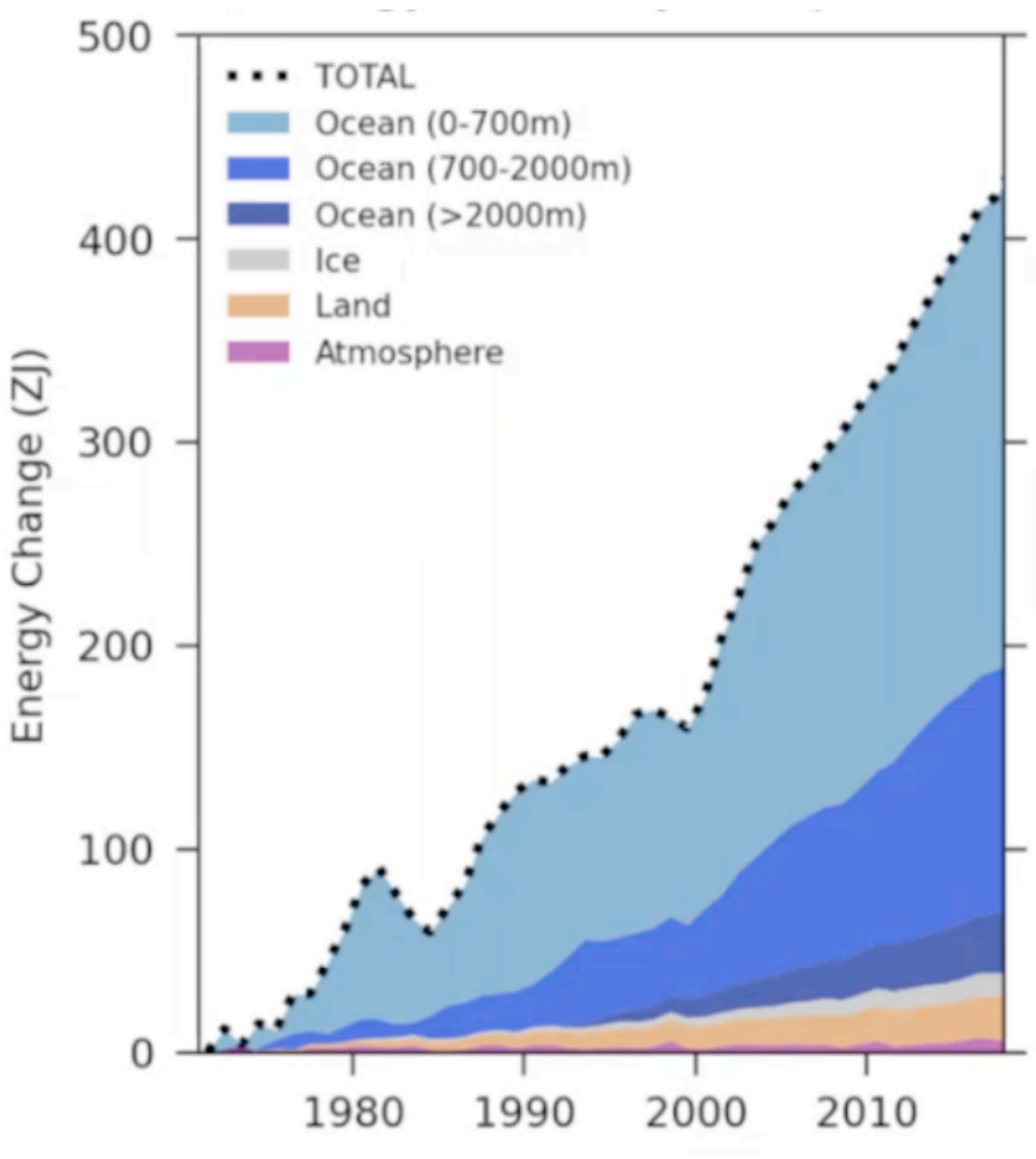
Toward improved global ocean heat content uncertainty quantification by modeling vertical spatio-temporal dependence

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

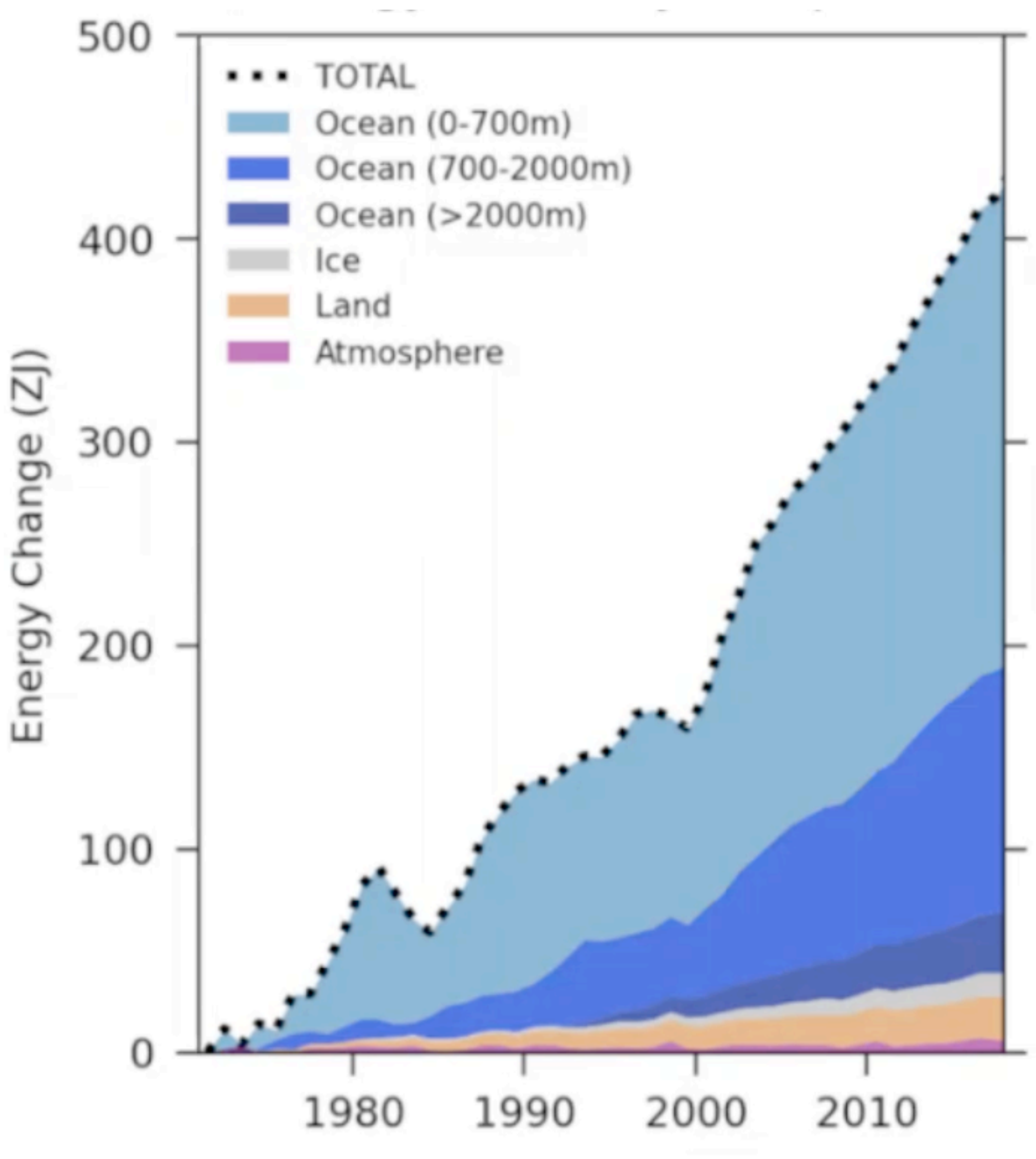
[1]Department of Statistics and Data Science, Carnegie Mellon University
[2]Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder

Most of the excess heat in the climate system has been stored in the ocean



(IPCC 2021)

Changes in ocean heat content (OHC) contribute to extreme climate events

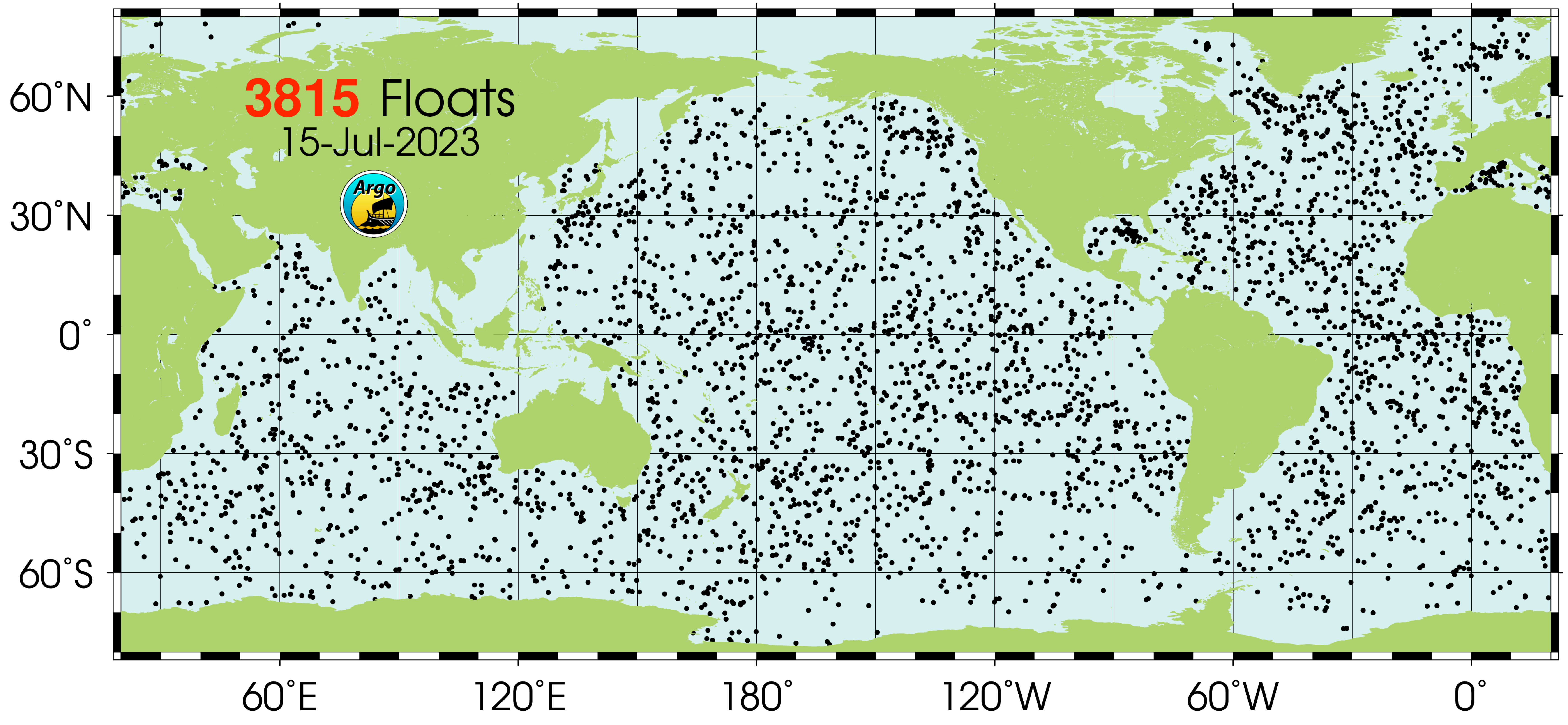


(IPCC 2021)

Intensified tropical storms

Rising sea levels

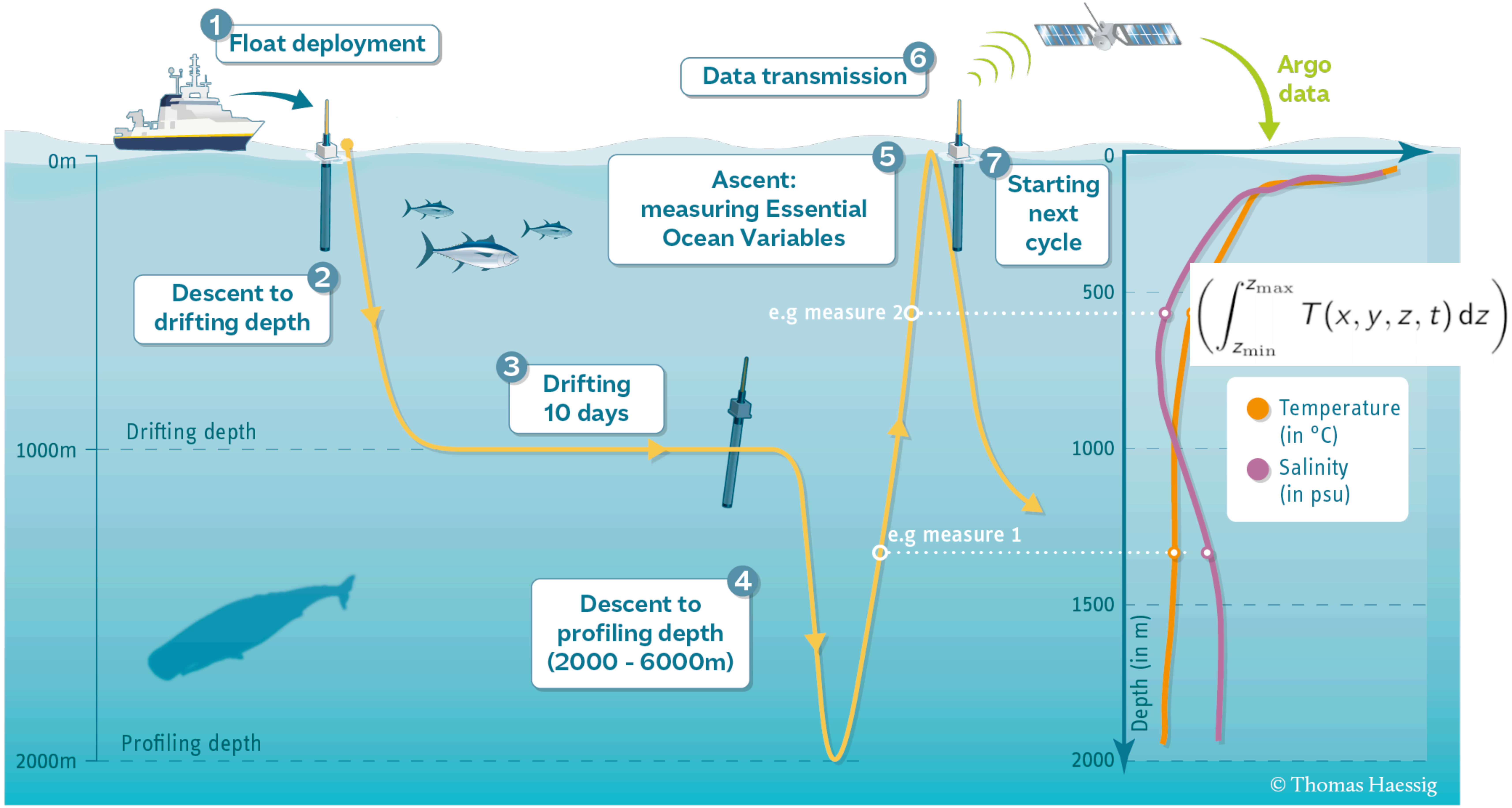
Argo floats are the state-of-the-art in ocean temperature measurements



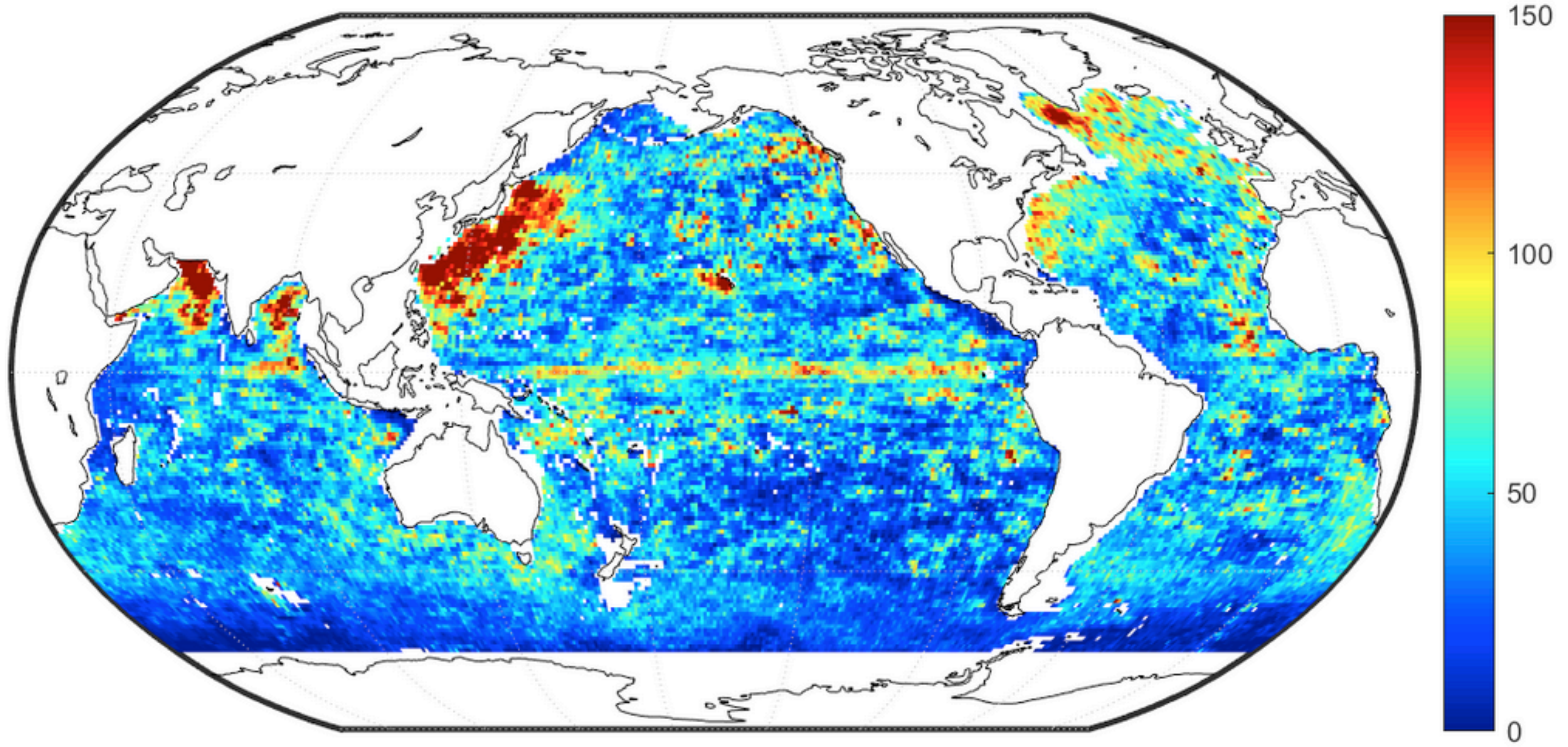
(Argo Program)

We use Argo data (2004-2021) to estimate 15-1850m OHC

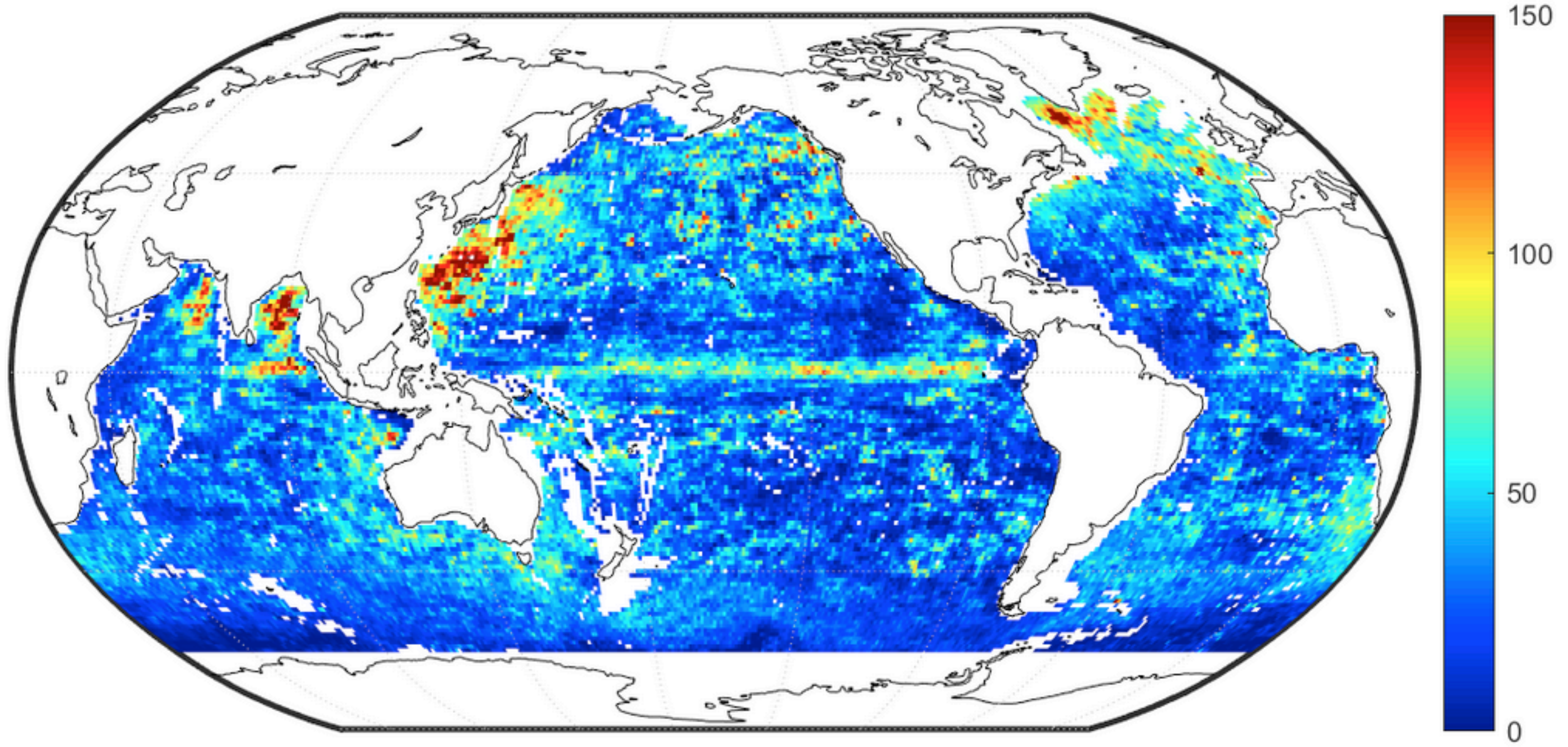
$$\mathrm{OHC}(t) = \rho_0 c_{p,0} \int\int \left(\int_{z_{\min}}^{z_{\max}} T(x,y,z,t)\,\mathrm{d}z \right) \mathrm{d}x\,\mathrm{d}y$$



We use Argo data (2004-2021) to estimate 15-1850m OHC

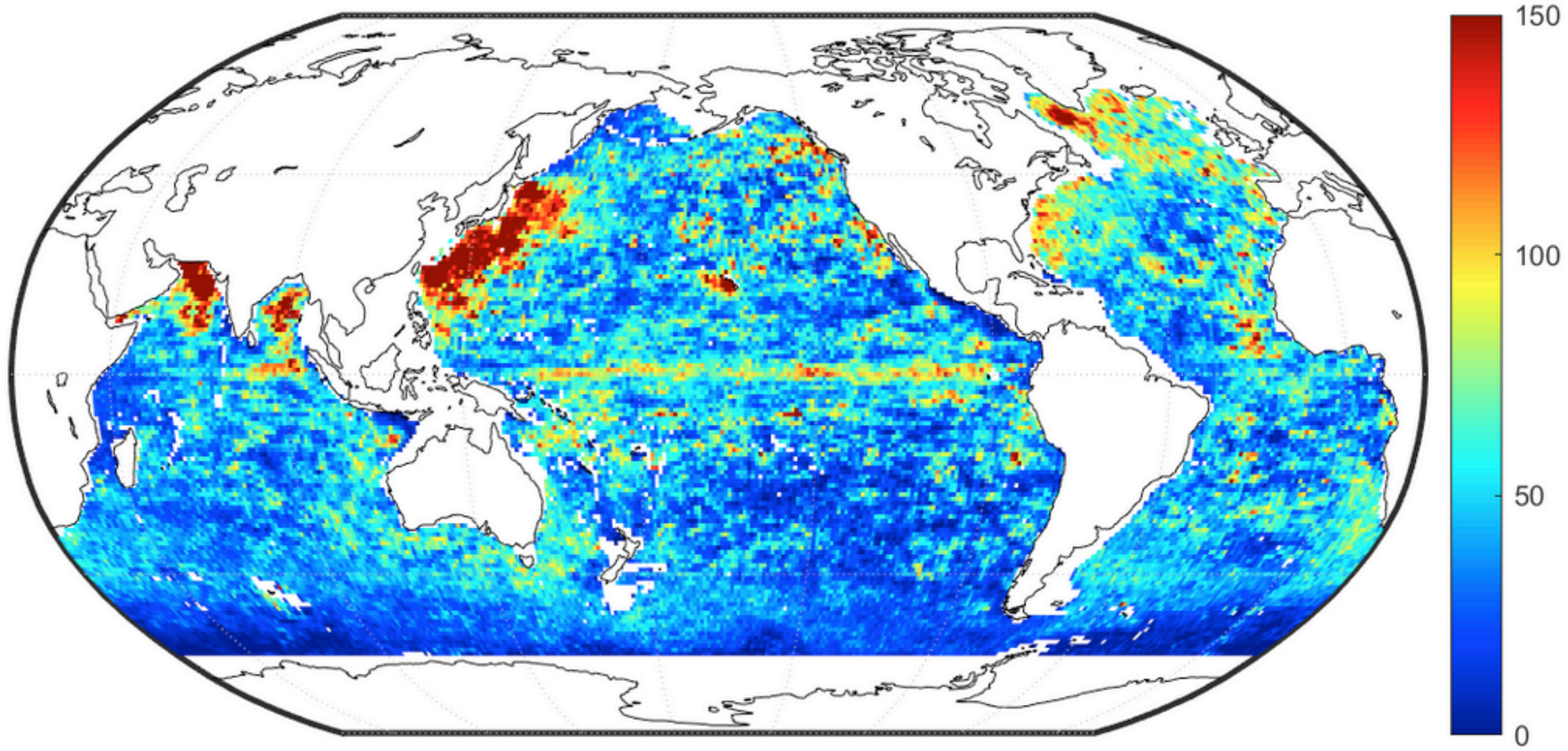


Top layer profiles (15-975m)

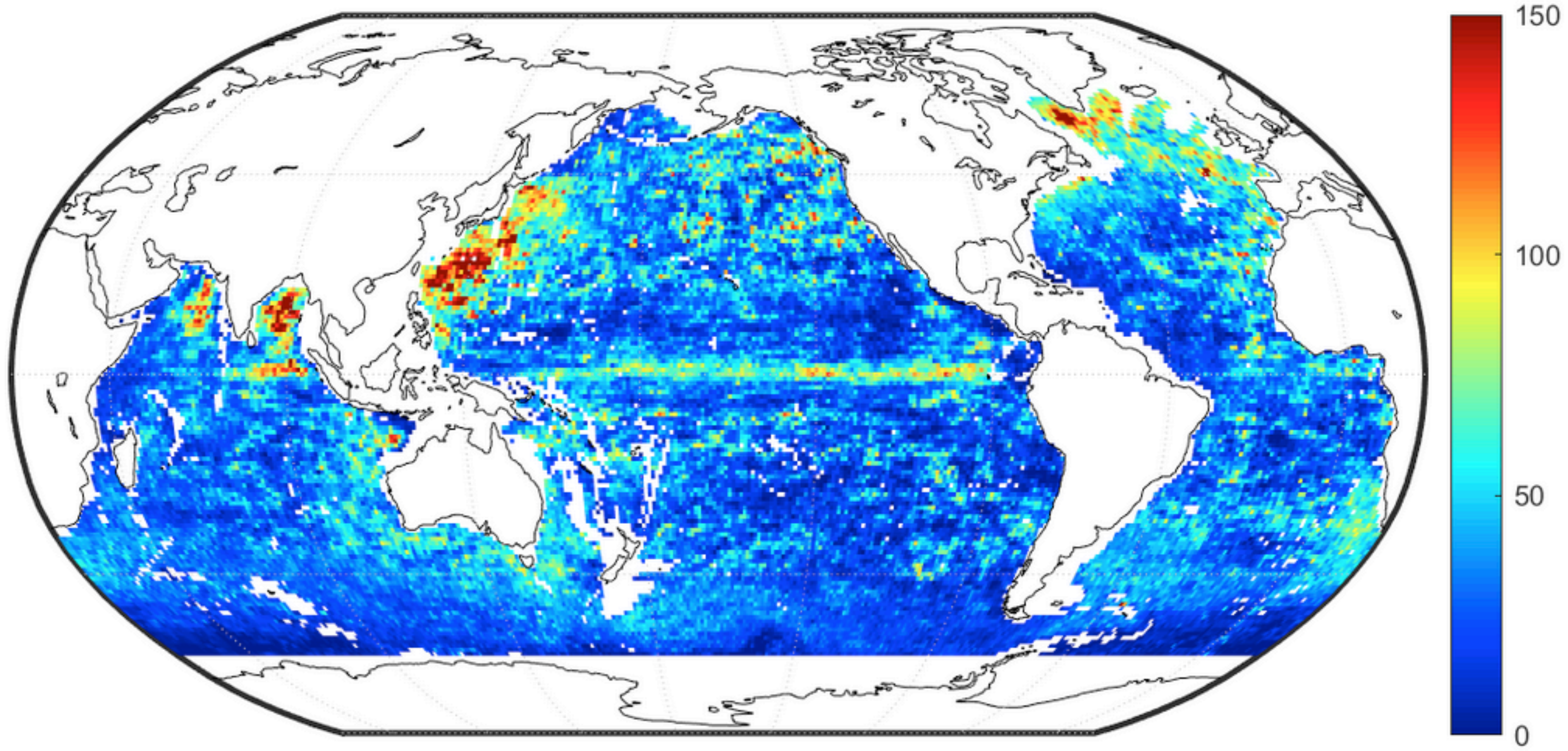


Bottom layer profiles (975-1850m)

We use Argo data (2004-2021) to estimate 15-1850m OHC



Top layer profiles (15-975m)



Bottom layer profiles (975-1850m)

Due to having fewer observations deeper in the water column, previous work has modeled the top and bottom layers **separately.**

Modeling challenges arise from the data size and nonstationarity

- **Size:** > 2.5 million Argo profiles (matrix inversion for covariance parameter estimation and kriging is infeasible)
- **Nonstationarity:** Challenging to define a nonstationary covariance function flexible enough to explain variability across entire global ocean
- **Uncertainties for a global integral?** Conditional simulations (extension of Nychka et.al. 2018)

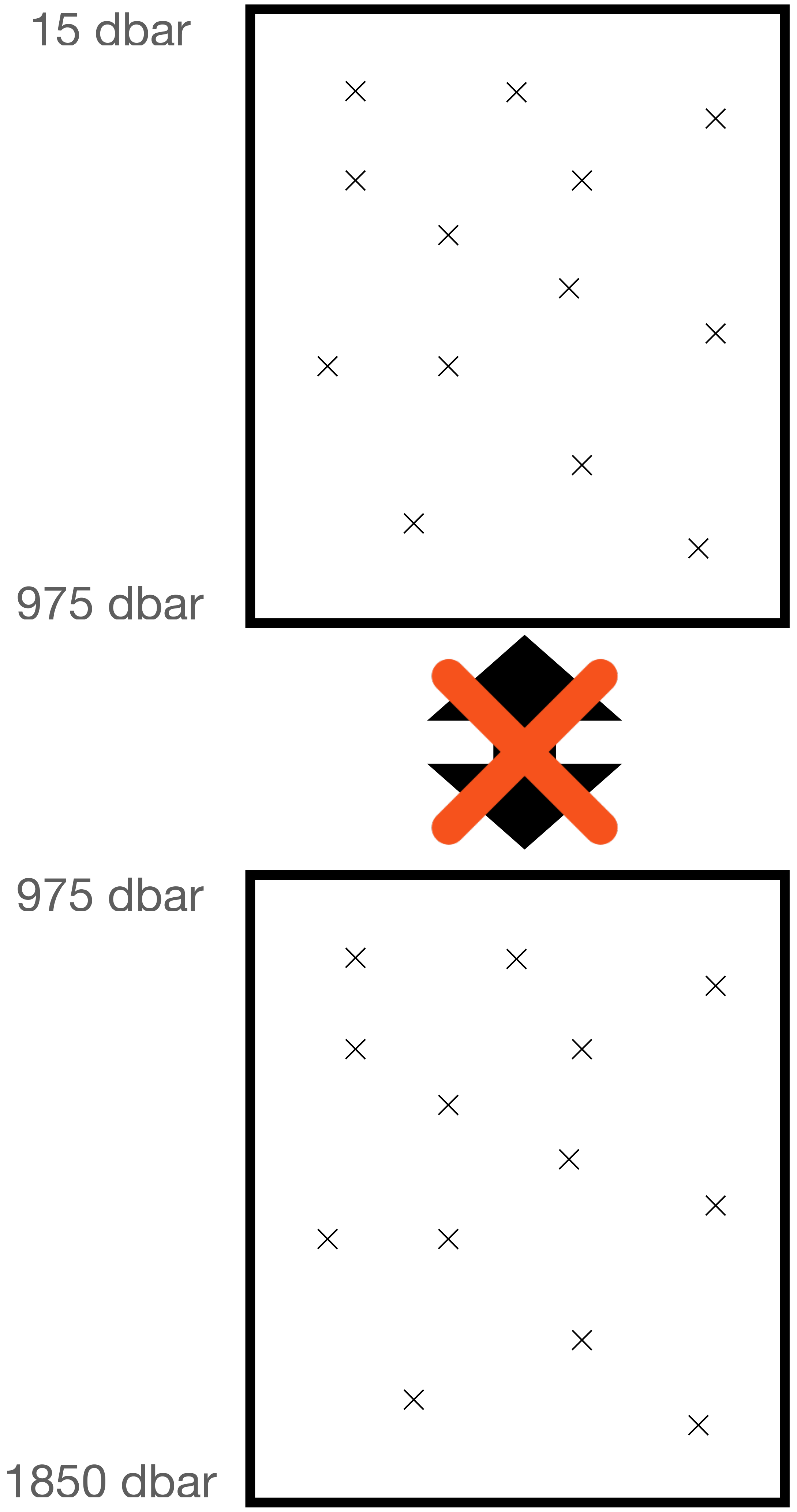
Modeling challenges arise from the data size and nonstationarity

- **Size:** > 2.5 million Argo profiles (matrix inversion for covariance parameter estimation and kriging is infeasible)
- **Nonstationarity:** Challenging to define a nonstationary covariance function flexible enough to explain variability across entire global ocean
- **Uncertainties for a global integral?** Conditional simulations (extension of Nychka et.al. 2018)

Local spatio-temporal modeling:

1. Model temperature mean field: local least squares regression (Roemmich and Gilson 2009) + linear time trend
2. Model residuals **(anomalies)**: **locally stationary** Gaussian process (GP) regression (Kuusela and Stein 2018)

We can improve the uncertainties by modeling the correlation



$$\mathrm{Var}(\mathrm{OHC}_{\mathrm{top}}|\mathrm{data})$$

$$2\mathrm{Cov}(\mathrm{OHC}_{\mathrm{top}}, \mathrm{OHC}_{\mathrm{bot}}|\mathrm{data})$$

$$\mathrm{Var}(\mathrm{OHC}_{\mathrm{bot}}|\mathrm{data})$$

$$\boxed{\mathrm{OHC}_{\mathrm{total}} = \mathrm{OHC}_{\mathrm{top}} + \mathrm{OHC}_{\mathrm{bot}}}$$

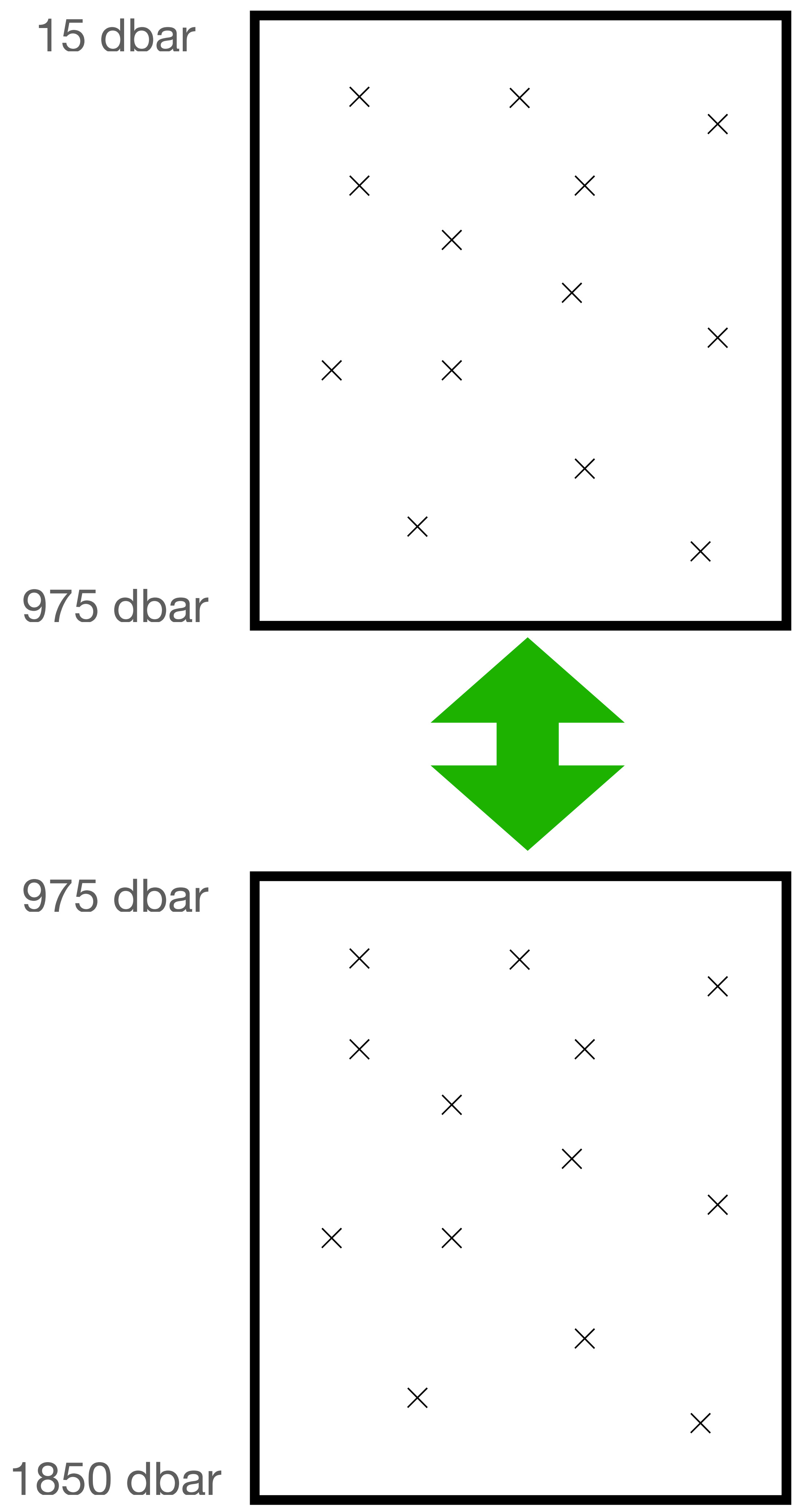
$$\mathrm{Var}(\cancel{\mathrm{OHC}_{\mathrm{total}}}|\mathrm{data})$$

(conservative upper bound)

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

We can improve the uncertainties by modeling the correlation

$$\mathrm{OHC}_{\mathrm{total}} = \mathrm{OHC}_{\mathrm{top}} + \mathrm{OHC}_{\mathrm{bot}}$$



$$\mathrm{Var}(\mathrm{OHC}_{\mathrm{top}}|\mathrm{data})$$

$$2\mathrm{Cov}(\mathrm{OHC}_{\mathrm{top}}, \mathrm{OHC}_{\mathrm{bot}}|\mathrm{data})$$

$$\mathrm{Var}(\mathrm{OHC}_{\mathrm{bot}}|\mathrm{data})$$

$$\mathrm{Var}(\mathrm{OHC}_{\mathrm{total}}|\mathrm{data})$$

A bivariate GP model accounts for cross-layer correlation

$$\begin{bmatrix} y_{\text{top}} \\ y_{\text{bot}} \end{bmatrix}_{i,j} = f_i(\begin{bmatrix} x_{\text{top}} \\ x_{\text{bot}} \end{bmatrix}_{i,j}, \begin{bmatrix} t_{\text{top}} \\ t_{\text{bot}} \end{bmatrix}_{i,j}) + \begin{bmatrix} \epsilon_{\text{top}} \\ \epsilon_{\text{bot}} \end{bmatrix}_{i,j}$$

Temperature residuals | **Latitude Longitude** | **Date** | **Nugget effect**

$$f_i \overset{\text{iid}}{\sim} \text{GP}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{K}(x_1, t_1, x_2, t_2; \boldsymbol{\theta}))$$

$$\begin{bmatrix} \epsilon_{\text{top}} \\ \epsilon_{\text{bot}} \end{bmatrix} \overset{\text{iid}}{\sim} \text{N}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}(\boldsymbol{\theta}_{\boldsymbol{\epsilon}}))$$

A bivariate GP model accounts for cross-layer correlation

$$\begin{bmatrix} y_{\text{top}} \\ y_{\text{bot}} \end{bmatrix}_{i,j} = f_i(\begin{bmatrix} x_{\text{top}} \\ x_{\text{bot}} \end{bmatrix}_{i,j}, \begin{bmatrix} t_{\text{top}} \\ t_{\text{bot}} \end{bmatrix}_{i,j}) + \begin{bmatrix} \epsilon_{\text{top}} \\ \epsilon_{\text{bot}} \end{bmatrix}_{i,j}$$

Temperature residuals | **Latitude Longitude** | **Date** | **Nugget effect**

$$f_i \stackrel{\text{iid}}{\sim} \text{GP}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{K}(x_1, t_1, x_2, t_2; \boldsymbol{\theta}))$$

$$\begin{bmatrix} \epsilon_{\text{top}} \\ \epsilon_{\text{bot}} \end{bmatrix} \stackrel{\text{iid}}{\sim} \text{N}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}(\boldsymbol{\theta}_{\boldsymbol{\epsilon}}))$$

Marginal covariance
(Kuusela and Stein 2018)

$$\boldsymbol{K}_{\text{ii}}(\boldsymbol{z_1}, \boldsymbol{z_2}; \boldsymbol{\theta}) = \frac{\delta_i^2}{\sqrt{|\boldsymbol{\Theta}_i|}} \exp(-\sqrt{(\boldsymbol{z_1} - \boldsymbol{z_2})^{\text{T}} \boldsymbol{\Theta}_i^{-1} (\boldsymbol{z_1} - \boldsymbol{z_2})})$$

Cross-covariance
(Kleiber and Nychka 2012)

$$\boldsymbol{K}_{\text{top,bot}}(\boldsymbol{z_1}, \boldsymbol{z_2}; \boldsymbol{\theta}) = \beta \frac{\delta_{\text{top}} \delta_{\text{bot}}}{\sqrt{|\boldsymbol{\Theta}_{\text{top,bot}}|}} \exp(-\sqrt{(\boldsymbol{z_1} - \boldsymbol{z_2})^{\text{T}} \boldsymbol{\Theta}_{\text{top,bot}} (\boldsymbol{z_1} - \boldsymbol{z_2})})$$

A bivariate GP model accounts for cross-layer correlation

$$\begin{bmatrix} y_{\text{top}} \\ y_{\text{bot}} \end{bmatrix}_{i,j} = f_i(\begin{bmatrix} x_{\text{top}} \\ x_{\text{bot}} \end{bmatrix}_{i,j}, \begin{bmatrix} t_{\text{top}} \\ t_{\text{bot}} \end{bmatrix}_{i,j}) + \begin{bmatrix} \epsilon_{\text{top}} \\ \epsilon_{\text{bot}} \end{bmatrix}_{i,j}$$

Temperature residuals | **Latitude Longitude** | **Date** | **Nugget effect**

$$f_i \stackrel{\text{iid}}{\sim} \text{GP}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{K}(x_1, t_1, x_2, t_2; \boldsymbol{\theta}))$$

$$\begin{bmatrix} \epsilon_{\text{top}} \\ \epsilon_{\text{bot}} \end{bmatrix} \stackrel{\text{iid}}{\sim} \text{N}(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \boldsymbol{\Sigma}_{\boldsymbol{\epsilon}}(\boldsymbol{\theta}_{\boldsymbol{\epsilon}}))$$

Marginal covariance
(Kuusela and Stein 2018)

$$\boldsymbol{K}_{\text{ii}}(\boldsymbol{z_1}, \boldsymbol{z_2}; \boldsymbol{\theta}) = \frac{\delta_i^2}{\sqrt{|\boldsymbol{\Theta}_i|}} \exp(-\sqrt{(\boldsymbol{z_1} - \boldsymbol{z_2})^{\text{T}} \boldsymbol{\Theta}_i^{-1} (\boldsymbol{z_1} - \boldsymbol{z_2})})$$

Cross-covariance
(Kleiber and Nychka 2012)

$$\boldsymbol{K}_{\text{top,bot}}(\boldsymbol{z_1}, \boldsymbol{z_2}; \boldsymbol{\theta}) = \beta \frac{\delta_{\text{top}} \delta_{\text{bot}}}{\sqrt{|\boldsymbol{\Theta}_{\text{top,bot}}|}} \exp(-\sqrt{(\boldsymbol{z_1} - \boldsymbol{z_2})^{\text{T}} \boldsymbol{\Theta}_{\text{top,bot}} (\boldsymbol{z_1} - \boldsymbol{z_2})})$$

This is the first time (to our knowledge) that a bivariate GP model is **fitted locally**.

Obtaining uncertainties is facilitated by local conditional simulations

Previously: modeling challenges from **data size and nonstationarity**

Uncertainties for $\mathrm{OHC}(t) = \rho_0 c_{p,0} \int\int \left(\int_{z_{\min}}^{z_{\max}} T(x, y, z, t)\, \mathrm{d}z \right) \mathrm{d}x\, \mathrm{d}y$?

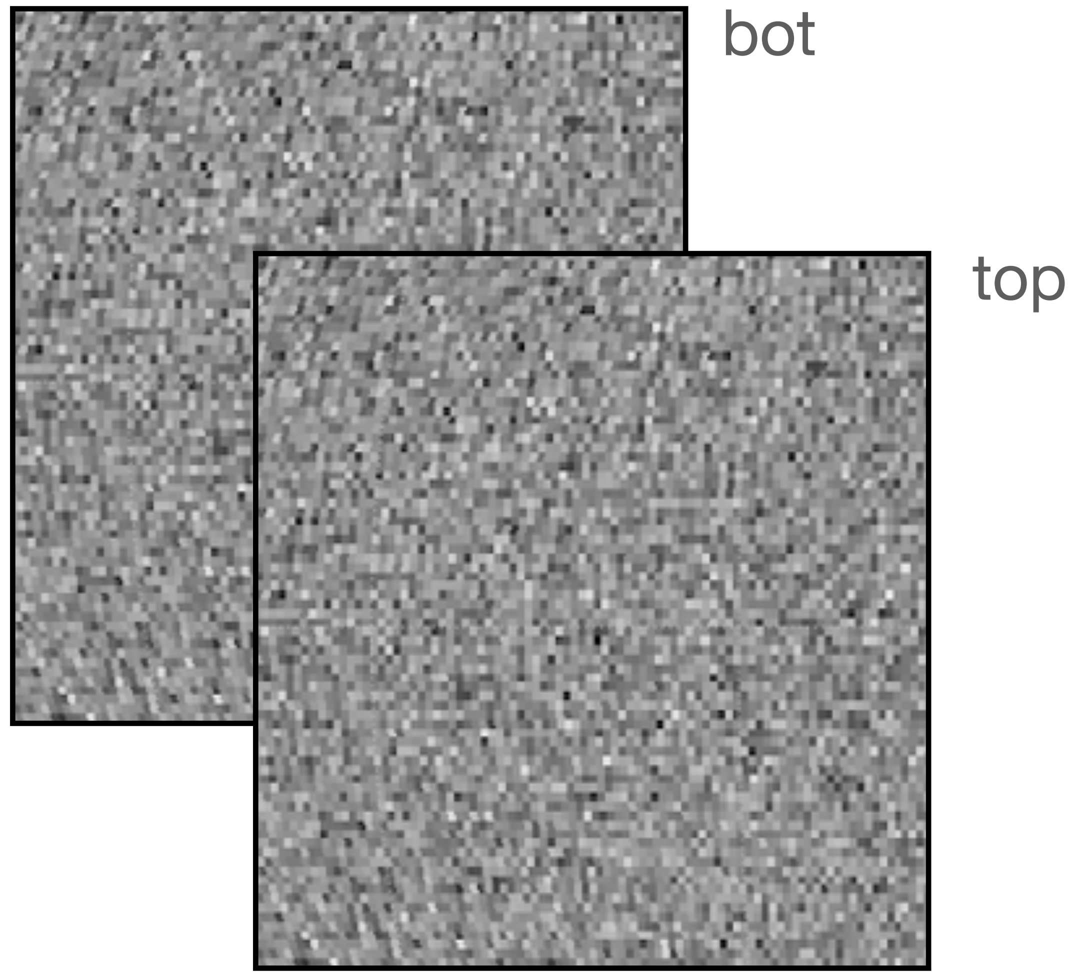
Obtaining uncertainties is facilitated by local conditional simulations

Previously: modeling challenges from **data size and nonstationarity**

Uncertainties for $\mathrm{OHC}(t) = \rho_0 c_{p,0} \int\int \left(\int_{z_{\min}}^{z_{\max}} T(x,y,z,t)\,\mathrm{d}z \right) \mathrm{d}x\,\mathrm{d}y$?

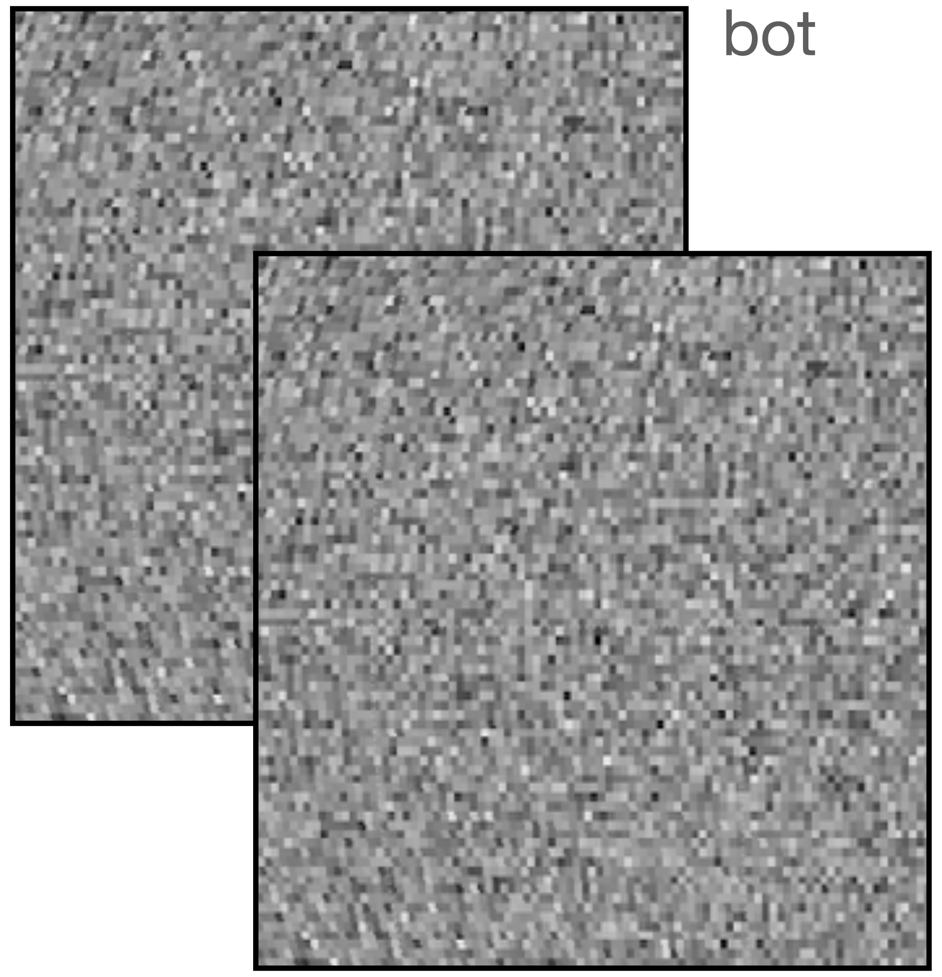
Local conditional simulations! (extension of Nychka et.al. 2018)

Obtaining uncertainties is facilitated by local conditional simulations



Simulate Gaussian
white noise over grid
(keep fixed)

Obtaining uncertainties is facilitated by local conditional simulations

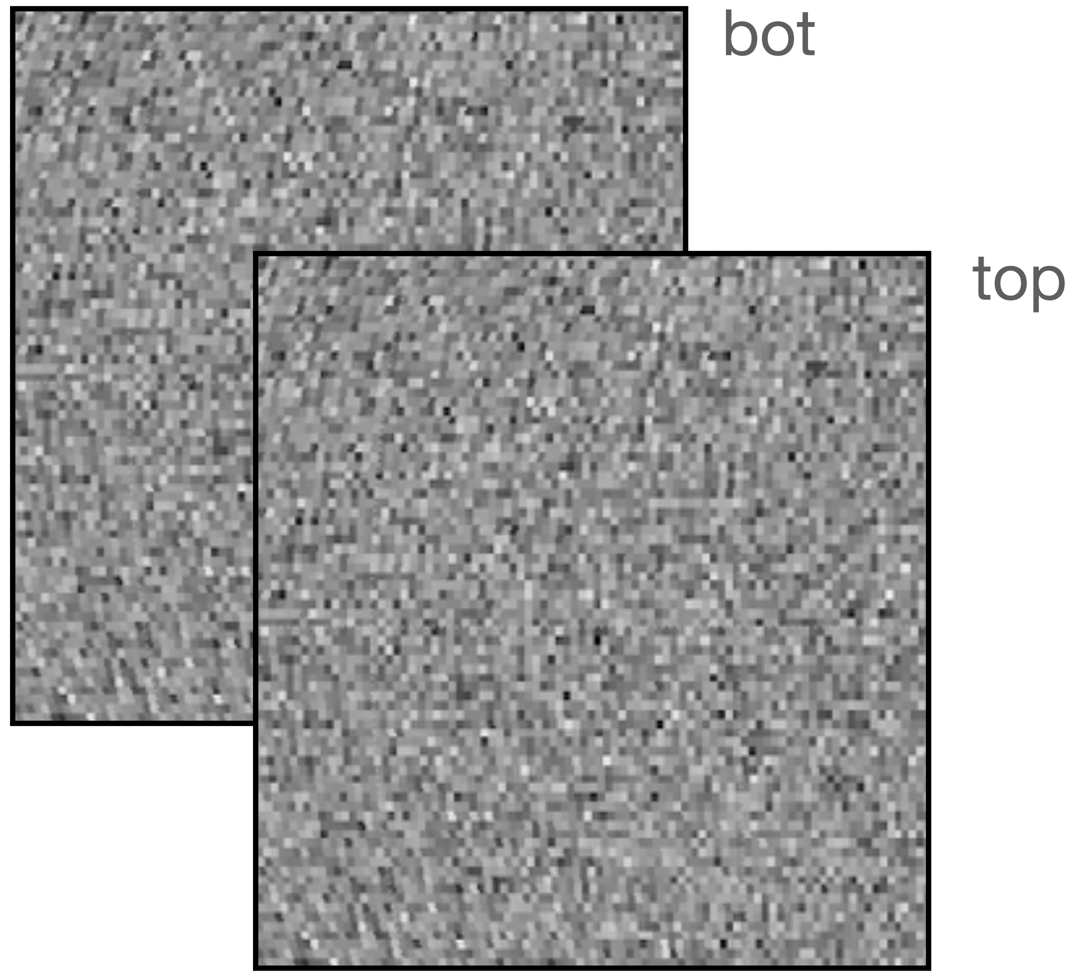


Simulate Gaussian
white noise over grid
(keep fixed)

Σ_i

Compute (local)
conditional covariance
matrix

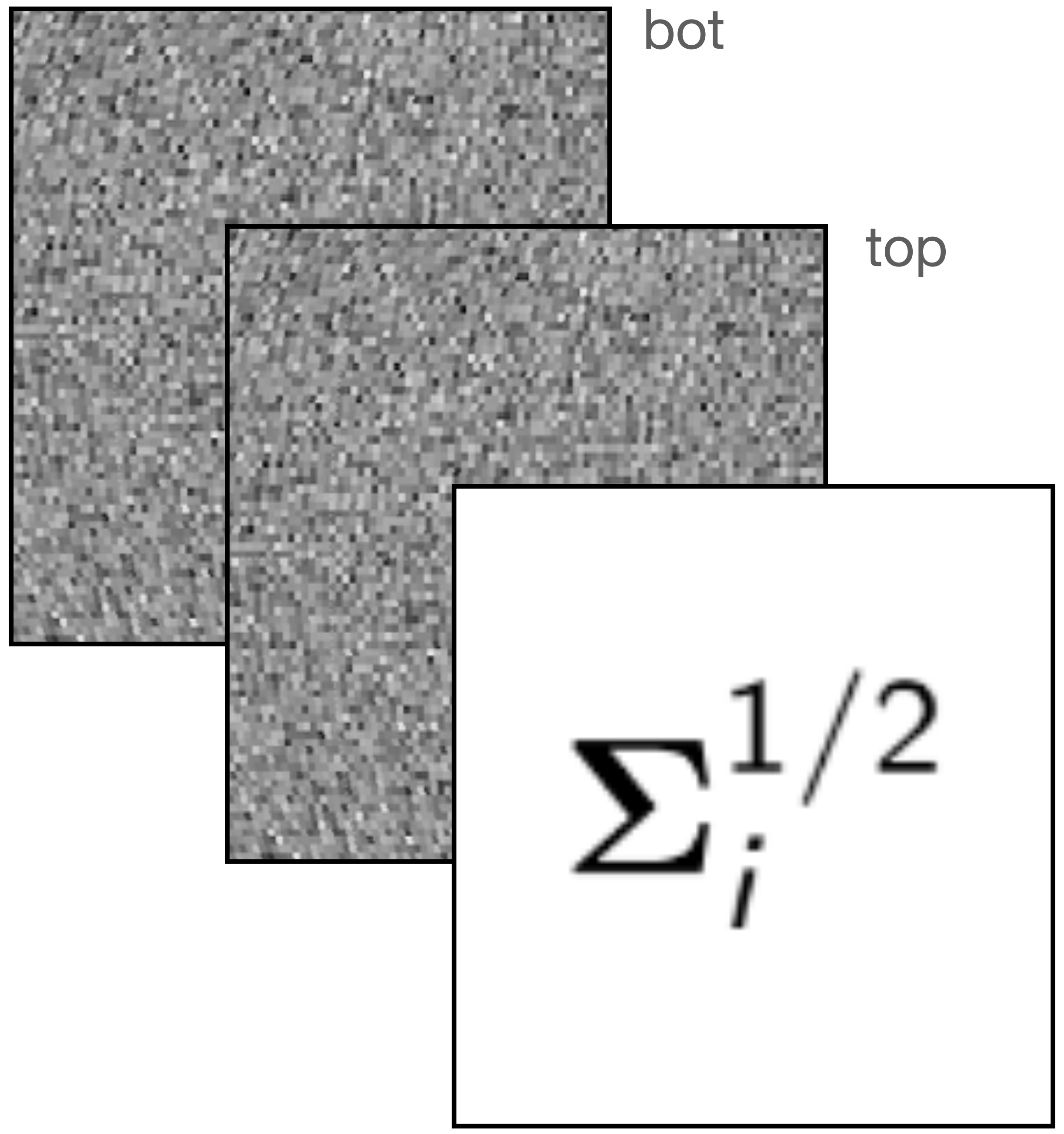
Obtaining uncertainties is facilitated by local conditional simulations



Simulate Gaussian white noise over grid (keep fixed)

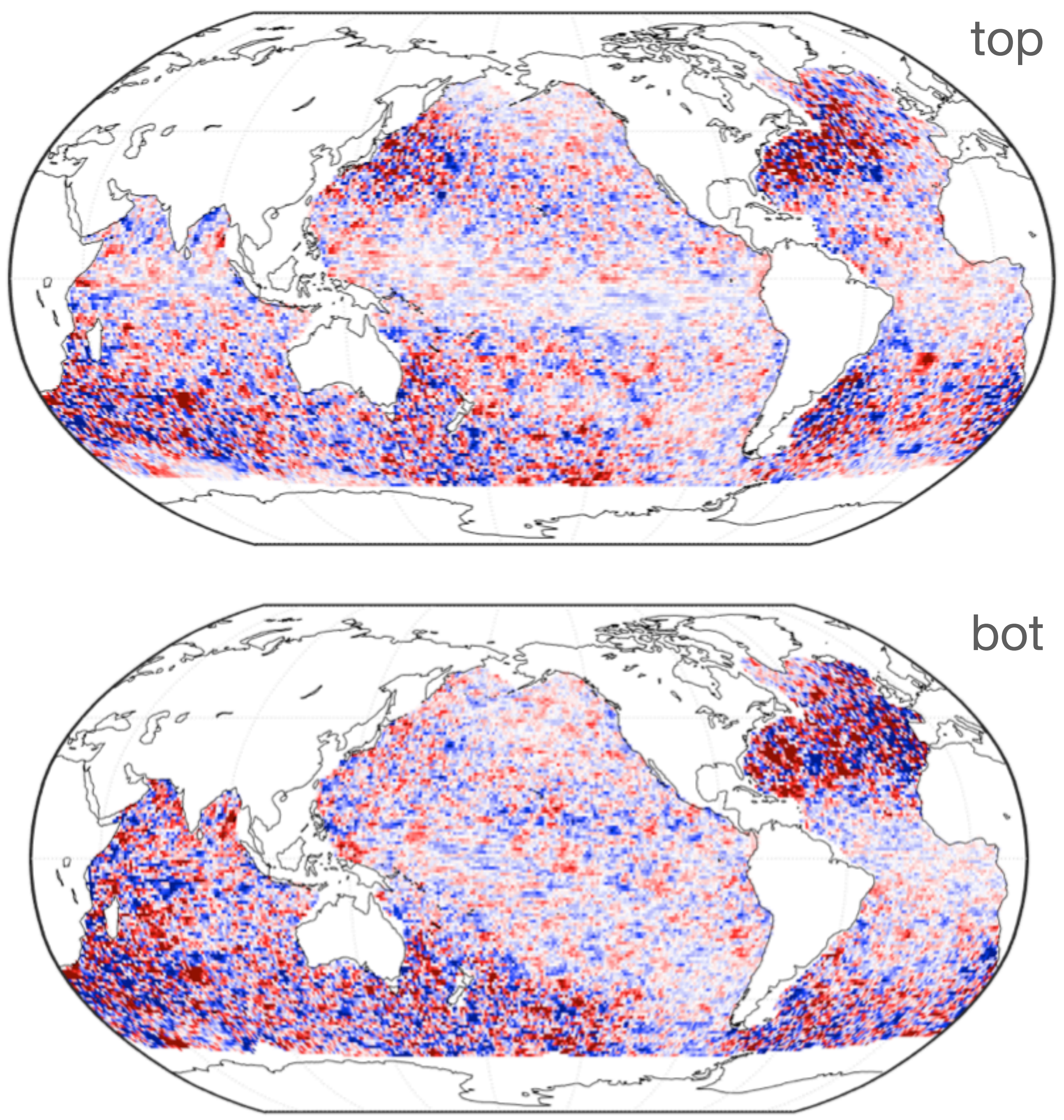
$$\Sigma_i$$

Compute (local) conditional covariance matrix



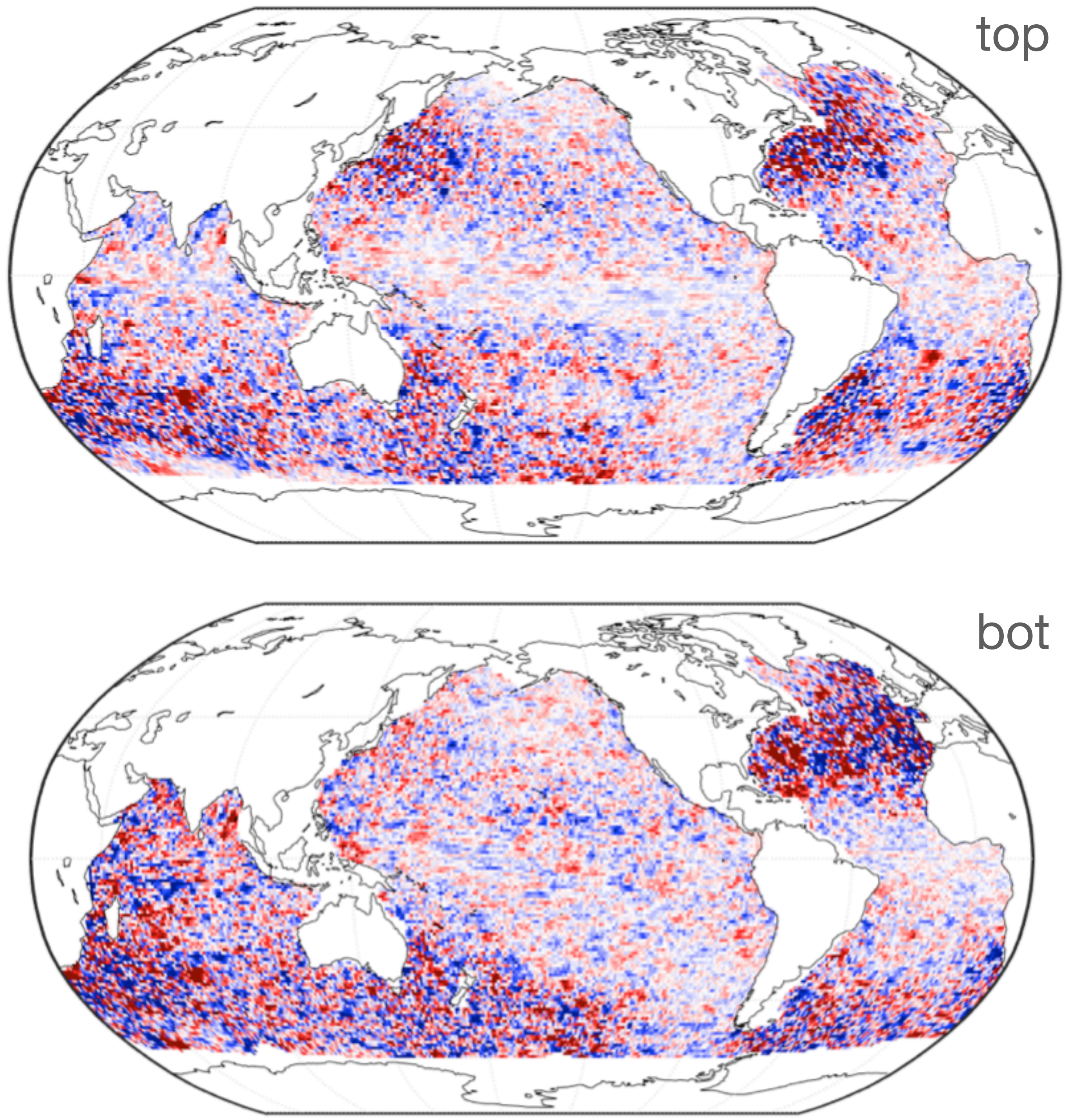
Find symmetric square root (eigendecomposition of previous step) and multiply by white noise

Obtaining uncertainties is facilitated by local conditional simulations

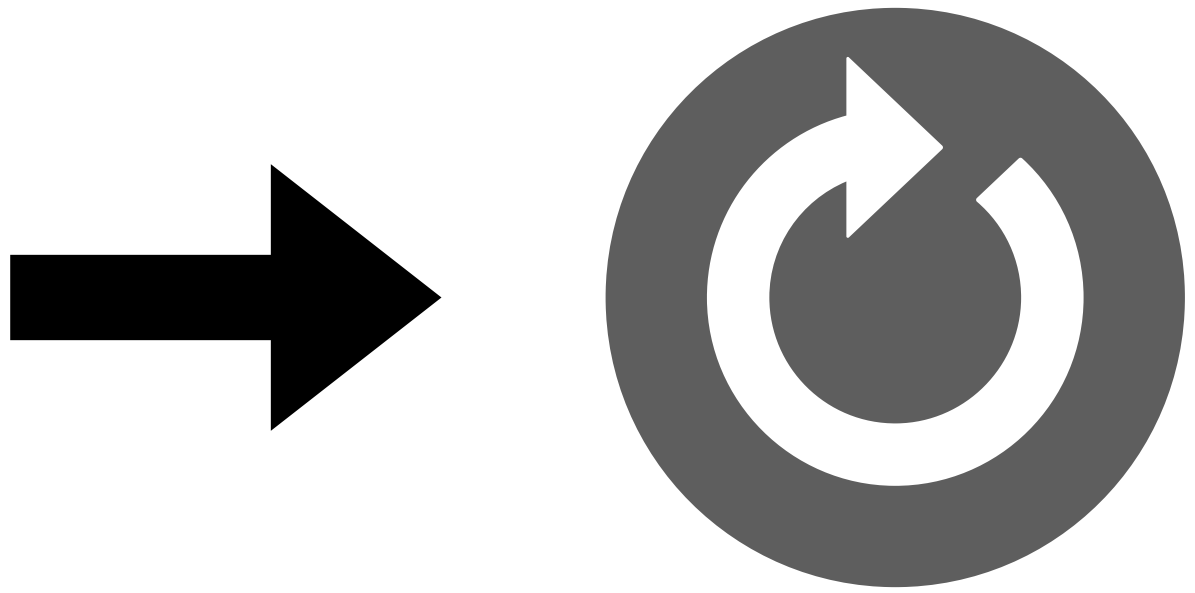


Keep the center point and repeat for all grid points

Obtaining uncertainties is facilitated by local conditional simulations

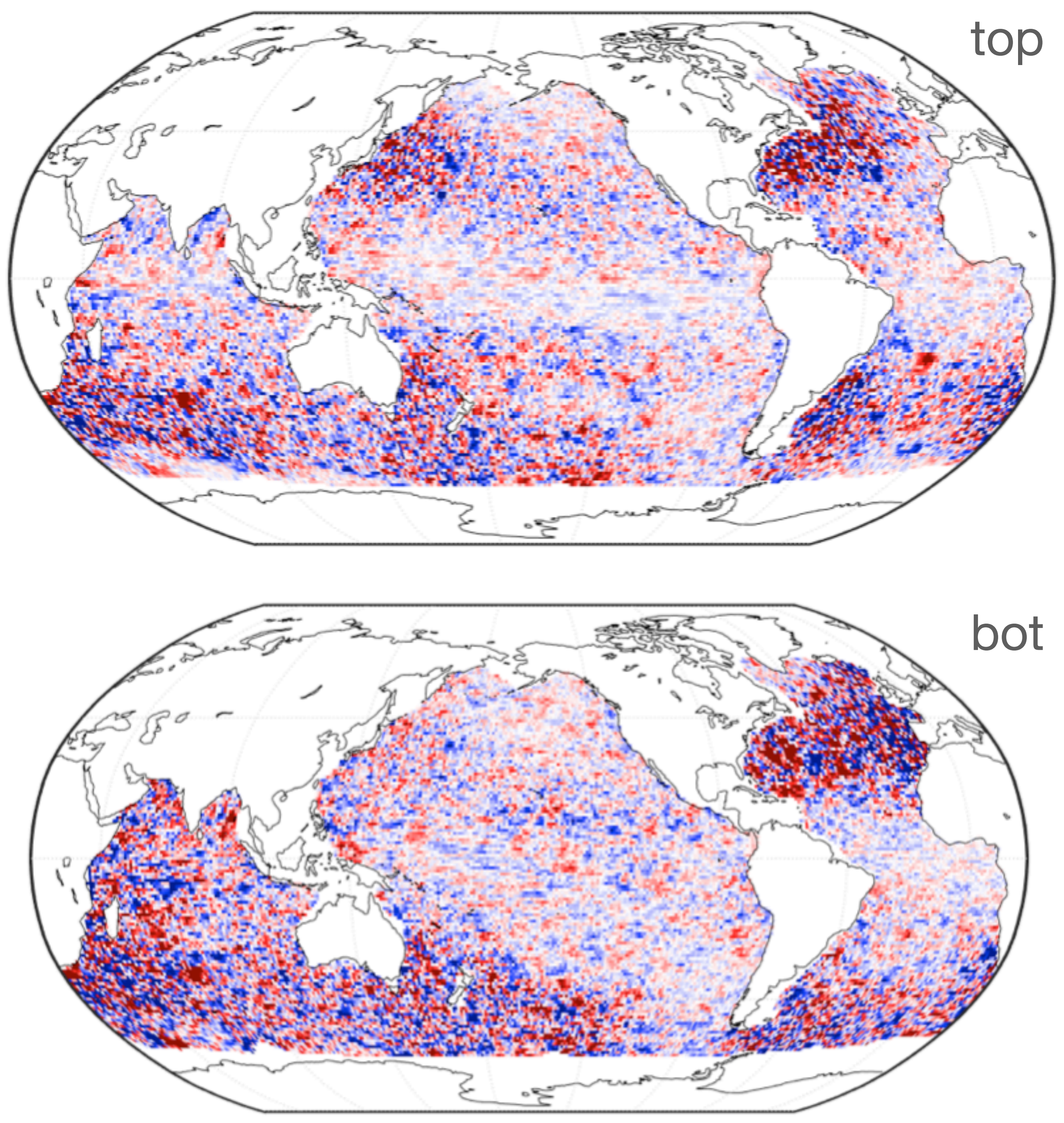


Keep the center point and repeat for all grid points



Repeat for desired number of samples

Obtaining uncertainties is facilitated by local conditional simulations



Keep the center point and repeat for all grid points

Repeat for desired number of samples

We modeled the correlation, so variance of top + bottom layer integrated samples is estimate of

$$\mathrm{Var}(\mathrm{OHC}_{\mathrm{total}}|\mathrm{data})$$

The implementation is still computationally challenging

- For every grid point, use data in a 10 degree/3 month window
- Estimate GP model parameters numerically with MLE + BFGS algorithm
- How many grid points?

The implementation is still computationally challenging

- For every grid point, use data in a 10 degree/3 month window
- Estimate GP model parameters numerically with MLE + BFGS algorithm
- How many grid points?
    - **On 1 x 1 deg grid = separate parameters for ~30,000 grid points (!)**

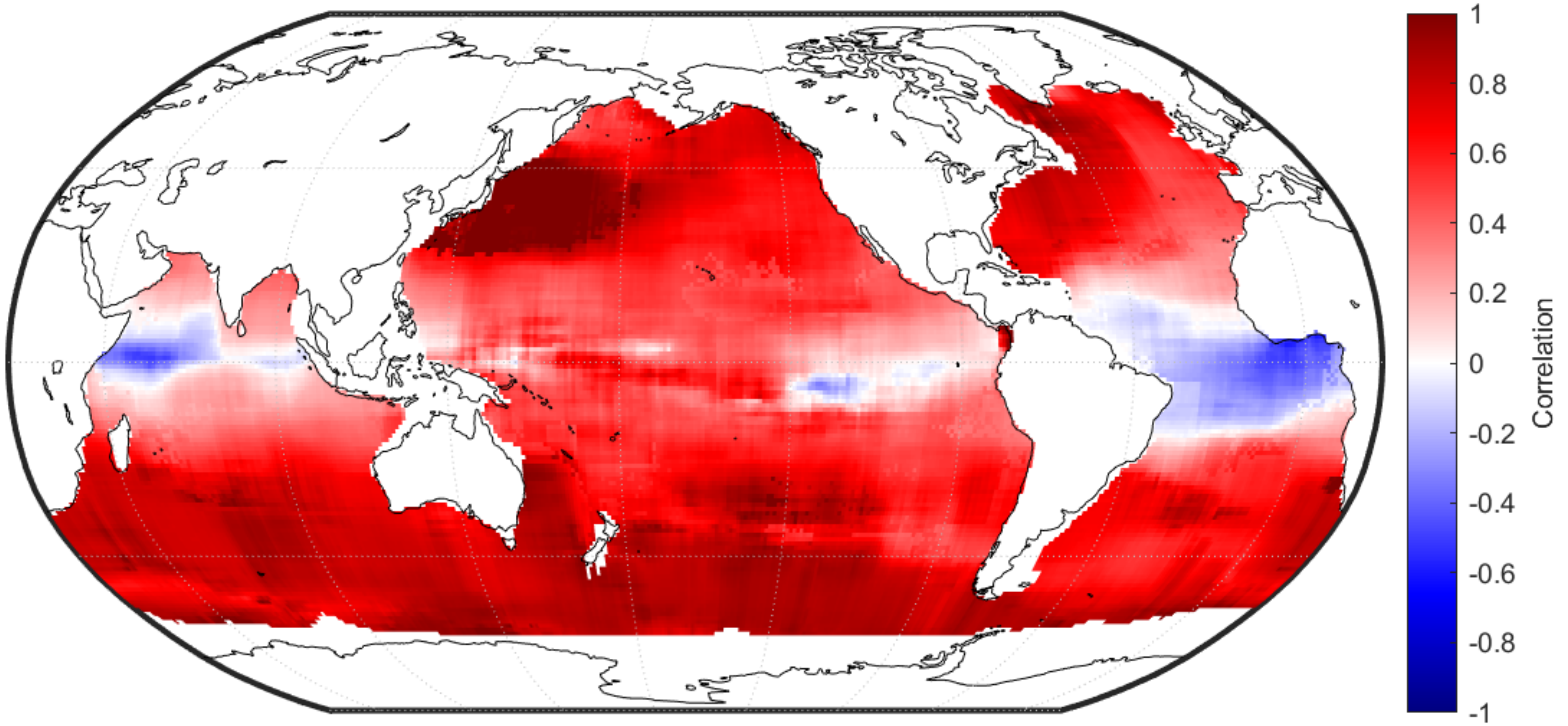
The implementation is still computationally challenging

- For every grid point, use data in a 10 degree/3 month window
- Estimate GP model parameters numerically with MLE + BFGS algorithm
- How many grid points?
    - **On 1 x 1 deg grid = separate parameters for ~30,000 grid points (!)**
- Embarrassingly parallel, but still computationally challenging

The implementation is still computationally challenging

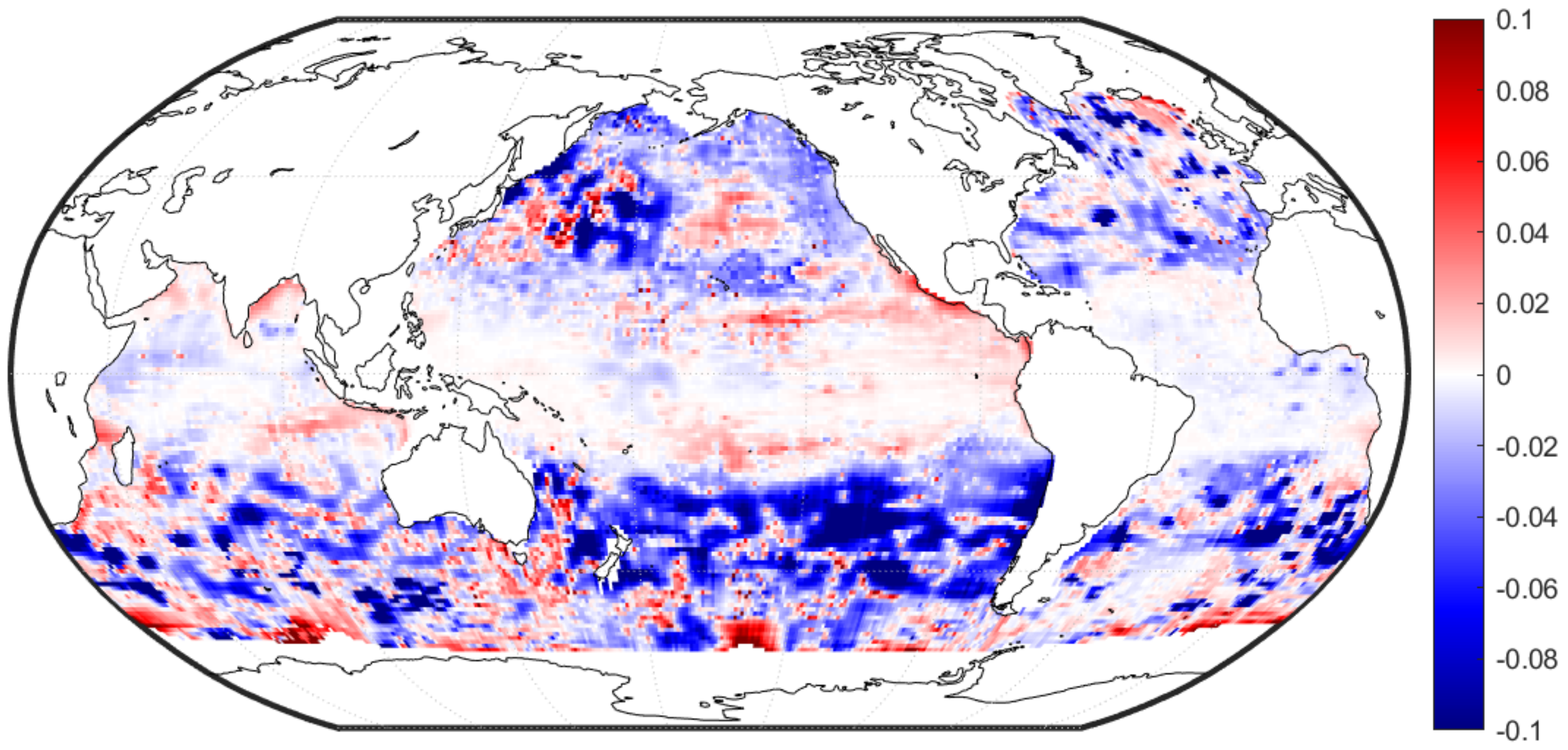
- For every grid point, use data in a 10 degree/3 month window
- Estimate GP model parameters numerically with MLE + BFGS algorithm
- How many grid points?
  - **On 1 x 1 deg grid = separate parameters for ~30,000 grid points (!)**
- Embarrassingly parallel, but still computationally challenging
  - Fit parameters for 180 x 20 test slice:
    - **67h (desktop, 24 cores)**
    - **8h (Pittsburgh Supercomputing Center, 128 cores)**
  - Obtain conditional simulations for Feb of every year: **~12h (desktop, 24 cores)**

Most ocean regions’ temperatures are positively correlated

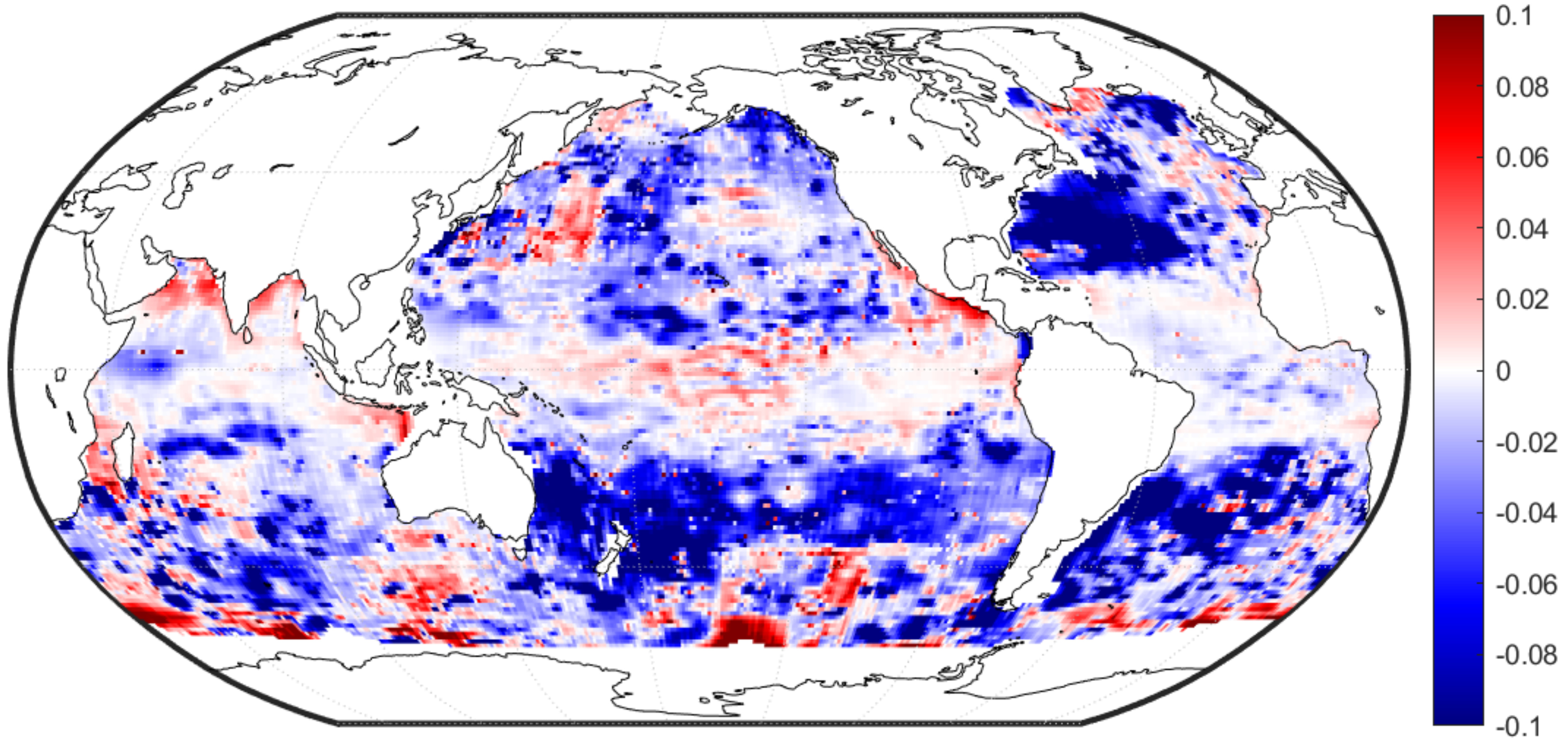


The bivariate model tends to produce lower kriging variances

$$\frac{\text{bivariate kriging variance - univariate kriging variance}}{\text{univariate kriging variance}} \quad (02/2010)$$



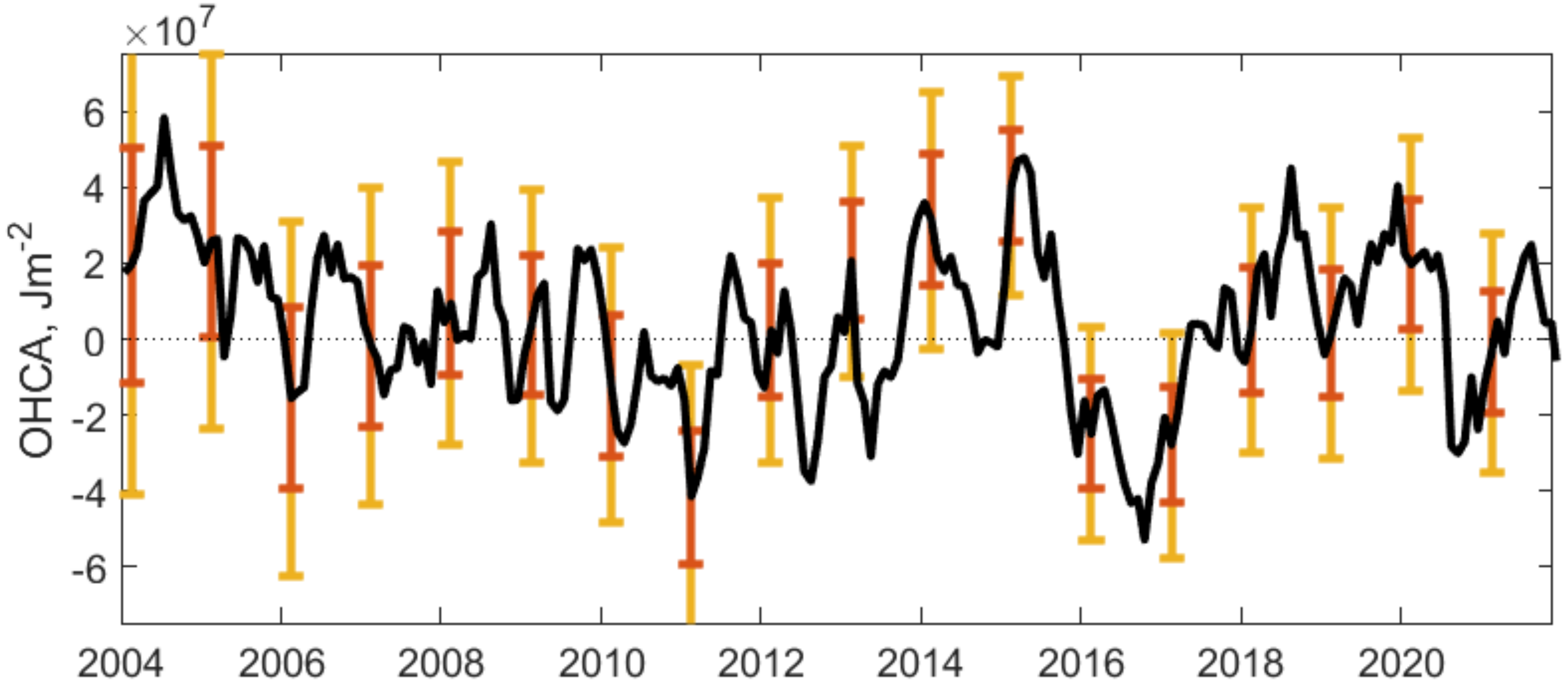
Top layer



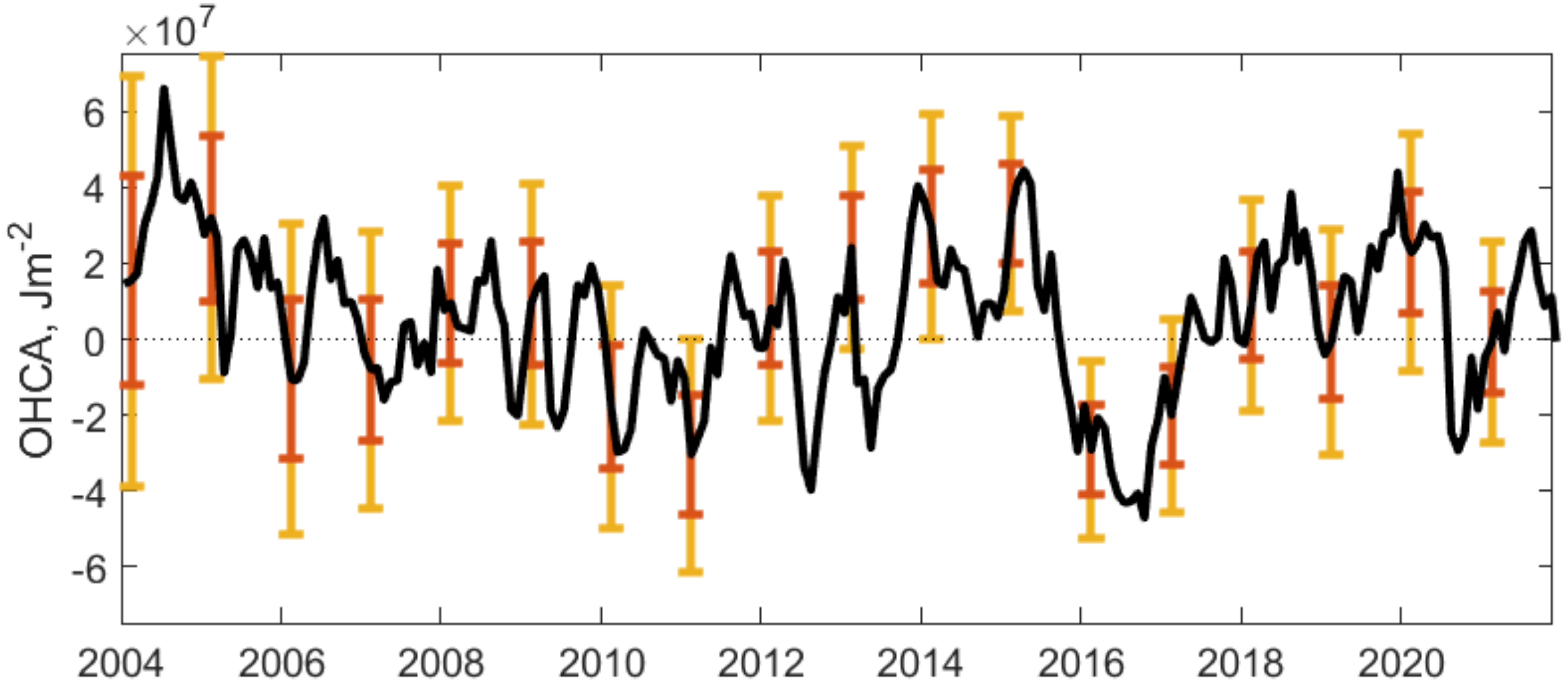
Bottom layer

Bivariate total OHC uncertainties tend to be ~15% smaller than univariate

Total OHC (global integral) anomaly estimates

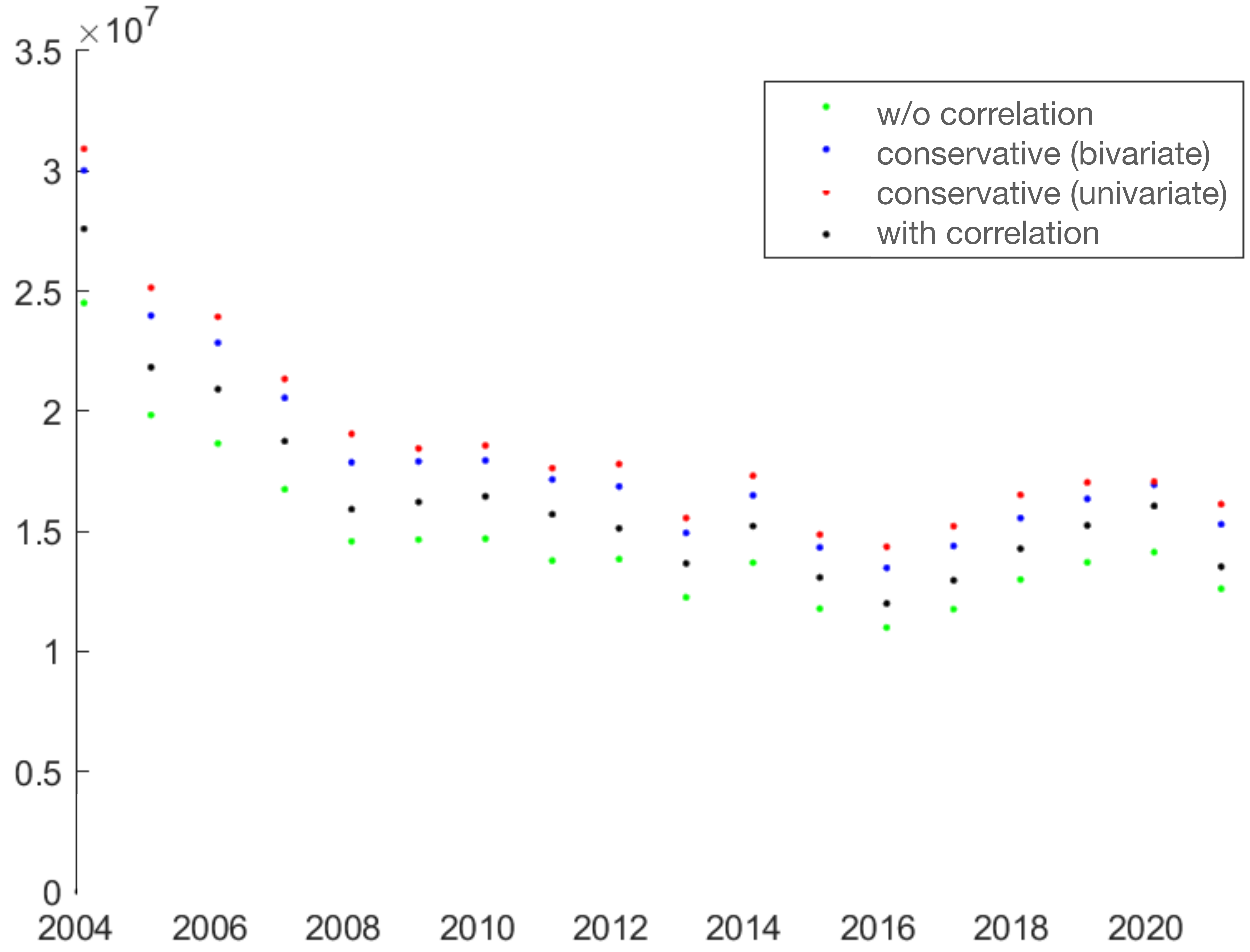


Conservative (univariate)



With correlation

Bivariate total OHC uncertainties tend to be ~15% smaller than univariate



(w/o correlation)

$$\boxed{\sqrt{\mathrm{Var}(\mathrm{OHC}_{\mathrm{top}}|\mathrm{data}) + \mathrm{Var}(\mathrm{OHC}_{\mathrm{bot}}|\mathrm{data})}}$$

(conservative)

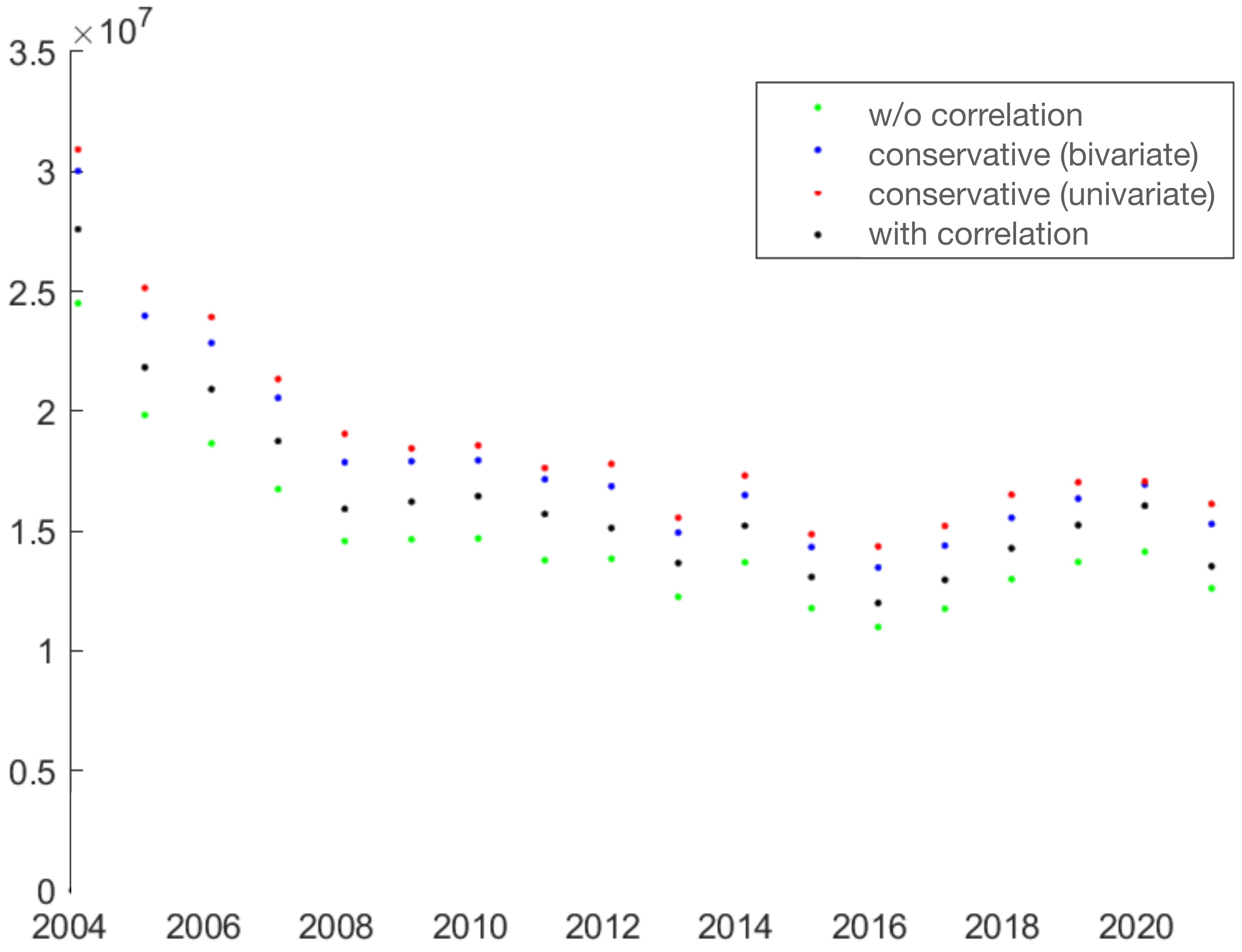
$$\boxed{\sqrt{\mathrm{Var}(\mathrm{OHC}_{\mathrm{top}}|\mathrm{data})} + \sqrt{\mathrm{Var}(\mathrm{OHC}_{\mathrm{bot}}|\mathrm{data})}}$$

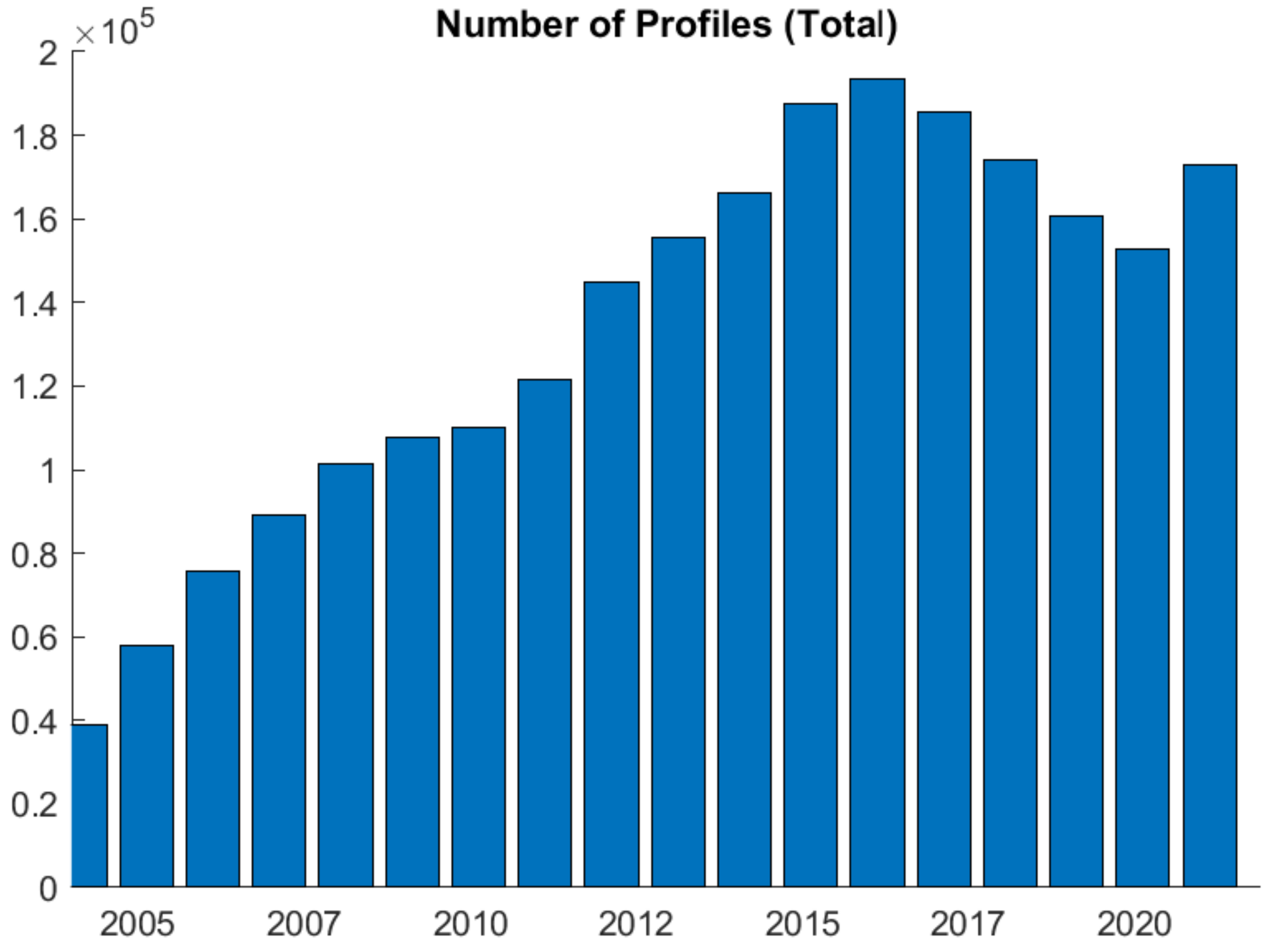
(with correlation)

$$\boxed{\sqrt{\mathrm{Var}(\mathrm{OHC}_{\mathrm{total}}|\mathrm{data})}}$$

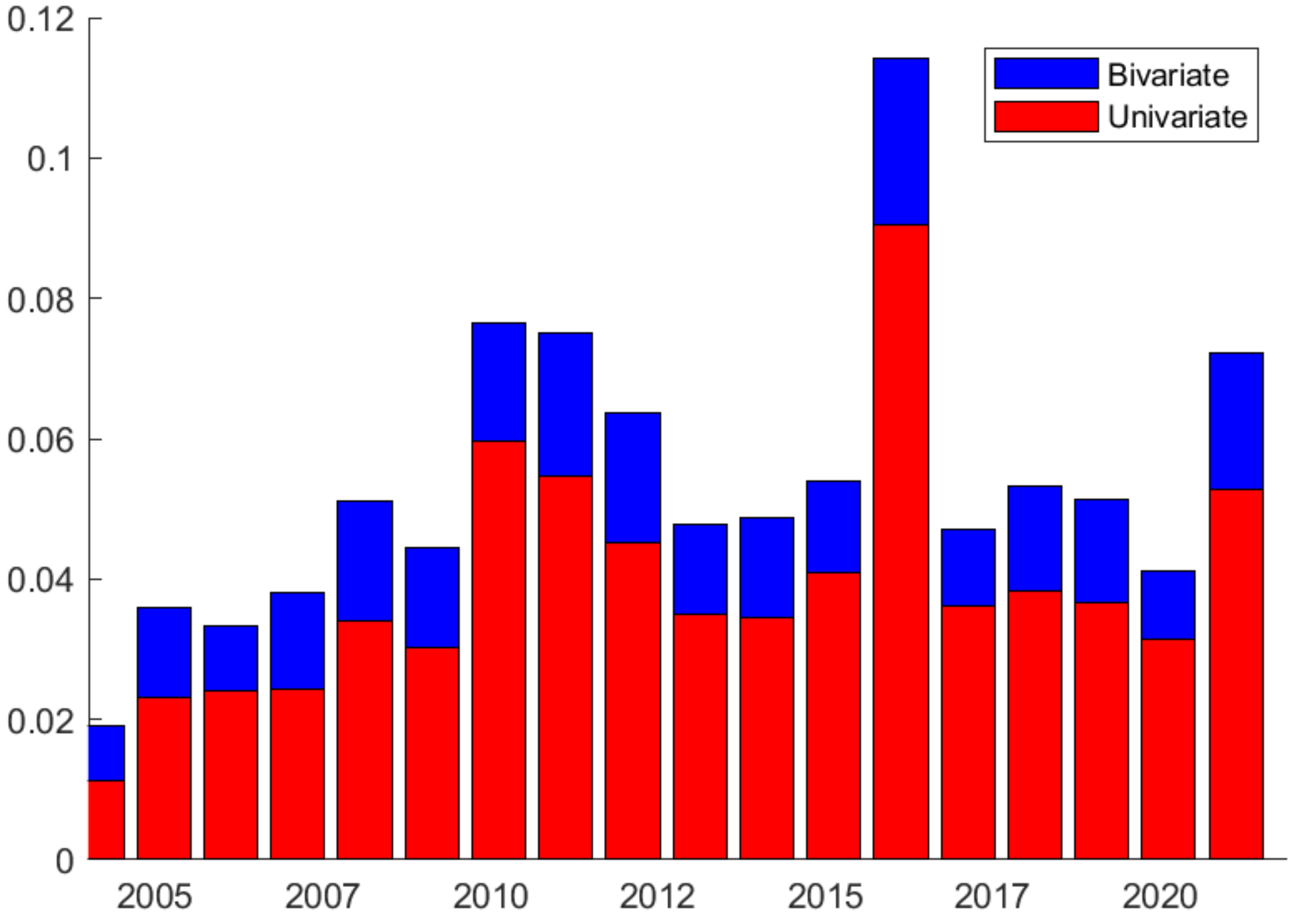
When we model the correlation, the uncertainties are **smaller** than the conservative and **larger** than those w/o the correlation.

The uncertainty trend follows the number of profiles/floats

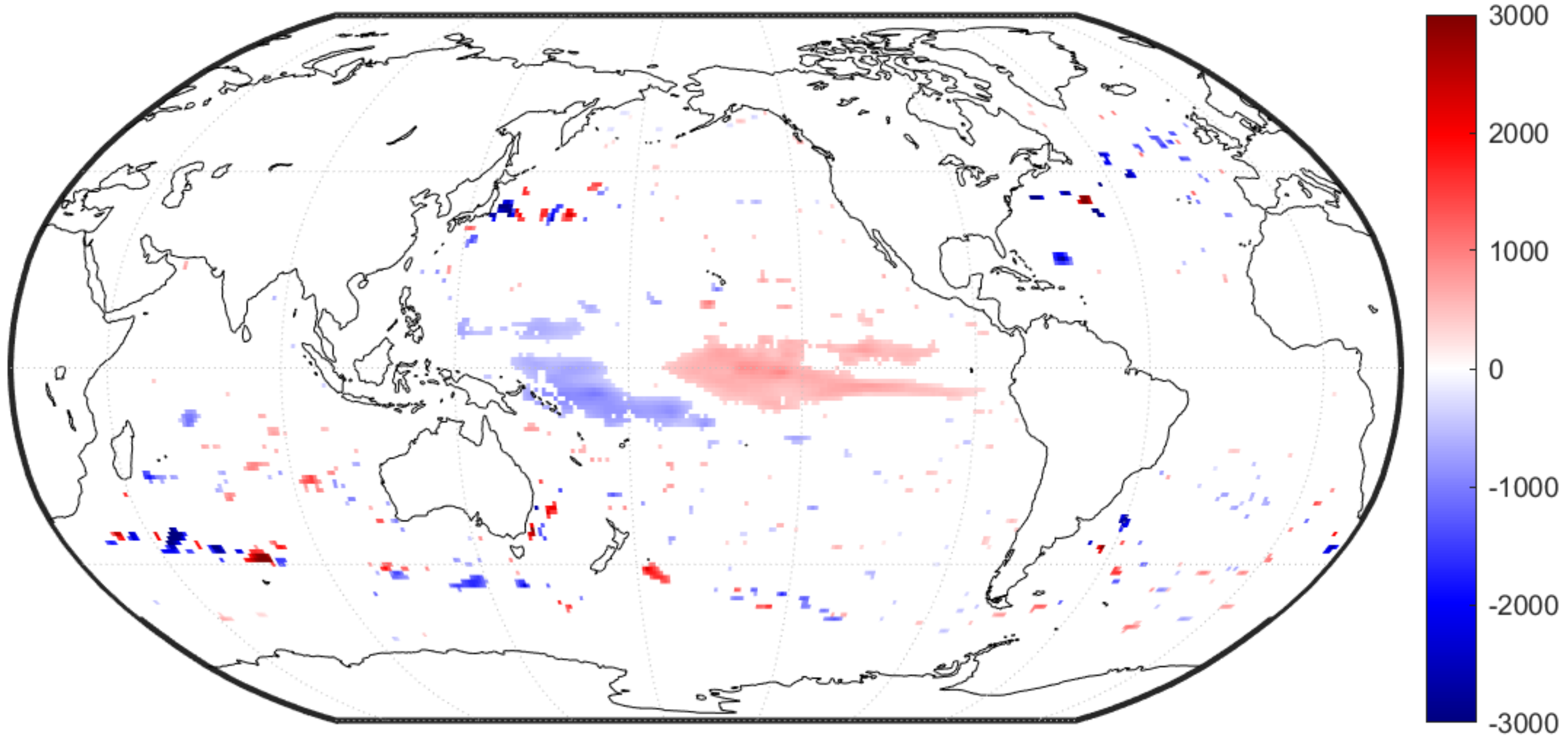




The equatorial OHC anomalies are consistently significant

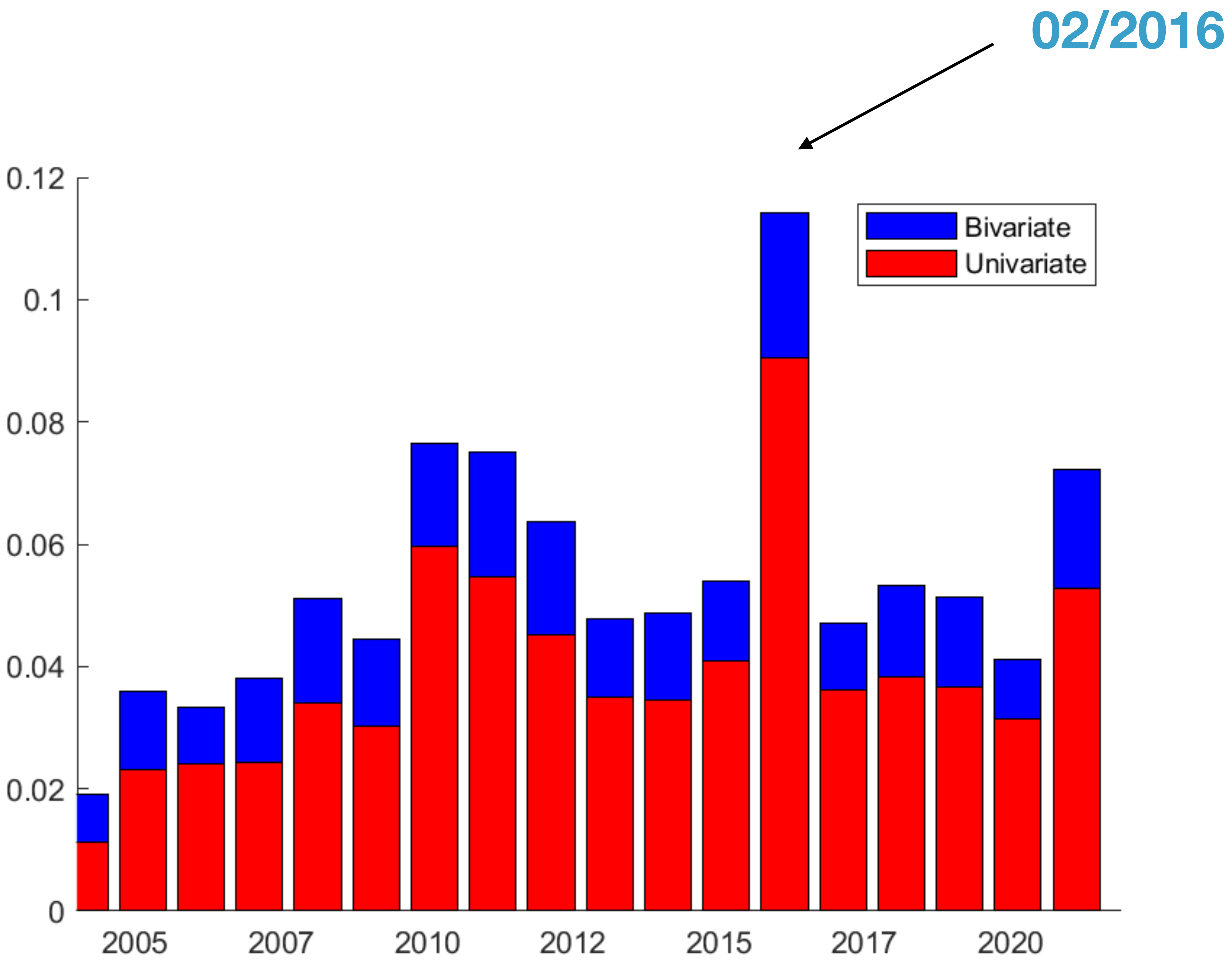


Proportion of grid points with significant anomalies over time (95% level)

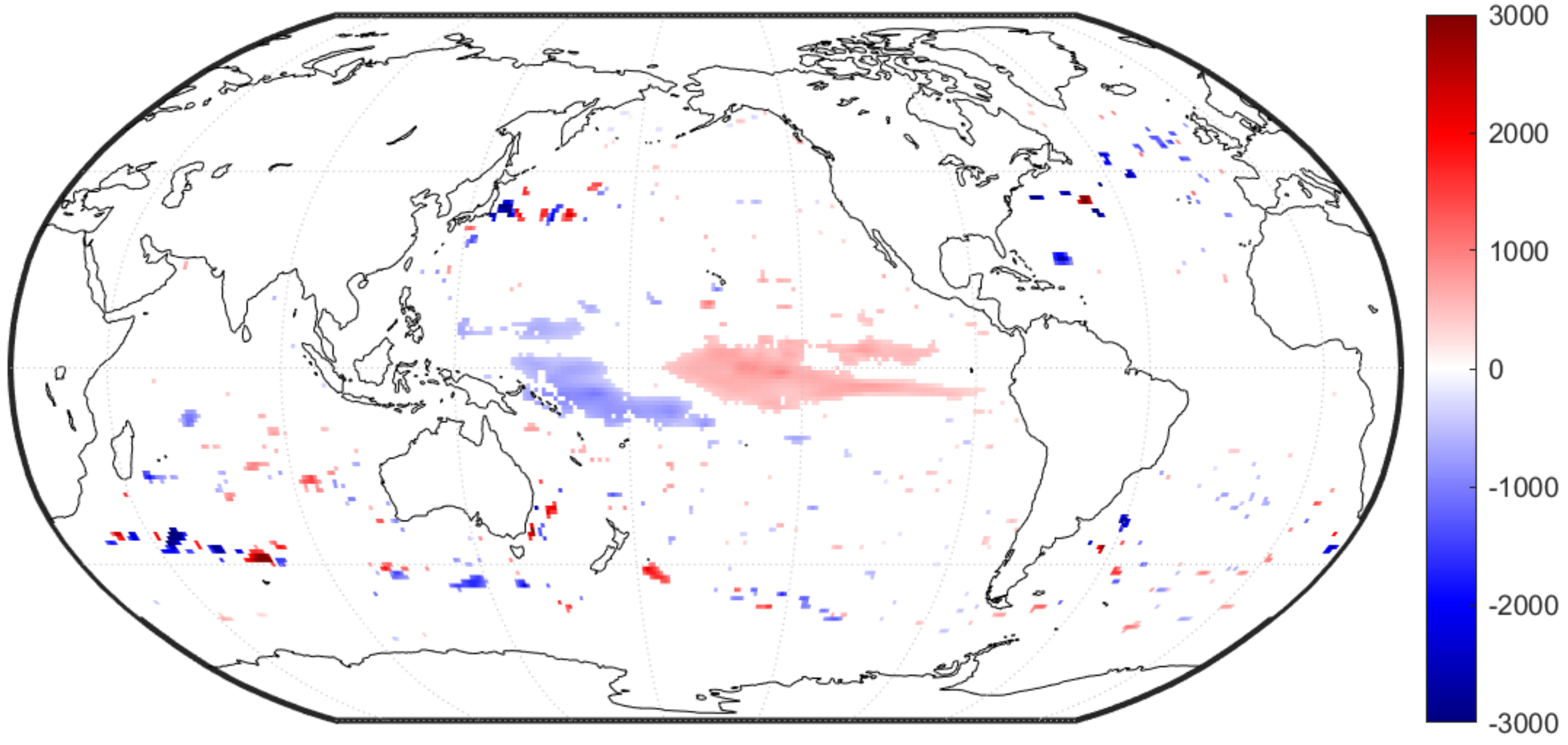


Significant temperature anomalies (02/2010)

The equatorial OHC anomalies are consistently significant

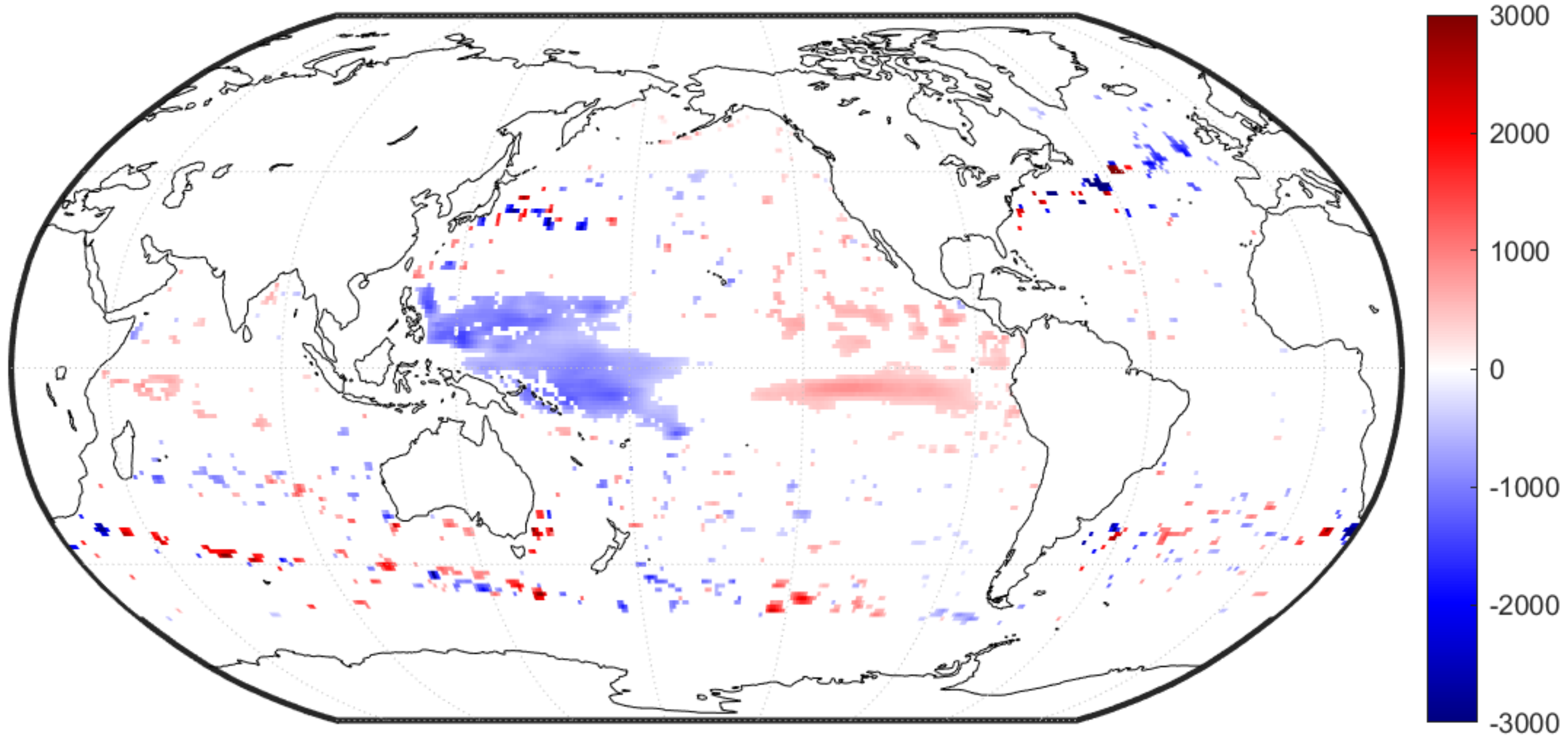


Proportion of grid points with significant anomalies over time (95% level)

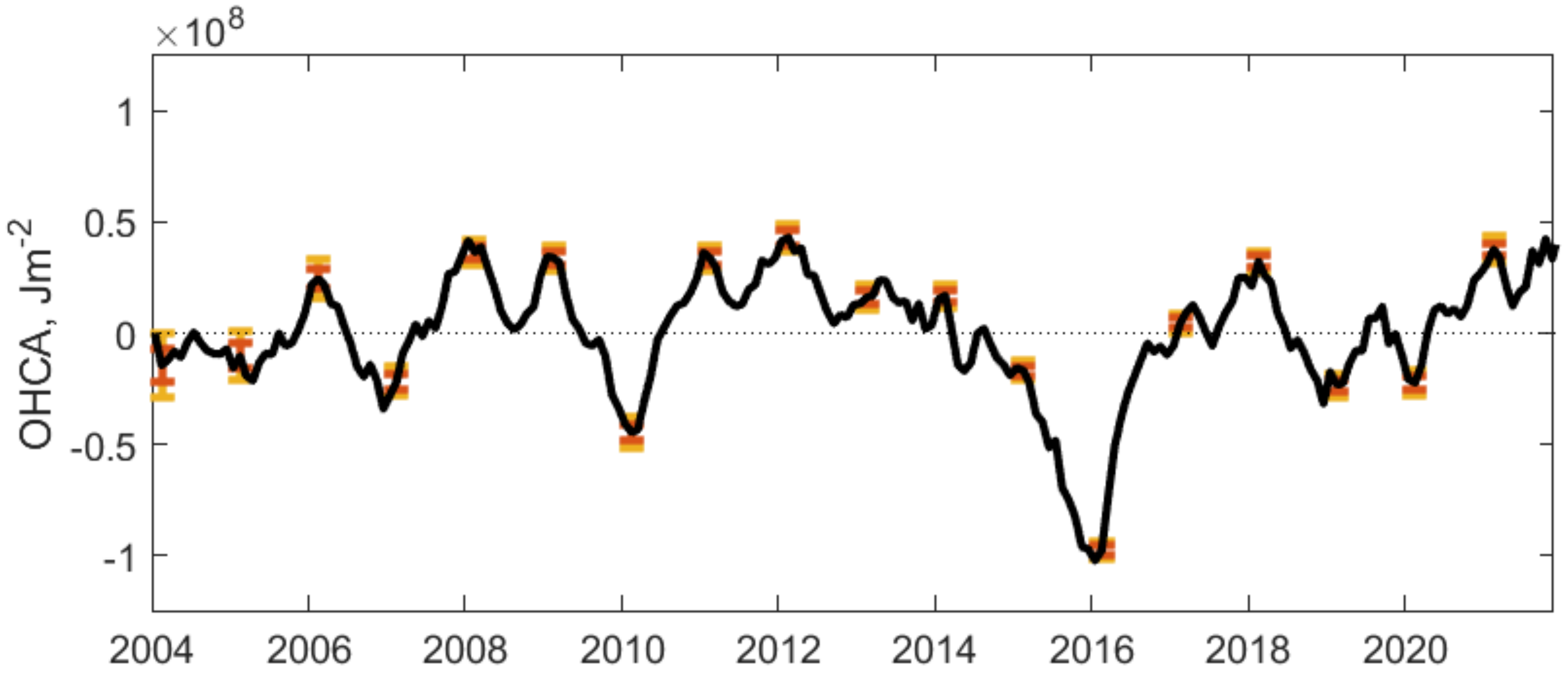


Significant temperature anomalies (02/2010)

The 2015-16 El Niño appears in the equatorial OHC anomalies

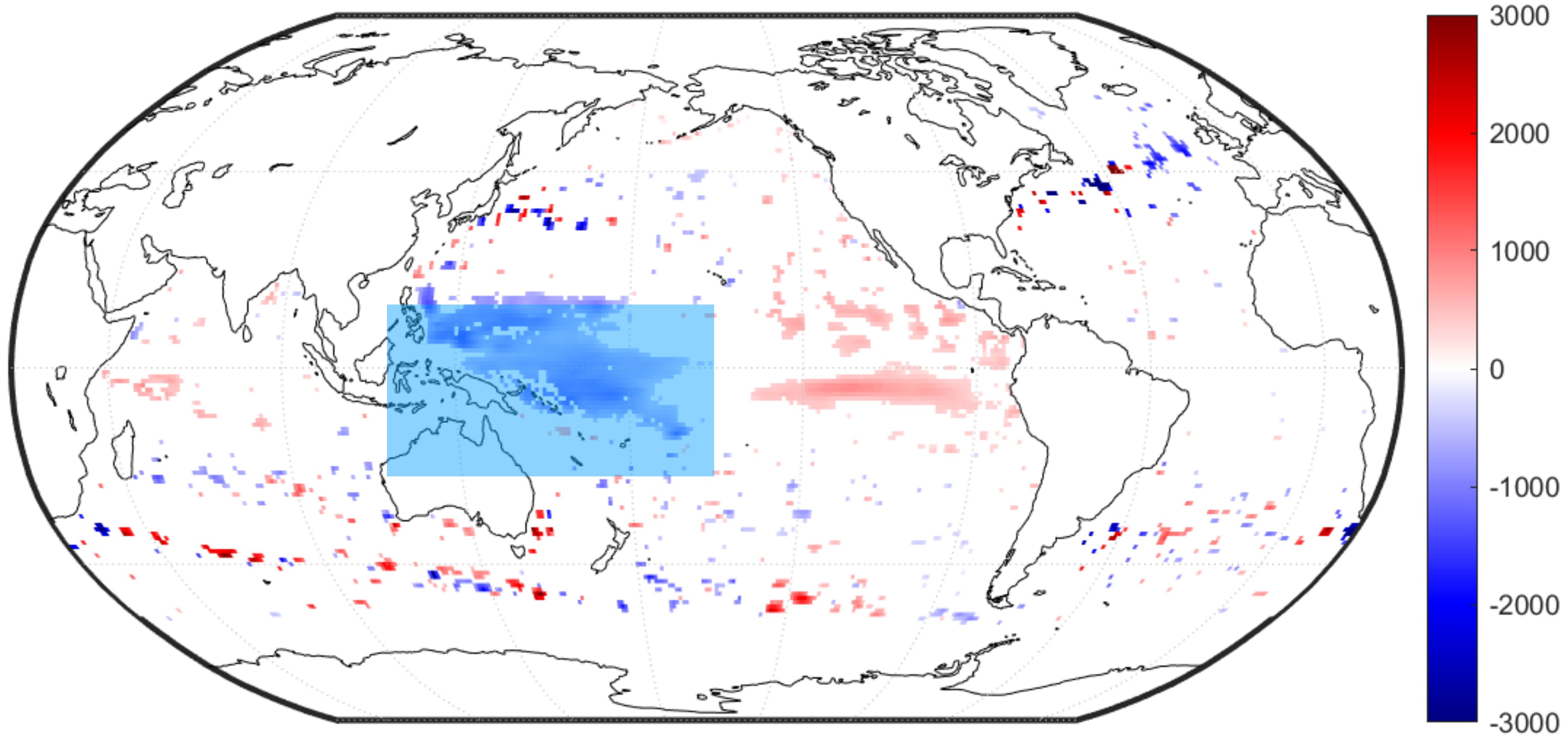


Significant temperature anomalies (02/2016)

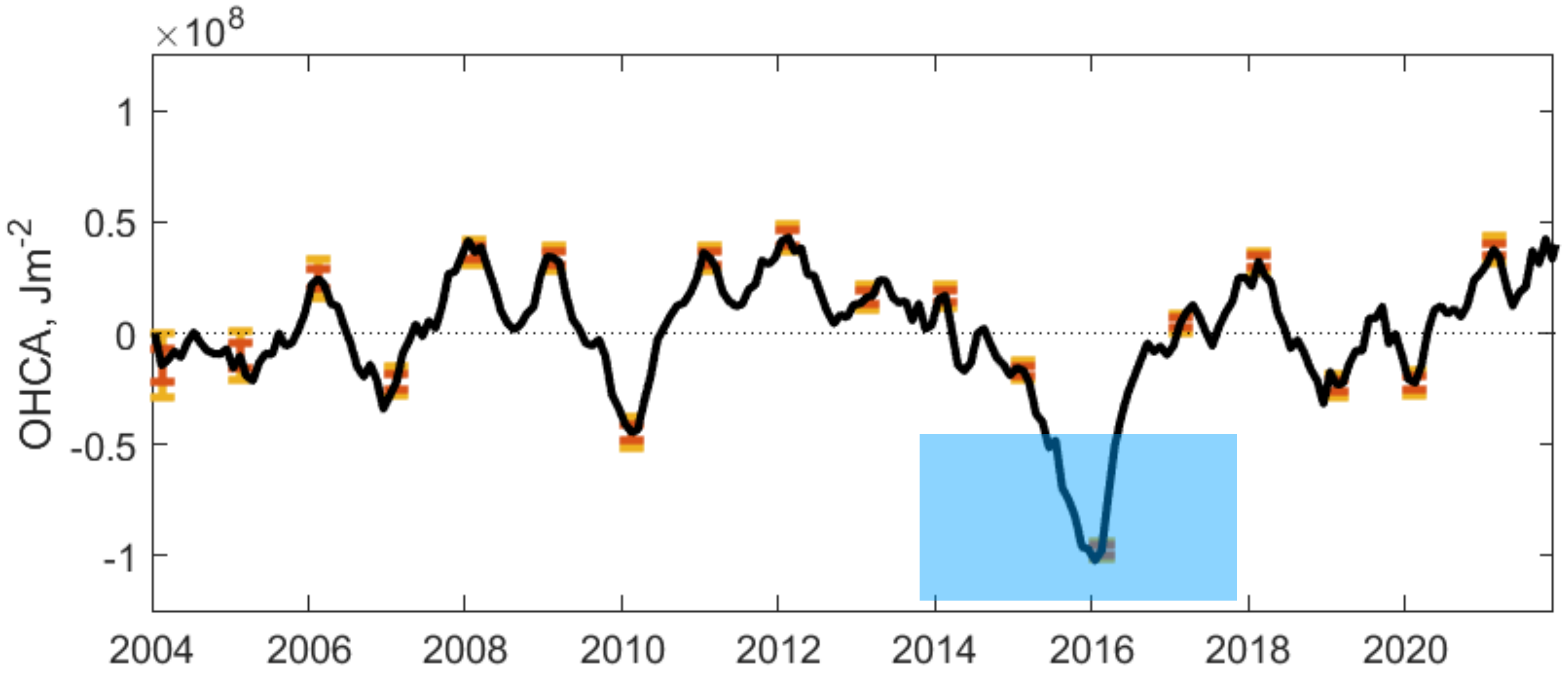


Western Equatorial Pacific OHC anomalies

The 2015-16 El Niño appears in the equatorial OHC anomalies



Significant temperature anomalies (02/2016)



Western Equatorial Pacific OHC anomalies

Future work

- Validate kriging variances and uncertainties
    - Cross-validation
    - Mapping synthetic profiles and comparing to model truth
- Account for non-Gaussianity
- Uncertainties for mean field and climatological time trend
- Generalize GP regression model to more than two layers
- Apply bivariate model to other fields (e.g. SSH and OHC, oxygen and T/S)

Summary

- Estimating ocean heat content (with reliable uncertainties) is crucial for **tracking climate change**
- Due to having fewer observations deeper in the water column, previous work has modeled the OHC in the top and bottom layers **separately**
- To model the uncertainties of the total OHC in the water column (top + bottom) we need to estimate the spatially-varying cross-layer **correlation**
- Empirically, using a bivariate GP model to estimate the correlation reduces the OHC anomaly (global integral) uncertainties both for each layer separately and ~15% for the total

Summary

- Estimating ocean heat content (with reliable uncertainties) is crucial for **tracking climate change**
- Due to having fewer observations deeper in the water column, previous work has modeled the OHC in the top and bottom layers **separately**
- To model the uncertainties of the total OHC in the water column (top + bottom) we need to estimate the spatially-varying cross-layer **correlation**
- Empirically, using a bivariate GP model to estimate the correlation reduces the OHC anomaly (global integral) uncertainties both for each layer separately and ~15% for the total

Thank you!