JGR Oceans

RESEARCH ARTICLE
10.1029/2018JC014225

Model-Derived Uncertainties in Deep Ocean Temperature Trends Between 1990 and 2010

F. K. Garry[1,2], E. L. McDonagh[3], A. T. Blaker[3], C. D. Roberts[4,5], D. G. Desbruyères[3,6], E. Frajka-Williams[2,3], and B. A. King[3]

[1]Now at Earth System Science, College of Life and Environmental Sciences, University of Exeter, Exeter, UK, [2]Ocean and Earth Science, University of Southampton, National Oceanography Centre, Southampton, UK, [3]National Oceanography Centre, Southampton, UK, [4]Met Office Hadley Centre, Exeter, UK, [5]Now at ECMWF, Reading, UK, [6]Now at Ifremer, University of Brest, CNRS, IRD, Laboratoire dOcéanographie Physique et Spatiale, IUEM, Plouzané, France

Key Points:
- We present a novel framework for estimating the uncertainties in deep ocean heat content trends using models
- Biases due to infrequent sampling in time and space lead to largest errors between 2,000 and 3,000 m
- There are substantial biases in regional change estimates calculated from hydrographic style sampling

Supporting Information:
- Data Set S1

Correspondence to:
F. K. Garry,
f.garry@exeter.ac.uk

Citation:
Garry, F. K., McDonagh, E. L., Blaker, A. T., Roberts, C. D., Desbruyères, D. G., Frajka-Williams, E., & King, B. A. (2019). Model-derived uncertainties in deep ocean temperature trends between 1990 and 2010. *Journal of Geophysical Research: Oceans*, *124*, 1155–1169. https://doi.org/10.1029/2018JC014225

Received 29 MAY 2018
Accepted 11 JAN 2019
Accepted article online 28 JAN 2019
Published online 22 FEB 2019

Abstract We construct a novel framework to investigate the uncertainties and biases associated with estimates of deep ocean temperature change from hydrographic sections and demonstrate this framework in an eddy-permitting ocean model. Biases in estimates from observations arise due to sparse spatial coverage (few sections in a basin), low frequency of occupations (typically 5–10 years apart), mismatches between the time period of interest and span of occupations, and from seasonal biases relating to the practicalities of sampling during certain times of year. Between the years 1990 and 2010, the modeled global abyssal ocean biases are small, although regionally some biases (expressed as a heat flux into the 4,000- to 6,000-m layer) can be up to 0.05 W/m^2. In this model, biases in the heat flux into the deep 2,000- to 4,000-m layer, due to either temporal or spatial sampling uncertainties, are typically much larger and can be over 0.1 W/m^2 across an ocean. Overall, 82% of the warming trend deeper than 2,000 m is captured by hydrographic section-style sampling in the model. At 2,000 m, only half the model global warming trend is obtained from observational-style sampling, with large biases in the Atlantic, Southern, and Indian Oceans. Biases due to different sources of uncertainty can have opposing signs and differ in relative importance both regionally and with depth, revealing the importance of reducing temporal and spatial uncertainties in future deep ocean observing design.

Plain Language Summary In recent decades, deep (below 2,000 m) ocean temperature trends have been measured when scientific research vessels repeat the same lines across an ocean basin. Repeats typically happen once or twice a decade, and there are only a few repeated lines across each basin. The sparsity of data in both space and time will result in errors in the multidecadal temperature trends calculated from this data. Here, we use a state-of-the-art ocean model to show how trends calculated from observational-style sampling compare to trends calculated using all model data. For the period 1990–2010, we estimate the error that may exist in observed deep ocean trend estimates. Overall, around 80% of the below 2,000-m warming trend was captured by observational-style sampling in the model, so deep ocean warming in recent decades may have been underestimated. However, our results are based on only one model simulation. The largest sources of sampling error are found in the Atlantic, Southern, and Indian Oceans. For each basin, we reveal whether limited sampling in time or space contributes most error to the temperature trend estimate, and therefore in which regions temperature trend estimates would benefit from additional deep ocean sampling.

1. Introduction

The deep (2,000–4,000 m) and abyssal (4,000–6,000 m) oceans are vast potential reservoirs for heat in the climate system, but they are sampled more sparsely in time and space than the ocean shallower than 2,000 m. On climate-relevant time scales, the oceans are by far the dominant sink for heat accumulating in the climate system (von Schuckmann et al., 2016), with the majority (> 93%) absorbed into the oceans (Rhein et al., 2013) and 15–20% sequestered deeper than 2,000 m (Gleckler et al., 2016; Talley et al., 2016). Monitoring of ocean heat content (OHC) change is required to properly characterize, quantify, and therefore predict the future of the climate system, with deep OHC playing an important role in decadal variability

(Palmer et al., 2011). Rapid climate change in response to sustained anthropogenic greenhouse gas emissions may lead to severe societal consequences, with an increase in OHC contributing to sea level rise (e.g., Church et al., 2013) and surface heat fluxes leading to extreme weather events (e.g., Grist et al., 2015). Knowledge of deep ocean temperature patterns and change is key to addressing open questions about the pathways of heat through the oceans and how they change over time (e.g., Xie, 2016).

Below 2,000 m, the primary in situ temperature observing system used for generating estimates of global and regional OHC trends is repeated hydrographic sections. This temperature observing system uniquely delivers the better than 0.005 °C accuracy required for OHC trend estimates. Over the past 30 years a global network of repeated hydrographic sections has been developed, coordinated by the Global Ocean Ship-Based Hydrographic Investigations Program (GO-SHIP), along which data are collected at least once a decade (Talley et al., 2016). Purkey and Johnson (2010) analyzed hydrographic section data from the time period 1990–2010, providing the first assessment of decadal changes in deep ocean temperature and identified statistically significant warming across most deep ocean basins, and particularly in the Southern Ocean. A later global estimate from in situ temperature data (Desbruyères et al., 2016) calculated heat uptake deeper than 2,000 m of 0.065 ± 0.040 W/m^2 over 1991–2010. Similar warming rates were observed globally during the 1990s and 2000s, though the spatial pattern of warming changed between decades (Desbruyères et al., 2016). In contrast, a residual method using upper ocean temperature observations and satellite measurements of ocean mass and sea surface height infers a slight cooling in the deep ocean between 2005 and 2012, though uncertainties in this method are large, and the cooling is not statistically significant (Llovel et al., 2014). Regionally, about two thirds of the warming occurs in the Southern Ocean (Desbruyères et al., 2016), with other regions of strong warming including deep western boundary currents (Kouketsu et al., 2011; Sloyan et al., 2013), western boundaries in general (Talley et al., 2016), and the western side of the Mid-Atlantic Ridge (Johnson et al., 2008). All are consistent with a slowdown in the northward flow of abyssal water from Antarctica (Talley et al., 2016). In addition, in some regions interannual variability can be large; in the Drake Passage isopycnal displacement (heave) is largely related to movement of the polar front (Firing et al., 2017).

Temperature change in the upper 2,000 m of the ocean is monitored by a fleet of Argo floats that, in 2007, achieved its target array size of 3,000 floats (approx every 3°) profiling every 10 days. Comparing temperature changes at 2,000 m derived from repeat hydrography and the Argo fleet shows that (over a basin) the mismatch was dominated by temporal rather than spatial differences in sampling (floats sample more frequently than repeat hydrography and are distributed around a basin rather than along individual sections; Desbruyères et al., 2017). Discrepancies due to differences in spatial sampling tended to be smaller than uncertainties on the hydrographic section estimate for a basin. The global average mismatch is small, but regionally, the differences can be large. The largest spatial and temporal mismatches were found in the Southern Ocean, where trends are large and temperature change is dominated by heave (rearrangement of isopycnals). Since heave-related variability decreases with ocean depth, Desbruyères et al. (2017) hypothesize that sampling-related uncertainty will also decrease down through the deep and abyssal oceans.

Hydrographic sections provide high accuracy temperature observations with high along-section resolution. The GO-SHIP reference sections are designed so that each section is repeated at least once a decade, and each deep ocean basin contains at least one section. Hydrographic sections therefore provide limited spatial coverage of the ocean and low temporal frequency data. In addition, they may also have seasonal biases to their occupations (e.g., sampling biased to summer in regions where sea ice forms). Here we describe a methodology that allows us to assess the biases and uncertainty in deep and abyssal temperature change estimates of GO-SHIP repeat hydrography within a model framework. We devise a method to subsample numerical ocean model output (available at all points in space and time, limited by model resolution) in a way that allows attribution of biases to the spatial distribution and temporal frequency of the observations. In the remainder of this paper the following specific questions are addressed: (a) Do estimates of deep OHC change from hydrographic sections during 1990–2010 represent global and regional trends accurately? (b) What are the reasons for inaccuracy and do they relate to temporal or spatial biases in the sampling? (c) What magnitude do these biases have in a model? (d) Do temporal and spatial biases vary regionally? (e) What implications might the results have for future deep ocean observing?

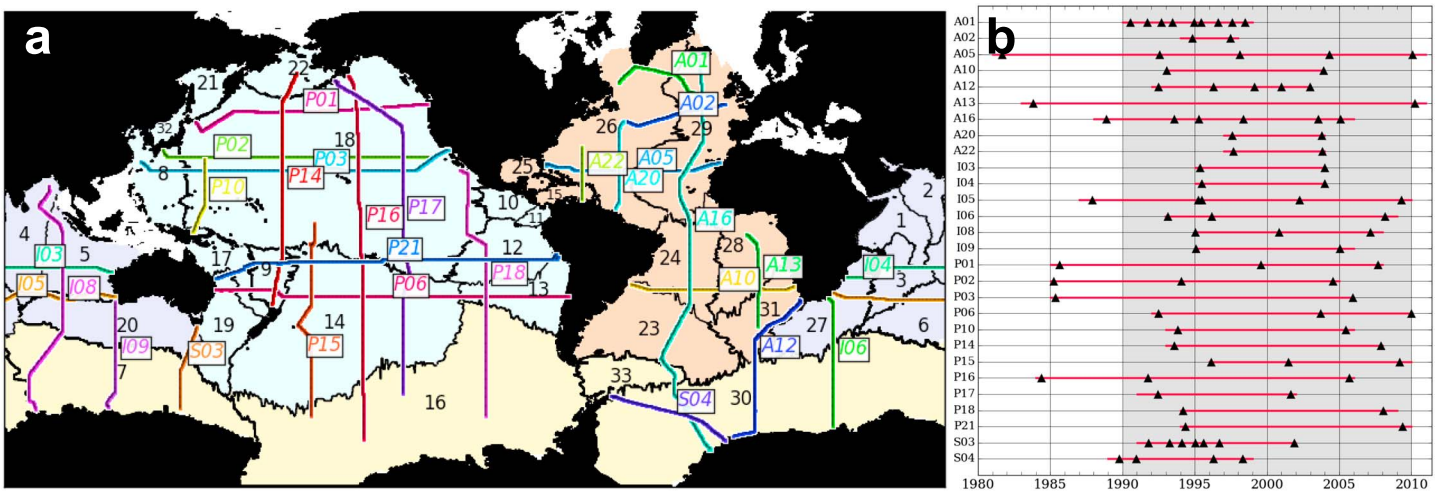

Figure 1. (a) Map of deep ocean basins and hydrographic sections occupied at least twice during the period 1980–2010 (following Purkey & Johnson, 2010), with the Atlantic, Pacific, Southern, and Indian Oceans defined using different colors. (b) Timings of hydrographic sections occupation (triangles), with red lines covering the time period for the "frequent sections" and the gray shading the time interval over which "section truth" and "basin truth" are extracted from model data (1990–2010 inclusive).

2. Data and Methods

2.1. Model

We use output from a simulation of the ocean model NEMO ORCA025 (N025_DFS_01), which has 1/4° horizontal resolution and 75 vertical levels. The model version is NEMO v3.6 (Madec, 2008) with sea ice represented by the Louvain-la-Neuve Ice Model version 2 (Timmermann et al., 2005). The simulation used the CORE bulk formulae and was surface forced with the Drakkar Forcing Set version 5.2, which prescribes surface air temperature, winds, humidity, surface radiative heat fluxes, and precipitation (Brodeau et al., 2010; Dussin et al., 2014). The model was initialized from rest in 1958 using climatological initial conditions for temperature and salinity taken in January from PHC2.1 (Steele et al., 2001) at high latitudes, MEDATLAS (Jourdan et al., 1998) in the Mediterranean, and Levitus et al. (2013) elsewhere and was integrated until December 2015. Similar configurations of the model used here have been used widely for research and operational applications (e.g., by the UK Met Office) on time scales from days to centuries. For further detail, we refer the reader to Megann et al. (2014) and Storkey et al. (2018).

As with any model-based study, the conclusions will include the influence of both atmospherically forced change and model drift, especially since only one model simulation is used here. Ocean models develop deep ocean biases partly due to imperfect knowledge of the initial state and surface fluxes, and partly due to deficiencies in the model such as spurious numerical mixing (Megann, 2018), and missing or inadequately parameterized physical processes and the grid resolution, which prevents realistic on-shelf deep water formation in the Southern Ocean (Heuzé et al., 2013, 2015). Deep water formation biases in other NEMO ORCA025 configurations are discussed for the North Atlantic by Katsman et al. (2018) and for the Southern Ocean by Heuzé et al. (2015). Model drift is strongest right at the start of an integration; since we analyze the period 1990–2010, this allows 32 years for the model to adjust, similar to the approach used in Hirschi et al. (2013). Analysis of deep ocean temperature trends will be affected more strongly by model drift than studies of the upper ocean or short-term variability. We present our framework in detail using a single high-resolution hindcast simulation. The data presented in this paper are available in the supporting information.

2.2. Hydrographic Sections

To replicate the observational analysis of decadal OHC change below 2,000 m in the model, the same repeat hydrographic sections (locations and timings; Figure 1) as Purkey and Johnson (2010) are extracted from the model. We focus our evaluation on the time period 1990–2010, but in following the observational methodology, occupations between 1980 and 1990 are also included and used to inform temperature changes. In the model we extract the full extent of each section during every occupation; in reality some sections are incomplete due to constraints on ship availability or delays (e.g., due to equipment failure or adverse weather). We

extract every model grid box on a section simultaneously at the midpoint in time of a ship's occupation. In reality, occupations of a section typically take weeks to months.

The oceans below 2,000 m are subdivided into basins (Figure 1) following Purkey and Johnson (2010) and Desbruyères et al. (2016). Each cruise is split into sections that are entirely contained in one basin. We analyze potential temperature (henceforth "temperature") from the model output that represent a 5-day mean.

2.3. Calculating Trends on Sections

Model output temperatures along a section are subsampled in time (as illustrated in Figure 1b) to generate the following:

- *Section truth.* Five-day mean model temperature between 1990 and 2010 (inclusive).
- *Frequent sections.* Five-day mean model temperature between the beginning of the first year in which a section was occupied and the end of the last year in which a section was occupied.
- *Pseudo sections.* Model temperatures only at the midpoint in time of each occupation—representative of the temporal frequency of observations.

The following analysis is repeated for the section truth, frequent sections, and pseudo sections model output. Following the analysis of Purkey and Johnson (2010), for each grid box along the section and at each vertical model level, a straight line of best fit is applied through temperature data θ (°C) with respect to time t as

$$\theta(t) = \beta t + c + \epsilon(t) \quad , \tag{1}$$

where β is the linear temperature trend, c the y intercept, and $\epsilon(t)$ is an error term derived from the residuals of the linear fit. Following the observational methodology, $\epsilon(t)$ is not used to calculate uncertainties.

For each vertical model level, the average trend on the section, $\hat{\beta}$ is calculated as

$$\hat{\beta} = \frac{\Sigma(w_g\ \beta)}{\Sigma(w_g)} \quad , \tag{2}$$

where w_g are weightings derived from the along-section distances between the grid box midpoints. Since there are three types of section data, the following three average trends are calculated for each section at each vertical level: $\hat{\beta}_{\text{SECTION TRUTH}}$, $\hat{\beta}_{\text{FREQUENT}}$, and $\hat{\beta}_{\text{PSEUDO}}$.

At each vertical model level we calculate the standard error, SE_S, of the average trend on a section for the pseudo section data to present an uncertainty calculation, following the method of Purkey and Johnson (2010)

$$\text{SE}_\text{S} = \frac{\sigma_\text{S}}{\sqrt{\text{DOF}_\text{S}}} \tag{3}$$

where the standard deviation of trends along the section, σ_S, is

$$\sigma_\text{S} = \sqrt{\frac{\sum w_g\left(\beta - \hat{\beta}\right)^2}{\sum w_g}} \quad , \tag{4}$$

and the effective degrees of freedom (DOF) for a section are calculated as

$$\text{DOF}_\text{S} = \frac{L}{163} \quad . \tag{5}$$

L_S (km) is the length of the section, and 163 (km) is the mean horizontal decorrelation length scale derived from observational data in Purkey and Johnson (2010). This decorrelation length scale is an estimate of how far apart two temperature measurements must be before they are uncorrelated. For simplicity, we choose to use the observation-derived length scale rather than recomputing one using the model output. Following the observational methodology, when sampled regions become separated at depth by seafloor topography longer than one decorrelation length scale, the data are assumed to be independent and contribute one DOF to the estimate.

Assuming a Student's t distribution, the two-sided 95% confidence intervals are estimated, with the endpoints calculated as

$$\hat{\beta} \pm \mathrm{SE_S} \;\; St(0.975, \mathrm{DOF_S}). \tag{6}$$

where St is a function which gives the desired quantile (here 0.975) from the probability density function for a Student's t continuous random variable using the DOF for the section (von Storch & Zwiers, 2001).

2.4. Basin Trends Estimated From Sections and Associated Errors

Following the observational methodology of Purkey and Johnson (2010), the section data are used to estimate the temperature trend for a specific basin, $\hat{\mathbf{B}}$. As there are three types of model data for each section, the following analysis is repeated to obtain three basin estimates: $\hat{\mathbf{B}}_{\mathrm{SECTION\ TRUTH}}$, $\hat{\mathbf{B}}_{\mathrm{FREQUENT}}$, and $\hat{\mathbf{B}}_{\mathrm{PSEUDO}}$.

At each vertical model level, the trend, $\hat{\beta}$, for each section contained within the basin is weighted by the total distance along the section, w_s to contribute to each basin estimate:

$$\hat{\mathbf{B}} = \frac{\sum(w_s\ \hat{\beta})}{\sum w_s} \quad , \tag{7}$$

where the SE for each basin is given by

$$\mathrm{SE_B} = \frac{\sigma_\mathrm{B}}{\sqrt{\mathrm{DOF_B}}}, \tag{8}$$

and the standard deviation for the basin estimate is

$$\sigma_\mathrm{B} = \frac{\sum(w_s\ \sigma_\mathrm{S})}{\sum w_s} \quad . \tag{9}$$

The DOF for the basin is given by

$$\mathrm{DOF_B} = \sum\ \mathrm{DOF_S} \quad . \tag{10}$$

Confidence intervals are calculated from the SE in the same way as they were for a section (equation (6)).

To assess the accuracy of the section-based estimates, the basin truth is calculated by averaging the temperatures from every point across each basin to obtain $\hat{\theta}_{\ \mathrm{B}}$. A linear fit is applied to $\hat{\theta}_\mathrm{B}$ through time as

$$\hat{\theta}_\mathrm{B}(t) = \hat{\mathbf{B}}_{\mathrm{BASIN\ TRUTH}}\ t + c + \varepsilon(t) \quad . \tag{11}$$

where the gradient of that fit gives the model truth basin temperature trend, $\hat{\mathbf{B}}_{\mathrm{BASIN\ TRUTH}}$. As before, c is the intercept with the y axis and $\varepsilon(t)$ is an error term derived from the residuals of the linear fit. This basin truth trend and all the basin estimates are weighted by the volume of the basin (at each vertical level) and integrated for each ocean (e.g., Atlantic) and globally.

2.5. Heat Fluxes

For each basin estimate from the model, a heat flux, HF, into the deep (2,000–4,000 m) and abyssal (4,000 – 6,000 m) ocean layers is calculated from the temperature change in the layer. Note that the model does not contain deep ocean trenches, and the seafloor is always shallower than 6,000 m; therefore, reference to 6,000 m is equivalent to measuring to the bottom of the ocean. First the heat gain as a function of depth, $H(z)$ (W/m), is calculated as

$$H(z) = \rho\ \ c_p\ \ \hat{\mathbf{B}}_{\mathrm{BASIN\ TRUTH}}\ \ A \quad , \tag{12}$$

where ρ is seawater density (kg/m^3), c_p the specific heat capacity of seawater (J·kg^{-1}·°C^{-1}), and A is the area of the model layer in that basin (m^2). The quantities used for ρ and c_p are depth-varying estimates from the World Ocean Atlas climatology. OHC trends are presented in this way to facilitate comparison to top of the atmosphere radiation measurements. The downward heat flux, HF (W/m^2), through depth z_1 is calculated as

$$\mathrm{HF} = \frac{1}{A(z_1)} \int_{z_{\max}}^{z_1} H(z)\,\mathrm{d}z \quad . \tag{13}$$

To calculate the heat flux into the deep ocean layer (2,000–4,000 m), the heat flux through the lower bound of the layer is subtracted from the heat flux through the upper bound of the layer.

2.6. Calculating Spatial and Temporal Biases

Comparing the simulated observational sampling and analysis of the model with the model truth, we estimate the total (combined temporal and spatial) bias of the observing system reproduced in that model:

$$\text{Total bias} = \hat{\mathbf{B}}_{\text{PSEUDO}} - \hat{\mathbf{B}}_{\text{BASIN TRUTH}} \ . \tag{14}$$

This total bias is also defined as

$$\text{Total bias} = \text{Spatial bias} + \text{ Temporal bias (frequency of occupation)} + \text{ Temporal bias (extrapolation)} \tag{15}$$

We isolate these terms using the intermediate calculations of $\hat{\mathbf{B}}_{\text{SECTION TRUTH}}$ and $\hat{\mathbf{B}}_{\text{FREQUENT}}$ made in section 2.4.

To estimate the spatial bias, we difference the basin estimate based on the relevant sections extracted from the model every 5 days and the model truth. This represents the additional information given to the temperature change from regions within the basin but away from the section or sections.

$$\text{Spatial bias} = \hat{\mathbf{B}}_{\text{SECTION TRUTH}} - \hat{\mathbf{B}}_{\text{BASIN TRUTH}}. \tag{16}$$

The temporal bias is split into two terms. Temporal bias (frequency of occupation) represents the bias related to the frequency of sampling between the first and last hydrographic sections, while temporal bias (extrapolation) represents the effect of extrapolation of the trend to cover the 1990–2010 period. To further elaborate, hydrographic sections are usually sampled at a frequency of approximately 5–10 years so trends calculated from this low temporal resolution data may not be a good representation of the trend over the time period between the first and last occupations. This uncertainty is quantified by subtracting the temperature trend or heat flux in a basin calculated from the pseudo sections in a model (same temporal coverage as observations) from that calculated from the frequent data (5-day mean model output taken between the beginning of the first year of occupation and the end of the last year of occupation):

$$\text{Temporal bias (frequency of occupation)} = \hat{\mathbf{B}}_{\text{PSEUDO}} - \hat{\mathbf{B}}_{\text{FREQUENT}}. \tag{17}$$

Note that the pseudo and frequent sections do not span exactly the same period of time because the frequent section data goes from the beginning of the first year of occupation to the end of the last year of occupation. Choosing whole years for the frequent data type also removes any uncertainty due to seasonal sampling bias that might be influencing the pseudo section trend. Therefore, the section truth and frequent section trends will not have any seasonal bias, though trends from pseudo sections might. A different methodology for evaluating the effect of potential seasonal biases is outlined in the next section.

Very few hydrographic sections have occupations spanning the full time period 1990–2010. Therefore, when comparing basin estimates from sections to the model basin truth, making the assumption that basin estimates from frequent sections are representative of 1990–2010 effectively extrapolates the calculated trend from the time period over which hydrographic sections occurred. The trend could also be calculated over a longer period by including part of the 1980s, as some sections use occupations earlier than 1990, and that error is also considered part of this uncertainty. We refer to this as temporal bias due to extrapolation, but it simply includes any error due to the trend being calculated over a different time period to 1990–2010. This uncertainty can be quantified by a bias calculated by taking the difference between frequent section data estimates (using 5-day mean data on a section between the first and last year of occupation) and the section truth data estimates (5-day mean data on a section during the whole time period, 1990–2010):

$$\text{Temporal bias (extrapolation)} = \hat{\mathbf{B}}_{\text{FREQUENT}} - \hat{\mathbf{B}}_{\text{SECTION TRUTH}}. \tag{18}$$

2.7. Sensitivity to Pseudo Section Timing

The timing of hydrographic sections is dictated by a number of constraints around the nominal sampling frequencies intended by observational scientists, such as ship time and availability, weather, and the availability of principal investigators, so sampling may be biased toward a particular season, particularly in polar regions. The sensitivity to the particular timings of section occupations is tested by analyzing the differences in basin estimates obtained by extracting sections from the model at different times randomly selected within 1 year of the actual occupation, a range by which the timing of a particular cruise might plausibly be

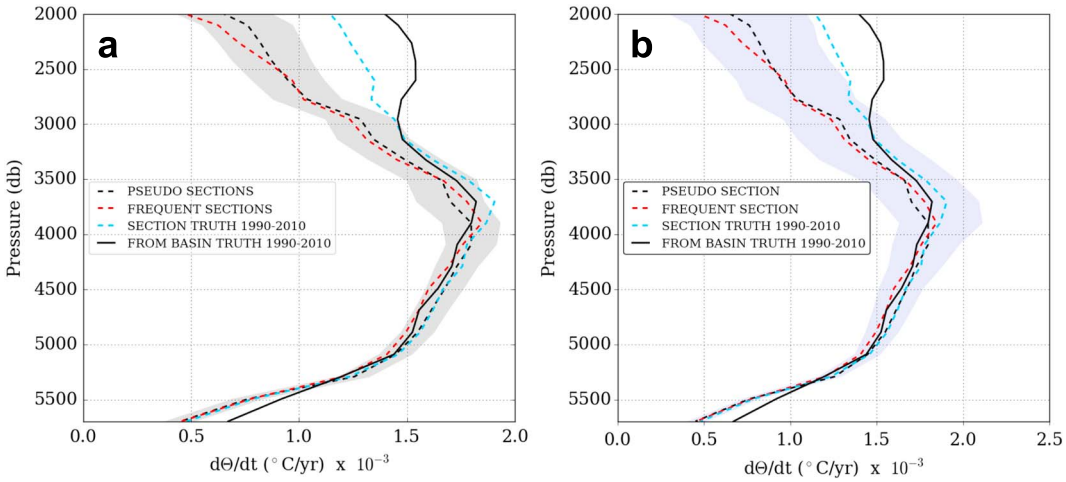

Figure 2. For the global oceans, horizontally averaged depth profiles of $d\Theta/dt$ over 1990–2010. In (a) gray shading indicates the uncertainty associated with averaging temperature trends horizontally along the sections (95% confidence interval). In (b) light blue shading indicates ±2 standard deviations from the mean of the trends from 1,000 timing test pseudo sections (dates shifted randomly between ±1 year).

moved. This random selection is repeated many times, giving new sets of timings for each section that have broadly the same time intervals as the original dates, and these are used in the trend analysis in the same way as the original timings. We generate an estimate of this uncertainty globally and for each ocean using 1,000 repeats between ± 1 year, but in addition for the Southern Ocean we present the following additional timing tests: 500 repeats between ± 2 months (sampling in only the same season), 500 repeats between −7 and −5 months (sampling in the opposite season in the preceding year), and 500 repeats between +5 and +7 months (sampling in the opposite season in the following year).

The uncertainty related to the pseudo section timings is quantified by calculating two standard deviations of the timing tests. Where the temperature trends are presented against depth, the uncertainty is presented as two standard deviations of the trends from the timing tests around the mean value of the timing tests. When the uncertainty is presented as a heat flux, the heat flux is calculated for all the timing tests and then the uncertainty reflects two standard deviations of those data.

3. Evaluating Global and Basin Temperature Trends in a Model

Globally averaged trends derived from sections all underestimate the true warming trend in the model between 2,000 and 3,000 m (Figure 2), with hydrographic subsampling failing to capture about a half of the warming signal in the model over recent decades at 2,000 m. Approximately two thirds of the total bias comes from extrapolation (section truth − frequent section), while one third is attributable to spatial bias (basin truth − section truth). Temporal bias due to time intervals between occupations is relatively unimportant globally. The magnitude of the bias decreases with depth, with around a third of the warming signal missed at 2,500 m. Between 3,000 and 5,300 m hydrographic subsampling captures the true temperature trend in the model well. Underestimation of the global trend deeper than 5,300 m arises from a spatial bias, though the total volume of the ocean basins below this depth is relatively small.

The volume-weighted average temperature trend below 2,000 m is computed using both pseudo section model output and the model truth (Figure 2, black dashed and solid black lines) to be 1.28 and 1.56 m·°C·year^{-1}, respectively. The total bias in the pseudo sections estimate is therefore −0.28 m·°C·year^{-1}, which is equivalent to pseudo sections underestimating the trend by 18%.

Observation-based analyses compute an estimate of uncertainty that is associated with averaging the section trends along sections (Desbruyères et al., 2016; Purkey & Johnson, 2010). We compute this for our pseudo sections (Figure 2a, gray shading). The uncertainty interval encompasses the frequent section estimate but not the section truth or the basin truth shallower than 3,500 m, highlighting the importance of other uncertainties (spatial, temporal due to extrapolation) not considered by observational analyses. Another measure of uncertainty, that due to sampling on particular dates (sensitivity of the trends to the actual cruise

Table 1
Uncertainties in Measuring Global and Oceanic Trends

Uncertainty	Global	Atlantic	Southern	Indian	Pacific
			10^{-3} W/m^2		
2,000–4,000 m					
Total bias	−92.5	−226.0	−143.7	−21.6	−49.5
Spatial bias	−25.5	−99.4	+26.4	−37.9	−8.8
Temporal bias (extrapolation)	−74.5	−144.2	−117.0	−106.4	−20.0
Temporal bias (frequency of occupation)	+7.5	+17.7	−53.2	+122.7	−20.7
Uncertainties when pseudo sections differ in timings by ±1 year	±62.1	±191.5	±180.8	±191.9	±32.7
Uncertainties due to averaging along sections	±38.3	±154.7	±157.1	±108.8	±26.3
4,000–6,000 m					
Total bias	+4.0	+10.7	−28.6	83.6	−9.8
Spatial bias	+4.3	−0.3	+21.8	+56.8	−7.5
Temporal bias (extrapolation)	−5.5	−9.6	−12.6	−1.6	−3.6
Temporal bias (frequency of occupation)	+5.2	+20.6	−37.7	+25.2	+1.3
Uncertainties when pseudo sections differ in timings by ±1 year	±12.3	±42.2	±35.0	±34.5	±4.7
Uncertainties due to averaging along sections	±9.6	±35.3	±40.1	±29.4	±6.2

timings), can be computed using the model output (Figure 2b, blue shading), and this is generally larger than the uncertainty associated with averaging the pseudo section trends along sections (Figure 2a, gray shading). However, between 2,000 and 2,700 m it still does not encompass the truth from basins or section truth.

Quantification of biases using the equivalent heat fluxes into the deep and abyssal layers (Table 1) reveals that globally the temporal bias due to extrapolation dominates in the deep layer. In the abyssal layer, the global biases are each approximately the same magnitude, while the temporal bias due to the frequency of occupation almost compensates the bias due to extrapolation, so the trend from pseudo sections is similar to that from the basin truth (Figure 2).

The Atlantic basin trends in the model are not well captured by section estimates, with substantial biases shallower than 3,500 m and deeper than 5,300 m (Figure 3a). The separation of each of the lines between 2,000 and 3,500 m in Figure 3a emphasizes the role of spatial and both types of temporal uncertainty, and relatively high values for the biases in heat flux into the deep ocean layer exist (Table 1): −0.14 W/m^2 for the temporal bias due to extrapolation and −0.1 W/m^2 for the spatial bias. However, in the abyssal ocean, the greatest Atlantic bias (+0.02 W/m^2) is due to the frequency of occupation; this is twice the magnitude of the other temporal bias (due to extrapolation), while the spatial bias is negligible in the abyssal Atlantic Ocean (Table 1). The ocean volumes (Figure 3) indicate how important trends at specific depths are to the layer averages in Table 1, since trends are weighted by the volume of ocean at each depth.

The Southern Ocean (Figure 3b) trends in the model are underestimated by pseudo sections due to temporal biases down to 4,500 m (section truth agrees well with the basin truth over the same depth range). The temporal bias in heat flux can also be attributed primarily to extrapolation (0.12 W/m^2, Table 1), with a smaller contribution due to the frequency of occupation (0.05 W/m^2), which might include some bias due to seasonal sampling. The spatial sampling bias for the Southern Ocean is much smaller in magnitude than the temporal uncertainties between 2,000 and 4,000 m (Table 1), suggesting that increased spatial coverage would not be as useful as more frequent sampling to capture the decadal trends accurately between 2,000 and 4,000 m. As in the Atlantic, bias due to extrapolation is much less important in the abyssal Southern

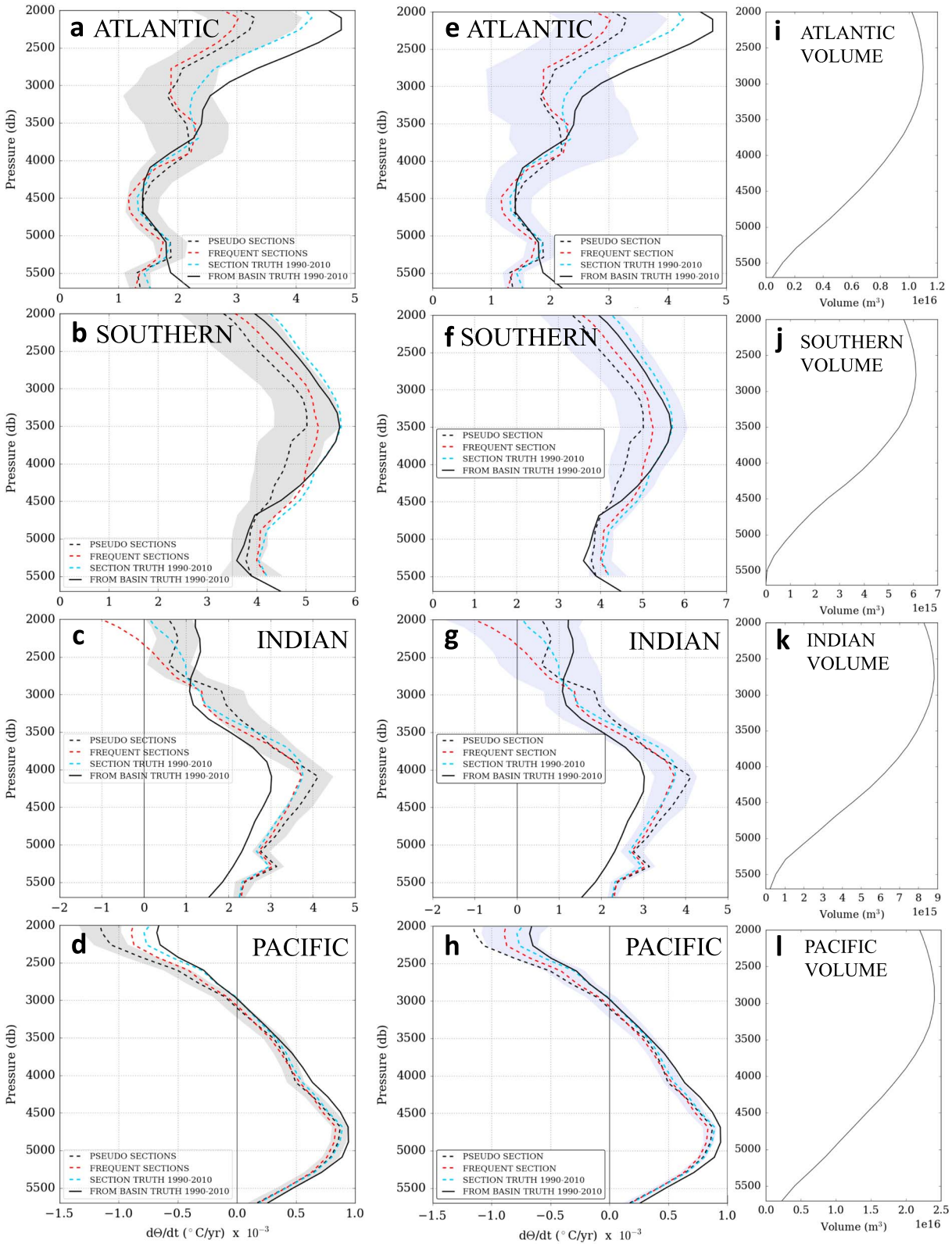

Figure 3. For each major ocean basin, horizontally averaged depth profiles of dΘ/dt over 1990–2010. In (a)–(d) gray shading indicates the uncertainty associated with averaging temperature trends horizontally along the sections (95% confidence interval). In (e)–(h) light blue shading indicates ±2 standard deviations from the mean of the trends from 1,000 timing test pseudo sections (dates shifted randomly between ±1~year). Panels (i)–(l) show the ocean basin volumes in the model.

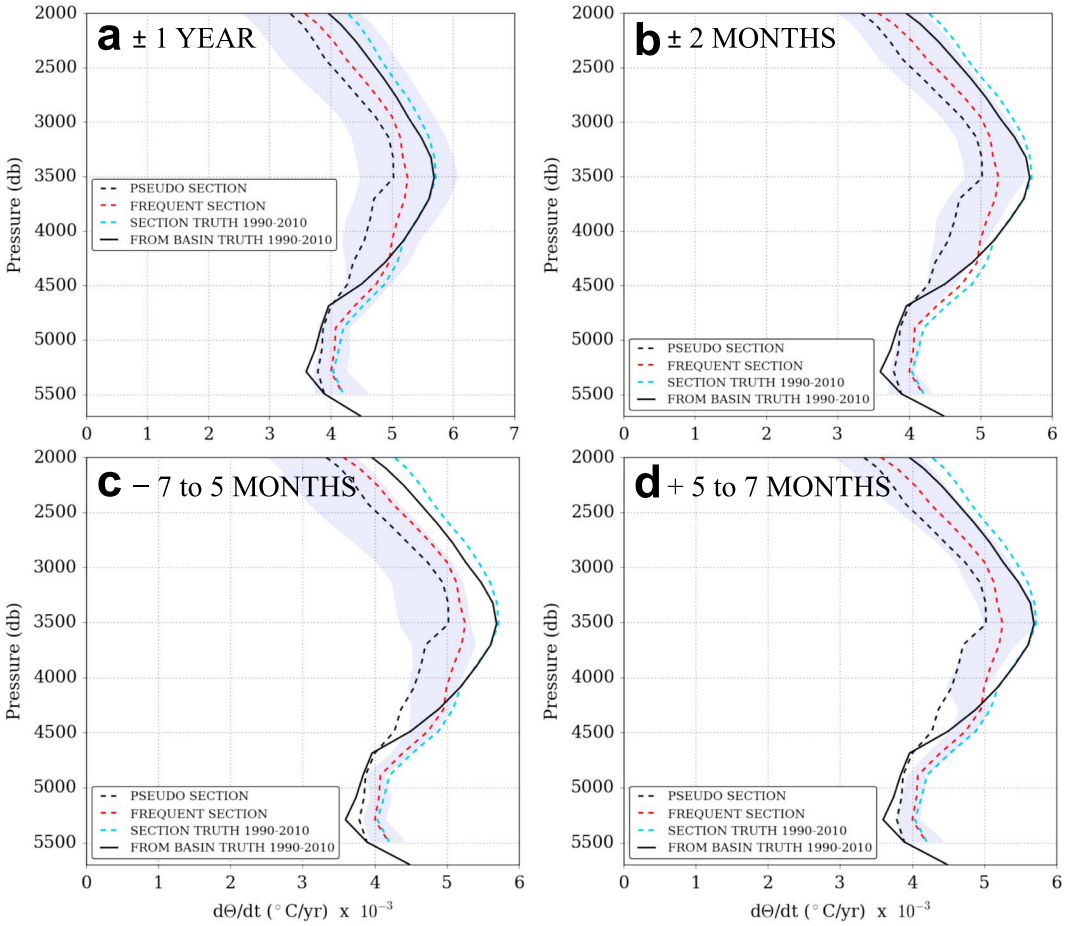

Figure 4. Horizontally averaged depth profiles of dΘ/d*t* in the Southern Ocean with shading indicating ±2 standard deviations of repeated pseudo sections, calculated using multiple random dates in timing windows (a: ±1 year; b: ±2 months; c and d: 5–7 months before and after the original pseudo section timings).

Ocean than for the deep layer, but together the temporal biases are more than twice as large as the spatial bias, with the spatial bias partly compensating the temporal biases (Table 1).

In the Indian Ocean (Figure 3c) spatial biases are substantial shallower than 3,000 m and deeper than 4,000 m. Between 2,000 and 4,000 m the temporal biases are much larger than the spatial bias (Table 1), but the negative spatial bias shallower than 3,000 m is compensated by a positive spatial bias deeper than 3,000 m (Figure 3c). In the 4,000- to 6,000-m layer the spatial bias is double the temporal bias due to the frequency of occupation. Bias due to extrapolation in this depth range is relatively unimportant.

Trends in the Pacific Ocean (Figure 3d) are generally well captured by section-based estimates but with a small systematic underestimate due to spatial sampling, and some contribution from temporal biases between 2,000 and 3,000 m.

The analysis of global trends by ocean (Figure 3) shows that biases arise largely due to a combination of spatial biases in the Atlantic and Indian Oceans and temporal biases in the Atlantic, Southern, and Indian Oceans. The gray shading in Figures 3a–3d reflects the uncertainty due to averaging pseudo section trends along sections and is broadest in the Southern and Atlantic Oceans. Figures 3e–3h show the same trend data, but the blue shading represents the uncertainty associated with sensitivity to specific pseudo section dates. This timing uncertainty is often larger than the along-section averaging uncertainty, particularly for the Atlantic Ocean and between 2,000 and 3,000 m in the Indian Ocean, indicating that the observational uncertainty derived from variability along sections is not always sufficiently broad. However, neither uncertainty encompasses the basin truth in the 2,000- to 3,000-m Atlantic and below 4,000-m Indian Oceans, highlighting the critical importance of spatial bias in these regions (Figure 3).

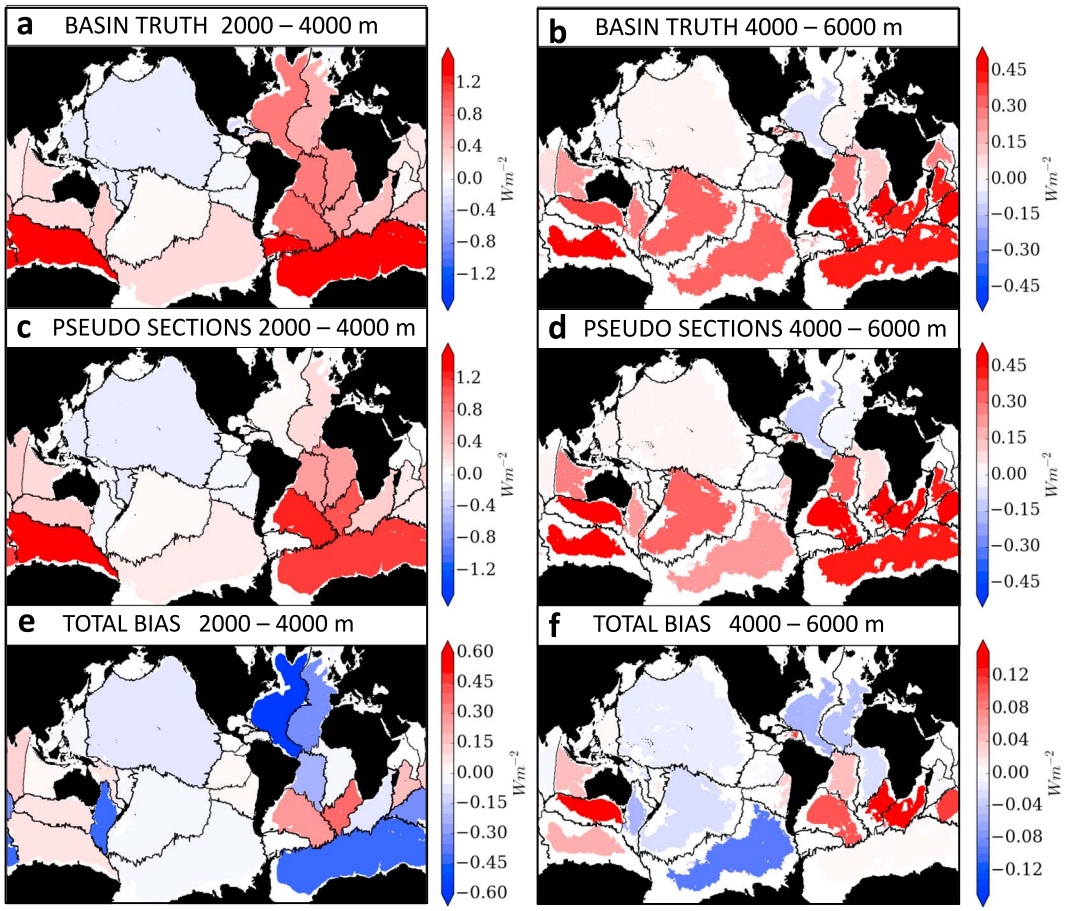

Figure 5. Local heat fluxes in each basin into the (a, c, and e) 2000 to 4000 m (b, d, and f) 4000 to 6000 m layers over the period 1990–2010. As labeled with (e) = (c) − (a) and (f) = (d) − (b).

To test the importance of sampling at different times of year in the Southern Ocean, we adjust the length of the interval in which we randomly shift the dates of pseudo sections; the Southern Ocean is a region where hydrographic sampling usually takes place in the summer months to avoid adverse weather conditions and sea ice during winter. The range of uncertainty that is generated when sampling at any time of year (Figure 4a) is larger than when sampling in the same season (Figure 4b). The trend from any adjusted timing pseudo section is likely to be an underestimate of the section and basin truth, because the shading is biased negative compared to the section truth and basin truth lines, and sampling in the same season means that observations are less likely to yield a representative trend than sampling throughout the year.

We explore how the uncertainty interval changes if the section occupations are shifted to the winter months. The dates of occupation are shifted backward in time (Figure 4c) and forward in time (Figure 4d). Shallower than 4,000 m, shifting the occupations 5–7 months back in time decreases the likelihood of capturing the true trend. Shifting the observations forward in time 5–7 months has little effect on the bias shallower than 4,000 m. The best representation of the true trend would be obtained when measuring all year round (Figure 4a). Shifting the occupations in time improves the temporal bias deeper than 4,000 m (section truth is more likely to lie within the blue shading). The trend calculated using the actual dates of occupations (Figure 4, black dashed line) lies toward the lower end of the distribution (blue shading). This means that the temporal bias in the abyssal ocean due to the frequency of occupation, which is calculated using the pseudo section observations (Table 1), is large by chance.

4. Uncertainties by Ocean Basin

We provide a regional breakdown of the biases and uncertainties that exist when sampling the deep and abyssal ocean, by expressing the trends calculated for each basin as heat fluxes into the deep and abyssal

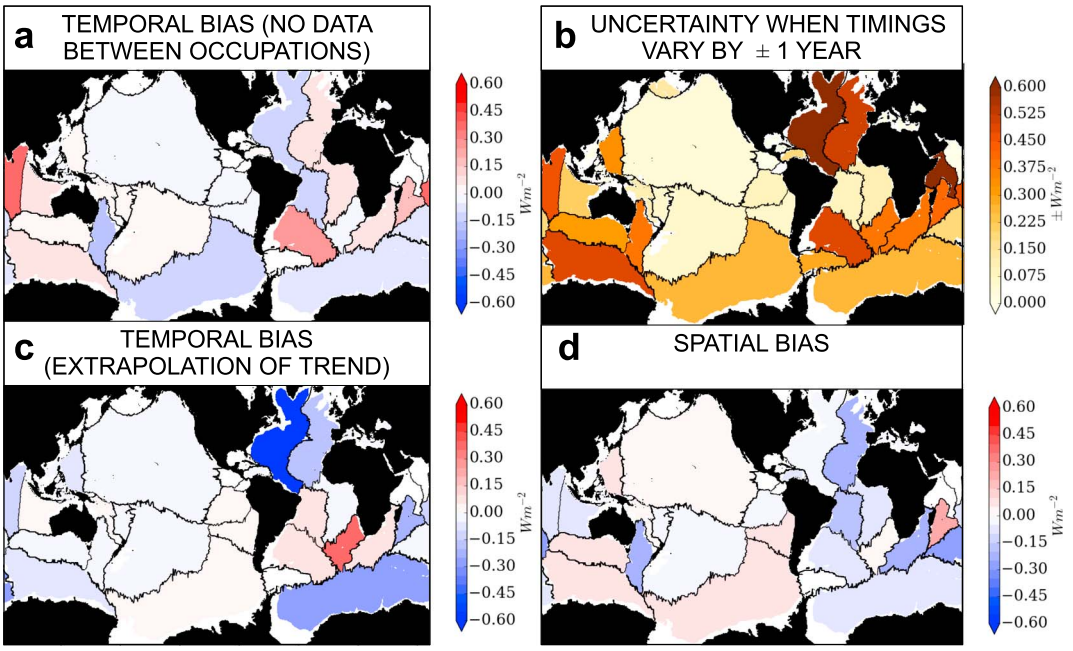

Figure 6. Biases in the heat fluxes into the deep ocean (2,000–4,000 m): (a) temporal bias due to frequency of occupation, (b) sensitivity to pseudo section timings, (c) temporal bias due to extrapolation, and (d) spatial bias.

ocean layers (Figure 5). The main features of the model ocean between 2,000 and 4,000 m are warming through the Atlantic, Indian, and Southern Oceans with strongest Southern Ocean warming in the Indian and Atlantic sectors, while the northern Pacific Ocean cools very slightly. The total bias is given by the difference between the fluxes derived using pseudo sections and the basin truth. The pseudo section estimates broadly match the pattern of the basin truth (Figures 5a and 5c), but panel (e) show that the pseudo sections underestimate the trend across the majority of the oceans. Basin fluxes across large areas of the Atlantic Ocean and Southern Ocean sector of the Atlantic are underestimated by 0.3–0.6 W/m^2.

In the abyssal oceans (4,000–6,000 m; Figures 5b and 5d) pseudo sections and basin truth both show warming across the Southern Hemisphere, with the Northern Hemisphere basins showing a much smaller warming trend, or cooling in some basins. The differences (Figure 5f) reveal that pseudo sections in this layer may either overestimate or underestimate the basin truth. The overestimates span the subpolar Indian and Atlantic basins, but total error for abyssal ocean basins is generally less than 0.12 W/m^2. The exceptions are basins 20 (southwest of Australia), 27 (southeast of Africa), and 31 (southwest of Africa), where the overestimates are higher, at 0.20, 0.22, and 0.20 W/m^2, respectively (Figure 5f).

The total bias is decomposed into the components representing (a) bias due to the frequency of occupation, (b) sensitivity to precise timings of occupations, (c) extrapolation bias due to observations not spanning 1990–2010, and (d) spatial bias (Figure 6). Between 2,000 and 4,000 m, biases due to the frequency of occupation (Figure 6a) are positive through the East Atlantic and Indian Oceans (trends overestimated by pseudo sections). The western Atlantic and much of the Southern Ocean show negative biases (trends underestimated by pseudo sections). This may include seasonal sampling bias or be due to higher frequency variability not captured by decadal sampling. However, Figure 6b indicates that uncertainties due to the timing of the pseudo sections can be relatively large and this uncertainty might result in changes to the biases in Figure 6a of at least the same magnitude as the original value. Therefore, the pattern of trend biases seen in Figure 6a may be largely a function of the precise timings of the pseudo sections.

Extrapolating trends from sections to assume representativity for the period 1990–2010 is relatively unimportant across the Pacific basins (Figure 6c) but can lead to substantial biases in some of Atlantic and Indian basins (including their Southern Ocean sectors). The largest bias due to extrapolation is in the northwest Atlantic. Insufficient spatial coverage leads to an underestimation of warming through the Atlantic and some Indian Ocean basins (Figure 6d), but a small overestimation in the Pacific and the Indian and Pacific sectors of the Southern Ocean. Some basins have bias contributions of opposing signs (e.g., the Pacific sec-

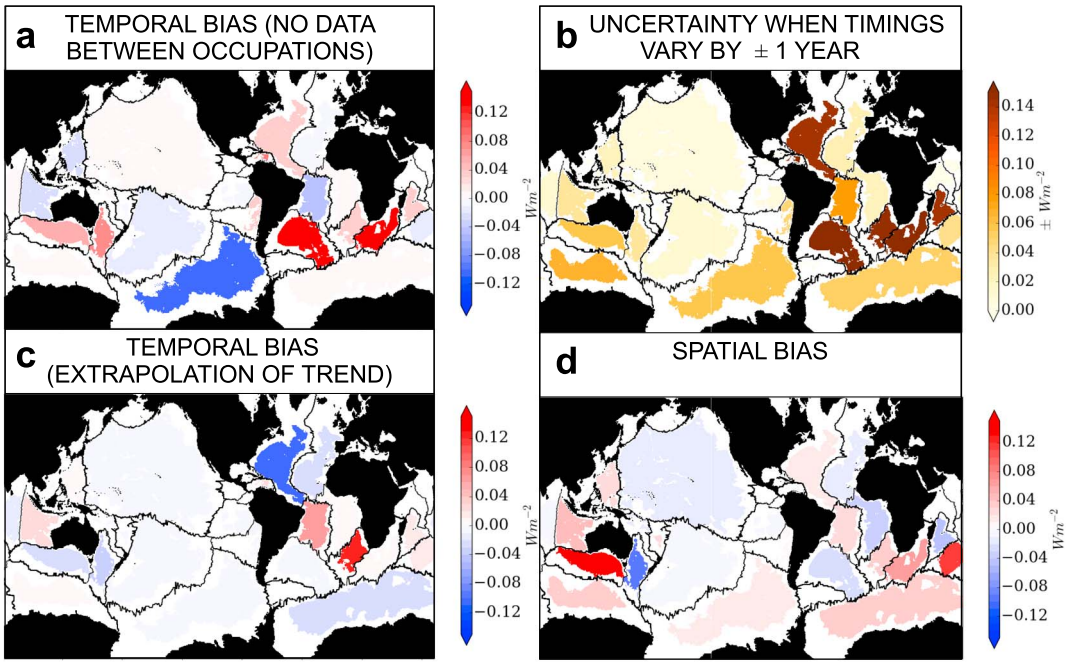

Figure 7. Biases in the heat fluxes into the abyssal ocean (4,000–6,000 m): (a) temporal bias due to frequency of occupation, (b) sensitivity to pseudo section timings, (c) temporal bias due to extrapolation, and (d) spatial bias.

tor of the Southern Ocean). The existence of compensating biases indicates that reducing either a temporal or spatial bias alone may not always improve the estimate in a basin.

Biases in heat fluxes into the abyssal layer do not exhibit spatially coherent patterns across oceans (Figure 7). For each bias there are some basins where the biases are comparatively strong compared to the rest of the global ocean. Uncertainty in the timing of pseudo sections (Figure 7b) is particularly large in several basins across the Atlantic and Indian Oceans, with moderate biases also in the Southern Ocean. As with 2,000–4,000 m, the uncertainty due to pseudo section timings between 4,000 and 6,000 m is of a similar order of magnitude to the calculated biases, in particular the temporal biases due to extrapolation and the frequency of occupations (Figure 7a), which could be significantly affected by relatively minor adjustments to the dates of the pseudo sections.

5. Comparison of Model Temperature Trends With Observational Estimates

Both our simulation and observation-derived temperature trends for the global ocean are positive (increasing temperature) at all depths deeper than 2,000 m. Model biases and uncertainty related to the hydrographic style subsampling are largest between 2,000 and 3,000 m, with about half the trend captured. The uncertainty in the observation-derived temperature trends is also largest between 2,000 and 3,000 m, but unlike the model the observed trend is not statistically different from zero.

At 2,000 m, the model truth yields a warming rate that is double that obtained from the pseudo sections. In contrast, between 2006 and 2014 the temperature trend at 2,000 m computed from Argo profiles was approximately half that computed from repeat hydrography (Desbruyères et al., 2017; Figure 2a). This difference was attributed to heave (vertical motion) of isopycnals (Desbruyères et al., 2017; Figure 6a), indicating that temperature change at 2,000 m is strongly influenced by dynamic rearrangement and that these changes occur on time scales that include interannual to multidecadal. Since the structure and timing of modeled dynamic adjustments is likely to differ from those in the real ocean, validation of modeled deep ocean temperature trends between 2,000 and 3,000 m with observational estimates may not be meaningful.

Deeper than 3,000 m, the global pseudo section estimate agrees well with the basin truth (Figure 2) and we expect the observation based estimates to be a good validation of the global temperature trend. We find a global temperature change that exhibits a similar structure with a maxima at about 4,500 m in the observations (Purkey & Johnson, 2010, Figure 9c) and at about 3,700 m in the model (Figure 2). The most striking

difference between these two estimates is that everywhere deeper than 2,000 m in the model is warming faster than observations (at least 0.5 m·°·C·year^{-1} but up to 1 m·°·C·year^{-1}). This accelerated warming is consistent with numerical diapycnal mixing in a z-level model that acts to diffuse heat into the deep and abyssal ocean unrealistically quickly (Megann, 2018).

Regionally, the Atlantic contributes a substantial difference between pseudo sections and the model truth between 2,000 and 3,000 m (Figure 3a). This is not unexpected given the large heave-related signal in the Atlantic observed by Desbruyères et al. (2017, their Figure 6b). The mismatch in the deep ocean between Argo and repeat hydrography is dominated by heave and is largest in the North Atlantic, Southern, and Indian Oceans (Desbruyères et al., 2017). The modeled trends in the Southern Ocean are around double those found in observations. However, as with the global trend, the vertical structure of the Southern Ocean trend deeper than 2,000 m in the model (Figure 3b) is similar to the trend deeper than 3,000 m in observations (Purkey & Johnson, 2010, Figure 9d), suggesting that Antarctic Bottom Water is too shallow in the model. Deeper than 3,000 m the basin truth temperature changes are systematically higher than pseudo sections in the Pacific Ocean (Figures 3d and 3h), and systematically low in the Indian Ocean (Figures 3c and 3g). In both cases these systematic and opposing biases arise due mainly to spatial sampling and this may reflect more isolated regional patterns of drift in deep and abyssal water volumes.

6. Conclusions

We construct a framework to investigate the uncertainties and biases in estimates of deep ocean temperature change from hydrographic sections due to frequency of occupation and spatial distribution. This framework is applied to a state-of-the-art ocean model hindcast. Overall, 82% of the simulated global temperature trend below 2,000 m is captured from pseudo hydrographic sections, and we find this result encouraging. However, in both the deep and abyssal oceans strong regional spatial or temporal biases can occur and these will adversely impact regional estimates of both heat content change and the thermosteric component of sea level rise. The largest contributors to global biases are biases in the deep Atlantic and Southern Oceans and abyssal Indian Ocean. Temporal biases are much larger than spatial bias in the Southern Ocean. Ship-based sampling of the Southern Ocean is, for practical reasons, seasonal, and this makes the possibility of accurately capturing the true basin trend highly unlikely.

Substantial biases in temperature trends between 2,000 and 3,500 m due to temporal and spatial sampling can occur, with less than 60% of the modeled global warming trend captured shallower than 2,500 m, and these biases are not found to be strongly sensitive to the precise timings of occupations. Modeled trends are well represented by hydrographic sections deeper than 3,500 m. However, details of how the different contributions to biases in temperature trends as a function of depth are unlikely to translate directly to observations due to known deficiencies in model representation of deep ocean water mass formation.

We should be cautious not to overinterpret our results, since they are based on analysis of only one simulation. However, if our priority is to better constrain global OHC estimates to improve estimates of deep and abyssal ocean temperature trends, sea level rise prediction, and our understanding of changes to the planetary energy imbalance, our analysis suggests focusing future new deep ocean sampling resources into higher-resolution sampling in the Atlantic, Southern, and Indian Oceans. This conclusion is broadly consistent with the results of Desbruyères et al. (2017).

Our framework can be used to make informative comparisons of observational estimates and model fields when validating deep and abyssal OHC evolution in models, though isopycnal heave due to dynamical processes can strongly influence the 2,000- to 3,000-m depth range in some basins. The framework could also provide a useful tool for intercomparison of biases between different models, for example, those that will contribute to Coupled Model Intercomparison Project Phase 6 (Eyring et al., 2016) and for the physical component of the Ocean Model Intercomparison Project 6 (Griffies et al., 2016).

Acknowledgments
While conducting this research, F. K. G. received funding from the University of Southampton through a Vice Chancellor's Award, the Marine Physics and Ocean Climate group at the National Oceanography Centre, and the Met Office Hadley Centre. A. T. B., E. L. M., and B. A. K. received NERC National Capability Programme Funding and ORCHESTRA, and E. L. M./B. A. K. were additionally funded through the NERC DEEP-C grant NE/K004387/1, which funded D.D.. A.T.B. and C. R. acknowledge the support of the EU Horizon 2020 research and innovation program under grant agreement 633211 (AtlantOS). We thank all those involved in development and creation of NEMO ocean model runs. The research data supporting this publication are openly available from the University of Exeter's institutional repository (https://doi.org/10.24378/exe.1104). We thank Nadia Pinardi and two anonymous reviewers for their help improving this article.

References

Brodeau, L., Barnier, B., Treguier, A.-M., Penduff, T., & Gulev, S. (2010). An ERA40-based atmospheric forcing for global ocean circulation models. *Ocean Modelling*, *31*, 88–104.

Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., et al. (2013). Sea Level Change. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution*

of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK and New York: Cambridge University Press.

Desbruyères, D., McDonagh, E. L., King, B. A., & Thierry, V. (2017). Global and full-depth ocean temperature trends during the early 21st century from Argo and repeat hydrography. *Journal of Climate*, *30*, 1985–1997.

Desbruyères, D. G., Purkey, S. G., McDonagh, E. L., Johnson, G. C., & King, B. A. (2016). Deep and abyssal ocean warming from 35 years of repeat hydrography. *Geophysical Research Letters*, *43*, 10,356–10,365. https://doi.org/10.1002/2016GL070413

Dussin, R., Barnier, B., & Brodeau, L. (2014). *The making of Drakkar Forcing Set DFS5*, DRAKKAR/MyOcean Rep. 05-10-14. Grenoble, France: LGGE.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental design and organization. *Geoscientific Model Development*, *9*, 1937–1958.

Firing, Y. L., Mc Donagh, E. L., King, B. A., & Desbruyères, D. G. (2017). Deep temperature variability in Drake Passage. *Journal of Geophysical Research: Oceans*, *122*, 713–725. https://doi.org/10.1002/2016JC012452

Gleckler, P. J., Durack, P. J., Stouffer, R. J., Johnson, G. C., & Forest, C. E. (2016). Industrial-era global ocean heat uptake doubles in recent decades. *Nature Climate Change*, *6*, 394–398.

Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Böning, C. W., et al. (2016). OMIP contribution to CMIP6: Experimental and diagnostic protocol for the physical component of the Ocean Model Intercomparison Project. *Geoscientific Model Development*, *9*, 3231–3296.

Grist, J. P., Josey, S. A., Jacobs, Z. L., Marsh, R., Sinha, B., & Van Sebille, E. (2015). Extreme air-sea interaction over the North Atlantic subpolar gyre during the winter of 2013–2014 and its sub-surface legacy. *Climate Dynamics*, *46*, 4027–4045.

Heuzé, C., Heywood, K. J., Stevens, D. P., & Ridley, J. K. (2013). Southern Ocean bottom water characteristics in CMIP5 models. *Geophysical Research Letters*, *40*, 1409–1414. https://doi.org/10.1002/grl.50287

Heuzé, C., Ridley, J. K., Calvert, D., Stevens, D. P., & Heywood, K. J. (2015). Increasing vertical mixing to reduce Southern Ocean deep convection in NEMO3.4. *Geoscientific Model Development*, *8*, 3119–3130.

Hirschi, J. J.-M., Blaker, A. T., Sinha, B., Coward, A., de Cuevas, B., Alderson, S., & Madec, G. (2013). Chaotic variability of the meridional overturning circulation on subannual to interannual timescales. *Ocean Science*, *9*, 805–823.

Johnson, G. C., Purkey, S. G., & Toole, J. M. (2008). Reduced Antarctic meridional overturning circulation reaches the North Atlantic Ocean. *Geophysical Research Letters*, *35*, L22601. https://doi.org/10.1029/2008GL035619

Jourdan, D., Balopoulos, E., Garcia-Fernandez, M., & Maillard, C. (1998). Objective analysis of temperature and salinity historical data set over the Mediterranean basin. In *IEEE Oceanic Engineering Society. OCEANS'98. Conference Proceedings (Cat. 98CH36259)*, *1*, IEEE, Nice, France, pp. 82–87.

Katsman, C. A., Drijfhout, S. S., Dijkstra, H. A., & Spall, M. A. (2018). Sinking of dense North Atlantic waters in a global ocean model: Location and controls. *Journal of Geophysical Research: Oceans*, *123*, 3563–3576. https://doi.org/10.1029/2017JC013329

Kouketsu, S., Kawano, T., Masuda, S., Sugiura, N., Sasaki, Y., Toyoda, T., et al. (2011). Deep ocean heat content changes estimated from observation and reanalysis product and their influence on sea level change. *Journal of Geophysical Research*, *116*, C03012. https://doi.org/10.1029/2010JC006464

Levitus, S., Antonov, J., Baranova, O., Boyer, T., Coleman, C., Garcia, H., et al. (2013). The world ocean database. *Data Science Journal*, *12*, WDS229–WDS234.

Llovel, W., Willis, J., Landerer, F., & Fukumori, I. (2014). Deep-ocean contribution to sea level and energy budget not detectable over the past decade. *Nature Climate Change*, *4*, 1031–1035.

Madec, G. (2008). NEMO ocean engine (No 27). France: Note du Pôle de modélisation, Institut Pierre-Simon Laplace(IPSL). ISSN No 1288-1619.

Megann, A. (2018). Estimating the numerical diapycnal mixing in an eddy-permitting ocean model. *Ocean Modelling*, *121*, 19–33.

Megann, A., Storkey, D., Aksenov, Y., Alderson, S., Calvert, D., Graham, T., et al. (2014). GO5.0: The joint NERC - Met Office NEMO global ocean model for use in coupled and forced applications. *Geoscientific Model Development*, *7*, 1069–1092.

Palmer, M. D., McNeall, D. J., & Dunstone, N. J. (2011). Importance of the deep ocean for estimating decadal changes in Earth's radiation balance. *Geophysical Research Letters*, *38*, L13707. https://doi.org/10.1029/2011GL047835

Purkey, S. G., & Johnson, G. C. (2010). Warming of global abyssal and deep Southern Ocean waters between the 1990s and 2000s: Contributions to global heat and sea level rise budgets. *Journal of Climate*, *23*, 6336–6351.

Rhein, M., Rintoul, S. R., Aoki, S., Campos, E., Chambers, D., Feely, R. A., et al. (2013). Observations: Ocean. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York: Cambridge University Press.

Sloyan, B. M., Wijffels, S. E., Tilbrook, B., Katsumata, K., Murata, A., & Macdonald, A. M. (2013). Deep Ocean Changes near the Western Boundary of the South Pacific Ocean. *Journal of Physical Oceanography*, *43*, 2132–2141.

Steele, M., Morley, R., & Ermold, W. (2001). PHC: A global ocean hydrography with a high quality Arctic Ocean. *Journal of Climate*, *14*, 2079–2087.

Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Yevgeny, A., Blockley, E. W., et al. (2018). UK Global Ocean GO6 and GO7: A traceable hierarchy of model resolutions. *Geoscientific Model Development*, *11*, 3187–3213.

Talley, L. D., Feely, R. A., Sloyan, B. M., Wanninkhof, R., Baringer, M. O., Bullister, J. L., et al. (2016). Changes in ocean heat, carbon content, and ventilation: A review of the first decade of GO-SHIP global repeat hydrography. *Annual Review of Marine Science*, *8*, 185–215.

Timmermann, R., Goosse, H., Madec, G., Fichefet, T., Ethe, C., & Duliere, V. (2005). On the representation of high latitude processes in the ORCA-LIM global coupled sea ice–ocean model. *Ocean Modelling*, *8*, 175–201.

von Schuckmann, K., Palmer, M., Trenberth, K., Cazenave, A., Chambers, D., Champollion, N., et al. (2016). An imperative to monitor Earth's energy imbalance. *Nature Climate Change*, *6*, 138–144.

von Storch, H., & Zwiers, F. W. (2001). *Statistical analysis in climate research*. Cambridge: Cambridge University Press.

Xie, S.-P. (2016). Oceanography: Leading the hiatus research surge. *Nature Climate Change*, *6*, 345–346.