Spatio-Temporal Local Interpolation for Quantifying Global Ocean Heat Transport from Autonomous Observations

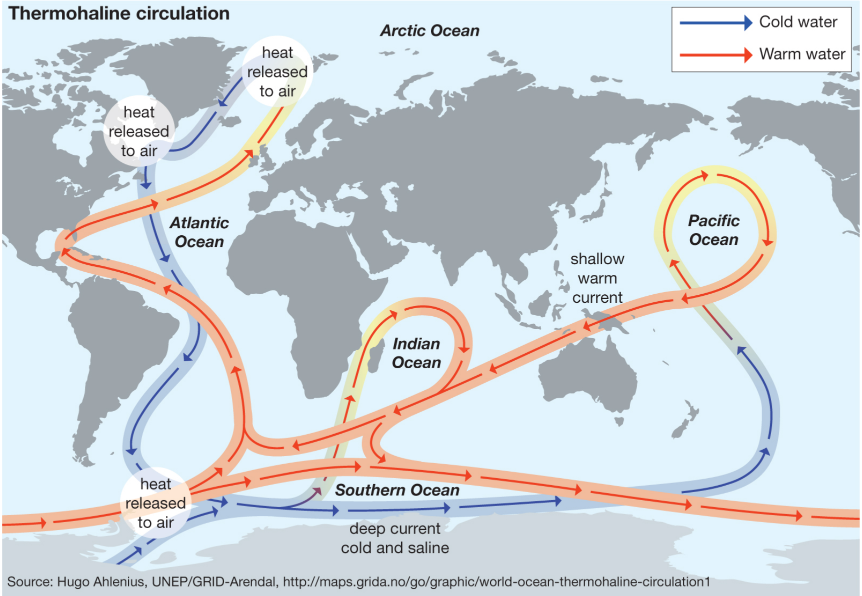
Beomjo Park[1] Mikael Kuusela[1]
Donata Giglio[2] Alison Gray[3]

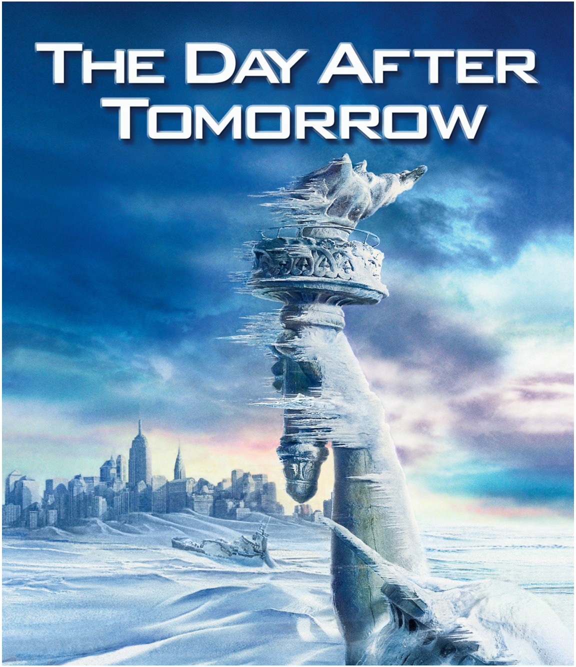
[1]Dept. of Statistics & Data Science, Carnegie Mellon University

[2]Dept. of Atmospheric and Oceanic Sciences, University of Colorado Boulder

[3]School of Oceanography, University of Washington

Aug. 6th, 2020


Thermohaline circulation
Cold water
Warm water
Arctic Ocean
heat released to air
heat released to air
Atlantic Ocean
Pacific Ocean
shallow warm current
Indian Ocean
heat released to air
Southern Ocean
deep current cold and saline
Source: Hugo Ahlenius, UNEP/GRID-Arendal, http://maps.grida.no/go/graphic/world-ocean-thermohaline-circulation1


THE DAY AFTER
TOMORROW

Ocean Heat Transport (OHT)

Temperature $\times$ **Velocity** integrated w.r.t. depth
across Latitude (Meridional) or Longitude (Zonal)

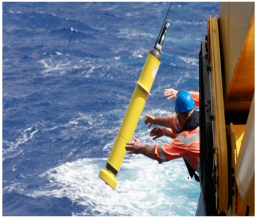
$\mathrm{OHT}(\boldsymbol{x}, t)$ at $\boldsymbol{x} = (x, y)$ where longitude x, latitude y, and time t:

$$\begin{aligned}\mathrm{OHT}(\boldsymbol{x}, t) =& C_p \int \frac{\theta(\boldsymbol{x}, t, p) \cdot \mathrm{v}(\boldsymbol{x}, t, p)}{g(\boldsymbol{x}, p)} \mathrm{d}p \\ \propto& \int \underbrace{\theta(\boldsymbol{x}, t, p)}_{\text{Temperature}} \cdot \underbrace{\mathrm{v}(\boldsymbol{x}, t, p)}_{\text{Velocity}} \mathrm{d}p\end{aligned}$$

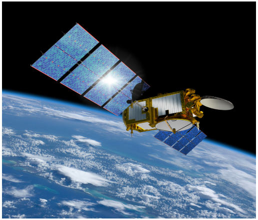
where g: gravitational acceleration, C_p: specific heat content.

OHT can be estimated from various data sources

Research Vessel

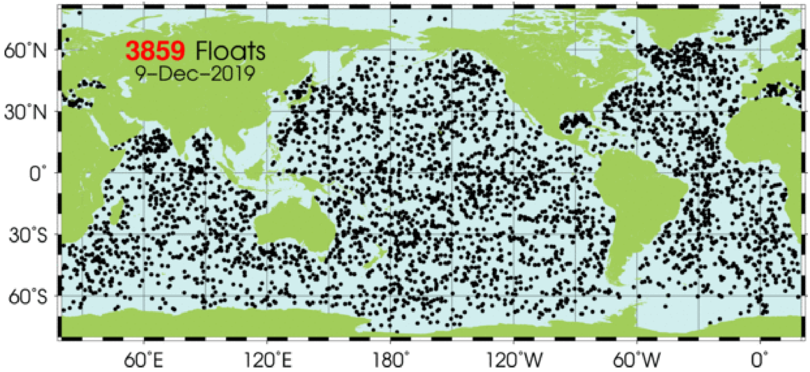


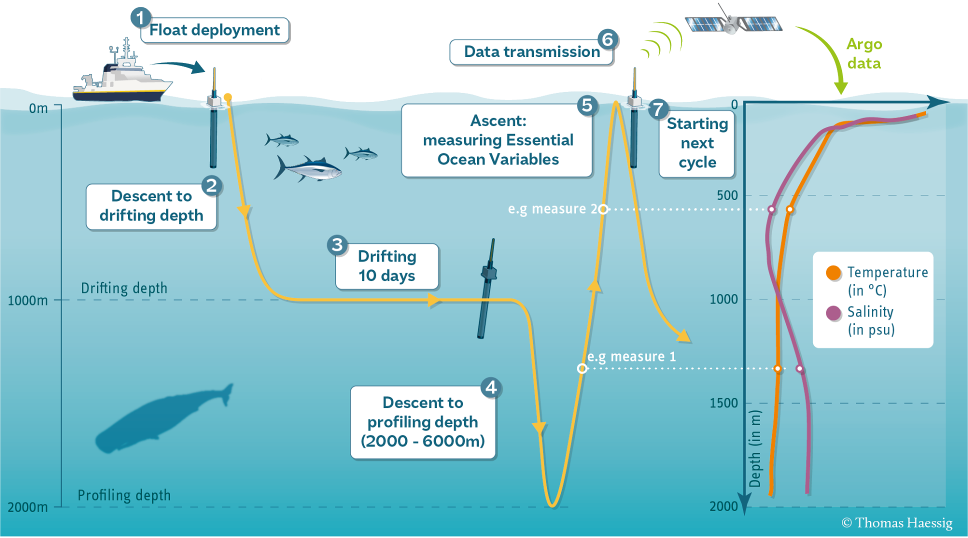
Underwater Probe



Satellite altimetry

Argo samples nearly uniform $3^\circ \times 3^\circ \times 10$ days





1 Float deployment
6 Data transmission
Argo data
0m
1000m
2000m
Drifting depth
Profiling depth
2 Descent to drifting depth
3 Drifting 10 days
4 Descent to profiling depth (2000 - 6000m)
5 Ascent: measuring Essential Ocean Variables
7 Starting next cycle
e.g measure 2
e.g measure 1
0
500
1000
1500
2000
Depth (in m)
Temperature (in °C)
Salinity (in psu)
© Thomas Haessig

Statistical Challenge

$$\mathrm{OHT}(\boldsymbol{x}, t) \propto \int \underbrace{\theta(\boldsymbol{x}, t, p)}_{\text{Temperature}} \cdot \underbrace{\mathrm{v}(\boldsymbol{x}, t, p)}_{\text{Velocity}} \, \mathrm{d}p$$

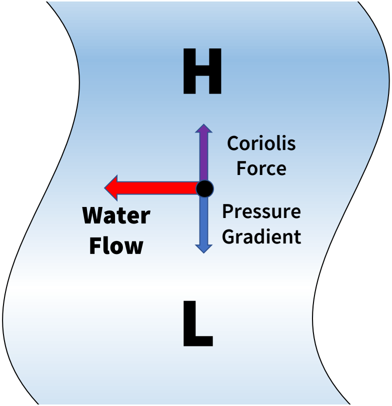
- Both are globally non-stationary spatio-temporal field
- Capture local structures from the sparse observation
- Computationally feasible method to handle massive in-situ data

Geostrophic Velocity $\mathbf{v}$

For fixed pressure p^*, relative velocity

$$\mathbf{v}_{\mathrm{rel}}(p^*) := \mathbf{v}(p^*) - \mathbf{v}_{\mathrm{ref}}(p_0) = \frac{1}{f}\mathbf{k} \times \nabla_{\boldsymbol{x}}\Psi(p^*)$$

where f: Coriolis parameter, Ψ: dynamic height anomaly $\left(\int_{p^*}^{p_{\mathrm{ref}}} \frac{1}{\rho}\mathrm{d}p\right)$



Modelling Ψ

Consider an additive model of $\{\Psi(\boldsymbol{x}, t)\}$ indexed by location $\boldsymbol{x} = (x, y)$ where latitude x, longitude y in degrees, and time t in days.

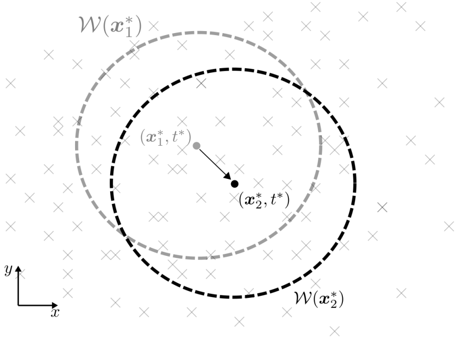
$$\Psi_{p^*}(\boldsymbol{x}, t) = \underbrace{m_{p^*}(\boldsymbol{x}, t)}_{\text{Mean Field}} + \underbrace{a_{p^*}(\boldsymbol{x}, t)}_{\text{Anomaly Field}} + \underbrace{\epsilon(\boldsymbol{x}, t)}_{\text{Nugget Effect}} \tag{1}$$

where $\mathbb{E}\Psi(\boldsymbol{x}, t) = m(\boldsymbol{x}, t)$, and $a(\boldsymbol{x}, t)$ is zero-mean, second-order stationary random field.

Modelling Ψ: Local Semiparametric Regression[1]

Within a small spatial window $\mathcal{W}(\boldsymbol{x}^*) = \{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{x}^*\| \leq \lambda_{\boldsymbol{x}}\}$,

$$m(\boldsymbol{x}, t) = \beta_0 + [\text{ 1st- and 2nd-order linear terms of } x \text{ and } y\,]$$
$$+ \sum_{l=1}^{L} \left[\beta_{c_l} \cos\left(\frac{2\pi l}{365} t\right) + \beta_{s_l} \sin\left(\frac{2\pi l}{365} t\right) \right], \qquad \forall \boldsymbol{x} \in \mathcal{W}(\boldsymbol{x}^*)$$



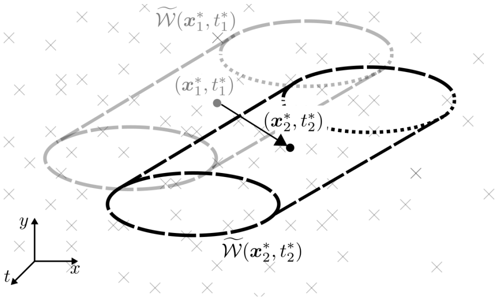
[1] Ridgway, K. R., Dunn, J. R., & Wilkin, J. L. (2002). Ocean interpolation by four-dimensional weighted least squares—Application to the waters around Australasia. *Journal of Atmospheric and Oceanic Technology*, 19(9), 1357–1375.

Modelling Ψ: Local Semiparametric Regression[2]

Within a small spatiotemporal window $\widetilde{\mathcal{W}}(\boldsymbol{x}^*, t^*) = \mathcal{W}(\boldsymbol{x}^*) \times [t^* \pm \lambda_t]$,

$$a_i \overset{\text{iid}}{\sim} \mathrm{GP}\left(0, k\left((\boldsymbol{x}_1, t_1), (\boldsymbol{x}_2, t_2); \boldsymbol{\xi}\right)\right), \qquad \text{year } i = 1, \cdots, I$$

for observations $j = 1, \ldots, J_i$ within $\widetilde{\mathcal{W}}(\boldsymbol{x}^*, t^*)$, where $k(\,\cdot\,; \boldsymbol{\xi})$ is an space-time Matern covariance function depending on parameters $\boldsymbol{\xi}$.



[2] Kuusela, M., & Stein, M. L. (2018). Locally stationary spatio-temporal interpolation of Argo profiling float data. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 474(2220)

Preditive v

Geostrophic velocity where the yearday time t is in year i:

$$\mathrm{v}_{\mathrm{rel}}(\boldsymbol{x}, t) = \frac{\mathbf{k} \times [\nabla_{\boldsymbol{x}} \Psi(\boldsymbol{x}, t) \mid \boldsymbol{\Psi}_i, \boldsymbol{\beta}, \boldsymbol{\xi}, \sigma]}{f(y) g(\boldsymbol{x})} \tag{2}$$

Joint process $[a_i, \nabla_{\boldsymbol{x}} a_i]$ is a multivariate Gaussian Process for year i:

$$\begin{bmatrix} a_i \\ \nabla_{\boldsymbol{x}} a_i \end{bmatrix} \overset{iid}{\sim} \mathrm{GP}\left(\mathbf{0}, \begin{bmatrix} k(\boldsymbol{s}, \boldsymbol{s}) & \nabla_{\boldsymbol{x}^*} k(\boldsymbol{s}, \boldsymbol{s}^*)^\top \\ \nabla_{\boldsymbol{x}} k(\boldsymbol{s}, \boldsymbol{s}^*) & \nabla_{\boldsymbol{x}} \nabla_{\boldsymbol{x}^*} k(\boldsymbol{s}, \boldsymbol{s}^*) \end{bmatrix}\right)$$

implies the Gaussian predictive distribution of $\nabla_{\boldsymbol{x}} \Psi$ with

$$\mathbb{E}(\nabla_{\boldsymbol{x}} \Psi^\star | \boldsymbol{\Psi}_i, \boldsymbol{\beta}, \boldsymbol{\xi}, \sigma) = \boldsymbol{\beta}(\boldsymbol{x}^\star) + \nabla_{\boldsymbol{x}} \boldsymbol{k}_i^\star(\boldsymbol{\xi})^\top (\boldsymbol{K}_i(\boldsymbol{\xi}) + \sigma^2 \boldsymbol{I})^{-1} [\boldsymbol{\Psi} - \boldsymbol{m}]_i$$

under the Gaussian nugget $\epsilon_{ij} \sim \mathrm{N}(0, \sigma^2)$.

Parameter Estimation: EM procedure

For spatio-temporal grid points $\{(\boldsymbol{x}^*, t^*) : \boldsymbol{x}^* \in \mathcal{X}, t^* \in [0, 365]\}$,

$$\log \mathcal{L}(\beta(\boldsymbol{x}^*), \boldsymbol{\xi}(\boldsymbol{x}^*, t^*)) = \sum_{i=1}^{I} \log \mathrm{N}(\boldsymbol{\Psi}_i; \tilde{\boldsymbol{\eta}}_{i\cdot}^\top \boldsymbol{\beta}, \boldsymbol{K}_i(\boldsymbol{\xi}))$$

where $\tilde{\boldsymbol{\eta}}_{ij}$ is the $\sum_{l=1}^{i-1} n_l + j$th column of the design matrix.

For iteration $\tau = 0, 1, \ldots,$

$$\begin{aligned} \boldsymbol{\beta}^{(\tau+1)} &= \mathrm{argmax}_{\boldsymbol{\beta}} \log \tilde{\mathcal{L}}(\boldsymbol{\beta} | \boldsymbol{\xi}^{(\tau)}), && \text{(E step)} \\ \boldsymbol{\xi}^{(\tau+1)} &= \mathrm{argmax}_{\boldsymbol{\xi}} \log \mathcal{L}(\boldsymbol{\xi} | \boldsymbol{\beta}^{(\tau+1)}), && \text{(M step)} \end{aligned}$$

where $\tilde{\mathcal{L}}$ is an approximated likelihood of $\mathcal{L}$ with Vecchia approximation

Debiasing Mean-field Misspecification

Suppose the mean-field m is mis-specified:
the anomaly field includes **systematic bias** $B(\boldsymbol{x}) = \mathbb{E}\left[a(\boldsymbol{x}, t)\right] \neq 0$

$$\Psi(\boldsymbol{x}, t) = [m(\boldsymbol{x}, t) + B(\boldsymbol{x})] + [a(\boldsymbol{x}, t) - B(\boldsymbol{x})] + \epsilon$$

We may estimate B from the conditional mean of predictive Ψ,

$$\mathbb{E}[a(\boldsymbol{x}^\star, t^\star)] \stackrel{p}{\leftarrow} \frac{1}{I}\sum_{i=1}^{I} \hat{a}_i(\boldsymbol{x}^\star, t^\star) \approx \frac{1}{I}\sum_{i=1}^{I}\left[\frac{1}{J_i}\sum_{j=1}^{J_i} \hat{a}_i(\boldsymbol{x}^\star, t^\star_{ij})\right] := \widehat{B}(\boldsymbol{x}^\star)$$

Consequently,

$$\nabla_{\boldsymbol{x}} \widehat{B}(\boldsymbol{x}^\star) = \frac{1}{I}\sum_{i=1}^{I}\left[\frac{1}{J_i}\sum_{j=1}^{J_i} \nabla_{\boldsymbol{x}} \boldsymbol{k}_i^\star(\boldsymbol{\xi}_{(j)})^\top (\boldsymbol{K}_i(\boldsymbol{\xi}_{(j)}) + \sigma^2_{(j)}\boldsymbol{I})^{-1}\left[\boldsymbol{\Psi} - \hat{\boldsymbol{m}}\right]_{ij}\right]$$

NCAR
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

5.34-petaflops, High Performance Cluster
with 145,152 Intel Xeon processors (36 cores/node)

Local model constructs the velocity field from sparse observations

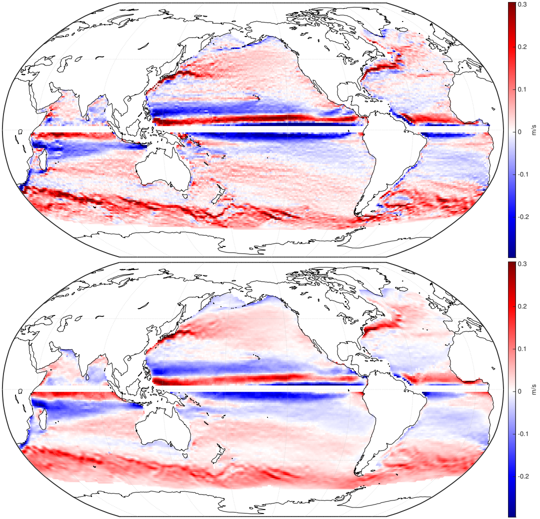
(East / West) Zonal velocity

Satellite

(Surface)

Argo

(10 m, Debiased)

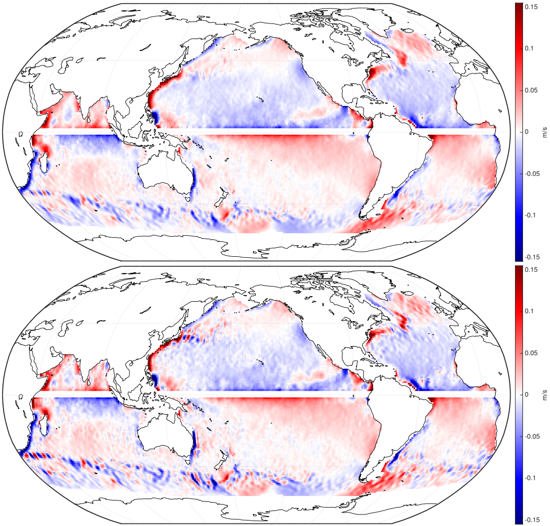


Debiasing procedure captures higher-order features

(North / South) Meridional Velocity at 10 m

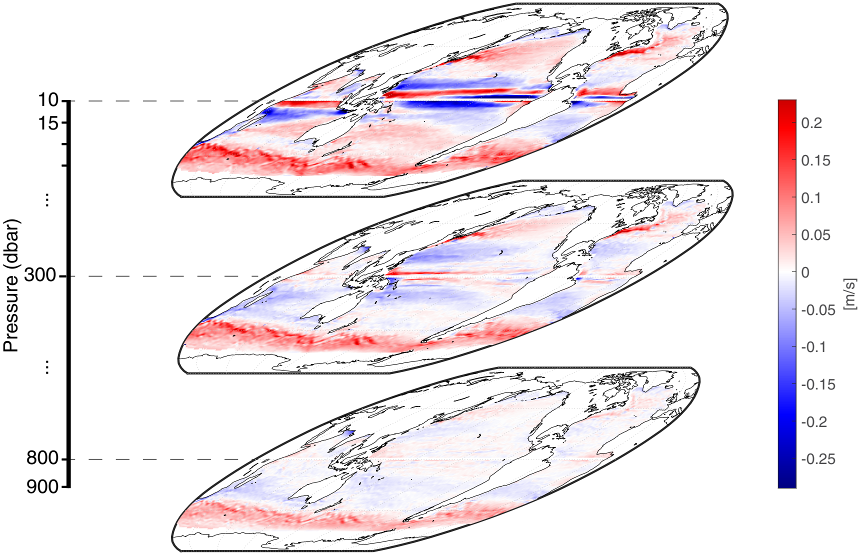
Initial

Debiased

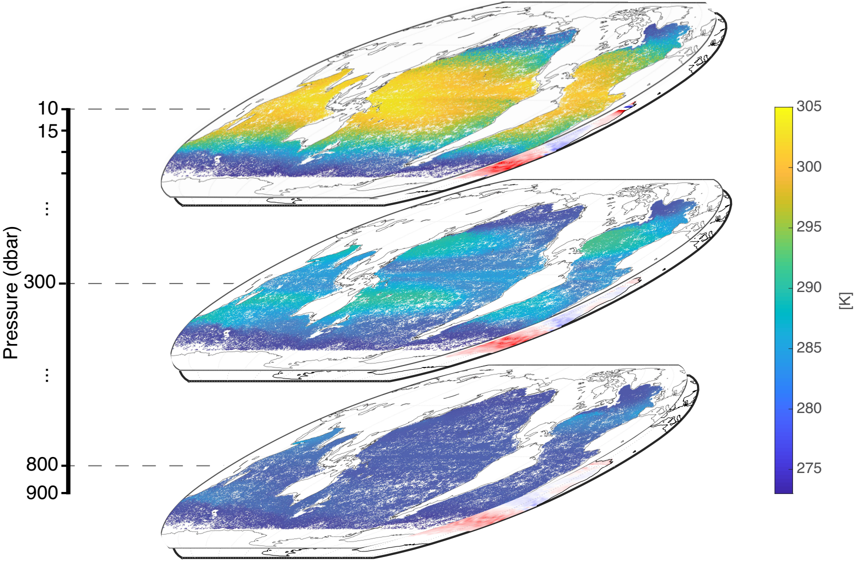


Satellite Altimetry

Argo provides velocity field at multiple depths



Argo provides in-situ temperature at multiple depths



OHT Field is mapped by Local Regression with plug-in estimator

Estimated OHT at any given location $\boldsymbol{x}_{ij}$ and time t_{ij}:

$$\begin{aligned}
\mathrm{OHT}(\boldsymbol{x}_{ij}, t_{ij}) &\propto \int \theta_{ij}(p) \cdot \mathbf{v}_{ij}(p)\, \mathrm{d}p \\
&\approxeq \int \theta_{ij}(p) \cdot \hat{\mathbf{v}}_{ij}(p)\, \mathrm{d}p \qquad \Big(\hat{\mathbf{v}} = \mathbb{E}(\mathbf{v} | \boldsymbol{\Psi}_i, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\theta}}, \hat{\sigma})\Big) \\
&\approxeq \sum_{k=0}^{N_{\mathrm{int}}} \widetilde{\theta\hat{\mathbf{v}}}(\boldsymbol{x}_{ij}, t_{ij}, p_k) \Delta_{p_k} := \widetilde{\mathrm{OHT}}(\boldsymbol{x}_{ij}, t_{ij})
\end{aligned}$$

where $\widetilde{\theta\hat{\mathbf{v}}}$ is the piecewise cubic Hermite interpolant (PCHIP).

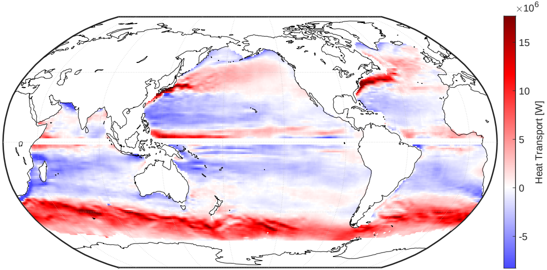
We map the OHT field again with Local Semiparametric Regression:

$$\widetilde{\mathrm{OHT}}(\boldsymbol{x}_{ij}, t_{ij}) = \underbrace{\widetilde{m}(\boldsymbol{x}_{ij}, t_{ij})}_{\text{Mean Field}} + \underbrace{\widetilde{a}_i(\boldsymbol{x}_{ij}, t_{ij})}_{\text{Anomaly Field}} + \underbrace{\epsilon_{ij}}_{\text{Nugget Effect}}$$

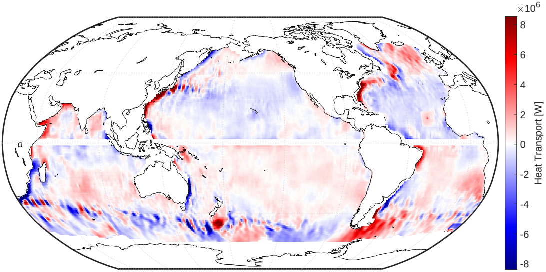
Upper-Ocean† Heat Transport mean field

† Upper Ocean: 10 to 900 dbar

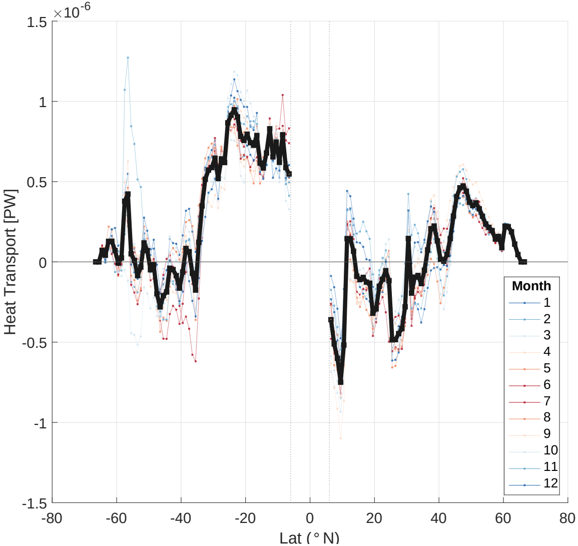
Zonal



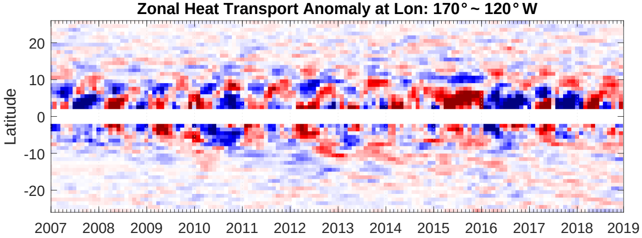
Meridional

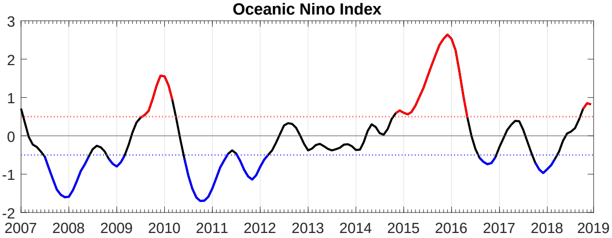


Upper-Ocean† mean Meridional Heat Transport driven by Geostrophy



Anomalous Heat Transport has connection to the anomalous climate phenomena, i.e., El Niño.





Map Area

Discussion

Statistical Aspects

- Parameter Tuning
  - Seasonal cycle harmonics
  - Bandwidth selection for spatio-temporal windows
- Uncertainty Quantification
  - Incorporate $\mathbb{V}(\mathbf{v}|\boldsymbol{\Psi}_i, \hat{\beta}, \hat{\theta}, \hat{\sigma})$ to the 2nd stage regression
  - Global confidence band for Two-stage estimate
- Relaxing the assumptions
  - Mapping the vertical dimension (4D map)
  - Multivariate joint process $(\boldsymbol{\theta}, \mathbf{v})$
  - Beyond Gaussian field

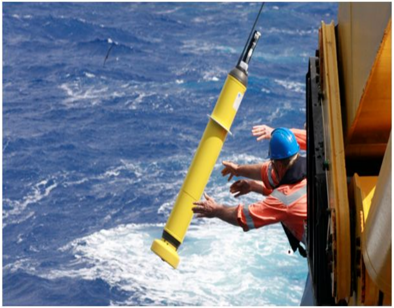
Oceanographic / Climatological Aspects

- Ocean contribution to Meridional Heat Transport
  - Other sources of heat transport, i.e., Ekman
  - Impact of Ageostrophic transport

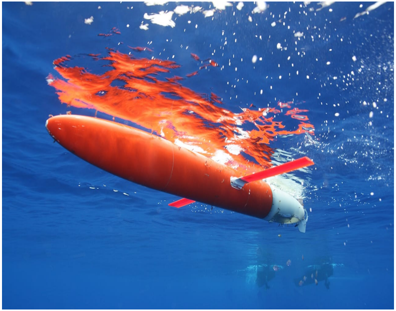
Thank You

Spray Glider

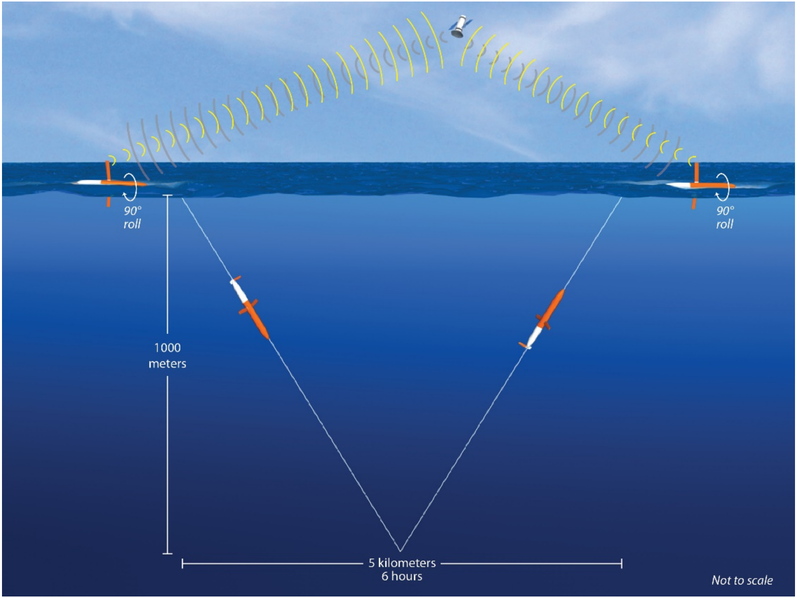
Autonomous Underwater Observations



Argo float

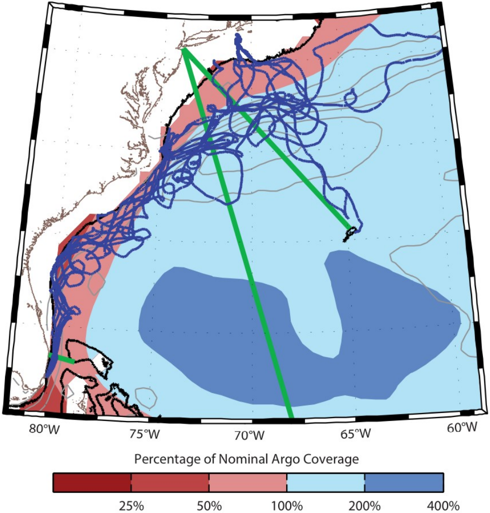


Spray glider



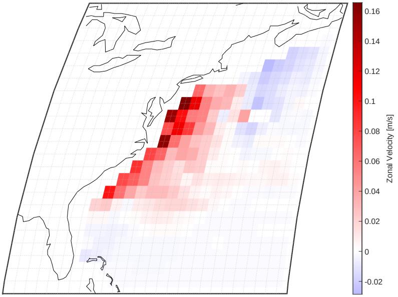
1 Woods Hole Oceanographic Institution

Gliders fill under-coverage along the US East Coast



[1] Todd, Robert E. and Locke-Wynn, Lea (2017). Underwater Glider Observations and the Representation of Western Boundary Currents in Numerical Models. *Oceanography*, 30(2), 88-89

Spray profiles correct under-estimated velocity of Gulf Stream

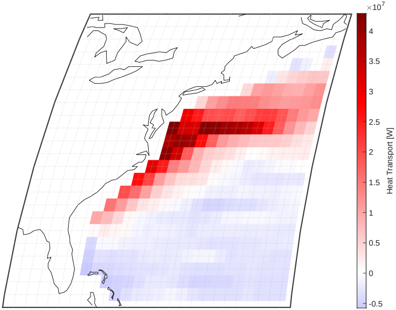


(a) Zonal (b) Meridional

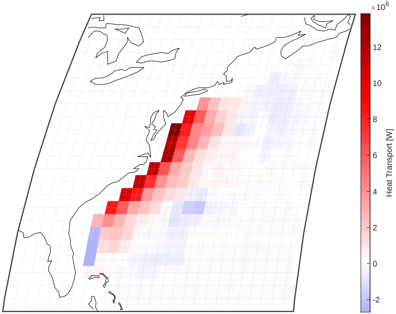
Figure: Mean velocity differences (Debiased) at 15 dbar

Mean Field

Spray profiles correct under-estimated Heat Transport of Gulf Stream



(a) Zonal

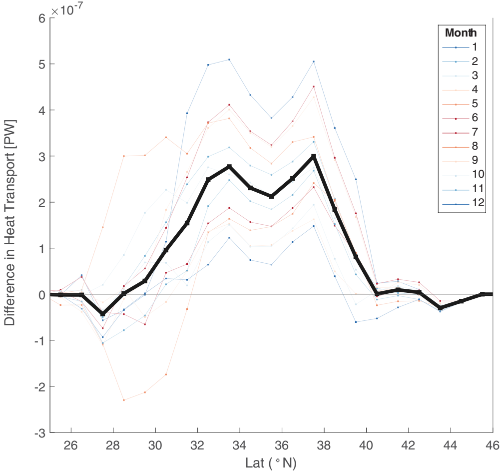


(b) Meridional

Figure: Mean absolute heat transport differences over 10 to 900 dbar

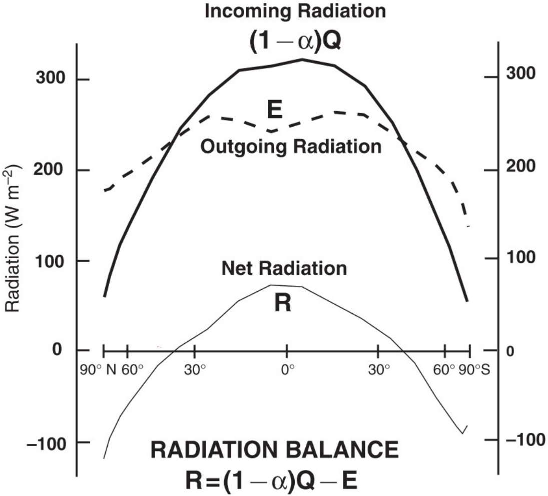
▸ Mean Field

Spray profiles correct under-estimated Meridional Heat Transport of Gulf Stream



▸ MHT

Earth’s Radiation Balance



[1] L. Bryden, H., & Imawaki, S. (2001). Chapter 6.1 Ocean heat transport., *International Geophysics* (pp. 455–474).

Choice of GP Kernel

Matèrn covariance function with smoothing parameter $\nu = 3/2$

$$k\left((\boldsymbol{x}_1, t_1), (\boldsymbol{x}_2, t_2); \xi\right) = \phi\left(1 + \sqrt{3}d\right)\exp\left(-\sqrt{3}d\right),$$
$$d\left((\boldsymbol{x}_1, t_1), (\boldsymbol{x}_2, t_2)\right) = \|(\boldsymbol{x}_1, t_1) - (\boldsymbol{x}_2, t_2)\|_{\mathrm{diag}(1/\xi)}$$

The partial derivative and hessian of the covariance function is given as

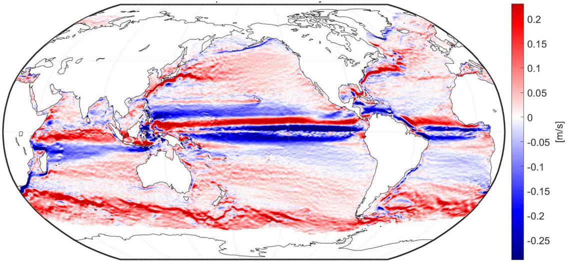
$$\frac{\partial k(\boldsymbol{x}^*, \boldsymbol{x}_i)}{\partial x^*} = \frac{-3\phi}{\xi_x^2}(x^* - x_i)\exp\left(-\sqrt{3}d\right)$$
$$\frac{\partial k(\boldsymbol{x}^*, \boldsymbol{x}_i)}{\partial y^*} = \frac{-3\phi}{\xi_y^2}(y^* - y_i)\exp\left(-\sqrt{3}d\right)$$
$$\frac{\partial^2}{\partial x_1 \partial x_2}k(x_1, x_2) = \frac{3\phi}{\xi_x^2}\left(1 - \sqrt{3}\frac{\Delta_x^2}{d\xi_x^2}\right)\exp(-\sqrt{3}d)$$

where $\Delta_x = x_1 - x_2$

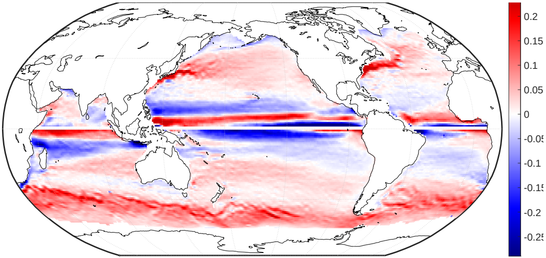
▸ Ψ Model

Surface Geostrophic Velocity from Satellite Altimetry

Satellite:



Debiased:

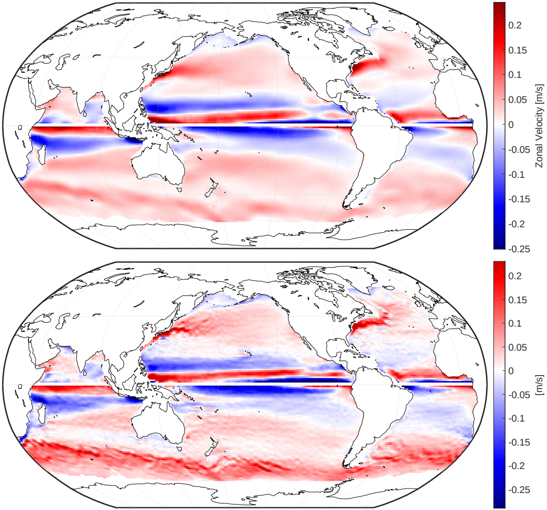


▸ Argo

Sub-surface mean Zonal Velocity field at 10 dbar

Initial

Debiased



Satellite Altimetry

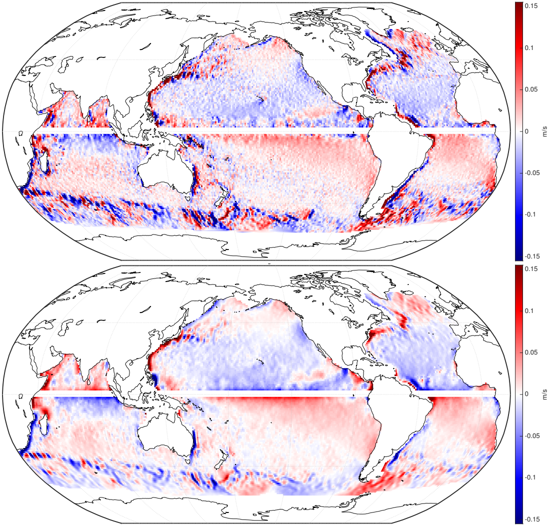
Geostrophic Velocity from Satellite Altimetry

Satellite

(Surface)

Argo

(10dbar, debiased)

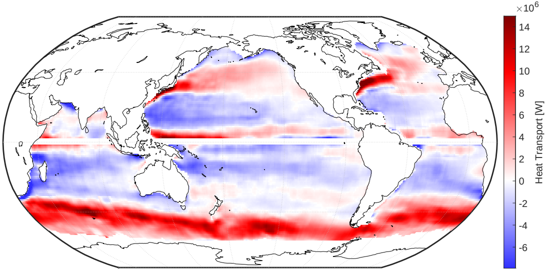


Argo

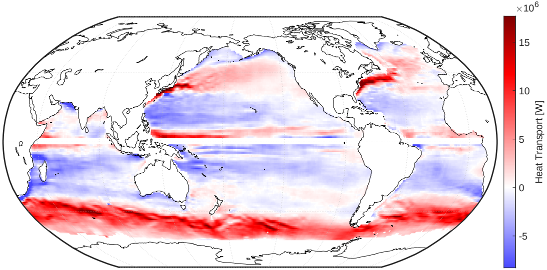
Upper-Ocean† mean Zonal Heat Transport field

† Upper Ocean: 10 to 900 dbar

Initial:



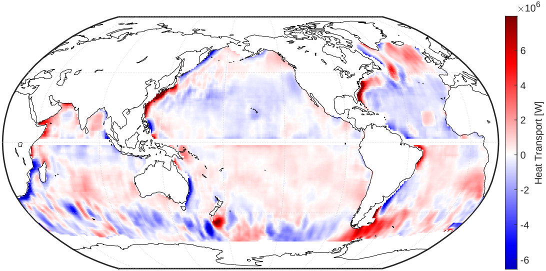
Debiased:



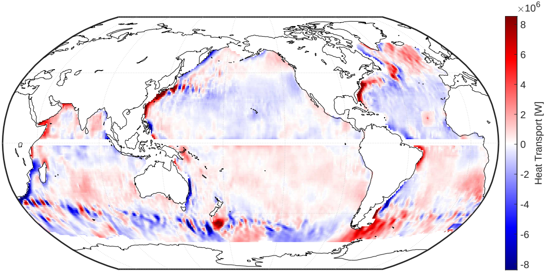
Upper-Ocean† mean Meridional Heat Transport field

† Upper Ocean: 10 to 900 dbar

Initial:



Debiased:



PCHIP Interpolation

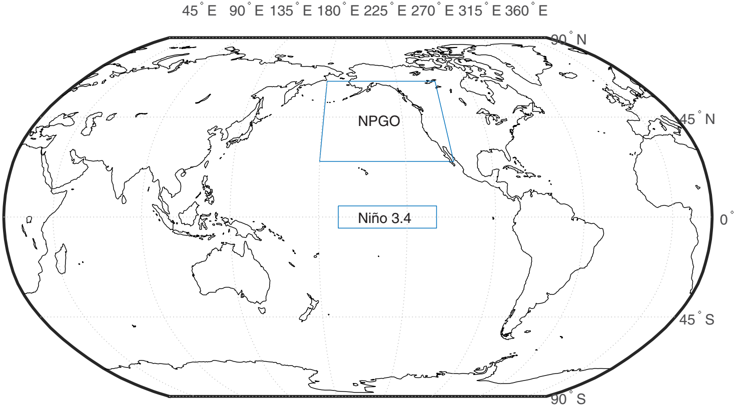
Fritsch, F. N. and Carlson, R. E. (1980). Monotone Piecewise Cubic Interpolation. *SIAM Journal on Numerical Analysis*, 17(2):238–246

Let $\theta\hat{\mathbf{v}} = \tilde{q}$

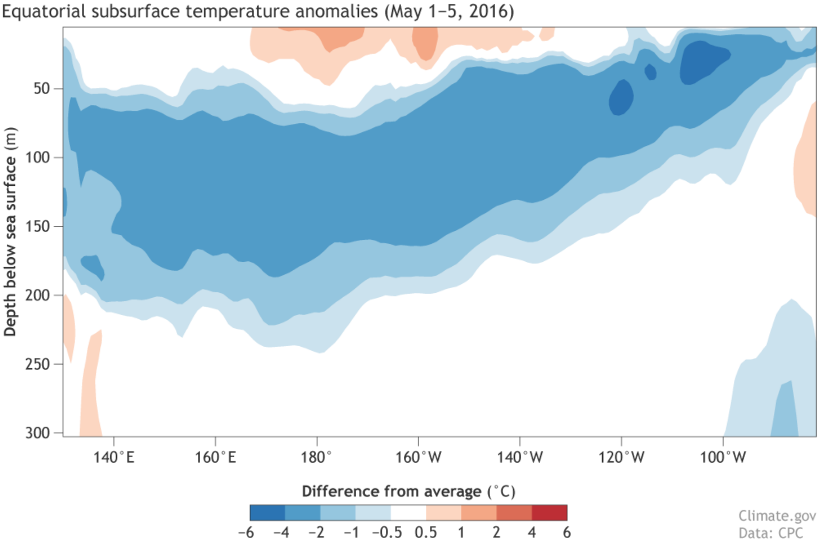
$$\begin{aligned}\widetilde{Q}(x_i, y_i, t_i) &= \int_{p_L}^{p_U} [\theta\hat{\mathbf{v}}](x_i, y_i, t_i, p)\mathrm{d}p \\ &\approxeq \sum_{k=0}^{N_{\mathrm{int}}} \Big[\tilde{q}(p_k)H_1(p_k^*) + \tilde{q}(p_{k+1})H_2(p_k^*) \\ &\qquad + \partial_p\tilde{q}(p_k)H_3(p_k^*) + \partial_p\tilde{q}(p_{k+1})H_4(p_k^*)\Big]\Delta_{p_k}\end{aligned}$$

where $p_k^* \in [p_k, p_{k+1}]$, $\Delta_{p_k} = p_{k+1} - p_k$, and $H_l(p)$ are the cubic Hermite basis functions for which interval $[p_k, p_{k+1}]$:
$H_1(p) = \phi\left(\frac{p_{k+1}-p}{\Delta_{p_k}}\right)$, $H_2 = \phi\left(\frac{p-p_k}{\Delta_{p_k}}\right)$, $H_3(p) = -\Delta_{p_k}\psi\left(\frac{p_{k+1}-p}{\Delta_{p_k}}\right)$,
$H_4(p) = \Delta_{p_k}\psi\left(\frac{p-p_k}{h_k}\right)$ where $\phi(p) = 3p^2 - 2p^3$, $\psi(p) = p^3 - p^2$.

Anomalous Heat Transport has connection to the unexplained ocean climate fluctuations.



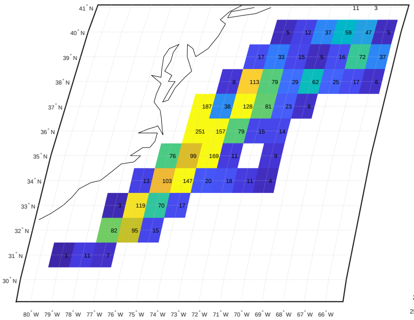
▸ ONI / Anomaly Heat Transport


Equatorial subsurface temperature anomalies (May 1–5, 2016)
Depth below sea surface (m)
50
100
150
200
250
300
140°E
160°E
180°
160°W
140°W
120°W
100°W
Difference from average (°C)
−6
−4
−2
−1
−0.5
0.5
1
2
4
6
Climate.gov
Data: CPC

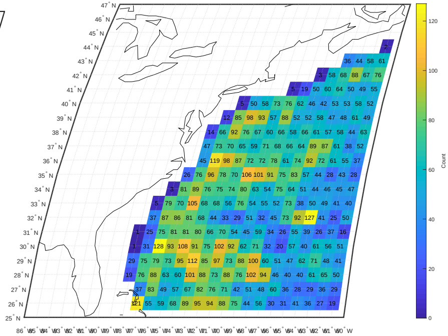
Map Area

Spray Profile Statistics

Out of 10,577 availabe profiles, only 2,791 profiles (26.4%) have measurements down to 800dbar.



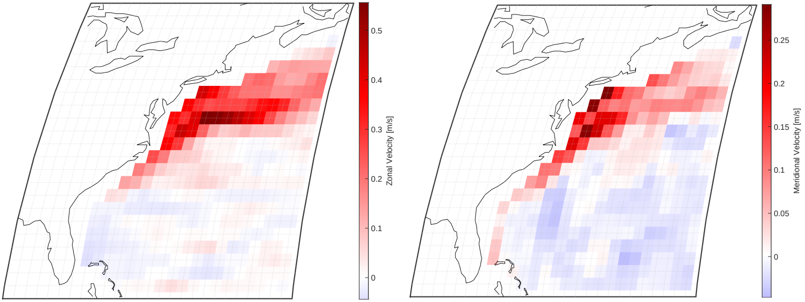
(a) Spray



(b) Argo

Figure: Number of profiles in $1^\circ \times 1^\circ$ grid at 15 dbar

Mean Velocity field from Argo profiles



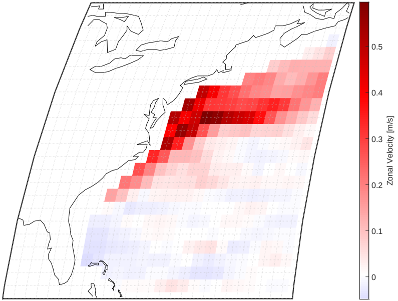
(a) Zonal

(b) Meridional

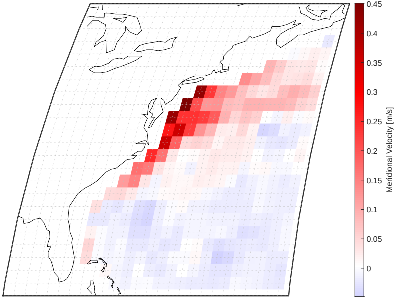
Figure: Mean velocity (Debiased) at 15 dbar

▸ Mean Field Difference

Mean Velocity field from Aggregated Spray and Argo



(a) Zonal

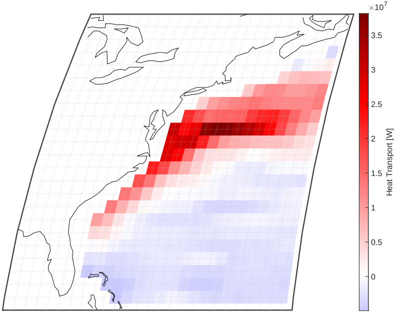


(b) Meridional

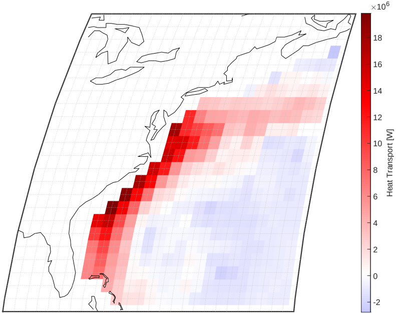
Figure: Mean velocity (Debiased) at 15 dbar

Mean Field Difference

Mean Heat Transport field from Argo profiles



(a) Zonal

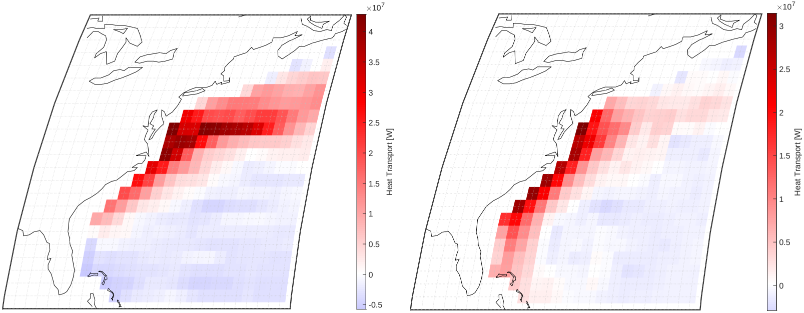


(b) Meridional

Figure: Mean absolute heat transport over 10 to 900 dbar

Mean Field Difference

Mean Heat Transport field from Aggregated profiles



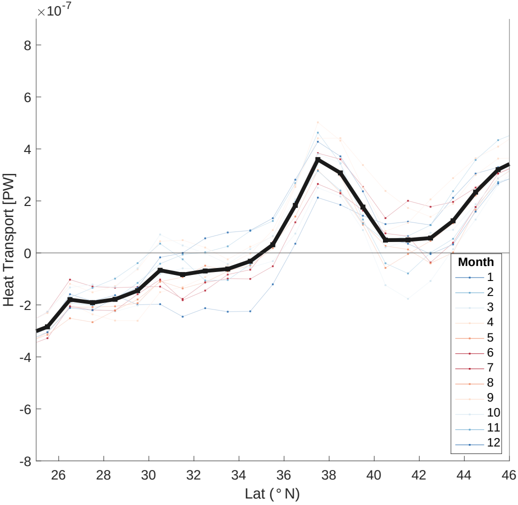
(a) Zonal

(b) Meridional

Figure: Mean absolute heat transport over 10 to 900 dbar

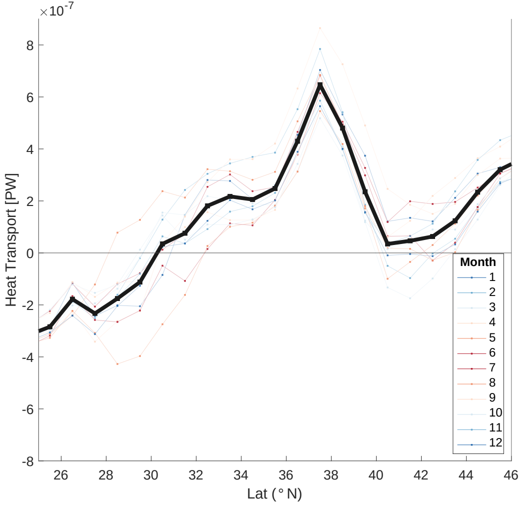
Mean Field Difference

Upper-Ocean mean Meridional Heat Transport in North Atlantic basin from Argo



Difference

Upper-Ocean mean Meridional Heat Transport in North Atlantic basin from Aggregated Spray and Argo



Difference