

# Supplement to “Locally stationary spatio-temporal interpolation of Argo profiling float data”

Mikael Kuusela and Michael L. Stein

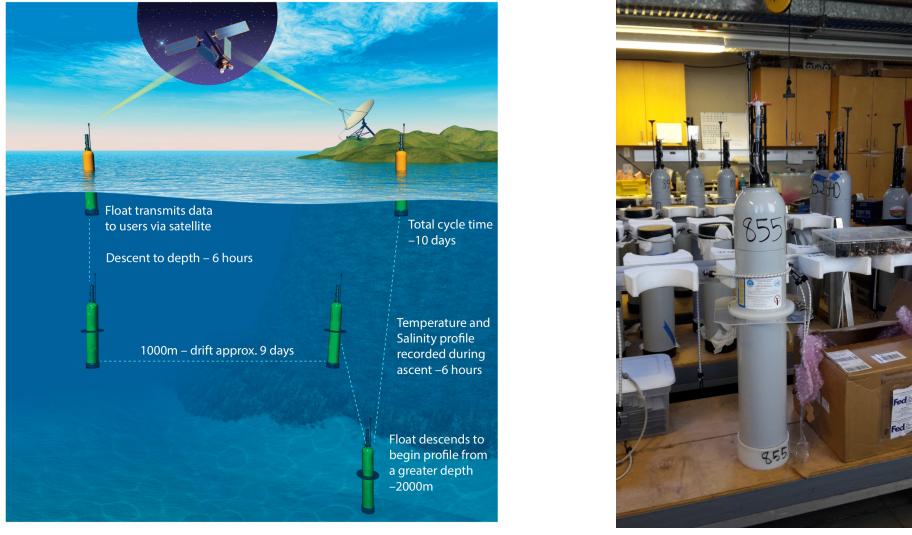
October 1, 2018

## 1 Overview of Argo floats and data

Argo [1, 2] is a global array of profiling floats for observing the subsurface ocean. The floats follow a regular cycle to measure temperature ( $T$ ), salinity ( $S$ ) and pressure ( $P$ ) in the upper 2,000 meters of the water column; see Figure 1(a). The cycle starts when a float at the surface changes its buoyancy and descends to the parking depth of 1,000 meters. The float stays at this depth for approximately 9 days, drifting with the current. After this, it descends further to 2,000 meters in preparation for measuring the vertical profiles of temperature and salinity. The float then ascends back to the surface recording *in situ* temperature, salinity and pressure triplets ( $T_i, S_i, P_i$ ) along the water column. This yields a temperature and salinity profile as a function of pressure for the upper 2,000 meters of the ocean. Once the float reaches the surface, it obtains a satellite position fix, sends its data via satellite to ground-based data centers and starts a new cycle. Each cycle takes approximately 10 days to complete.

There are currently almost 4,000 Argo floats operating in the world’s oceans. This results in nearly uniform global coverage at roughly  $3^\circ \times 3^\circ$  spatial resolution and 10 days temporal resolution. Figure 1 in the main text illustrates a monthly aggregate of Argo data at 300 dbar. The first Argo floats were deployed in 1999, sparse global coverage was achieved in 2004 and the  $3^\circ \times 3^\circ$  design resolution in 2007. While satellites provide plenty of data of the ocean surface, there are few other sources of data besides Argo for the subsurface ocean. Indeed, the pre-Argo ship-based subsurface temperature and salinity observations are sparse in both space and time with heavy sampling biases towards the Northern Hemisphere, coastal areas and summer months, while Argo provides much larger quantities of data and much more uniform sampling [2]. This improvement is particularly dramatic in the Southern Ocean during winter months, where pre-Argo subsurface data are extremely scarce [3, Figure 1.1].

Argo data are transmitted using either the Argos or Iridium satellite systems. Newer floats (Figure 1(b) shows a SOLO-II float at the Scripps Institution of Oceanography) primarily use Iridium, while roughly a half of the current array uses the older Argos system. The transmission system has important implications on the vertical resolution of Argo profiles: the limited bandwidth of the Argos system allows these floats to transmit only 50–100 data points per profile, while the Iridium floats can transmit up to a 1,000 vertical observations at a time. To ensure successful data transmission, the Argos floats need to remain on the surface for 6–12 hours, while Iridium data transmission takes only 15 minutes. The shorter surface time leads the Iridium floats to have a longer lifetime and less undesired surface drift than the Argos floats. The Argos floats also rely on the same satellite system for location information, while the Iridium floats use GPS, improving the accuracy of the position data. (The Argos location uncertainty can be



(a) Park and profile cycle

(b) Argo float

**Figure 1:** Figure (a) shows a schematic of the typical cycle of an Argo float. Most of the time, the float drifts at the 1,000 meters parking depth. Every 10 days, it descends further to 2,000 meters, turns on its sensors and measures a temperature and salinity profile as it ascends to the surface. (Figure courtesy of the Argo Program, <http://www.argo.ucsd.edu>, <http://doi.org/10.17882/42182>, used with permission.) Figure (b) shows an Argo float at the Scripps Institution of Oceanography (SIO). The measurement sensors and an antenna for satellite communications can be seen at the top of the float. The cylindrical part houses batteries, electronics and a hydraulic pump which is used to inflate and deflate an external bladder at the bottom of the float to adjust its buoyancy. There are a few different models of Argo floats. The one shown in the photo is a SOLO-II float designed and built at SIO. (Photo by Mikael Kuusela.)

neglected when mapping global temperature and salinity fields, but must be accounted for in analyses that involve Argo trajectory data.)

The nominal accuracies of Argo temperature, salinity and pressure sensors are 0.005 °C, 0.01 and 2.5 dbar, respectively [2]. (Salinity is measured on the unitless *practical salinity scale*.) The temperature data are considered very reliable, but pressure and salinity measurements may need to be adjusted for sensor offsets and drift. Argo data are distributed in two ways: real-time data are made available within 24 hours and have gone through various automatic checks and adjustments, while delayed-mode data have been manually checked by an oceanographic expert and may contain finer corrections for sensor offsets and drift [4]. Depending on the float model and configuration, the reported temperature and salinity values may be spot samples or pressure-bin averages. The vertical sampling is also often nonuniform, with finer sampling close to the surface. Some floats also sample at a finer temporal resolution than the nominal 10 days, in particular in the Kuroshio region near Japan.

Two extensions of the basic Argo program are currently under development [2]: The Deep Argo program aims to extend the vertical coverage of the Argo array to 6,000 meters, while biogeochemical Argo floats are to be equipped with extra sensors for measuring biogeochemical variables, such as dissolved oxygen, nitrate, chlorophyll and pH. Pilot arrays of both programs have been deployed in certain areas of the ocean.

## 2 Statistical methodology for the Student nugget model

We provide here a description of using the Laplace approximation to make inferences under the Student nugget model. This is an overview only—further details can be found in [5, 6].

Denoting  $\mathbf{f}_i = [f_i(\mathbf{x}_{i,1}, t_{i,1}), \dots, f_i(\mathbf{x}_{i,m_i}, t_{i,m_i})]^T$ , the  $i$ th year likelihood can be written as

$$p(\mathbf{y}_i|\boldsymbol{\theta}, \sigma^2, \nu) = \int p(\mathbf{y}_i|\mathbf{f}_i, \sigma^2, \nu)p(\mathbf{f}_i|\boldsymbol{\theta}) d\mathbf{f}_i = \int \exp(\Psi_i(\mathbf{f}_i)) d\mathbf{f}_i, \quad (1)$$

where  $\Psi_i(\mathbf{f}_i) = \log(p(\mathbf{y}_i|\mathbf{f}_i, \sigma^2, \nu)p(\mathbf{f}_i|\boldsymbol{\theta}))$ . Denote the mode of  $\Psi_i(\mathbf{f}_i)$  by  $\hat{\mathbf{f}}_i$ . Since

$$\hat{\mathbf{f}}_i = \arg \max_{\mathbf{f}_i \in \mathbb{R}^{m_i}} \Psi_i(\mathbf{f}_i) = \arg \max_{\mathbf{f}_i \in \mathbb{R}^{m_i}} \{p(\mathbf{y}_i|\mathbf{f}_i, \sigma^2, \nu)p(\mathbf{f}_i|\boldsymbol{\theta})\} \quad (2)$$

$$= \arg \max_{\mathbf{f}_i \in \mathbb{R}^{m_i}} \{p(\mathbf{f}_i|\mathbf{y}_i, \boldsymbol{\theta}, \sigma^2, \nu)p(\mathbf{y}_i|\boldsymbol{\theta}, \sigma^2, \nu)\} = \arg \max_{\mathbf{f}_i \in \mathbb{R}^{m_i}} p(\mathbf{f}_i|\mathbf{y}_i, \boldsymbol{\theta}, \sigma^2, \nu), \quad (3)$$

$\hat{\mathbf{f}}_i$  is also the mode of the conditional distribution  $p(\mathbf{f}_i|\mathbf{y}_i, \boldsymbol{\theta}, \sigma^2, \nu)$ . The Laplace approximation is based on taking a second-order Taylor expansion of  $\Psi_i(\mathbf{f}_i)$  around  $\hat{\mathbf{f}}_i$ , which gives

$$\Psi_i(\mathbf{f}_i) \approx \Psi_i(\hat{\mathbf{f}}_i) - \frac{1}{2}(\mathbf{f}_i - \hat{\mathbf{f}}_i)^T \left( -\frac{\partial^2}{\partial \mathbf{f}_i^2} \Psi_i(\mathbf{f}_i) \Big|_{\mathbf{f}_i=\hat{\mathbf{f}}_i} \right) (\mathbf{f}_i - \hat{\mathbf{f}}_i). \quad (4)$$

With this approximation, Equation (1) becomes a Gaussian integral that can be evaluated analytically. The resulting approximate log-likelihood is [5, 6]

$$\begin{aligned} & \log q(\mathbf{y}_i|\boldsymbol{\theta}, \sigma^2, \nu) \\ &= \log p(\mathbf{y}_i|\hat{\mathbf{f}}_i, \sigma^2, \nu) - \frac{1}{2}\hat{\mathbf{f}}_i^T \mathbf{K}_i(\boldsymbol{\theta})^{-1} \hat{\mathbf{f}}_i - \frac{1}{2} \log \det \mathbf{K}_i(\boldsymbol{\theta}) - \frac{1}{2} \log \det \left( \mathbf{K}_i(\boldsymbol{\theta})^{-1} + \mathbf{W}_i(\hat{\mathbf{f}}_i, \sigma^2, \nu) \right), \end{aligned} \quad (5)$$

where  $\mathbf{W}_i(\hat{\mathbf{f}}_i, \sigma^2, \nu) = -\frac{\partial^2}{\partial \mathbf{f}_i^2} \log p(\mathbf{y}_i|\mathbf{f}_i, \sigma^2, \nu) \Big|_{\mathbf{f}_i=\hat{\mathbf{f}}_i}$ . The  $n$ -year approximate log-likelihood is then the sum of the one-year log-likelihoods,  $\log q(\mathbf{y}|\boldsymbol{\theta}, \sigma^2, \nu) = \sum_{i=1}^n \log q(\mathbf{y}_i|\boldsymbol{\theta}, \sigma^2, \nu)$ , and this is maximized numerically with respect to the model parameters  $(\boldsymbol{\theta}, \sigma^2, \nu)$ .

Denote  $f_i^* = f_i(\mathbf{x}^*, t^*)$ . The predictive distribution for  $f_i^*$  is

$$p(f_i^*|\mathbf{y}_i, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2, \hat{\nu}) = \int p(f_i^*|\mathbf{f}_i, \hat{\boldsymbol{\theta}})p(\mathbf{f}_i|\mathbf{y}_i, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2, \hat{\nu}) d\mathbf{f}_i, \quad (6)$$

where we have plugged in the Laplace approximated MLEs for the model parameters. Here one can again employ the Laplace approximation by taking a second-order Taylor expansion of  $\log p(\mathbf{f}_i|\mathbf{y}_i, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2, \hat{\nu})$  around its mode. This results in the Gaussian approximation  $(\mathbf{f}_i|\mathbf{y}_i, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2, \hat{\nu}) \stackrel{a}{\sim} N(\hat{\mathbf{f}}_i, (\mathbf{K}_i(\hat{\boldsymbol{\theta}})^{-1} + \mathbf{W}_i(\hat{\mathbf{f}}_i, \hat{\sigma}^2, \hat{\nu}))^{-1})$ . Equation (6) then becomes a Gaussian integral, which can be evaluated to give

$$(f_i^*|\mathbf{y}_i, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2, \hat{\nu}) \stackrel{a}{\sim} N((\mathbf{k}_i^*(\hat{\boldsymbol{\theta}}))^T \mathbf{K}_i(\hat{\boldsymbol{\theta}})^{-1} \hat{\mathbf{f}}_i, (\mathbf{k}_i^*(\hat{\boldsymbol{\theta}}))^T (\mathbf{K}_i(\hat{\boldsymbol{\theta}}) + \mathbf{W}_i(\hat{\mathbf{f}}_i, \hat{\sigma}^2, \hat{\nu})^{-1})^{-1} \mathbf{k}_i^*(\hat{\boldsymbol{\theta}})). \quad (7)$$

The Laplace approximated predictive distribution for  $y_i^* = f_i^* + \varepsilon_i^*$  is therefore

$$(y_i^*|\mathbf{y}_i, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2, \hat{\nu}) \stackrel{a}{\sim} P_{Z_1+Z_2}, \quad (8)$$

where  $P_{Z_1+Z_2}$  denotes the distribution of  $Z_1 + Z_2$ , with  $Z_1$  following the Gaussian distribution in Equation (7),  $Z_2 \sim t_{\hat{\nu}}(\hat{\sigma}^2)$  and  $Z_1$  and  $Z_2$  are independent. The point predictions are made

using the approximate conditional mean  $\hat{y}_i^* = (\mathbf{k}_i^*(\hat{\boldsymbol{\theta}}))^T \mathbf{K}_i(\hat{\boldsymbol{\theta}})^{-1} \hat{\mathbf{f}}_i \approx \mathbb{E}(y_i^* | \mathbf{y}_i, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2, \hat{\nu})$  and the  $1 - \alpha$  predictive intervals are  $[\underline{y}_i^*, \bar{y}_i^*] = [\hat{y}_i^* - q_{1-\alpha/2}, \hat{y}_i^* + q_{1-\alpha/2}]$ , where  $q_{1-\alpha/2}$  is the  $1 - \alpha/2$  quantile of the distribution of  $(Z_1 - \hat{y}_i^*) + Z_2$ , where  $Z_1$  and  $Z_2$  are as in Equation (8). Since  $q_{1-\alpha/2}$  is not available in closed form, we compute it by Monte Carlo sampling.

Software implementations of Gaussian processes with a Student nugget differ in terms of how  $\hat{\mathbf{f}}_i = \arg \max_{\mathbf{f}_i \in \mathbb{R}^{m_i}} p(\mathbf{f}_i | \mathbf{y}_i, \boldsymbol{\theta}, \sigma^2, \nu)$  is computed and how the approximate log-likelihood  $\log q(\mathbf{y} | \boldsymbol{\theta}, \sigma^2, \nu)$  is maximized. In this work, we compute the Student fits using the GPML toolbox [7, 8], which we have adapted to include the multi-year likelihood and the Student predictive intervals. GPML uses conjugate gradient descent for finding the model parameters  $(\boldsymbol{\theta}, \sigma^2, \nu)$  and Newton's method for maximizing  $p(\mathbf{f}_i | \mathbf{y}_i, \boldsymbol{\theta}, \sigma^2, \nu)$  with respect to  $\mathbf{f}_i$ , with special care for the nonconcavity of the latter problem [7]. When finding the approximate MLEs, we initialize  $\boldsymbol{\theta} = [\phi, \theta_{\text{lat}}, \theta_{\text{lon}}, \theta_t]^T$  to the MLEs from the Gaussian nugget fit. This initialization was found to significantly improve the convergence and stability of the conjugate gradient descent.

### 3 Quality control criteria

We use the following criteria to filter out profiles that may contain bad data due to technical issues:

- We require quality control (QC) flags 1 (“good data”) or 2 (“probably good data”) [4] for position, time, pressure, temperature and salinity. If any observation in the profile has any other QC flag, the entire profile is discarded.
- We reject profiles that are shorter than 100 dbar or have gaps larger than 200 dbar.
- We require that the pressure be strictly monotonic across the profile.
- We require each profile to have at least two observations and the same number of observations for both temperature and salinity.
- We identify and reject profiles from problematic APEX floats as recommended in [9].
- We reject profiles with cycle number 0 (the launch cycle) or with unrealistic coordinate values.

When available, we use delayed (mode D) or adjusted real-time (mode A) data. If mode D or A data are not available, we accept real-time profiles (mode R) provided that the above-mentioned criteria are satisfied. Argo User’s Manual [10] provides more information about the different data types. These selection criteria are designed to provide a sufficiently clean and regular data set that enables a realistic comparison between the different interpolation procedures—we acknowledge that more efficient criteria can probably be developed.

### 4 Cross-validation of the predictive distribution

This section explains the computation of the cross-validated predictive sample quantiles for the uncertainty quantification studies in Section 4(c) in the main text. With the Gaussian nugget, the cross-validated predictive distribution for  $y_{i,j}$  is  $(y_{i,j} | \mathbf{y}_{i \setminus (i,j)}, \hat{\boldsymbol{\theta}}, \hat{\sigma}^2) \sim N(\hat{m}_{-(i,j)}^G, \hat{v}_{-(i,j)}^G)$ , where  $\mathbf{y}_{i \setminus (i,j)}$  is  $\mathbf{y}_i$  with either  $y_{i,j}$  removed (LOOO) or all observations with the same float ID as  $y_{i,j}$  removed (LOFO). The predictive mean  $\hat{m}_{-(i,j)}^G$  and variance  $\hat{v}_{-(i,j)}^G$  are as in Equation (3.6)

in the main text, but with the appropriate data points removed. The covariance parameters  $\hat{\theta}$  and  $\hat{\sigma}^2$  are taken from the MLEs at the grid point closest to  $y_{i,j}$ . Conditional on  $(\mathbf{y}_{i \setminus (i,j)}, \hat{\theta}, \hat{\sigma}^2)$ , it then follows that  $(y_{i,j} - \hat{m}_{-(i,j)}^G) / \sqrt{\hat{v}_{-(i,j)}^G} \sim N(0, 1)$ , so we can compare the sample quantiles of  $\{(y_{i,j} - \hat{m}_{-(i,j)}^G) / \sqrt{\hat{v}_{-(i,j)}^G}, i = 1, \dots, n, j = 1, \dots, m_i\}$  to the corresponding quantiles of the standard Gaussian distribution.

With the Student nugget, the cross-validated Laplace approximated predictive distribution is  $(y_{i,j} | \mathbf{y}_{i \setminus (i,j)}, \hat{\theta}, \hat{\sigma}^2, \hat{\nu}) \stackrel{a}{\sim} P_{Z_1+Z_2}$ , where  $P_{Z_1+Z_2}$  denotes the distribution of  $Z_1 + Z_2$ , with  $Z_1 \sim N(\hat{m}_{-(i,j)}^S, \hat{v}_{-(i,j)}^S)$ , where  $\hat{m}_{-(i,j)}^S$  and  $\hat{v}_{-(i,j)}^S$  are as in Equation (7), but with the appropriate data points removed when computing the Laplace approximation,  $Z_2 \sim t_{\hat{\nu}}(\hat{\sigma}^2)$  and  $Z_1$  and  $Z_2$  are independent. Let  $F$  denote the cumulative distribution function of  $P_{Z_1+Z_2}$ . Then, conditional on  $(\mathbf{y}_{i \setminus (i,j)}, \hat{\theta}, \hat{\sigma}^2, \hat{\nu})$ , we have that  $F(y_{i,j}) \sim U(0, 1)$ , where  $U(0, 1)$  is the uniform distribution on  $(0, 1)$ . Letting  $\Phi$  denote the standard Gaussian cumulative distribution function, it follows that  $\Phi^{-1}(F(y_{i,j})) \sim N(0, 1)$  and we can again compare the sample quantiles of  $\{\Phi^{-1}(F(y_{i,j})), i = 1, \dots, n, j = 1, \dots, m_i\}$  to the quantiles of the standard Gaussian distribution. Since  $F$  is not available in closed form, we compute its values by Monte Carlo sampling. The model parameters  $(\hat{\theta}, \hat{\sigma}^2, \hat{\nu})$  are taken to be the Laplace approximated MLEs at the grid point closest to  $y_{i,j}$ .

## 5 Supplementary results

The following pages provide supplementary results as described in the main text.

## 5.1 Point prediction performance with the Student nugget

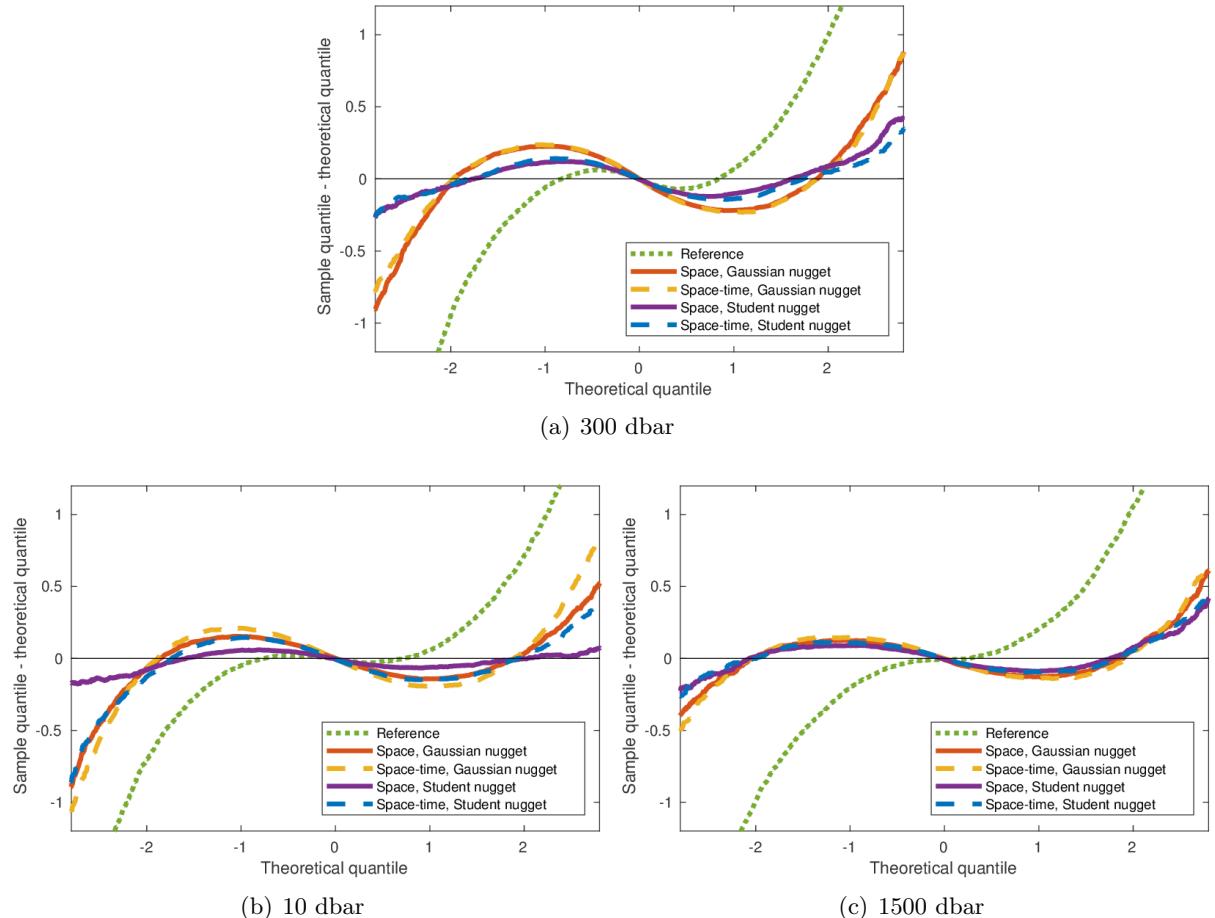
**Table 1:** Point prediction performance for the Student nugget with LOOO cross-validation. The percentages in the parentheses are differences with respect to the corresponding model with a Gaussian nugget.

Pressure level	Performance metric	Space (1 month)	Space-time (3 months)
10 dbar	RMSE	0.6159 (-4.8%)	0.5583 (-10.1%)
	Q <sub>3</sub> AE	0.4868 (-0.9%)	0.3920 (-5.0%)
	MdAE	0.2509 (-0.8%)	0.1890 (-4.9%)
300 dbar	RMSE	0.6223 (-9.3%)	0.5873 (-14.6%)
	Q <sub>3</sub> AE	0.4248 (-2.4%)	0.3813 (-3.5%)
	MdAE	0.1995 (-2.0%)	0.1779 (-2.2%)
1500 dbar	RMSE	0.1027 (-3.1%)	0.0912 (-3.3%)
	Q <sub>3</sub> AE	0.0736 (-1.6%)	0.0653 (-1.8%)
	MdAE	0.0361 (-1.6%)	0.0316 (-1.7%)

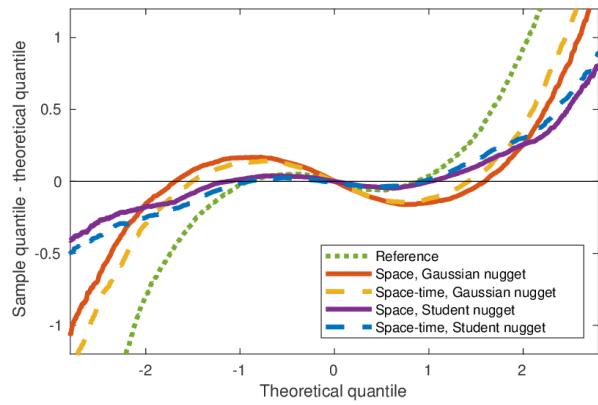
**Table 2:** Point prediction performance for the Student nugget with LOFO cross-validation. The percentages in the parentheses are differences with respect to the corresponding model with a Gaussian nugget.

Pressure level	Performance metric	Space (1 month)	Space-time (3 months)
10 dbar	RMSE	0.6894 (-1.0%)	0.6637 (-2.3%)
	Q <sub>3</sub> AE	0.5757 (0.3%)	0.5269 (-0.9%)
	MdAE	0.2963 (0.6%)	0.2589 (-1.4%)
300 dbar	RMSE	0.7617 (-1.8%)	0.7661 (-3.7%)
	Q <sub>3</sub> AE	0.5712 (0.4%)	0.5565 (-0.2%)
	MdAE	0.2754 (-0.0%)	0.2670 (-0.2%)
1500 dbar	RMSE	0.1316 (-0.6%)	0.1325 (-1.4%)
	Q <sub>3</sub> AE	0.0977 (-0.2%)	0.0966 (-0.8%)
	MdAE	0.0501 (-0.5%)	0.0486 (-0.5%)

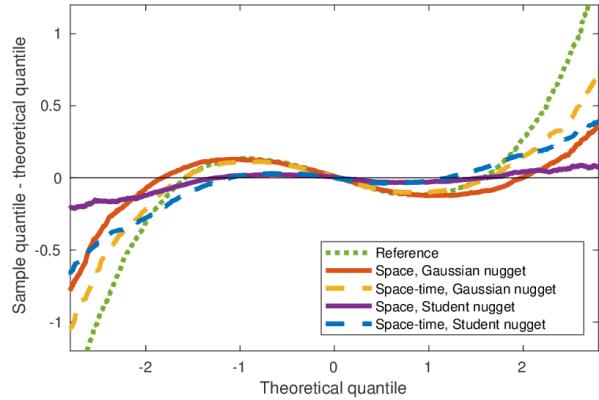
## 5.2 Uncertainty quantification performance



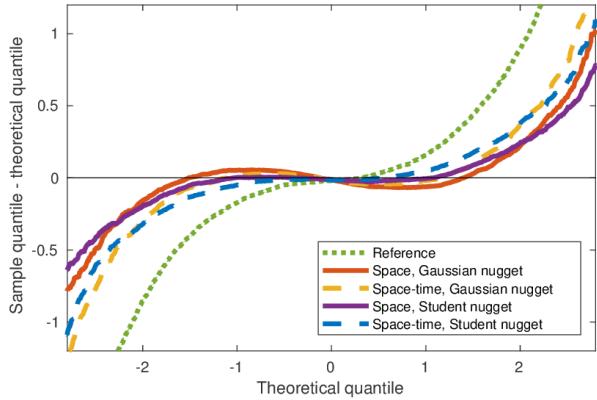
**Figure 2:**  $q_{\text{sample}} - q_{\text{theory}}$  vs.  $q_{\text{theory}}$  for LOOO cross-validation; Figure (a) is the same as Figure 5 in the main paper



(a) 300 dbar



(b) 10 dbar



(c) 1500 dbar

**Figure 3:**  $q_{\text{sample}} - q_{\text{theory}}$  vs.  $q_{\text{theory}}$  for LOFO cross-validation

**Table 3:** Coverage and interval lengths, LOOO, 10 dbar

Confidence level	Method	Empirical coverage	Mean length	Median length
68 %	Reference	0.6622	0.8517	0.7797
	Space, Gaussian nugget	0.7471	1.0272	0.9229
	Space-time, Gaussian nugget	0.7683	0.8477	0.7362
	Space, Student nugget	0.7065	0.9139	0.8474
	Space-time, Student nugget	0.7423	0.8043	0.7304
95 %	Reference	0.8911	1.6693	1.5282
	Space, Gaussian nugget	0.9471	2.0132	1.8088
	Space-time, Gaussian nugget	0.9474	1.6614	1.4429
	Space, Student nugget	0.9461	2.0457	1.8453
	Space-time, Student nugget	0.9443	1.7695	1.5541
99 %	Reference	0.9470	2.1938	2.0084
	Space, Gaussian nugget	0.9804	2.6458	2.3772
	Space-time, Gaussian nugget	0.9780	2.1834	1.8962
	Space, Student nugget	0.9866	3.3607	2.7012
	Space-time, Student nugget	0.9800	2.9097	2.1930

**Table 4:** Coverage and interval lengths, LOOO, 300 dbar

Confidence level	Method	Empirical coverage	Mean length	Median length
68 %	Reference	0.6607	0.7511	0.6435
	Space, Gaussian nugget	0.7745	0.9749	0.8728
	Space-time, Gaussian nugget	0.7800	0.8722	0.7816
	Space, Student nugget	0.7261	0.8427	0.7621
	Space-time, Student nugget	0.7389	0.7697	0.7049
95 %	Reference	0.8755	1.4721	1.2612
	Space, Gaussian nugget	0.9482	1.9108	1.7107
	Space-time, Gaussian nugget	0.9490	1.7095	1.5320
	Space, Student nugget	0.9432	1.9918	1.7525
	Space-time, Student nugget	0.9452	1.8543	1.6192
99 %	Reference	0.9329	1.9347	1.6575
	Space, Gaussian nugget	0.9777	2.5112	2.2483
	Space-time, Gaussian nugget	0.9793	2.2466	2.0134
	Space, Student nugget	0.9835	3.9385	2.6417
	Space-time, Student nugget	0.9844	3.8086	2.3976

**Table 5:** Coverage and interval lengths, LOOO, 1500 dbar

Confidence level	Method	Empirical coverage	Mean length	Median length
68 %	Reference	0.6073	0.1200	0.0935
	Space, Gaussian nugget	0.7392	0.1603	0.1316
	Space-time, Gaussian nugget	0.7465	0.1429	0.1196
	Space, Student nugget	0.7199	0.1525	0.1274
	Space-time, Student nugget	0.7266	0.1369	0.1157
95 %	Reference	0.8596	0.2353	0.1833
	Space, Gaussian nugget	0.9492	0.3142	0.2579
	Space-time, Gaussian nugget	0.9504	0.2800	0.2344
	Space, Student nugget	0.9477	0.3198	0.2609
	Space-time, Student nugget	0.9477	0.2875	0.2396
99 %	Reference	0.9273	0.3092	0.2410
	Space, Gaussian nugget	0.9828	0.4129	0.3389
	Space-time, Gaussian nugget	0.9823	0.3680	0.3081
	Space, Student nugget	0.9838	0.4809	0.3528
	Space-time, Student nugget	0.9839	0.4329	0.3246

**Table 6:** Coverage and interval lengths, LOFO, 10 dbar

Confidence level	Method	Empirical coverage	Mean length	Median length
68 %	Reference	0.7348	1.2470	1.2045
	Space, Gaussian nugget	0.7391	1.2054	1.0651
	Space-time, Gaussian nugget	0.7283	1.0598	0.9156
	Space, Student nugget	0.6895	1.0346	0.9647
	Space-time, Student nugget	0.6874	0.9310	0.8530
95 %	Reference	0.9308	2.4441	2.3609
	Space, Gaussian nugget	0.9480	2.3626	2.0876
	Space-time, Gaussian nugget	0.9372	2.0773	1.7945
	Space, Student nugget	0.9408	2.2553	2.0571
	Space-time, Student nugget	0.9278	1.9998	1.7867
99 %	Reference	0.9680	3.2120	3.1027
	Space, Gaussian nugget	0.9825	3.1049	2.7436
	Space-time, Gaussian nugget	0.9756	2.7300	2.3584
	Space, Student nugget	0.9849	3.5920	2.9646
	Space-time, Student nugget	0.9764	3.1814	2.5303

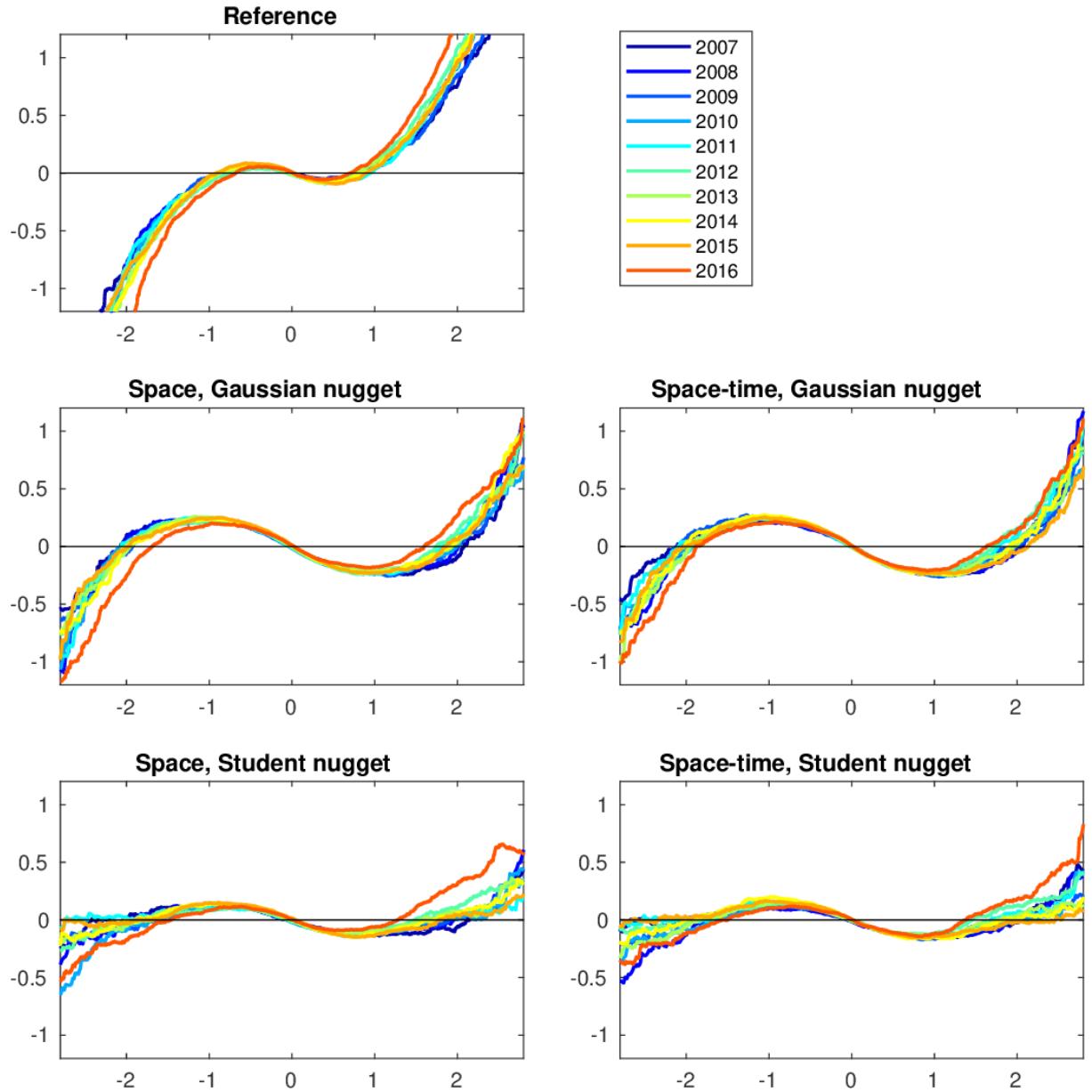
**Table 7:** Coverage and interval lengths, LOFO, 300 dbar

Confidence level	Method	Empirical coverage	Mean length	Median length
68 %	Reference	0.6707	1.0853	0.9497
	Space, Gaussian nugget	0.7475	1.2438	1.1084
	Space-time, Gaussian nugget	0.7320	1.1586	1.0393
	Space, Student nugget	0.6839	1.0152	0.9129
	Space-time, Student nugget	0.6743	0.9535	0.8731
95 %	Reference	0.8870	2.1271	1.8613
	Space, Gaussian nugget	0.9354	2.4379	2.1725
	Space-time, Gaussian nugget	0.9256	2.2708	2.0370
	Space, Student nugget	0.9255	2.3047	2.0537
	Space-time, Student nugget	0.9171	2.1846	1.9605
99 %	Reference	0.9396	2.7954	2.4462
	Space, Gaussian nugget	0.9717	3.2039	2.8551
	Space-time, Gaussian nugget	0.9658	2.9843	2.6770
	Space, Student nugget	0.9766	4.3110	3.0468
	Space-time, Student nugget	0.9729	4.1819	2.8874

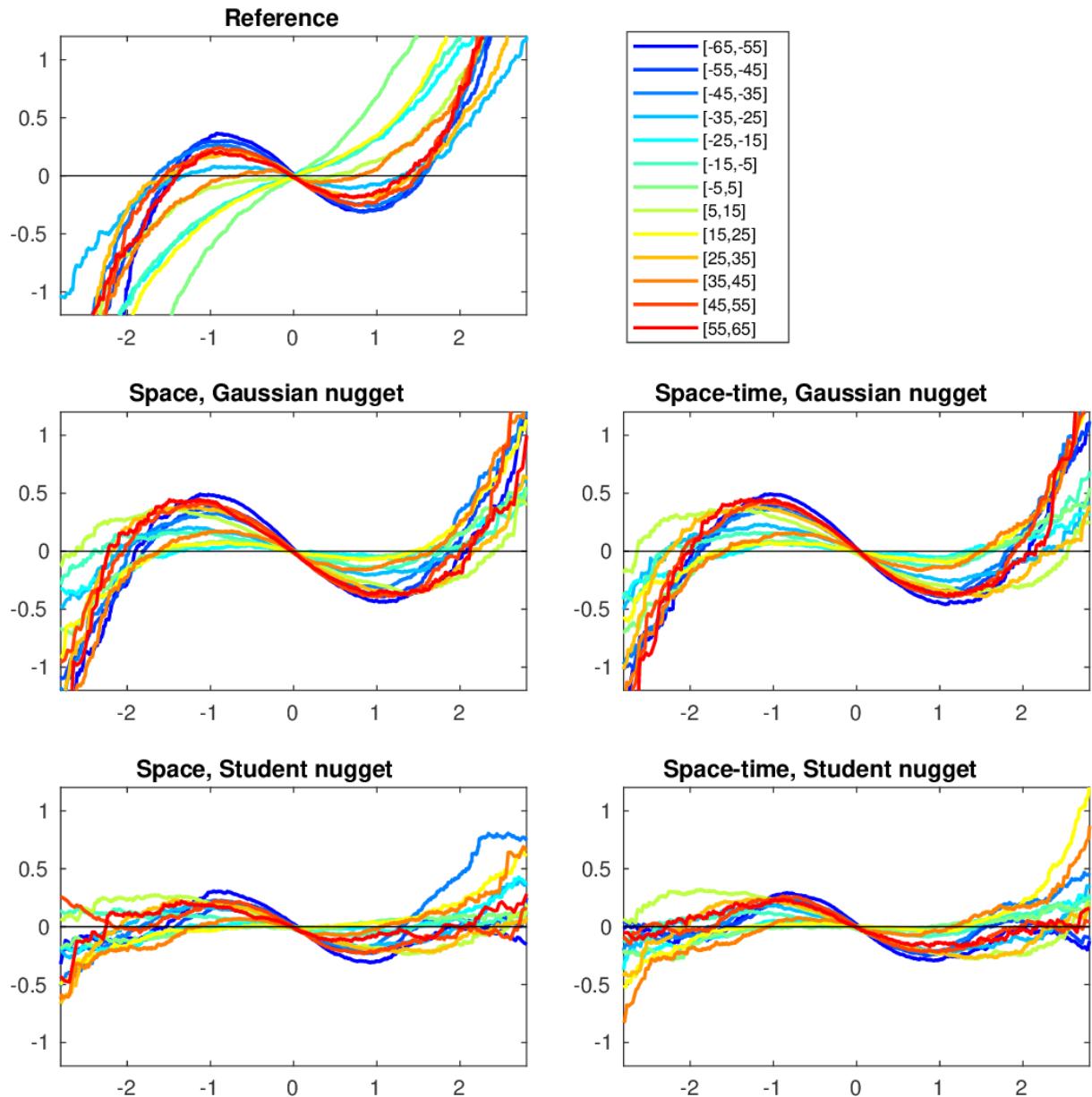
**Table 8:** Coverage and interval lengths, LOFO, 1500 dbar

Confidence level	Method	Empirical coverage	Mean length	Median length
68 %	Reference	0.6186	0.1695	0.1376
	Space, Gaussian nugget	0.7089	0.1991	0.1577
	Space-time, Gaussian nugget	0.6890	0.1860	0.1487
	Space, Student nugget	0.6803	0.1841	0.1491
	Space-time, Student nugget	0.6581	0.1720	0.1398
95 %	Reference	0.8725	0.3322	0.2697
	Space, Gaussian nugget	0.9339	0.3903	0.3091
	Space-time, Gaussian nugget	0.9234	0.3645	0.2914
	Space, Student nugget	0.9287	0.3803	0.3041
	Space-time, Student nugget	0.9156	0.3542	0.2866
99 %	Reference	0.9372	0.4366	0.3545
	Space, Gaussian nugget	0.9750	0.5129	0.4062
	Space-time, Gaussian nugget	0.9680	0.4791	0.3830
	Space, Student nugget	0.9750	0.5578	0.4161
	Space-time, Student nugget	0.9672	0.5175	0.3911

### 5.3 Stratified quantile plots



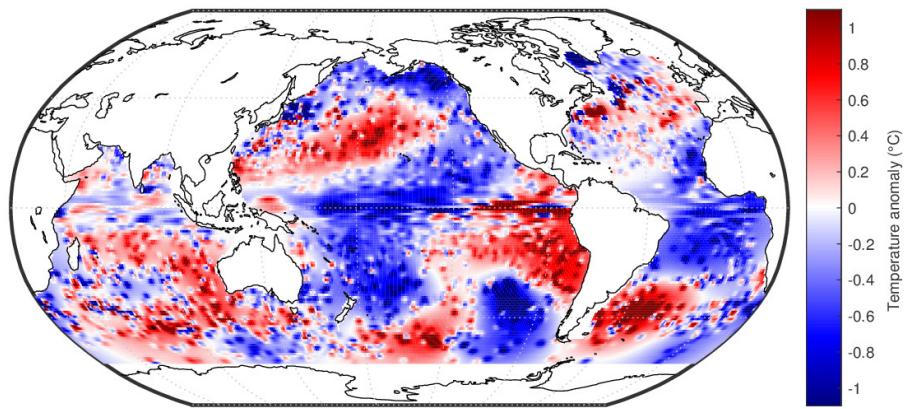
**Figure 4:** Figure 2(a) stratified by year.



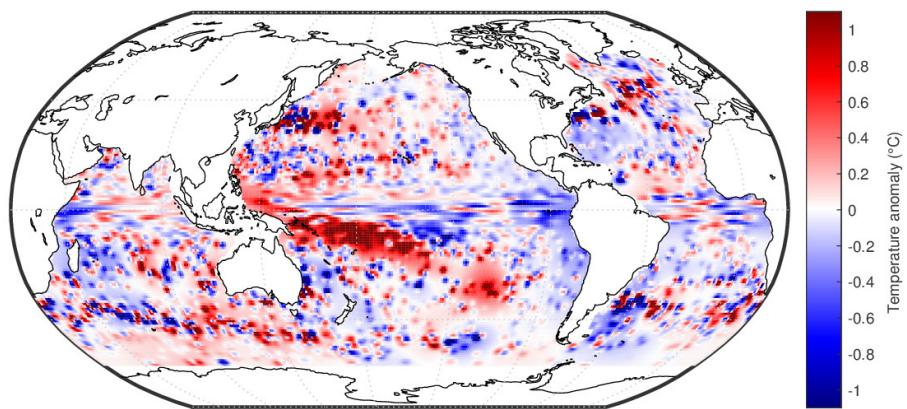
**Figure 5:** Figure 2(a) stratified by latitude bins.

#### **5.4 Model 1: Anomalies and model parameters**

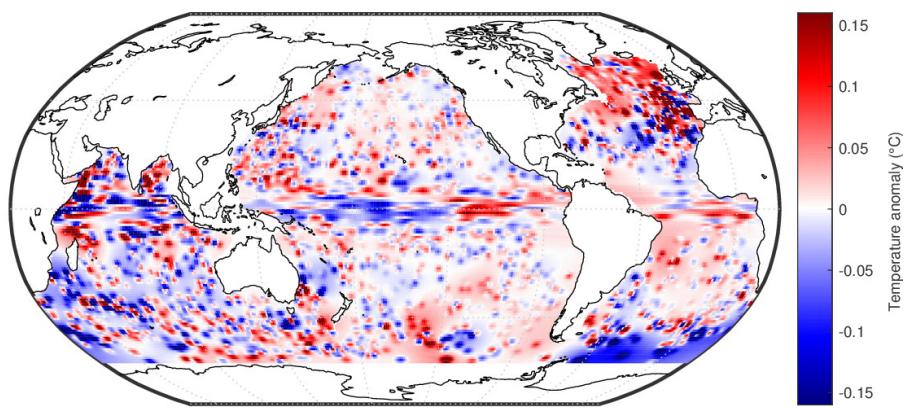
The following pages illustrate the Roemmich–Gilson-like reference model (Model 1).



(a) 10 dbar

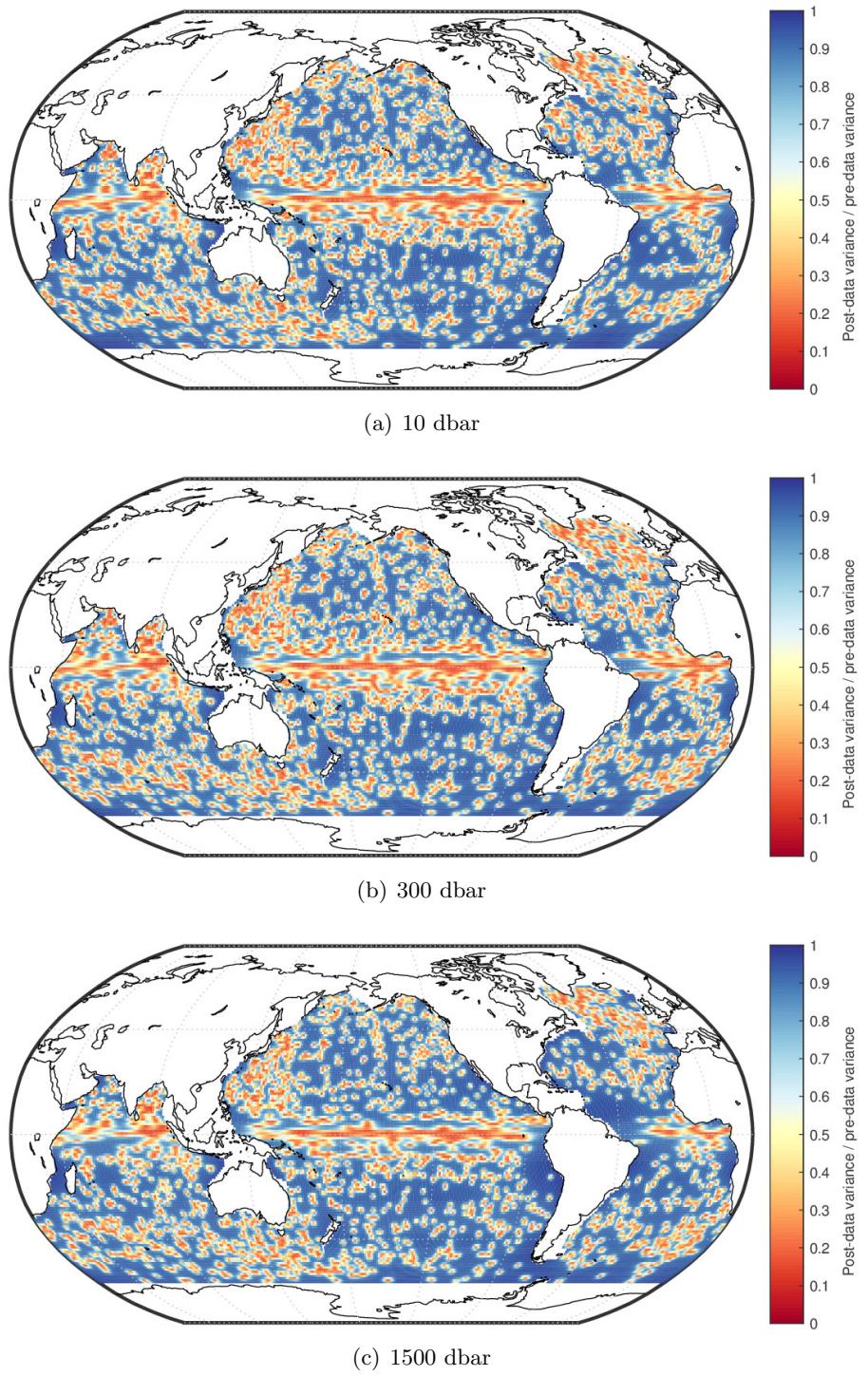


(b) 300 dbar

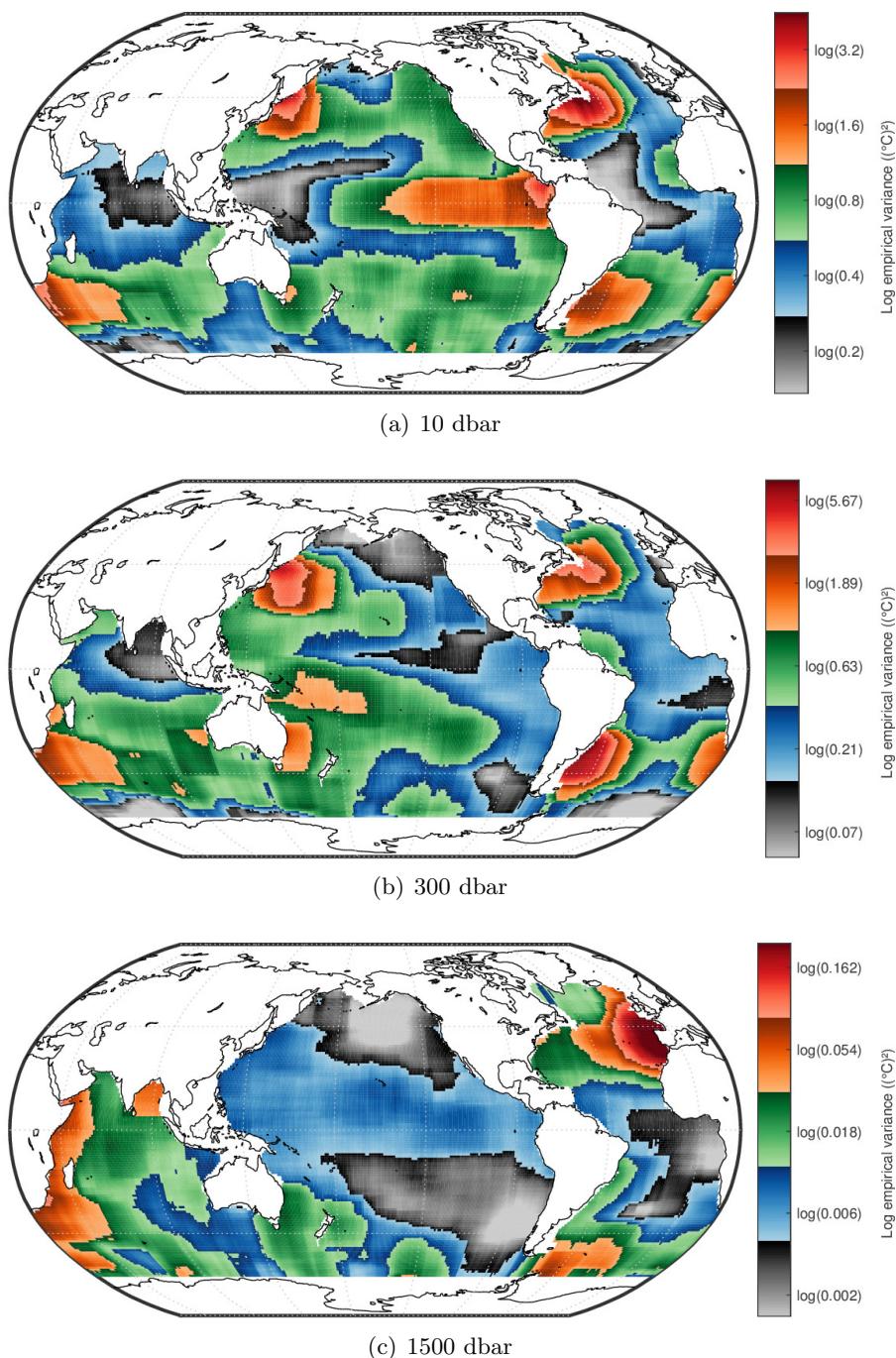


(c) 1500 dbar

**Figure 6:** February 2012 temperature anomalies



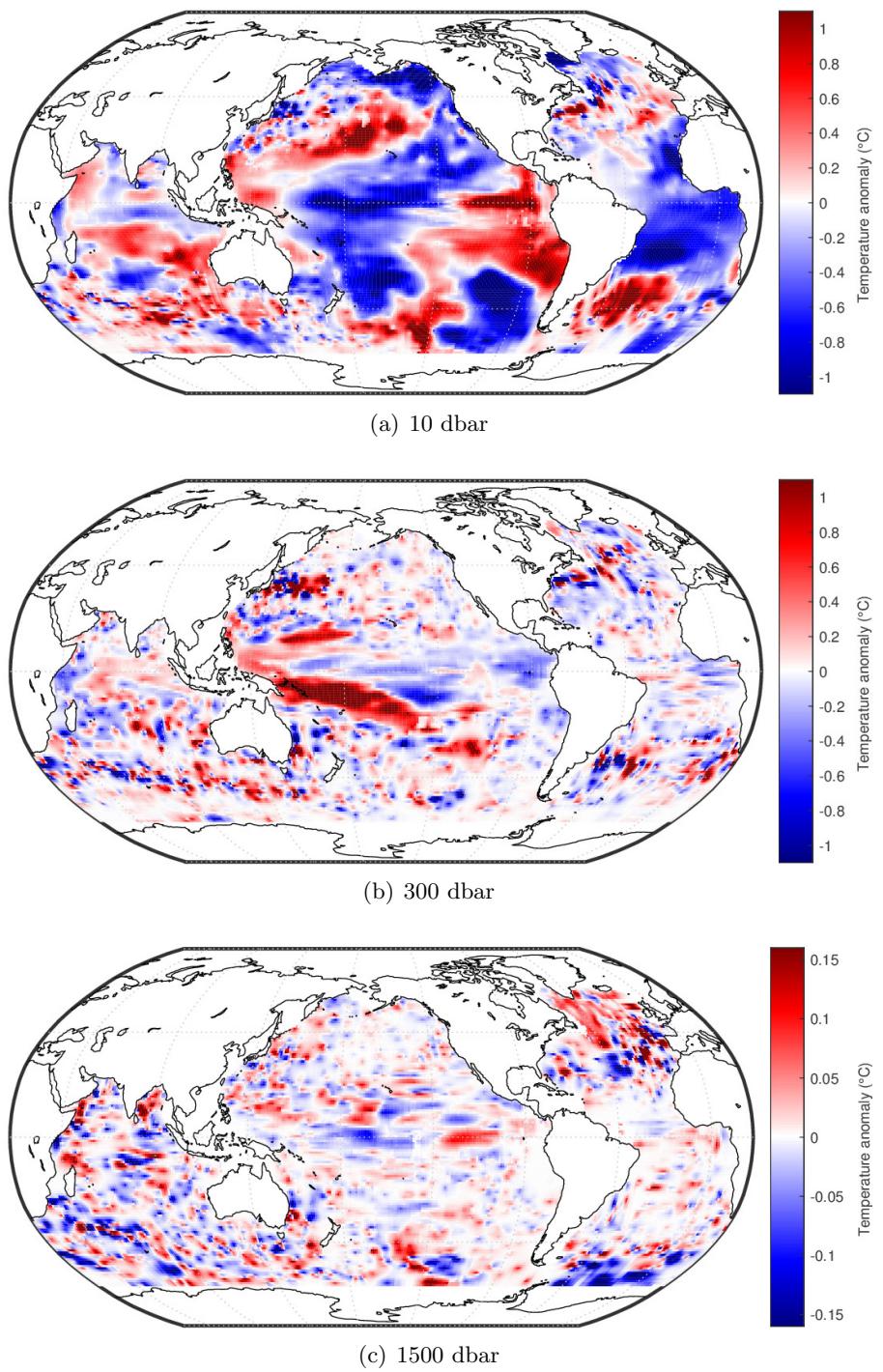
**Figure 7:** February 2012 post-data-to-pre-data variance ratios. The uncertainty patterns are almost the same at all depths since the Roemmich–Gilson covariance does not vary with depth.



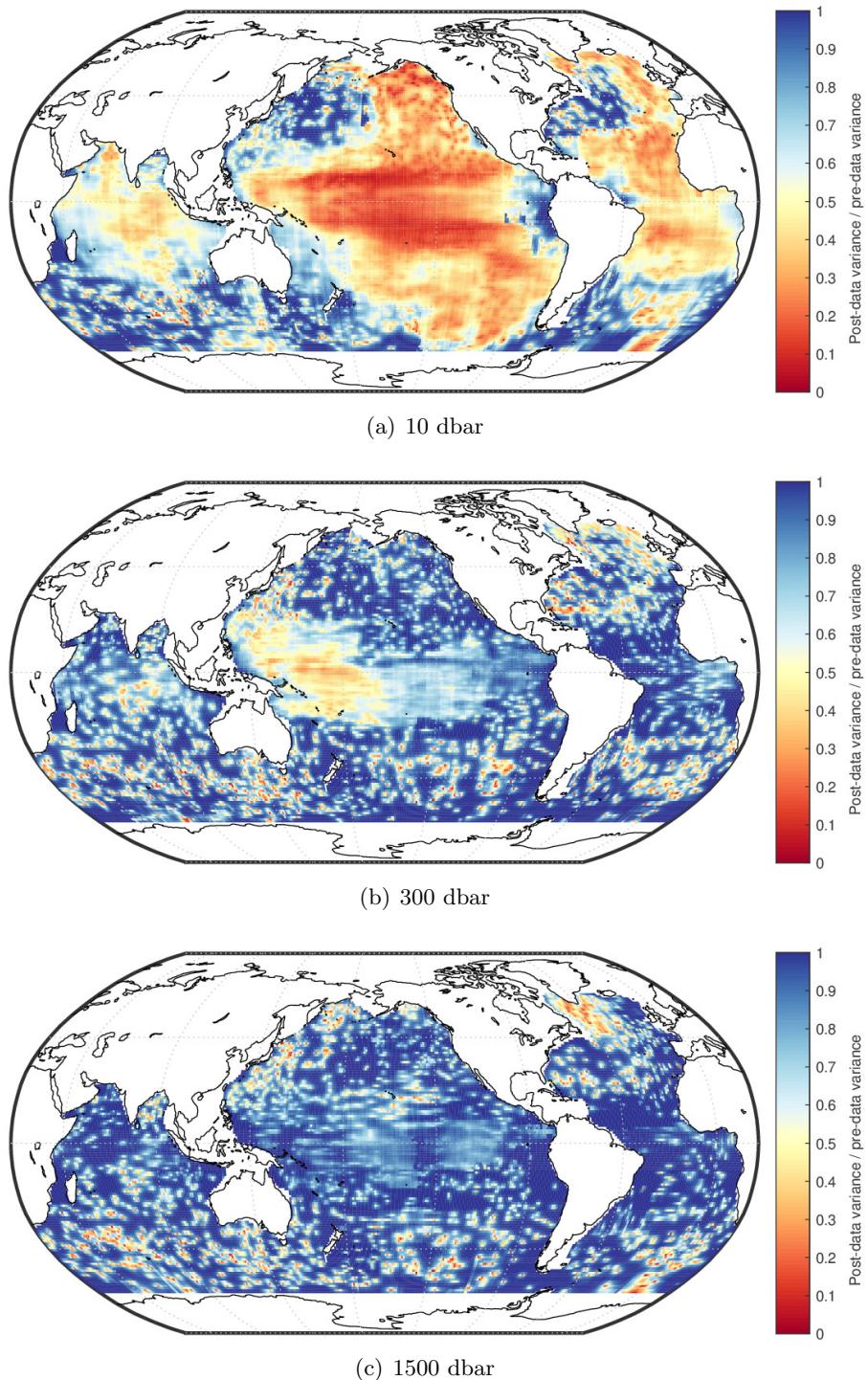
**Figure 8:** Moving-window empirical variance for estimation of  $\phi$  through  $\hat{\phi} = \text{empirical variance}/1.15$ .

## **5.5 Model 2: Anomalies and model parameters**

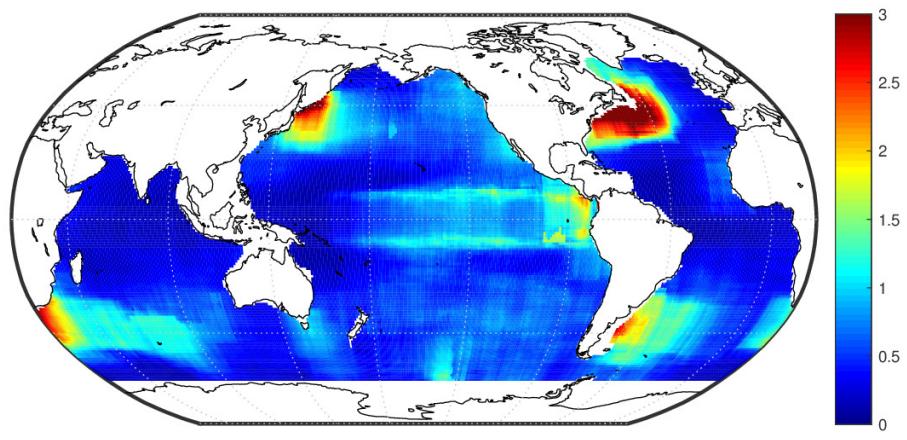
The following pages illustrate the 1-month spatial model with a Gaussian nugget (Model 2).



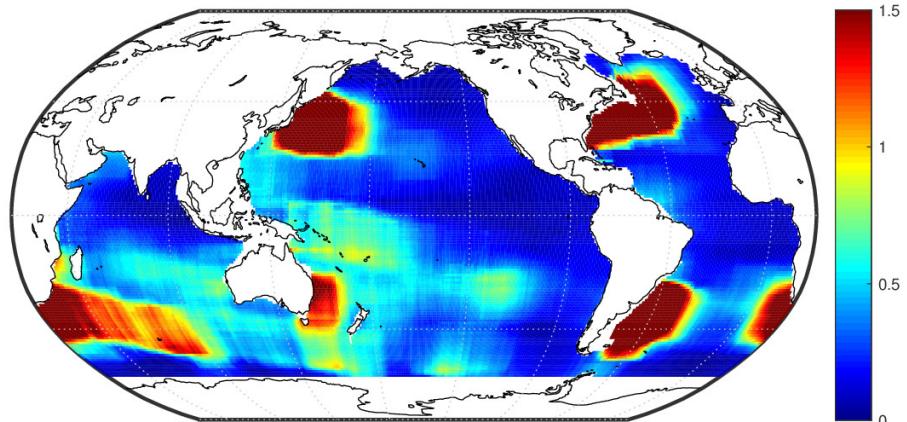
**Figure 9:** February 2012 temperature anomalies



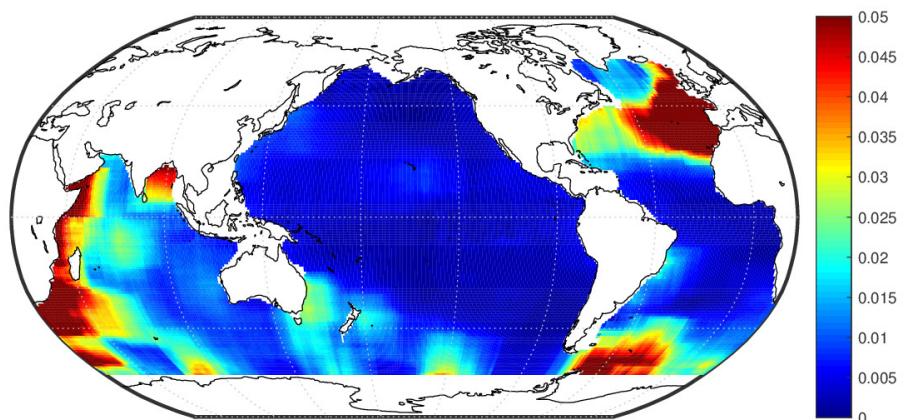
**Figure 10:** February 2012 post-data-to-pre-data variance ratios



(a) 10 dbar

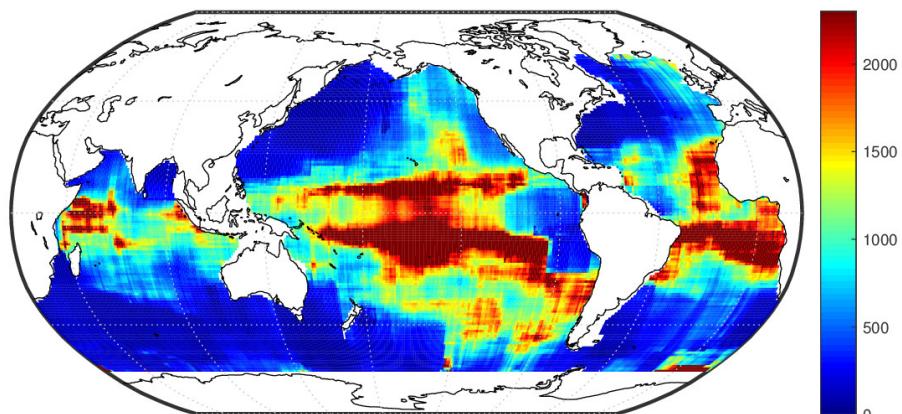


(b) 300 dbar

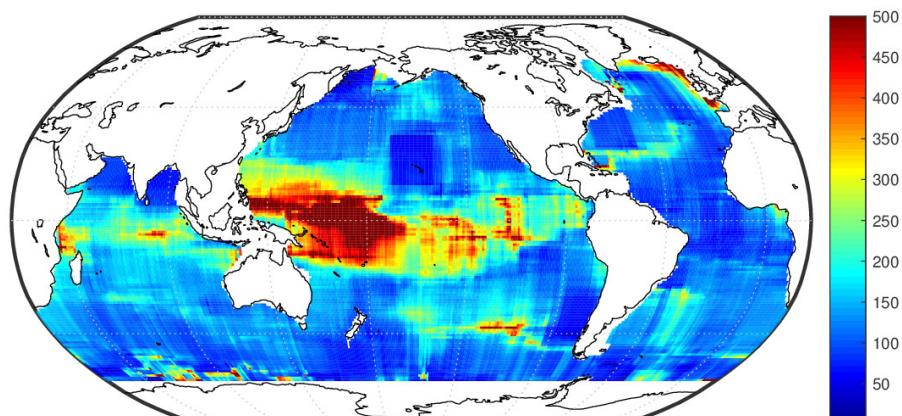


(c) 1500 dbar

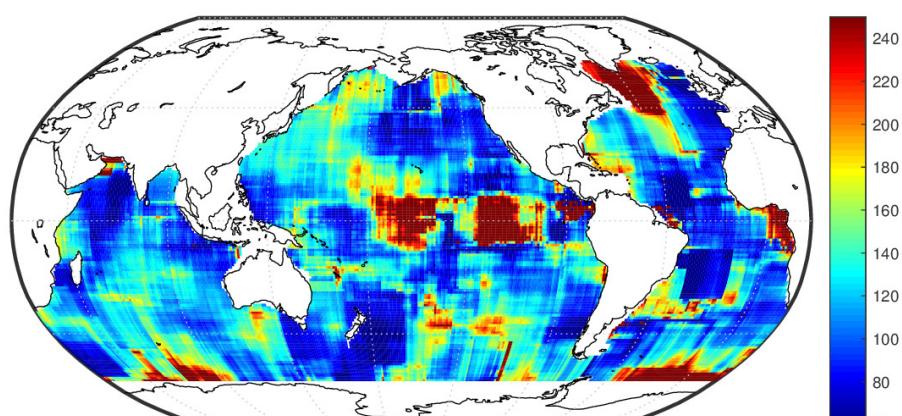
**Figure 11:**  $\phi$



(a) 10 dbar

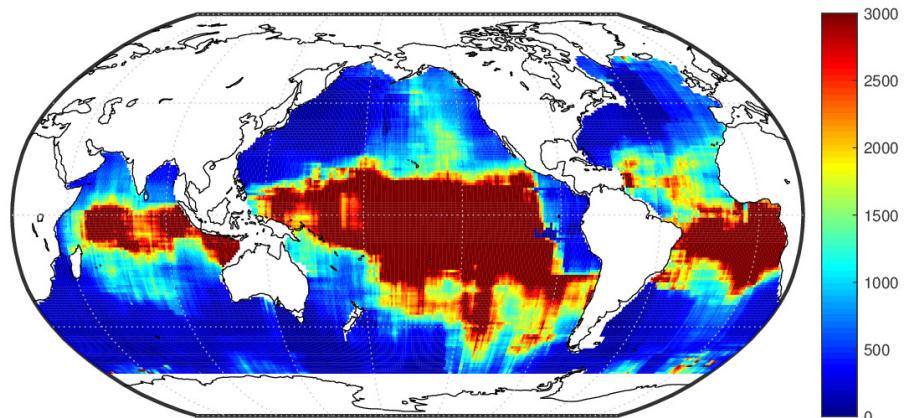


(b) 300 dbar

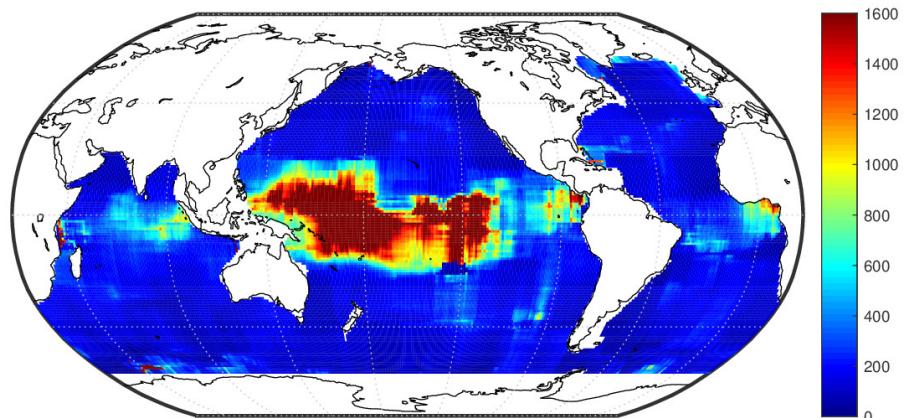


(c) 1500 dbar

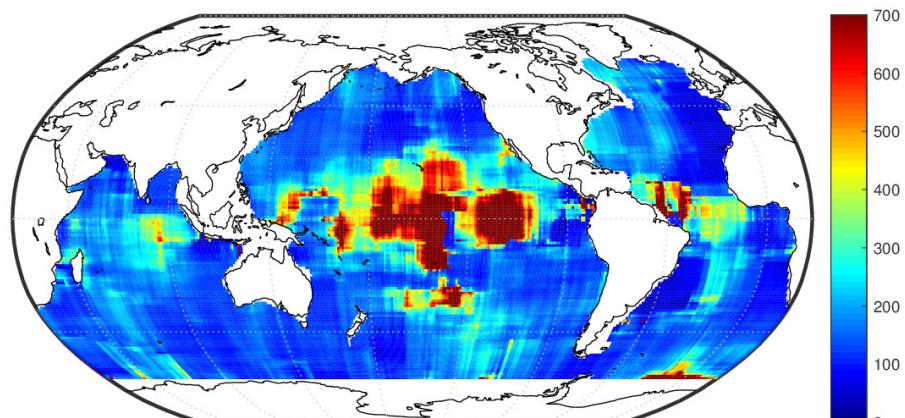
**Figure 12:**  $\theta_{\text{lat}}$  (in km)



(a) 10 dbar

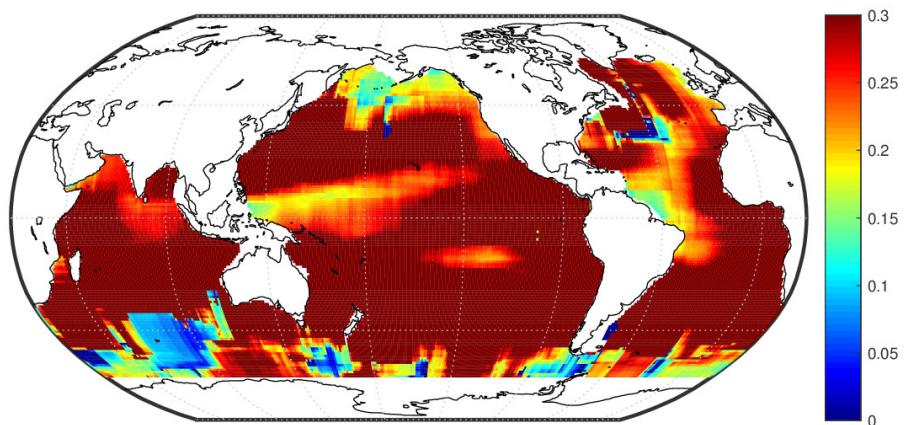


(b) 300 dbar

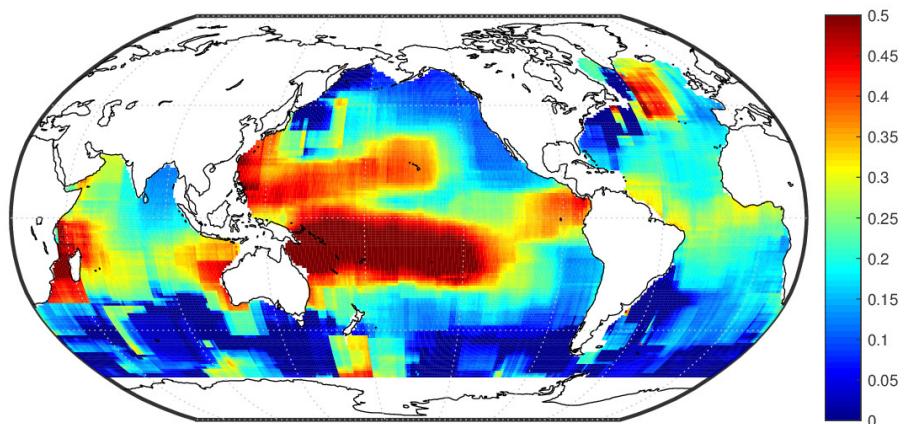


(c) 1500 dbar

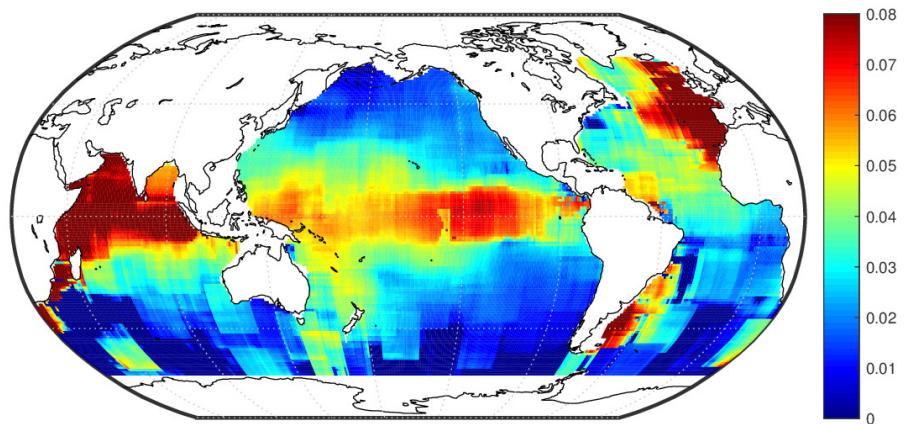
**Figure 13:**  $\theta_{\text{ion}}$  (in km)



(a) 10 dbar

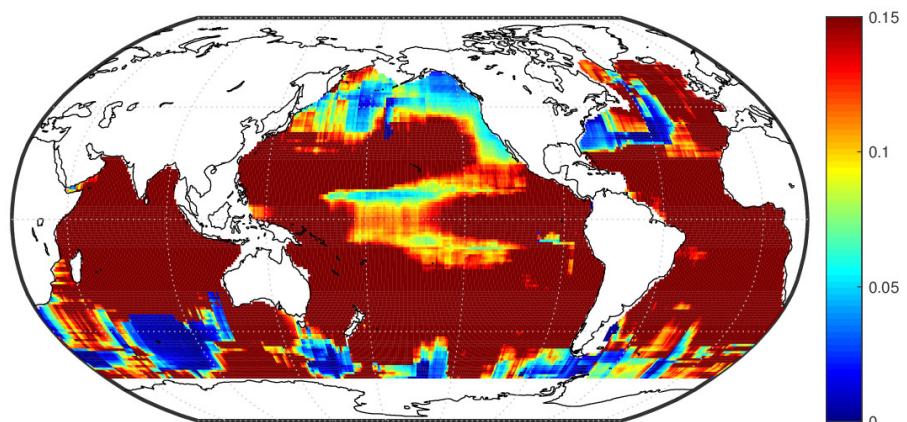


(b) 300 dbar

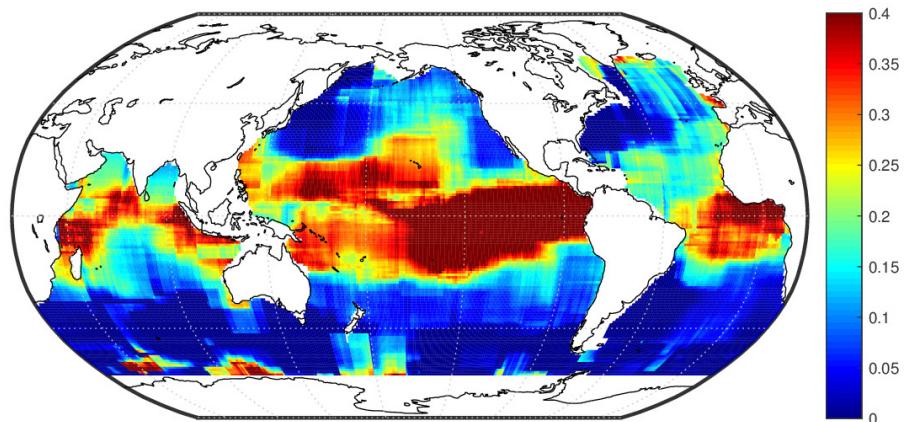


(c) 1500 dbar

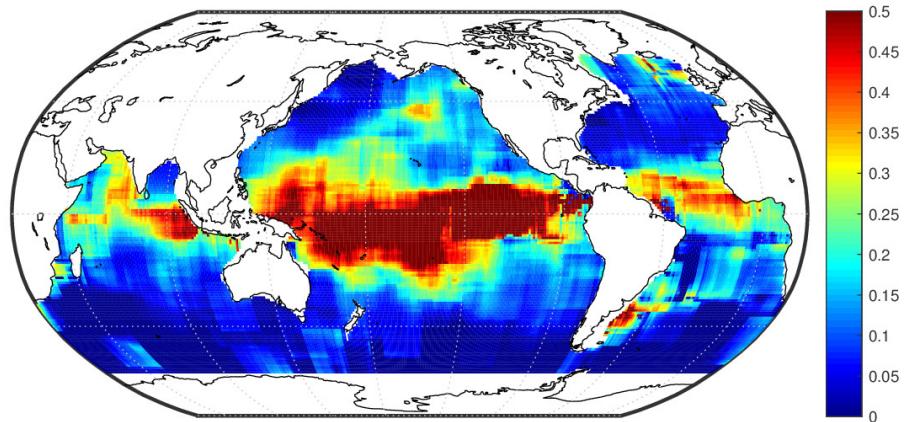
**Figure 14:**  $\sigma$



(a) 10 dbar

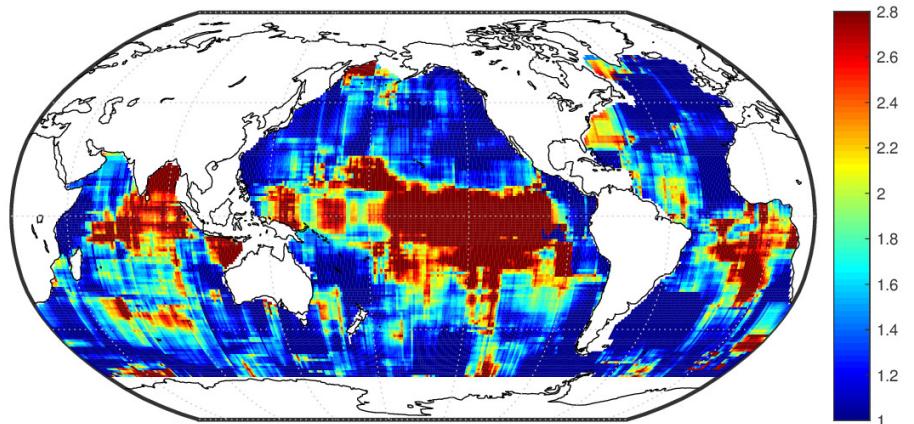


(b) 300 dbar

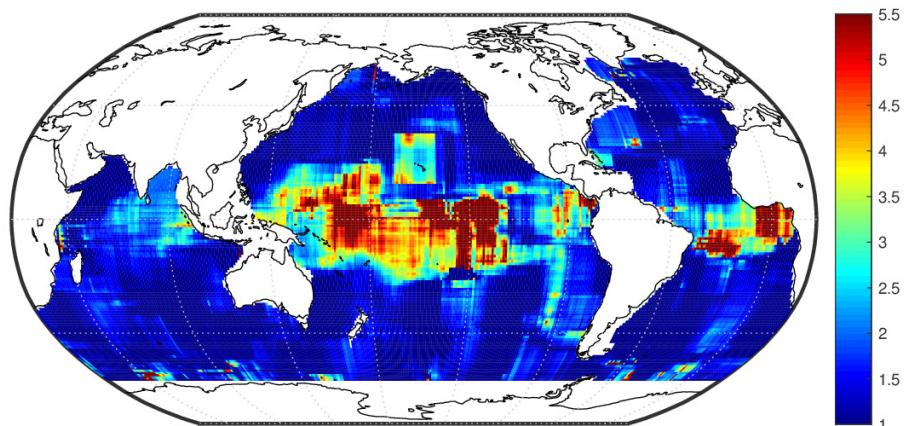


(c) 1500 dbar

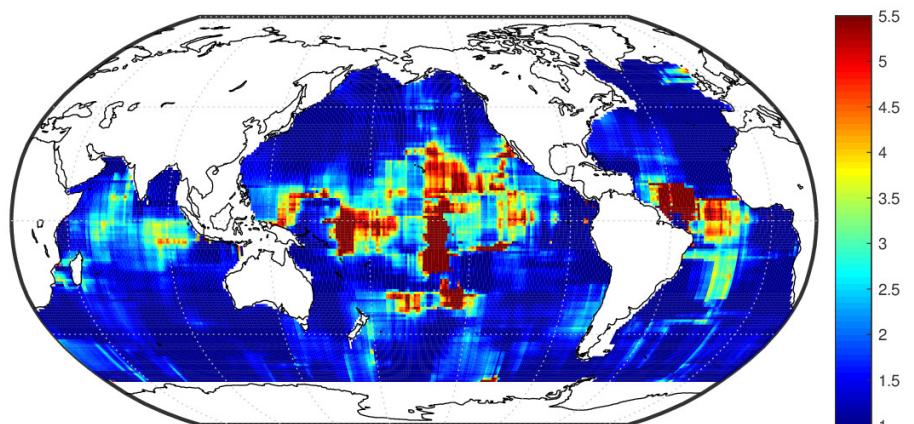
**Figure 15:**  $\sigma^2 / (\phi + \sigma^2)$



(a) 10 dbar

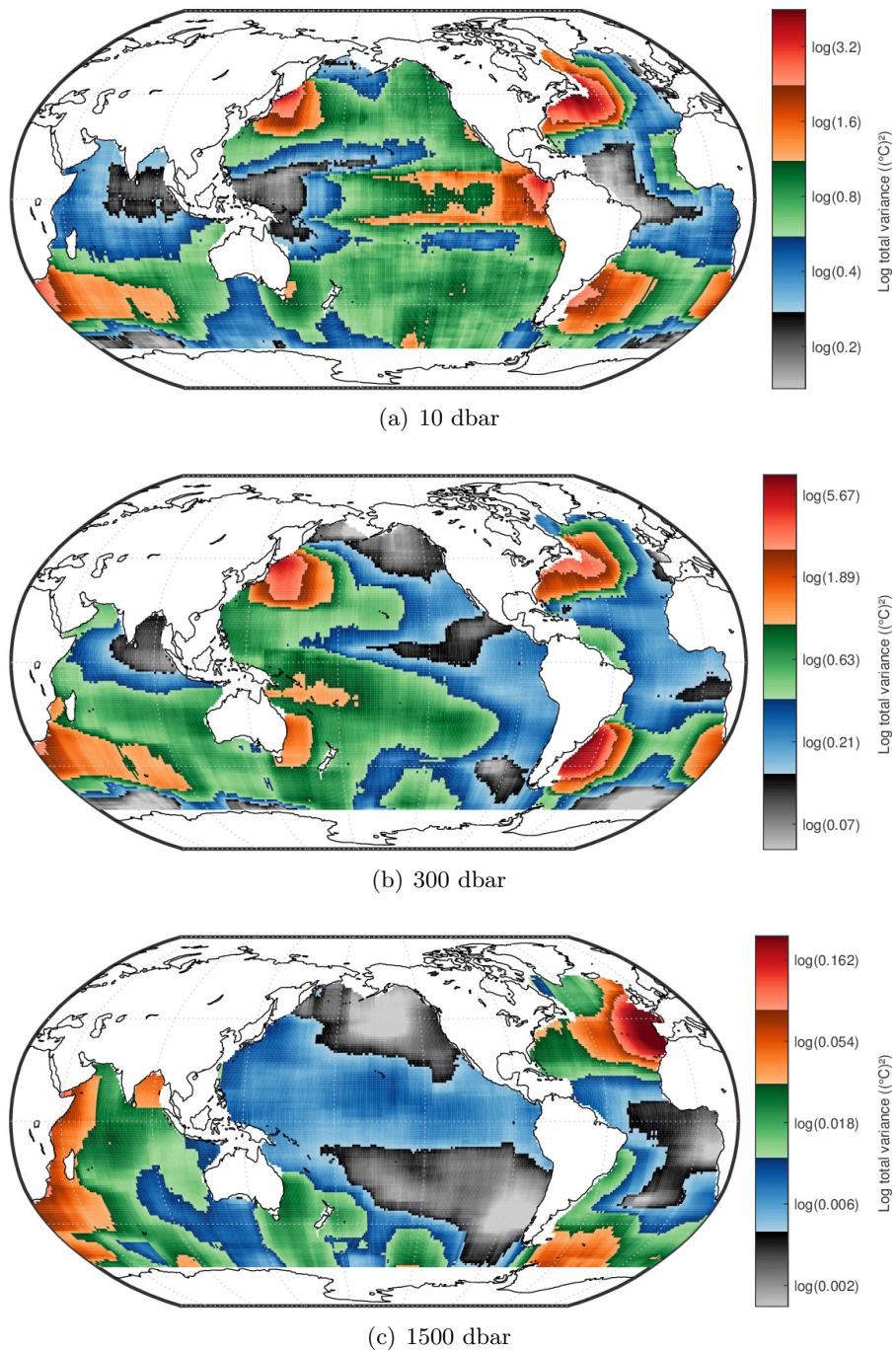


(b) 300 dbar

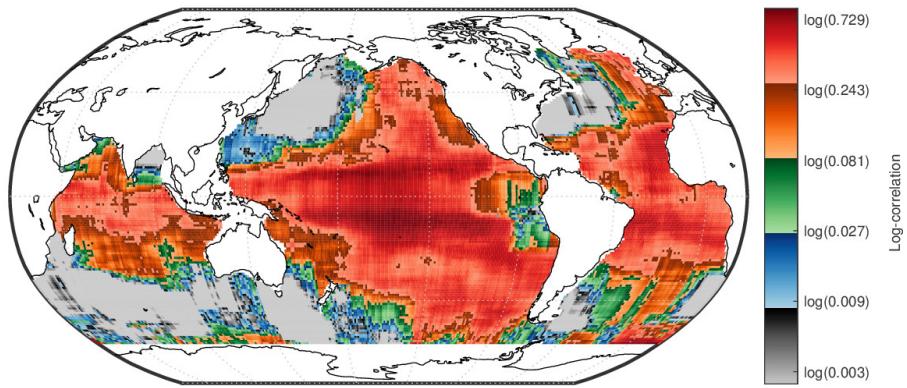


(c) 1500 dbar

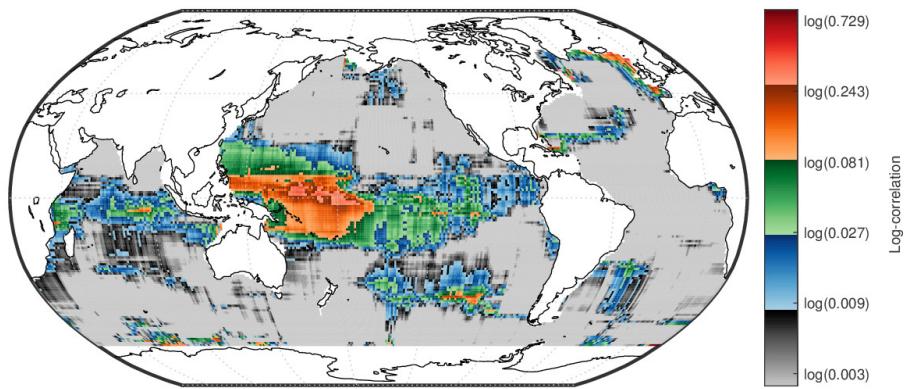
**Figure 16:**  $\theta_{\text{lon}}/\theta_{\text{lat}}$



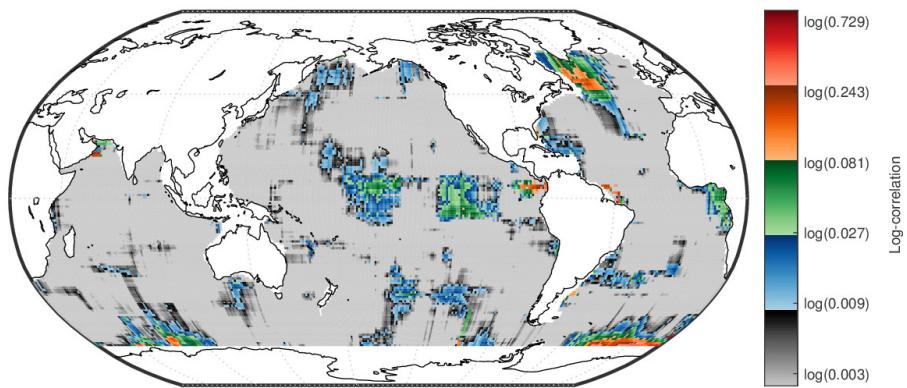
**Figure 17:**  $\phi + \sigma^2$



(a) 10 dbar

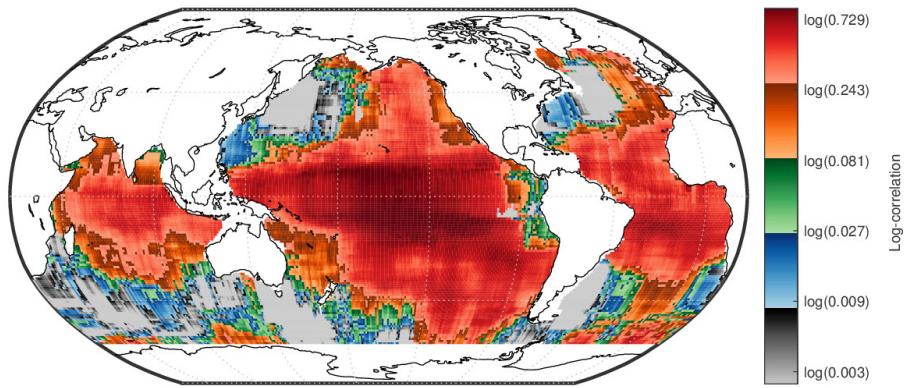


(b) 300 dbar

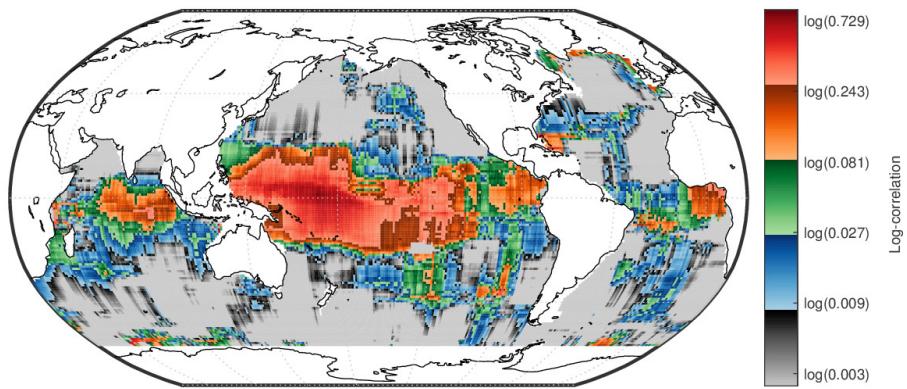


(c) 1500 dbar

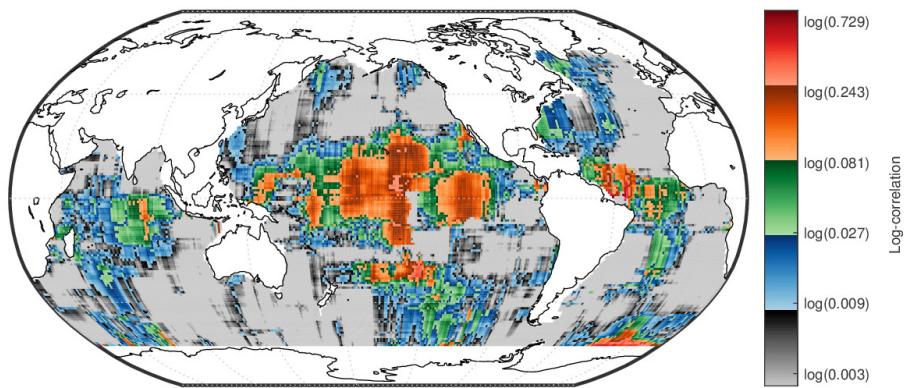
**Figure 18:** Correlation at  $\Delta x_{\text{lat}} = 800 \text{ km}$



(a) 10 dbar



(b) 300 dbar

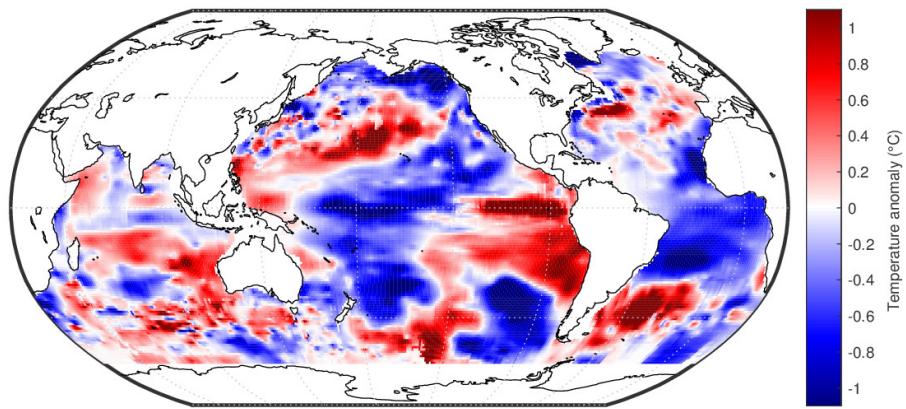


(c) 1500 dbar

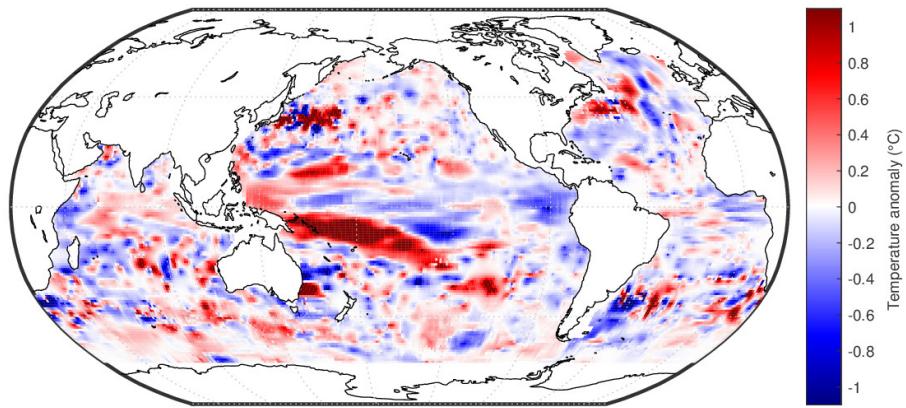
**Figure 19:** Correlation at  $\Delta x_{\text{lon}} = 800 \text{ km}$

## 5.6 Model 3: Anomalies and model parameters

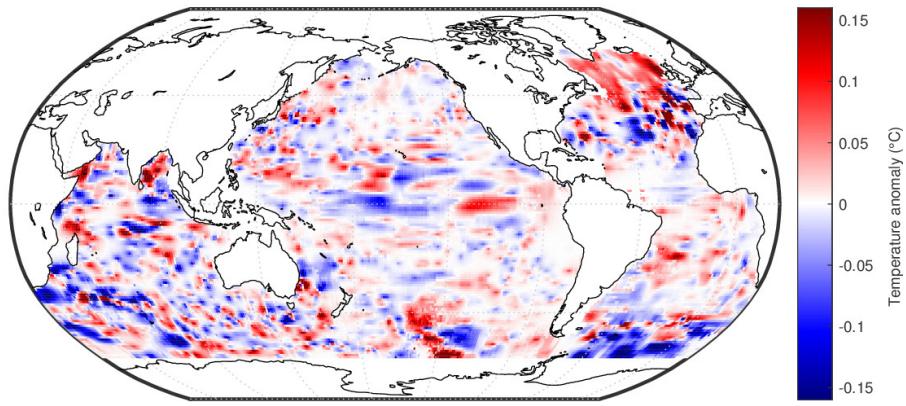
The following pages illustrate the 1-month spatial model with a Student nugget (Model 3). Quantities that require the existence of the second moment for the Student nugget are masked out when  $\hat{\nu} \leq 2$ .



(a) 10 dbar

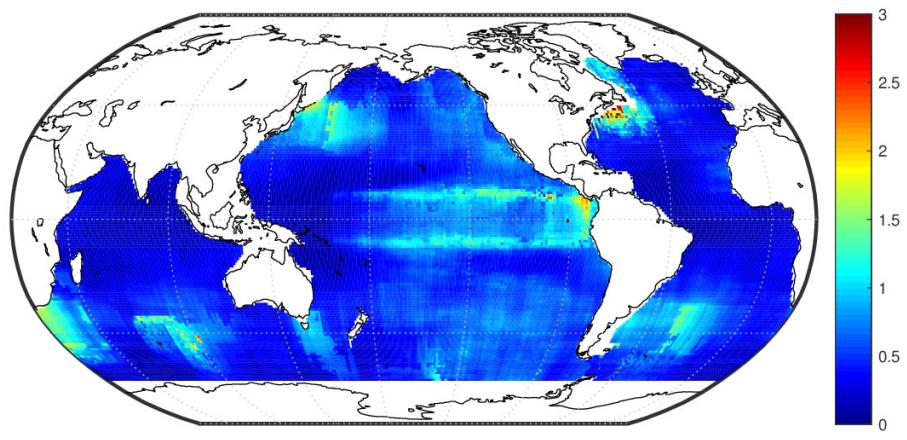


(b) 300 dbar

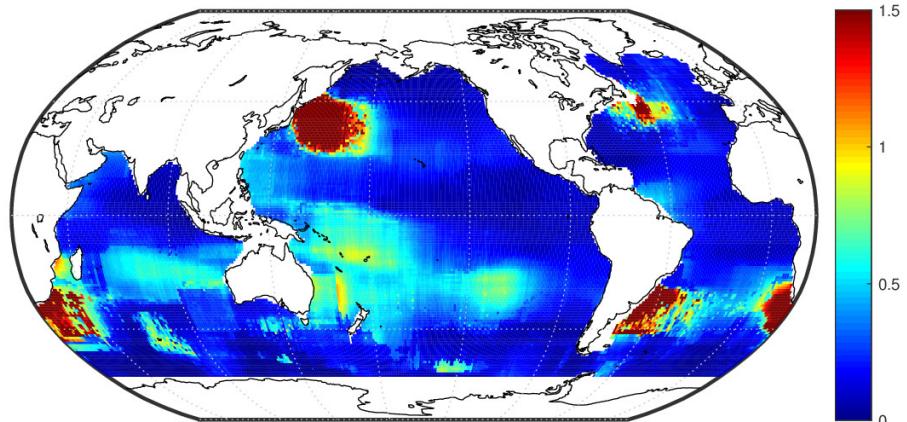


(c) 1500 dbar

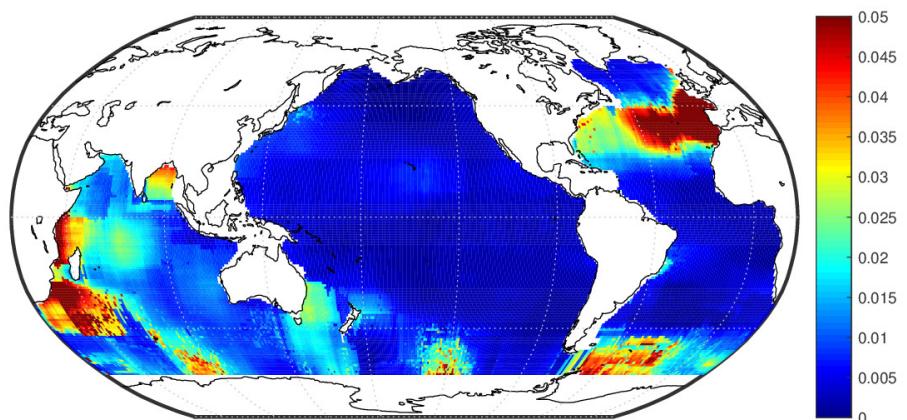
**Figure 20:** February 2012 temperature anomalies



(a) 10 dbar

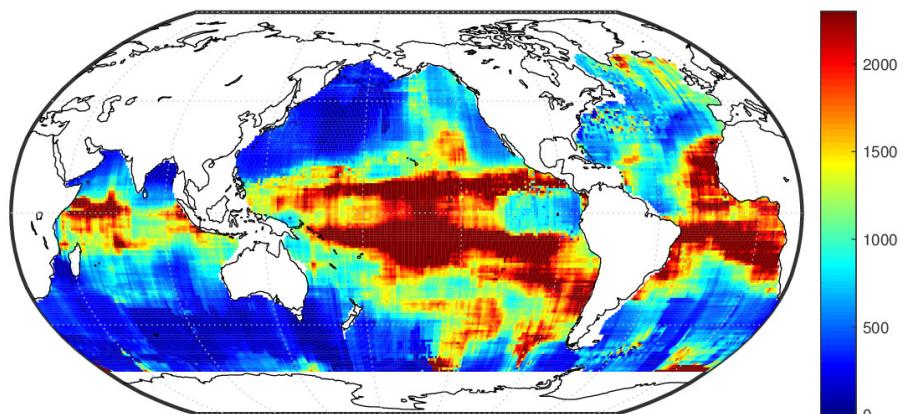


(b) 300 dbar

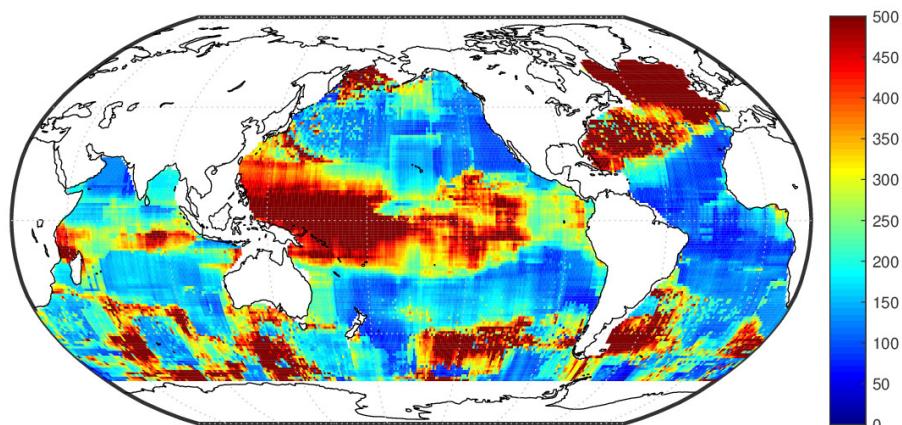


(c) 1500 dbar

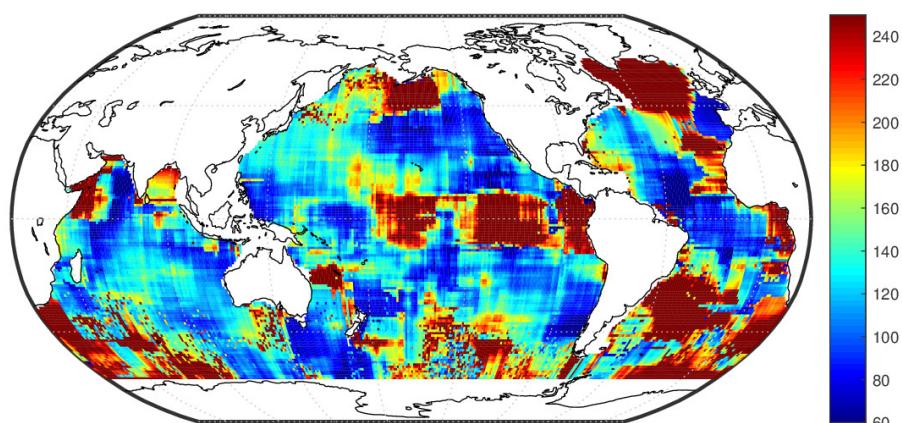
**Figure 21:**  $\phi$



(a) 10 dbar

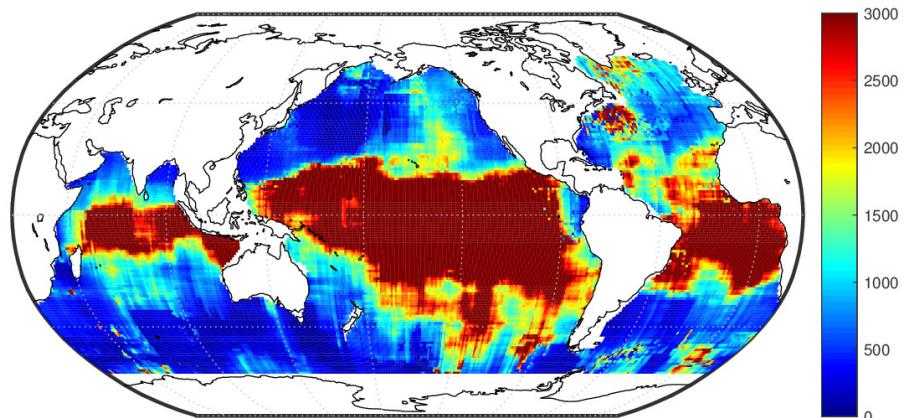


(b) 300 dbar

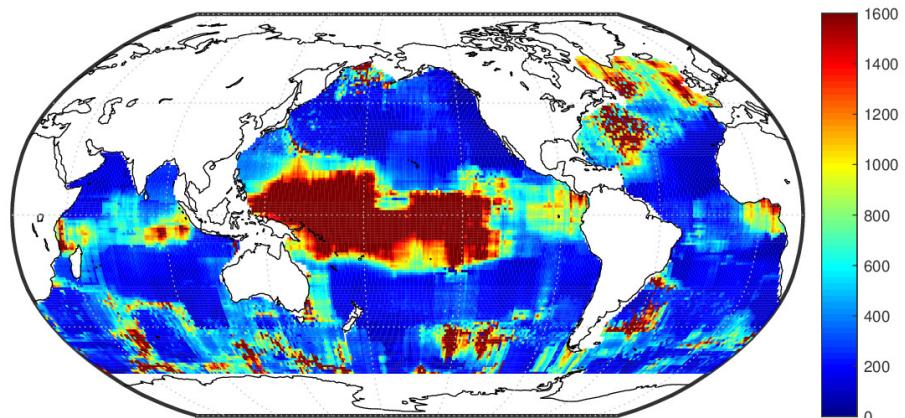


(c) 1500 dbar

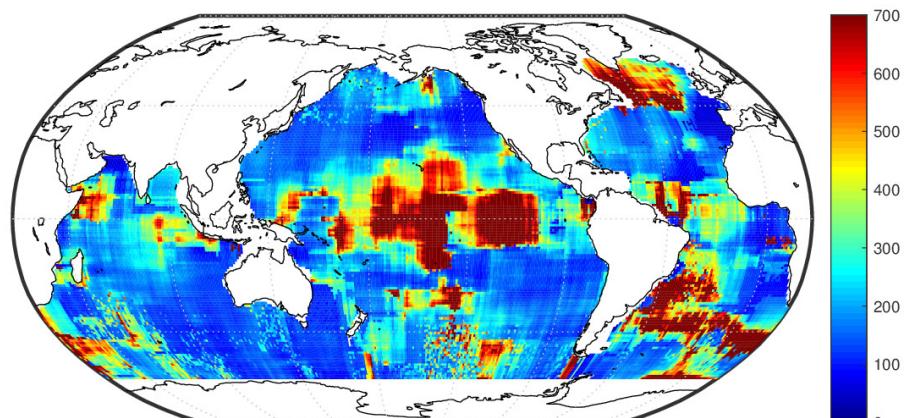
**Figure 22:**  $\theta_{\text{lat}}$  (in km)



(a) 10 dbar

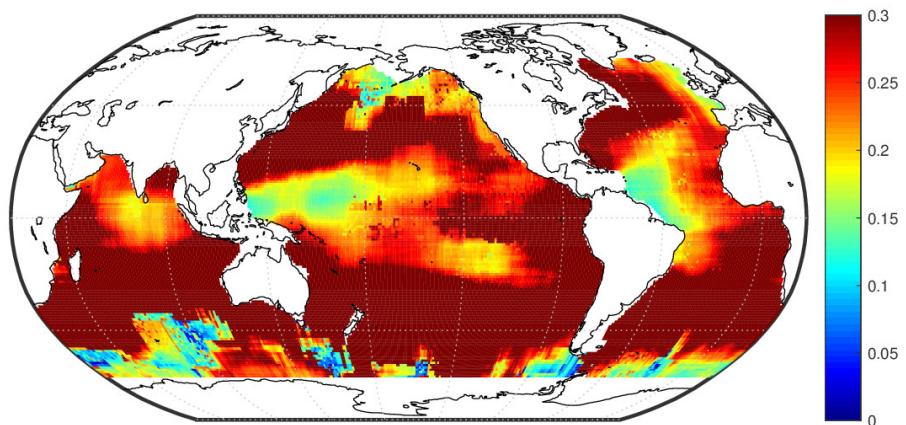


(b) 300 dbar

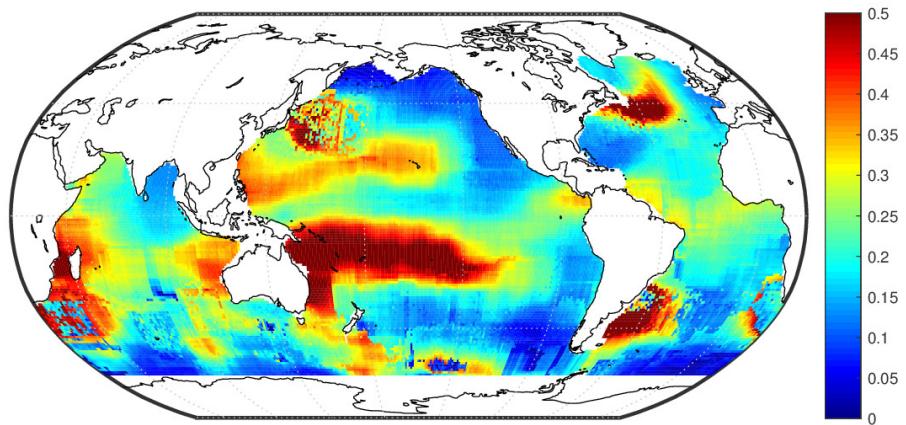


(c) 1500 dbar

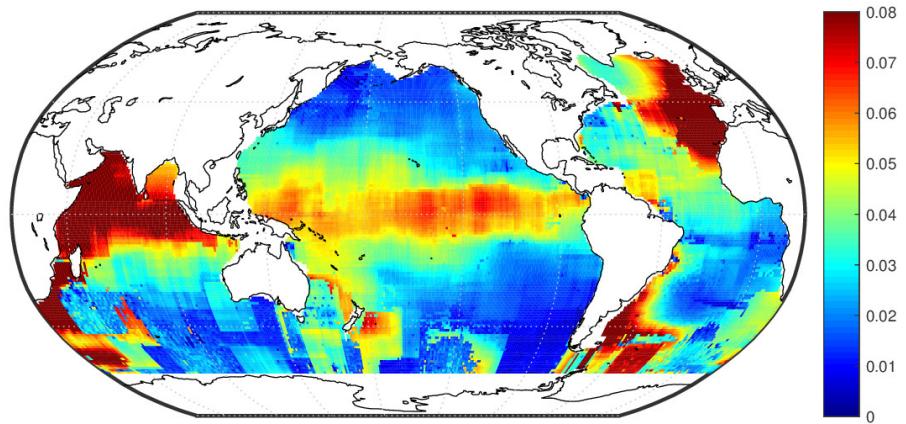
**Figure 23:**  $\theta_{\text{ion}}$  (in km)



(a) 10 dbar

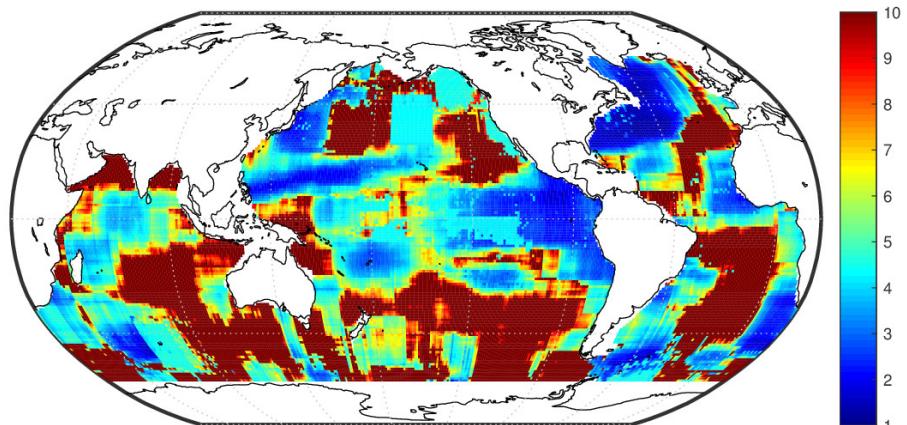


(b) 300 dbar

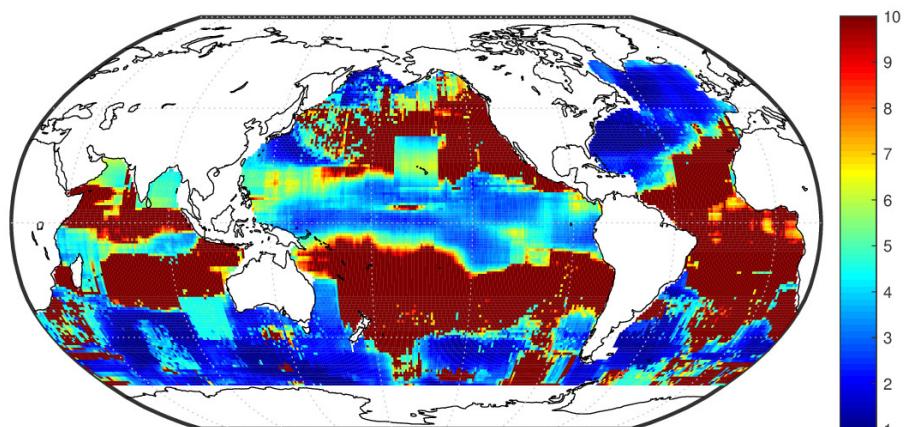


(c) 1500 dbar

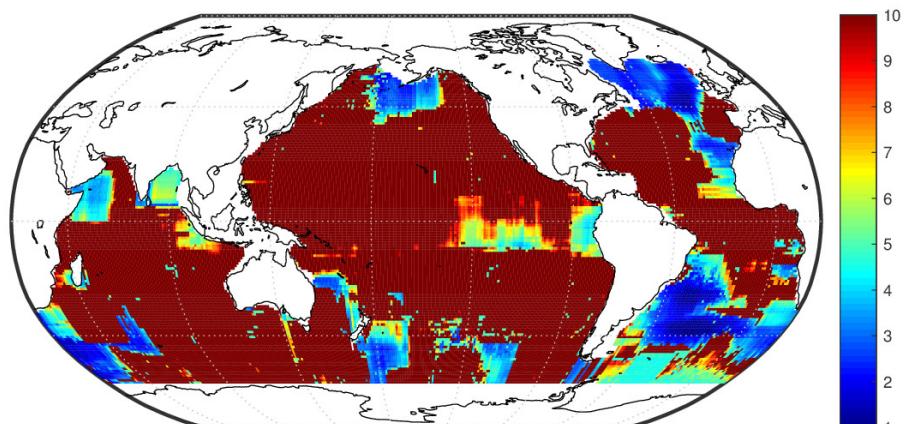
**Figure 24:**  $\sigma$



(a) 10 dbar

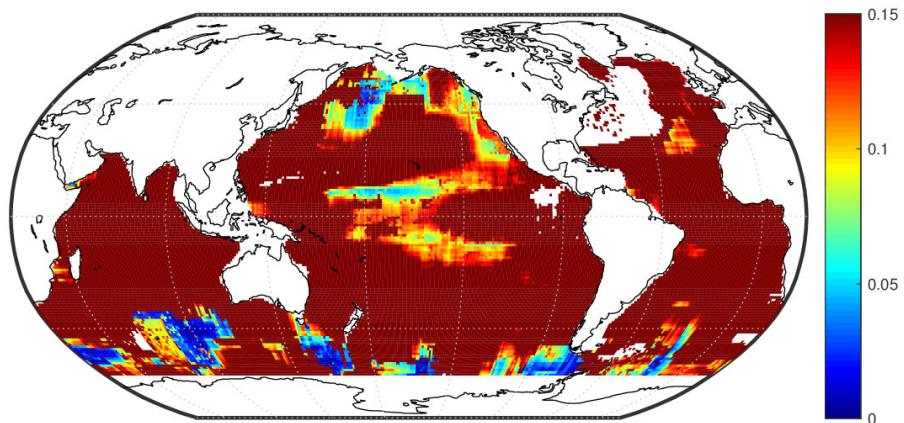


(b) 300 dbar

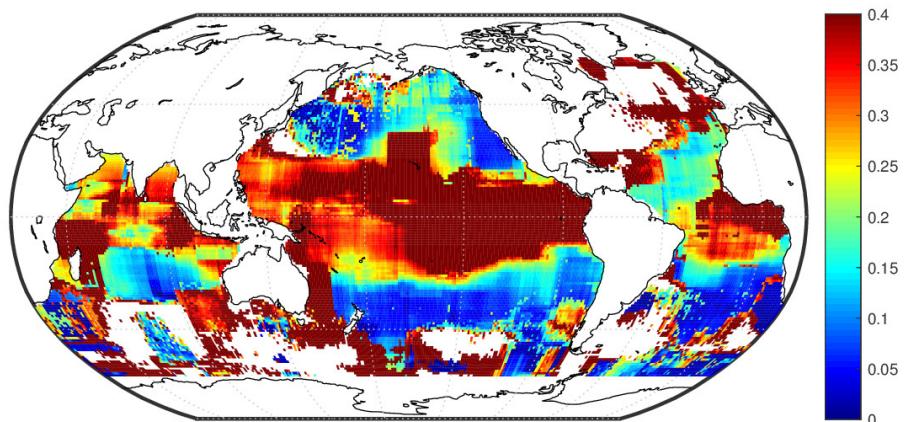


(c) 1500 dbar

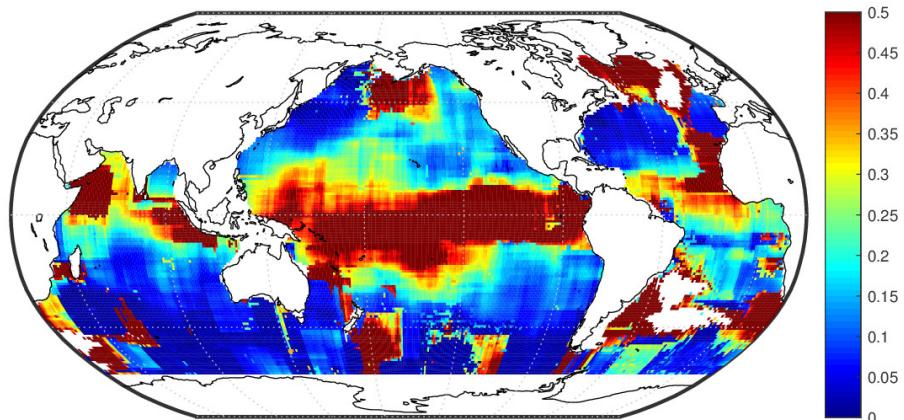
**Figure 25:**  $\nu$



(a) 10 dbar

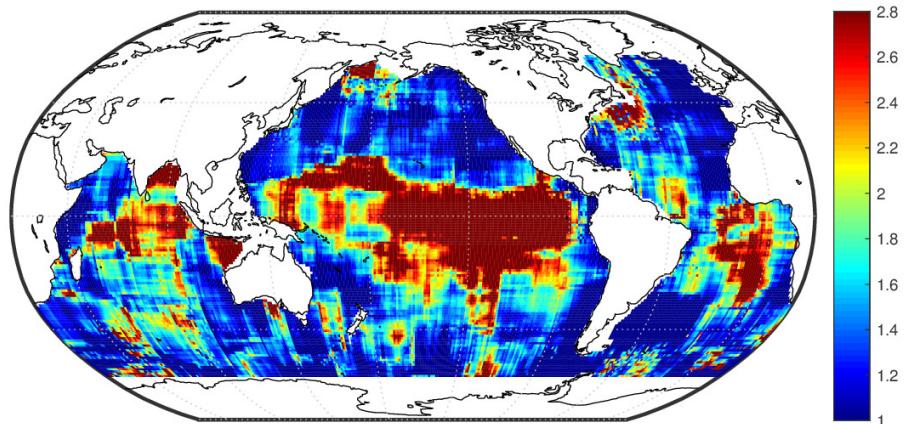


(b) 300 dbar

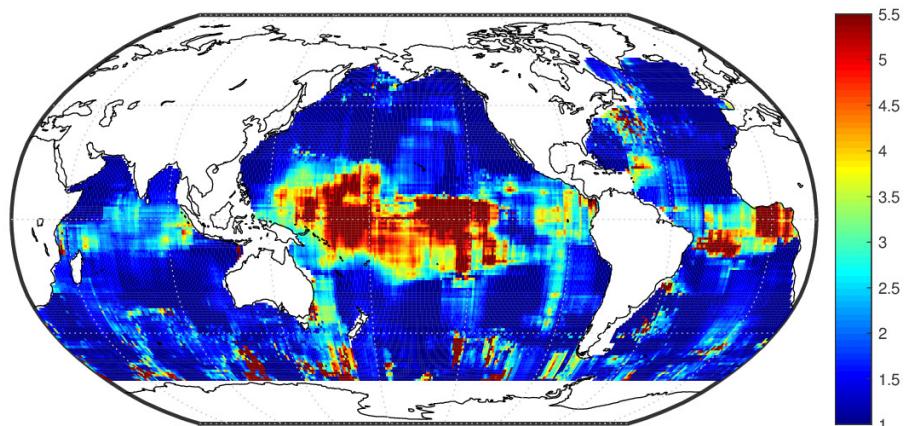


(c) 1500 dbar

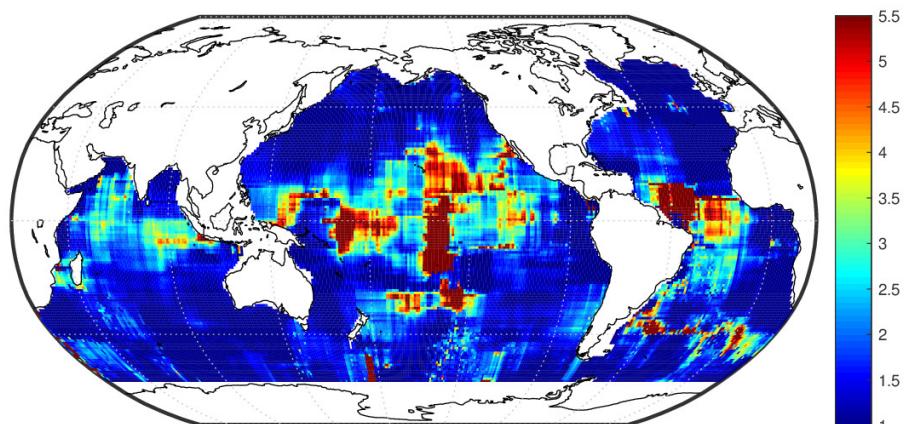
**Figure 26:**  $\frac{\nu}{\nu-2}\sigma^2 / \left( \phi + \frac{\nu}{\nu-2}\sigma^2 \right)$



(a) 10 dbar

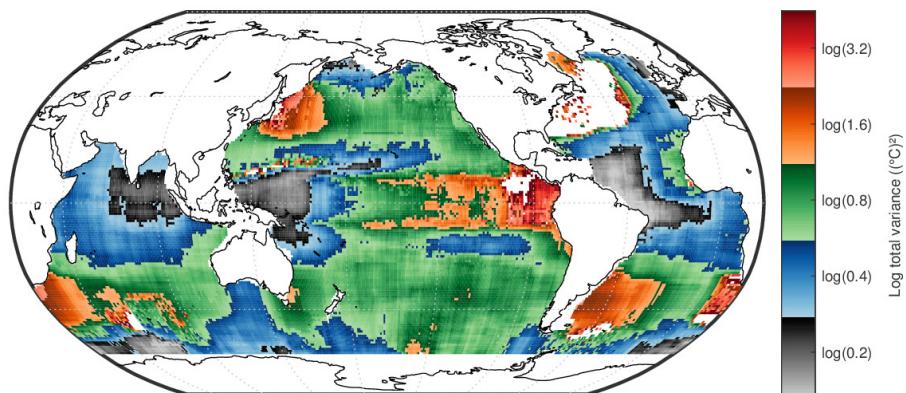


(b) 300 dbar

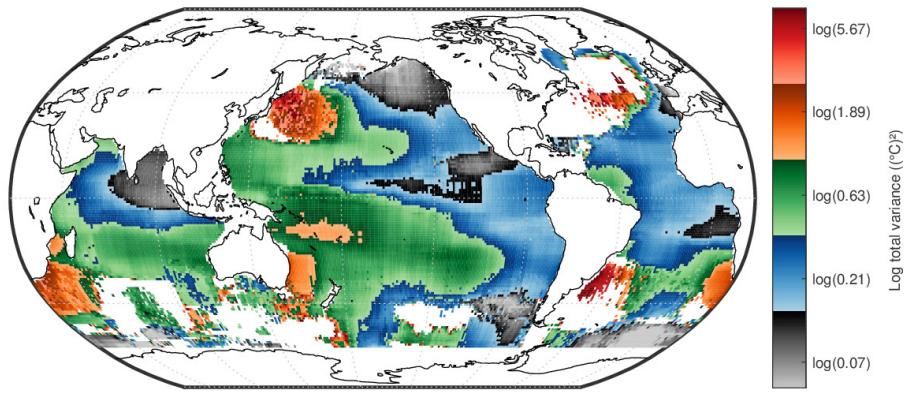


(c) 1500 dbar

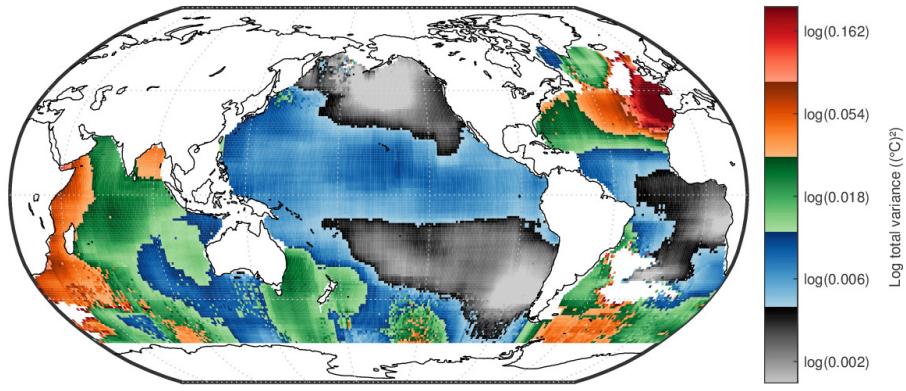
**Figure 27:**  $\theta_{\text{lon}}/\theta_{\text{lat}}$



(a) 10 dbar

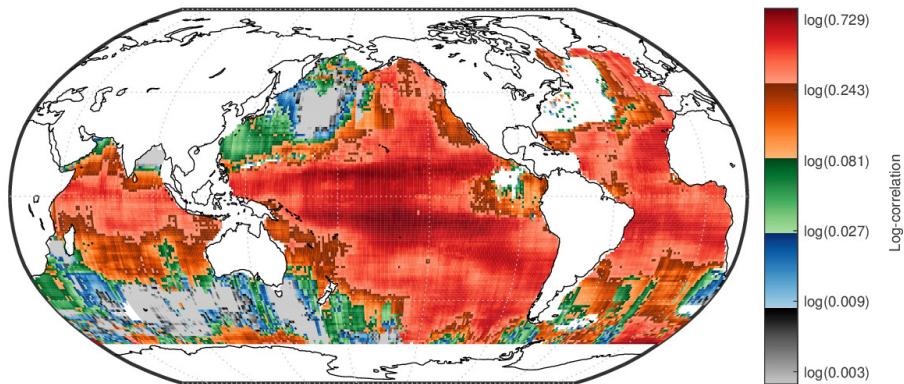


(b) 300 dbar

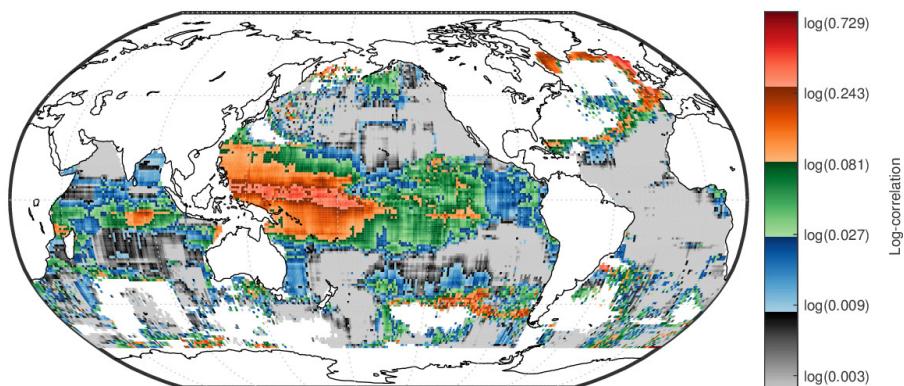


(c) 1500 dbar

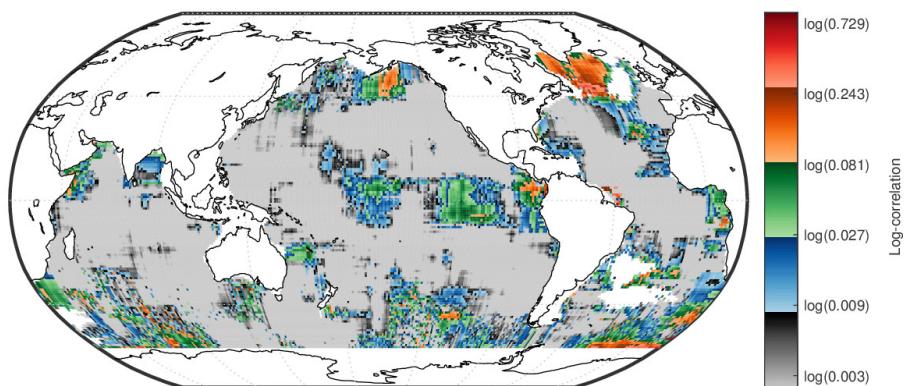
**Figure 28:**  $\phi + \frac{\nu}{\nu-2} \sigma^2$



(a) 10 dbar

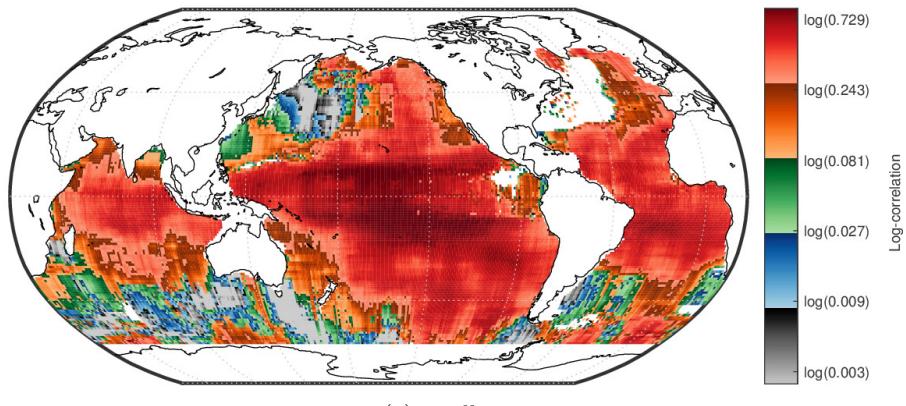


(b) 300 dbar

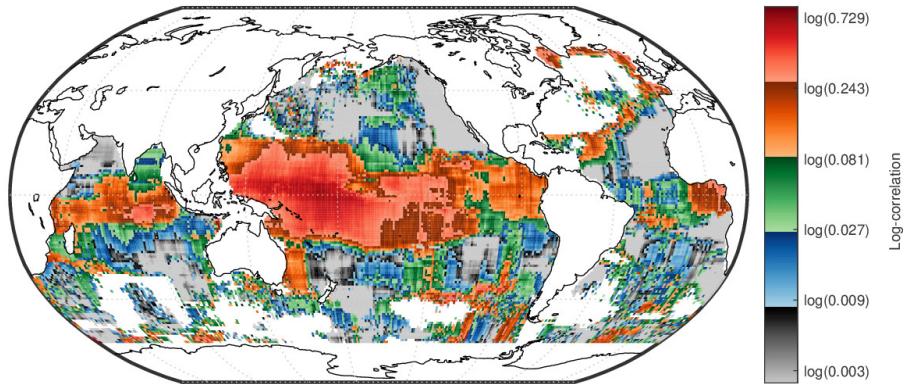


(c) 1500 dbar

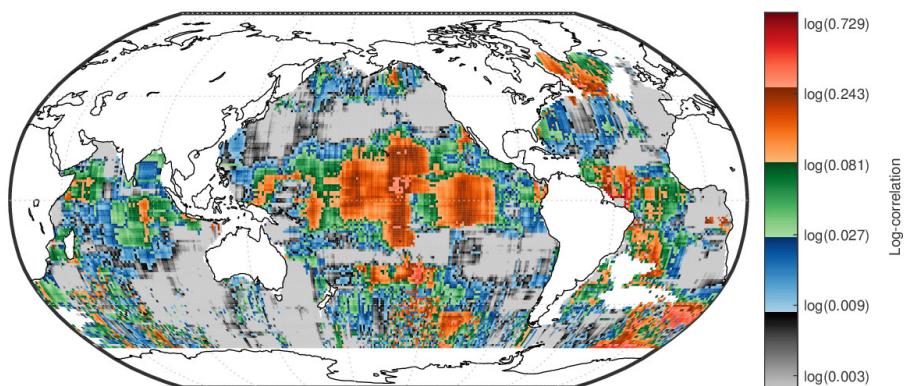
**Figure 29:** Correlation at  $\Delta x_{\text{lat}} = 800 \text{ km}$



(a) 10 dbar



(b) 300 dbar

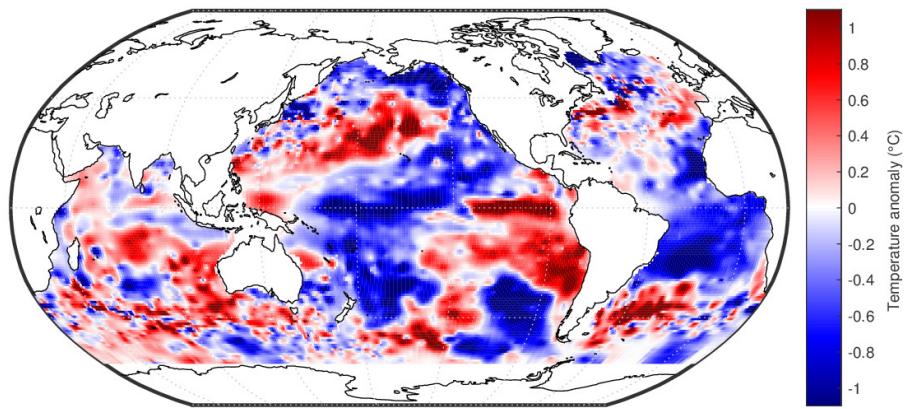


(c) 1500 dbar

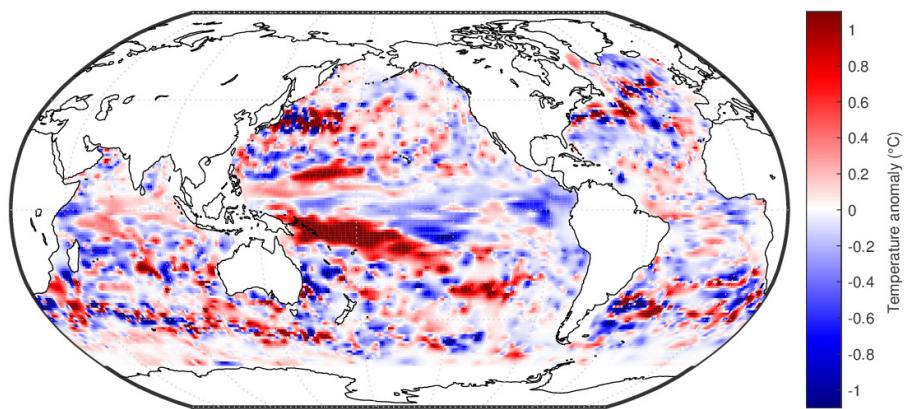
**Figure 30:** Correlation at  $\Delta x_{\text{lon}} = 800 \text{ km}$

## **5.7 Model 5: Anomalies and model parameters**

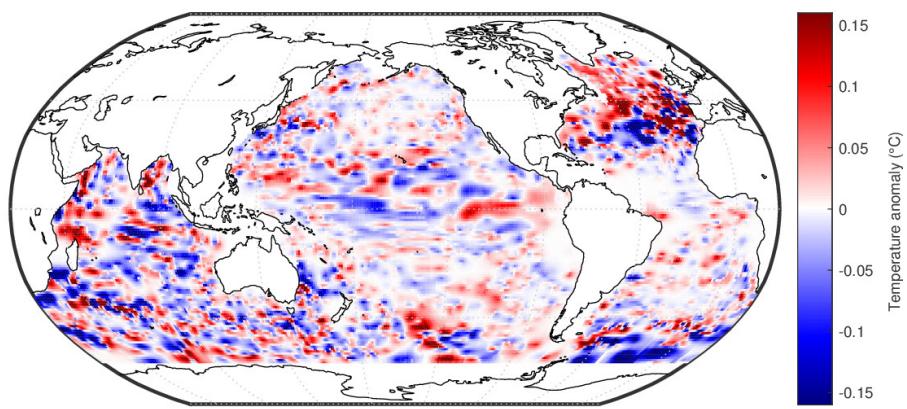
The following pages illustrate the 3-month spatio-temporal model with a Gaussian nugget (Model 5).



(a) 10 dbar

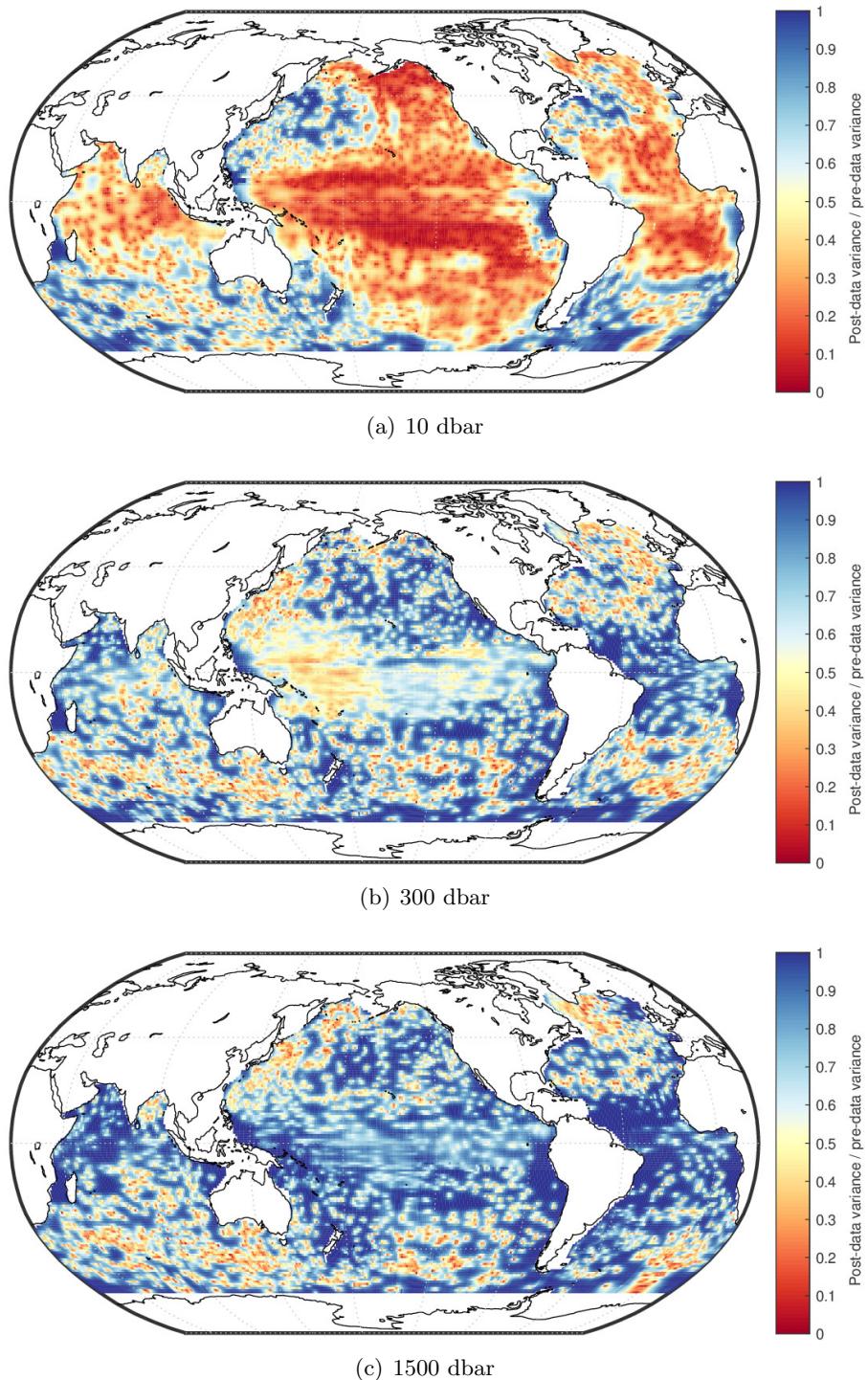


(b) 300 dbar

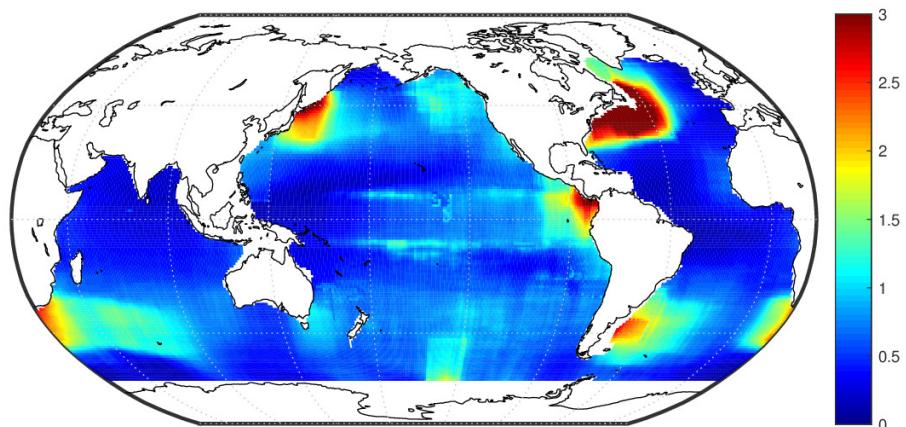


(c) 1500 dbar

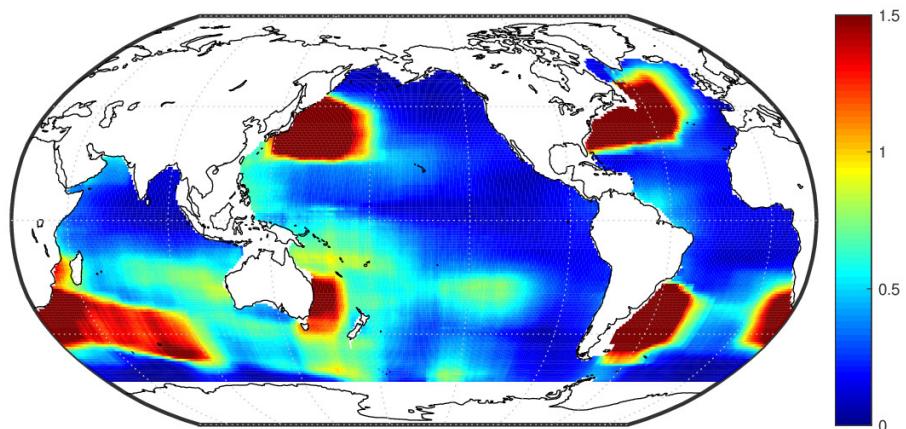
**Figure 31:** February 2012 temperature anomalies



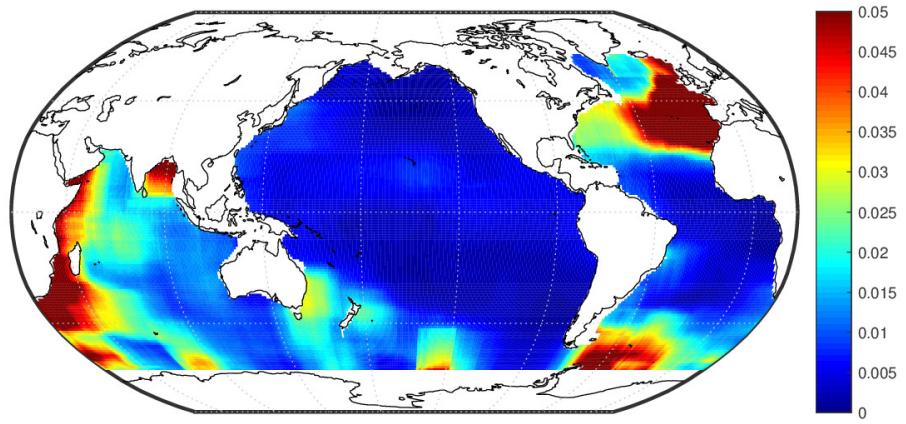
**Figure 32:** February 2012 post-data-to-pre-data variance ratios



(a) 10 dbar

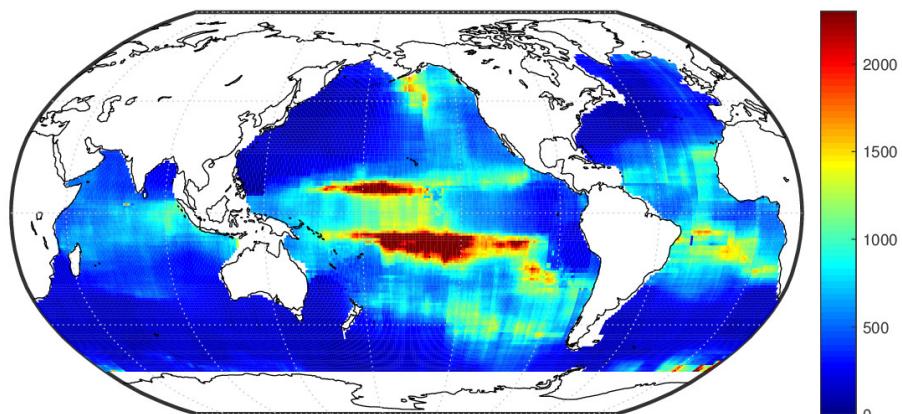


(b) 300 dbar

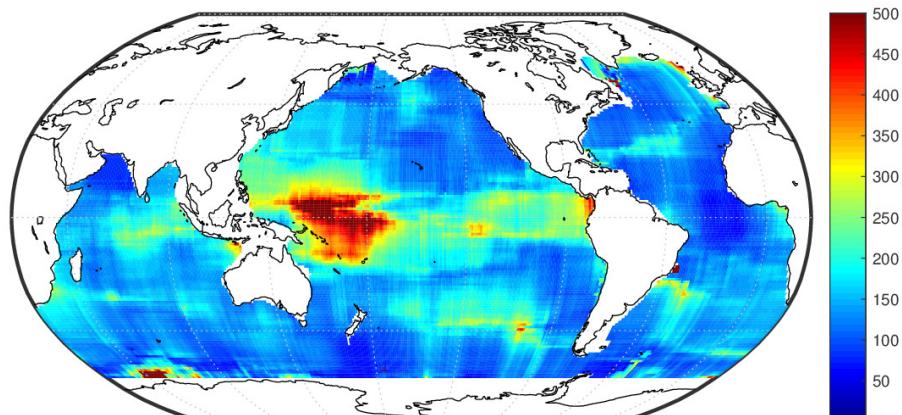


(c) 1500 dbar

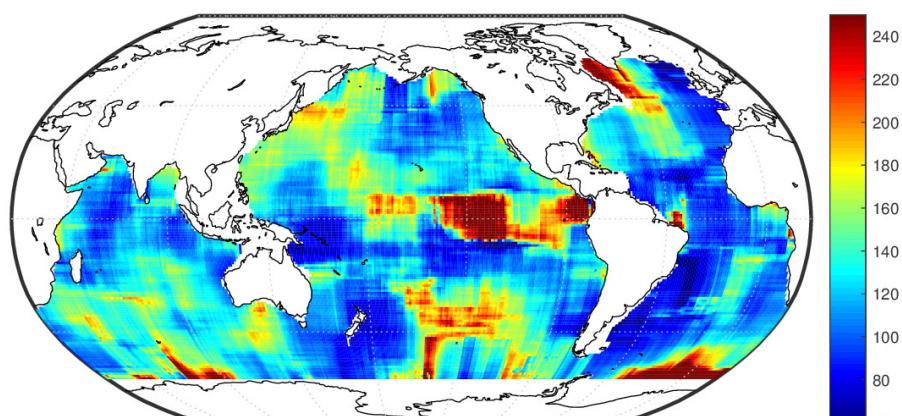
**Figure 33:**  $\phi$



(a) 10 dbar

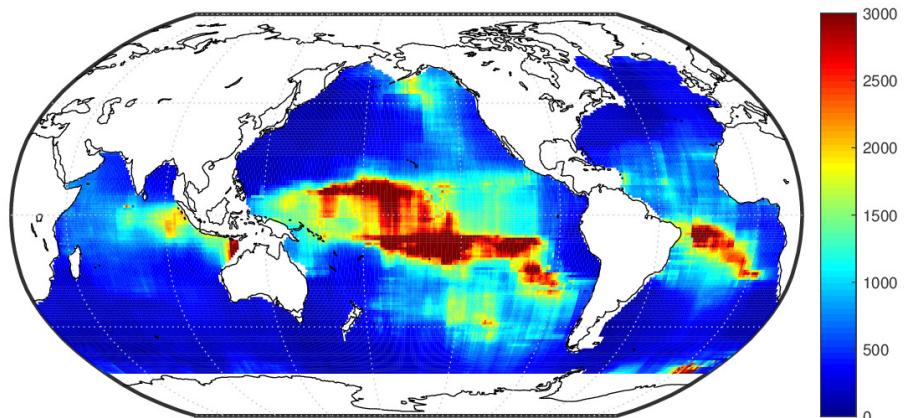


(b) 300 dbar

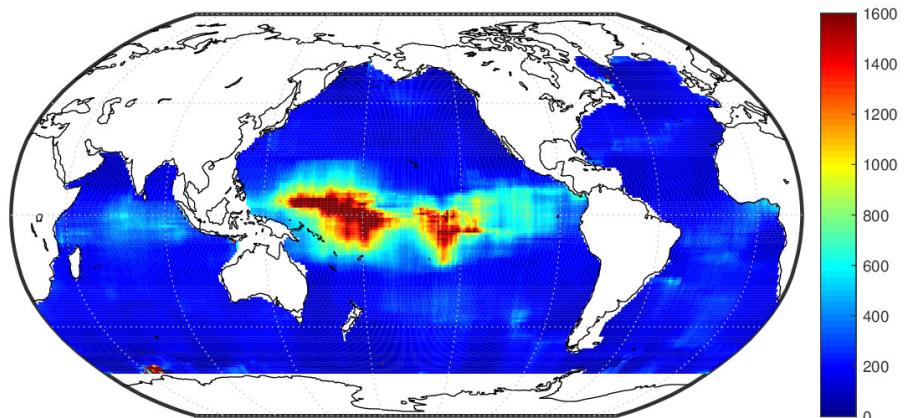


(c) 1500 dbar

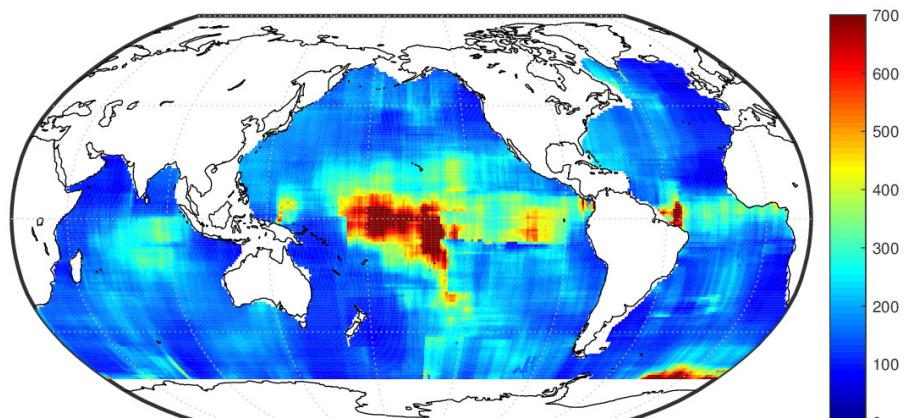
**Figure 34:**  $\theta_{\text{lat}}$  (in km)



(a) 10 dbar

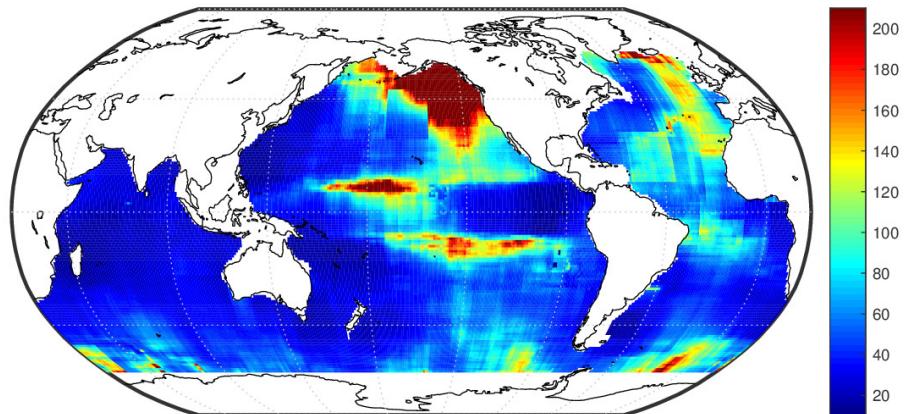


(b) 300 dbar

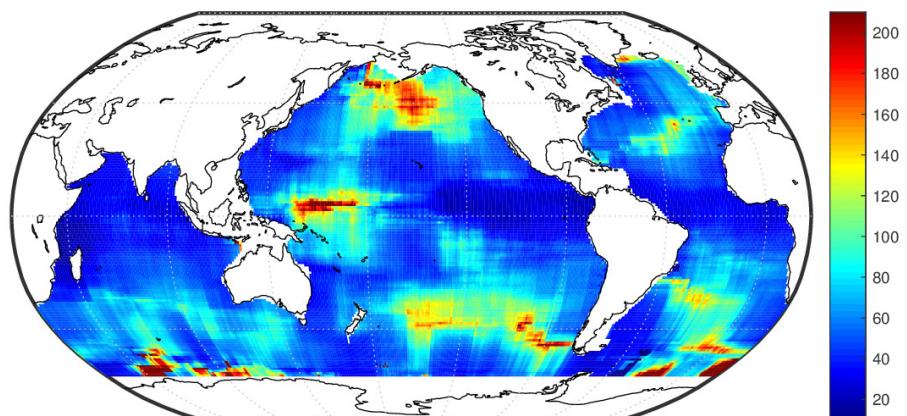


(c) 1500 dbar

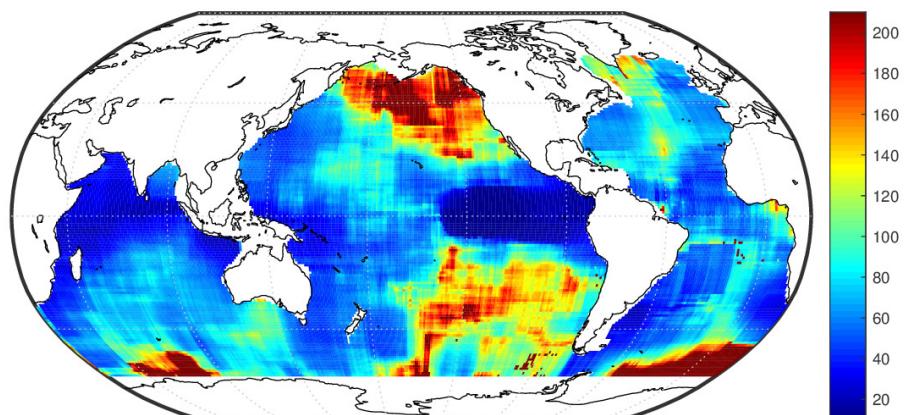
**Figure 35:**  $\theta_{\text{ion}}$  (in km)



(a) 10 dbar

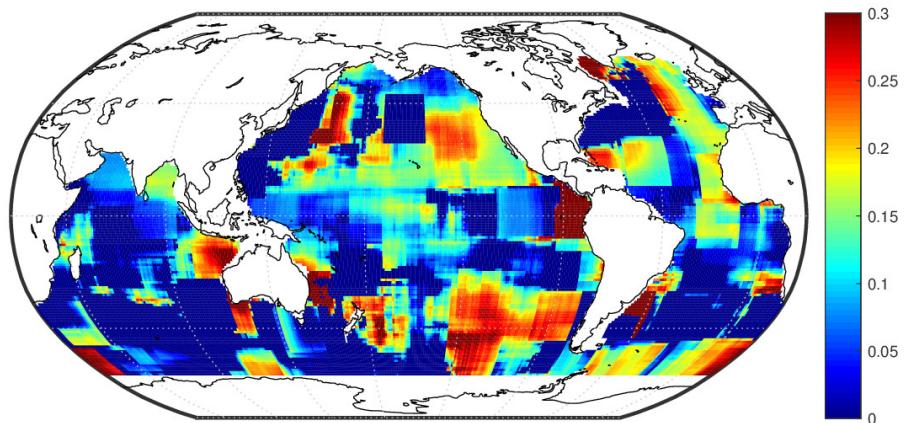


(b) 300 dbar

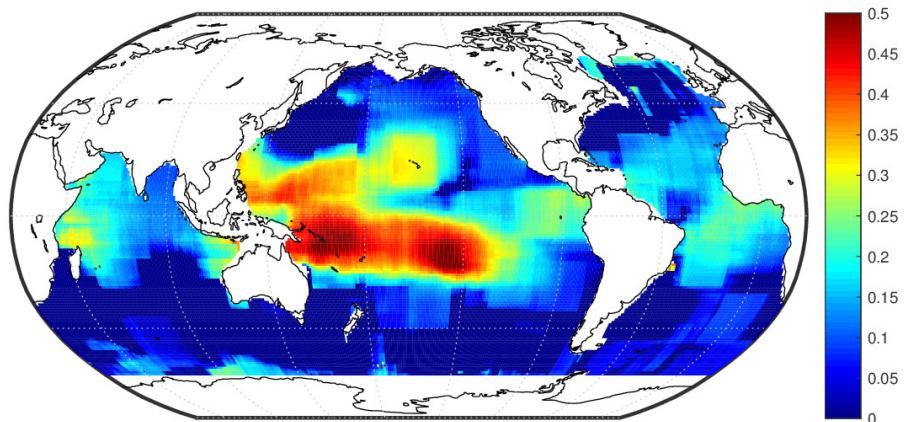


(c) 1500 dbar

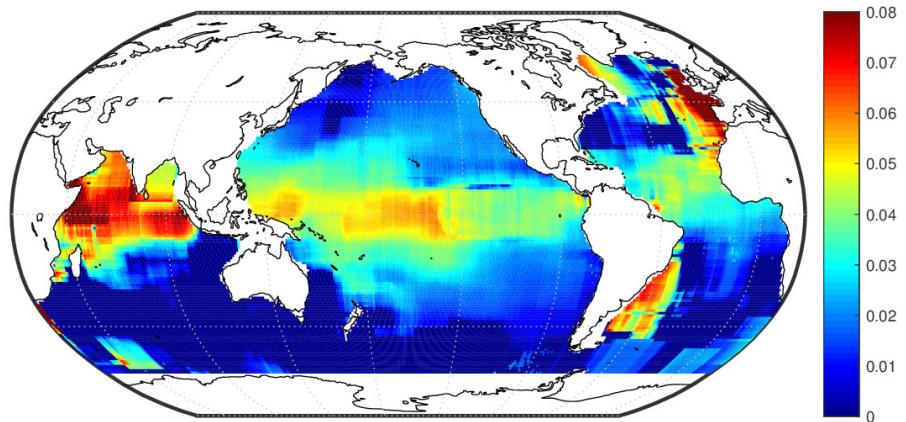
**Figure 36:**  $\theta_t$  (in days)



(a) 10 dbar

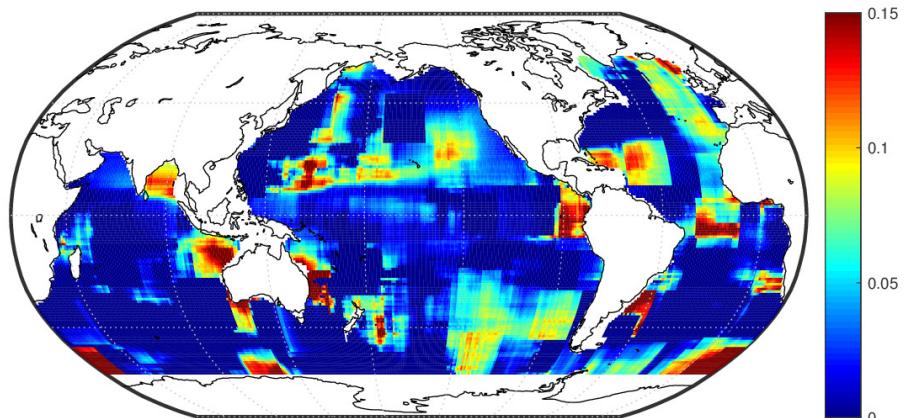


(b) 300 dbar

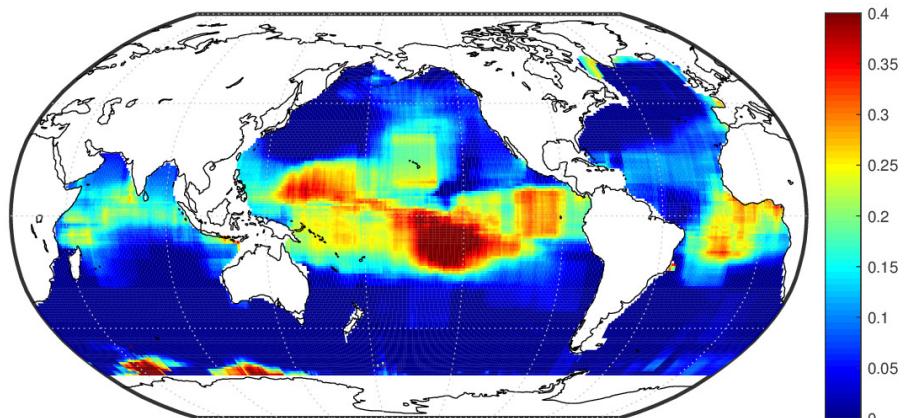


(c) 1500 dbar

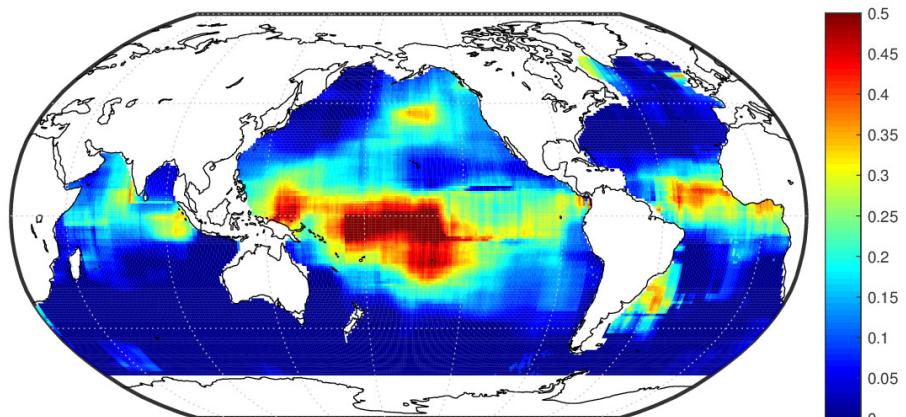
**Figure 37:**  $\sigma$



(a) 10 dbar

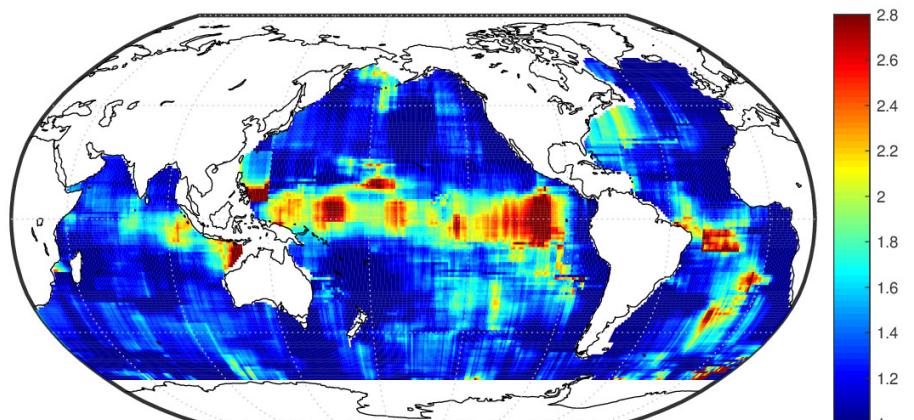


(b) 300 dbar

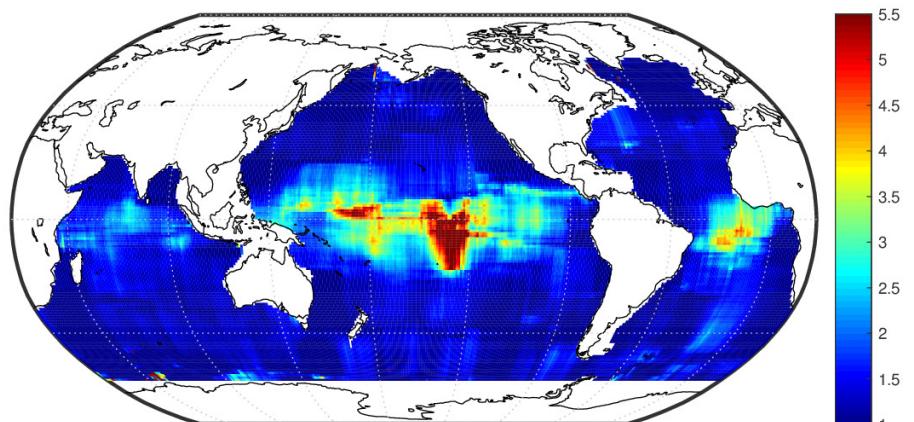


(c) 1500 dbar

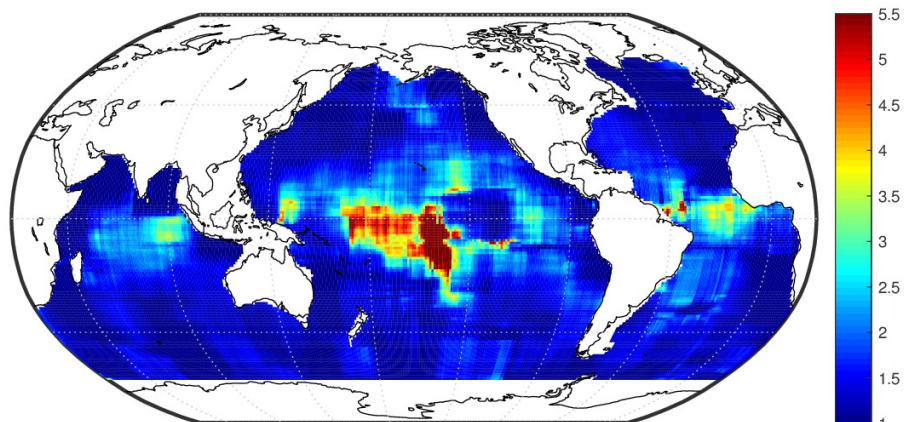
**Figure 38:**  $\sigma^2 / (\phi + \sigma^2)$



(a) 10 dbar

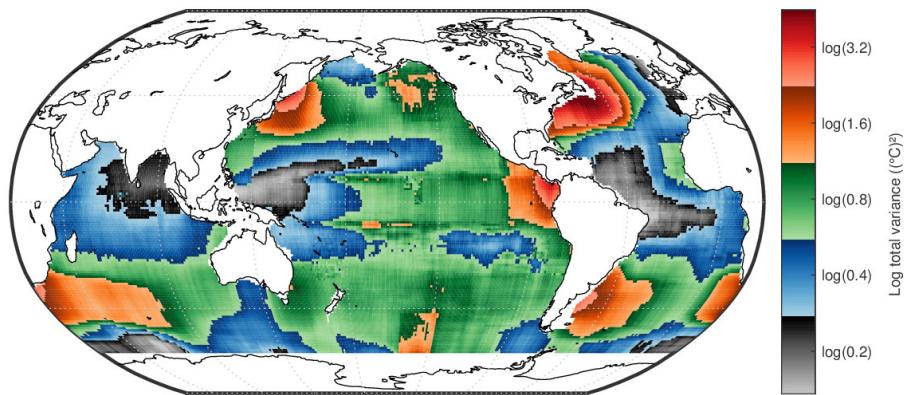


(b) 300 dbar

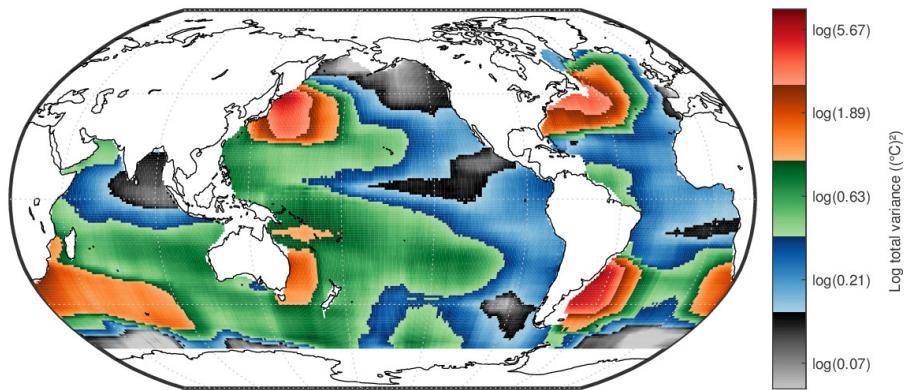


(c) 1500 dbar

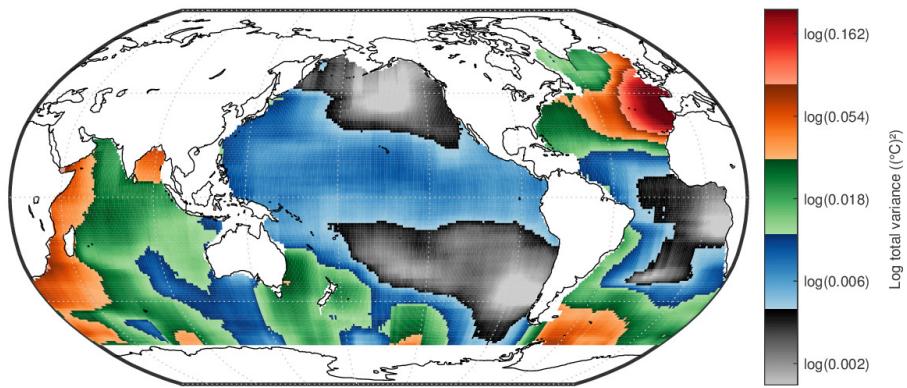
**Figure 39:**  $\theta_{\text{lon}}/\theta_{\text{lat}}$



(a) 10 dbar

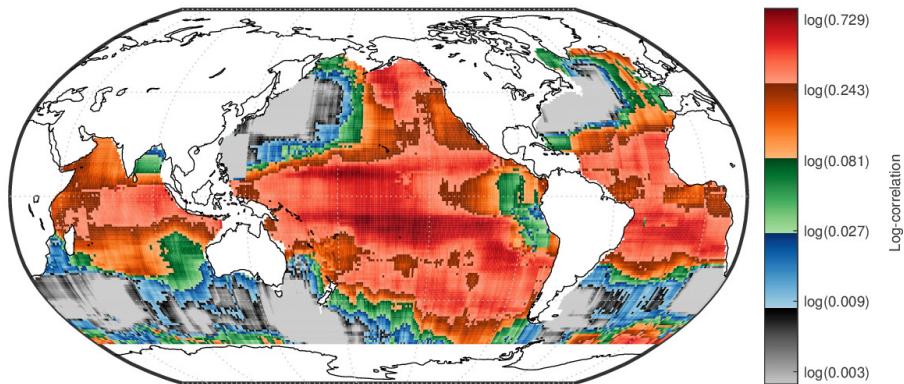


(b) 300 dbar

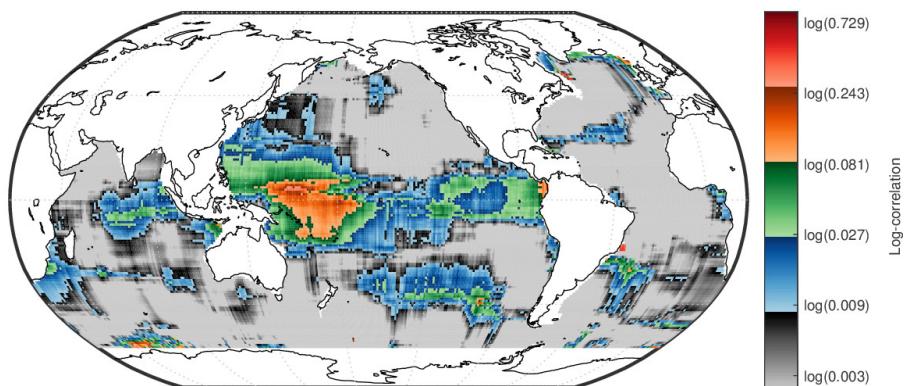


(c) 1500 dbar

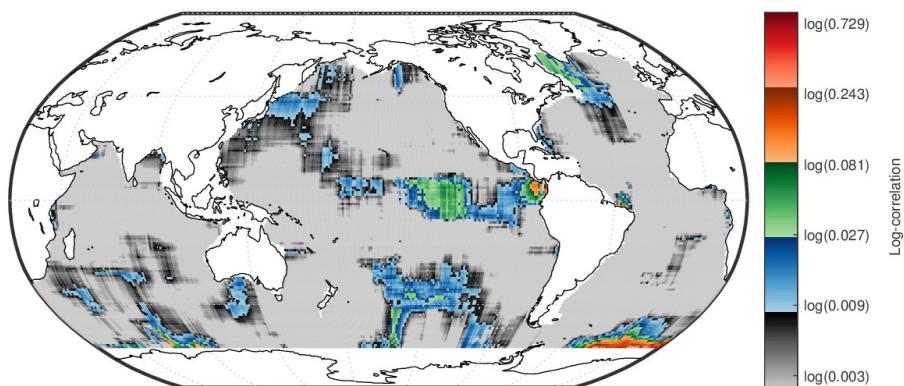
**Figure 40:**  $\phi + \sigma^2$



(a) 10 dbar

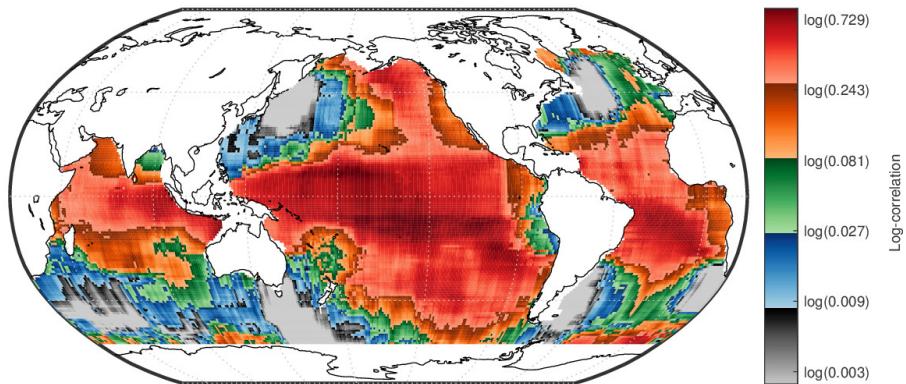


(b) 300 dbar

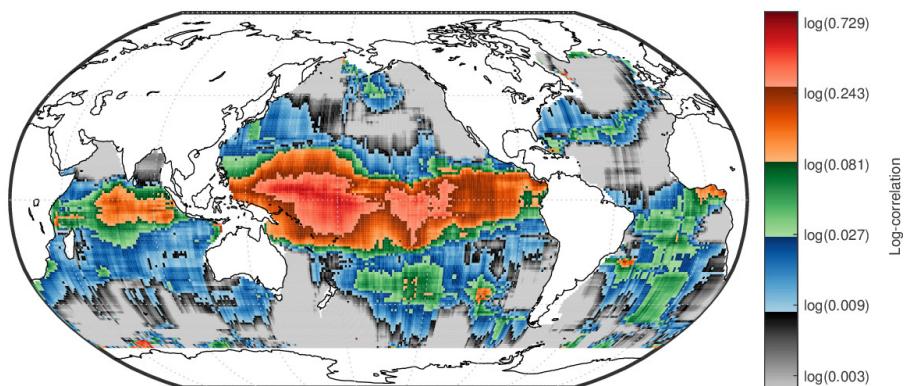


(c) 1500 dbar

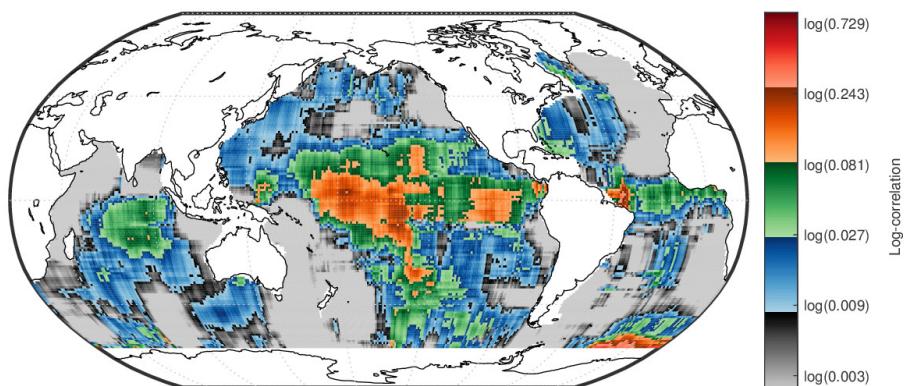
**Figure 41:** Correlation at  $\Delta x_{\text{lat}} = 800 \text{ km}$



(a) 10 dbar

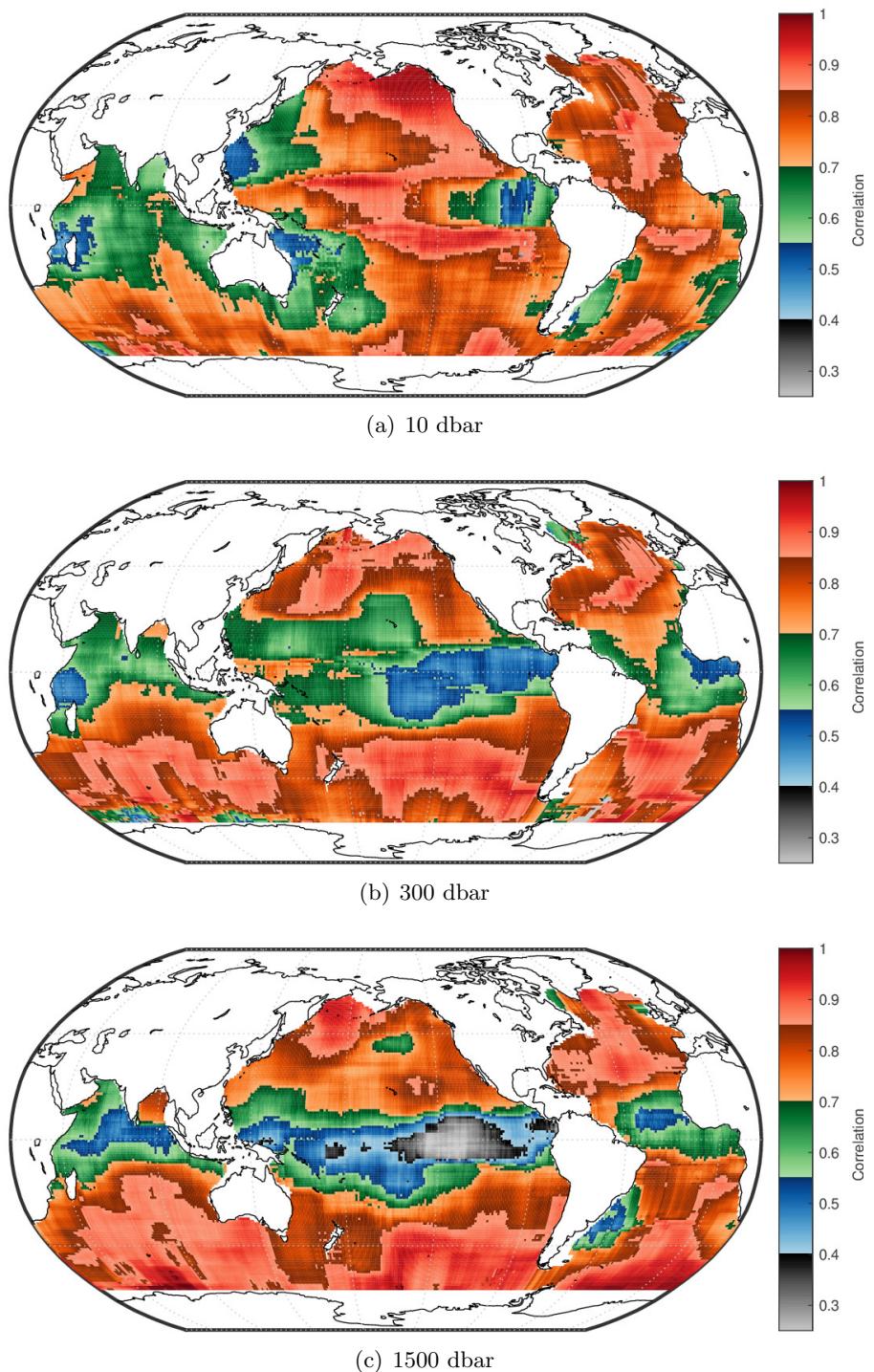


(b) 300 dbar



(c) 1500 dbar

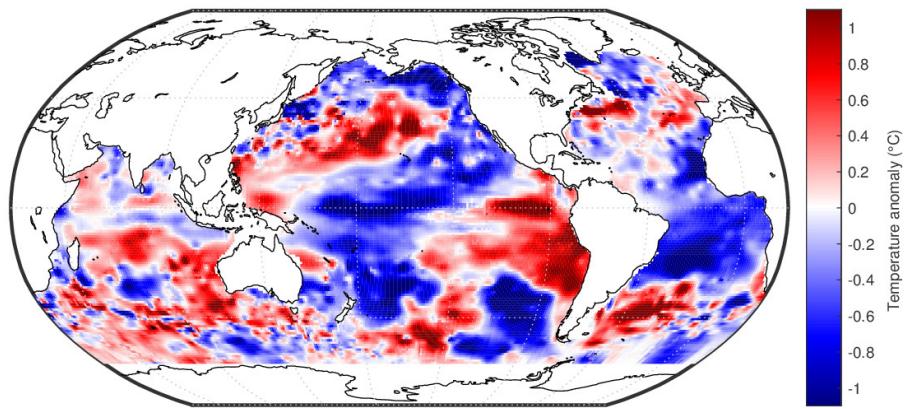
**Figure 42:** Correlation at  $\Delta x_{\text{lon}} = 800 \text{ km}$



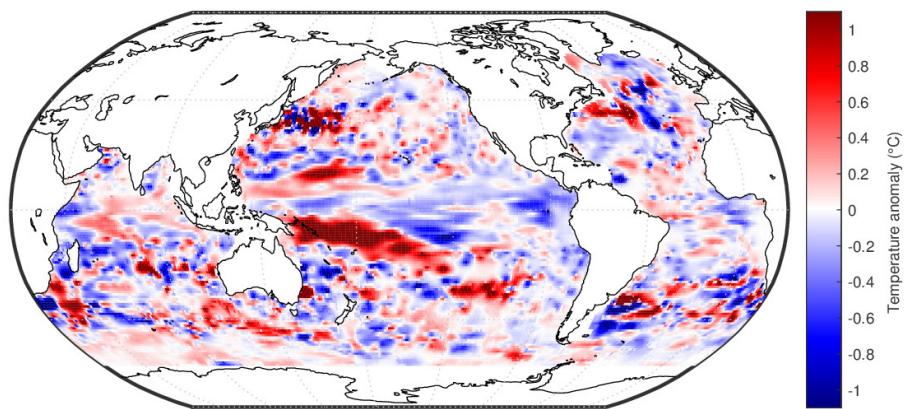
**Figure 43:** Correlation at  $\Delta t = 10$  days

## 5.8 Model 6: Anomalies and model parameters

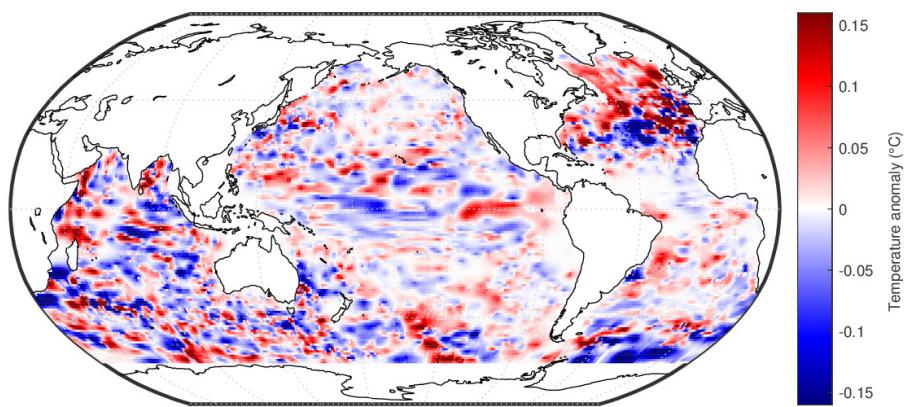
The following pages illustrate the 3-month spatio-temporal model with a Student nugget (Model 6). Quantities that require the existence of the second moment for the Student nugget are masked out when  $\hat{\nu} \leq 2$ .



(a) 10 dbar

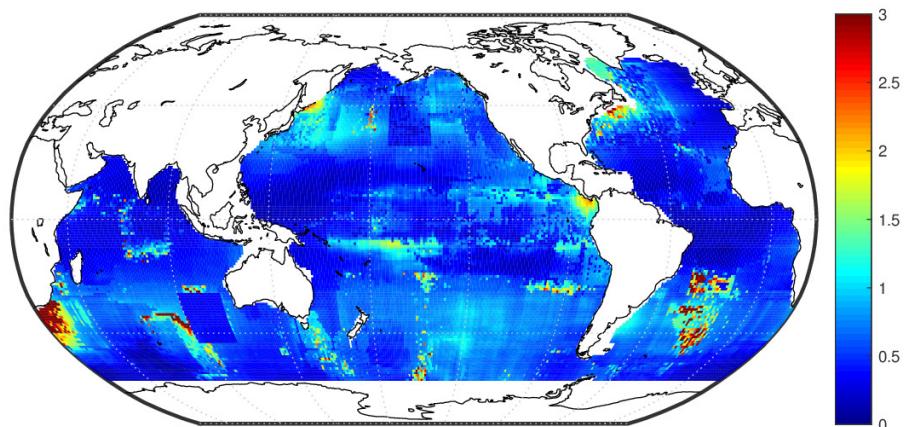


(b) 300 dbar

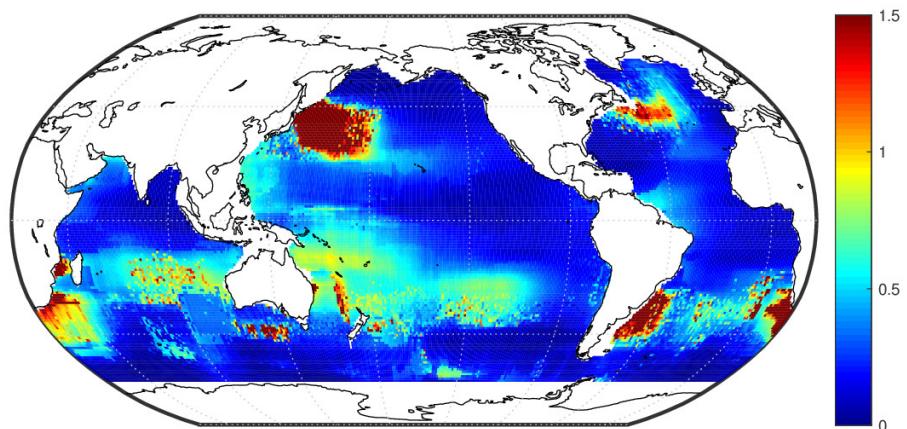


(c) 1500 dbar

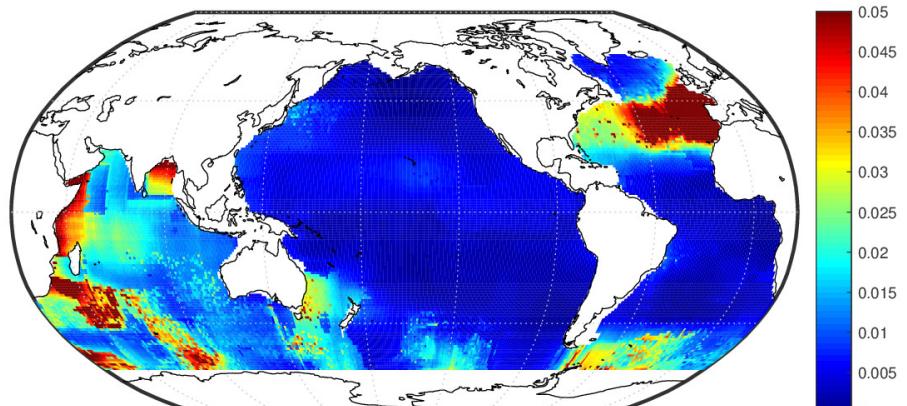
**Figure 44:** February 2012 temperature anomalies



(a) 10 dbar

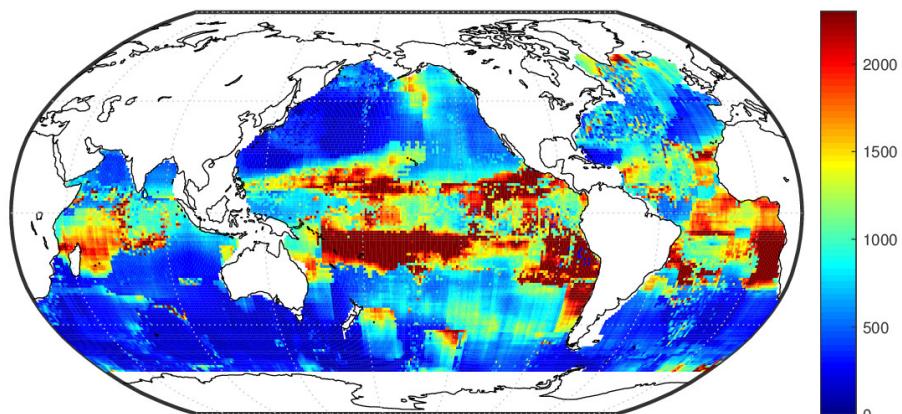


(b) 300 dbar

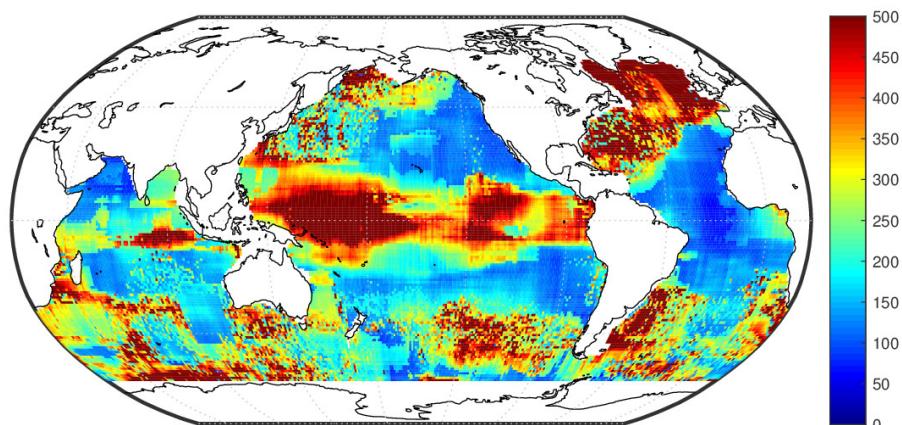


(c) 1500 dbar

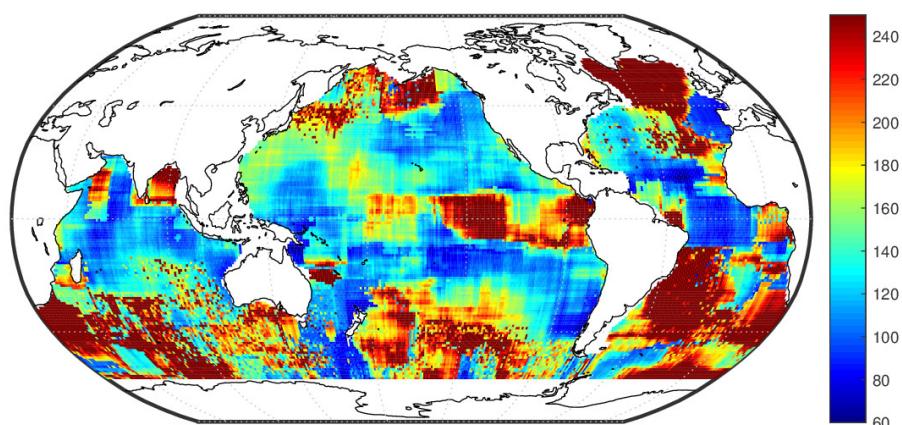
**Figure 45:**  $\phi$



(a) 10 dbar

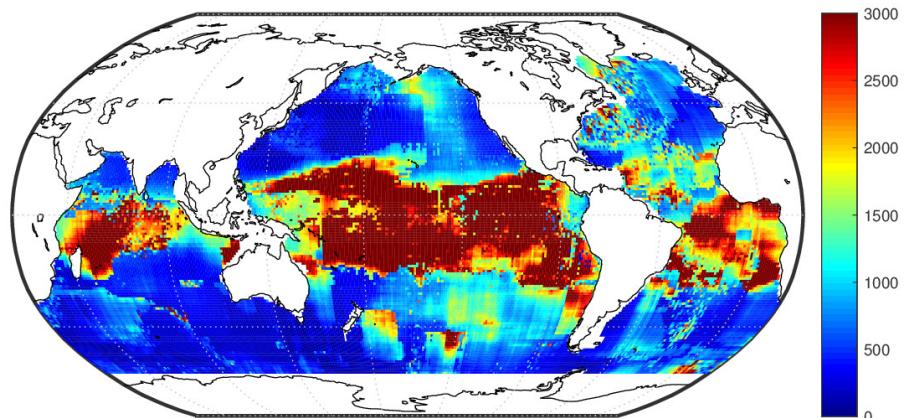


(b) 300 dbar

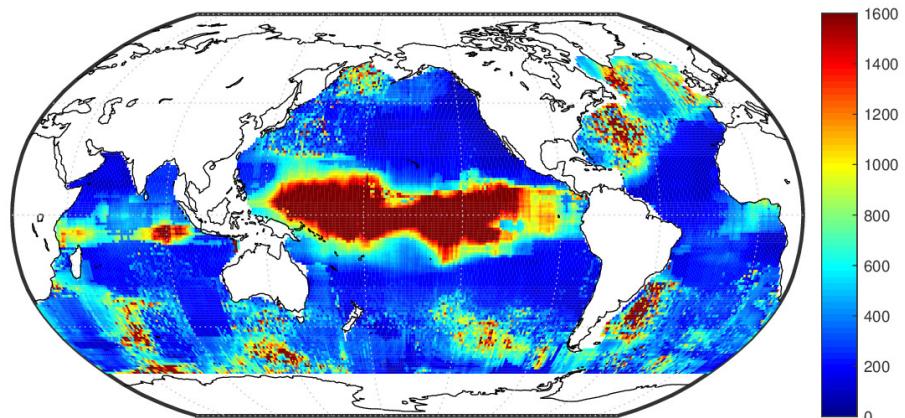


(c) 1500 dbar

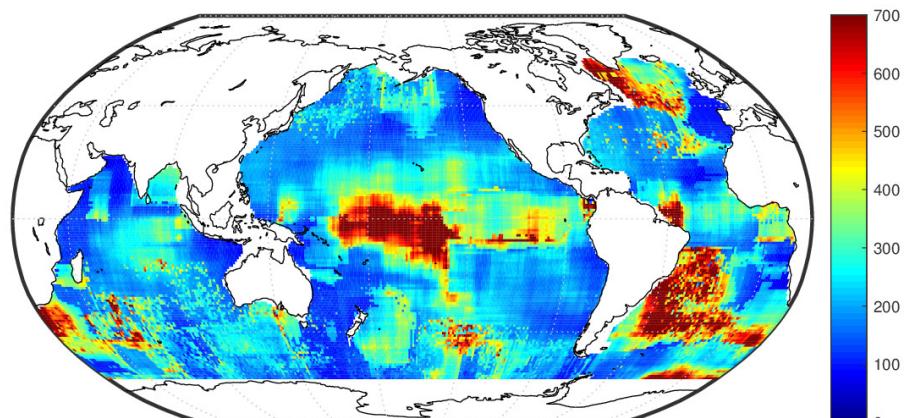
**Figure 46:**  $\theta_{\text{lat}}$  (in km)



(a) 10 dbar

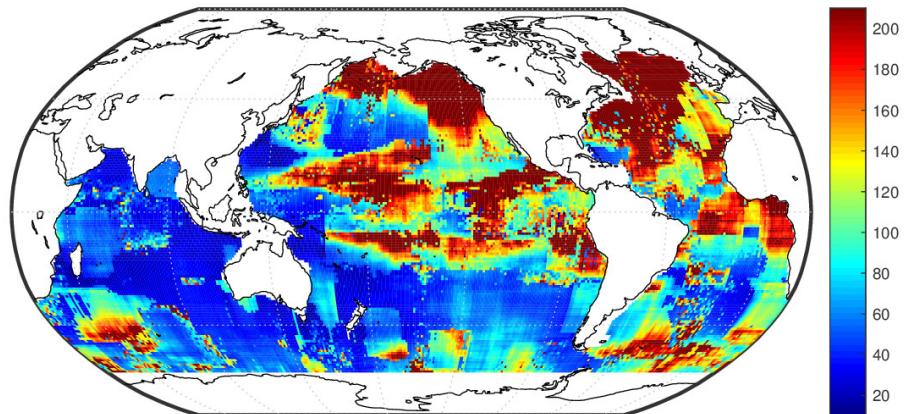


(b) 300 dbar

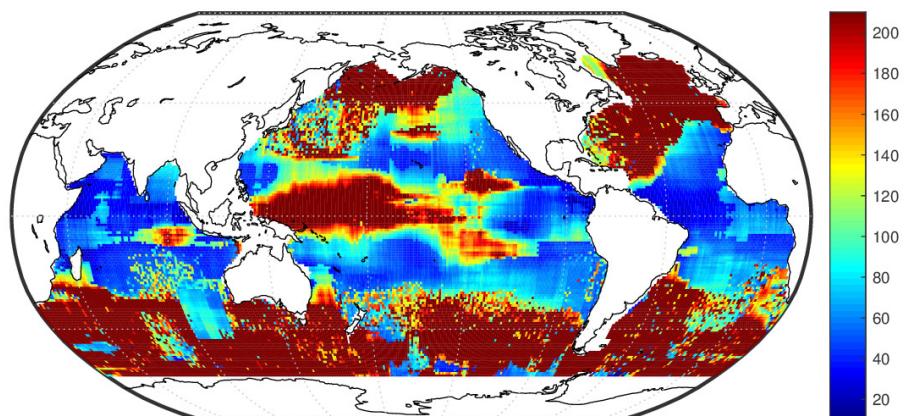


(c) 1500 dbar

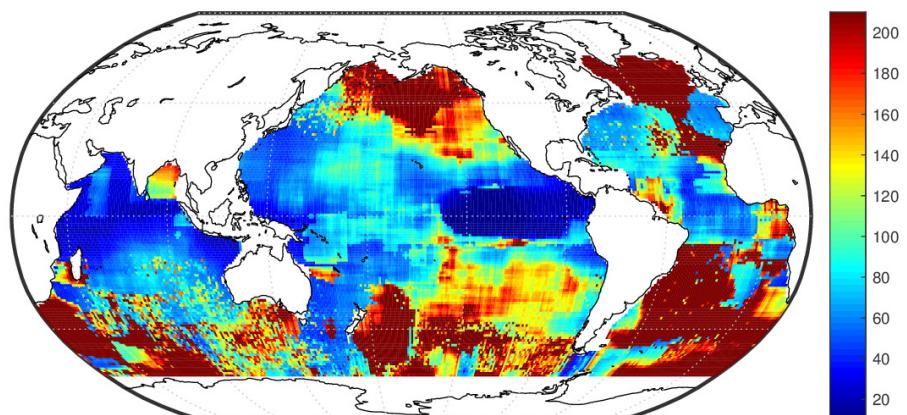
**Figure 47:**  $\theta_{\text{lon}}$  (in km)



(a) 10 dbar

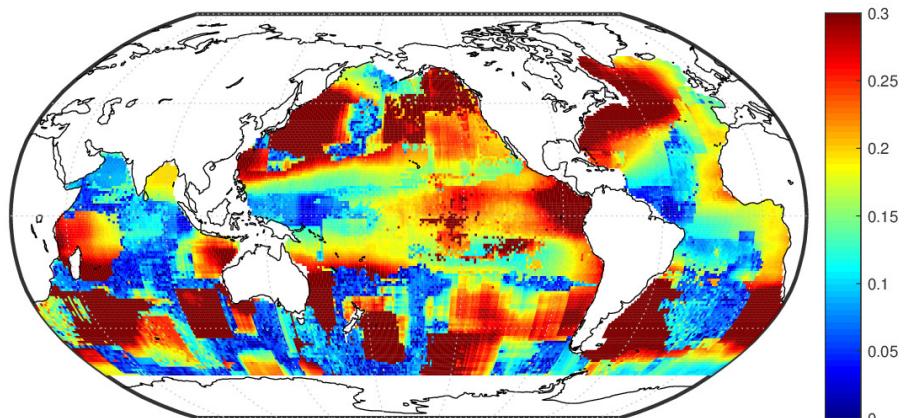


(b) 300 dbar

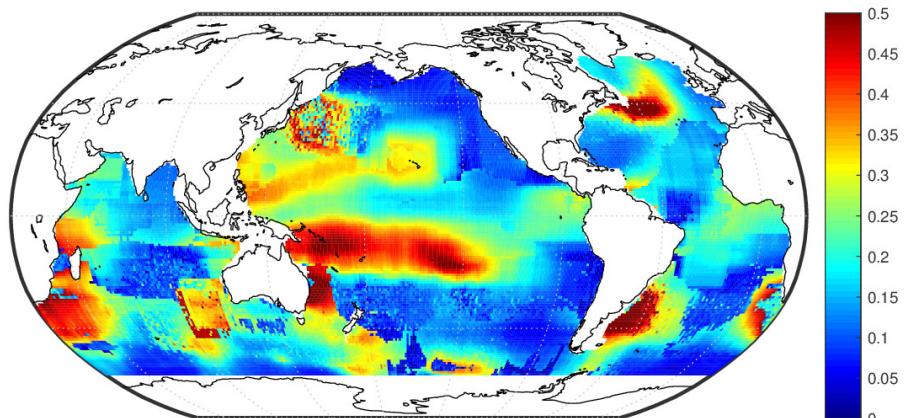


(c) 1500 dbar

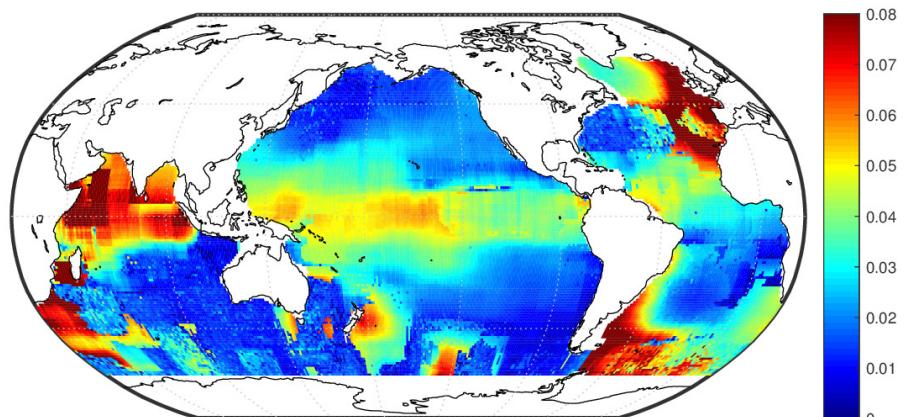
**Figure 48:**  $\theta_t$  (in days)



(a) 10 dbar

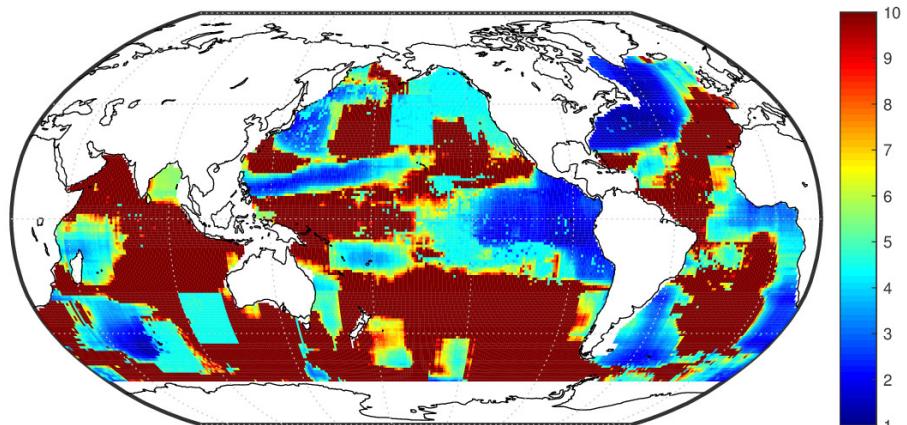


(b) 300 dbar

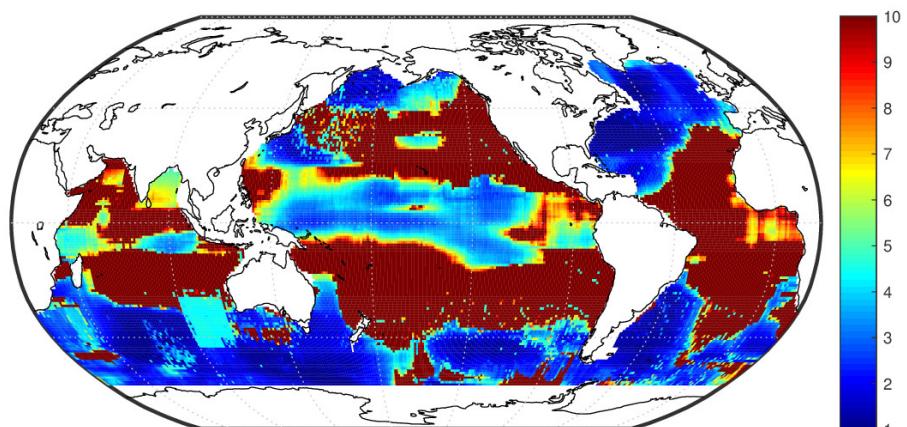


(c) 1500 dbar

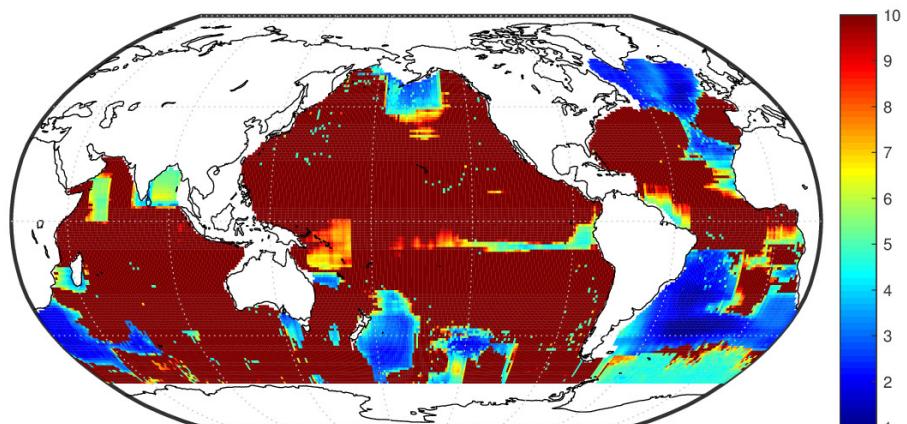
**Figure 49:**  $\sigma$



(a) 10 dbar

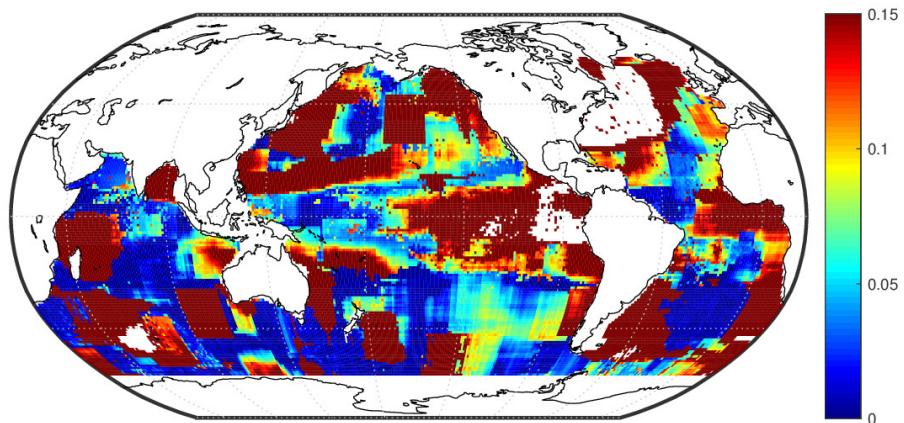


(b) 300 dbar

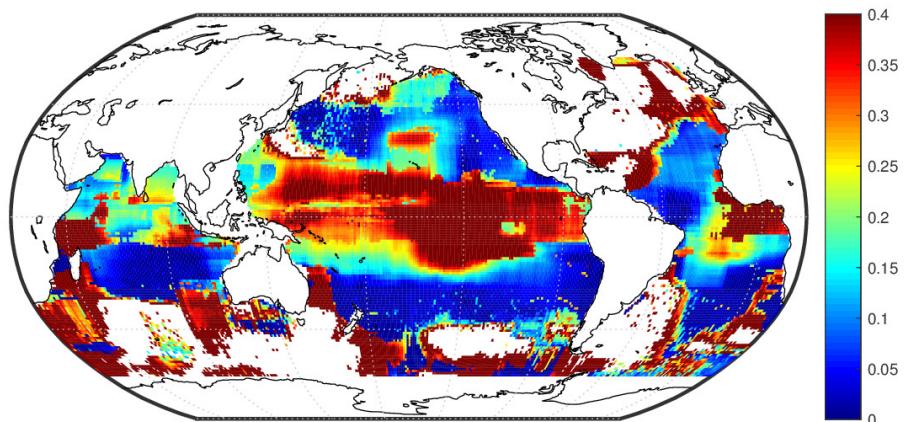


(c) 1500 dbar

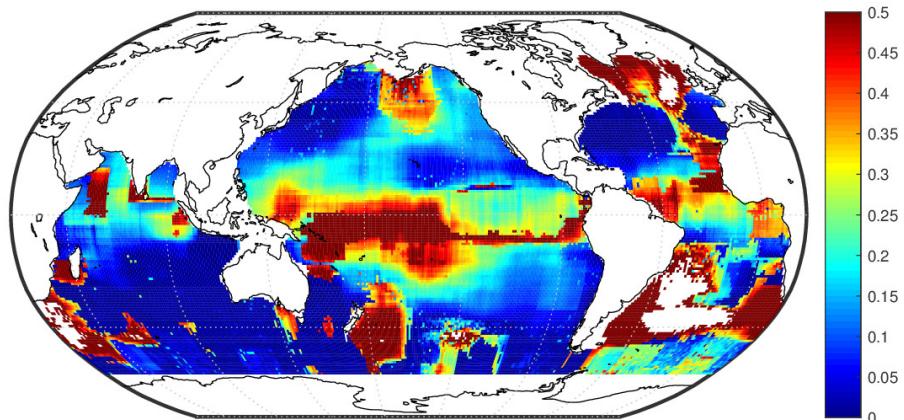
**Figure 50:**  $\nu$



(a) 10 dbar

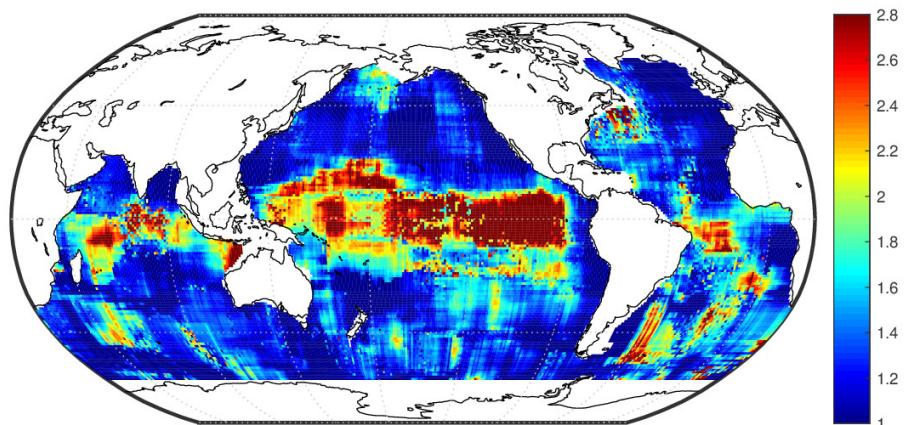


(b) 300 dbar

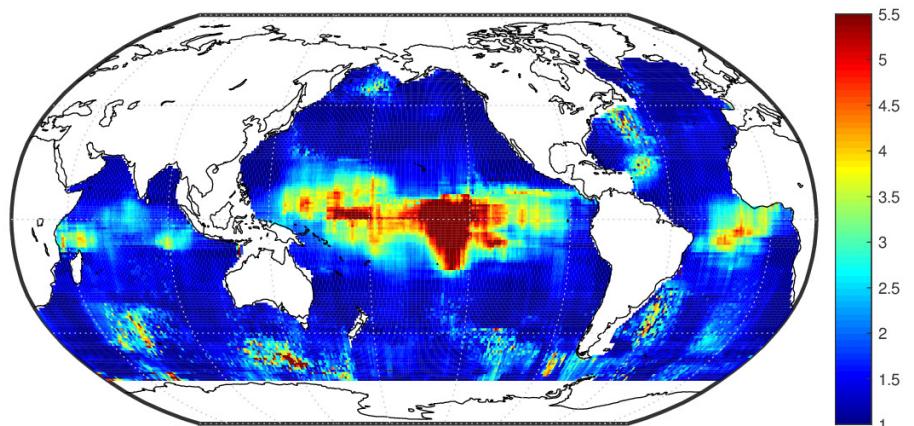


(c) 1500 dbar

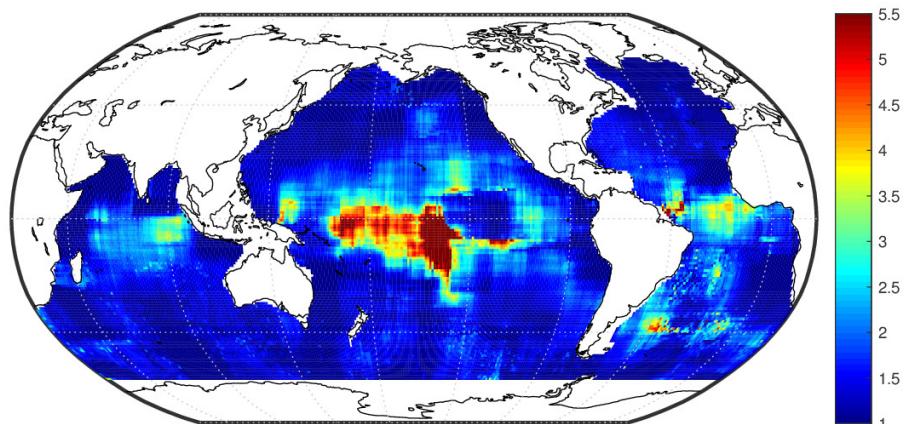
**Figure 51:**  $\frac{\nu}{\nu-2}\sigma^2 / \left( \phi + \frac{\nu}{\nu-2}\sigma^2 \right)$



(a) 10 dbar

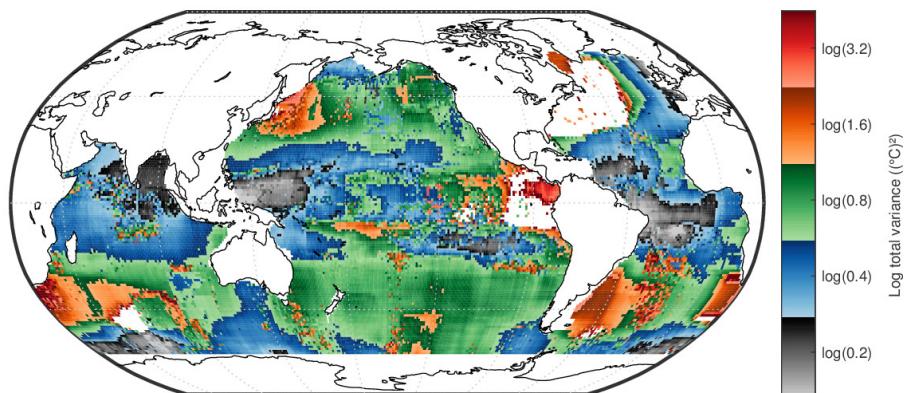


(b) 300 dbar

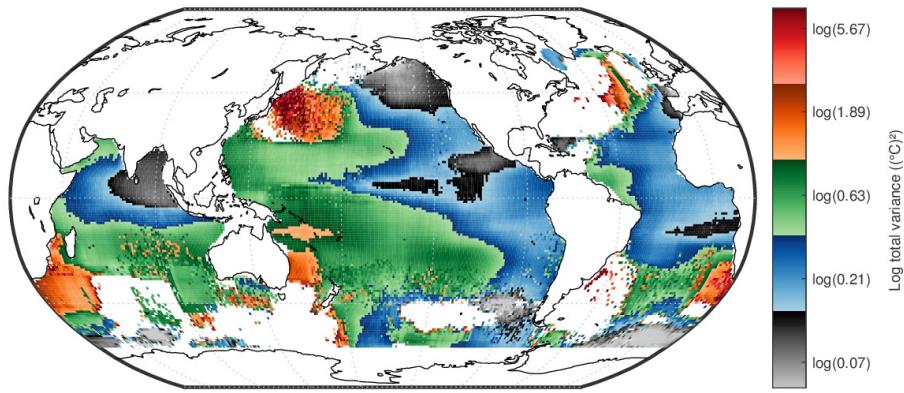


(c) 1500 dbar

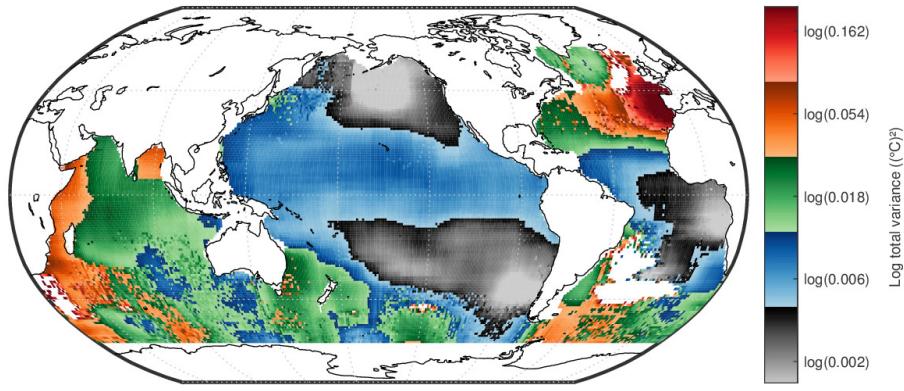
**Figure 52:**  $\theta_{\text{lon}}/\theta_{\text{lat}}$



(a) 10 dbar

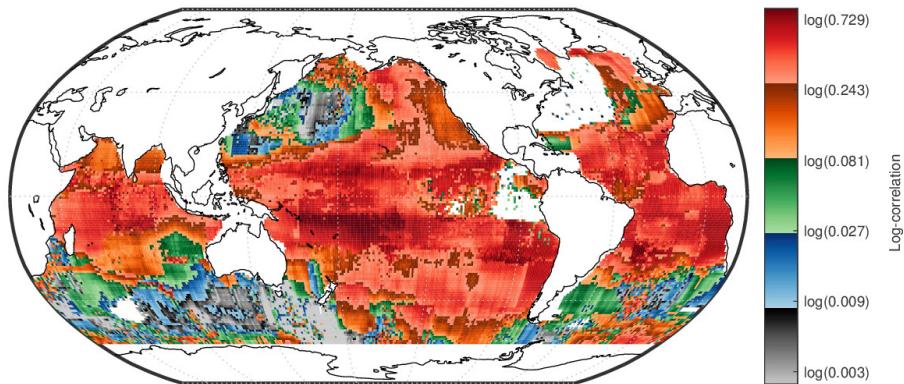


(b) 300 dbar

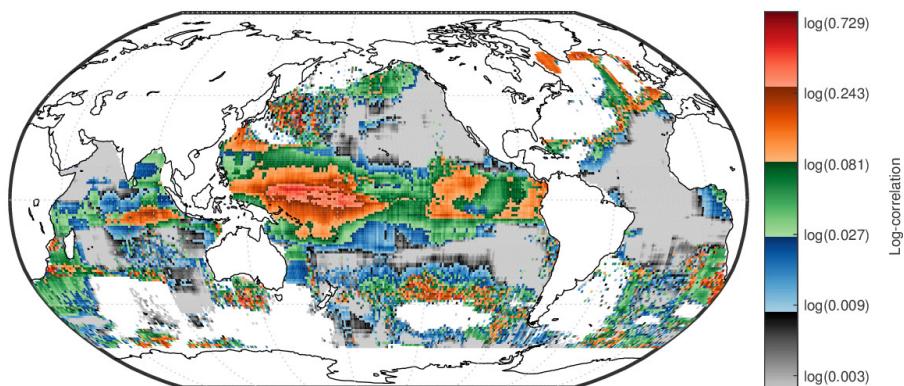


(c) 1500 dbar

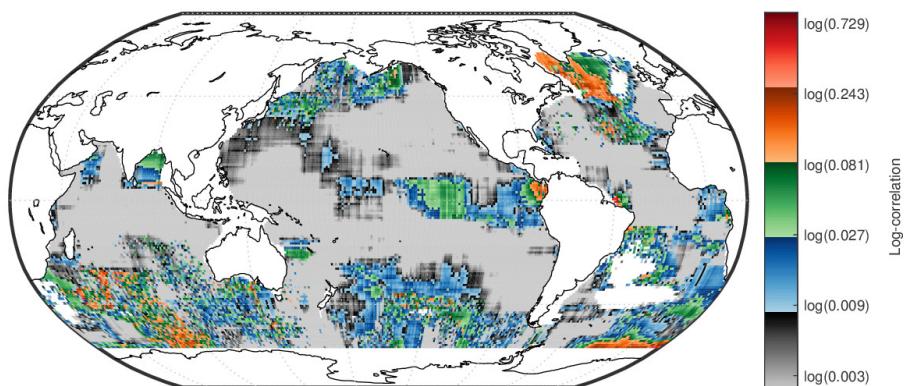
**Figure 53:**  $\phi + \frac{\nu}{\nu-2} \sigma^2$



(a) 10 dbar

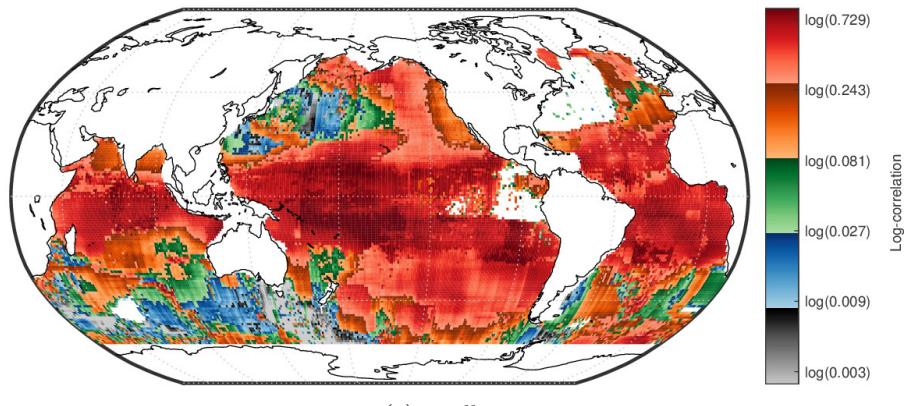


(b) 300 dbar

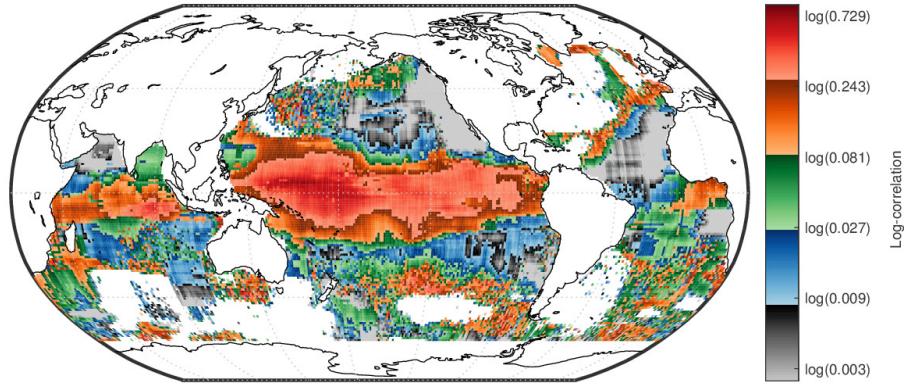


(c) 1500 dbar

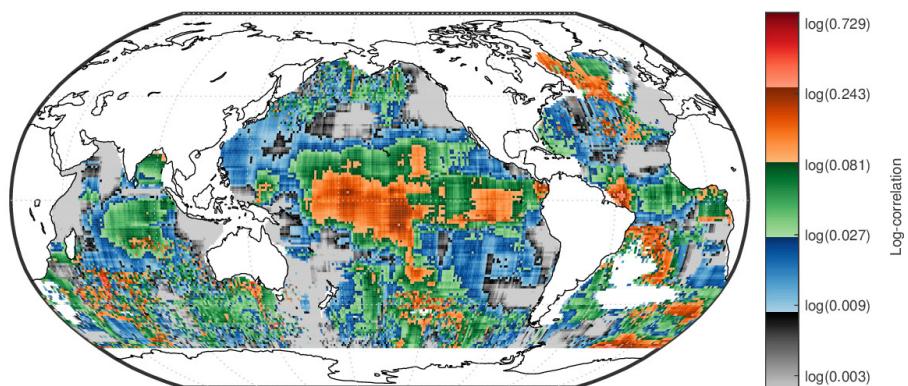
**Figure 54:** Correlation at  $\Delta x_{\text{lat}} = 800 \text{ km}$



(a) 10 dbar

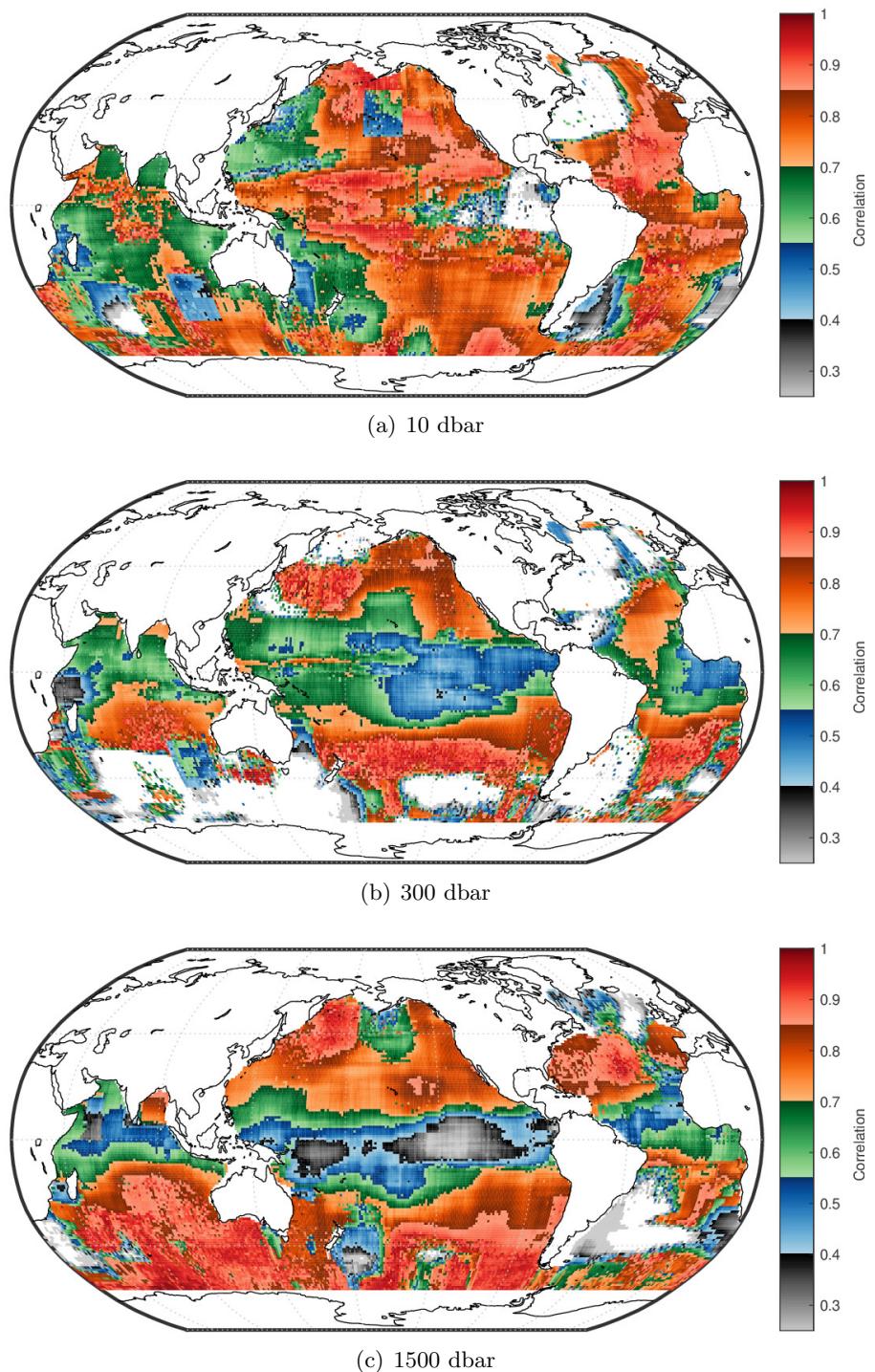


(b) 300 dbar



(c) 1500 dbar

**Figure 55:** Correlation at  $\Delta x_{\text{lon}} = 800 \text{ km}$



**Figure 56:** Correlation at  $\Delta t = 10$  days

## References

- [1] Dean Roemmich et al. On the design and implementation of Argo: A global array of profiling floats. International CLIVAR Project Office Report 21, 1998. Available online at <http://www.argo.ucsd.edu/argo-design.pdf>.
- [2] Stephen C. Riser et al. Fifteen years of ocean observations with the global Argo array. *Nature Climate Change*, 6:145–153, 2016.
- [3] Dean Roemmich and John Gilson. The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program. *Progress in Oceanography*, 82:81–100, 2009.
- [4] Annie Wong, Robert Keeley, Thierry Carval, and the Argo Data Management Team. Argo quality control manual for CTD and trajectory data, 2015. <http://dx.doi.org/10.13155/33951>.
- [5] Jarno Vanhatalo, Pasi Jylänki, and Aki Vehtari. Gaussian process regression with Student- $t$  likelihood. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, and A. Culotta, editors, *Advances in Neural Information Processing Systems 22*, pages 1910–1918. 2009.
- [6] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.
- [7] Carl Edward Rasmussen and Hannes Nickisch. Gaussian Processes for Machine Learning (GPML) toolbox. *Journal of Machine Learning Research*, 11(Nov):3011–3015, 2010.
- [8] Carl E. Rasmussen and Hannes Nickisch. GPML Matlab code, version 4.0, October 19, 2016. Available at <http://www.gaussianprocess.org/gpml/code/matlab>.
- [9] Argo Program. Frequently asked questions about Argo data, retrieved October 14, 2017. [http://www.argo.ucsd.edu/Data\\_FAQ.html](http://www.argo.ucsd.edu/Data_FAQ.html).
- [10] Argo Data Management Team. Argo User’s Manual V3.2, 2017. <http://doi.org/10.13155/29825>.