A statistical method for improving continental shelf and near-shore marine climate predictions

Eric C. J. Oliver and Neil J. Holbrook

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

1 Introduction

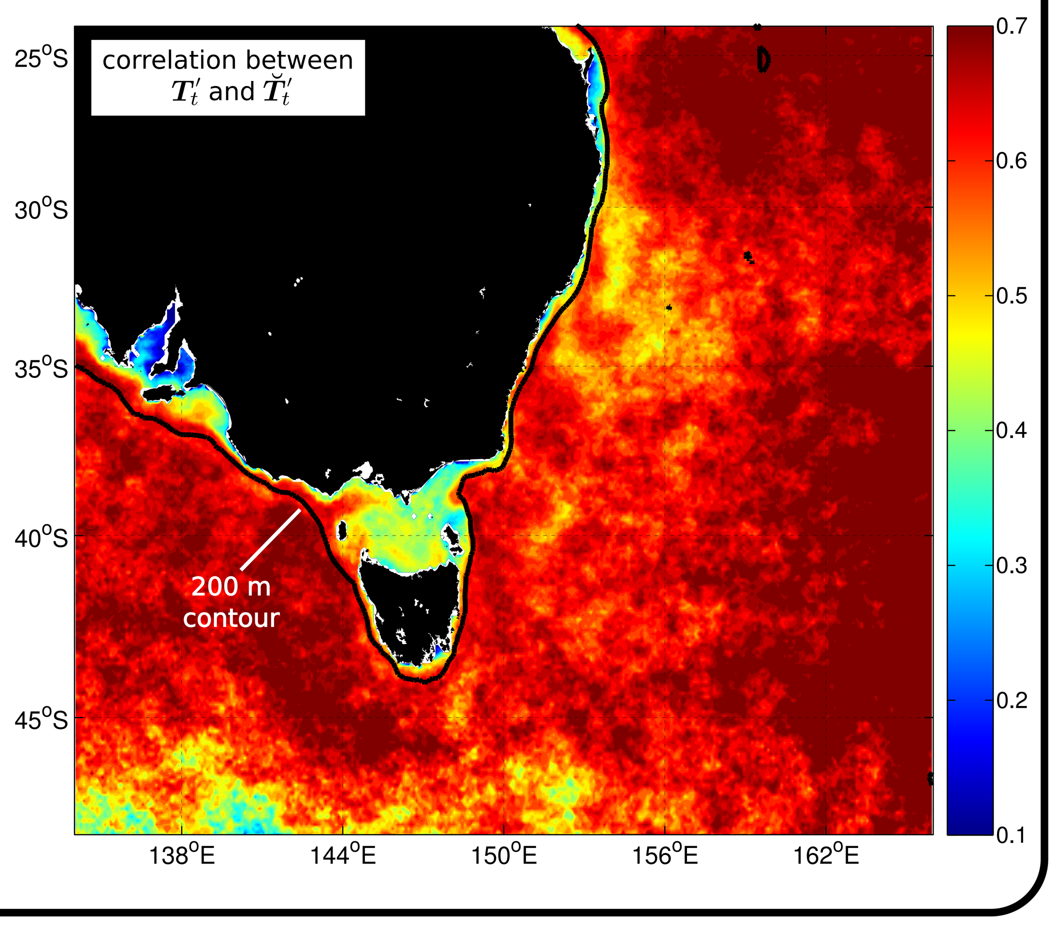
- There is a scarcity of spatially and temporally homogenous measurements of **ocean variability** on the Australian continental shelf.
- The ocean reanalysis product **Bluelink ReANalysis (BRAN)** provides estimates of ocean variability around Australia at 1/10 degree resolution. BRAN reproduces the large-scale patterns of sea surface temperature (SST) in deep water[1,2], such as those associated with the East Australian Current and the Leeuwin Current, but performs poorly over the **continental shelf**.
- We have developed a **linear statistical model** to more accurately estimate in-shore SST using off-shore SST from BRAN. SST variability is separated into the mean, seasonal cycle, and the residual variability and separate models are developed for each component.
- Model performance is demonstrated at a point location in Bass Strait and then it is extended onto the continental shelf around **southeastern Australia**.

2 Data and Problem

Observed and reanalysed SST:

- Daily 4 km fields of **observed SST** were obtained from the Advanced Very High Resolution Radiometer (AVHRR) for the period 14/10/1992 to 13/5/2008. Denoted by $\boldsymbol{T}_t$.
- Daily ~0.1 degree fields of **reanalysed SST** were obtained from the Bluelink ReANalysis (BRAN) for the period 14/10/1992 to 13/5/2008. Denoted by $\breve{\boldsymbol{T}}_t$.

Sea surface temperature (SST) is poorly represented over much of the continental shelf (see **correlation** between $\boldsymbol{T}_t$ and $\breve{\boldsymbol{T}}_t$ to the right) including coastal South Australia, the Bass Strait, and parts of coastal New South Wales and Queensland.



3 Statistical Model

Let $\boldsymbol{T}_t$ denote observed SST and $\breve{\boldsymbol{T}}_t$ denote SST output from BRAN where t is a time index. Both $\boldsymbol{T}_t$ and $\breve{\boldsymbol{T}}_t$ are $m \times n$ matrices where m and n are the number of grid points in the latitudinal and longitudinal directions respectively. Consider SST time series' at two locations: an **off-shore** location $(Y_t, \breve{Y}_t)$ and an **in-shore** location $(X_t, \breve{X}_t)$. Assume the time series' can be written as

$$X_t = \underbrace{\bar{X}}_{\text{mean}} + \underbrace{X_t^{\mathrm{S}}}_{\text{seasonal cycle}} + \underbrace{X_t'}_{\text{residual}}$$

In-shore SST will be **predicted** ($\hat{Y}_t$) from off-shore SST. Separate models for each component are developed and model parameters are trained on $\boldsymbol{T}_t$ and then informed by $\breve{\boldsymbol{T}}_t$.

(i) **Mean**: The time-mean values can be related as

$$\bar{Y} = a\bar{X} \longrightarrow \hat{\bar{Y}} = a\bar{\breve{X}}$$

(ii) **Seasonal Cycle**: write as a sum of harmonics

$$X_t^{\mathrm{S}} = \sum_{k=1}^{K} A_k^{\mathrm{X}} \cos(\omega_k t - \phi_k^{\mathrm{X}})$$

and the parameters can be related linearly

$$\left.\begin{array}{l} A_k^{\mathrm{Y}} = \gamma_k A_k^{\mathrm{X}} \\ \phi_k^{\mathrm{Y}} = \phi_k^{\mathrm{X}} + \Delta_k \end{array}\right\} \quad \text{for} \quad k = 1, \ldots, K$$

$$\downarrow$$

$$\hat{Y}_t^{\mathrm{S}} = \sum_{k=1}^{K} \gamma_k A_k^{\breve{\mathrm{X}}} \cos\left[\omega_k t - (\phi_k^{\breve{\mathrm{X}}} + \Delta_k)\right]$$

(iii) **Residual**: model using linear regression

$$Y_t' = \beta X_t' + \epsilon_t$$

with the predicted residual given by

$$\hat{Y}' = \hat{\beta} \breve{X}_t'$$

The choice of off-shore **predictor location** $(i_{\mathrm{Y}}, j_{\mathrm{Y}})$ is given by the minimum of a cost function $\boldsymbol{J}$:

$$J_{ij} = \frac{1}{\underbrace{R_{ij}}_{\text{localisation matrix}} \underbrace{S_{ij}}_{\text{strength matrix}} \underbrace{Q_{ij}}_{\text{quality matrix}}} \qquad (0 < J_{ij} < 1)$$

The $\boldsymbol{R}$, $\boldsymbol{S}$, and $\boldsymbol{Q}$ matrices are defined as follows:

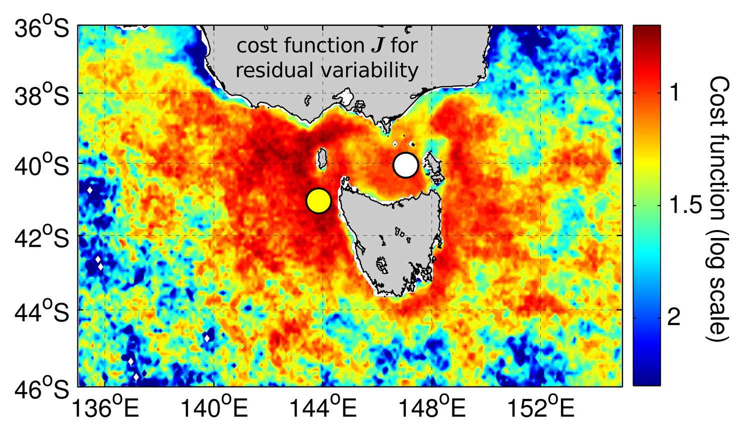
- $R_{ij} = \exp\left\{-\left[(i - i_{\mathrm{Y}})^2 + (j - j_{\mathrm{Y}})^2\right] / 2\sigma^2\right\}$ — simple 2D Gaussian
- $S_{ij} = 1 - \left|\dfrac{\bar{T}_{ij} - \bar{T}_{i_{\mathrm{Y}}j_{\mathrm{Y}}}}{\bar{T}_{i_{\mathrm{Y}}j_{\mathrm{Y}}}}\right|$, $S_{ij} = \rho(T_{ijt}', T_{i_{\mathrm{Y}}j_{\mathrm{Y}}t}')$ — measures strength of connection between in-shore and off-shore SST
- $Q_{ij} = 1 - \left|\dfrac{\bar{\breve{T}}_{ij} - \bar{T}_{ij}}{\bar{T}_{ij}}\right|$, $Q_{ij} = \rho(T_{ijt}', \breve{T}_{ijt}')$ — measures quality of off-shore SST predictor

The **cost function** decreases with (i) proximity and (ii) connection strength of predictor and predictand and (iii) data quality at the possible predictor location.

4 A point location in Bass Strait

Statistical model applied to SST at (147°E,40°S) (○) in Bass Strait.

Predictor location (●) is minimum of $\boldsymbol{J}$.



Results:

mean

	$\bar{Y}$	$\bar{\breve{Y}}$	$\hat{\bar{Y}}$
	15.1	14.7	15.2

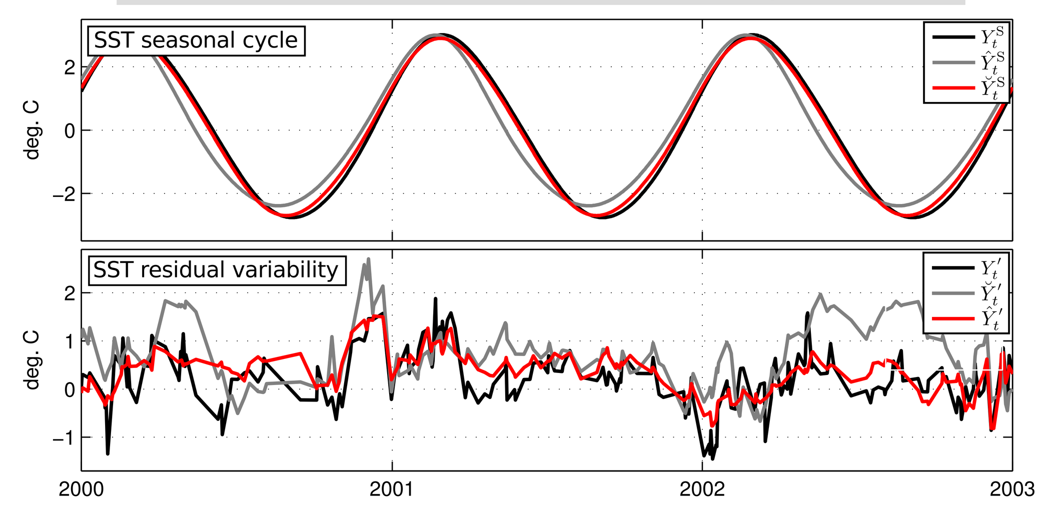
seasonal cycle

	$\breve{Y}_t^{\mathrm{S}}$	$\hat{Y}_t^{\mathrm{S}}$	
Y_t^{S}	0.97	0.99	corr.
	0.50	0.16	RMSE

residual

	$\breve{Y}_t'$	$\hat{Y}_t'$	
Y_t'	0.49	0.69	corr.
	1.01	0.20	RMSE

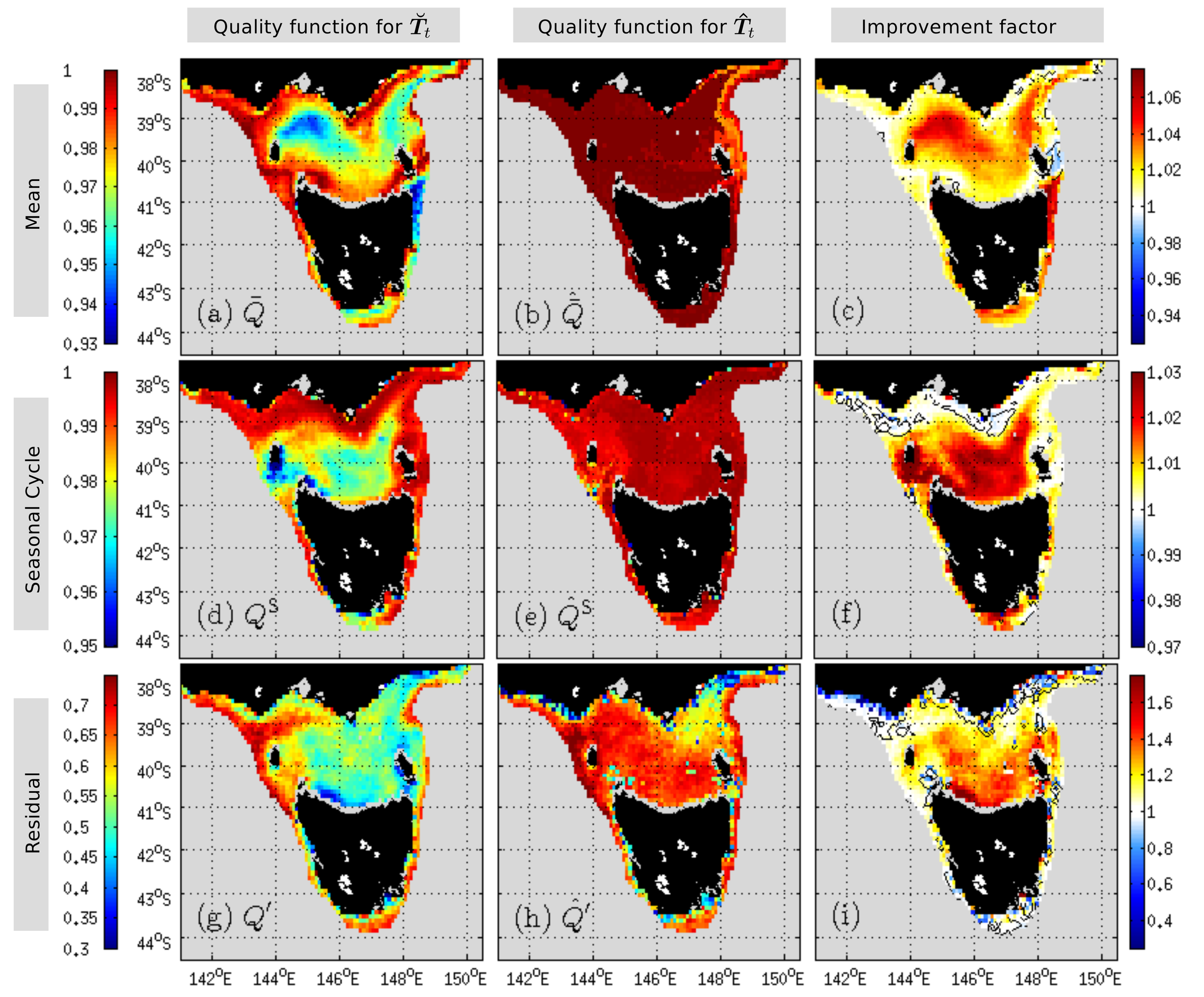
time series: observed (—), BRAN (—), statistical prediction (—)



The statistical model provides **better estimates** of SST variability in Bass Strait than those provided by BRAN.

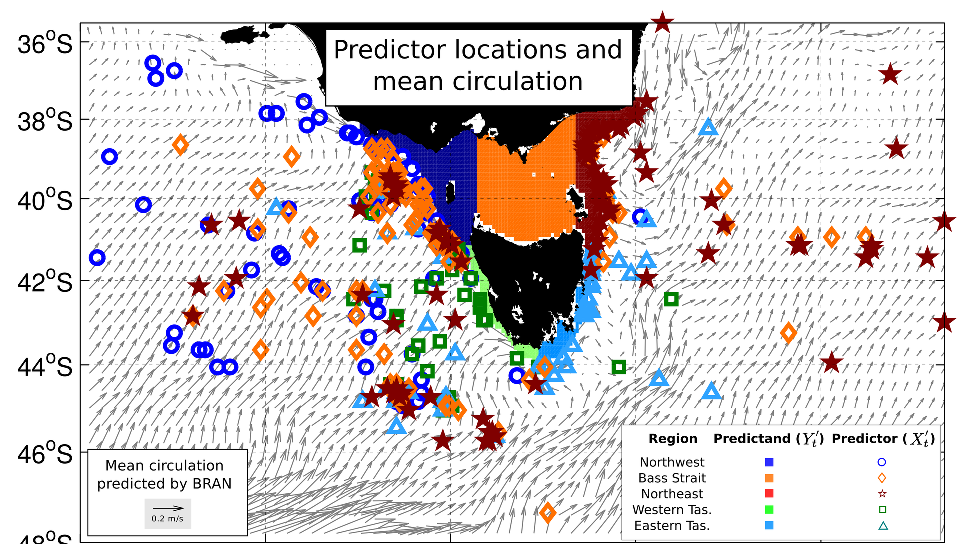
5 The Continental Shelf

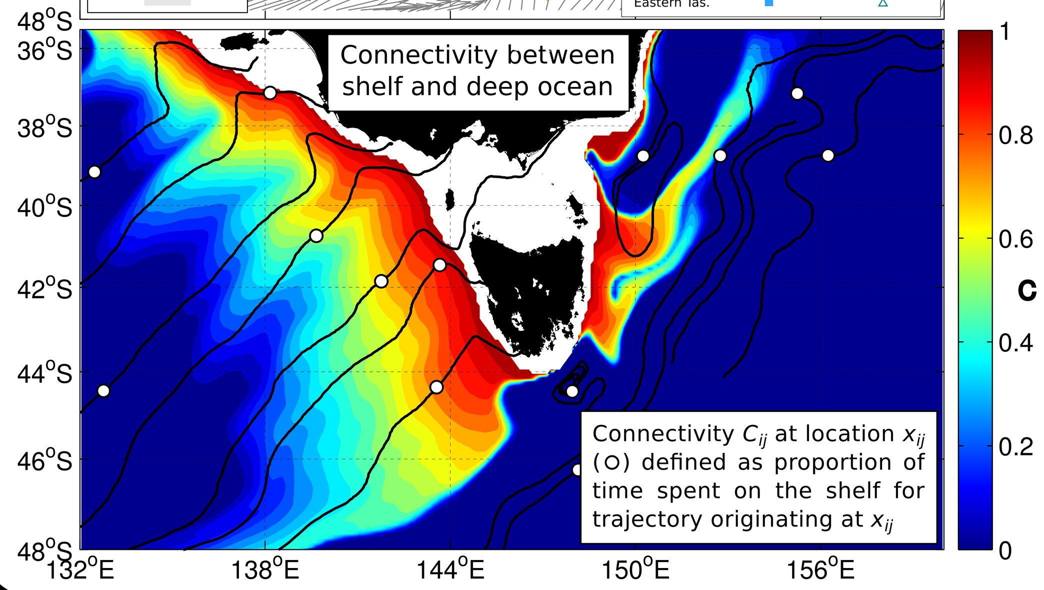
- The statistical downscaling model was **systematically** applied to each location with water depth less than 200 m off the coasts of **Victoria and Tasmania**.
- The cost function was masked for all locations in water depth less than 200 m to exclude those locations as potential predictors.



In general, the statistical model provides much better estimates of SST on the **continental shelf**.

6 Influence of the Mean Circulation



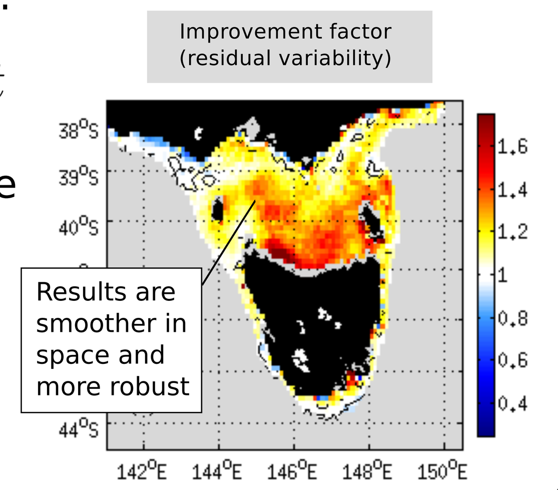


- The shelf is divided into **five regions** and colour coded along with the associated off-shore predictor locations
- Predictors tend to be located within regions of **strong mean flow**.
- Over 85% of the predictors are located **west** of the southern tip of Tasmania.
- The influence of the **South Australian and Zeehan Currents**[3] and Ekman transport appear to connect the shelf to the off-shore.
- Connectivity between the deep ocean and the shelf quantified by calculating Lagrangian trajectories:

$$\vec{x}(t + \Delta t) = \vec{x}(t) + \vec{u}(\vec{x}(t))\Delta t$$

- Cost matrix updated to include connectivity matrix **C**:

$$J_{ij} = \frac{1}{R_{ij} S_{ij} Q_{ij} \underbrace{C_{ij}}_{\text{connectivity between shelf and deep ocean}}}$$



7 Conclusions

- The statistical model presented here provides a **systematic method** of estimating SST variability on the **continental shelf** across a range of time scales in southeastern Australia. The **hybrid statistical-connectivity model** shows promise for **more accurate** and useful estimates of inshore SST.
- Future work will include downscaling projected model **simulations** representing the 1990s and **2060's** onto the shelf including means, variances, seasonal cycles and **extremes**.

8 References

1 Oke P. et al., 2008: The Bluelink ocean data assimilation system (BODAS), Ocean Modelling, **21** (1-2), 46-70.

2 Schiller A. et al., 2008: Eddy-resolving ocean circulation in the Asian-Australian region inferred from an ocean reanalysis effort, Progress in Oceanography, **76** (3), 334-365.

3 Ridgway K.R. and Condie S.A., 2004: The 5500-km-long boundary flow off western and southern Australia, J. Geophysical Research, **109** (C04017), 1-18.

Oliver E.C.J. and Holbrook N.J.: A statistical method for improving continental shelf and near-shore marine climate predictions, J. Atmos. Ocean. Tech., *subm.*