual maxima** at all J locations as a list of vectors $Y = \{y_t | j = 1, 2, \ldots J\}$
- We model the extremes using a bayesian hierarchical model (**BHM**): a model with several nested layers:

$$p(\theta | Y, X) \propto p(Y | \theta_2)\, p(\theta_2 | \theta_1, X)\, p(\theta_1)$$

climate process ... ution

- i. Data layer. ... stribution:

where ϕ ... $| j = 1, 2, \ldots, J\}$

- ii. Climate pr... modeled as a function of th...

- The **marine climate statistics** (mean SST, SST variance, eddy kinetic energy, etc) are collected together into the covariate matrix ***X***, e.g.,

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

and the linear regression takes on the form:

$$\boldsymbol{a} = \beta_{a,0} + \beta_{a,1}\boldsymbol{\mu} + \beta_{a,2}\boldsymbol{\sigma}^2 + \beta_{a,3}\boldsymbol{K} + \boldsymbol{\epsilon}_a$$

$$\boldsymbol{\phi} = \beta_{\phi,0} + \beta_{\phi,1}\boldsymbol{\mu} + \beta_{\phi,2}\boldsymbol{\sigma}^2 + \beta_{\phi,3}\boldsymbol{K} + \boldsymbol{\epsilon}_\phi$$

... $\tau_a^{-1}\mathbf{I})$

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

where **β**s and τs are parameters of the regression model

– Since the models for ***a*** and ***b*** are independent, we can factor the climate process layer as

$$p(\theta_2 | \theta_1, X) = p(a | \beta_a, \tau_a, X)\, p(\phi | \beta_\phi, \tau_\phi, X)$$

Bayesian Hierarchical Model

- Define **<u>observed</u> annual maxima** at all J locations as a list of vectors $\boldsymbol{Y} = \{\boldsymbol{y}_t | j = 1, 2, \ldots J\}$
- 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

- **iii. Priors.** Assume that the parameters $\boldsymbol{\theta}_1$ are independent

$$p(\boldsymbol{\theta}_1) = p(\boldsymbol{\beta}_a)p(\boldsymbol{\beta}_\phi)p(\tau_a)p(\tau_\phi)$$

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

- Samples from the **posterior distribution** are estimated numerically using *Markov chain Monte Carlo, Metropolis rule and Gibbs sampling methods*

Bayesian Hierarchical Model

- Define **<u>observed</u> annual maxima** at all J locations as a list of vectors $\boldsymbol{Y} = \{\boldsymbol{y}_t | j = 1, 2, \ldots J\}$
- 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

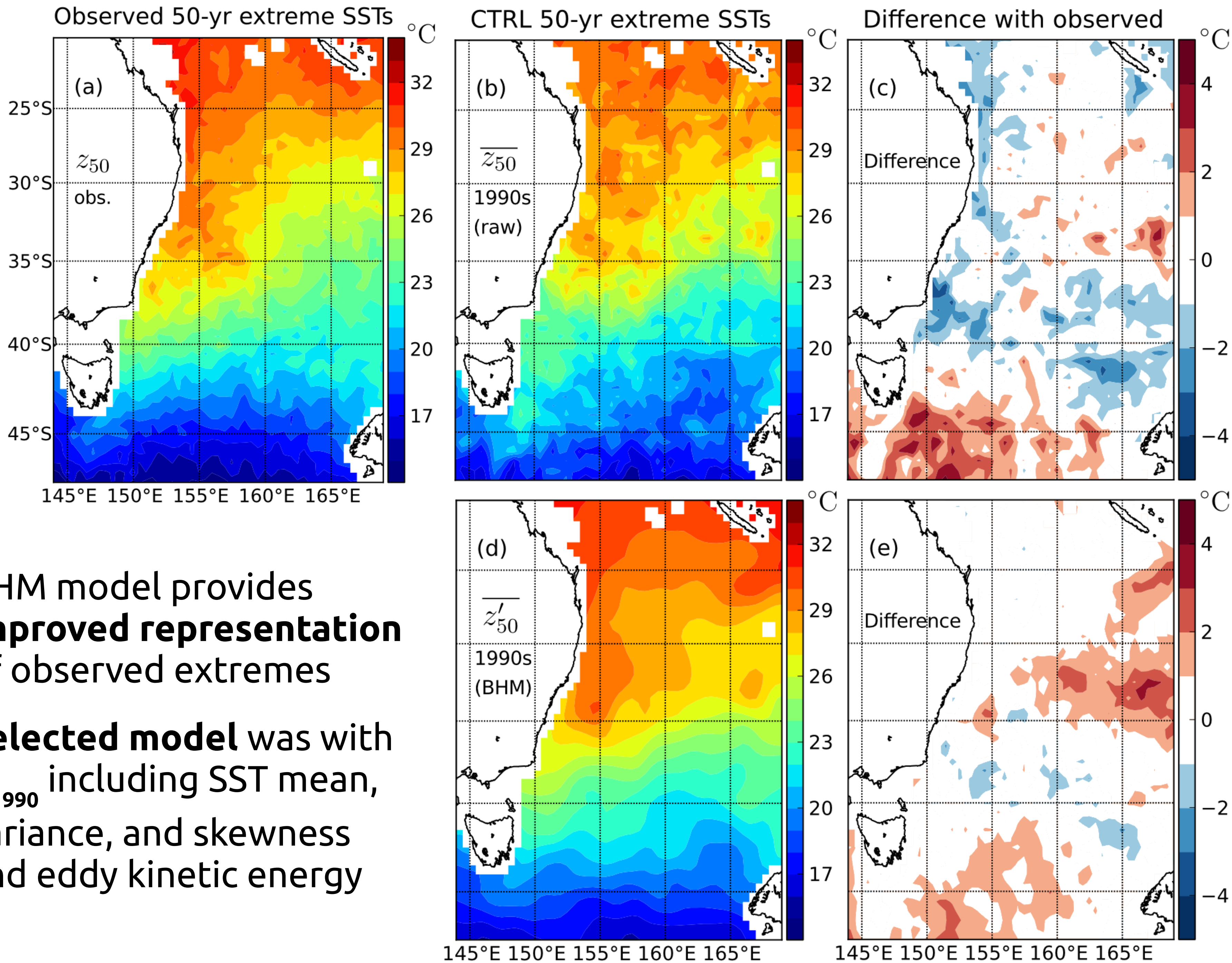
- **iii. Priors.** Assume that the parameters $\boldsymbol{\theta}_1$ are independent

$$p(\boldsymbol{\theta}_1) = p(\boldsymbol{\beta}_a)p(\boldsymbol{\beta}_\phi)p(\tau_a)p(\tau_\phi)$$

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

- Samples from the **posterior distribution** are estimated numerically using *Markov chain Monte Carlo, Metropolis rule and Gibbs sampling methods*
- **Procedure:**
  - Fit the model using the 1990s model climate ($\boldsymbol{X}_{1990}$)
  - Use fitted model and 2060s model climate ($\boldsymbol{X}_{2060}$) to estimate future extremes
  - Assumes **model stationarity**

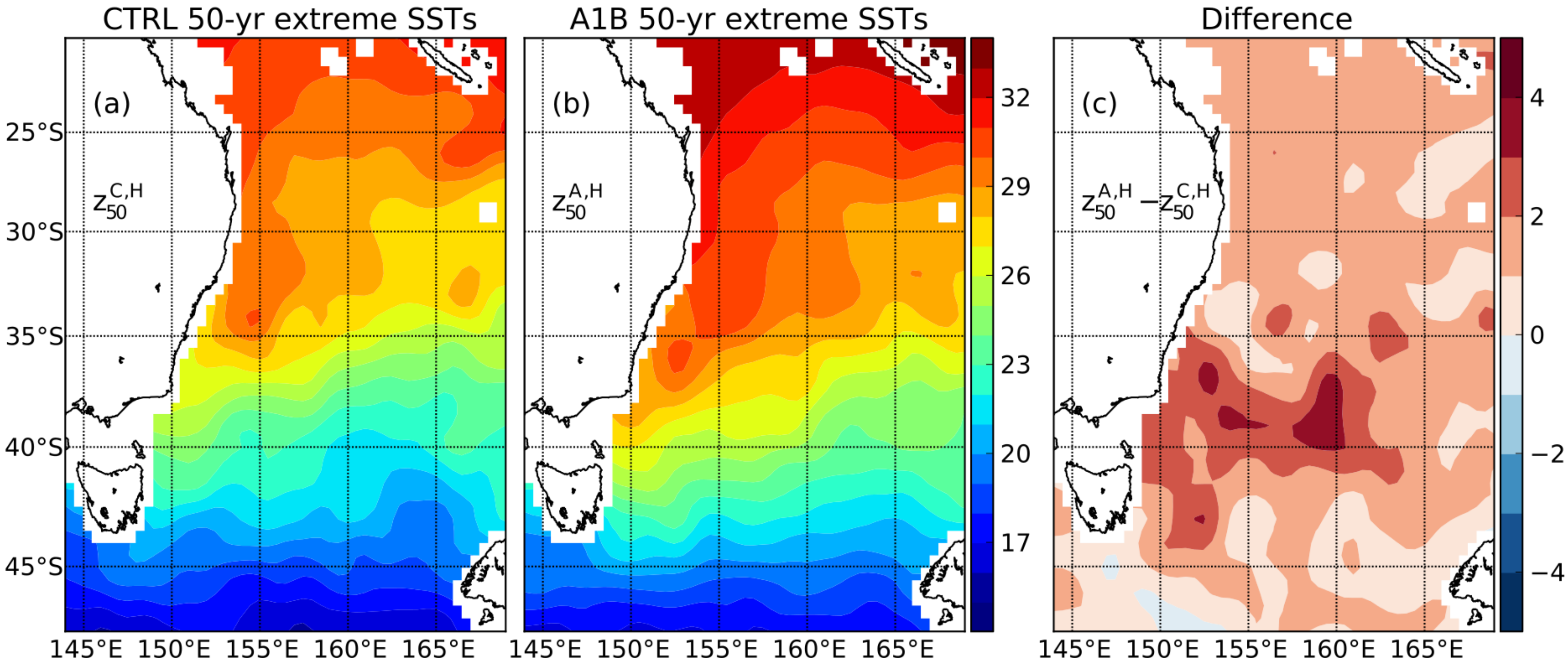
BHM Extremes: 1990s



- BHM model provides **improved representation** of observed extremes
- **Selected model** was with $\boldsymbol{X}_{1990}$ including SST mean, variance, and skewness and eddy kinetic energy

BHM Extremes: 2060s

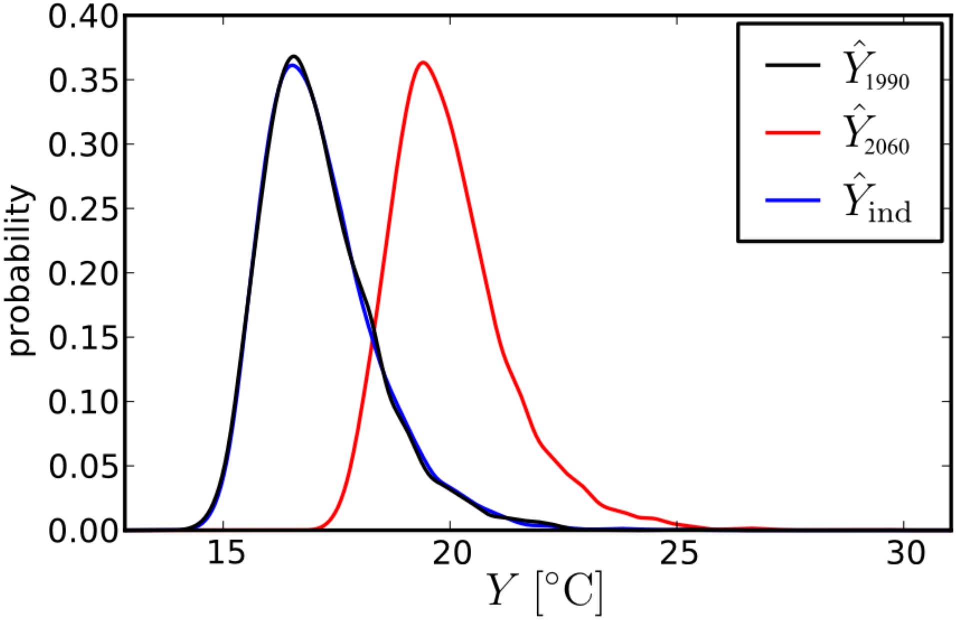
- BHM extremes model applied using fitted model parameters 2060s climate statistics as predictors ($\boldsymbol{X}_{2060}$)

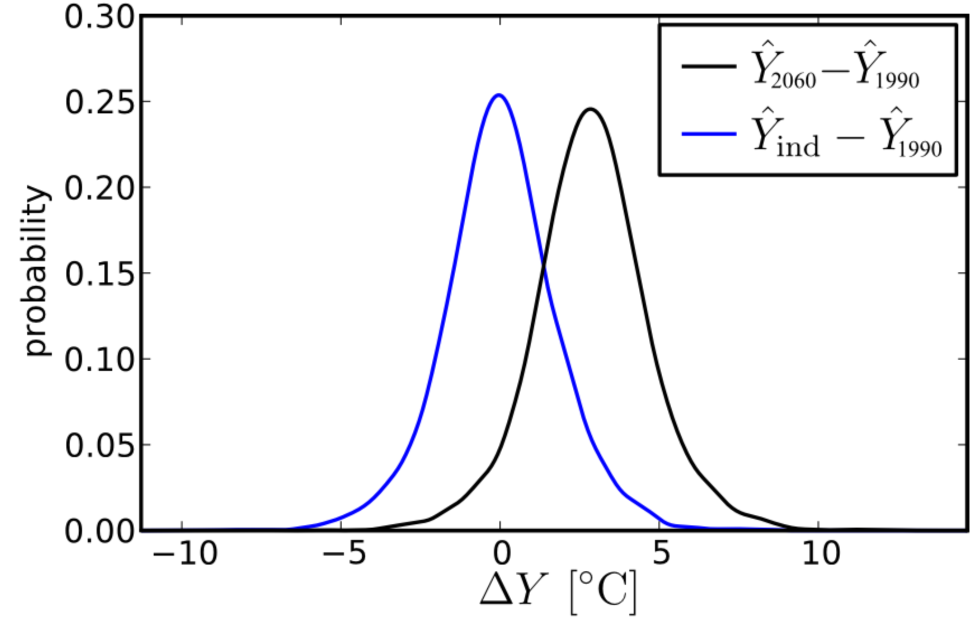


- However, the **power of the technique** allows for
  - significance testing, use as toy model, etc...

Confidence

- The extremes model is **probabilistic** in nature (Bayesian) and so we can put **confidence limits** on our projections
- This type of information is very helpful when making statements about climate change
- By **drawing independent samples** of ***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 (i.e., due to random variations)

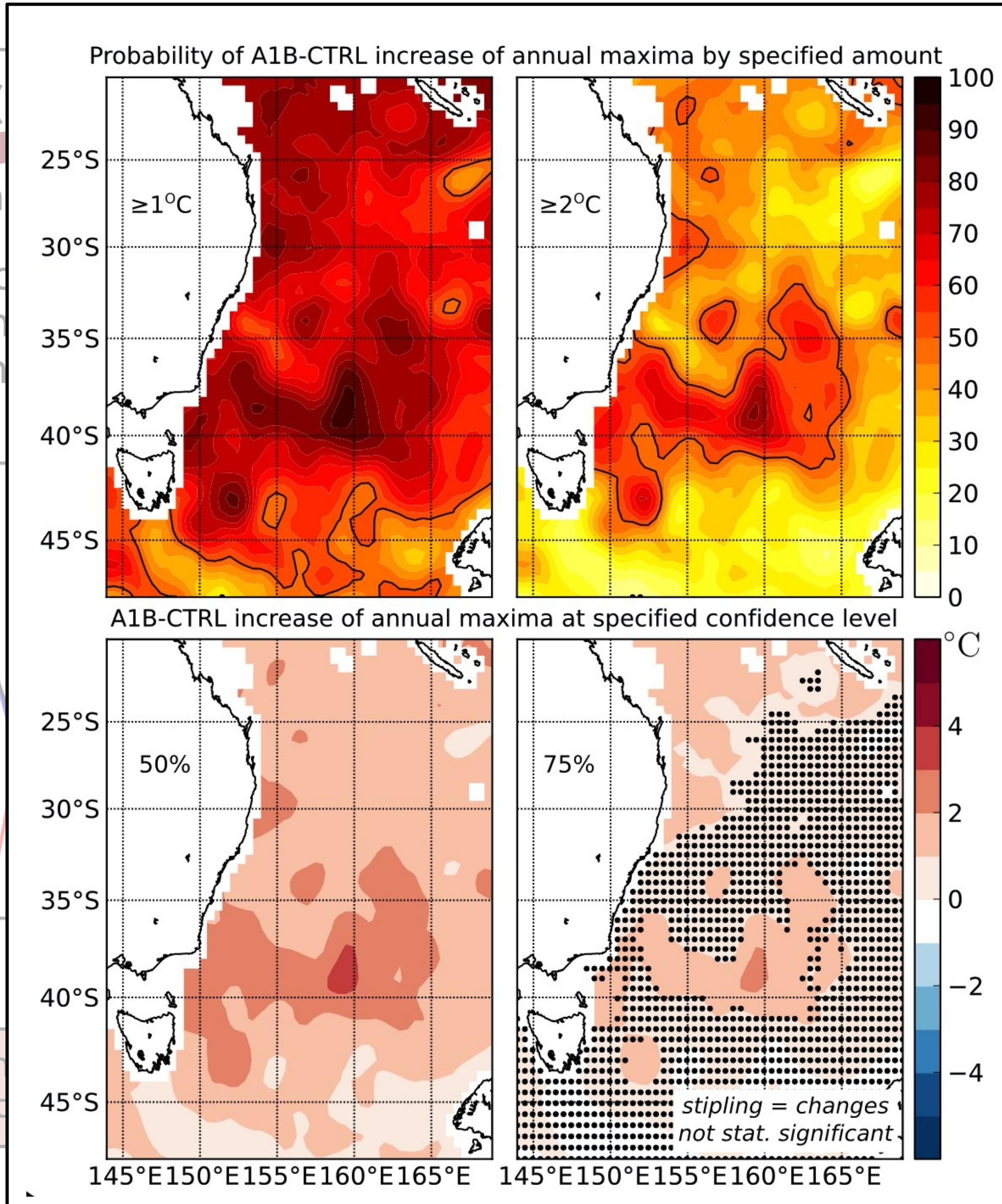




East of Tasmania (152°E, 43°S) with 75% confidence a projected change of 1.78°C is significant (compared to randomness in an unchanged climate) but with 90% confidence a a change of 0.67°C is not significant.

Confidence

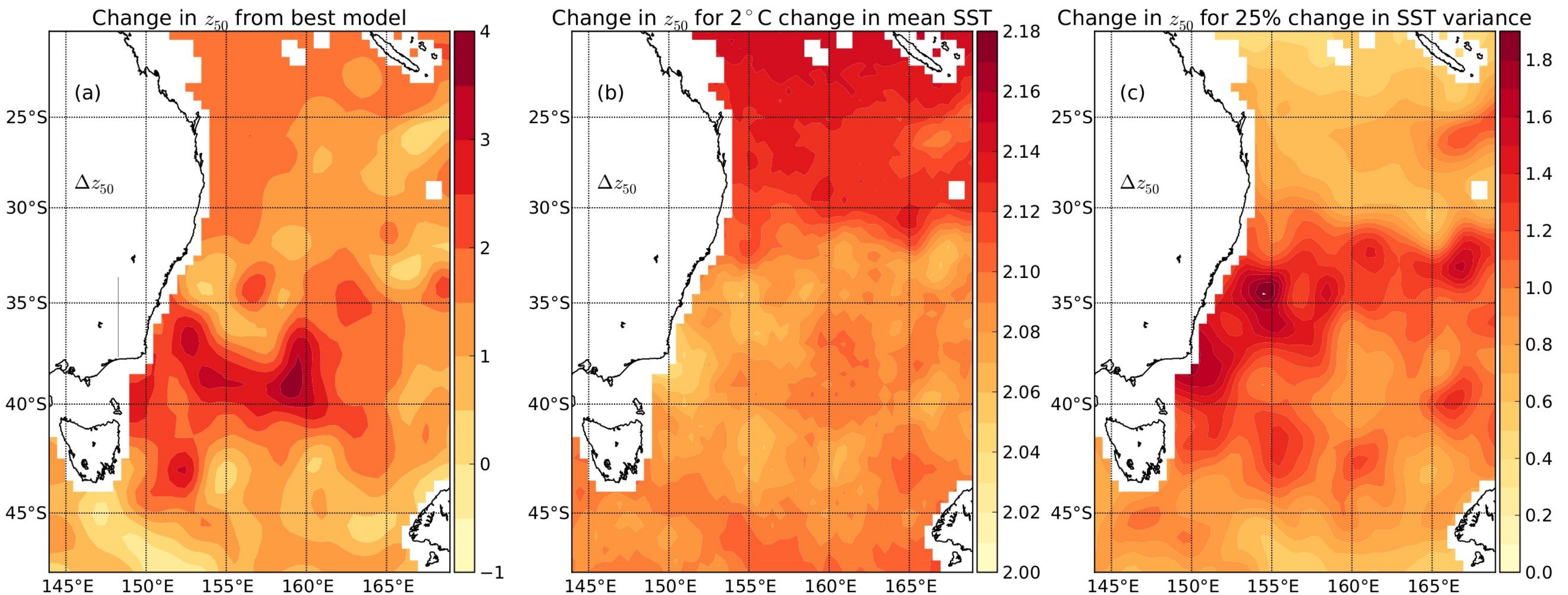
- The extren … ndent samples
  in nature (… rior for the
  put confid … the projected
  projection … is statistically … mpared to the … n unchanged … ue to random
- This type c
  helpful wh
  about clim



Off eastern Tasman … al maxima will be positive,
a ~70% chance that … e will exceed 2°C

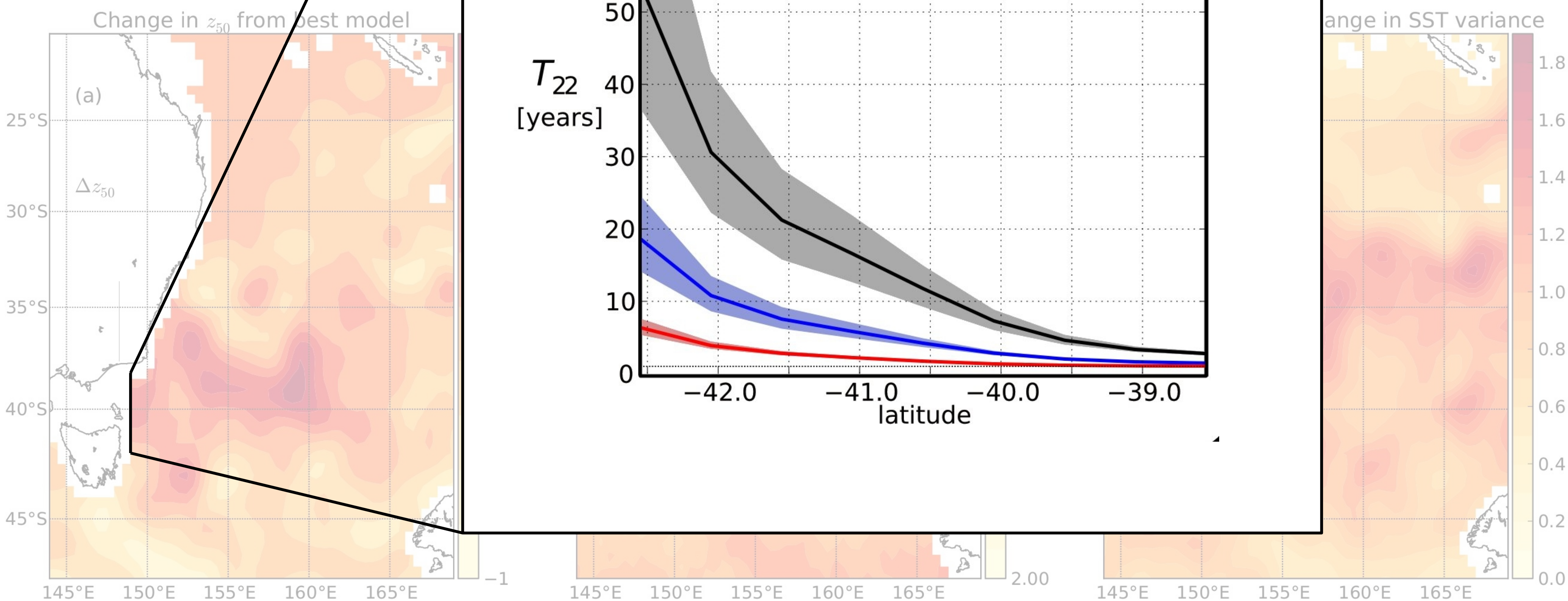
Toy Model

- Use the extremes model as a **"toy model"** to test the response of the extremes to **specified changes in climate**:
  - We can specify a particular climate ($\boldsymbol{X}_{spec}$), 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



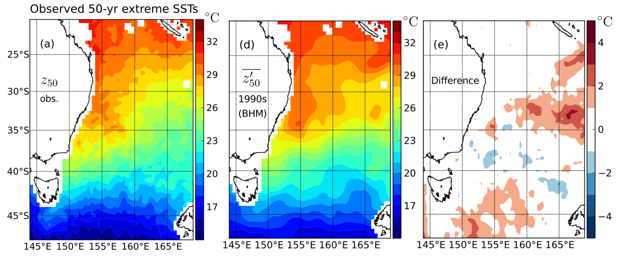
Toy Model

- Use the extremes model as a **“toy model”** to test the response of the extremes to **specified changes in c**
  - We can specify a p   s a 2°C warming of the me
  - Then drawing from   te we can test what the resp   he climate



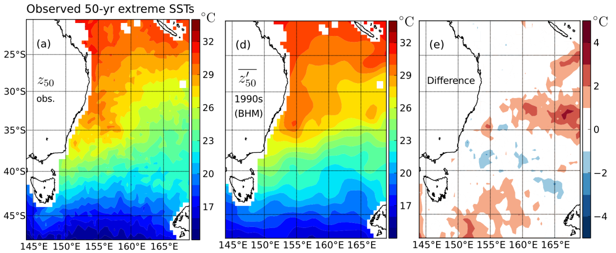
Summary

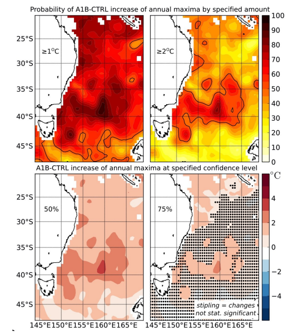
- The BHM extreme model provides a technique to **improve estimates of extremes** from global ocean and climate models – can be applied to other variables (precipitation?)



Summary

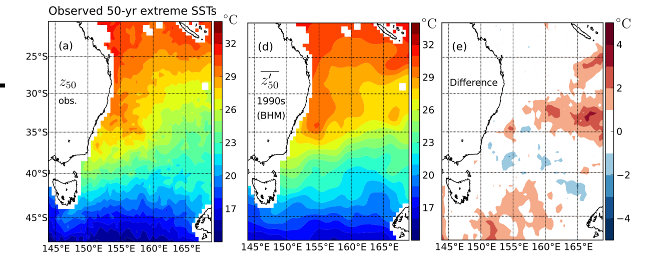
- The BHM extreme model provides a technique to **improve estimates of extremes** from global ocean and climate models – can be applied to other variables (precipitation?)
- The basic approach is to **model observed extremes as a function of the historical climate statistics**, assume stationarity, and then use climate projections or specified climates to estimate extremes for other climate scenarios

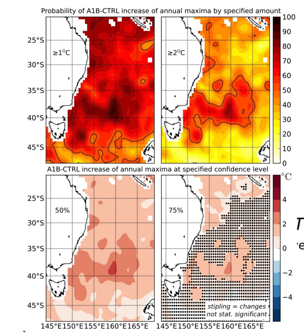


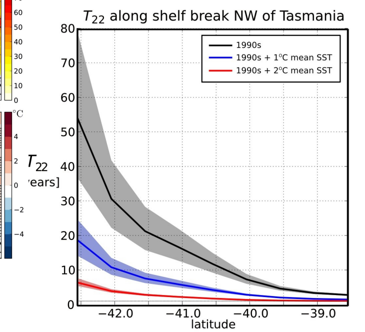


Summary

- The BHM extreme model provides a technique to **improve estimates of extremes** from global ocean and climate models – can be applied to other variables (precipitation?)
- 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

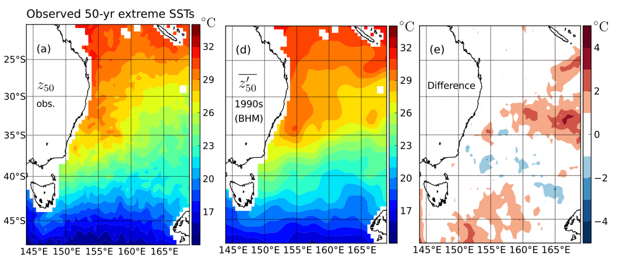


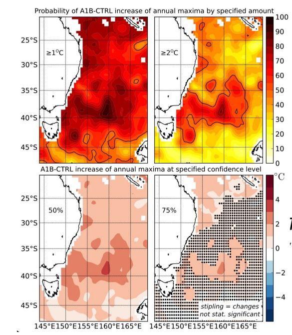


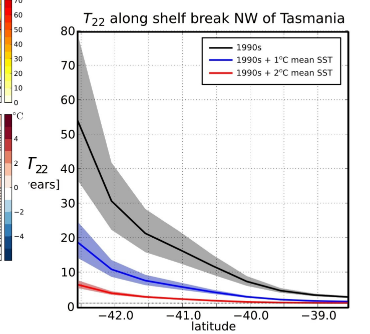


Summary

- The BHM extreme model provides a technique to **improve estimates of extremes** from global ocean and climate models – can be applied to other variables (precipitation?)
- 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







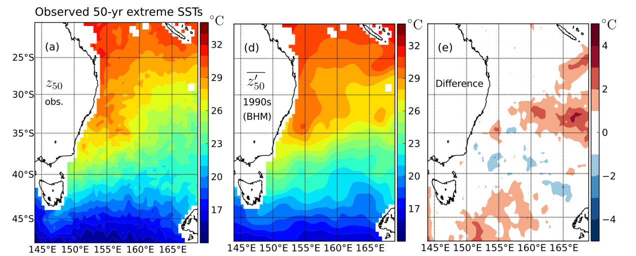
Publications

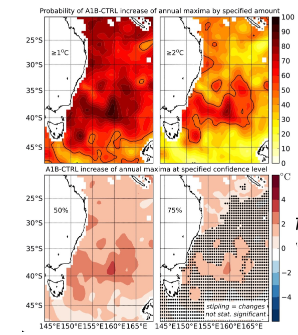
Oliver, E. C. J., S. J. Wotherspoon and N. J. Holbrook (2014), Estimating extremes from global ocean and climate models: A Bayesian hierarchical model approach, *Progress in Oceanography*, 122, 77-91

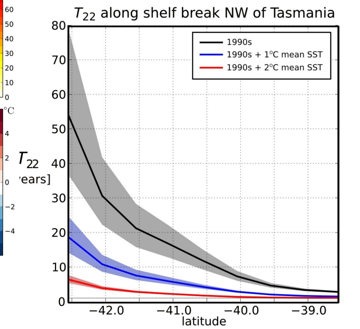
Oliver, E. C. J., S. J. Wotherspoon, M. A. Chamberlain and N. J. Holbrook (2014), Projected Tasman Sea extremes in sea surface temperature through the 21st century, *Journal of Climate*, 27(5), 1980-1998

Summary

- The BHM extreme model provides a technique to **improve estimates of extremes** from global ocean and climate models – can be applied to other variables (precipitation?)
- 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







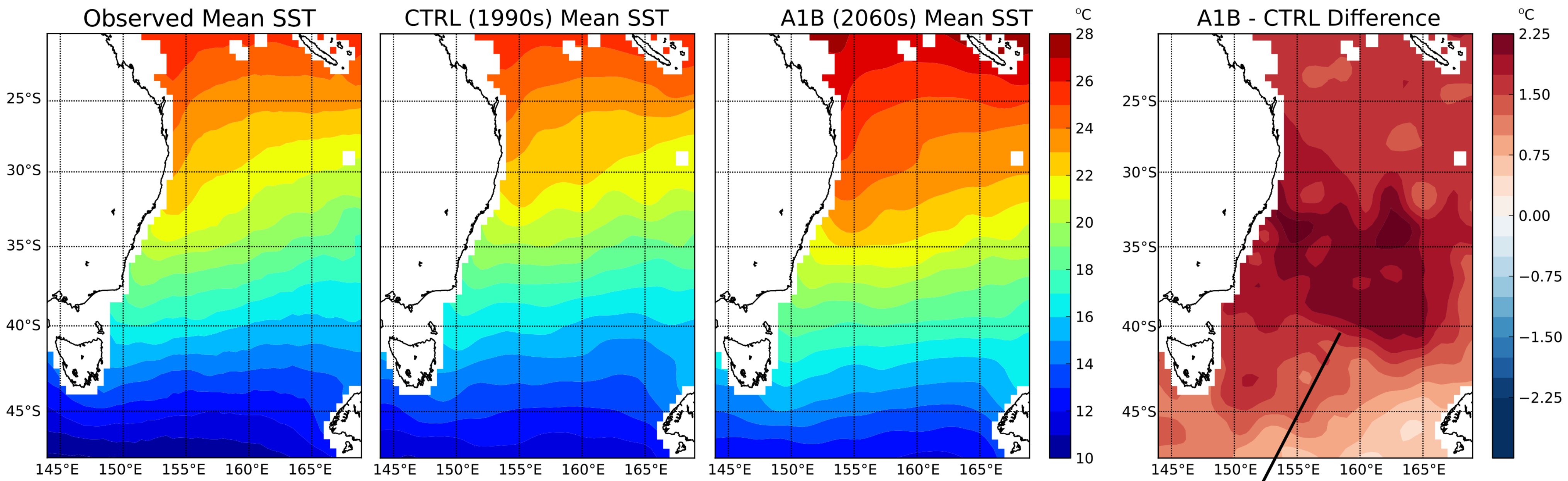
Publications

Oliver, E. C. J., S. J. Wotherspoon and N. J. Holbrook (2014), Estimating extremes from global ocean and climate models: A Bayesian hierarchical model approach, *Progress in Oceanography*, 122, 77-91

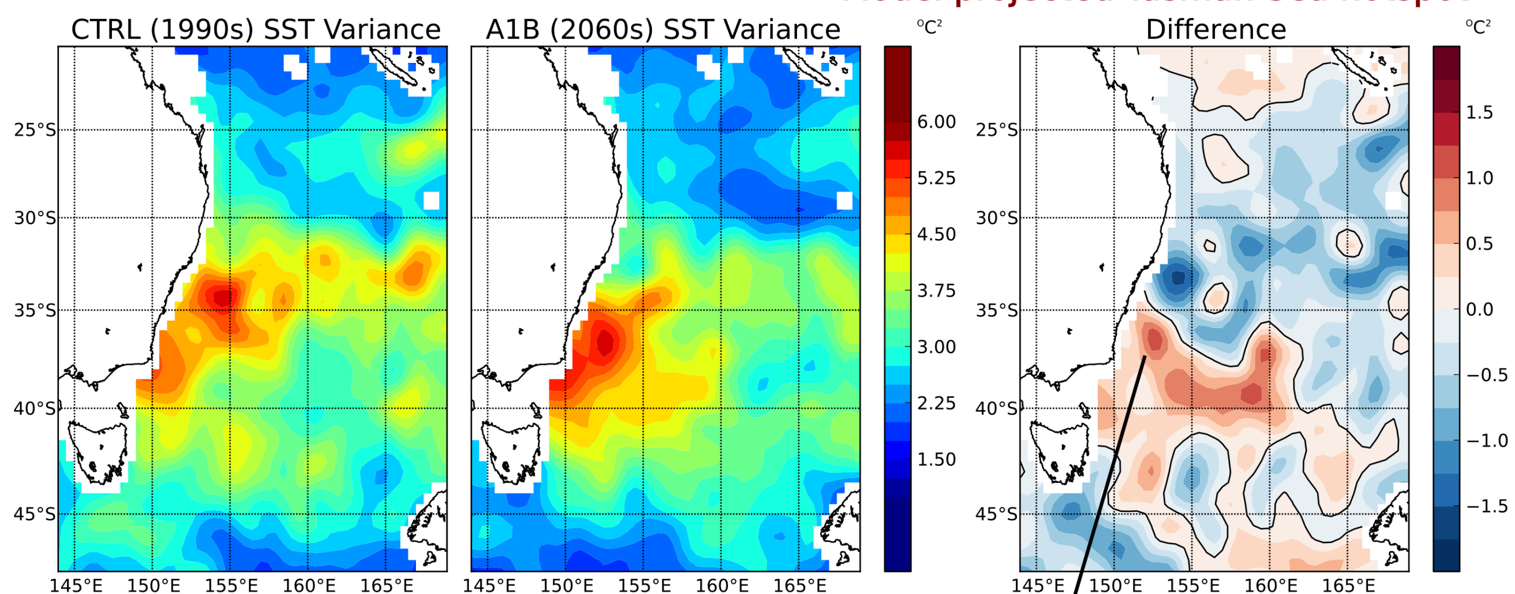
Oliver, E. C. J., S. J. Wotherspoon, M. A. Chamberlain and N. J. Holbrook (2014), Projected Tasman Sea extremes in sea surface temperature through the 21st century, *Journal of Climate*, 27(5), 1980-1998

Acknowledgements *Neil Holbrook* (IMAS, post-doc supervisor), *Simon Wotherspoon* (IMAS, statistics mathemagician), *Matthew Chamberlain*, *Richard Matear* (CSIRO, OFAM modelling). The *Super Science Fellowship* and the *Centre of Excellence for Climate System Science* (Australia Research Council), *IMAS* and *UTAS* for support, financial and otherwise

SST Mean and Variance

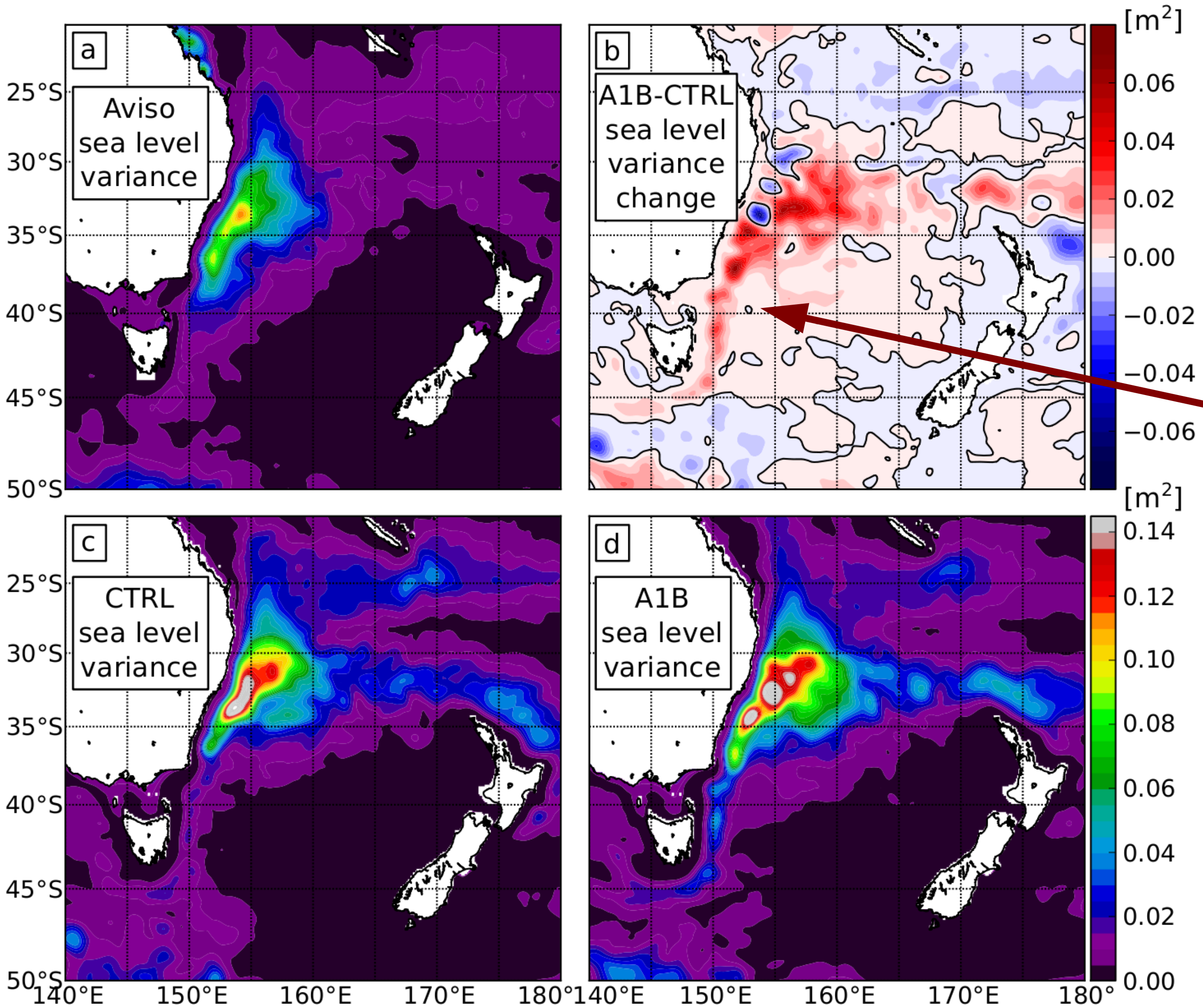


Model projected Tasman Sea hotspot



Increase of SST variability
in EAC extension region

Eddy Kinetic Energy



- Sea level variance (~eddy kinetic energy) consistent between model and observations
- Significant increase in eddy kinetic energy in EAC Extension region, where flow is not steady but in fact consists of a train of mesoscale eddies…

Model Stationarity

Fundamental relationship

We posit that there exists a relationship between the extremes and climate parameters $\mathbf{X}$:

$$\text{“extremes“} = f(\mathbf{X})$$

This relationship expresses fundamental aspects of the climate system which do not change with time.

Role of βs and τs

Effectively, we have performed a linear approximation to $f(\mathbf{X})$:

$$f(\mathbf{X}) = \mathbf{X}\beta + O(\mathbf{X}^2)$$

Therefore, the βs (and τs) are stationary since $f(\mathbf{X})$ is stationary