

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

Mikael Kuusela and Michael L. Stein

November 1, 2017

1 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”) [1] 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 db or have gaps larger than 200 db.
- 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 [2].
- 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 [3] 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.

2 Supplementary results

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

2.1 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 db	RMSE	0.6159 (-4.8%)	0.5583 (-10.1%)
	Q ₃ AE	0.4868 (-0.9%)	0.3920 (-5.0%)
	MdAE	0.2509 (-0.8%)	0.1890 (-4.9%)
300 db	RMSE	0.6223 (-9.3%)	0.5873 (-14.6%)
	Q ₃ AE	0.4248 (-2.4%)	0.3813 (-3.5%)
	MdAE	0.1995 (-2.0%)	0.1779 (-2.2%)
1500 db	RMSE	0.1027 (-3.1%)	0.0912 (-3.3%)
	Q ₃ 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 db	RMSE	0.6894 (-1.0%)	0.6637 (-2.3%)
	Q ₃ AE	0.5757 (0.3%)	0.5269 (-0.9%)
	MdAE	0.2963 (0.6%)	0.2589 (-1.4%)
300 db	RMSE	0.7617 (-1.8%)	0.7661 (-3.7%)
	Q ₃ AE	0.5712 (0.4%)	0.5565 (-0.2%)
	MdAE	0.2754 (-0.0%)	0.2670 (-0.2%)
1500 db	RMSE	0.1316 (-0.6%)	0.1325 (-1.4%)
	Q ₃ AE	0.0977 (-0.2%)	0.0966 (-0.8%)
	MdAE	0.0501 (-0.5%)	0.0486 (-0.5%)

2.2 Uncertainty quantification performance

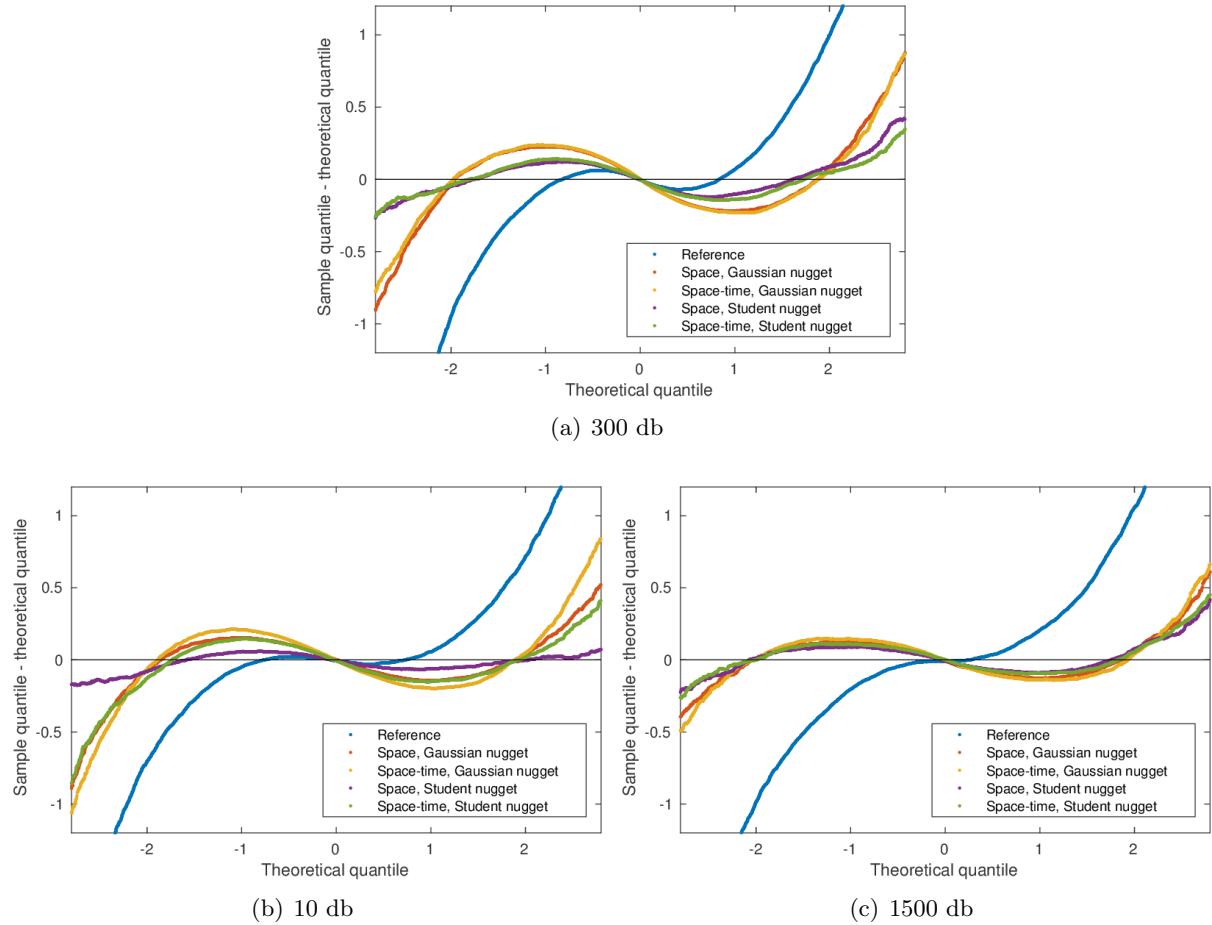
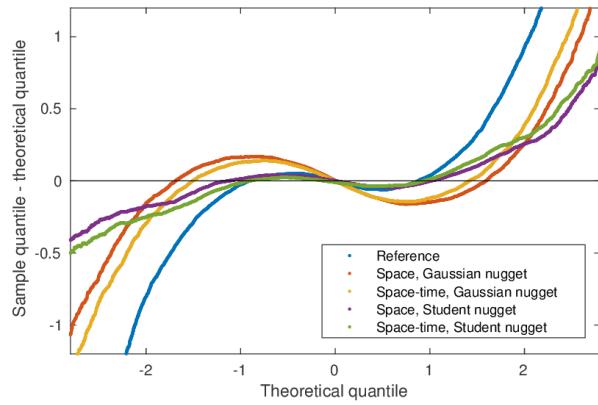
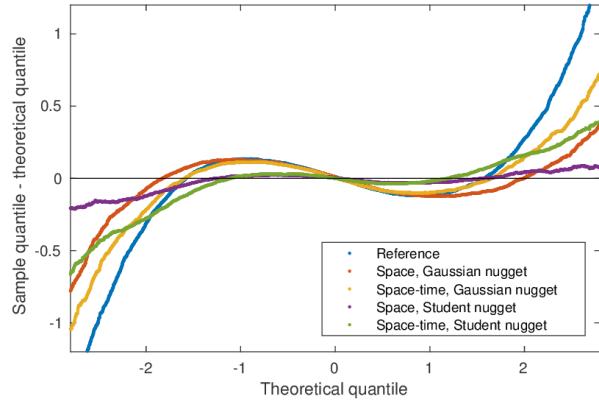


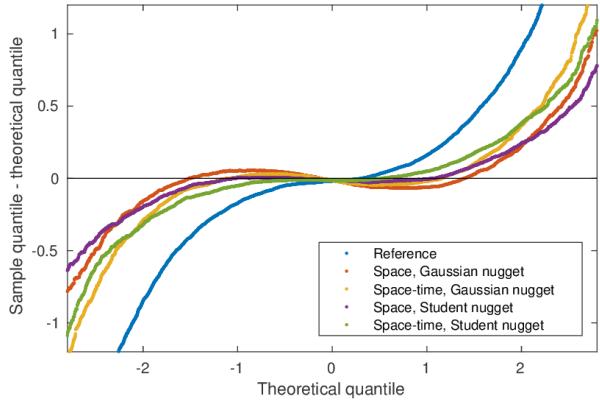
Figure 1: $q_{\text{sample}} - q_{\text{theory}}$ vs. q_{theory} for LOOO cross-validation



(a) 300 db



(b) 10 db



(c) 1500 db

Figure 2: $q_{\text{sample}} - q_{\text{theory}}$ vs. q_{theory} for LOFO cross-validation

Table 3: Coverage and interval lengths, LOOO, 10 db

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 db

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 db

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 db

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 db

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 db

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

2.3 Model 1: Anomalies and model parameters

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

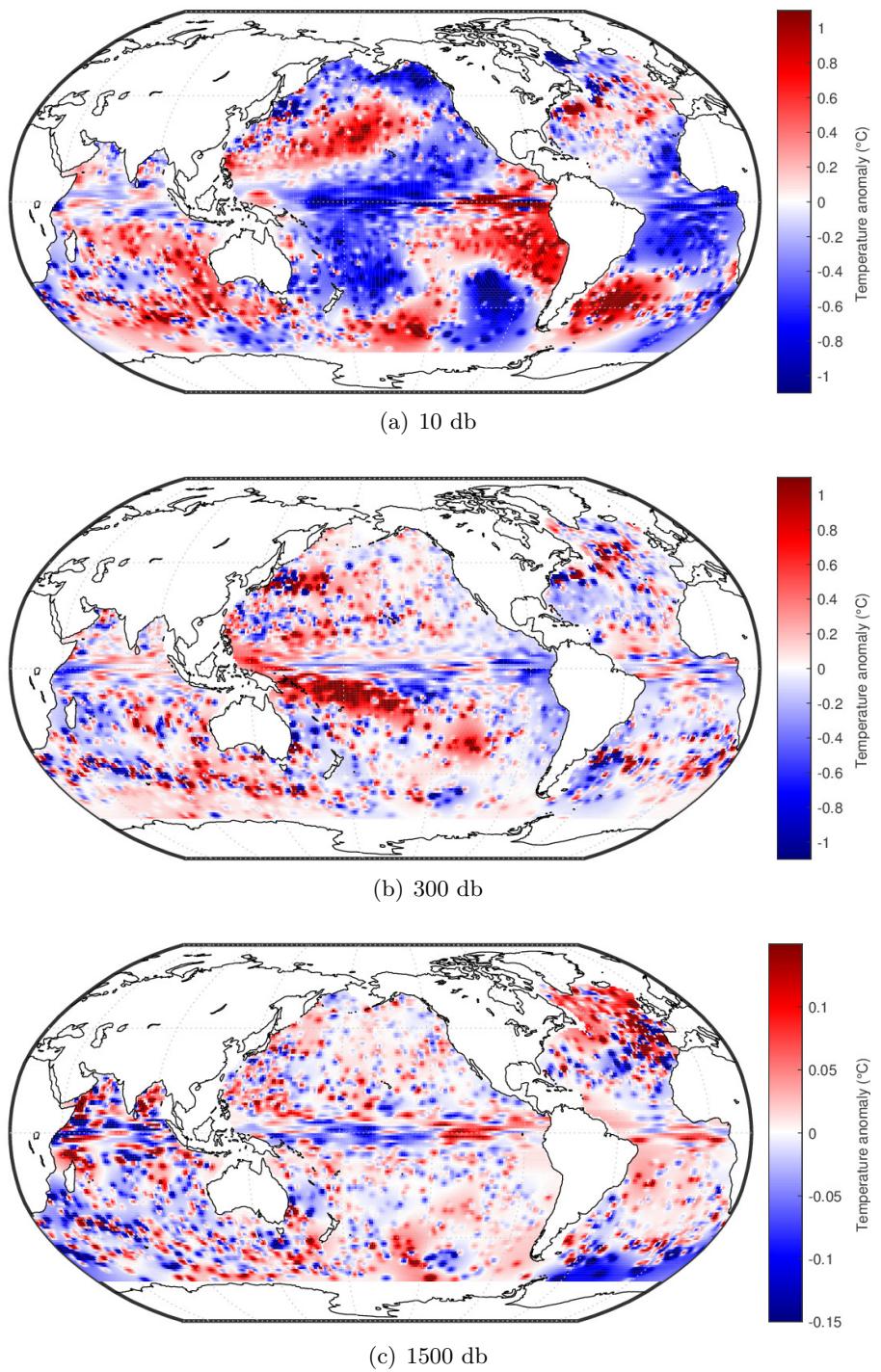


Figure 3: February 2012 temperature anomalies

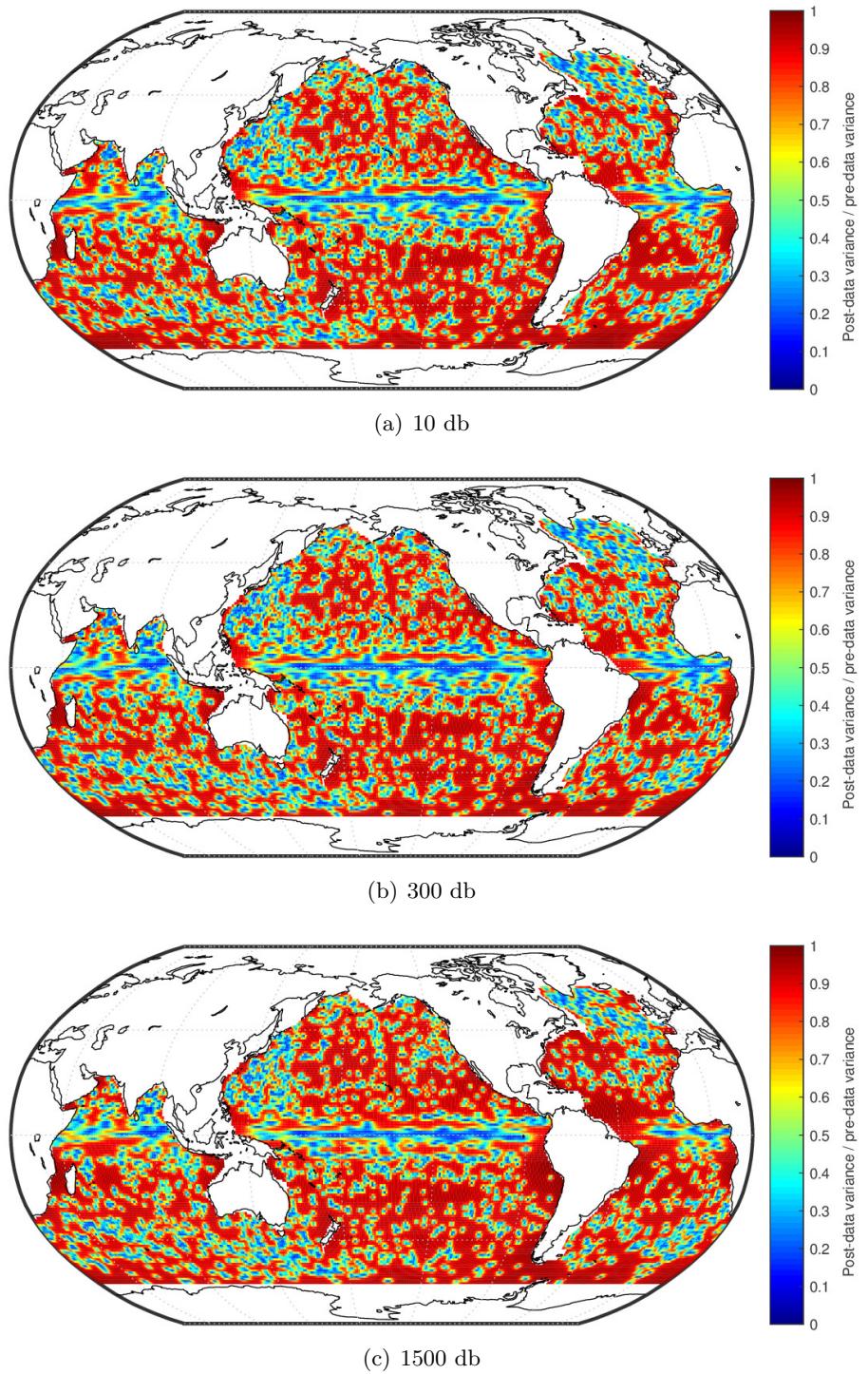


Figure 4: February 2012 post-data-to-pre-data variance ratios. The uncertainty patterns are almost the same at all depths since the Roemmich–Gilson covariance does not vary with depth.

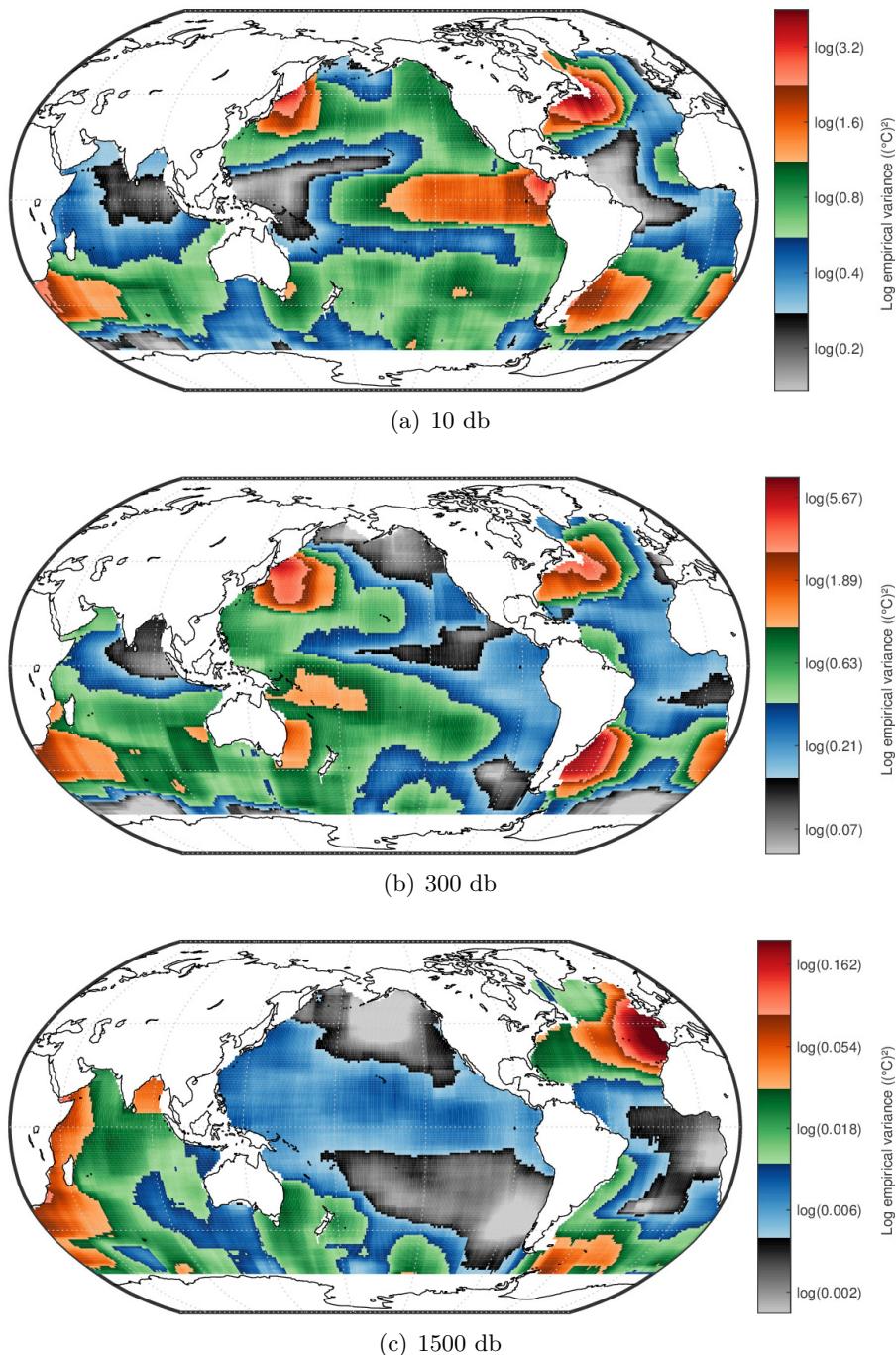


Figure 5: Moving-window empirical variance for estimation of ϕ through $\hat{\phi} = \text{empirical variance}/1.15$.

2.4 Model 2: Anomalies and model parameters

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

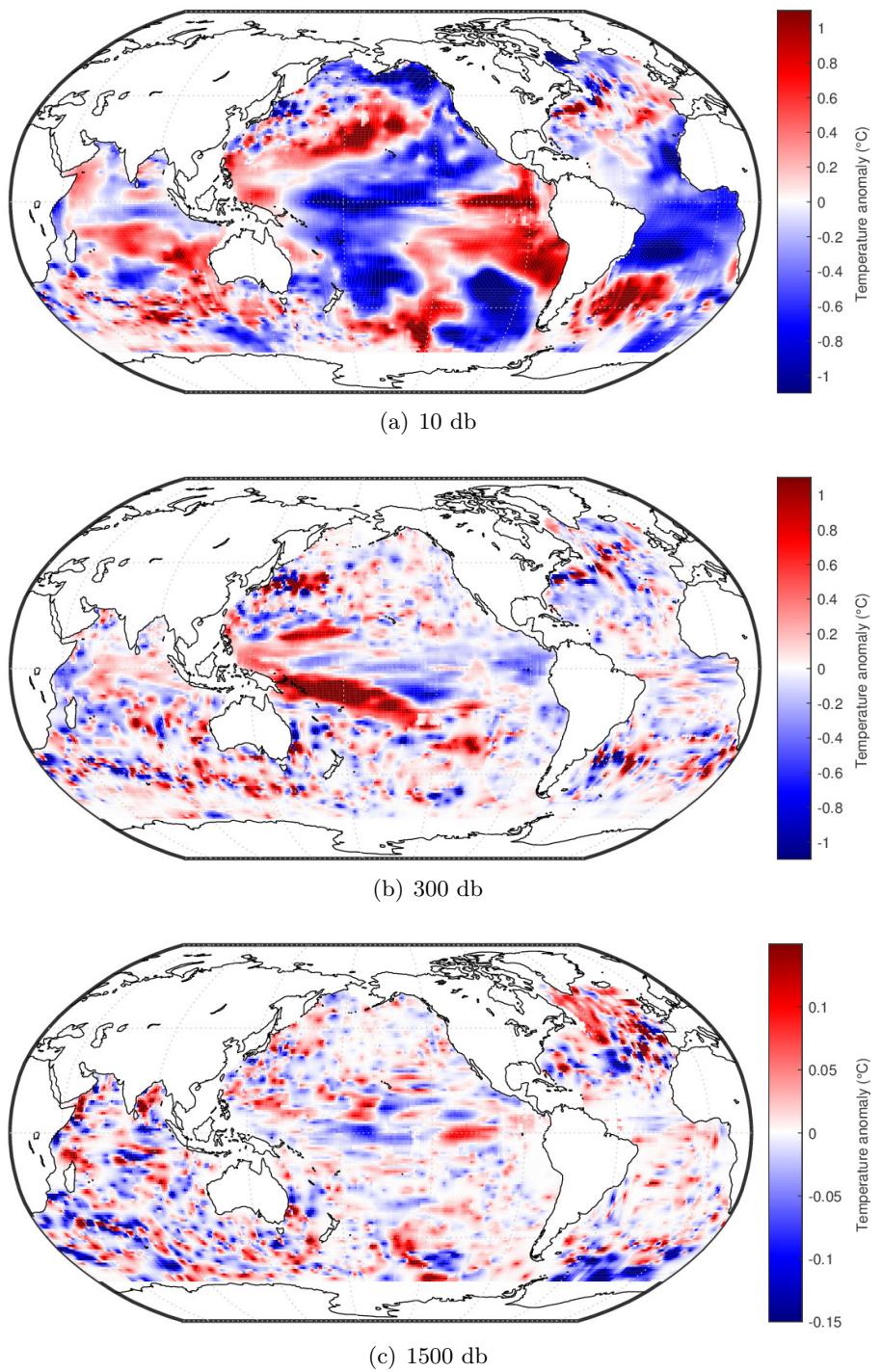


Figure 6: February 2012 temperature anomalies

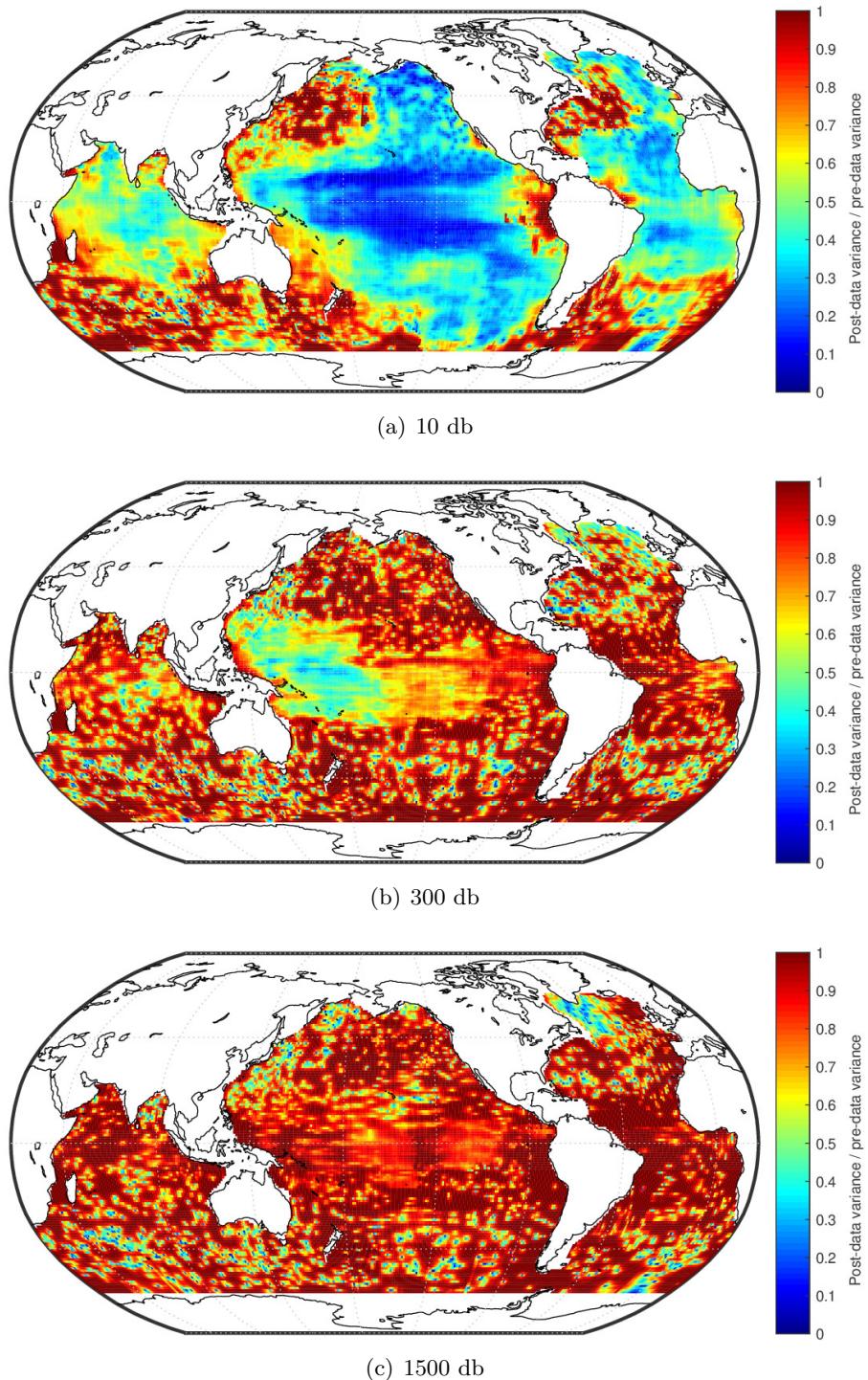
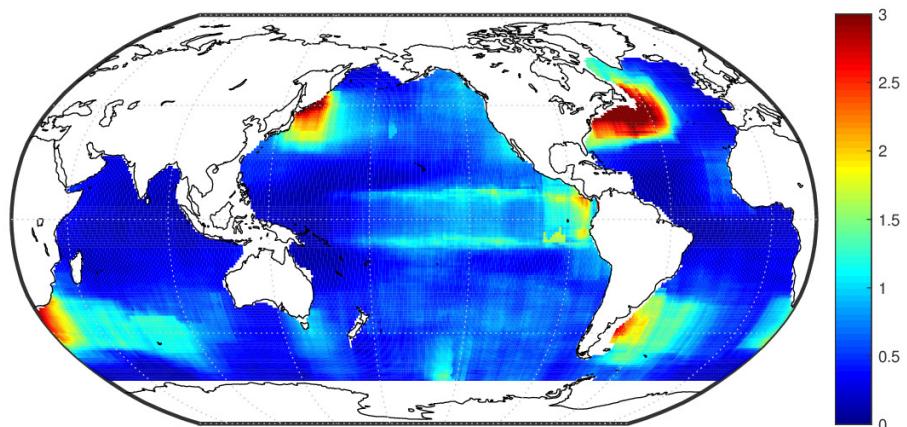
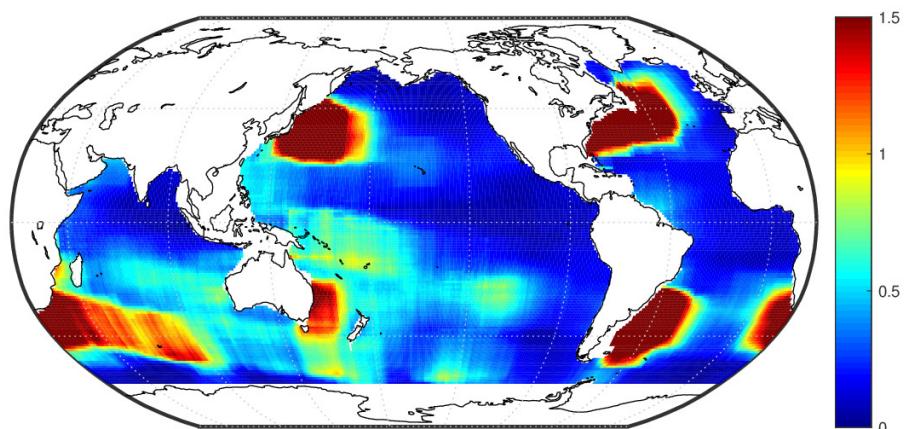


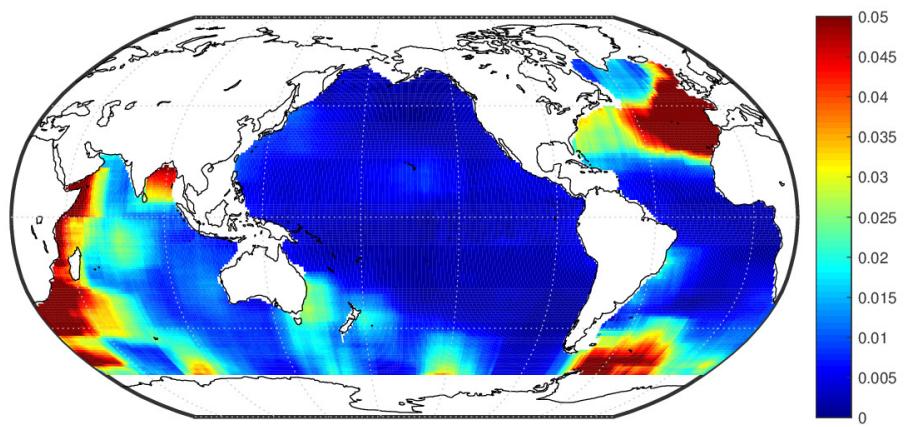
Figure 7: February 2012 post-data-to-pre-data variance ratios



(a) 10 db

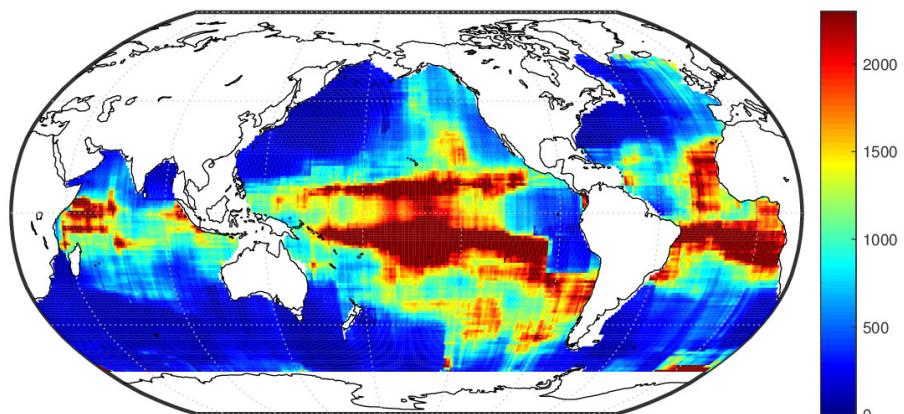


(b) 300 db

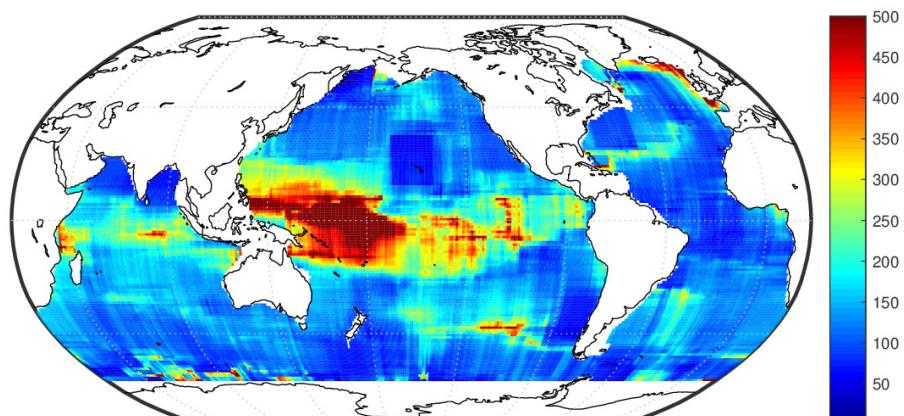


(c) 1500 db

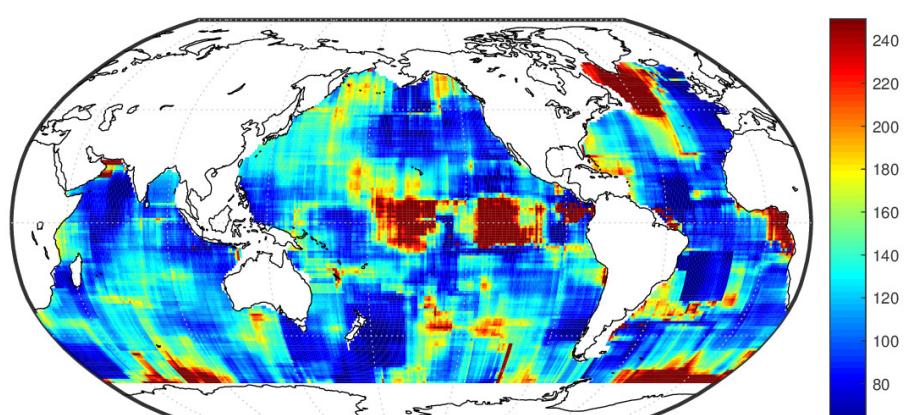
Figure 8: ϕ



(a) 10 db

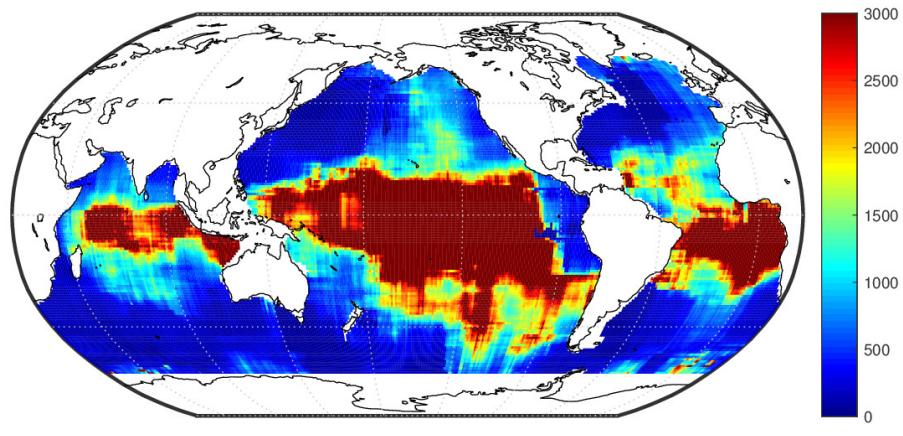


(b) 300 db

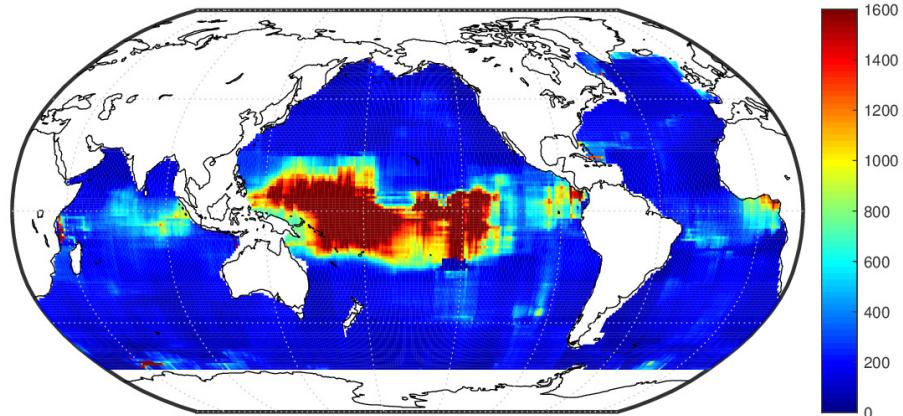


(c) 1500 db

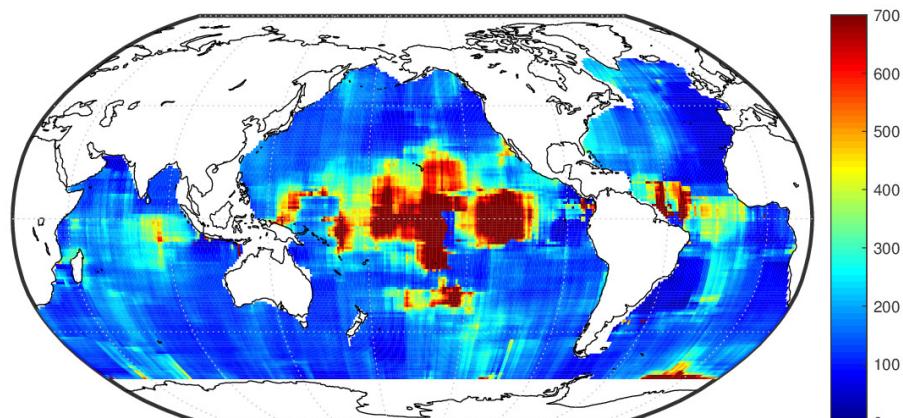
Figure 9: θ_{lat} (in km)



(a) 10 db

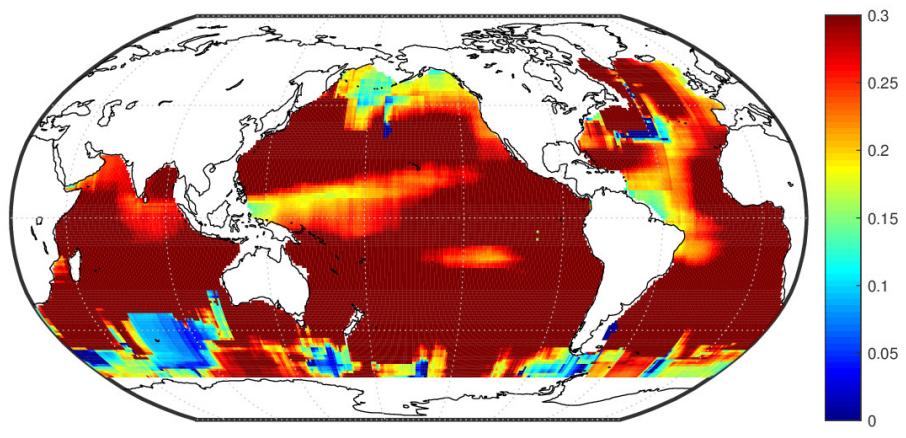


(b) 300 db

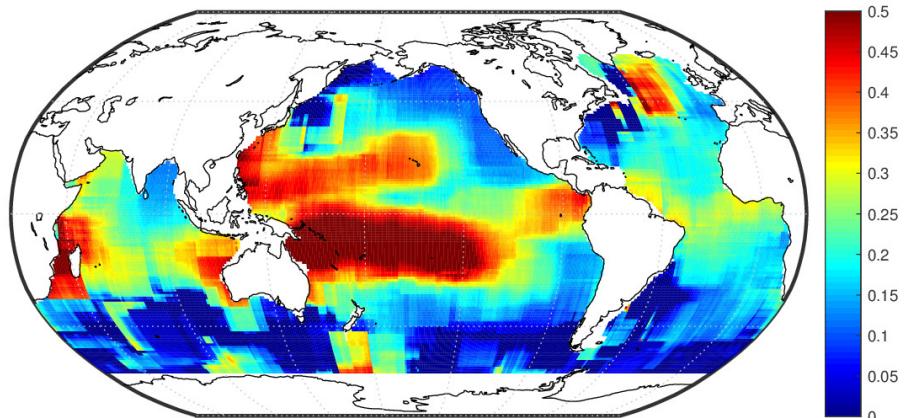


(c) 1500 db

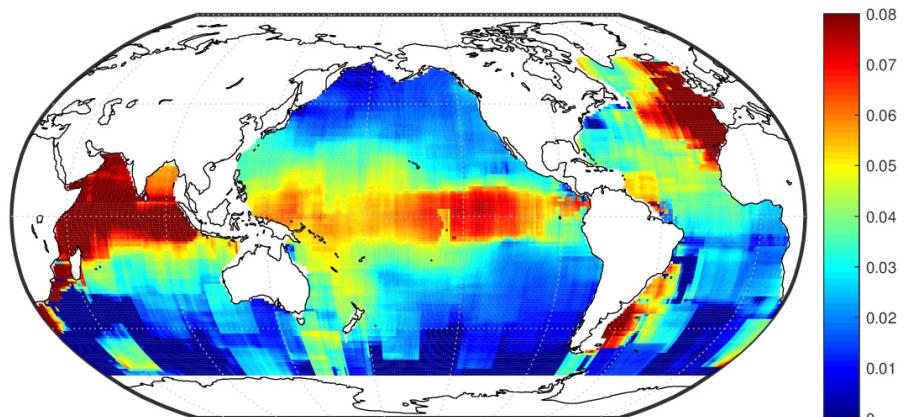
Figure 10: θ_{lon} (in km)



(a) 10 db

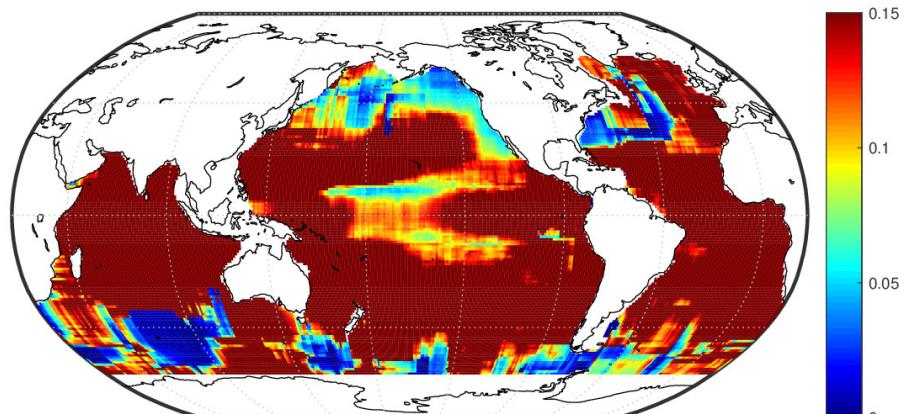


(b) 300 db

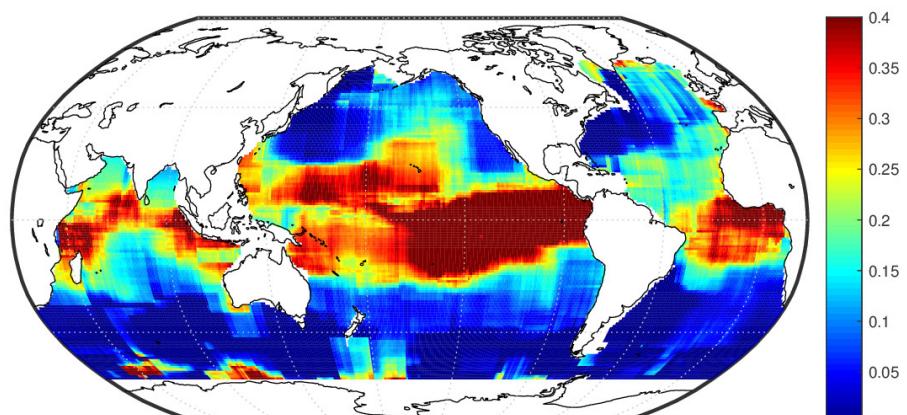


(c) 1500 db

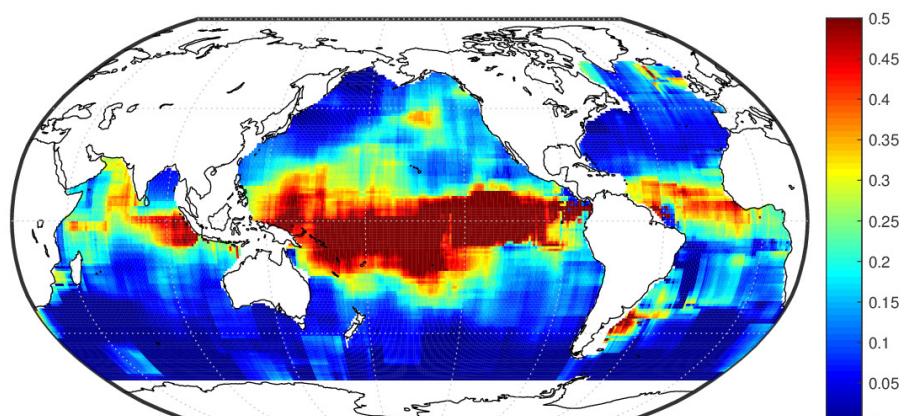
Figure 11: σ



(a) 10 db

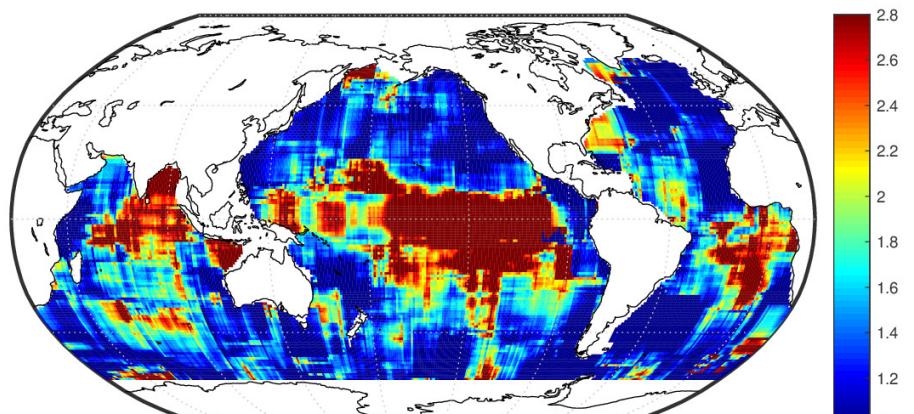


(b) 300 db

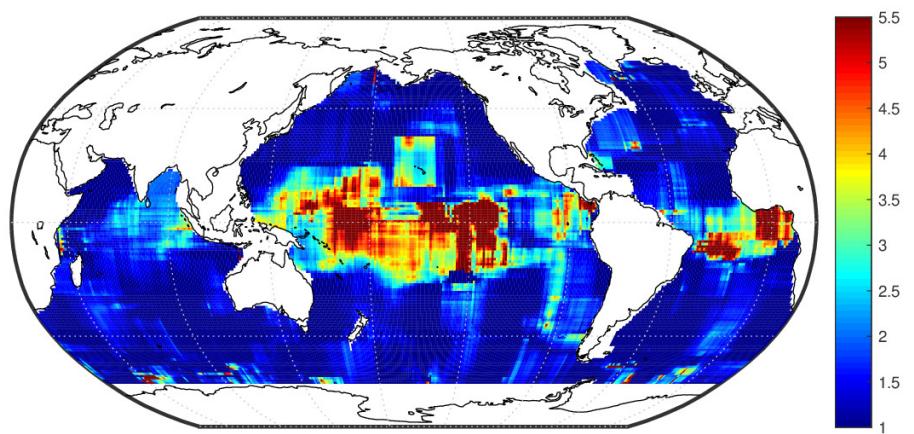


(c) 1500 db

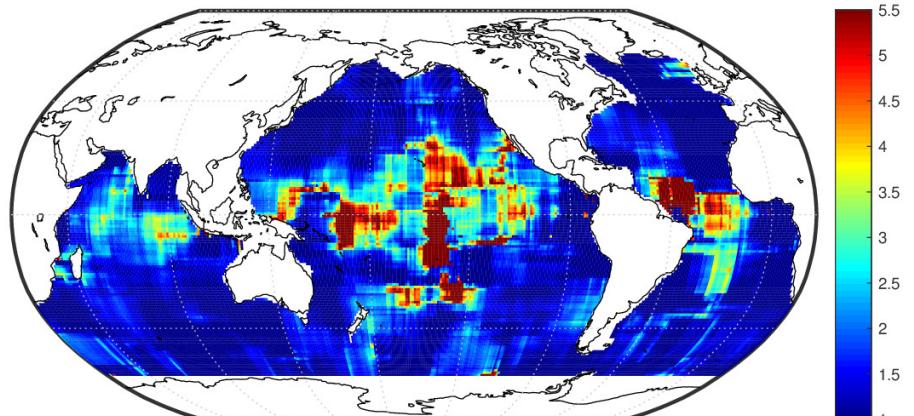
Figure 12: $\sigma^2 / (\phi + \sigma^2)$



(a) 10 db



(b) 300 db



(c) 1500 db

Figure 13: $\theta_{\text{lon}}/\theta_{\text{lat}}$

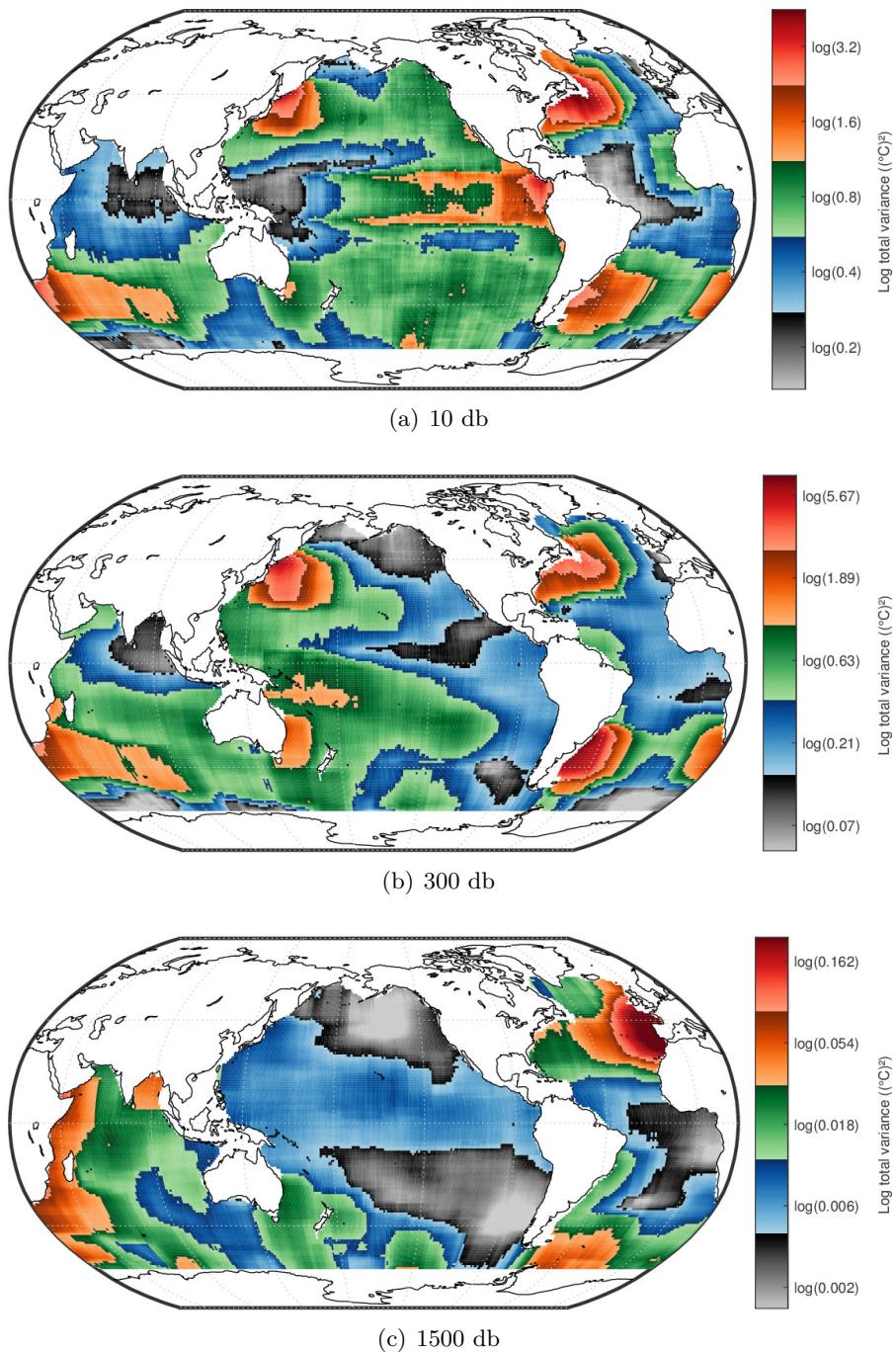
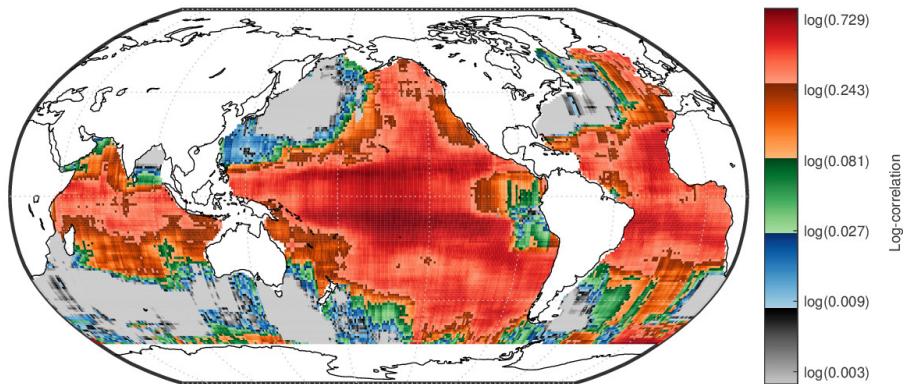
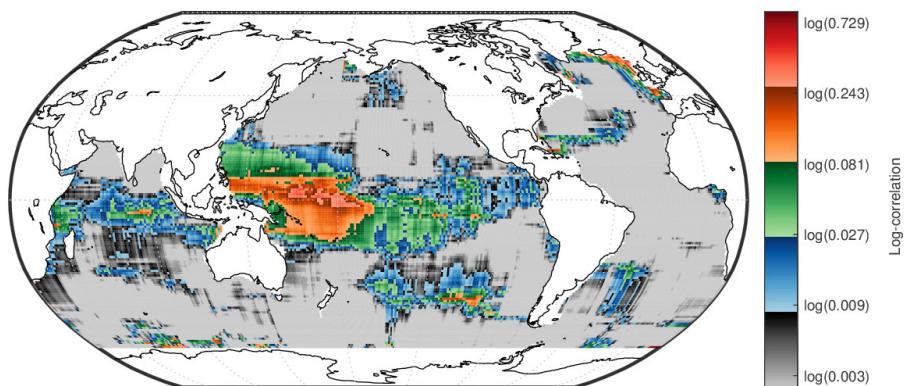


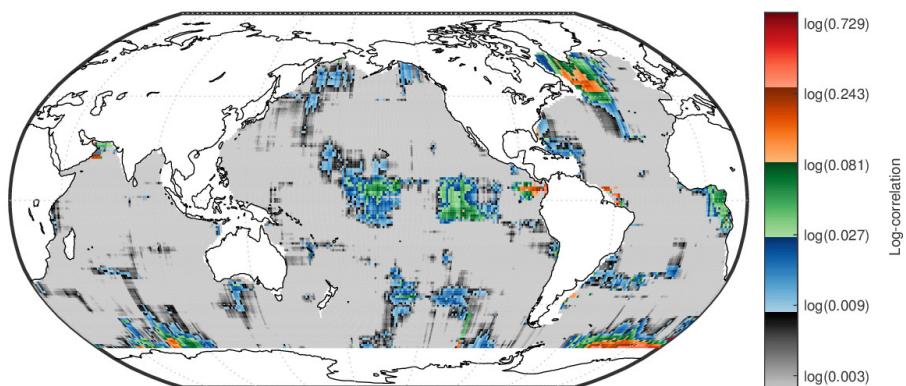
Figure 14: $\phi + \sigma^2$



(a) 10 db

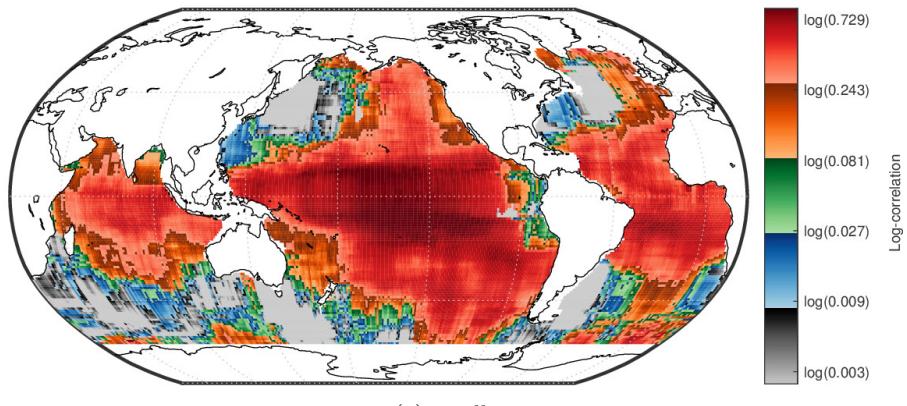


(b) 300 db

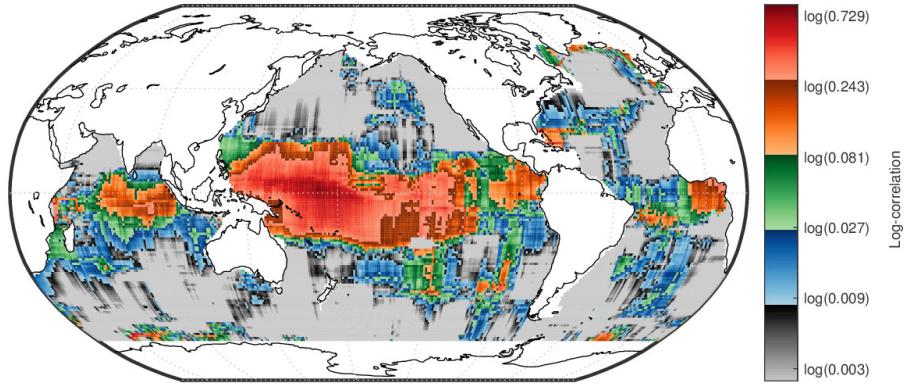


(c) 1500 db

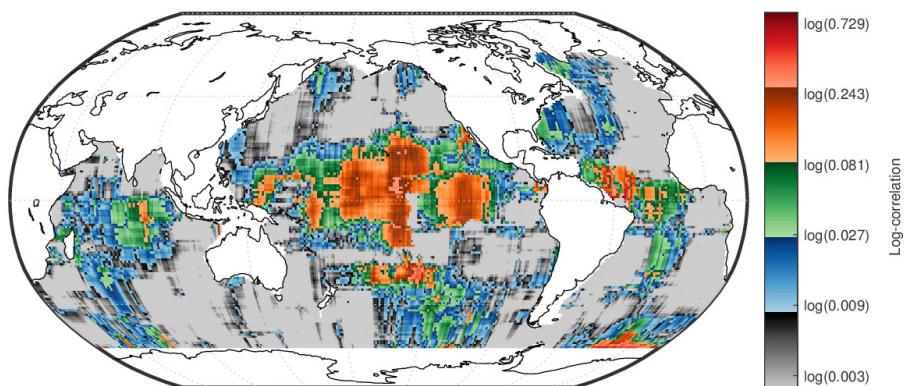
Figure 15: Correlation at $\Delta x_{\text{lat}} = 800 \text{ km}$



(a) 10 db



(b) 300 db



(c) 1500 db

Figure 16: Correlation at $\Delta x_{\text{lon}} = 800 \text{ km}$

2.5 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 $\hat{\nu} \leq 2$.

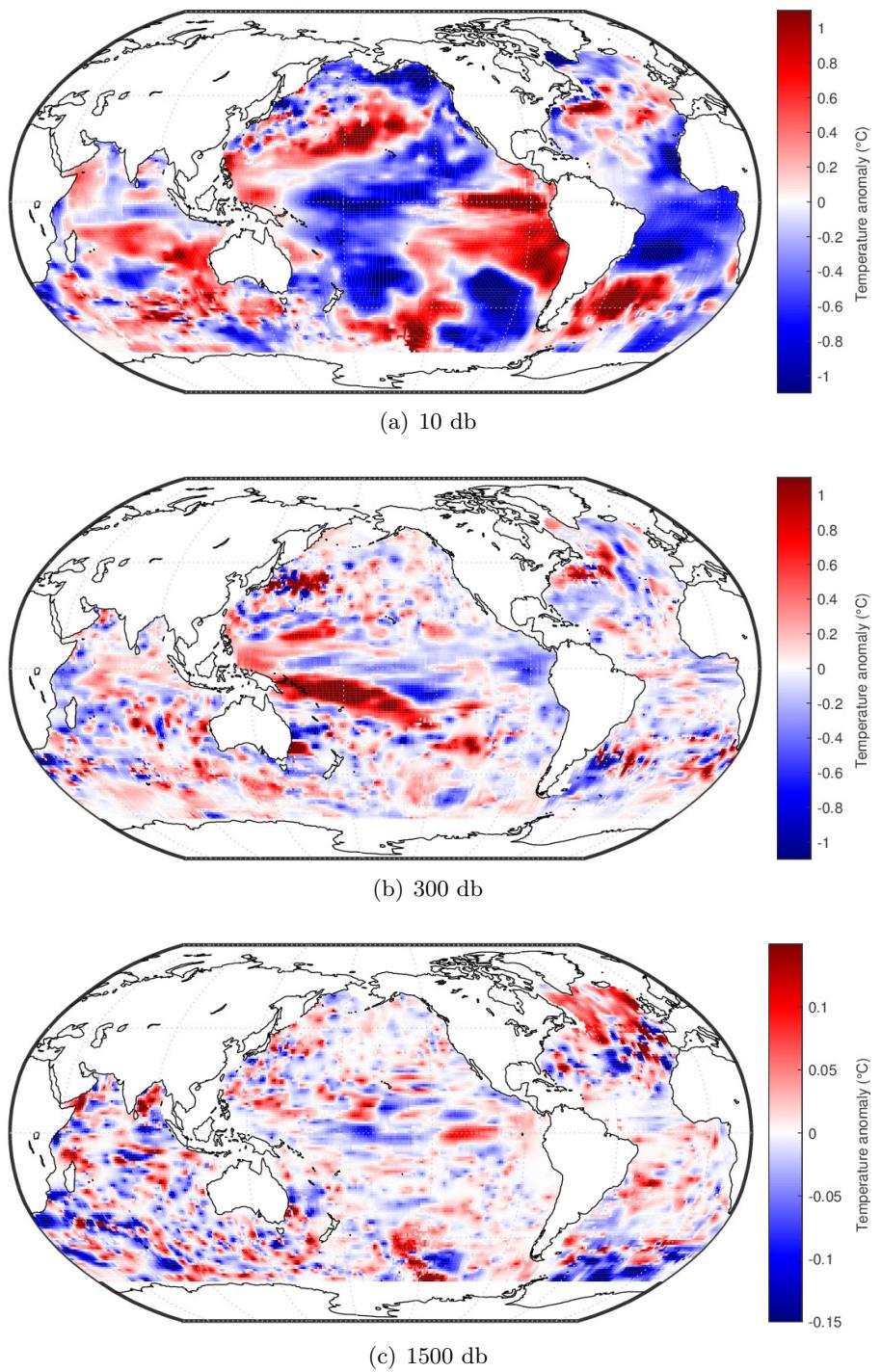
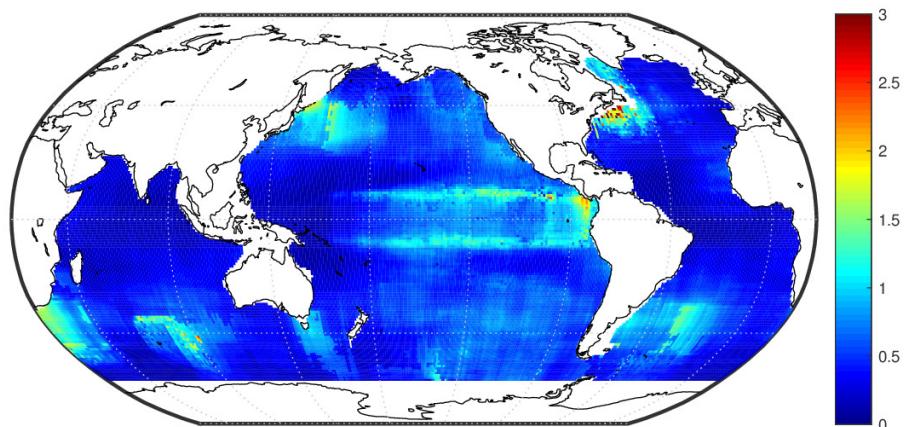
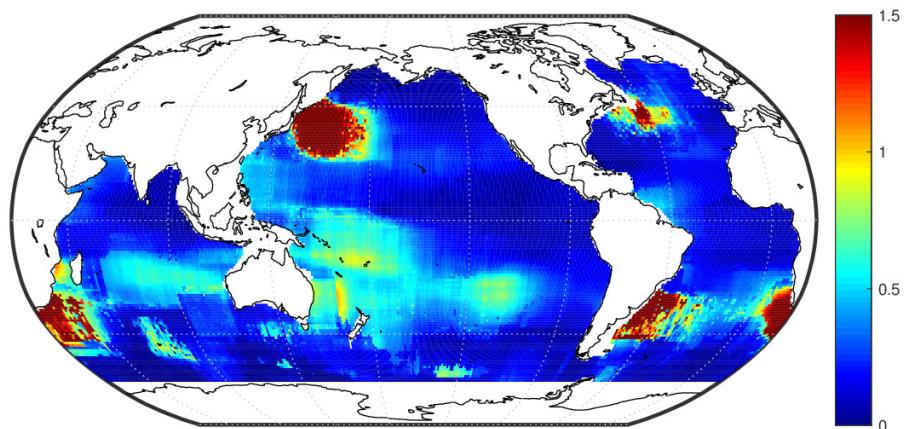


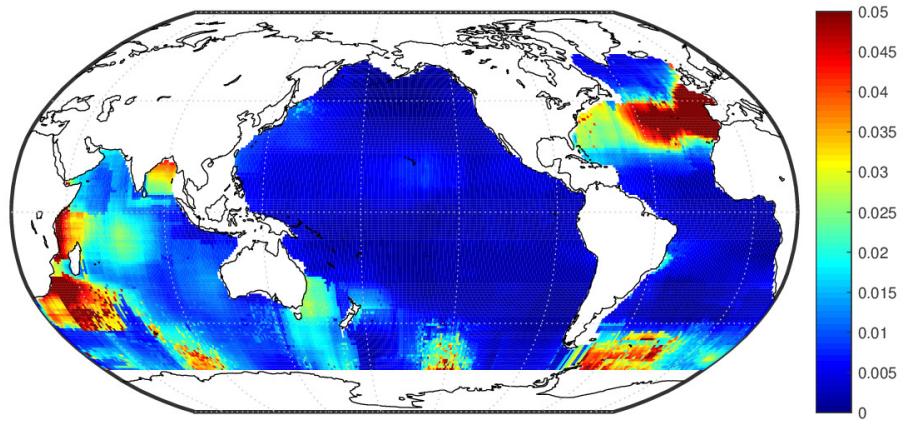
Figure 17: February 2012 temperature anomalies



(a) 10 db

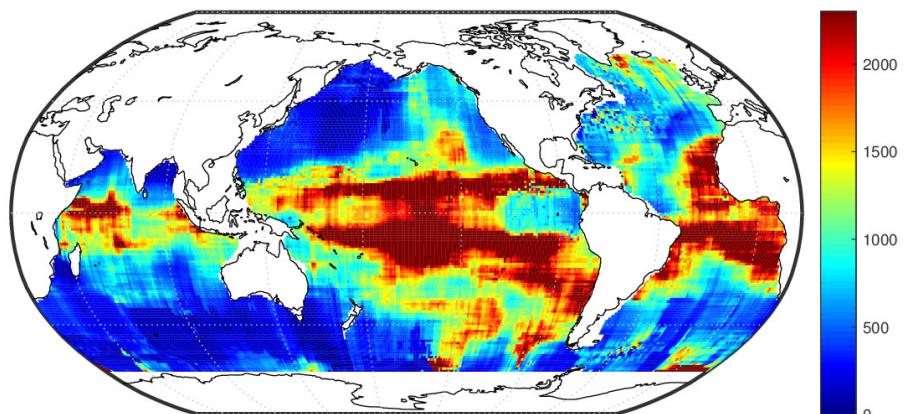


(b) 300 db

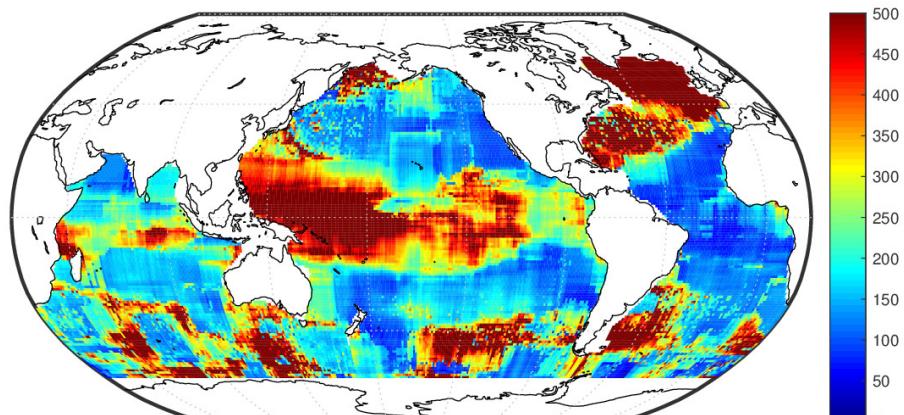


(c) 1500 db

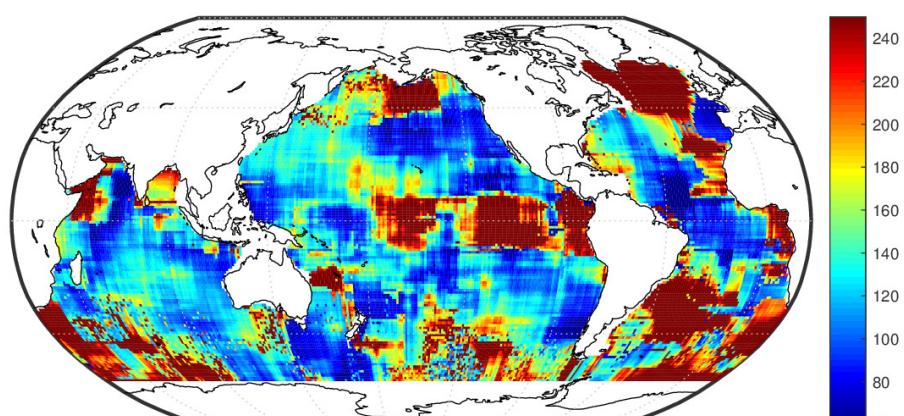
Figure 18: ϕ



(a) 10 db

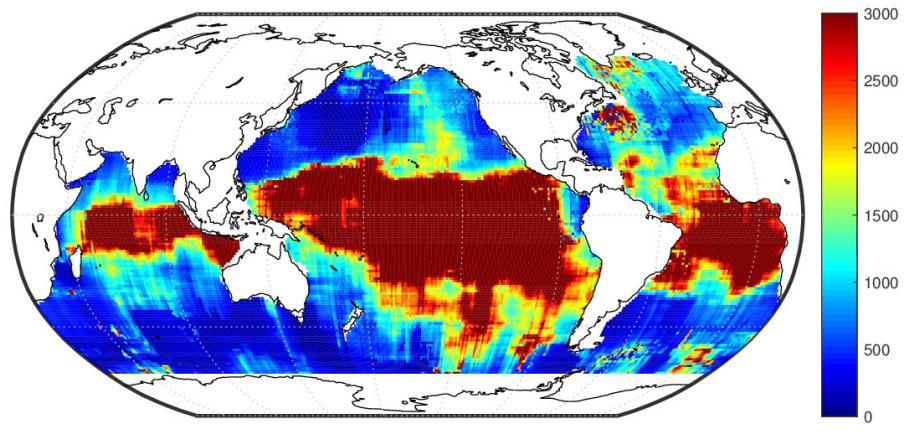


(b) 300 db

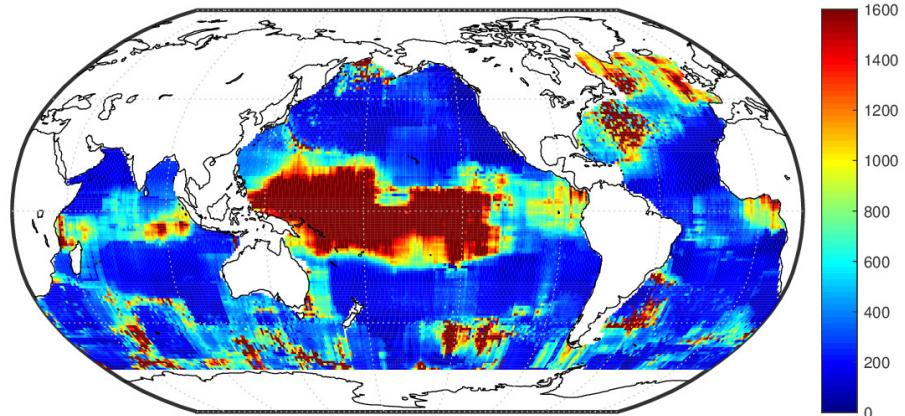


(c) 1500 db

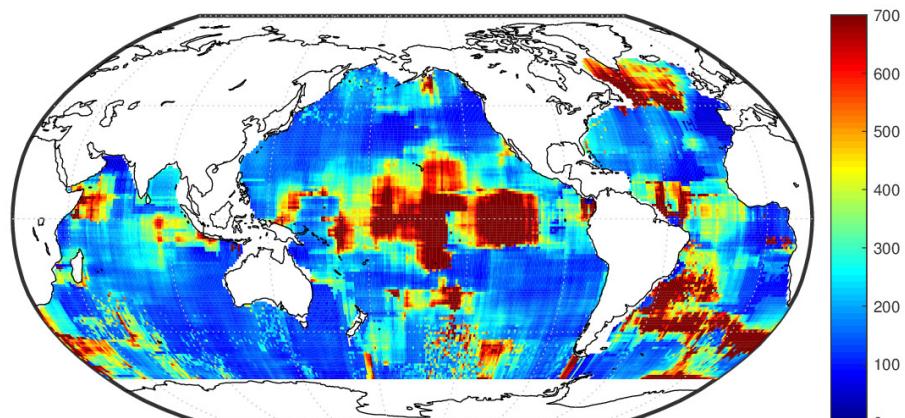
Figure 19: θ_{lat} (in km)



(a) 10 db

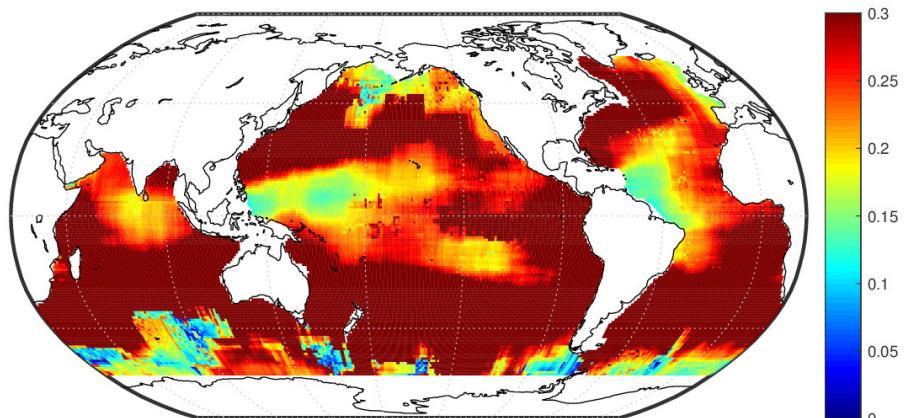


(b) 300 db

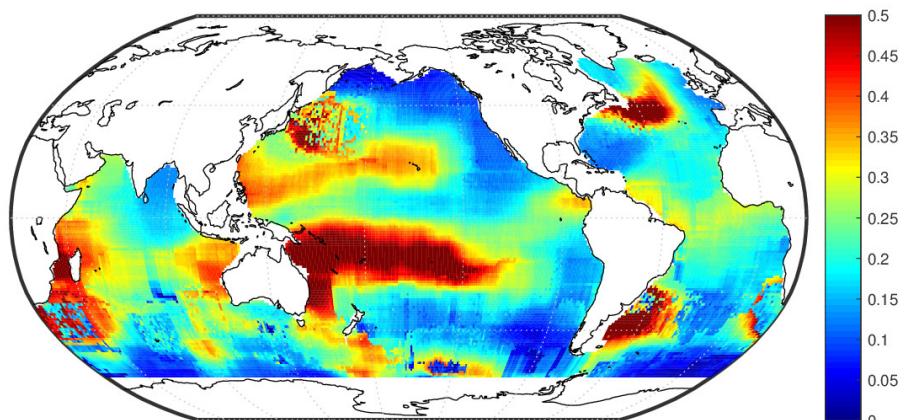


(c) 1500 db

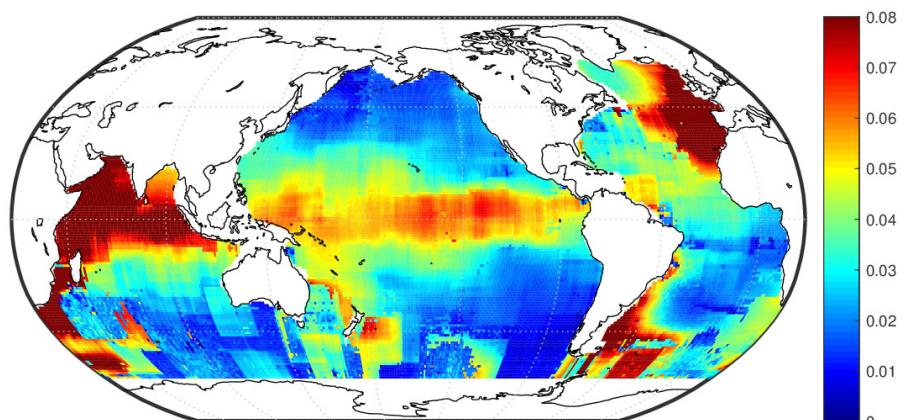
Figure 20: θ_{lon} (in km)



(a) 10 db

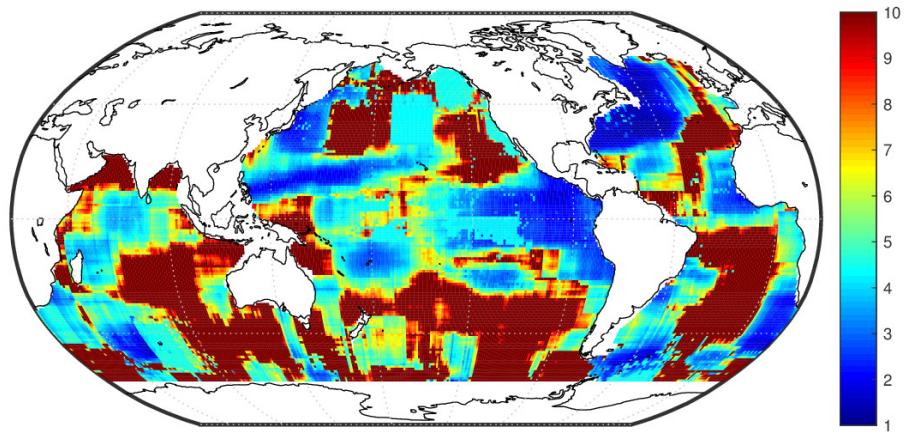


(b) 300 db

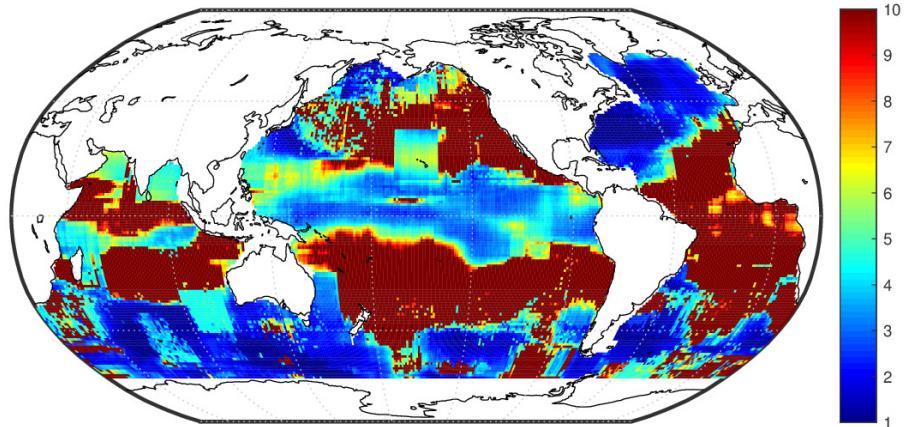


(c) 1500 db

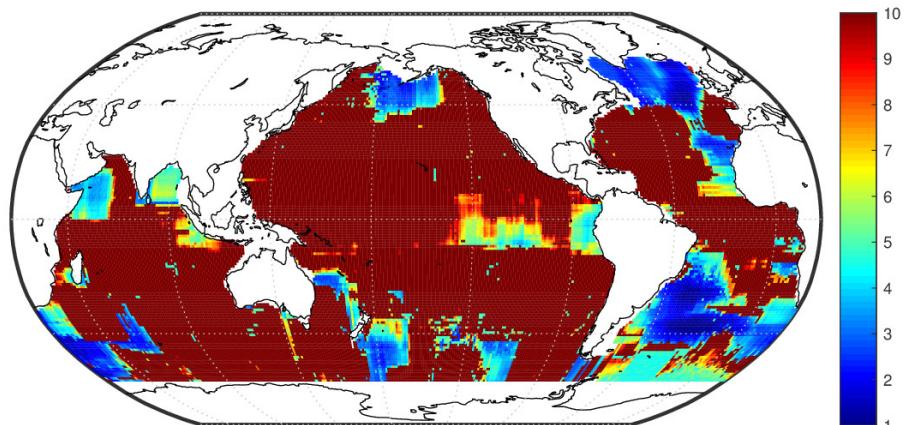
Figure 21: σ



(a) 10 db

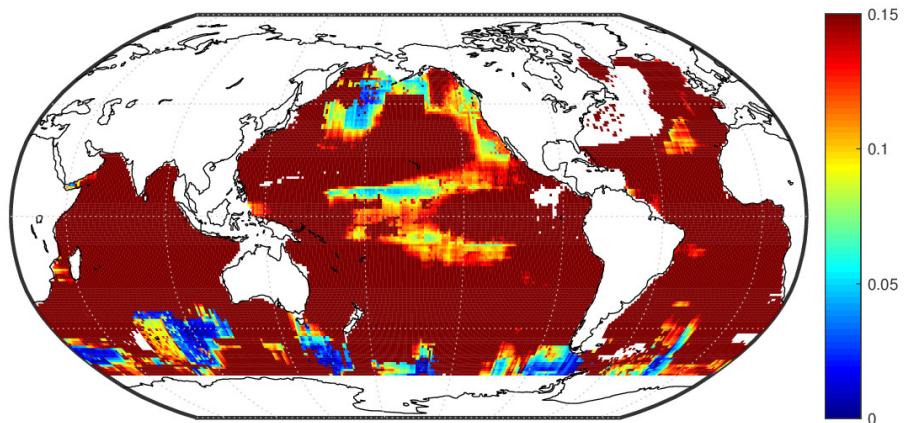


(b) 300 db

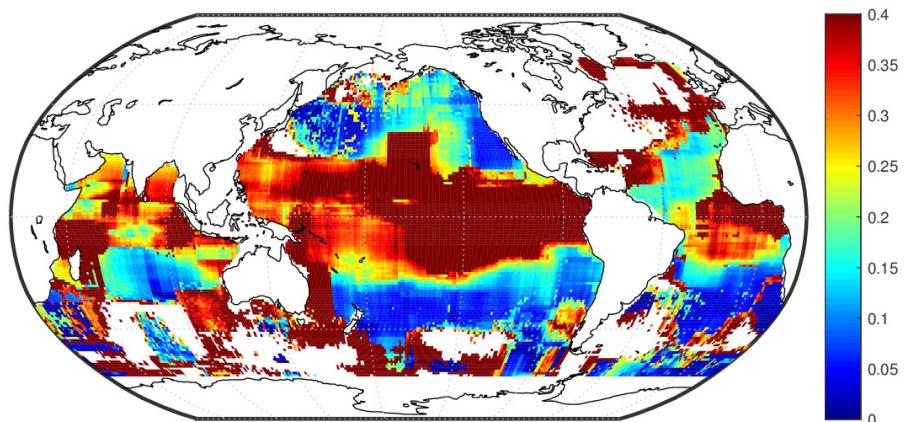


(c) 1500 db

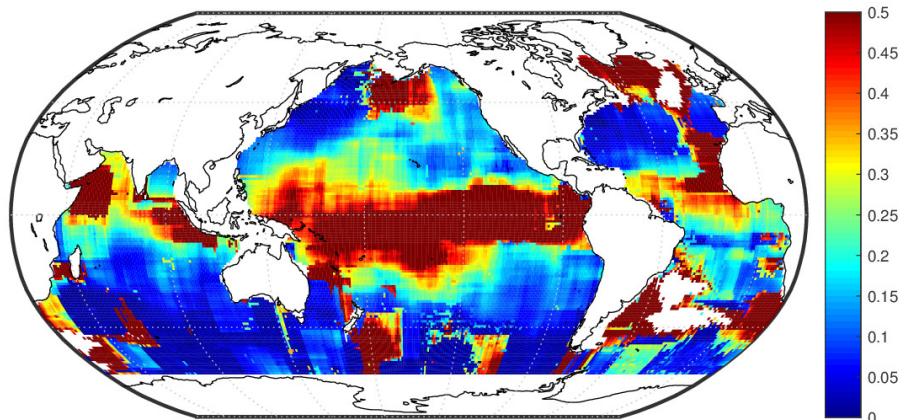
Figure 22: ν



(a) 10 db

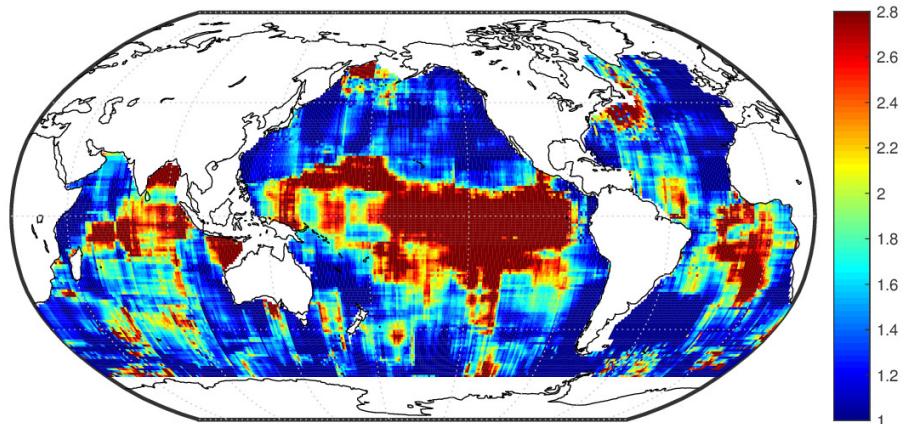


(b) 300 db

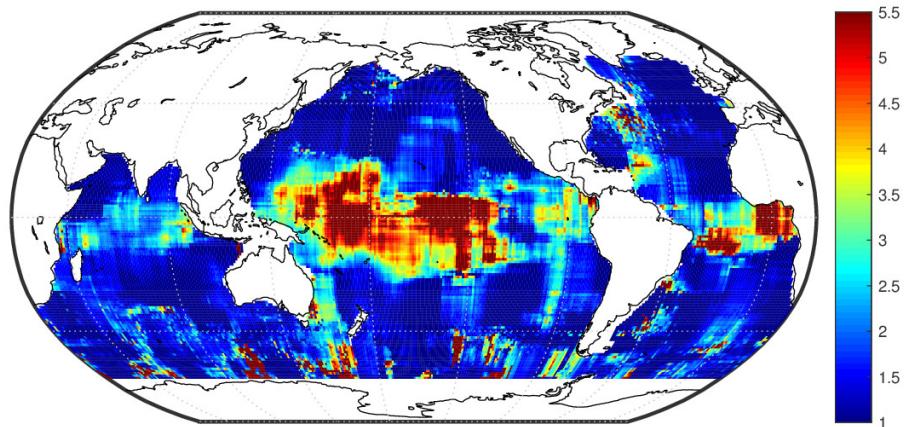


(c) 1500 db

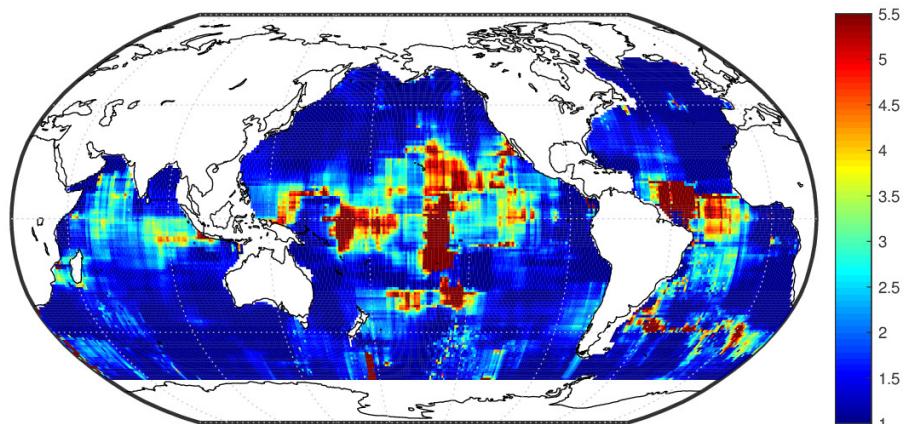
Figure 23: $\frac{\nu}{\nu-2}\sigma^2 / \left(\phi + \frac{\nu}{\nu-2}\sigma^2 \right)$



(a) 10 db



(b) 300 db



(c) 1500 db

Figure 24: $\theta_{\text{lon}}/\theta_{\text{lat}}$

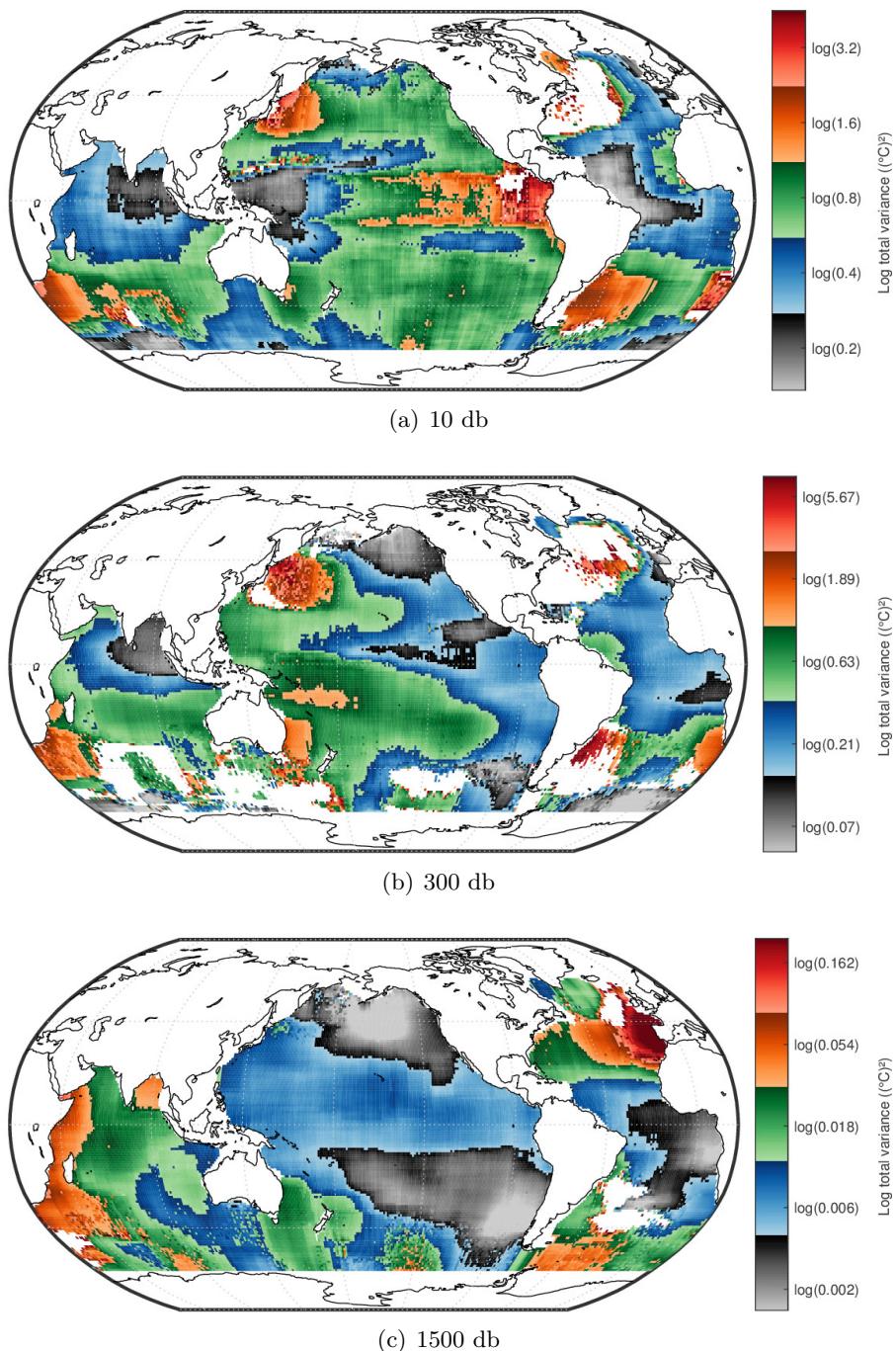
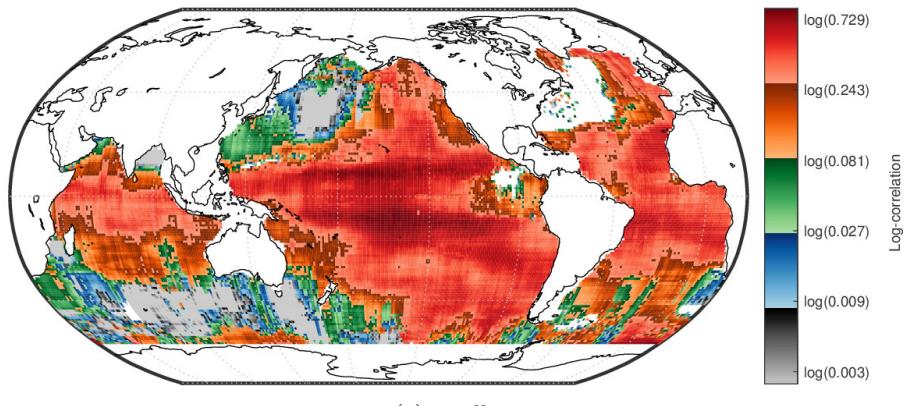
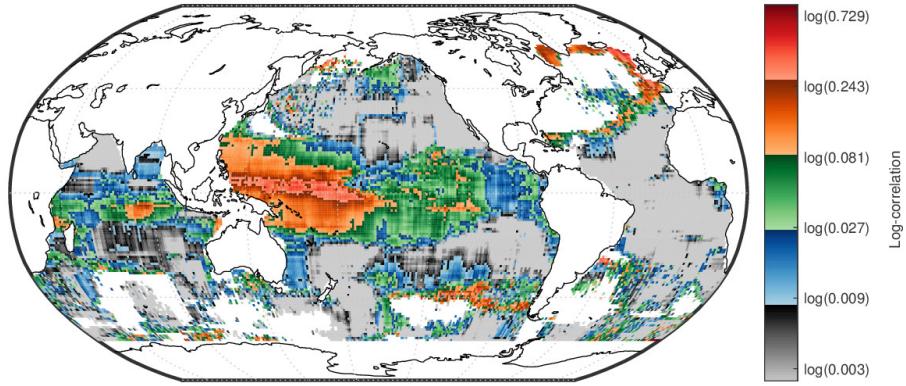


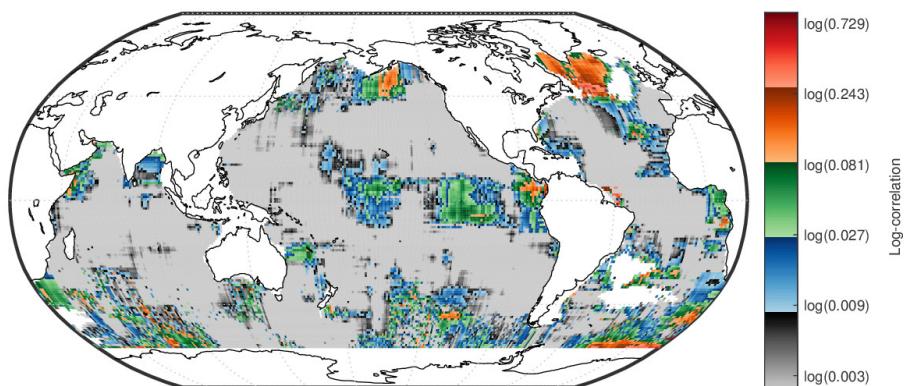
Figure 25: $\phi + \frac{\nu}{\nu-2} \sigma^2$



(a) 10 db

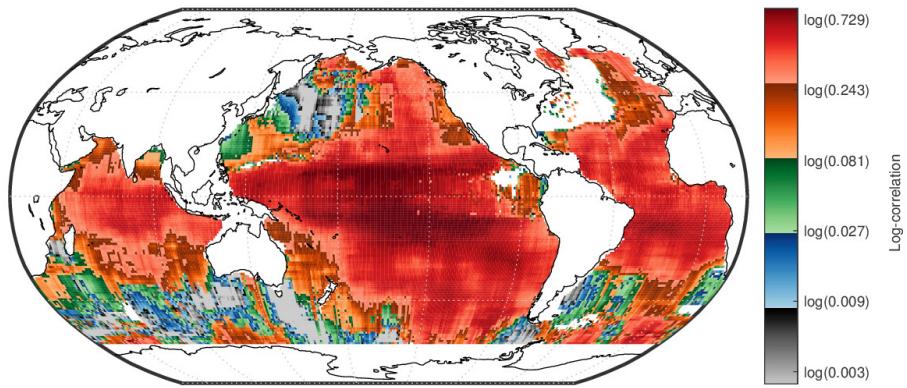


(b) 300 db

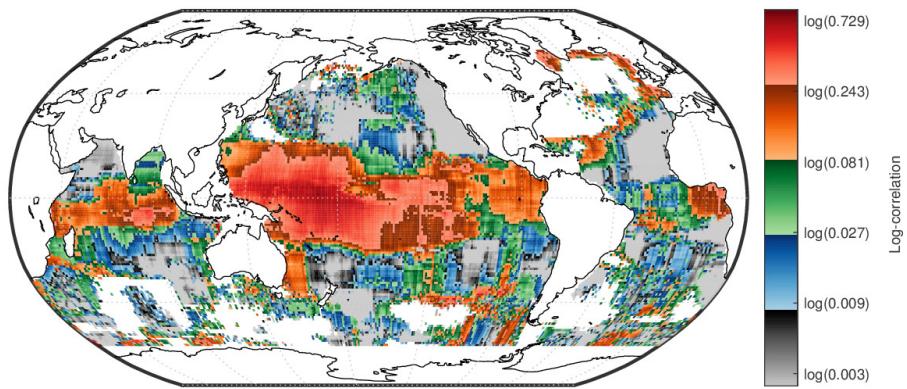


(c) 1500 db

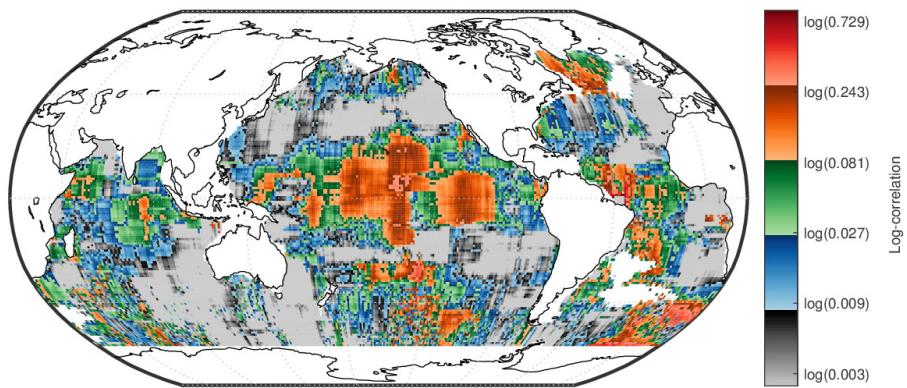
Figure 26: Correlation at $\Delta x_{\text{lat}} = 800 \text{ km}$



(a) 10 db



(b) 300 db



(c) 1500 db

Figure 27: Correlation at $\Delta x_{\text{lon}} = 800 \text{ km}$

2.6 Model 5: Anomalies and model parameters

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

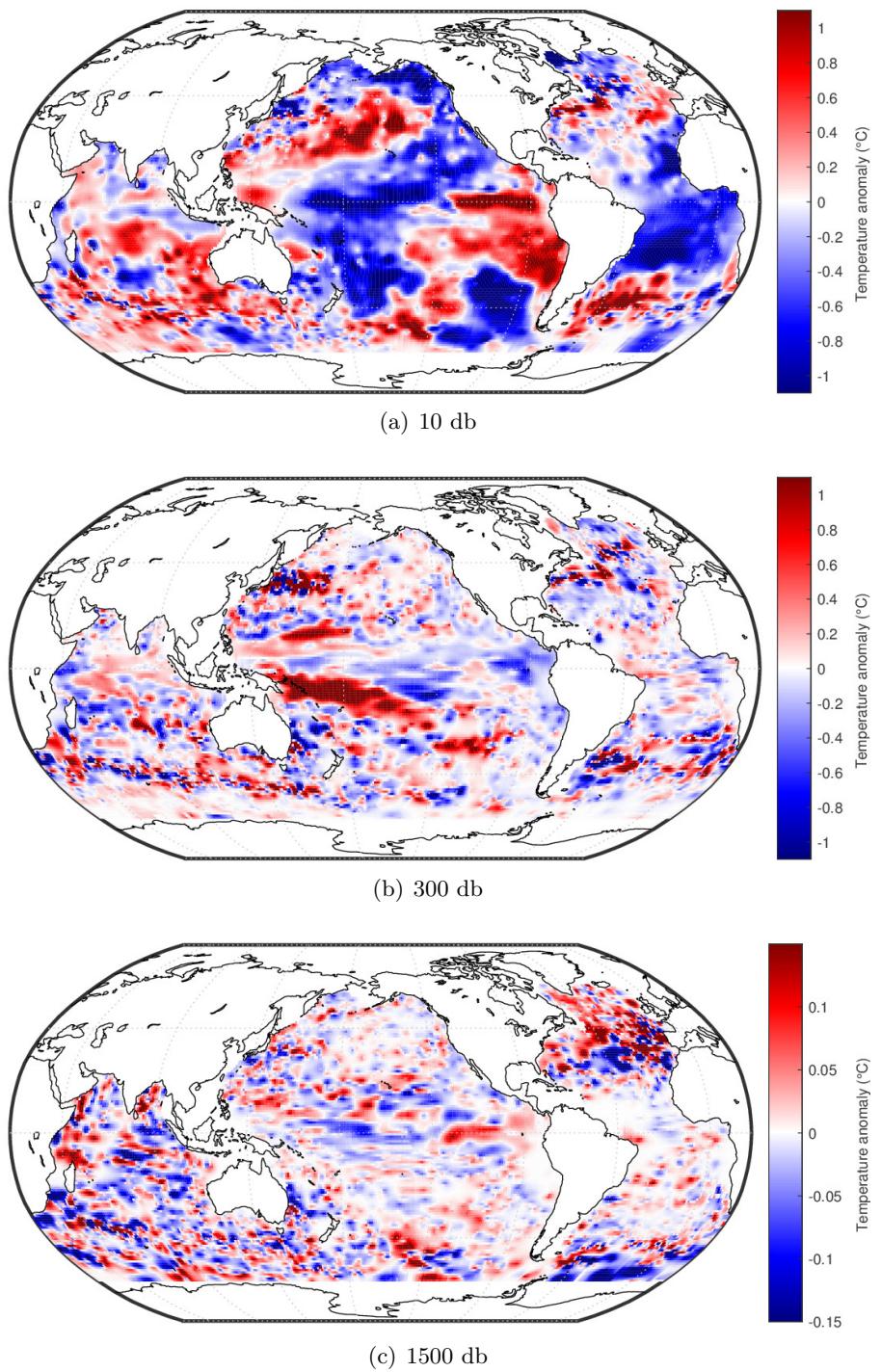


Figure 28: February 2012 temperature anomalies

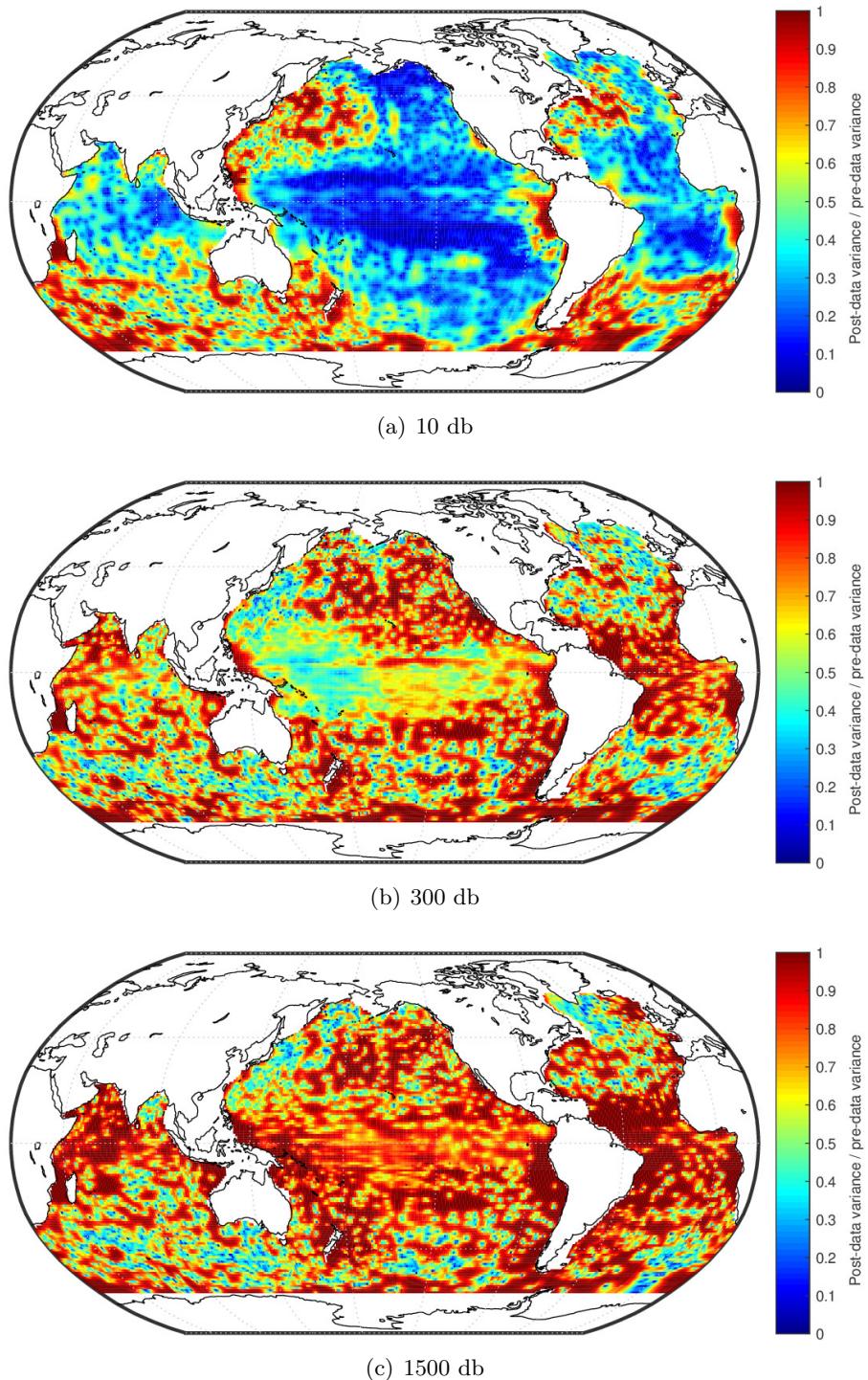
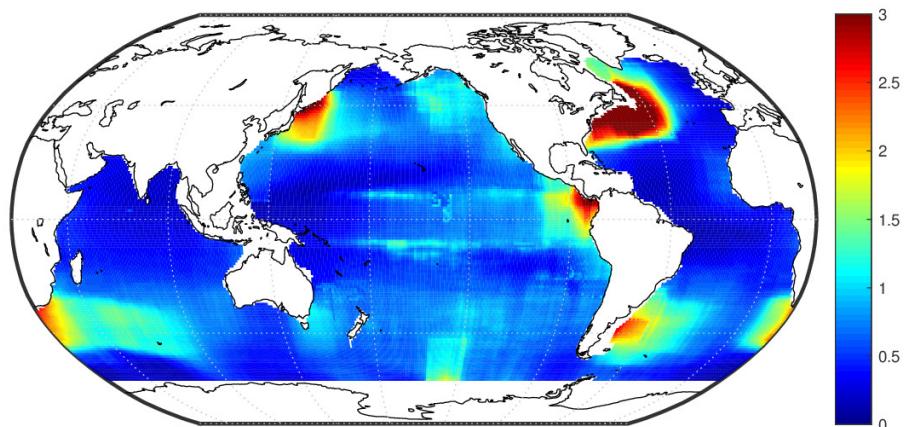
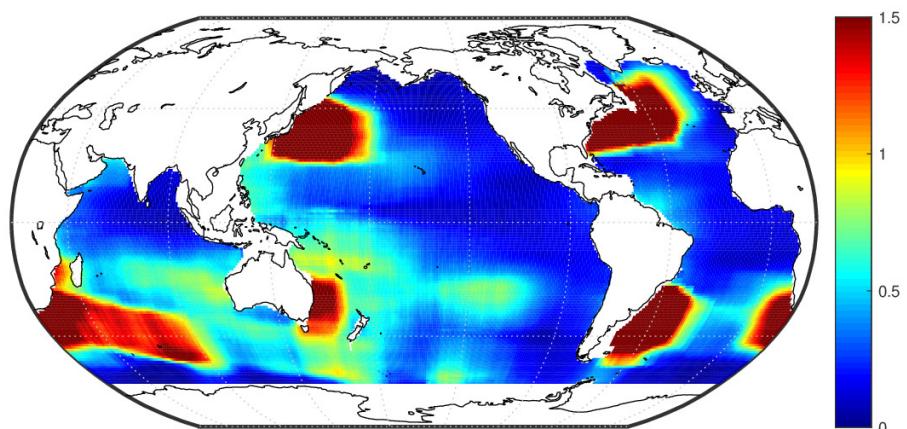


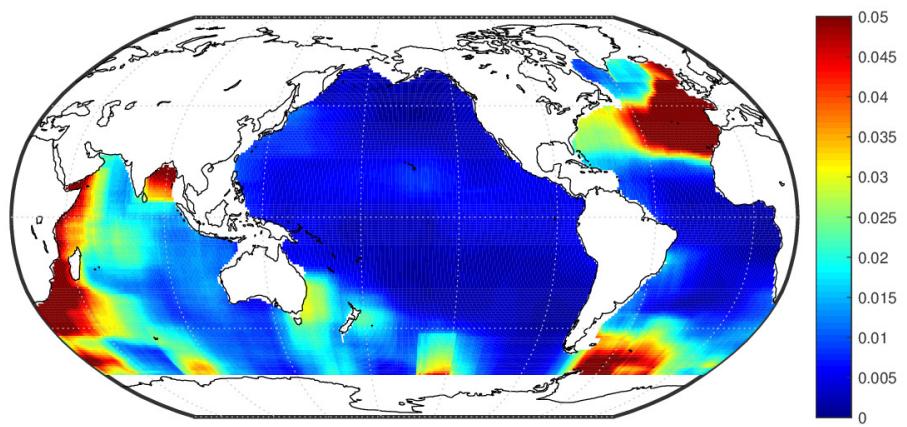
Figure 29: February 2012 post-data-to-pre-data variance ratios



(a) 10 db

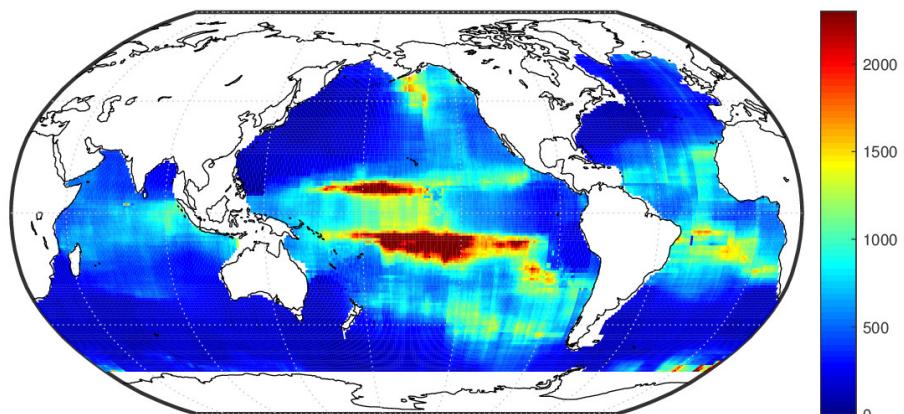


(b) 300 db

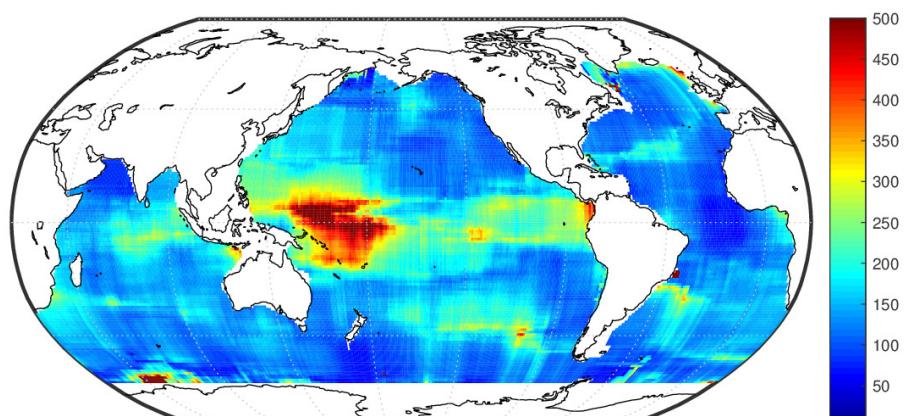


(c) 1500 db

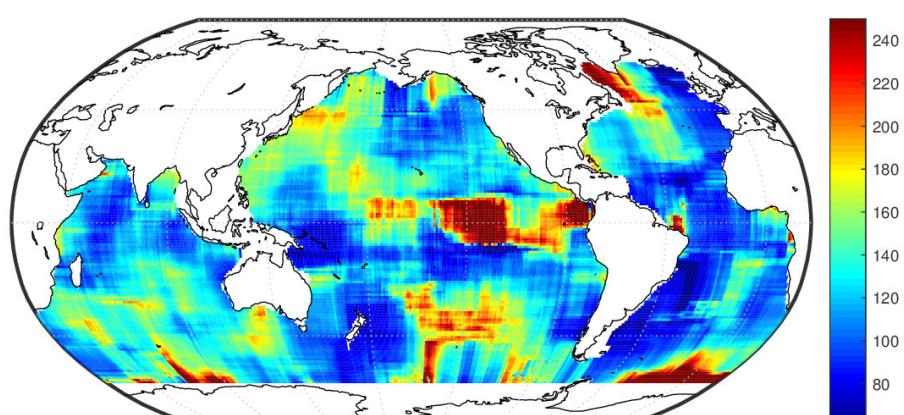
Figure 30: ϕ



(a) 10 db

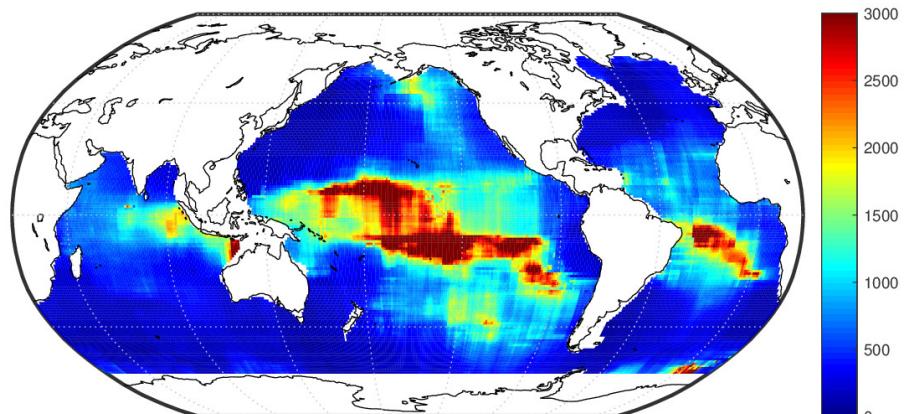


(b) 300 db

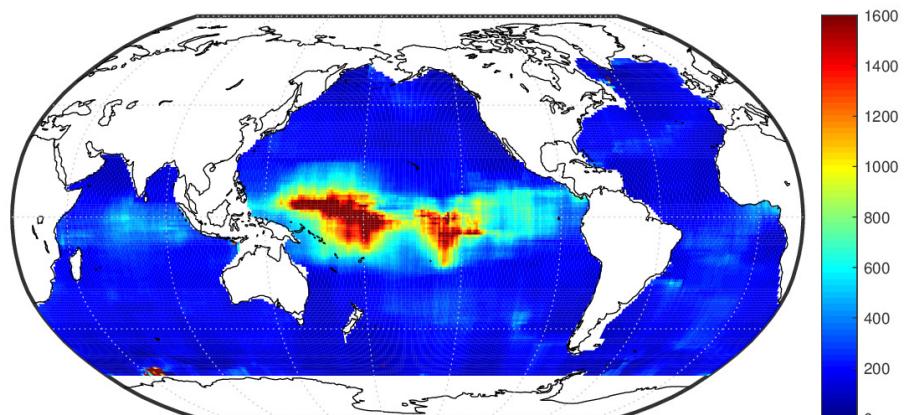


(c) 1500 db

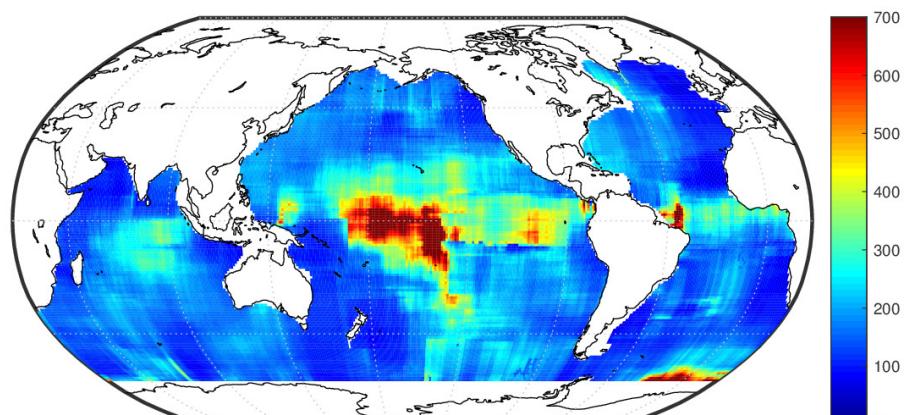
Figure 31: θ_{lat} (in km)



(a) 10 db

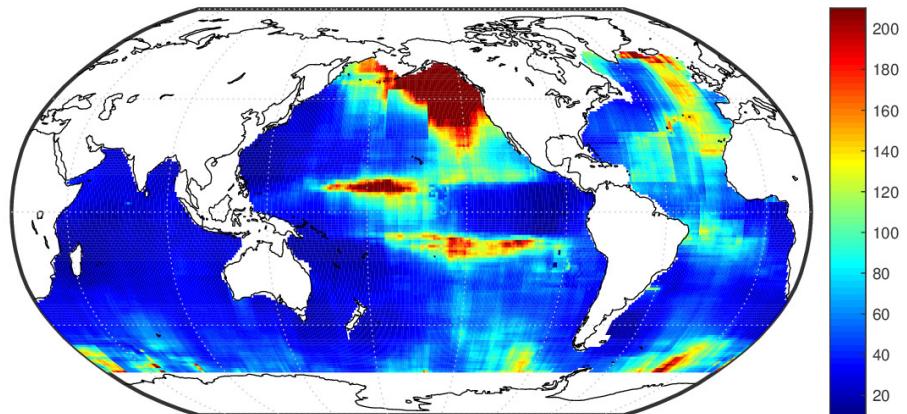


(b) 300 db

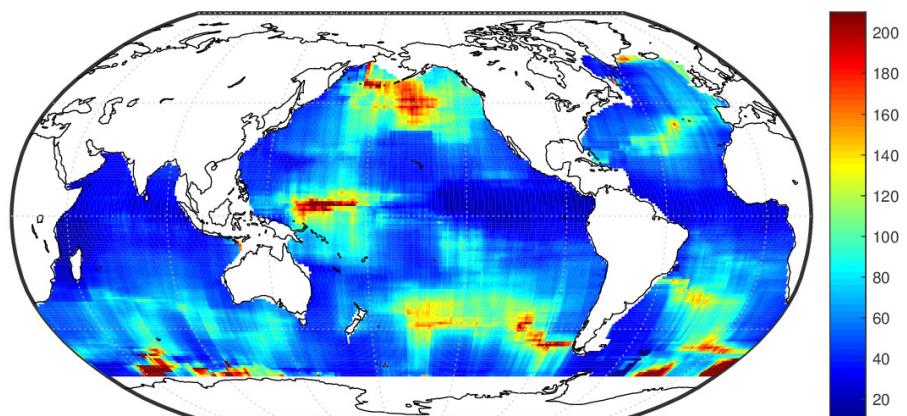


(c) 1500 db

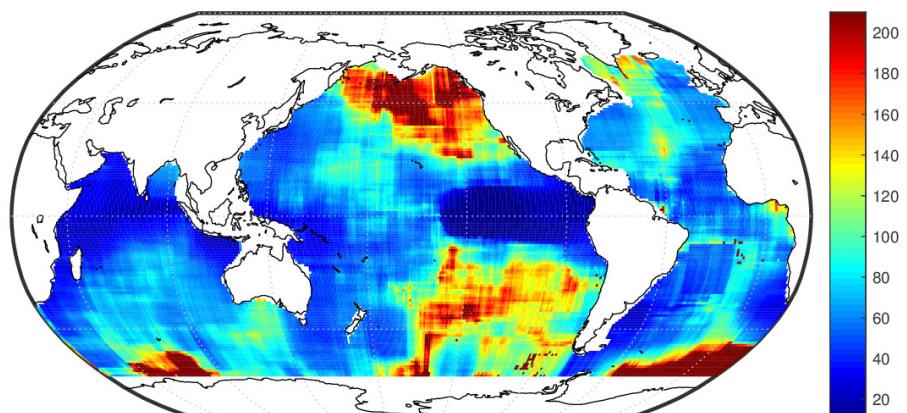
Figure 32: θ_{ion} (in km)



(a) 10 db

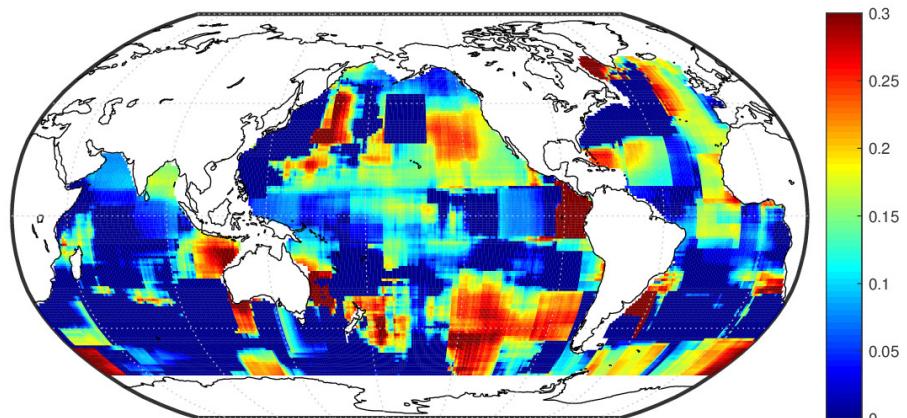


(b) 300 db

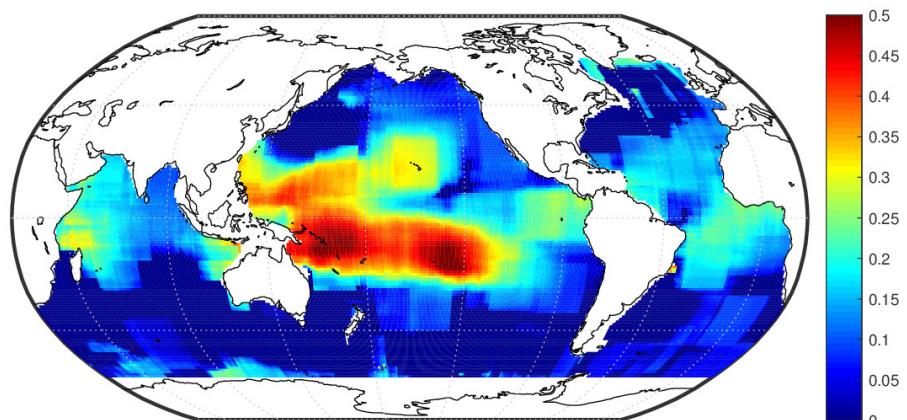


(c) 1500 db

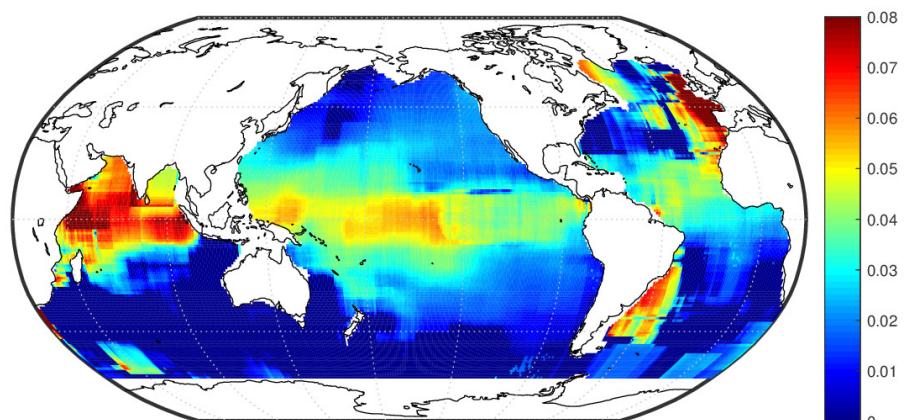
Figure 33: θ_t (in days)



(a) 10 db

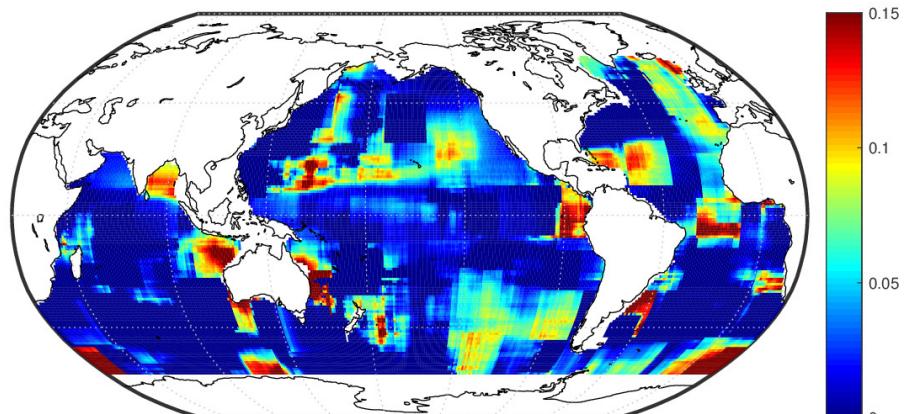


(b) 300 db

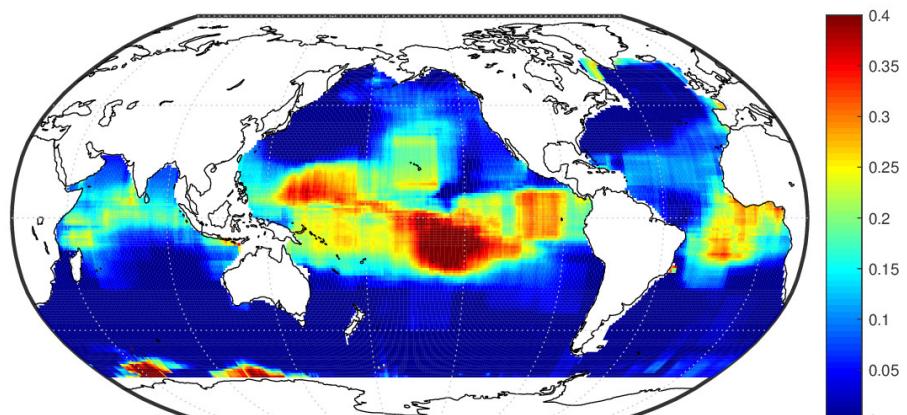


(c) 1500 db

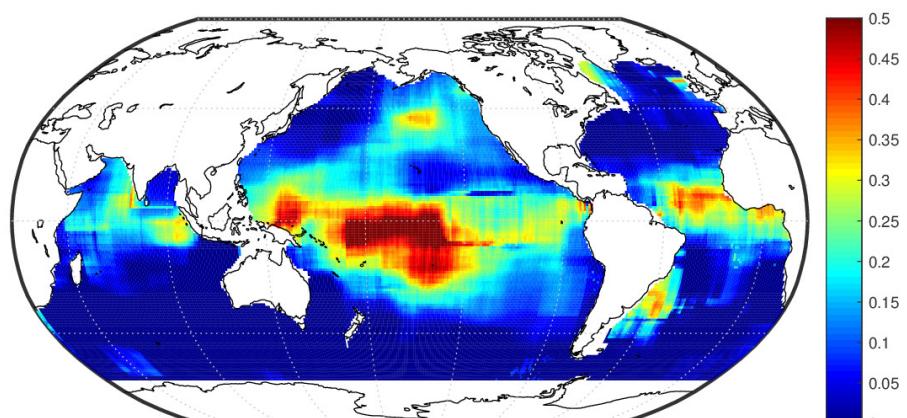
Figure 34: σ



(a) 10 db

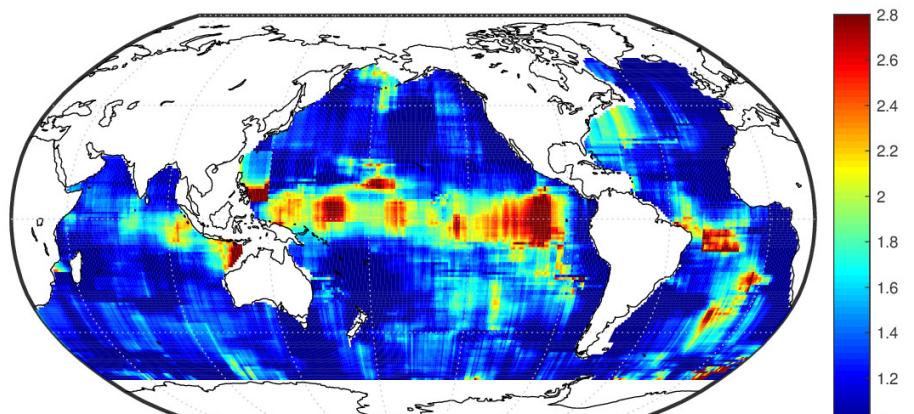


(b) 300 db

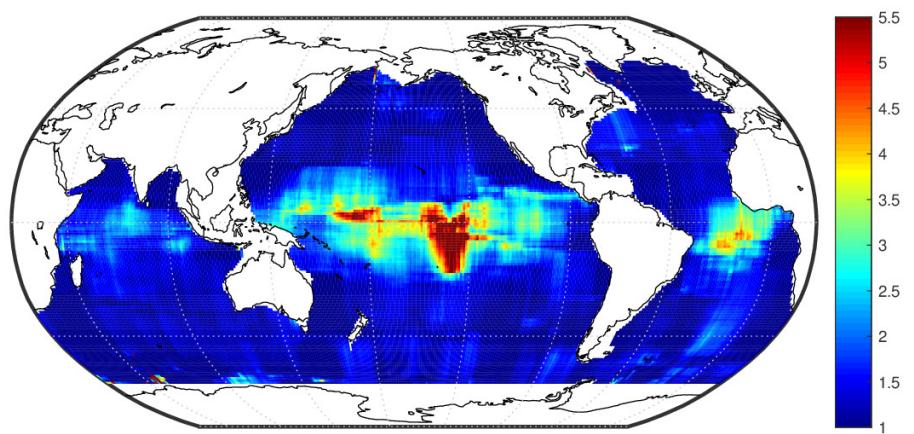


(c) 1500 db

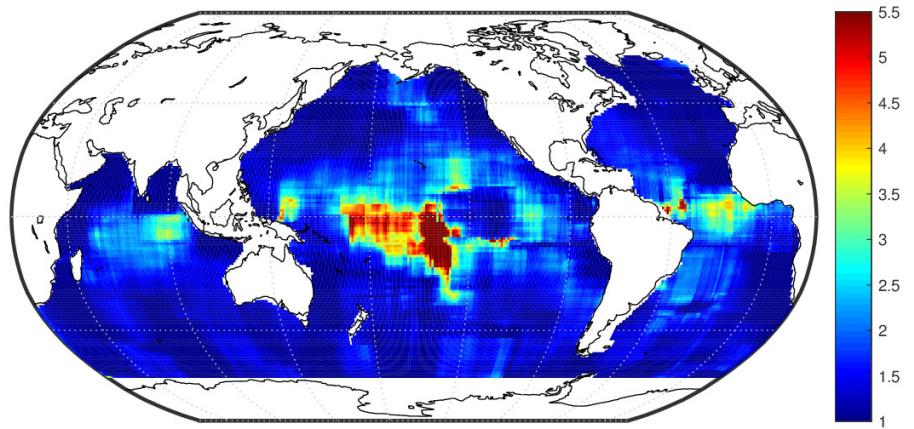
Figure 35: $\sigma^2 / (\phi + \sigma^2)$



(a) 10 db



(b) 300 db



(c) 1500 db

Figure 36: $\theta_{\text{lon}}/\theta_{\text{lat}}$

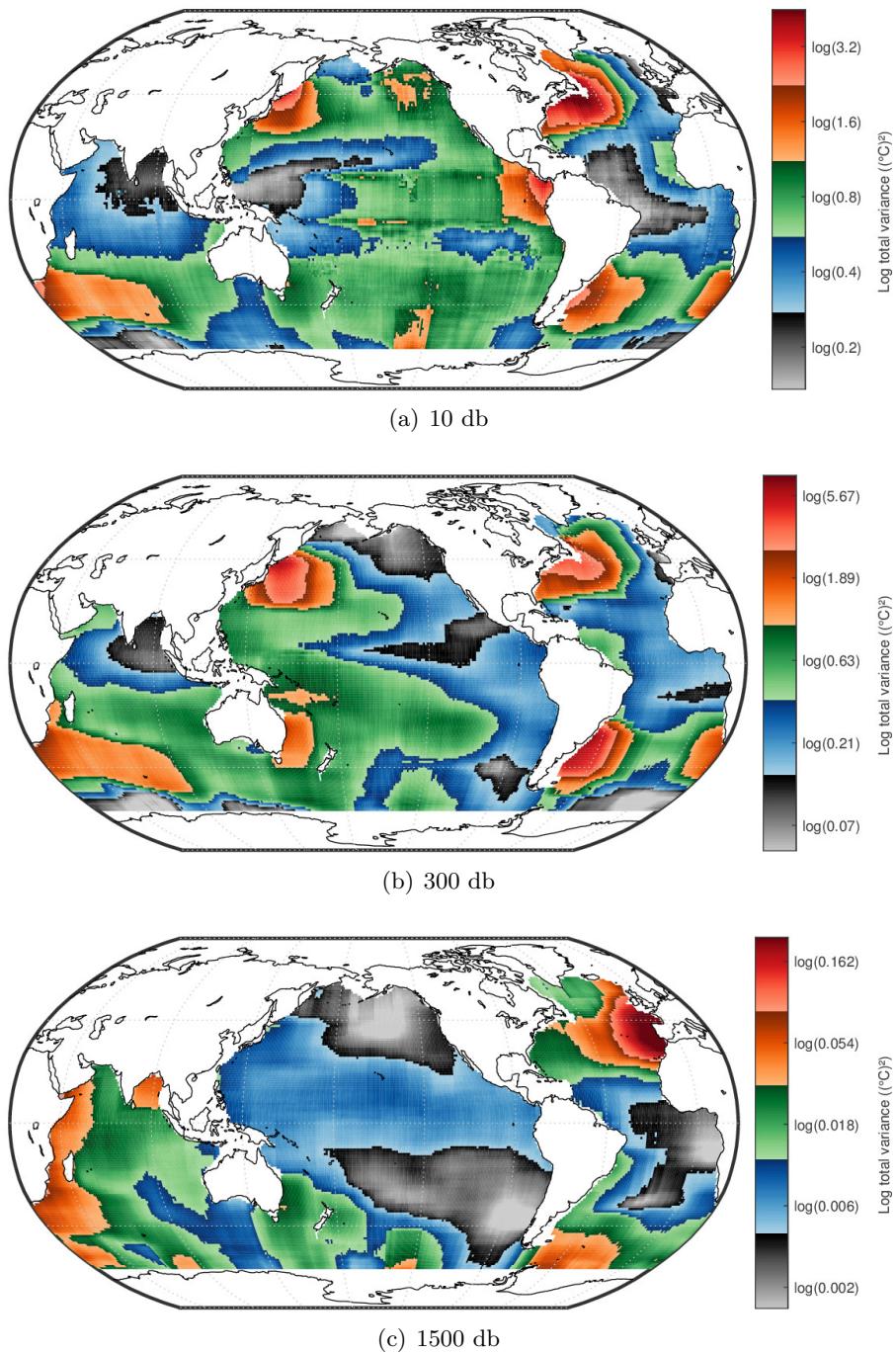
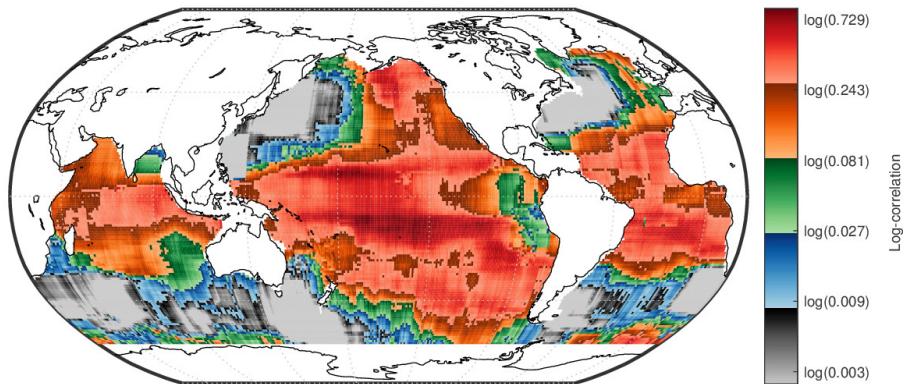
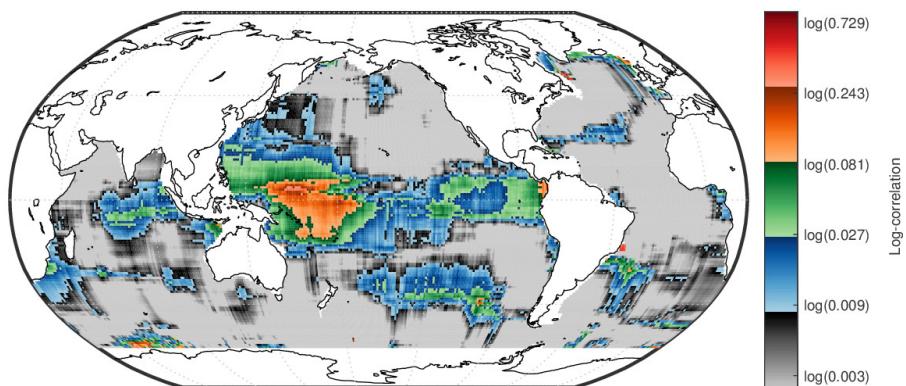


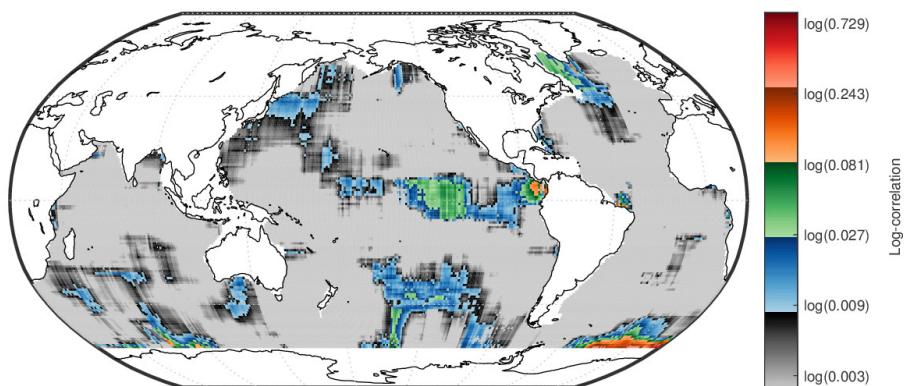
Figure 37: $\phi + \sigma^2$



(a) 10 db

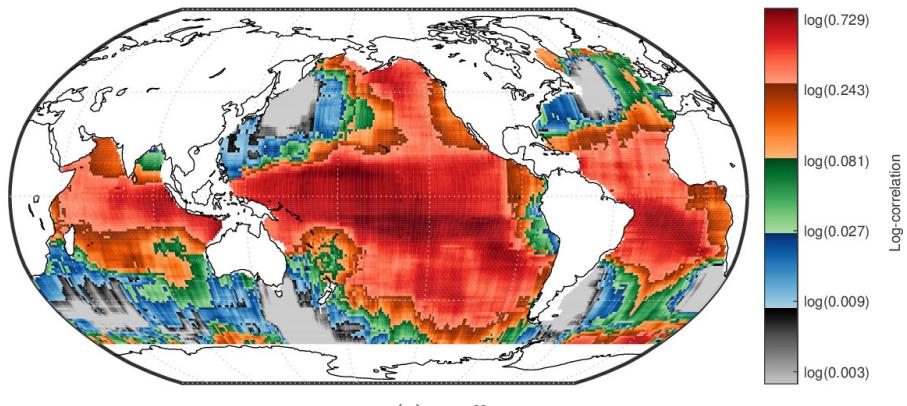


(b) 300 db

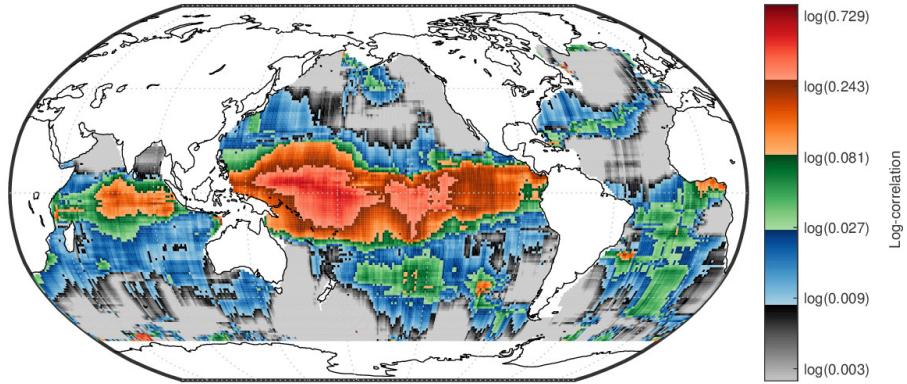


(c) 1500 db

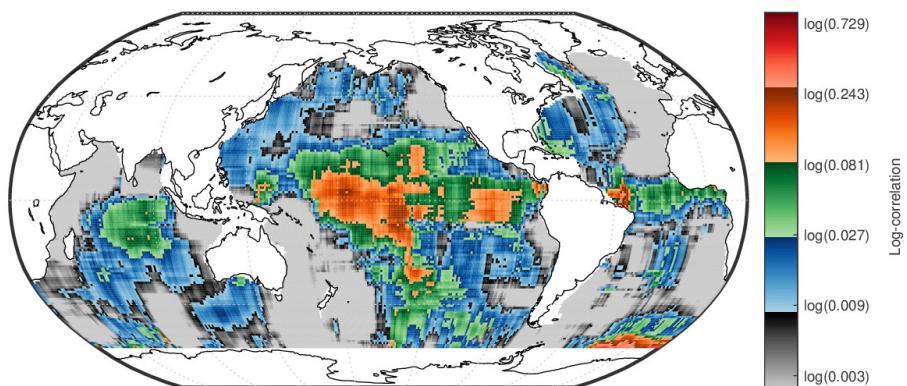
Figure 38: Correlation at $\Delta x_{\text{lat}} = 800 \text{ km}$



(a) 10 db



(b) 300 db



(c) 1500 db

Figure 39: Correlation at $\Delta x_{\text{lon}} = 800 \text{ km}$

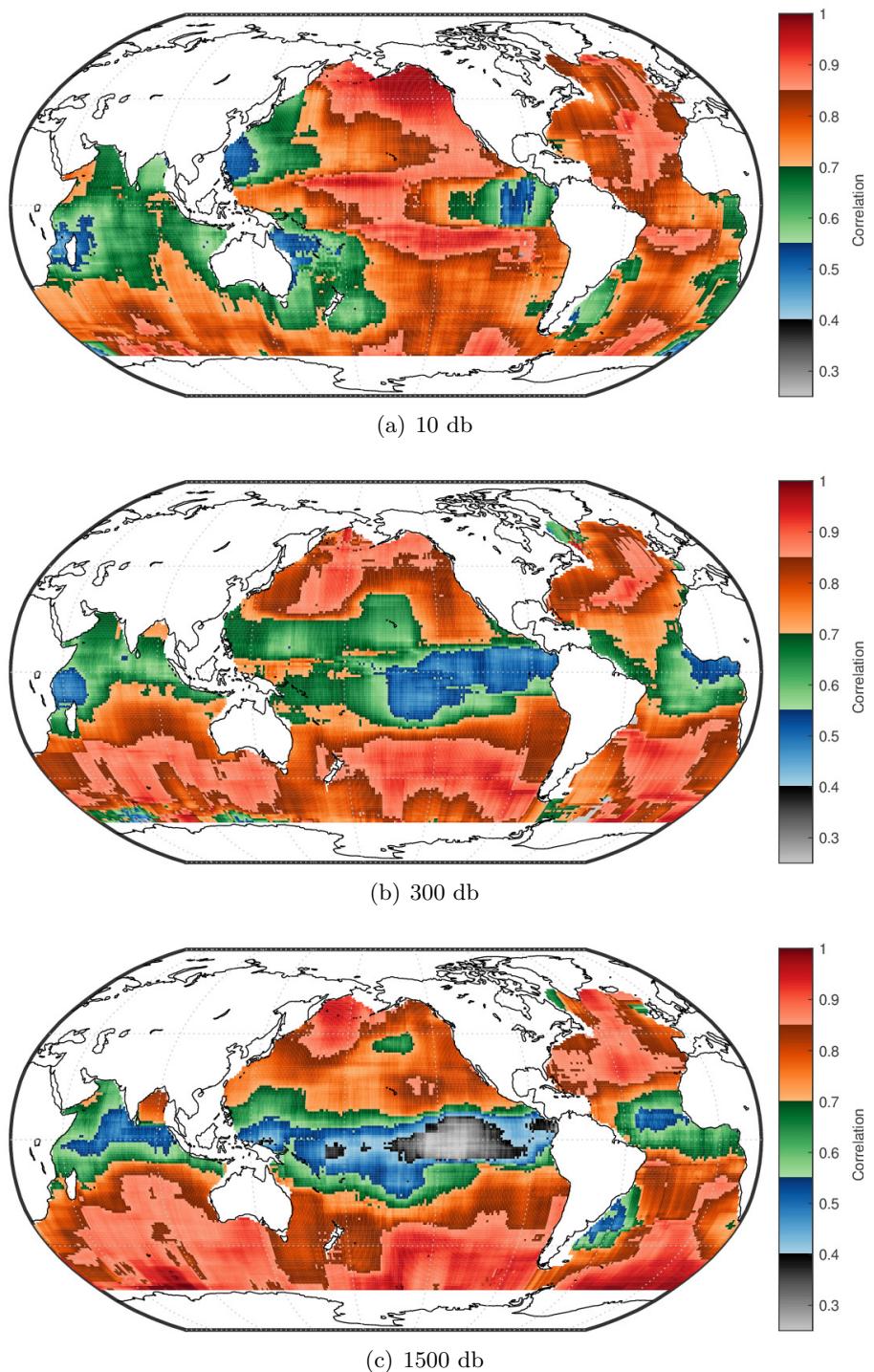


Figure 40: Correlation at $\Delta t = 10$ days

2.7 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 $\hat{\nu} \leq 2$.

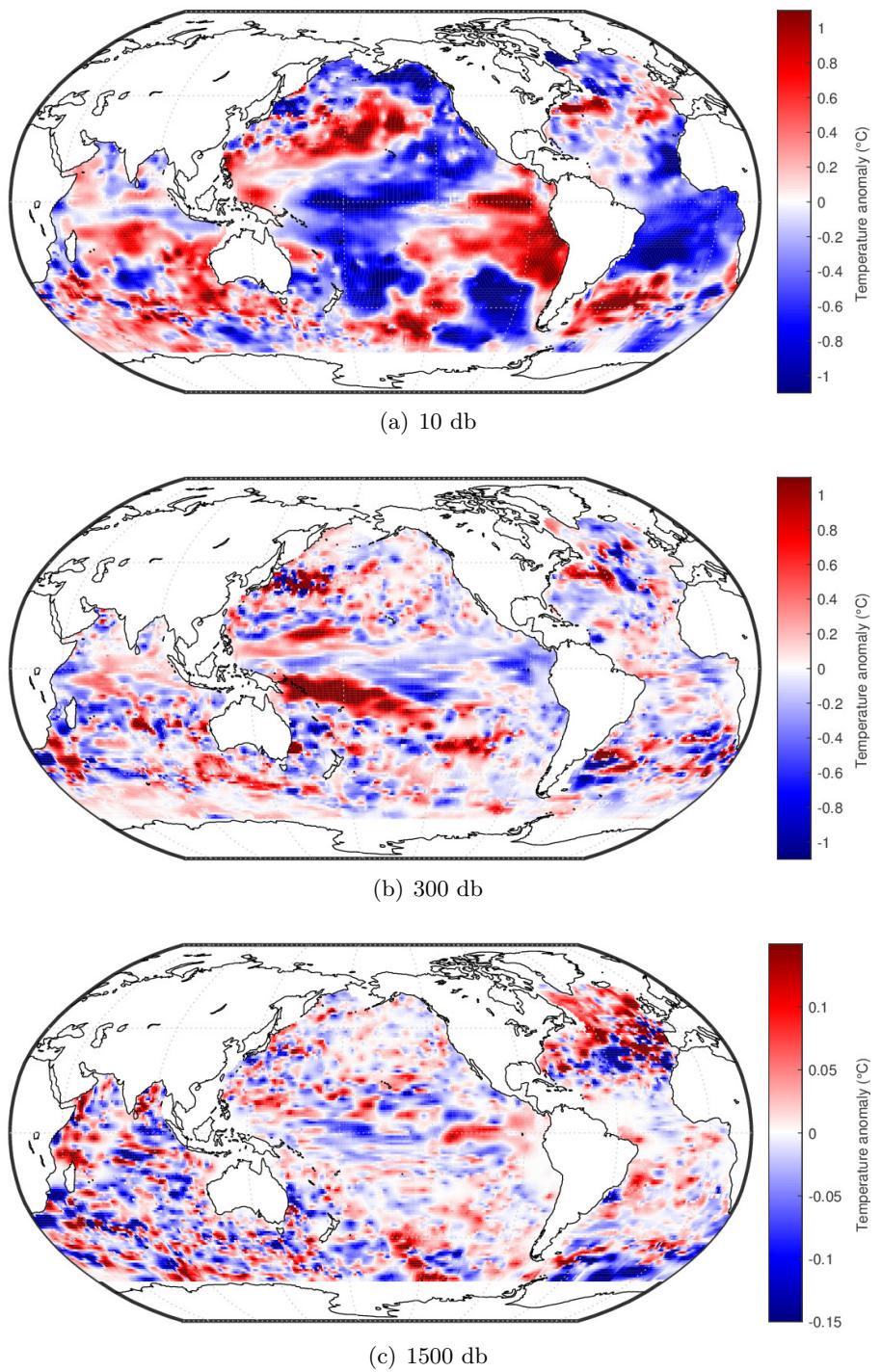
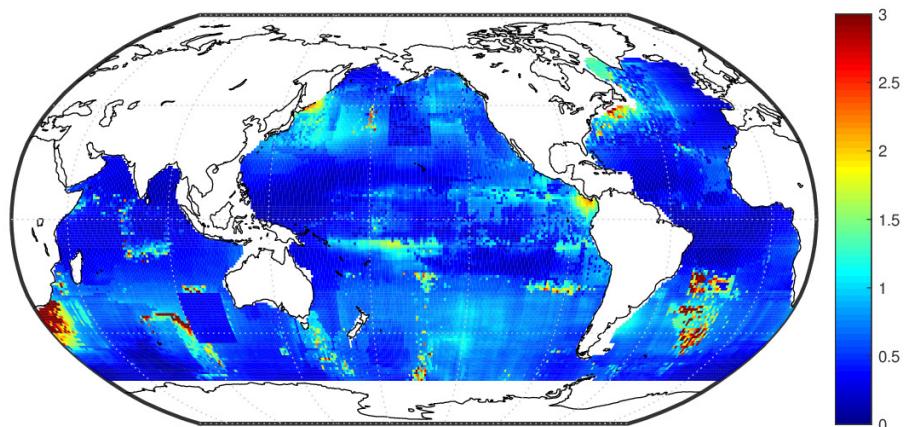
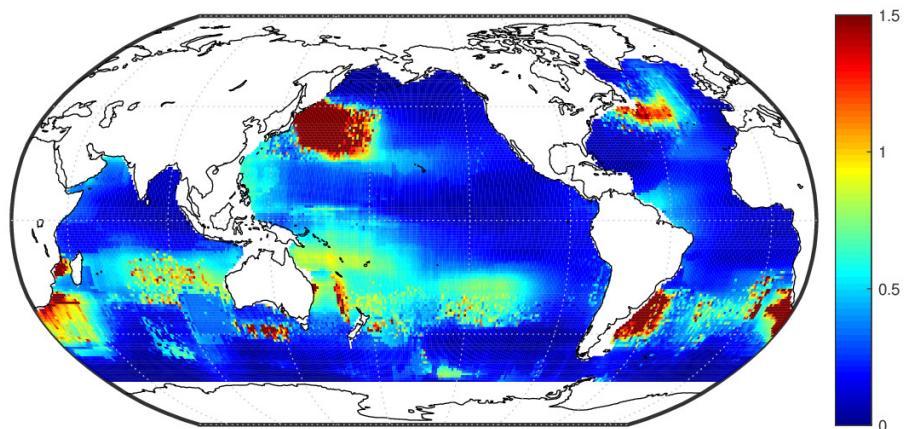


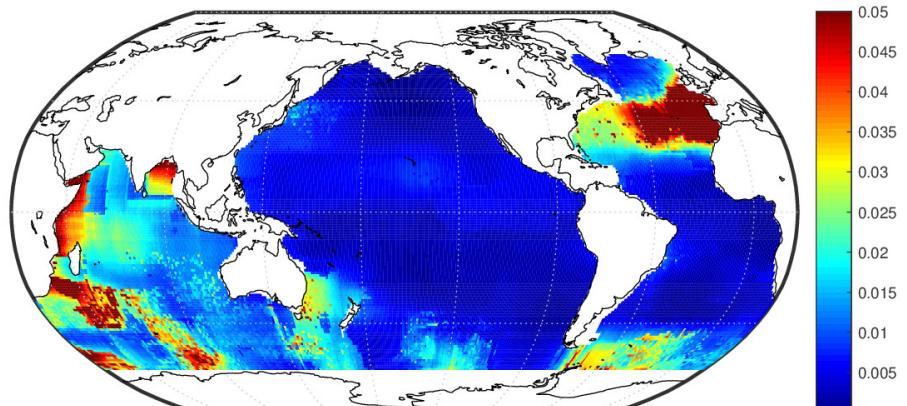
Figure 41: February 2012 temperature anomalies



(a) 10 db

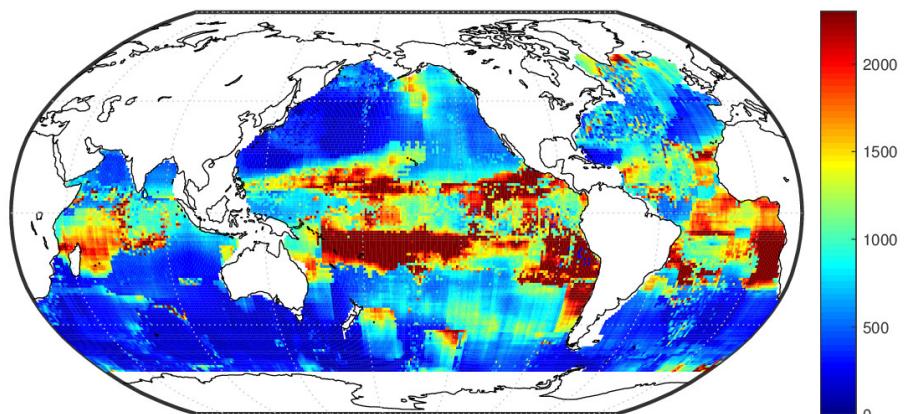


(b) 300 db

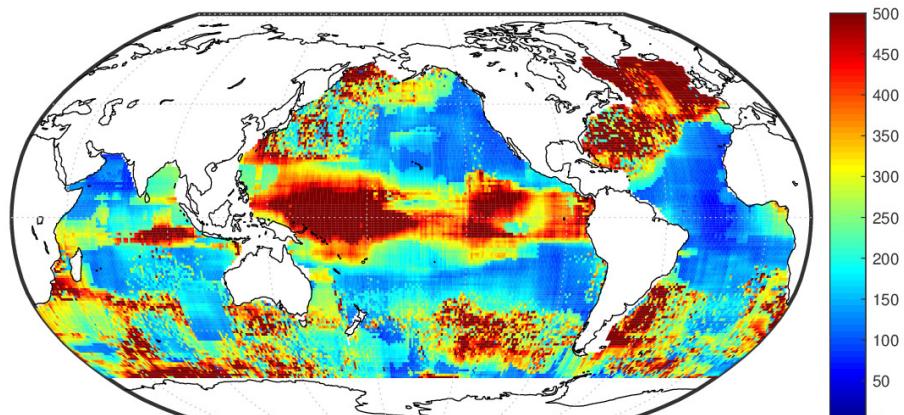


(c) 1500 db

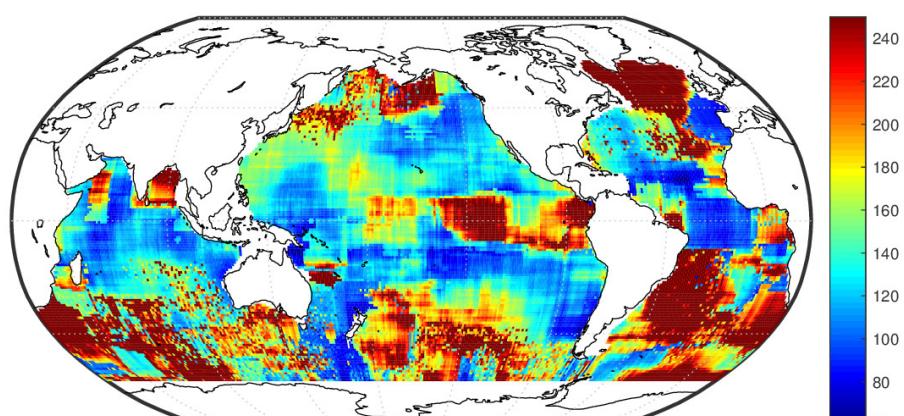
Figure 42: ϕ



(a) 10 db

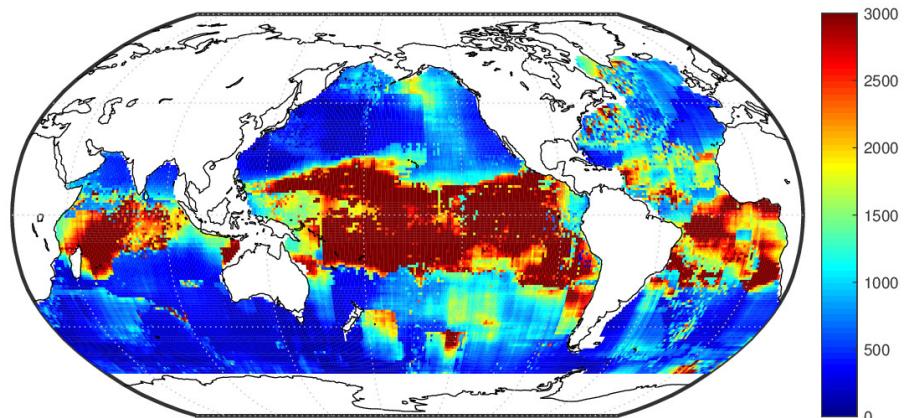


(b) 300 db

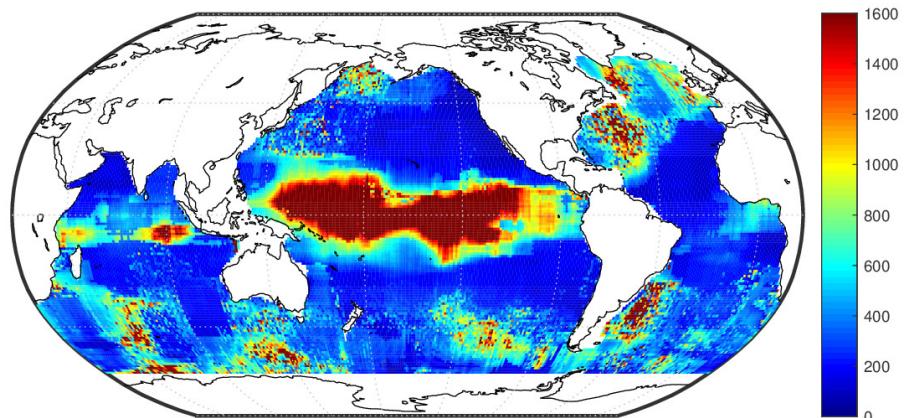


(c) 1500 db

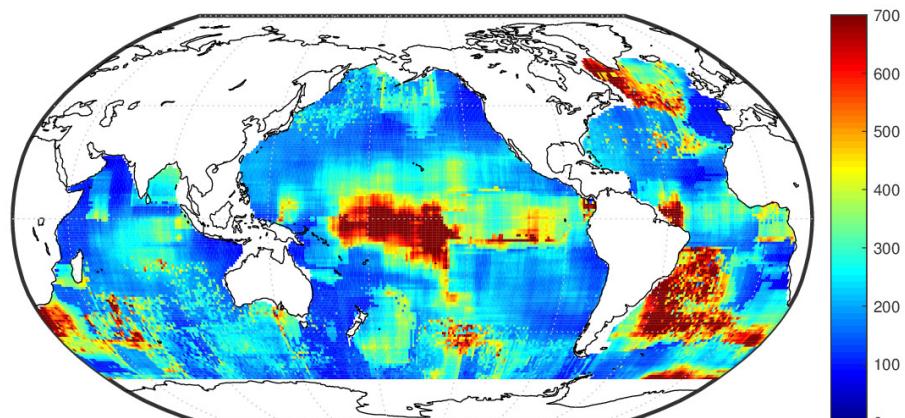
Figure 43: θ_{lat} (in km)



(a) 10 db

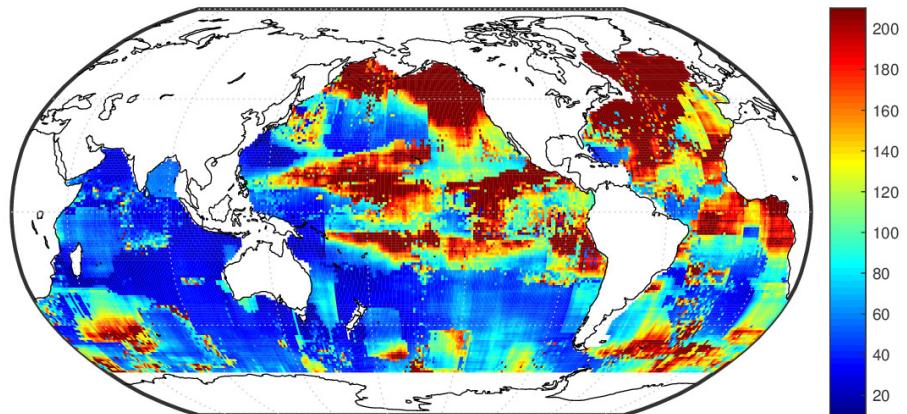


(b) 300 db

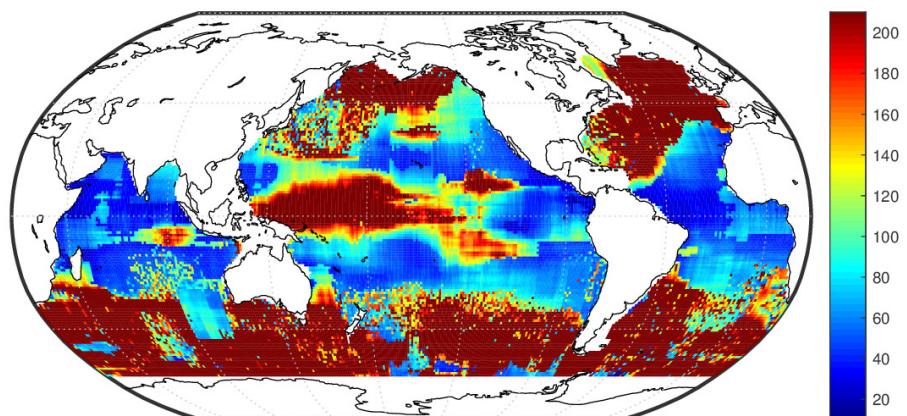


(c) 1500 db

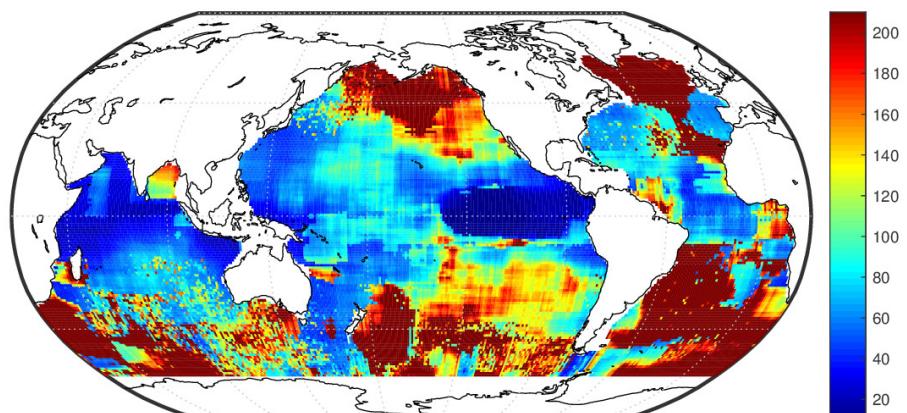
Figure 44: θ_{lon} (in km)



(a) 10 db

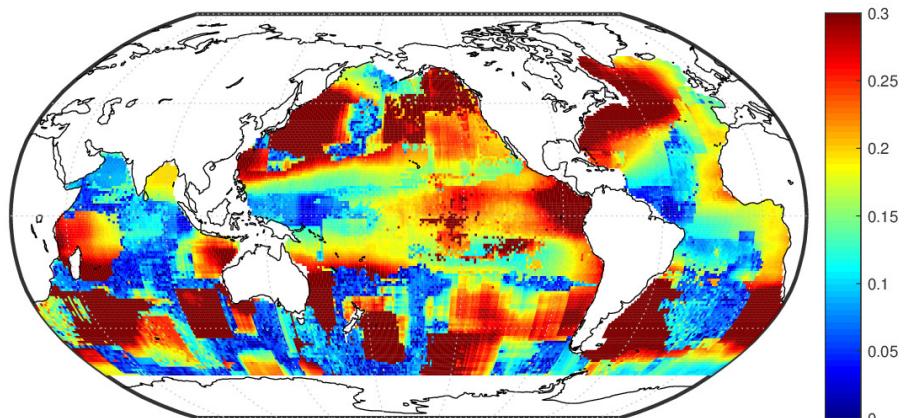


(b) 300 db

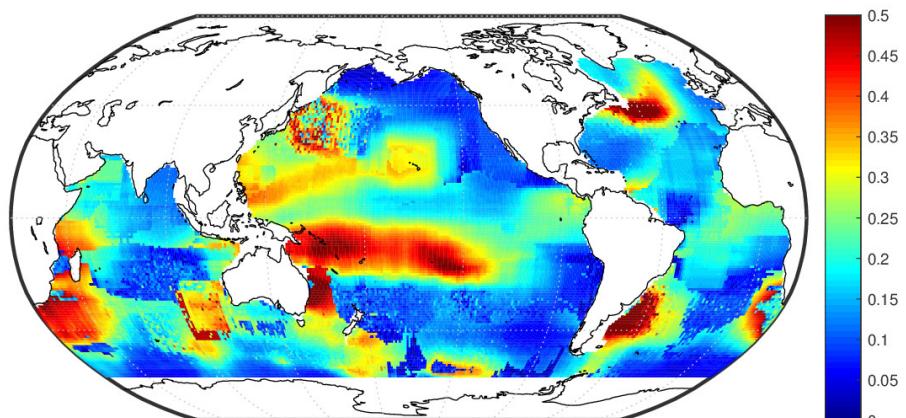


(c) 1500 db

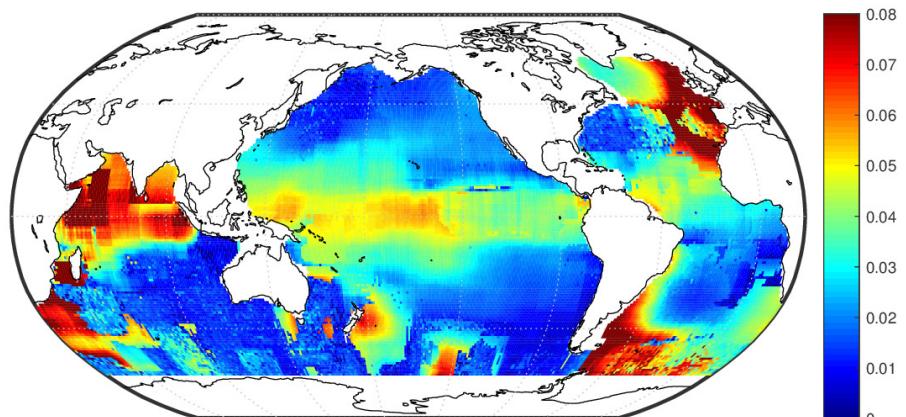
Figure 45: θ_t (in days)



(a) 10 db

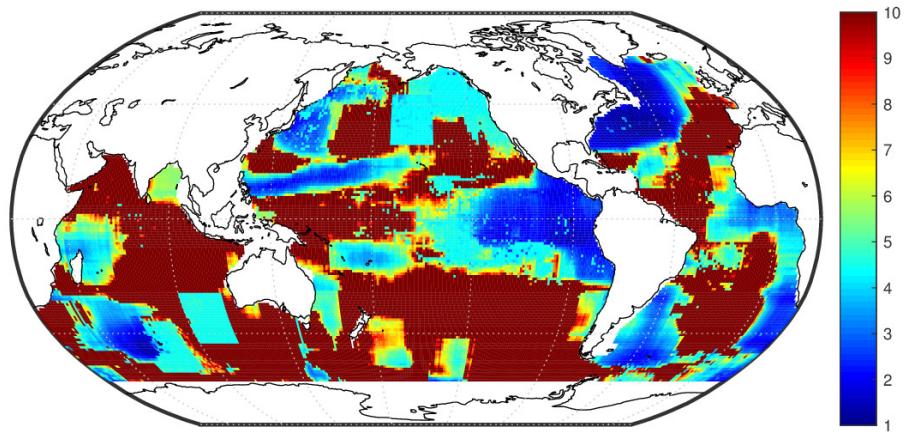


(b) 300 db

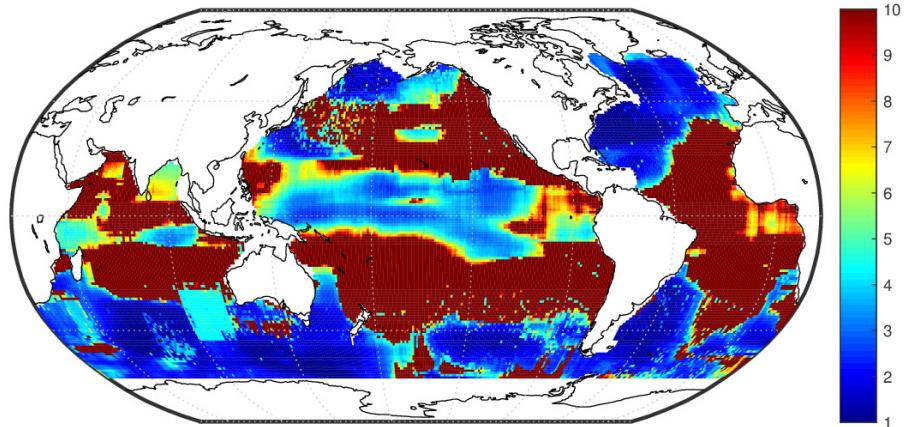


(c) 1500 db

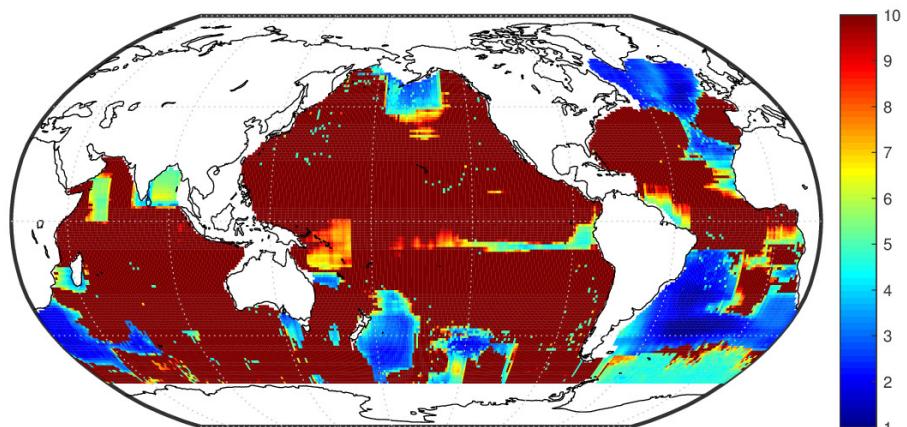
Figure 46: σ



(a) 10 db

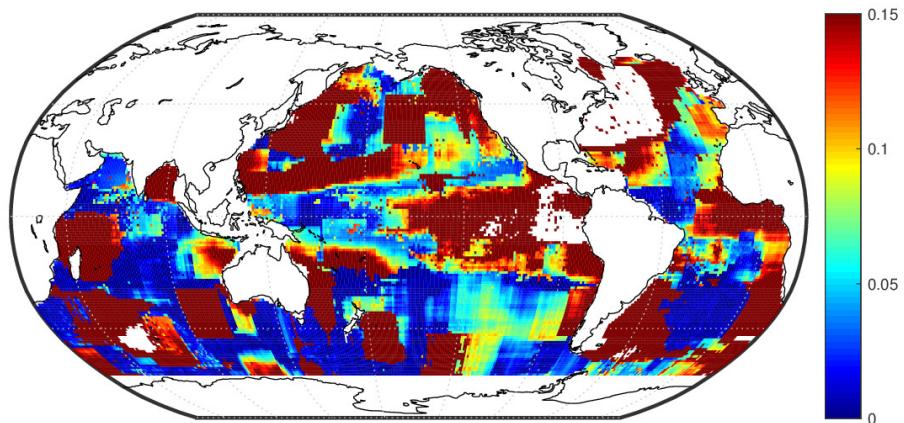


(b) 300 db

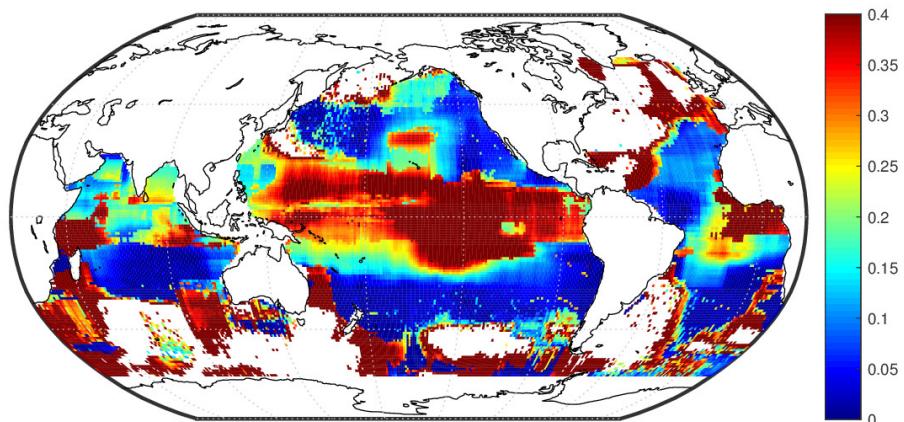


(c) 1500 db

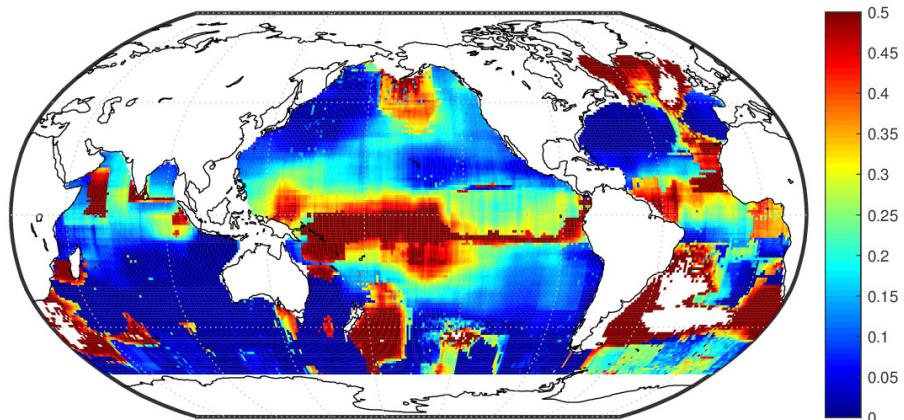
Figure 47: ν



(a) 10 db

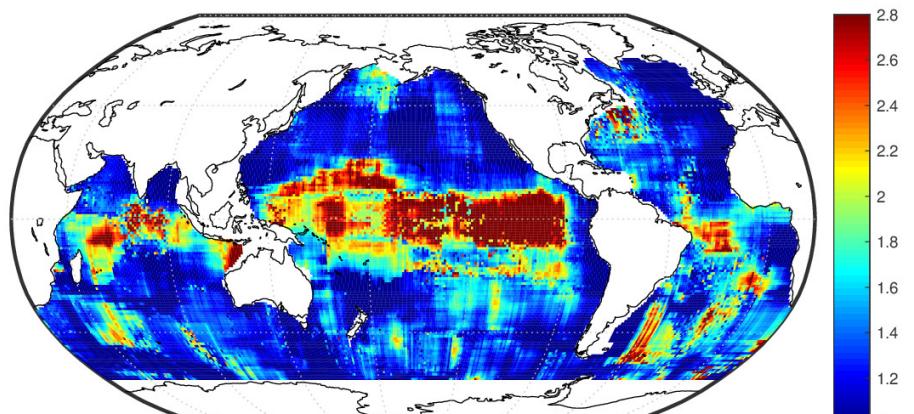


(b) 300 db

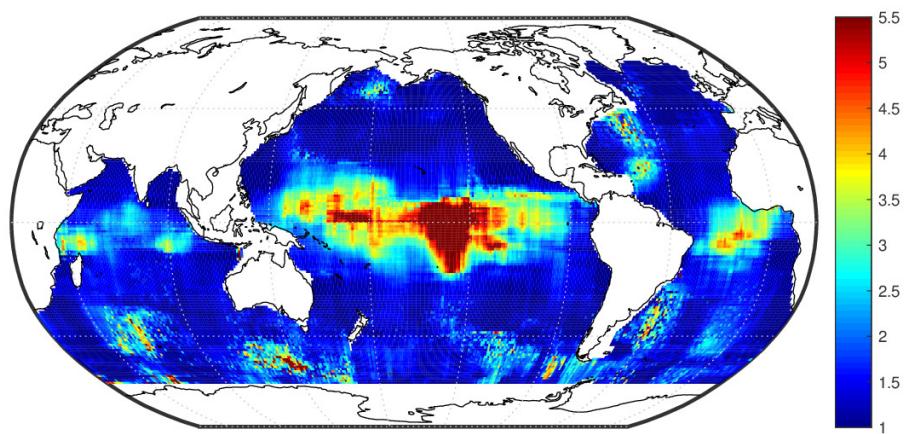


(c) 1500 db

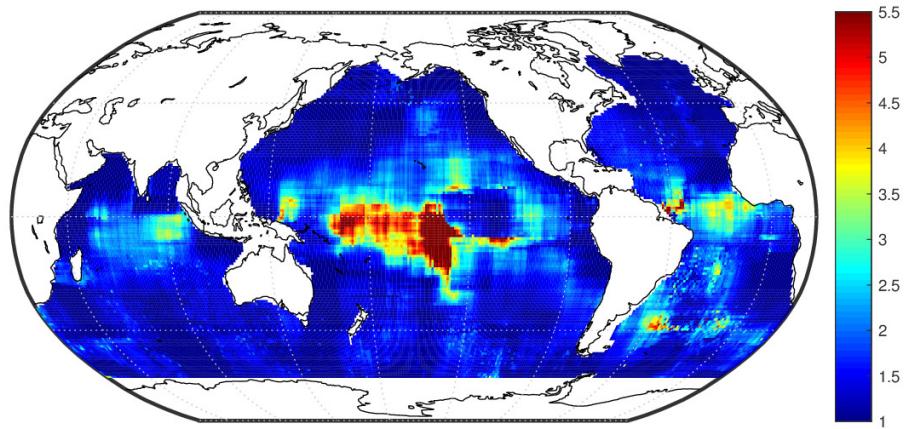
Figure 48: $\frac{\nu}{\nu-2}\sigma^2 / \left(\phi + \frac{\nu}{\nu-2}\sigma^2 \right)$



(a) 10 db

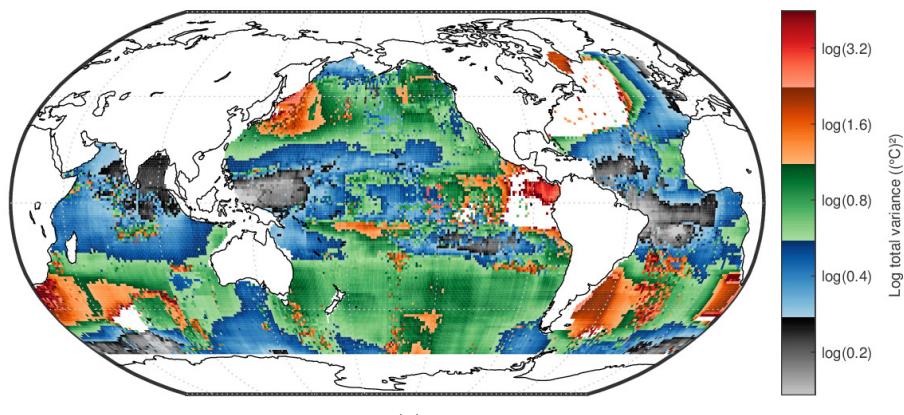


(b) 300 db

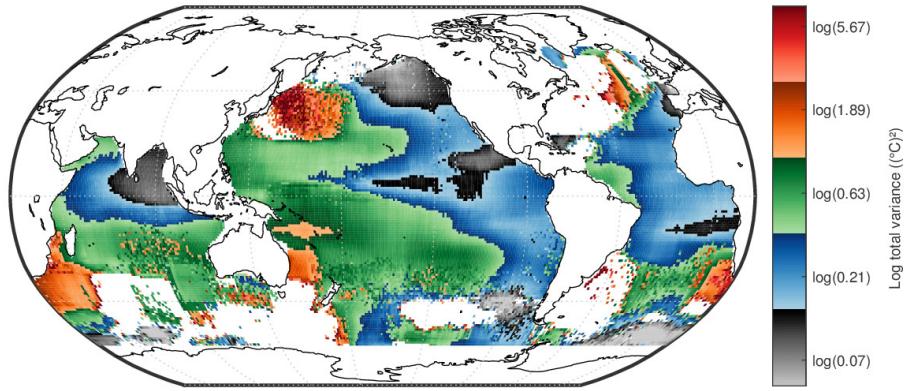


(c) 1500 db

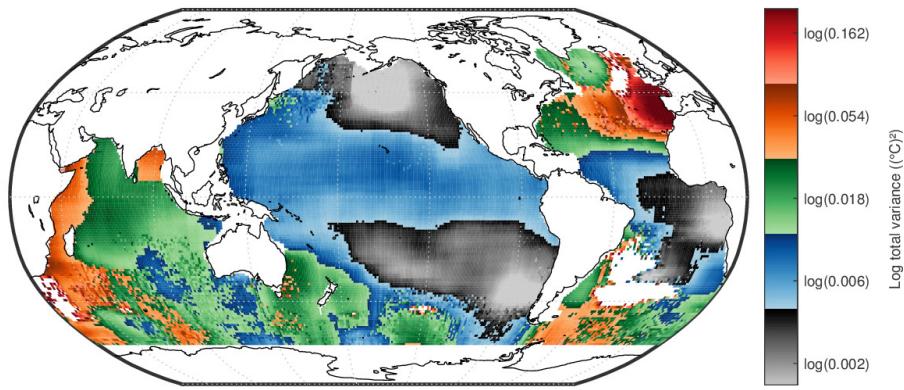
Figure 49: $\theta_{\text{lon}}/\theta_{\text{lat}}$



(a) 10 db

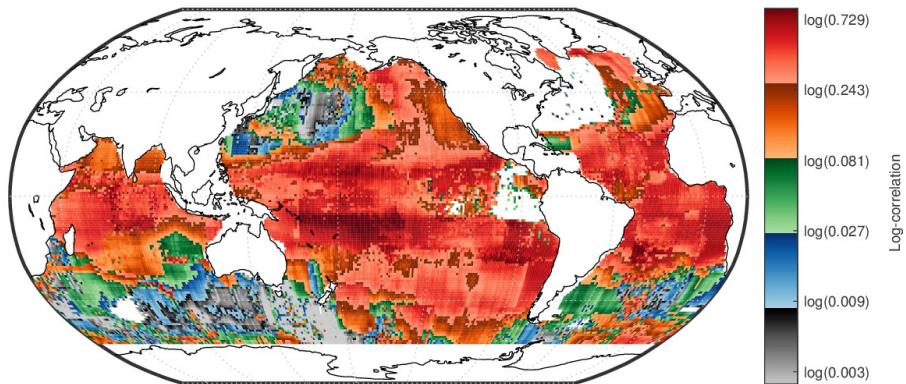


(b) 300 db

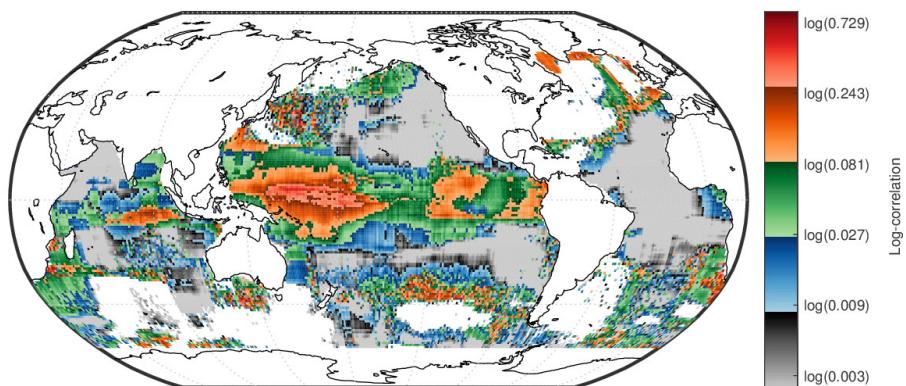


(c) 1500 db

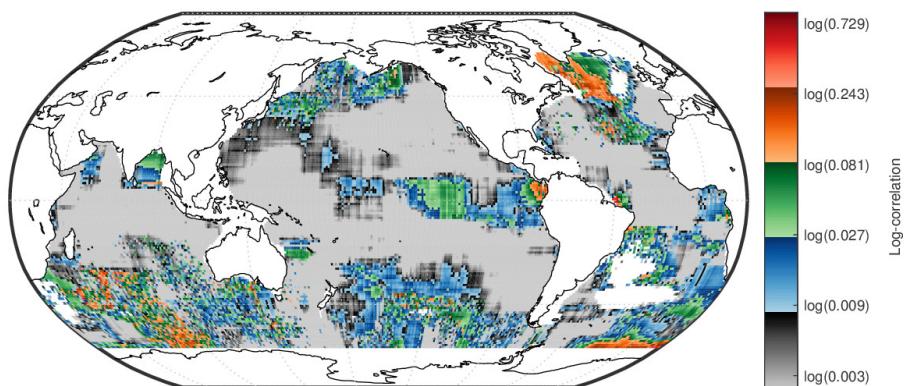
Figure 50: $\phi + \frac{\nu}{\nu-2} \sigma^2$



(a) 10 db

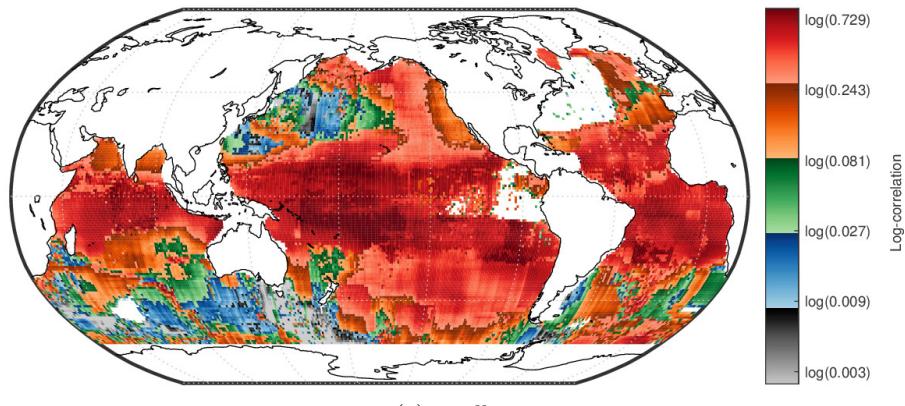


(b) 300 db

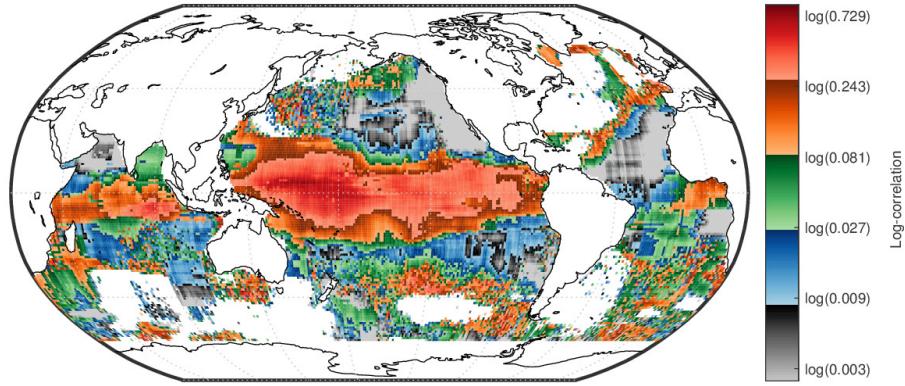


(c) 1500 db

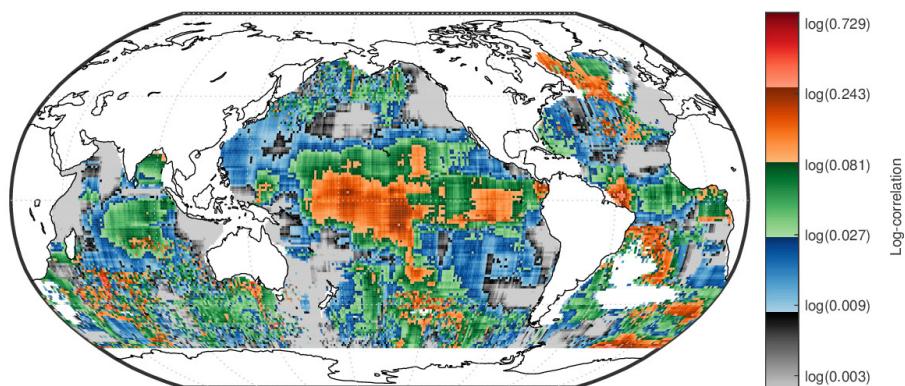
Figure 51: Correlation at $\Delta x_{\text{lat}} = 800 \text{ km}$



(a) 10 db



(b) 300 db



(c) 1500 db

Figure 52: Correlation at $\Delta x_{\text{lon}} = 800 \text{ km}$

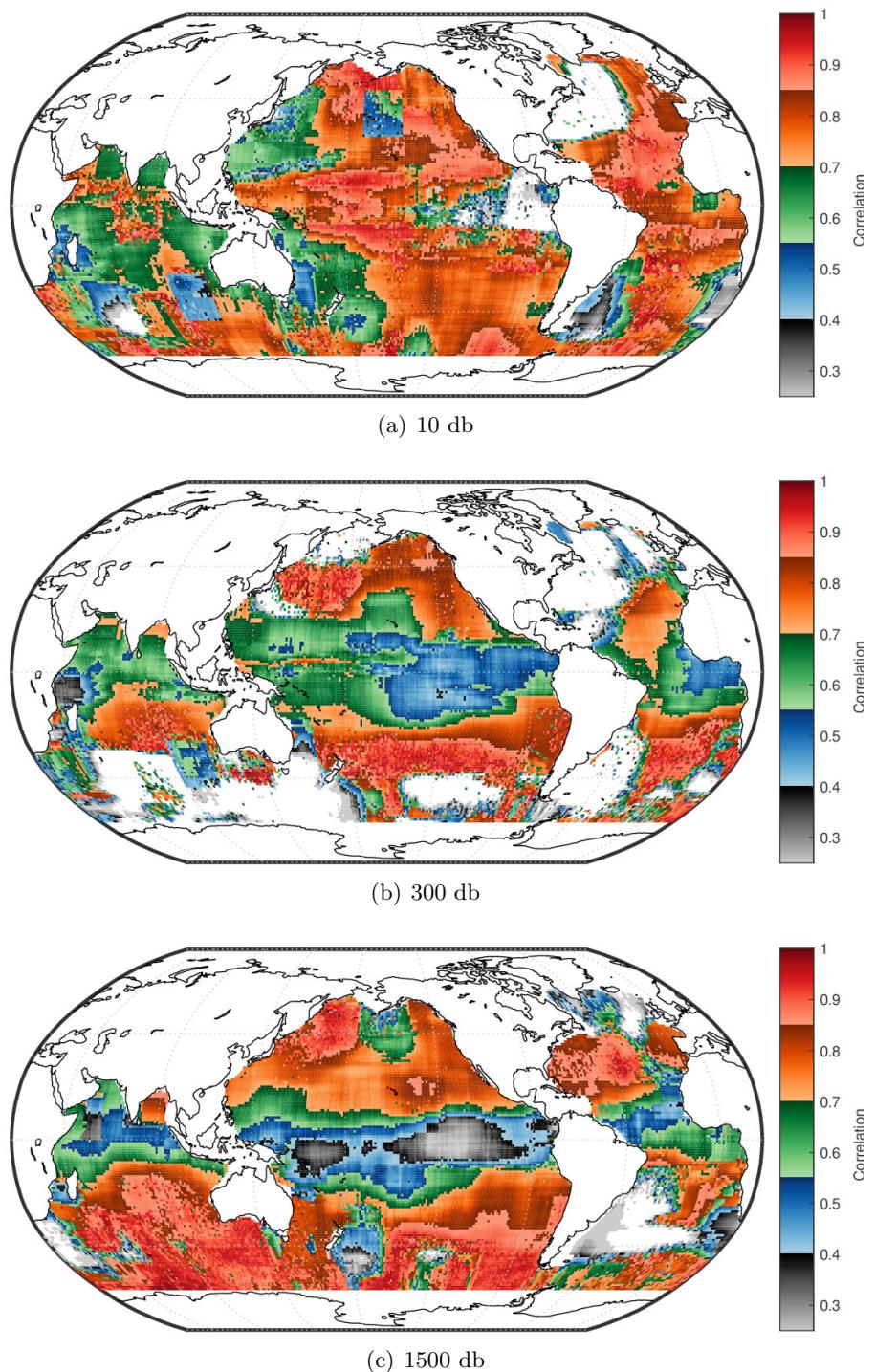


Figure 53: Correlation at $\Delta t = 10$ days

References

- [1] 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>.
- [2] Argo Program. Frequently asked questions about Argo data, retrieved October 14, 2017. http://www.argo.ucsd.edu/Data_FAQ.html.
- [3] Argo Data Management Team. Argo User's Manual V3.2, 2017. <http://doi.org/10.13155/29825>.