Improving global ocean heat content uncertainties by modeling vertical spatio-temporal dependence

Thea Sukianto[1], **Mikael Kuusela**[1], **and Donata Giglio**[2]

[1]Carnegie Mellon University, [2]University of Colorado Boulder

Introduction

- Obtaining reliable uncertainties for ocean heat content (OHC) is crucial for tracking climate change ($\sim$ 91% of excess heat in the Earth system is stored in the ocean)
- Estimating OHC and its uncertainties from Argo temperature profiles presents exciting statistical and computational challenges from nonstationarity and large dataset size
- Due to data availability, we split the analysis into two vertical sections (top: 15–975 m and bottom: 975–1850 m), which produces a conservative uncertainty estimate:

$$\mathrm{Var}(\mathrm{OHC}_{\mathrm{total}}|\mathrm{data}) \le \left(\sqrt{\mathrm{Var}(\mathrm{OHC}_{\mathrm{top}}|\mathrm{data})} + \sqrt{\mathrm{Var}(\mathrm{OHC}_{\mathrm{bot}}|\mathrm{data})}\right)^2$$

- Here we improve the uncertainty estimates by extending the locally stationary spatio-temporal interpolation method (Kuusela and Stein, 2018) to model the vertical dependence

Mapping methodology

- **Quantity of interest**

$$\mathrm{OHC}(t) = \rho_0 c_{p,0} \iiint T(x,y,z,t)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \rho_0 c_{p,0} \iint \Big(\underbrace{\int_{z_{\min}}^{z_{\max}} T(x,y,z,t)\,\mathrm{d}z}_{:=\tilde{T}(x,y,t)} \Big)\,\mathrm{d}x\,\mathrm{d}y,$$

where $T(x,y,z,t)$ is the potential temperature at space-time location (x,y,z,t)

We map the mean-subtracted temperature residual $\tilde{R}(x,y,t)$ in two vertical sections jointly as realizations from a bivariate Gaussian process (GP)

- **Bivariate locally stationary GP regression**

Let $s = (x,y,t)$ and i = year, j = observation. Then we model the temperature residuals as:

$$\begin{bmatrix} \tilde{R}_{\mathrm{top}} \\ \tilde{R}_{\mathrm{bot}} \end{bmatrix}_{i,j} = f_i\left(\begin{bmatrix} s_{\mathrm{top}} \\ s_{\mathrm{bot}} \end{bmatrix}_{i,j}\right) + \begin{bmatrix} \epsilon_{\mathrm{top}} \\ \epsilon_{\mathrm{bot}} \end{bmatrix}_{i,j}$$

$$f_i \overset{\mathrm{iid}}{\sim} \mathrm{GP}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, K(s_1, s_2; \theta)\right), \quad \begin{bmatrix} \epsilon_{\mathrm{top}} \\ \epsilon_{\mathrm{bot}} \end{bmatrix}_{i,j} \overset{\mathrm{iid}}{\sim} \mathrm{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma(\theta)\right)$$

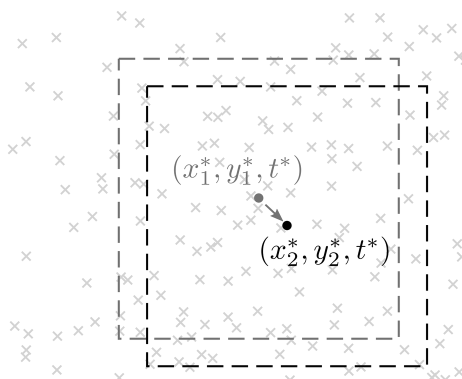
with cross-covariance function (Kleiber and Nychka, 2012)

$$K_{\mathrm{top,bot}}(s_1, s_2; \theta) = \beta \frac{\delta_{\mathrm{top}}\delta_{\mathrm{bot}}}{\sqrt{|\Theta_{\mathrm{top,bot}}|}} \exp\left(-\sqrt{(s_1 - s_2)^\top \Theta_{\mathrm{top,bot}}^{-1} (s_1 - s_2)}\right)$$

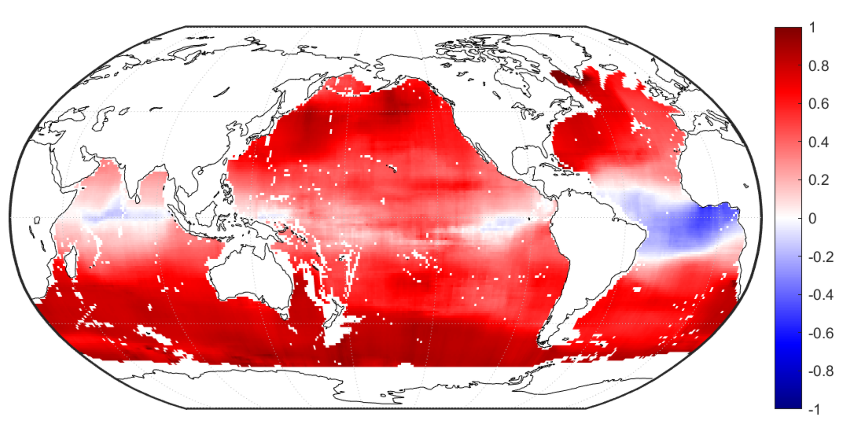
$\Rightarrow$ *β is an estimate of the dependence between vertical sections!*

Parameter estimation and kriging

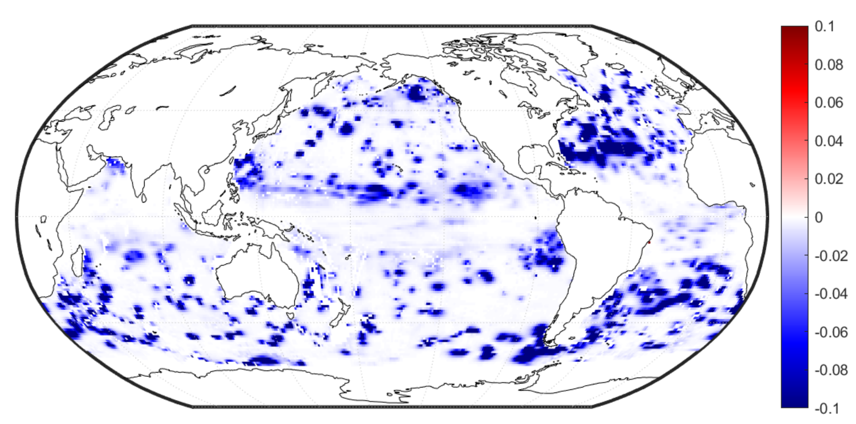
- We focus on data in a small neighborhood around a spatio-temporal grid point $s^* = (x^*, y^*, t^*)$ and regard each year as an i.i.d. replicate from the above stationary Gaussian process within the neighborhood



- We compute the MLE of the covariance parameters and the kriging prediction at s^* for overlapping moving windows (Haas, 1990, 1995), which allows for spatially-varying parameter estimates and predictions



Estimated β (cross-correlation) from 2004-2022 Argo data

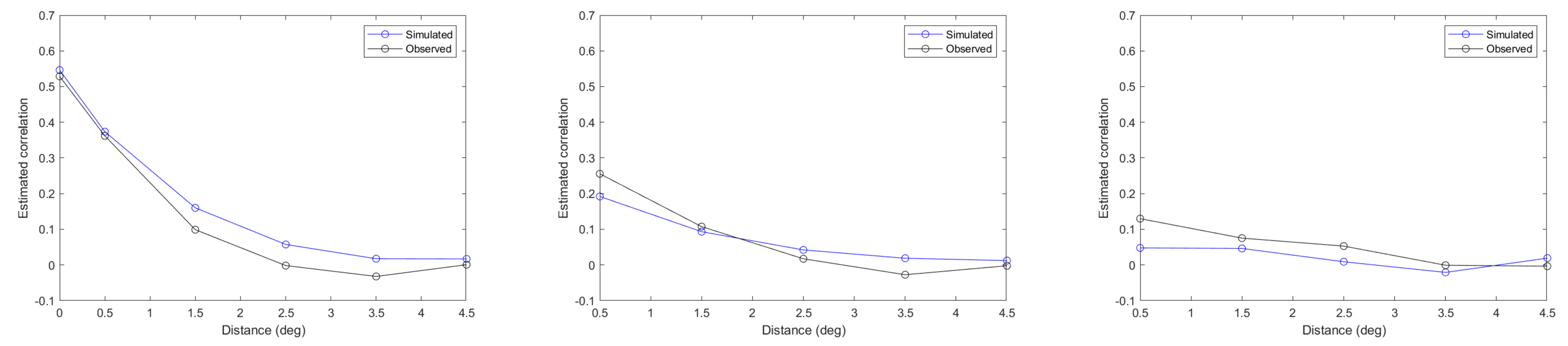


Relative kriging variance difference between bivariate & univariate models (975–1850 m)

$\Rightarrow$ Since the sparser bottom layer is able to borrow strength from the top, we see that the bivariate model provides a 10%+ reduction in predictive uncertainty

Uncertainty quantification with local conditional simulations

- Getting uncertainties for $\mathrm{OHC}(t) = \rho_0 c_{p,0} \int\int \tilde{T}(x,y,t)\,\mathrm{d}x\,\mathrm{d}y$ requires taking into account the predictive covariance between $\tilde{T}(x_1,y_1,t_1)$ and $\tilde{T}(x_2,y_2,t_2)$
- We do this by generating ensembles of conditional simulation realizations from the predictive distribution $p(\tilde{T}(\cdot,\cdot,\cdot)|\mathrm{data})$ by convolving a bivariate white noise field with the square roots of local predictive covariance matrices



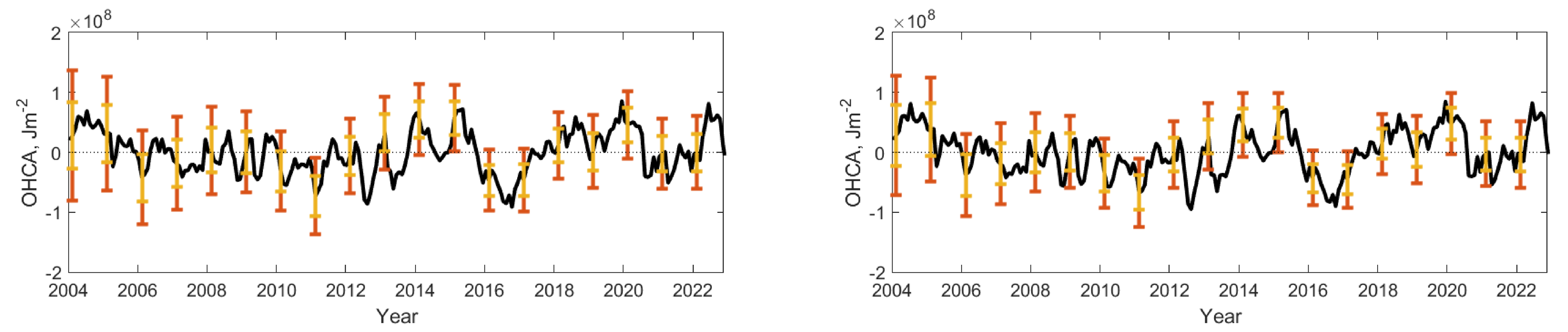
(a) $\Delta t = 0$ months (b) $\Delta t = 1$ month (c) $\Delta t = 2$ months

Cross-validated cross-correlation at varying spatial/temporal distances (x-axis, panels a,b,c) show conditional simulations (blue) capture observed dependence (black) well between vertical sections

Uncertainty quantification results

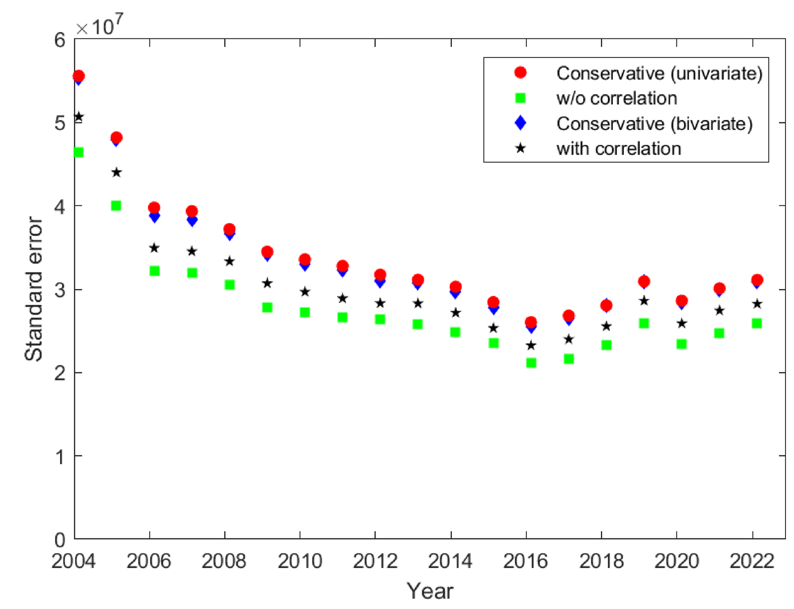
- By producing bivariate conditional simulations, we can now include the cross-dependence term effect in the uncertainty estimate:
$\mathrm{Var}(\mathrm{OHC}_{\mathrm{total}}|\mathrm{data}) = \mathrm{Var}(\mathrm{OHC}_{\mathrm{top}}|\mathrm{data}) + \mathrm{Var}(\mathrm{OHC}_{\mathrm{bot}}|\mathrm{data}) + 2\mathrm{Cov}(\mathrm{OHC}_{\mathrm{top}}, \mathrm{OHC}_{\mathrm{bot}}|\mathrm{data})$
- Below we provide uncertainty estimates for the total OHC anomaly (15–1850 m) for Feb of each year from 2004-2022

Global ocean heat content anomalies



(a) Univariate (b) Bivariate

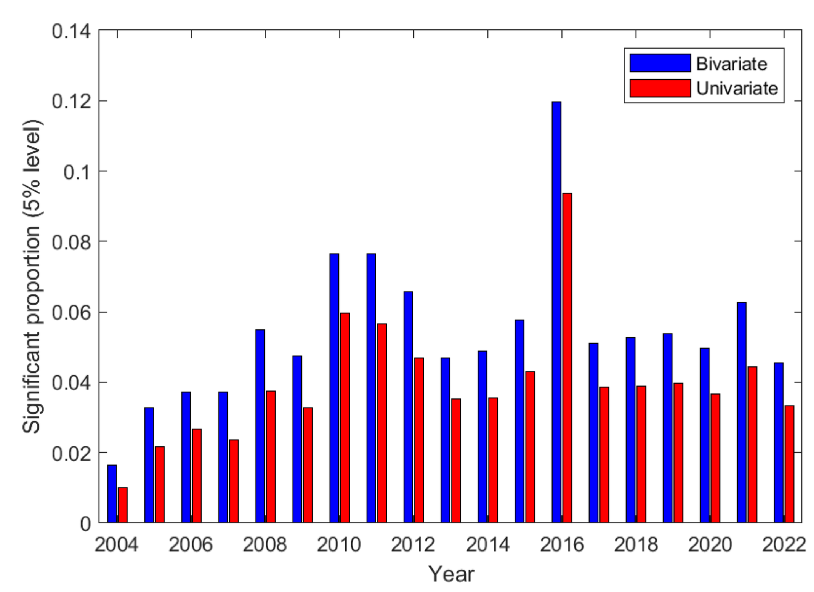
Global OHC anomalies with 68% (orange) and 95% (red) uncertainties



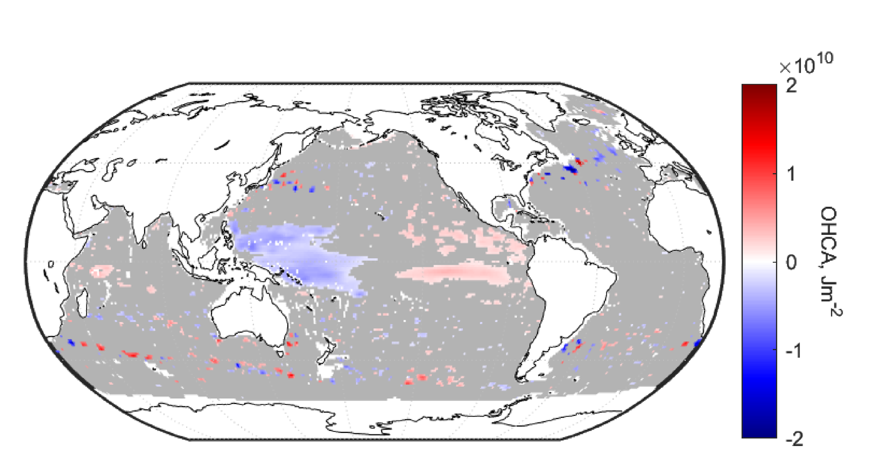
Global OHC anomaly standard errors

The bivariate uncertainties with vertical dependence (black) are 15% smaller than the univariate conservative (red)

Regional ocean heat content anomalies



Proportion of grid points with significant OHC anomalies at the 5% level



Regional OHC anomalies in Feb 2016; only trends that are significant at 5% level are shown

Ongoing work

- This framework can also be used to improve uncertainties for OHC trends, ocean heat uptake (OHU), and other quantities of interest to ocean/climate scientists
- Our univariate estimates with uncertainties were recently featured in the 2023 State of the Climate:

This work was partially supported by NOAA awards NA21OAR4310258 and NA21OAR4310261. We used Bridges-2 at the Pittsburgh Supercomputing Center through ACCESS allocation MTH230014.

Contact: tsukiant@andrew.cmu.edu