Variability of ocean heat uptake: Reconciling observations and models

K. M. AchutaRao,[1] B. D. Santer,[1] P. J. Gleckler,[1] K. E. Taylor,[1] D. W. Pierce,[2] T. P. Barnett,[2] and T. M. L. Wigley[3]

[1]   This study examines the temporal variability of ocean heat uptake in observations and in climate models. Previous work suggests that coupled Atmosphere-Ocean General Circulation Models (A-OGCMs) may have underestimated the observed natural variability of ocean heat content, particularly on decadal and longer timescales. To address this issue, we rely on observed estimates of heat content from the 2004 World Ocean Atlas (available at http://www.nodc.noaa.gov/OC5/indprod.html, hereinafter referred to as WOA-2004) compiled by Levitus et al., 2005. Given information about the distribution of observations in WOA-2004, we evaluate the effects of sparse observational coverage and the infilling that Levitus et al. use to produce the spatially complete temperature fields required to compute heat content variations. We first show that in ocean basins with limited observational coverage, there are important differences between ocean temperature variability estimated from observed and infilled portions of the basin. We then employ data from control simulations performed with eight different A-OGCMs as a test bed for studying the effects of sparse, space-varying and time-varying observational coverage. Subsampling model data with actual observational coverage has a large impact on the inferred temperature variability in the top 300 and 3000 m of the ocean. This arises from changes in both sampling depth and in the geographical areas sampled. Our results illustrate that subsampling model data at the locations of available observations increases the variability, reducing the discrepancy between models and observations.

Citation: AchutaRao, K. M., B. D. Santer, P. J. Gleckler, K. E. Taylor, D. W. Pierce, T. P. Barnett, and T. M. L. Wigley (2006), Variability of ocean heat uptake: Reconciling observations and models, *J. Geophys. Res.*, *111*, C05019, doi:10.1029/2005JC003136.

1. Introduction

[2]   Increases in observed ocean heat content over the second half of the 20th century were first reported by *Levitus et al.* [2000]. The World Ocean Atlas was compiled and released by Levitus and colleagues in 2000 and facilitated the first systematic comparisons between modeled and observed ocean heat content changes. Prior to that time, most of the formal detection and attribution studies seeking to identify human effects on climate had focused on temperatures near the Earth's surface. The availability of the 2000 World Ocean Atlas (hereinafter referred to as WOA-2000) allowed climate scientists to perform detection and attribution work with temperature changes in the global ocean. This provided a useful consistency check on model estimates of ocean heat uptake, as well as on detection and attribution results that had been obtained previously with atmospheric variables (see *Mitchell et al.* [2001] for a review).

[3]   The first ocean detection studies were by *Barnett et al.* [2001] and *Levitus et al.* [2001]. Barnett et al. analyzed output from the Parallel Climate Model (PCM; *Washington et al.* [2000]) and showed that the PCM "fingerprint" of ocean heat content changes in response to increases in well-mixed greenhouse gases (GHGs) was statistically identifiable in the WOA-2000 data. Similar conclusions were reached by *Levitus et al.* [2001] and *Reichert et al.* [2002], using A-OGCMs developed at the Geophysical Fluid Dynamics Laboratory (GFDL-R30) and at the Max-Planck Institute for Meteorology (ECHAM4/OPYC3). In both A-OGCMs, it was found that observed ocean heat content changes could be successfully reproduced, but only by including anthropogenic forcing. *Levitus et al.* [2001] also noted that the observed ocean heat content changes were far larger than those in other components of the Earth's heat budget.

[4]   The availability of observed estimates of ocean heat content changes provides an important constraint on A-OGCM simulations of late 20th century climate change. Ocean heat content data from WOA-2000 are now used routinely to constrain estimates of climate sensitivity [e.g., *Forest et al.*, 2000; *Gregory et al.*, 2002] and to evaluate the fidelity with which models simulate the secular changes and variability in ocean heat content [*Hansen et al.*, 2002; *Sun*

---

[1]Program For Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, California, USA.
[2]Scripps Institution of Oceanography, La Jolla, California, USA.
[3]National Center for Atmospheric Research, Boulder, Colorado, USA.

and Hansen, 2003; *Gent and Danabasoglu*, 2004; *Gregory et al.*, 2004; *Gent et al.*, 2006; *Hansen et al.*, 2005].

[5] Most models used in the above-mentioned studies successfully capture the long-term trends in observed ocean heat content, but have not been able to reproduce the observed variability on interannual to decadal timescales. For the uppermost 300 m of the global ocean, *Levitus et al.* [2000] found interannual variability in heat content of the order of 3×10^{22} J, which corresponds to a volume–mean temperature change of 0.075°C. Between the mid-1970s and mid-1980s, the WOA-2000 data indicate a decrease in the heat content of the 0–300 m layer of nearly 6×10^{22} J, corresponding to a volume-mean temperature decrease of ca. 0.15°C. Over the same period, the heat content of the 0–3000 m layer decreases by 7.5×10^{22} J.

[6] Ocean heat content changes are considerably less variable in most models. Part of this difference is related to the neglect of volcanic forcing in certain model runs (e.g., in *Barnett et al.* [2001] and *Reichert et al.* [2002]). There is evidence from WOA-2000 of some synchronicity in the timing of explosive volcanic eruptions and global-scale decreases in ocean heat content. The variability of ocean heat content is slightly enhanced in model climate change experiments that incorporate some representation of volcanic forcing. However, inclusion of volcanic effects and solar irradiance changes cannot reconcile modeled and observed variability differences [*Levitus et al.*, 2001; *Hansen et al.*, 2002]. This discrepancy has raised questions [*Hegerl and Bindoff*, 2005] about the reliability of model-based estimates of natural variability, which are a key component of detection and attribution studies. The results of such work could be biased if current A-OGCMs significantly underestimated the unforced variability of ocean heat content.

[7] It is therefore important to evaluate how well current climate models simulate forced and unforced ocean heat content changes. Assessing the reliability of model simulations requires an understanding of uncertainties in both climate models and in the observations themselves [*Santer et al.*, 2003]. A key question here is whether estimates of observed ocean heat content variability are significantly affected by the way in which ocean temperatures have been sampled. In the present study, we investigate whether the variability differences between models and data are partly related to sparse coverage of ocean observations, systematic changes in the depth and geographical location of observations, and the infilling methods used to generate spatially complete temperature fields.

[8] Our analysis considers commonly used measures of heat content, integrated over two different depths (the top 300 and 3000 m of the ocean). The observational data that we use are from a new and updated version of the World Ocean Atlas (available at http://www.nodc.noaa.gov/OC5/indprod.html, hereinafter referred to as WOA-2004) recently released by *Levitus et al.*, 2005. This new data set includes observations not available at the time of the earlier release of WOA-2000. Relative to WOA-2000, the updated heat content time series show smaller increases in ocean heat content in the late 1990s. The heat content variability on interannual to interdecadal timescales is very similar to the variability in WOA-2000.

[9] Alternate estimates of the time evolution of ocean heat content are available. Examples include the independent observational analysis of *Ishii et al.* [2003] and the ocean reanalysis products of *Carton et al.* [2000a, 2000b] and *Stammer et al.* [2002, 2003], which employ ocean models to assimilate in situ data. We rely here on the WOA-2004 data, which remains our best source of information on long-term ocean climate change and the data set that is most frequently used to evaluate models.

[10] Section 2 provides a brief introduction to the observed estimates of heat content. We examine the observations and their implied variability in detail in section 3. Section 4 uses a suite of A-OGCMs to assess the effect of incomplete observational coverage on the simulated variability of ocean temperatures. We present some conclusions in section 5 and address the possible implications of our work for climate model evaluation and for climate change detection studies.

2. Background

[11] The world ocean has been poorly observed, with systematic variations in coverage over space and time. The Northern Hemisphere (NH) oceans are generally better observed than the Southern Hemisphere (SH) oceans, since most observations are concentrated along NH commercial shipping routes. The number of observations is low in the early part of the record (1950s), reaches maximum values in the 1980s and 1990s, and declines slightly in the past few years, because not all recent observations have been incorporated into the database. In tandem with changes in the geographical coverage of ocean measurements, advances in instrumentation and the expansion of monitoring programs have systematically improved our ability to monitor temperature changes in deeper portions of the ocean.

[12] Ocean heat content (HC) is calculated from temperature data using the relation

$$HC = C_p \times \rho \times T \times V \qquad (1)$$

where C_p is the specific heat of seawater at constant pressure, ρ is the density of seawater, and T and V are ocean temperature and volume, respectively.

[13] In order to compute changes in HC over time, temperature must be measured over the entire volume of the ocean. If this condition is not fulfilled, temperatures must be estimated in the "unobserved" portions of the ocean. In both WOA-2000 and WOA-2004, infilling was performed with an objective analysis technique [*Stephens et al.*, 2002]. The heat content calculations are therefore dependent on the coverage and representativeness of the observations and on the reliability of the analysis technique used for infilling.

[14] Recent work by *Gregory et al.* [2004] suggests that sparse, time-varying data coverage in several ocean basins contributes to the apparent mismatch in ocean heat content variability between the HadCM3 A-OGCM [*Gordon et al.*, 2000] and WOA-2000. To study coverage effects, Gregory et al. relied on the raw observations available in the World Ocean Database (WOD-1998; *Levitus et al.* [1998]) and interpolated these onto the HadCM3 Ocean grid (nominally 1.25° latitude × 1.25° longitude and 20 vertical levels). They then subsampled a HadCM3 simulation of 20th century climate change (driven by combined anthropogenic

and natural forcings) at model grid points corresponding to the locations of actual observations.

[15] An innovative aspect of the *Gregory et al.* [2004] investigation was their use of two different methods to infill the model results in ocean areas and levels with no observations. The first method assumed that for each ocean model layer, the average model temperature anomaly of the "observed" portion of the layer was representative of the average anomaly of the entire layer. The second method simply assumed zero temperature anomaly in the "unobserved" portion of each ocean model layer. These two methods (which are identical when data coverage is complete) help to quantify the possible effects of incomplete coverage on observed estimates of ocean heat content changes.

[16] In the relatively well observed top 360 m of the NH oceans (between 0°–65°N), *Gregory et al.* [2004] found that the variability in ocean heat content was comparable in WOD-1998 and HadCM3. In this region, the observed and simulated changes in heat content anomalies (calculated using the two infilling strategies described above) were virtually identical after ca. 1970. Over the more sparsely observed 0–3000 m layer of the global ocean (between 65°N–65°S), there were large differences between the heat content anomalies in the model and in observations, even after 1970. The vertical structure of ocean temperature variability was also different. Subsampling the model data enhanced the subsurface variability, pointing toward sparse data coverage as a contributory factor to the model data variability discrepancies.

[17] *Gregory et al.* [2004] concluded that analysts must be cognizant of such data coverage differences and exercise caution in using the WOA-2000 observational estimates to assess the fidelity with which A-OGCMs simulate heat content changes. In a followup study, *Allison* [2004] compared the *Levitus et al.* [2000] heat content data to results from the ENACT project (Enhanced ocean data Assimilation and Climate prediction), which applied data assimilation techniques to generate spatially complete ocean temperature fields. The global-scale heat content variability in WOA-2000 and ENACT was similar, except for the subsurface variability maximum at 500 m depth in the former, which was absent in ENACT. *Allison* [2004] also examined a climate change experiment performed with a high-resolution OGCM (HadCEM, with a 40-level, 1/3° latitude × 1/3° longitude eddy-permitting ocean model). The variability of ocean heat content in HadCEM was higher than in HadCM3, suggesting that model resolution may also contribute to model-observed variability differences.

3. Observed Data

3.1. Data Sources

[18] WOA-2004 is a gridded data set available on a regular 1° latitude × 1° longitude grid at 33 standard levels. Data are provided as annual, seasonal, and monthly climatologies calculated for the 1957–1990 period and as anomalies from this climatology. For the upper 700 m of the ocean (16 standard levels), anomalies are in the form of annual means for the 49-year period 1955–2003. Over the top 3000 m of the oceans (28 standard levels), running 5-year

mean anomalies are provided for the 40 overlapping pentads between 1955–1998.

[19] WOA-2004 is based on many millions of temperature observations that have been made with a variety of different instruments over the 1955–2003 time period. These observations have been collected in the World Ocean Database (WOD). Raw observations in the WOD are quality controlled and binned into grid cells. The arithmetic means of individual grid cells are then objectively analyzed to fill in grid cells that do not contain data. The infilling method employs both climatological mean information and temperatures from a "region of influence" around grid cells with missing data. The quality control procedures and analysis method are described in detail in *Stephens et al.* [2002].

[20] In the present study, we employ the "*dd*" (data distribution) field reported in the World Ocean Atlas. For each year or pentad, grid cells with *dd* values ≥ 1 (i.e., with at least one observation in the grid cell) were used to define the observed data coverage mask. This focuses attention on areas of the ocean that are actually observed. The coverage mask varies in time and space (latitude, longitude and depth).

3.2. Description of Observed Coverage Changes

[21] To investigate the effect of observed coverage changes, we consider (at each time and grid point) the standard levels from the surface down to the depth of interest (300 and 3000 m in our case) and sum the thickness associated with each standard level if at least one observation exists at that level. Levels with no observations are skipped. The levels and their thicknesses in this summation are the same as the standard levels used in the heat content calculations of *Levitus et al.* [2005]. The results provide a measure of the column of water represented by observations. Figure 1 shows this effective depth of coverage for four individual years: 1964, 1974, 1984, and 1994 (years 10, 20, 30, and 40 of WOA-2004).

[22] Figure 1 (left) is derived from annual mean data for the top 300 m of the ocean. Figure 1 (right) is based on pentadal mean data (centered on 1964, 1974, etc.) for the 0–3000 m range and provide information on the effective depth of coverage over a much larger volume of the ocean. The larger spatial extent for data in the 3000 m range is a result of using pentadal means, for which the $dd \geq 1$ criterion is more easily fulfilled.

[23] Results for both depth ranges show systematic changes in the areal extent and depth of ocean temperature observations. The extent of observational coverage has increased over time. Even in 1994, however, there is sparse coverage of the Arctic and Southern Oceans. The mean depth of coverage for the 0–300 m layer increased from 226 m in 1964 to 286 m in 1994. For the 0–3000 m layer, the bulk of the measurements of deeper portions of the ocean are restricted to the North Atlantic and (in recent decades) to individual transects in the North Pacific and South Atlantic. The effective depth of coverage increased from 763 to 1277 m over this 40-year period.

[24] To provide a more detailed picture of data coverage changes in the upper layers of the ocean, we use the *dd* criterion to compute (separately for each ocean basin and each depth level from 0 to 300 m) the fractional volume of
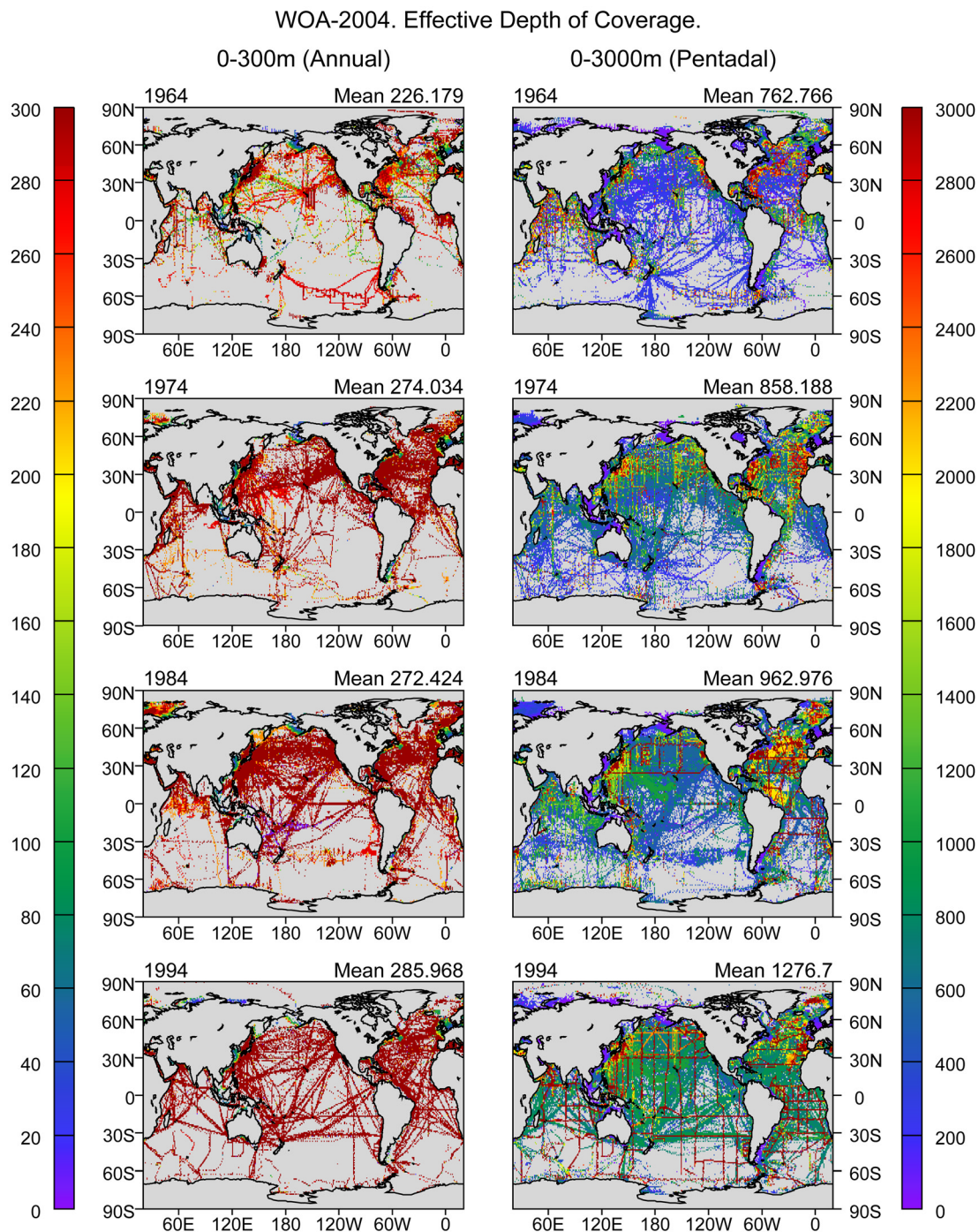
Figure 1. Changes over time in the effective depth (in meters) represented by observations in WOA-2004. (left) Years 1964, 1974, 1984, and 1994 for the top 300 m. (right) Effective depth of coverage for the top 3000 m for pentads centered on the same 4 years. Note the different color scales used for the two depth ranges.

each basin and level that is observed. The time evolution of these coverage changes is shown in Figure 2. In the first decade of the WOA-2004 data set, the coverage in less than 20% in most basins and layers. Coverage is systematically higher in the NH ocean basins, increasing to maximum values of 60–70% in North Atlantic and North Pacific in the 1980s and then declining to ca. 50% in the last decade. Data coverage in the SH ocean basins never exceeds 40%

and is often substantially less than this. One curious feature of the Indian Ocean results (both NH and SH) is that biennial measurement campaigns are clearly identifiable in the spatially averaged coverage data.

[25] Figure 2 also illustrates the systematic increase in the effective depth of coverage in each ocean basin. This is largely due to the introduction of expendable bathythermographs (XBTs) in the 1970s. These large, nonrandom
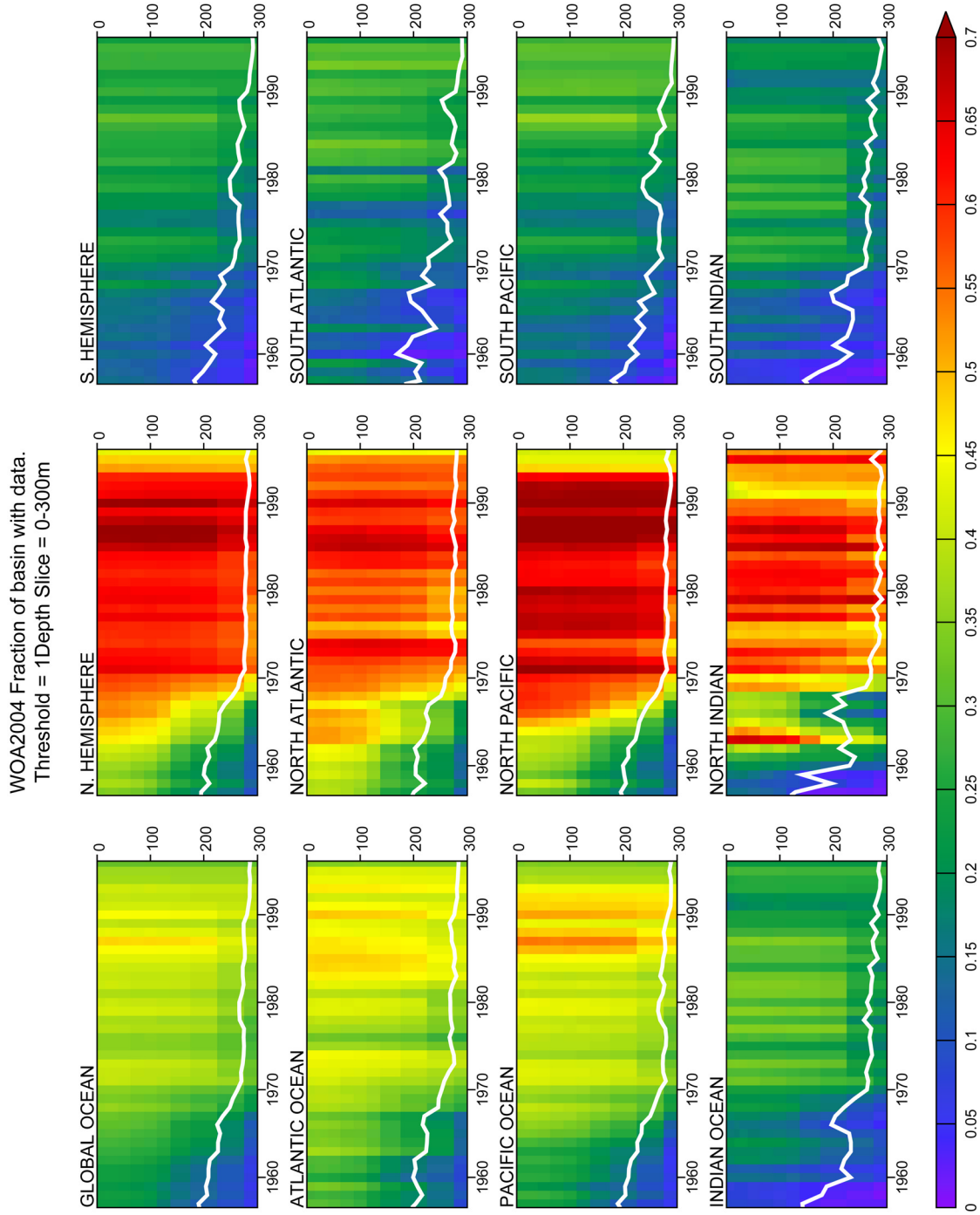
WOA2004 Fraction of basin with data.
Threshold = 1Depth Slice = 0-300m

Figure 2. Time evolution of the volume fraction of observational coverage in the top 300 m of the WOA-2004 data set. Results are shown for individual ocean basins. Time evolution of the effective depth of observations in each basin is indicated by the white line.
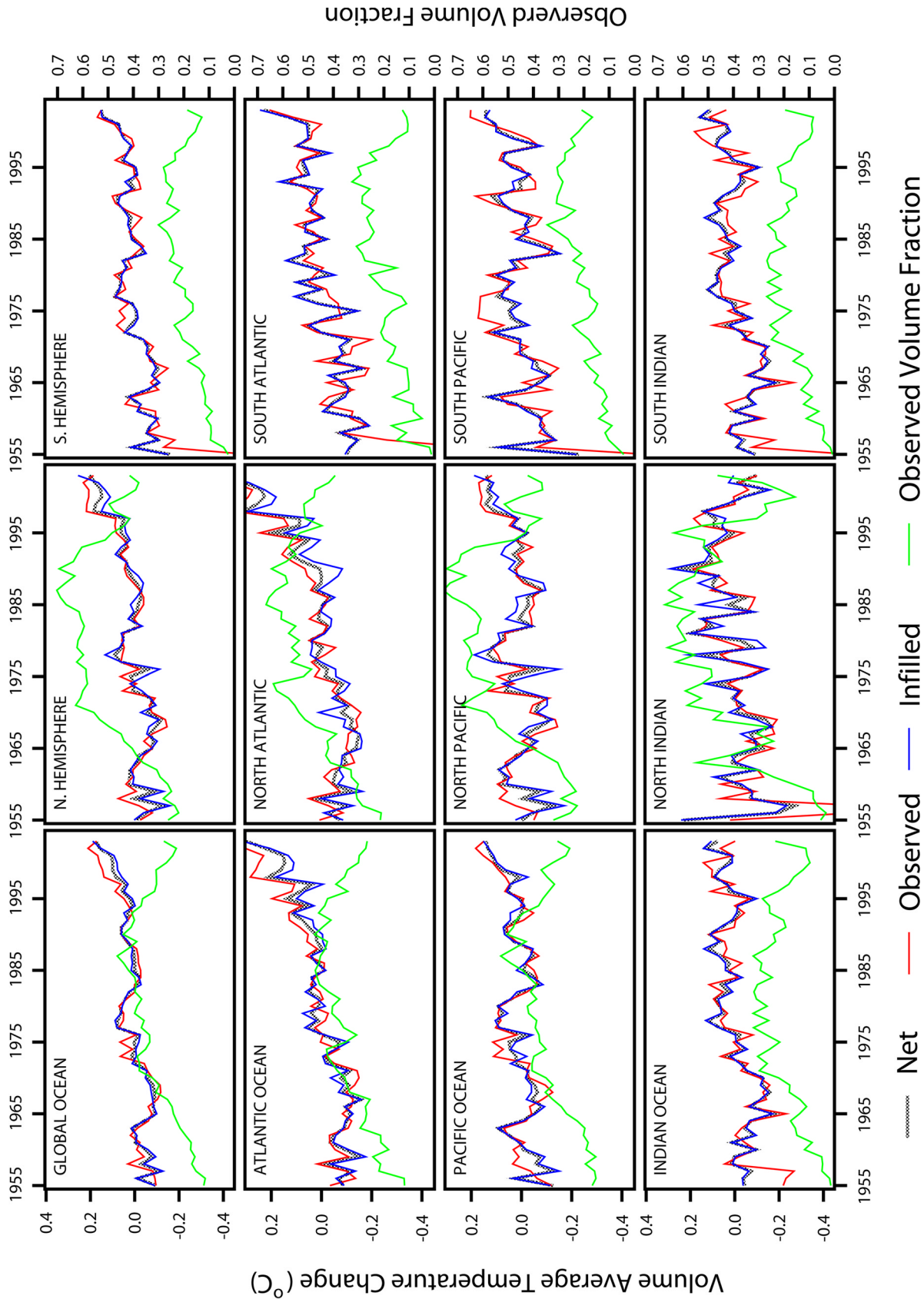
Figure 3. Time evolution of the volume−mean temperature change in the top 300 m of individual ocean basins. Results are for (black line) $T_{Net}(t)$, (red line) $T_o(t)$, and (blue line) $T_I(t)$ (see section 3.3). For each basin the volume fraction of observational data coverage in WOA-2004 is shown in the green line (with the scale on the right).

Table 1. Summary Statistics for WOA-2004 Volume−Mean Temperature Change Over the Top 300 m[a]

Basin	Fractional Volume Coverage			Volume Mean Temperature Change						
				Standard Deviation (Detrended Series)		Linear Trend (Standard Error) $\times 10^{-3}$ °C per year.		Lag 1 Autocorrelation	
	Max	Min	Mean	Net	Obs.	Net	Obs.	Net	Obs.
1	2	3	4	5	6	7	8	9	10
S. Atlantic	0.33	0.01	0.18	0.09	0.14	0.415 (±0.050)	0.626 (±0.136)	0.61	0.75
N. Atlantic	0.67	0.21	0.49	0.12	0.13	0.604 (±0.118)	0.660 (±0.162)	0.87	0.90
ATLANTIC OCEAN	0.47	0.12	0.35	0.09	0.12	0.517 (±0.057)	0.630 (±0.102)	0.84	0.89
S. Indian	0.27	0.01	0.16	0.08	0.11	0.266 (±0.107)	0.370 (±0.094)	0.67	0.38
N. Indian	0.67	0.03	0.42	0.11	0.14	0.229 (±0.110)	0.428 (±0.140)	0.37	0.45
INDIAN OCEAN	0.32	0.01	0.20	0.07	0.09	0.261 (±0.093)	0.358 (±0.083)	0.66	0.55
S. Pacific	0.35	0.04	0.20	0.08	0.12	0.224 (±0.076)	0.312 (±0.150)	0.36	0.55
N. Pacific	0.76	0.23	0.52	0.07	0.08	0.179* (±0.105)	0.153* (±0.138)	0.62	0.65
PACIFIC OCEAN	0.53	0.15	0.34	0.06	0.07	0.204 (±0.085)	0.192* (±0.146)	0.65	0.77
S. Hemisphere	0.30	0.03	0.19	0.06	0.10	0.279 (±0.060)	0.399 (±0.098)	0.69	0.68
N. Hemisphere	0.70	0.22	0.50	0.08	0.09	0.338 (±0.090)	0.360 (±0.150)	0.79	0.85
WORLD OCEAN	0.46	0.12	0.31	0.06	0.07	0.304 (±0.071)	0.349 (±0.111)	0.84	0.86

[a]Asterisks denote trends that are not significantly different from zero at the 5% level. Standard errors (adjusted using the lag-1 autocorrelation of regression residuals, as in *Santer et al.* [2000b]) are shown in parentheses.

changes in the vertical and areal extent of observational coverage suggest that estimates of the mean changes and variability in global-scale ocean heat content may be sensitive to details of the selected infilling method.

3.3. Estimating Effects of Observed Coverage Changes and Infilling

[26] To address this issue, we partition the WOA-2004 temperature data into "observed" and "infilled" subsets, with the former defined by the "*dd*" criterion. The total heat content is a linear combination of the heat content in these two subsets, weighted by the volume fractions of each subset. Because the coverage varies with time, the fractional weights also change over time:

$$T_{Net}(t) = T_o(t) \times f_o(t) + T_i(t) \times f_i(t) \qquad (2)$$

where T_o and T_i are the mean temperatures over the "observed" and "infilled" volumes, respectively, at time t and the weights $f_o(t) + f_i(t) = 1$.

[27] Figure 3 shows values of $T_{Net}(t)$, $T_o(t)$, and $T_i(t)$ for the top 300 m of various ocean basins. Values of $f_o(t)$ are also plotted. The latter help to highlight the systematically lower observational coverage in the SH and the large coverage changes over time (see Figure 2). Apart from the increased coverage in NH basins from the mid-1970s to the mid-1990s, the "observed" fraction of the upper ocean has been consistently less than 50% and is close to zero in SH basins at the beginning of the WOA-2004 record. $T_{Net}(t)$ is therefore strongly influenced by the behavior of $T_i(t)$, and time series of $T_{Net}(t)$ and $T_i(t)$ are virtually superposed for the SH ocean basins. In contrast, $T_o(t)$ is noticeably different from $T_{Net}(t)$ in SH oceans, particularly when $f_o(t)$ is very small in the 1950s. The better observed NH basins show much closer agreement between the $T_{Net}(t)$ and $T_o(t)$ time series, particularly during times of maximum coverage.

[28] Table 1 summarizes some of the key statistical properties of these time series. As noted previously, the best observed basins are the North Atlantic and North Pacific, with time-mean coverage of 49 and 52%, respectively, and maximum coverage of 67 and 76%. The tempo-
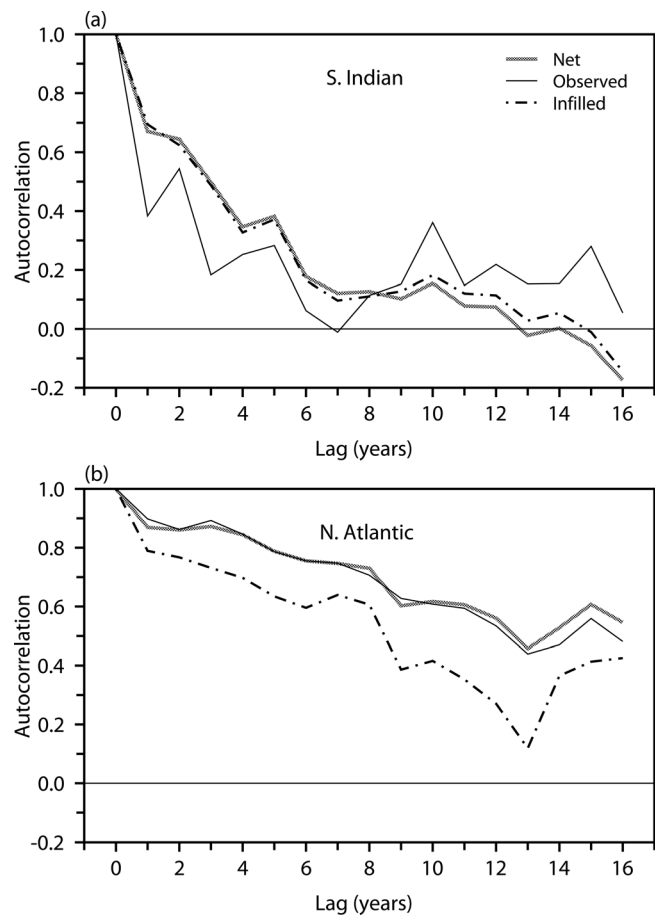

Figure 4. Temporal autocorrelation as a function of lag (from 0 to 16 years) for volume−mean temperature changes in the top 300 m of the (a) South Indian Ocean and (b) North Atlantic Ocean. Results are for $T_{Net}(t)$, $T_o(t)$, and $T_i(t)$ (see section 3.3).

Table 2. Table of CMIP2+ Models Analyzed in This Study, Listed Alphabetically by Model Acronym[a]

| Model | Sponsor, Country | Years Archived | |
		Control	1% CO_2
BCM_Version1, 2002	University of Bergen (UB), Norway	1–300	101–180
CCCma_CGCM2, 2002	Canadian Centre for Climate Modeling and Analysis (CCCma), Canada	51–130	51–130
CSIRO_Mk2, 1997	Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia	351–450	351–450
ECHO-G, 1999	Model and Data Group (MD),Germany	310–409	310–387
GFDL_R30_c, 1996	Geophysical Fluid Dynamics Laboratory (GFDL), USA	1–300	51–130
HadCM3, 1997	Meteorological Office (MO),UK	101–180	101–180
MRI_CGCM2.3, 2002	Meteorological Research Institute (MRI), Japan	1–150	1–150
PCM, 1999	Department of Energy (DOE), USA	1–300	151–232

[a]Information is also provided on the vintage of the simulations and the years of the archived control and 1% CO_2 runs.

ral standard deviation of $T_{Net}(t)$ is consistently smaller than that of $T_o(t)$ in all ocean basins considered. Similarly, the linear trend in $T_{Net}(t)$ over 1955–2003 is smaller than the trend in $T_o(t)$ (in 10 of 12 cases). Such differences must be related to the infilling of large volumes of the ocean with zero anomalies, which tends to damp the positive trends in $T_o(t)$.

[29] In terms of both trends and temporal standard deviations, the differences between $T_{Net}(t)$ and $T_o(t)$ tend to be largest for poorly observed ocean basins. This implies that for these portions of the ocean, infilling can noticeably alter the overall changes and temporal variability of the $T_o(t)$ data.

[30] Differences in r_1, the lag-1 temporal autocorrelation of the $T_{Net}(t)$ and $T_o(t)$ anomaly data, are also largest in basins with sparse observational coverage (Table 1). This is particularly evident for the South Indian Ocean, where $r_1 = 0.67$ for $T_{Net}(t)$ and $r_1 = 0.38$ for $T_o(t)$. The r values for the South Indian Ocean $T_{Net}(t)$ data are systematically higher than those of $T_o(t)$ out to lag 7 (see Figure 4a) and systematically lower than those of $T_o(t)$ from lags 9 to 16. For this ocean basin, the temporal autocorrelation structure of $T_{Net}(t)$ is dominated by temperature changes in infilled rather than in observed areas. In contrast, in a well-observed basin like the North Atlantic (Figure 4b), the temporal autocorrelation structure of $T_{Net}(t)$ is largely driven by temperature changes in $T_o(t)$.

[31] In summary, the results presented in this section suggest that infilling can have a nonnegligible effect on basic statistical properties of the $T_{Net}(t)$ time series, particularly for poorly observed basins. Our focus has been on the top 300 m of the oceans, for which data coverage is considerably higher than in the deep ocean. The $T_{Net}(t)$ versus $T_o(t)$ differences identified for the upper ocean are therefore likely to be larger and more serious for the 0–3000 m layer.

4. Simulated Variability of Ocean Temperatures
4.1. Model Data

[32] Climate models with spatially complete data provide a useful test bed for exploring the effects of changing observational coverage [Santer et al., 2000a; Duffy et al., 2001]. Here, we rely on results from eight different A-OGCMs that participated in the Coupled Model Intercomparison Project (CMIP; Meehl et al. [2000]). Under the CMIP2+ phase of this project, modeling groups contributed

output from a pair of simulations. The first was a control simulation with no changes in external forcings. In the second experiment (hereafter referred to as the 1% CO_2 run), CO_2 was increased at a compounded rate of 1% per year, leading to a doubling of atmospheric CO_2 by year 70.

[33] The models involved in CMIP2+ differ in many important aspects, including horizontal and vertical resolution, dynamics, the underlying physical parameterizations, and in their use of flux adjustment. Unlike the 1% CO_2 runs, model control simulations were not carried out with specific output requirements and vary in length from 80 to 300 years. Table 2 lists the models used in this study with details of their vintage, lengths of individual runs, etc. The CMIP2+ data are available through the Program for Climate Model Diagnosis and Intercomparison (PCMDI) at http://www-pcmdi.llnl.gov. This Web site provides detailed references documenting key features of the A-OGCMs participating in CMIP2+.

[34] The climatological mean performance of the OGCM components of the CMIP2+ models has been documented in Gleckler et al. [2004]. In the present study, we analyze the effect of observed coverage changes on the simulated variability of ocean temperatures in the CMIP2+ control runs. Our focus is on interannual and decadal timescale variability in the CMIP2+ control runs, although we also consider the effects of coverage changes in the context of the CO_2 increase experiments.

4.2. Decorrelation Times in Model Control Runs

[35] Before exploring the effects of observational coverage changes on model-based variability estimates, it is useful to briefly compare the temporal variability of ocean temperatures in the eight CMIP2+ control runs. We use the decorrelation time r_t as the basis for this comparison. We define r_t by computing annual mean temperature anomalies (relative to the initial year of the control run), vertically averaging these anomalies over the top 300 m of the ocean, and then determining the lag t (in years) at which the temporal autocorrelation falls below 0.5. In Figure 4a, for example, the decorrelation time is three years for upper ocean (top 300 m) temperature anomalies in the South Indian Ocean.

[36] Values of r_t are calculated at each grid point for each model control run. Because the length of these integrations is variable, we stipulated that the maximum lag could not exceed $n/3$, where n is the total length of the control run, thus yielding an upper bound on r_t. Since the spectrum is
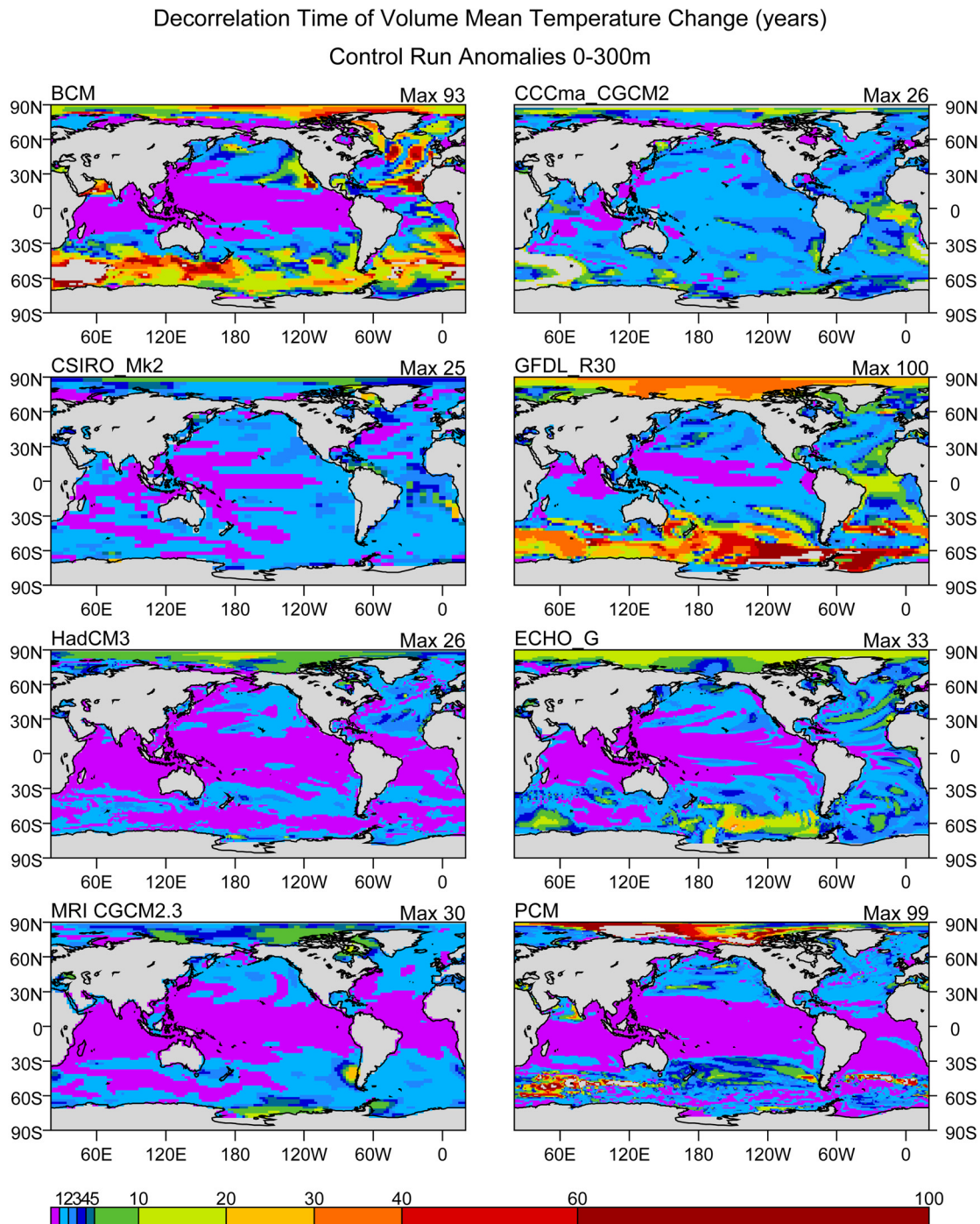
Figure 5. Geographical distribution of decorrelation time (in years) for the annual mean volume–mean temperature anomalies calculated over the top 300 m of the ocean in the CMIP2+ model control runs. Maximum lag (shown at the upper right corner for each panel) is $n/3$, where n is the total length of the control run. Note nonlinear color scale.

related to the Fourier transform of the autocorrelation function [*Jenkins and Watts*, 1968], maps of r_t provide basic information on the spatial distribution of ocean temperature variability at different timescales.

[37] Figure 5 shows that there is a large range in r_t, both geographically and across the eight CMIP2+ models. A few features are common to all models, such as the short decorrelation times ($\leq 1-2$ years) in the tropics and in the

vicinity of the western boundary currents. Large intermodel differences in r_t are evident in the Southern Ocean, where the BCM, GFDL R30, ECHO-G, and PCM models have areas with decorrelation times of decades or longer. The other CMIP2+ models have much shorter decorrelation times in this region.

[38] Some aspects of the spatial distribution of r_t are closely linked to features of each model's oceanic circula-
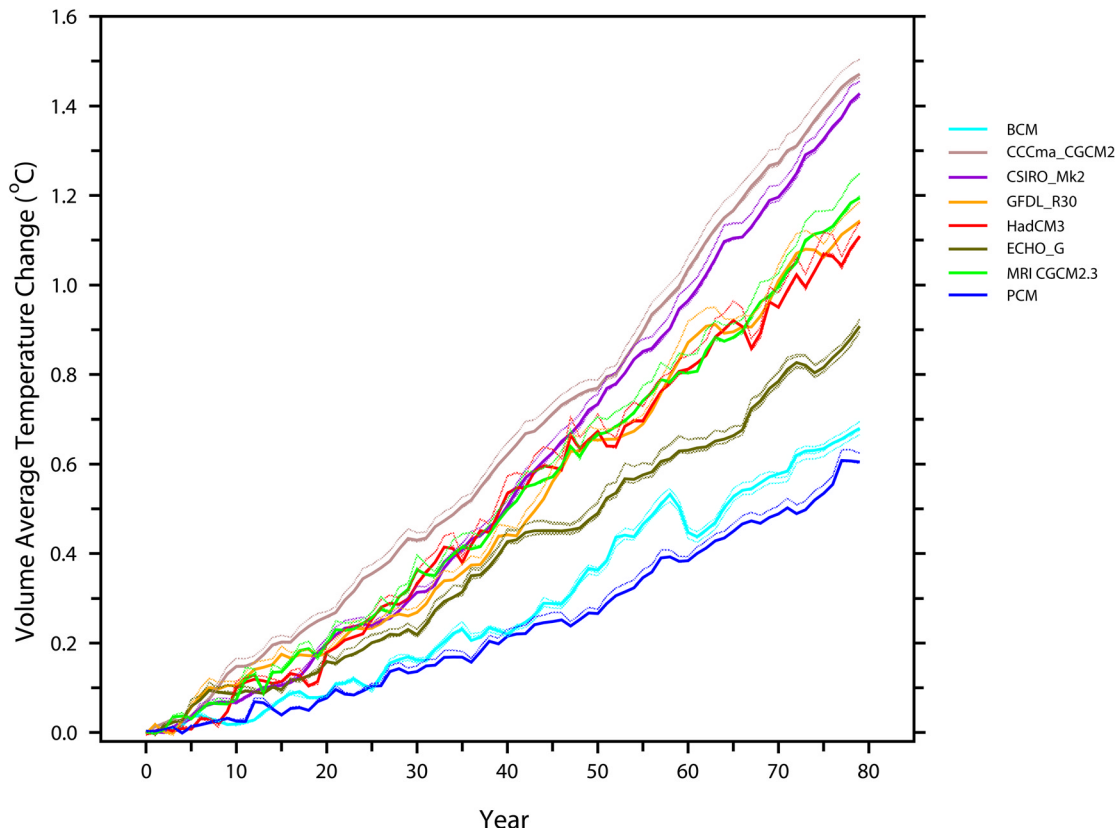
Figure 6. Effect of regridding on the volume–mean temperature change in the top 300 m of the CMIP2+ 1% CO_2 runs. Results are for four different regridding methods (thin lines). Bold lines denote changes calculated on the original model grid.

tion, as is evident when surface currents are superimposed on Figure 5 (not shown). Other aspects of the r_t fields (such as some of the long decorrelation times in the Southern Ocean) are more difficult to diagnose and may arise from some combination of bona fide low-frequency variability of the coupled system and/or residual model drift in the control runs. Here, our primary interest is not in the causes of this variability, but rather in a gross characterization of its spatial structure, timescales, and intermodel differences. Such information will be useful in understanding how the sampling of A-OGCM upper ocean temperature fields with incomplete observational coverage (Figure 1) may alter the simulated variability.

4.3. Regridding of Model Ocean Temperature Data

[39] The simulated ocean temperature data are on grids of varying resolutions and geometries. In sampling model output with the observed data coverage mask, either the observations must be transformed to the model grid, or the model output must be transformed to the grid used in WOA-2004. In the first approach, the individual WOD observations that have been incorporated into the WOA-2004 are "binned" as in the work of *Levitus et al.* [2005], but now on the model grid. This generates a coverage mask unique to each model, a somewhat cumbersome process when dealing with multiple models.

[40] The second approach, transforming individual model grids to the WOA-2004 grid, has the advantage that it allows different sampling strategies to be implemented in

a consistent way across a range of models. However, regridding can change the resolution and even the geometry of the grid and thereby alter both the volume and temperature and hence the heat content. Since we are attempting to quantify the effects of subsampling on ocean heat content, it is important to verify that errors introduced by the selected regridding procedure are within acceptable limits.

[41] Ocean model output is not archived at the standard WOA-2004 levels and must be regridded both vertically and horizontally. We performed the regridding in two separate steps (first horizontally and then vertically). Results are not affected by the order of operation. Two different horizontal and two different vertical regridding procedures were examined, yielding a total of four different regridding combinations. We used all of these combinations to transform the control and 1% CO_2 runs from each of the eight CMIP2+ models to the WOA-2004 grid.

[42] Figure 6 shows the volume-weighted temperature changes (1% CO_2 run minus control) over the top 300 m of the global ocean for individual models. Results are displayed on the original model grid and after regridding. The effect of regridding is small relative to the overall temperature change at the end of the climate change experiment. Additionally, regridding does not distort the variability of 0–300 m temperature changes. These results hold for different depth ranges and ocean basins and for both ocean heat content and volume-averaged temperature change. In all cases, the effect of regridding is no more than 3% of the overall temperature or heat content change (on the
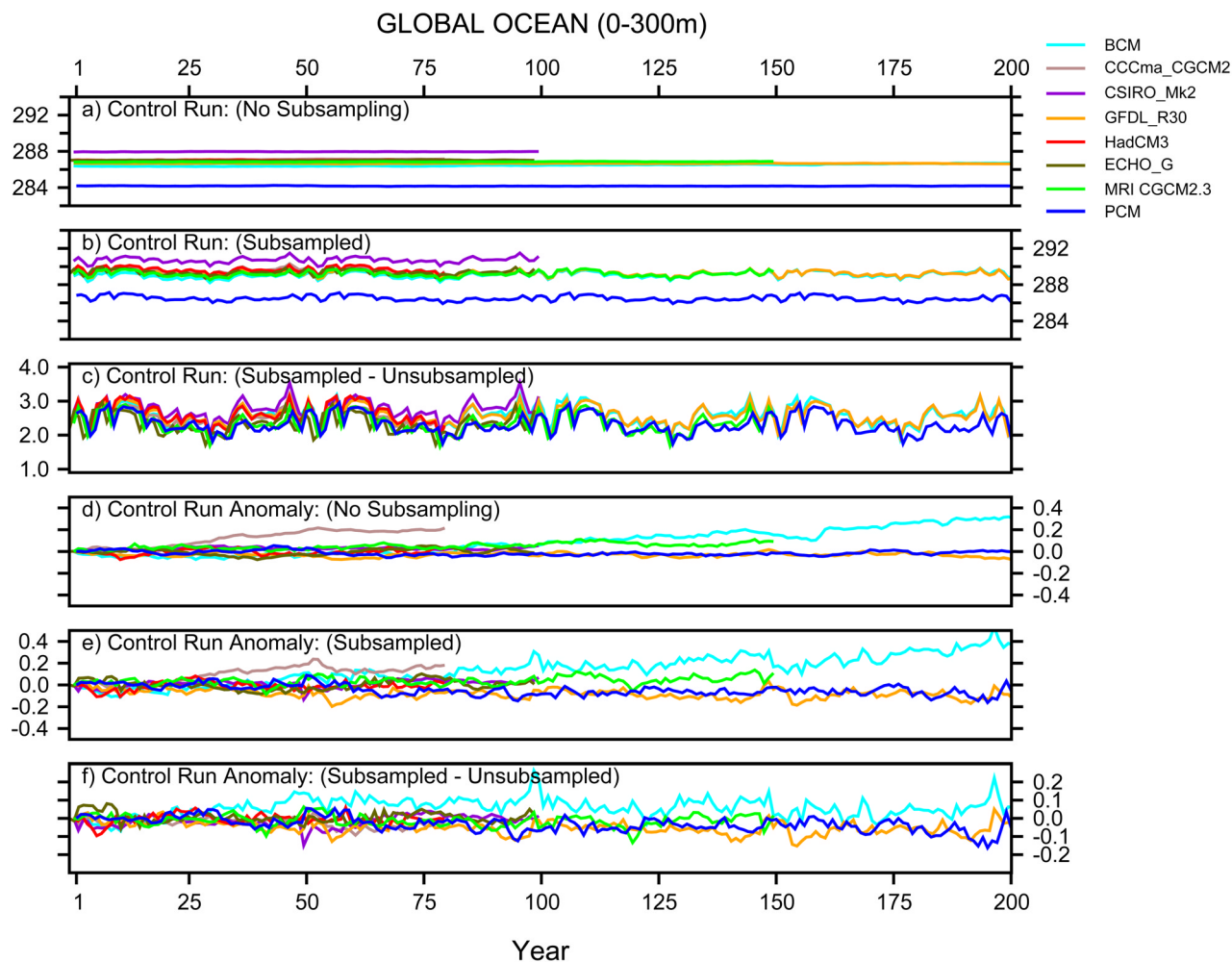
Figure 7. Time series of annual mean volume–mean temperature in (a) $T'_{Net}(t)$, (b) $T'_{Sub}(t)$, and (c) $T'_{Sub}(t) - T'_{Net}(t)$. Results are from control runs of the eight individual CMIP2+ models and are for the top 300 m of the global ocean. Analogous time series calculated with anomaly data rather than absolute data are shown for (d) $\Delta T'_{Net}(t)$, (e) $\Delta T'_{Sub}(t)$, and (f) $\Delta T'_{Sub}(t) - \Delta T'_{Net}(t)$.

original model grid) at the end of the 1% CO$_2$ experiment and is generally ≪3% of the final change. Because of this very small sensitivity to the choice of regridding method, we use only one method when sampling model output with observational coverage.

4.4. Subsampling Model Data With Observed Coverage

[43] We rely on two *dd*-based observational coverage masks (see section 3.1 and Figure 1). The first is for the 0–300 m layer and utilizes 49 years of annual mean data. The second is for the 0–3000 m layer and uses 40 overlapping pentadal means (with overlap by all but one year). After transforming model control and 1% CO$_2$ run output to the WOA-2004 grid, we apply the observational masks to the regridded annual and pentadal data. Since the model experiments are longer than the WOA-2004 temperature records, masking was repeated cyclically (that is, year 50 of the annual mean 0–300 m temperature data from the control run was sampled with the observational coverage in the first year (1955) of the WOA-2004 data, etc.)

[44] Consider first the evolution of global mean ocean temperatures, averaged over 0–300 m in the eight CMIP2+

control runs. Consistent with the terminology used for the observations, we denote this by $T'_{Net}(t)$, where the prime denotes a simulated result and "*Net*" signifies no subsampling. The subsampled version of this is $T'_{Sub}(t)$. Values of $T'_{Net}(t)$ range from 284 K in PCM to 288 K in the CSIRO_Mk2 model (Figure 7a). For each individual model, the variability of absolute values of $T'_{Net}(t)$ is small relative to the intermodel differences in $T'_{Net}(t)$.

[45] Subsampling model control run data with observed coverage increases both the mean and variability of $T'_{Sub}(t)$ relative to $T'_{Net}(t)$ (see Figure 7b). The increase in the mean is by roughly 2–3 K and arises from preferential sampling of the warmer near-surface layers of the ocean (see Figure 2). The effect of subsampling is illustrated by subtracting $T'_{Net}(t)$ from $T'_{Sub}(t)$ (Figure 7c). The temporal variability in this difference time series is highly correlated in all models, a strong indication that it is induced by the subsampling.

[46] We next examine the annual mean anomalies of the 0–300 m temperature data, denoted by $\Delta T'_{Net}(t)$ and $\Delta T'_{Sub}(t)$. Anomalies are defined as the departures of $T'_{Net}(t)$ and $T'_{Sub}(t)$ from their respective values in the first year of
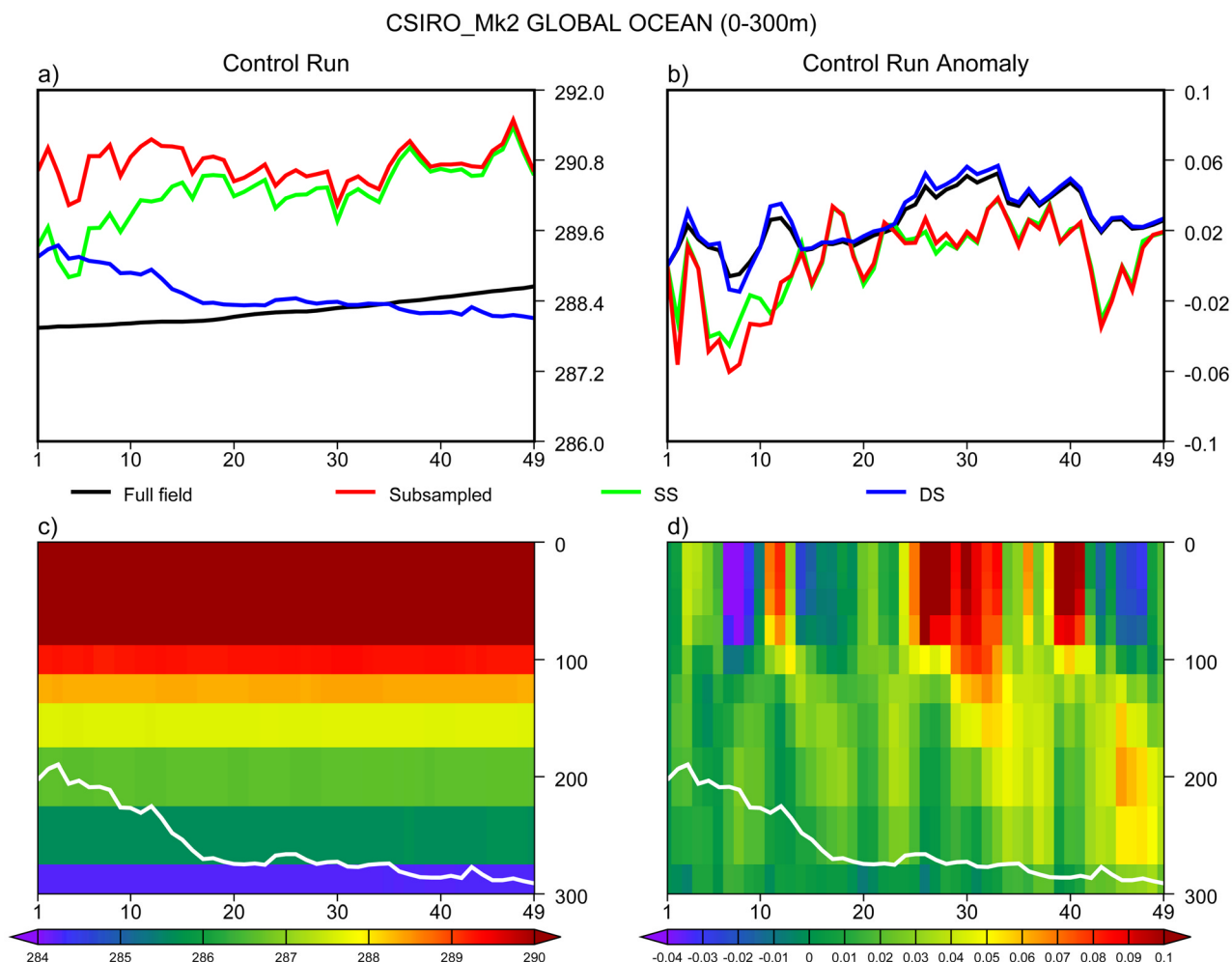
Figure 8. Changes in global ocean temperature in the first 49 years of the CSIRO control run. Results are for (a, c) absolute temperatures and (b, d) anomalies. (top) Volume-averaged changes for the top 300 m of the ocean. (bottom) Temperature changes are given at discrete levels; the white lines denote the effective depth of coverage. For details regarding the SS and DS sampling methods (green and blue lines), refer to section 4.5.1. Black and red lines in Figures 8a and 8b denote unsubsampled and subsampled data, respectively.

the control run. This definition helps to illustrate the residual climate drift in $\Delta T'_{Net}(t)$ in the CCCma, BCM and MRI control runs (Figure 7d).

[47] As in the case of the absolute temperature data (Figures 7a and 7b), subsampling the 0–300 m temperature anomaly fields increases the variability in $\Delta T'_{Sub}(t)$ relative to $\Delta T'_{Net}(t)$ (Figures 7d and 7e). However, the variability induced by subsampling the anomaly fields is not strongly correlated between models, as it was for absolute temperatures (compare with Figures 7c and 7f). We attempt to understand this result in section 4.5.

4.5. Effect of Subsampling: CSIRO_Mk2 Control Run
4.5.1. Global Ocean Results
[48] To investigate in more detail the enhanced variability induced by subsampling, we confine our attention to a single model (CSIRO_Mk2). Figure 8 shows the time evolution of temperatures in the upper 300 m of the global ocean in the CSIRO_Mk2 control run. Results are for absolute temperatures (Figures 8a and 8c) and anomalies (Figures 8b and 8d). Figure 8 (top) gives the vertically

integrated 0–300 m temperature results, while Figure 8 (bottom) displays changes at discrete levels. The white lines in Figures 8c and 8d indicate the effective depth of observational coverage (see Figure 2).

[49] As noted in section 4.4, subsampling model data with WOA-2004 coverage introduces a warm bias in $T'_{Sub}(t)$ relative to $T'_{Net}(t)$. This bias arises because $T'_{Sub}(t)$ preferentially samples warmer upper layers, as is clearly evident in Figures 8a and 8c. Subsampling also amplifies the variability of both $T'_{Net}(t)$ and $\Delta T'_{Net}(t)$ (Figures 8a and 8b).

[50] The observational results in Figures 1 and 2 indicate that there have been significant temporal changes in both the areal extent of observational coverage and the depth at which observations are taken. One advantage of using spatially complete model data is that we can deconvolve these two effects and estimate the relative contributions of depth and areal coverage changes to the variability differences induced by subsampling.

[51] We perform this deconvolution in two different ways. In the "Spatial Sampling" (SS) strategy we designate the entire 0–300 m water column (at any given model grid
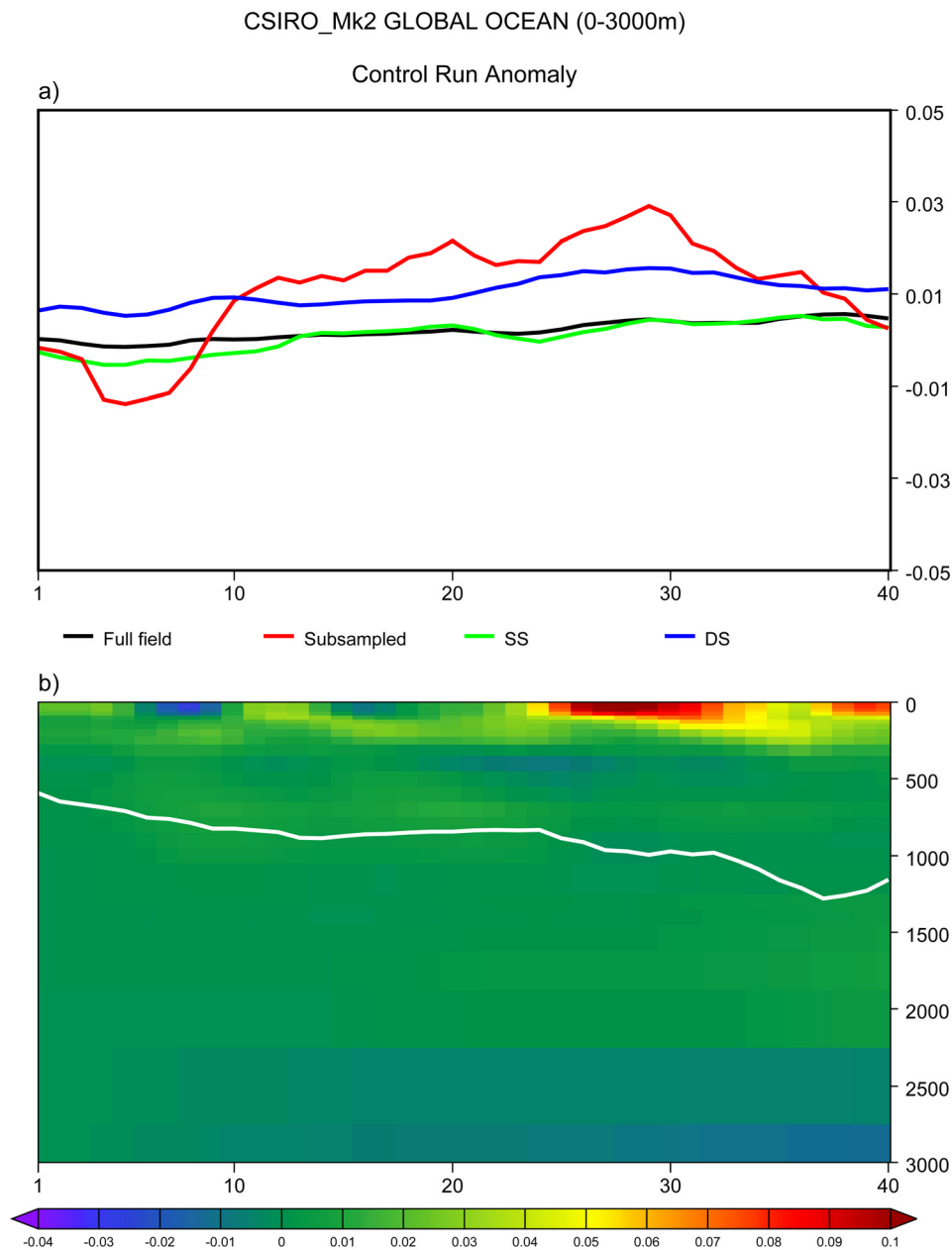
Figure 9. As for Figures 8b and 8d, but for the first 40 pentads of the 0–3000 m layer.

point and in any given year) as "observed" if the annual mean *dd* mask indicates that a valid observation was present at any level between 0 and 300 m for that grid point and year. This strategy minimizes the effect of temporal changes in the depth of coverage and isolates the effect of changes in areal coverage. In the "Depth Sampling" (DS) strategy, we assume that areal coverage of the 0–300 m layer is spatially complete and time-invariant. The only change over time is in the average depth of observational coverage (see white lines in Figures 8c and 8d).

[52] Consider first the effects of subsampling absolute temperatures with the SS and DS approaches (Figure 8a). For the first 30 years of the CSIRO control, DS yields temperatures that are warmer than in $T'_{Net}(t)$, supporting our earlier conclusion that this bias is introduced by preferential

sampling of warmer near-surface layers in the early years of WOA-2004. A comparison of the SS, DS, and $T'_{Sub}(t)$ results shows that the systematic increase in sampling depth is responsible for much of the low-frequency variance in $T'_{Sub}(t)$. Conversely, comparison of the SS and $T'_{Sub}(t)$ time series reveals that the high-frequency variability in $T'_{Sub}(t)$ is largely dictated by changes in geographical coverage and not by changes in sampling depth.

[53] In the case of the anomaly fields (which is the relevant field for a direct comparison with WOA-2004), the DS and $\Delta T'_{Net}(t)$ time series are virtually identical, as are SS and $\Delta T'_{Sub}(t)$ (Figure 8b). This illustrates that the temporal variability in $\Delta T'_{Sub}(t)$ arises primarily from changes in the location and areal extent of sampling. The fact that sampling depth changes have relatively little
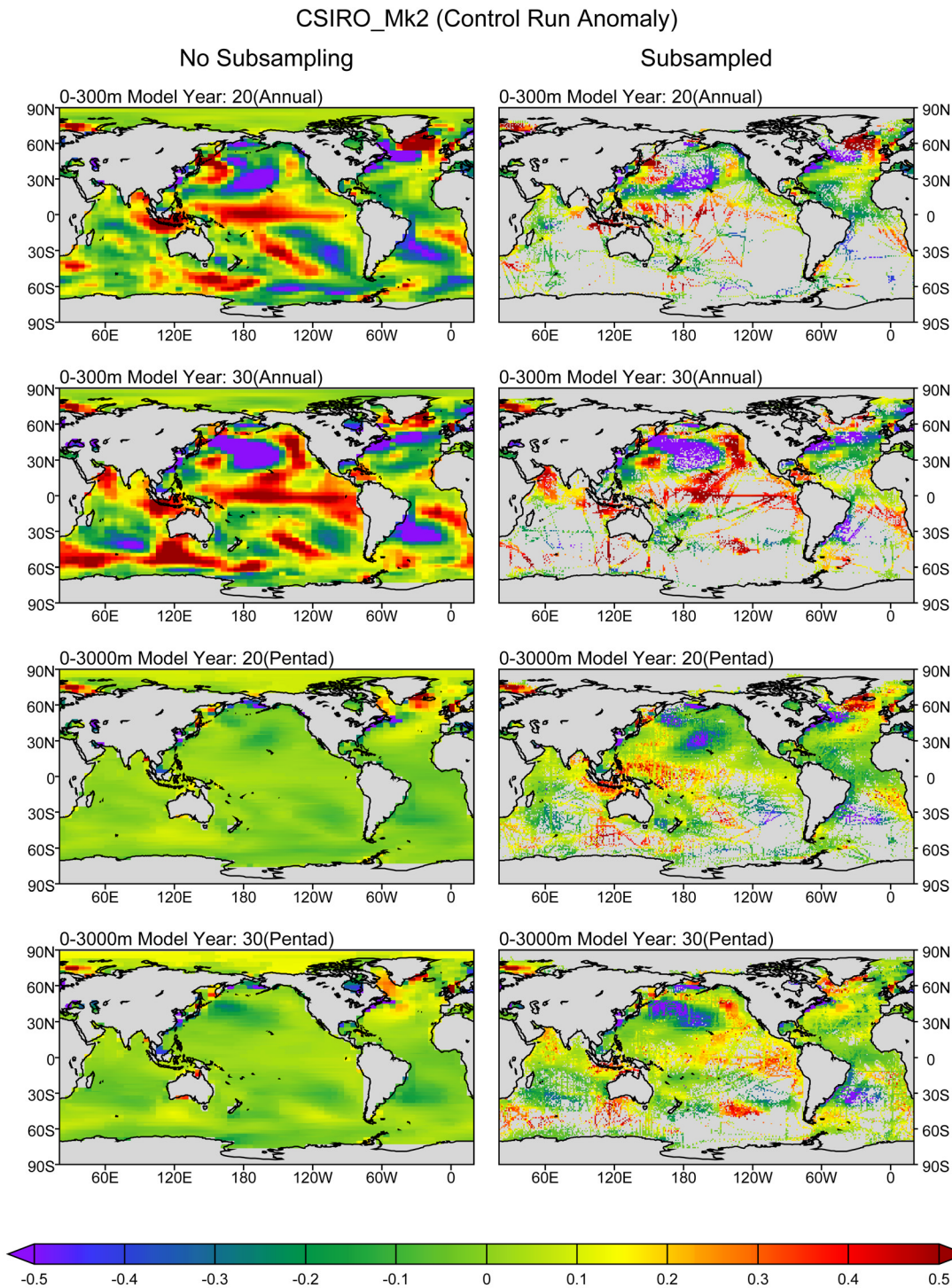
Figure 10. Spatial distribution of volume–mean temperature anomalies in the CSIRO_Mk2 control run. Results are for (left) $\Delta T'_{Net}(t)$ and (right) $\Delta T'_{Sub}(t)$ and are given for two different times and depth ranges (0–300 m and 0–3000 m). Coverage masks used to create the subsampled data for years 20 and 30 of the CSIRO data correspond to calendar years 1974 and 1984 in Figure 1.

impact on the variability of $\Delta T'_{Sub}(t)$ is due to the broad vertical coherence of temperature anomalies in the upper 300 m of the CSIRO control run (Figure 8d).

[54] We performed a similar subsampling exercise for global ocean temperature anomalies in the top 3000 m of the CSIRO control run (Figure 9). Results differ markedly from those obtained for the 0–300 m anomaly data. For the

0–3000 m data, $\Delta T'_{Net}(t)$ is highly similar to the SS time series (Figure 9a). The DS results are consistently warmer than $\Delta T'_{Net}(t)$. This is because the effective subsampling depth never exceeds 1500 m and therefore fails to capture the slight cooling below 2000 m (Figure 9b).

[55] While $\Delta T'_{Sub}(t)$ and SS results were very similar for the 0–300 m layer, the variability in $\Delta T'_{Sub}(t)$ for the 0–
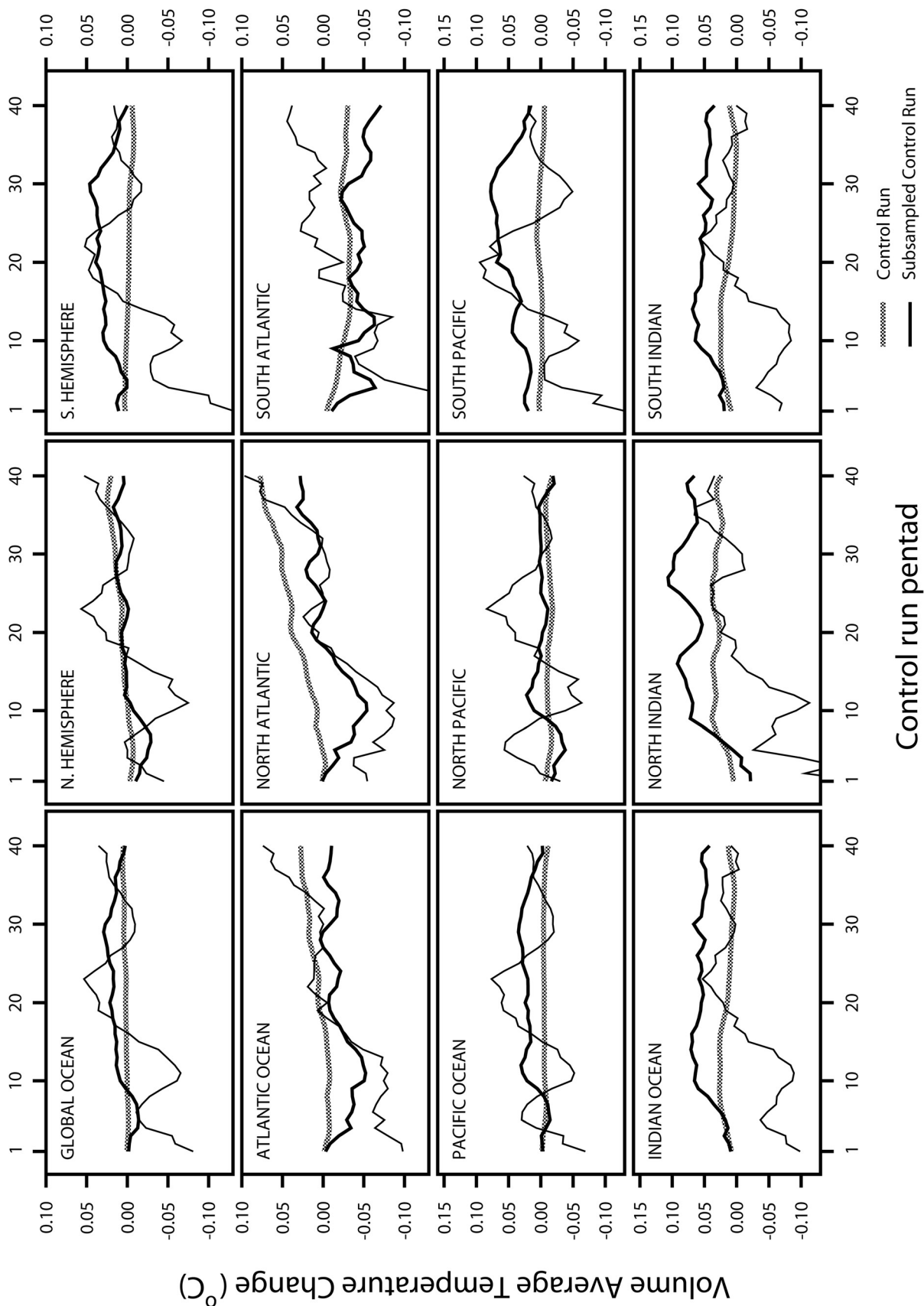
Figure 11. Time series of volume−mean temperature anomalies in the CSIRO Mk2 control run. Results are for the top 3000 m of 12 different ocean basins. Comparisons of $\Delta T'_{Net}(t)$ and $\Delta T'_{Sub}(t)$ illustrate the effect of subsampling on the simulated variability. $T_o(t)$ values from WOA-2004 are overlaid for each basin.
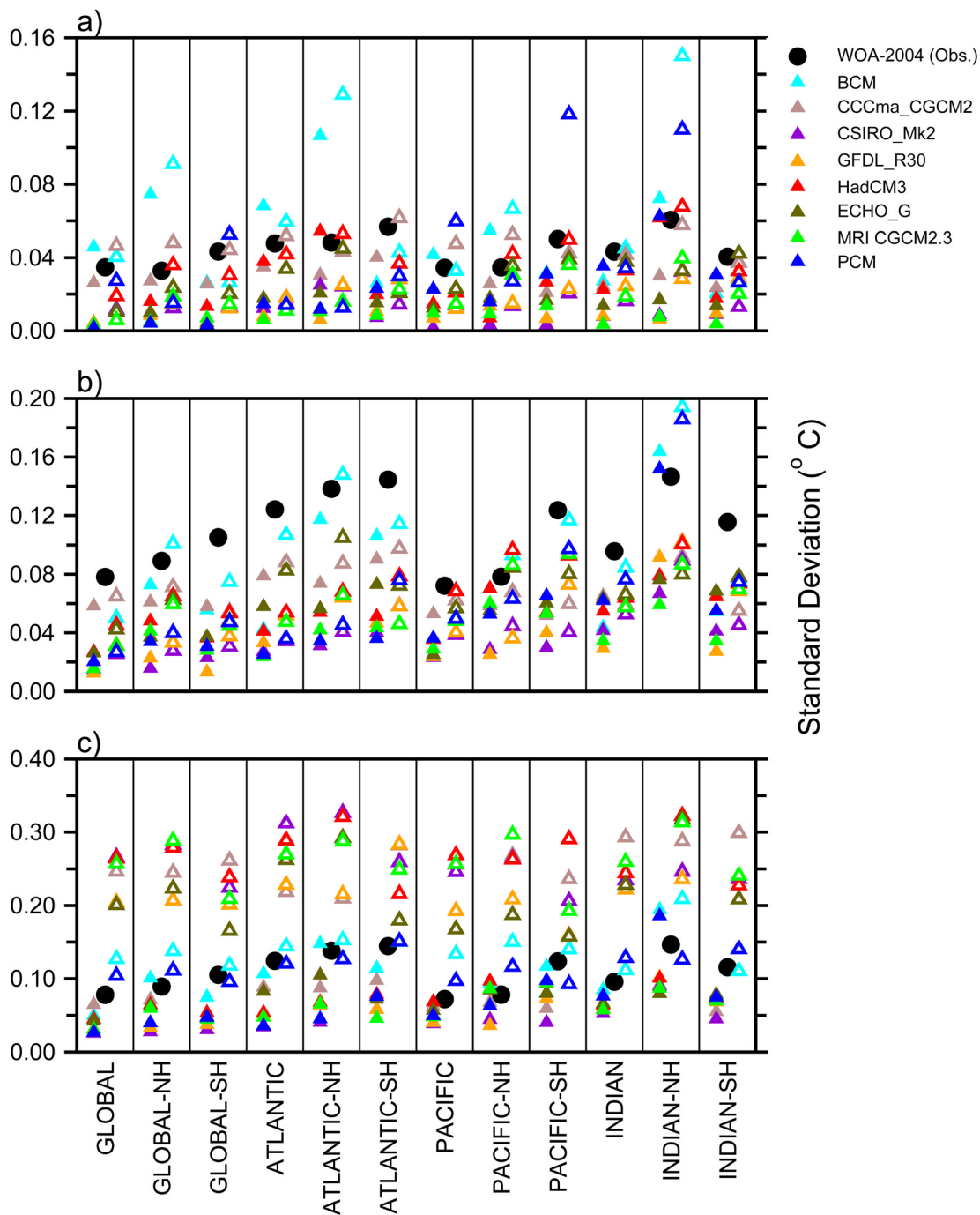
Figure 12. Temporal standard deviations of the volume–mean control run temperature anomalies in (solid triangles) $\Delta T'_{Net}(t)$ and (open triangles) $\Delta T'_{Sub}(t)$. Results are shown for 12 different ocean basins and for the 0–3000 and 0–300 m ranges (a and b). Standard deviations of $\Delta T'_{Sub}(t)$ from the (solid triangles) control run anomalies and (open triangles) 1% CO_2 run anomalies for the 0–300 m range are compared in (c). Standard deviations of $T_o(t)$ from the WOA-2004 data are denoted by solid circles.

3000 m layer is not as clearly related to either SS or DS results. This suggests that for this deeper layer, which has a more complex vertical structure of temperature anomalies (Figure 9b), the variability in $\Delta T'_{Sub}(t)$ must arise from the combined effects of changes in sampling depth and areal coverage rather than from changes in the latter only.

[56] Figure 10 illustrates the effect of subsampling for the 0–300 m layer (top four panels) and the 0–3000 m layer

(bottom four panels). This example uses anomaly data for years 20 and 30 of the CSIRO control run (or pentads centered on these years). These selected times coincide with the observed coverage maps for 1974 and 1984 (Figure 1).

[57] Consider first the 0–300 m results. A comparison of $\Delta T'_{Net}(t)$ and $\Delta T'_{Sub}(t)$ over areas of observational coverage shows that subsampling successfully reproduces the gross structure and size of the $\Delta T'_{Net}(t)$ anomalies. This is largely

due to the vertical coherence of temperature anomalies in the 0–300 m layer in the CSIRO control run. The sparse observational coverage fails to capture large, coherent anomalies that lie outside the coverage mask. This is why the SS approach yielded results that were closest to $\Delta T'_{Sub}(t)$ (Figure 8b).

[58] For the 0–3000 m layer, it is evident that subsampling can distort the size (and sometimes even the sign) of the temperature anomalies on the spatially complete field. (Figure 10; bottom four panels). $\Delta T'_{Net}(t)$ is cooler than $\Delta T'_{Sub}(t)$ in many locations, particularly where the effective sampling depth is less than 1000 m (see Figure 1). In these areas, the relatively shallow observational measurements do not reliably portray the vertically complex structure of the temperature changes in the full 0–3000 m water column. The closest agreement between $\Delta T'_{Net}(t)$ and $\Delta T'_{Sub}(t)$ is in regions such as the North Atlantic, where the effective sampling depth often exceeds 2000 m.

4.5.2. Results for Individual Ocean Basins

[59] In attempting to elucidate the effects of observational coverage changes on the simulated variability of ocean temperatures, our initial focus has been on the global ocean. However, as shown in Figure 2, coverage changes have different characteristics in different ocean basins. In section 4.6, we use the CSIRO control run to study the effects of subsampling on the variability of 0–3000 m temperatures in individual basins.

[60] Comparison of $\Delta T'_{Net}(t)$ and $\Delta T'_{Sub}(t)$ indicates that the variability of 0–3000 m temperature on the original model grid is amplified by subsampling with observed coverage changes (Figure 11). This amplification occurs in each ocean basin and confirms that the previously described global ocean result was in no way anomalous.

[61] Figure 11 also includes observed pentadal mean 0–3000 m temperature anomalies from WOA-2004 (defined with the "dd" mask). As noted previously, there is no direct time correspondence between temperature anomalies in WOA-2004 and in the CSIRO control run. We include the WOA-2004 results to provide a simple comparison of the amplitude of internally generated variability in the model and variability in observations (in observed portions of the ocean).

[62] This comparison shows that in most ocean basins, subsampling does not fully reconcile the variability of 0–3000 m temperatures in WOA-2004 and in the CSIRO control run. The observed variability is consistently larger than the variability in $\Delta T'_{Sub}(t)$. This discrepancy is partly attributable to the fact that the control run does not include external forcings that have contributed to the trend and low-frequency variability in observed ocean temperatures [e.g., Hansen et al., 2002; Barnett et al., 2005; Pierce et al., 2006].

[63] One curious aspect of Figure 11 relates to the WOA-2004 and $\Delta T'_{Sub}(t)$ time series in the North Atlantic, where the CSIRO control run has a residual warming trend. Sampling this spatially coherent warming with the observed coverage mask yields low-frequency temperature changes that are highly correlated with observed results. This correspondence is either purely fortuitous, or is some way related to the effects of coverage changes on a coherent warming signal in the observations and coherent drift in the control run.

4.6. Effect of Subsampling: CMIP2+ Control and 1% CO$_2$ Runs

[64] In section 4.5, we demonstrated that the variability of 0–300 and 0–3000 m temperatures in the CSIRO model was invariably enhanced by subsampling spatially complete model output with actual observational coverage. This was evident in all ocean basins examined. To determine whether this is a general result or unique to the CSIRO_Mk2 model, we repeated the subsampling of 0–300 and 0–3000 m anomaly fields with control run output from the seven other CMIP2+ models.

[65] Figure 12 shows the temporal standard deviation of $\Delta T'_{Net}(t)$ and $\Delta T'_{Sub}(t)$ for both ocean layers and for 12 different ocean basins. Results were calculated from the first 49 years (40 overlapping pentads) of the 0–300 m (0–3000 m) control run anomaly fields (Figures 12a and 12b). In both layers, and for virtually every ocean basin and model control run, the standard deviation of $\Delta T'_{Sub}(t)$ exceeds that of $\Delta T'_{Net}(t)$, that is, subsampling consistently amplifies the variability of ocean temperatures.

[66] Standard deviations of the dd-masked WOA-2004 results are also plotted in Figure 12. For the 0–3000 m layer, the observed variability of $\Delta T'_{Sub}(t)$ lies within the range of model values. This is a noteworthy result, since it is the variability of unsubsampled $\Delta T'_{Net}(t)$ data that differs markedly in models and observations [Barnett et al., 2001; Sun and Hansen, 2003; Gregory et al., 2004]. Note that observed variability for the 0–3000 m layer is generally in better agreement with model results in the NH than in the more poorly observed SH (Figure 12a).

[67] For temperatures in the 0–300 m layer, the observed variability exceeds that of most CMIP2+ control runs (Figure 12b). The exception is the BCM, in which the variability is inflated by the large drift in the control run (Figure 7b). The fact that model-versus-observed variability differences are larger for the 0–300 m layer than for 0–3000 m is probably due to the absence of anthropogenic forcing in the model control runs. As shown by Barnett et al. [2005] and Pierce et al. [2006], the gradual warming induced by historical changes in greenhouse gas forcing is most prominent in the upper 700 m of the global ocean. This warming trend enhances the variability of observed temperatures for the 0–300 m layer.

[68] To assess the contribution of external forcing to model data variability differences, we use the observational dd mask to sample the CMIP2+ climate change signals. The latter are defined by subtracting the contemporaneous state of the control from the climate change experiment. The observed mask (which is only 49 years in length), is applied to the first 49 years of the climate change signal. For both depth integrals considered here (0–300 and 0–3000 m), the trend in volume-averaged temperature in $T'_{Sub}(t)$ is invariably larger than in $T'_{Net}(t)$ (Table 3). This is probably due to preferential sampling of the larger warming of upper layers of the ocean. Note also that the simulated ocean temperature trends in the CMIP2+ perturbed runs are almost always larger than the corresponding trends in the WOA-2004.

[69] Figure 12c compares standard deviations of $\Delta T'_{Sub}(t)$ calculated from both the control and 1% CO$_2$ run data. The

Table 3. Effect of Subsampling on Linear Trends in Volume-Mean Temperature Change ($\times 10^{-3}$°C per year) for WOA-2004 and for the CMIP2+ 1% CO_2 Runs[a]

Depth Range	Model	Global Ocean		N. Hemisphere		S. Hemisphere	
		$T'_{Net}(t)$	$T'_{Sub}(t)$	$T'_{Net}(t)$	$T'_{Sub}(t)$	$T'_{Net}(t)$	$T'_{Sub}(t)$
0–300 m	Obs.	0.304	0.349	0.338	0.360	0.279	0.399
	BCM	0.557	0.714	0.594	0.762	0.532	0.616
	CCCma_CGCM2	1.337	1.444	1.284	1.426	1.373	1.528
	CSIRO_Mk2	1.177	1.519	1.367	1.626	1.044	1.272
	GFDL_R30	0.977	1.158	1.038	1.160	0.935	1.152
	HadCM3	1.189	1.503	1.220	1.553	1.168	1.352
	ECHO_G	0.830	1.134	1.059	1.270	0.670	0.888
	MRI CGCM2.3	1.107	1.468	1.305	1.624	0.967	1.185
	PCM	0.470	0.587	0.445	0.583	0.488	0.493
0–3000 m	Obs.	0.095	0.169	0.126	0.146	0.073	0.202
	BCM	0.141	0.309	0.208	0.355	0.094	0.260
	CCCma_CGCM2	0.330	0.708	0.340	0.678	0.323	0.733
	CSIRO_Mk2	0.240	0.574	0.285	0.620	0.209	0.525
	GFDL_R30	0.210	0.441	0.212	0.407	0.208	0.479
	HadCM3	0.217	0.551	0.221	0.444	0.215	0.668
	ECHO_G	0.184	0.336	0.252	0.437	0.136	0.220
	MRI CGCM2.3	0.228	0.590	0.269	0.656	0.199	0.514
	PCM	0.126	0.220	0.138	0.245	0.118	0.198

[a]Values shown for observations are the $T_{Net}(t)$ and $T_o(t)$ trends from Table 1.

control run variability is generally lower than in observations, while the variability in the 1% CO_2 run is almost always higher than observed results. Since linear forcing leads to a nearly linear temperature response, the warming trend in the 1% CO_2 run inflates the standard deviation relative to that in the control run. This inflation is approximately described by

$$\hat{s} \approx \sqrt{\left[s^2 + \left((\Delta T)^2 \right)/12 \right]} \qquad (3)$$

where ΔT is the total trend and s is the standard deviation of the control run data. This overestimate in variability arises because the linear forcing change due to a 1% per year CO_2 increase is likely to be substantially larger than the total forcing that has actually occurred over 1955–2003 (although the latter is uncertain, primarily because of uncertainties in aerosol forcing [*Ramaswamy et al.*, 2001]). It is therefore preferable to perform model-observed variability comparisons with climate change experiments that use more realistic historical forcing [see, e.g., *Barnett et al.*, 2005; *Pierce et al.*, 2006].

4.7. The Vertical Profile of Variability

[70] In section 4.6, we showed that the variability of 0–300 m temperatures in the 1% CO_2 runs is substantially higher than in the control runs. This is a necessary consequence of the addition of a near-linear warming signal. Over time, this signal penetrates into the deeper ocean with magnitude that decreases with depth [see *Gleckler et al.*, 2006]. This leads to an enhancement of variability that also decreases with depth. Such behavior is illustrated in Figure 13, which shows the vertical profile (down to 700 m) of the temporal standard deviation of $\Delta T'_{Sub}(t)$ from the control and 1% CO_2 runs. The inclusion of a CO_2-induced warming signal markedly increases the variability in the upper layers of the global ocean. As noted in section 4.6, the higher than observed variability in the 1% CO_2 runs (Figure 13b) is partly related to the fact that these

experiments involve changes in radiative forcing that are larger than observed.

5. Conclusions

[71] The aim of this study has been to investigate differences in the variability of ocean temperatures in observational estimates (such as WOA-2004) and in coupled atmosphere-ocean climate models. We find that

[72] 1. Sparse data coverage has led to inflated estimates of observed temperature variability in virtually all ocean basins.

[73] 2. To study observed ocean heat content changes, spatially complete temperature fields are required. These are generated by infilling temperatures in data-sparse regions. The infilling method used in WOA-2004 may bias the statistical properties of the temperature data.

[74] 3. To circumvent problems associated with statistical infilling procedures, it is preferable to compare modeled and observed changes in volume–average temperature only at locations and depths where observations exist (rather than to compare changes in total ocean heat content). The *dd* (data distribution) fields in the WOA-2004 provide information that facilitates the subsampling of spatially complete model data.

[75] We used a suite of model control and 1% CO_2 runs from the CMIP2+ experiment to investigate the impact of observational data coverage changes on the simulated variability of ocean temperatures. Temperatures were averaged over the top 300 and 3000 m of the ocean. This subsampling study yielded the following key findings:

[76] 1. Subsampling spatially complete model control run data with the data "mask" of actual observational coverage amplifies the temporal variability of the model ocean temperature data. This increase in variability is a robust result. It occurs in all CMIP2+ models, in both the 0–300 and the 0–3000 m layers, and in virtually all ocean basins. Subsampling brings model-based variability estimates into better accord with observations.
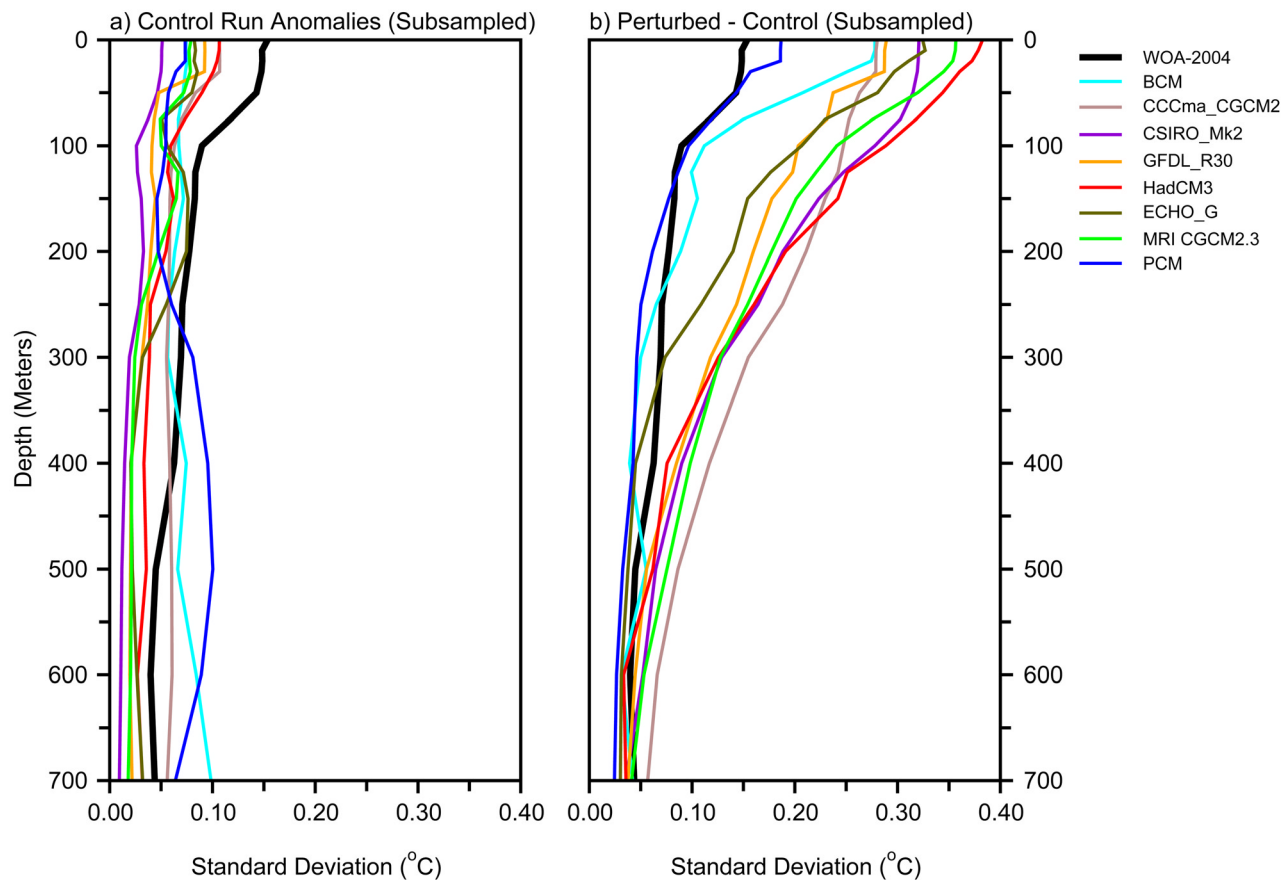
GLOBAL Ocean (Annual Data)

Figure 13. Vertical profile of the temporal standard deviations of temperature anomalies in the top 700 m of the global ocean. Results are for the (a) subsampled control and (b) 1% CO_2 runs of the eight CMIP2+ models. Standard deviations were calculated from the first 49 years of the $\Delta T'_{Sub}(t)$ time series. The standard deviation of $\Delta T_o(t)$ values from WOA-2004 (thick black line) is overlaid on each figure.

[77] 2. The causes of the enhanced temporal variability introduced by subsampling model temperature anomalies differ for the 0–300 and 0–3000 m layers. For the former, the variability increase is mainly due to changes in the areal coverage of observations. For the 0–3000 m layer, the larger variability is due to changes in both spatial extent and the depth of observational coverage.

[78] 3. The CMIP2+ runs with (compounded) annual CO_2 increase of 1% per year provide estimates of the ocean warming signal arising from increases in greenhouse gases. This warming signal is largely confined to the uppermost 1000 m of the oceans. It introduces a trend that enhances the simulated variability in the temperature of the 0–3000 m and 0–300 m layers.

[79] Our investigation has shown that using volume-averaged temperature at the actual location of observations partially explains the apparent mismatch between modeled and observed variability of ocean temperatures. However, the idealized climate change experiments analyzed here do not allow us to determine whether the remaining discrepancy in variability is primarily due to model error. The idealized greenhouse gas forcing in the CMIP2+ 1% CO_2 runs leads to consistently larger than observed ocean warm-

ing trends, which amplifies the simulated variability. In the real world, atmospheric CO_2 has not increased at a rate of 1% per year, and other forcings (such as the cooling effects of anthropogenic sulfate aerosol particles and volcanic eruptions) have offset some of the greenhouse gas induced warming of the world's oceans. The experiments recently performed in support of the IPCC's Fourth Assessment Report, which include more realistic historical changes in natural and anthropogenic forcings, are more appropriate for direct comparison with observations.

References

Allison, L. (2004), Analysis of ocean heat content change in ENACT, and study into the effect of data sampling resolution on an eddy-permitting ocean model (HadCEM), *Hadley Centre Tech. Note 58*, 15pp., Hadley Centre, Met Office, Exeter, U. K.

Barnett, T. P., D. W. Pierce, and R. Schnur (2001), Detection of anthropogenic climate change in the world's oceans, *Science*, *292*, 270–274, doi:10.1126/science.1058304.

Barnett, T. P., D. W. Pierce, K. AchutaRao, P. J. Gleckler, B. D. Santer, J. M. Gregory, and W. M. Washington (2005), Penetration of human-induced warming into the world's oceans, *Science*, *309*, 284–287, doi:10.1126/science.1112418.

Carton, J. A., G. Chepurin, X. Cao, and B. S. Giese (2000a), A simple ocean data assimilation analysis of the global upper ocean 1950–1995. Part I: Methodology, *J. Phys. Oceanogr.*, *30*, 294–309.

Carton, J. A., G. Chepurin, and X. Cao (2000b), A simple ocean data assimilation analysis of the global upper ocean 1950–1995. Part II: Results, *J. Phys. Oceanogr.*, *30*, 311–326.

Duffy, P. B., C. Doutriaux, I. K. Fodor, and B. D. Santer (2001), Effect of missing data on estimates of near-surface temperature change since 1900, *J. Clim.*, *14*, 2809–2814.

Forest, C. E., M. R. Allen, P. H. Stone, and A. P. Sokolov (2000), Constraining uncertainties in climate models using climate change detection techniques, *Geophys. Res. Lett.*, *27*(4), 569–572.

Gent, P. R., and G. Danabasoglu (2004), Heat uptake and the thermohaline circulation in the Community Climate System Model, version 2, *J. Clim.*, *17*, 4058–4069.

Gent, P. R., F. O. Bryan, G. Danabasoglu, K. Lindsay, D. Tsumune, M. W. Hecht, and S. C. Doney (2006), Ocean chlorofluorocarbon and heat uptake during the 20th century in the CCSM3, *J. Clim.*, in press.

Gleckler, P. J., K. AchutaRao, and K. R. Sperber (2004), The mean state of the ocean and cryosphere, in *An Appraisal of Coupled Climate Model Simulations*, edited by D. Bader, *Program Clim. Model Diagnosis Intercomparison Publ. UCRL-TR-202550*, chap. 5, Program for Clim. Model Diagnosis and Intercomparison, Lawrence Livermore Natl. Lab., Livermore, Calif. (available at http://www-pcmdi.llnl.gov).

Gleckler, P. J., T. M. L. Wigley, B. D. Santer, J. M. Gregory, K. AchutaRao, and K. E. Taylor (2006a), Volcanoes and climate: Krakatoa's signature persists in the ocean, *Nature*, *439*, 675, doi:10.1038/439675a.

Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood (2000), The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, *Clim. Dyn.*, *16*, 147–168.

Gregory, J. M., R. J. Stouffer, S. C. B. Raper, P. A. Stott, and N. A. Rayner (2002), An observationally based estimate of the climate sensitivity, *J. Clim.*, *15*, 3117–3121.

Gregory, J. M., H. T. Banks, P. A. Stott, J. A. Lowe, and M. D. Palmer (2004), Simulated and observed decadal variability in ocean heat content, *Geophys. Res. Lett.*, *31*, L15312, doi:10.1029/2004GL020258.

Hansen, J., et al. (2002), Climate forcings in Goddard Institute for Space Studies SI2000 simulations, *J. Geophys. Res.*, *107*(D18), 4347, doi:10.1029/2001JD001143.

Hansen, J., et al. (2005), Earth's energy imbalance: Confirmation and implications, *Science*, *308*, 1431–1435, doi:10.1126/science.1110252.

Hegerl, G. C., and N. L. Bindoff (2005), Warming the world's oceans, *Science*, *309*, 254–255, doi:10.1126/science.1114456.

Ishii, M., M. Kimoto, and M. Kachi (2003), Historical ocean subsurface temperature analysis with error estimates, *Mon. Weather Rev.*, *131*, 51–73.

Jenkins, G. M., and D. G. Watts (1968), *Spectral Analysis and Its Applications*, Holden-Day, Boca Raton, Fla.

Levitus, S., T. P. Boyer, M. E. Conkright, T. O'Brien, J. Antonov, C. Stephens, L. Stathoplos, D. Johnson, and R. Gelfeld (1998), *World Ocean database 1998*, vol. 1, *Introduction, NOAA Atlas NESDIS 18*, NOAA, Silver Spring, Md.

Levitus, S., J. I. Antonov, T. P. Boyer, and C. Stephens (2000), Warming of the world ocean, *Science*, *287*, 2225–2229, doi:10.1126/science.287.5461.2225.

Levitus, S., J. I. Antonov, J. Wang, T. L. Delworth, K. W. Dixon, and A. J. Broccoli (2001), Anthropogenic warming of Earth's climate system, *Science*, *292*, 267–270, doi:10.1126/science.1058154.

Levitus, S., J. Antonov, and T. Boyer (2005), Warming of the world ocean, 1955–2003, *Geophys. Res. Lett.*, *32*, L02604, doi:10.1029/2004GL021592.

Meehl, G. A., G. J. Boer, C. Covey, M. Latif, and R. J. Stouffer (2000), The Coupled Model Intercomparison Project (CMIP), *Bull. Am. Meteorol. Soc.*, *81*, 313–318.

Mitchell, J. F. B., et al. (2001), Detection of climate change and attribution of causes, in *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J. T. Houghton et al., pp. 695–738, Cambridge Univ. Press, New York.

Pierce, D. W., T. P. Barnett, K. AchutaRao, P. J. Gleckler, B. D. Santer, J. M. Gregory, and W. M. Washington (2006), Anthropogenic warming of the oceans: Observations and model results, *J. Clim.*, in press.

Ramaswamy, V., et al. (2001), Radiative forcing of climate change, in *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J. T. Houghton et al., chapter 6, Cambridge Univ. Press, New York.

Reichert, B. K., R. Schnur, and L. Bengtsson (2002), Global ocean warming tied to anthropogenic forcing, *Geophys. Res. Lett.*, *29*(11), 1525, doi:10.1029/2001GL013954.

Santer, B. D., et al. (2000a), Interpreting differential temperature trends at the surface and in the lower troposphere, *Science*, *287*, 1227–1232, doi:10.1126/science.287.5456.1227.

Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor (2000b), Statistical significance of trends and trend differences in layer-average atmospheric temperature time series, *J. Geophys. Res.*, *105*(D6), 7337–7356.

Santer, B. D., et al. (2003), Influence of satellite data uncertainties on the detection of externally forced climate change, *Science*, *300*, 1280–1284, doi:10.1126/science.1082393.

Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C. N. Hill, and J. Marshall (2002), Global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model, *J. Geophys. Res.*, *107*(C9), 3118, doi:10.1029/2001JC000888.

Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C. N. Hill, and J. Marshall (2003), Volume, heat, and freshwater transports of the global ocean circulation 1993–2000, estimated from a general circulation model constrained by World Ocean Circulation Experiment (WOCE) data, *J. Geophys. Res.*, *108*(C1), 3007, doi:10.1029/2001JC001115.

Stephens, C., J. I. Antonov, T. P. Boyer, M. E. Conkright, R. A. Locarnini, T. D. O'Brien, and H. E. Garcia (2002), *World Ocean Atlas 2001*, vol. 1, *Temperature, NOAA Atlas NESDIS*, vol. 49, NOAA, Silver Spring, Md.

Sun, S., and J. E. Hansen (2003), Climate simulations for 1951–2050 with a coupled atmosphere-ocean model, *J. Clim.*, *16*, 2807–2826.

Washington, W. M., et al. (2000), Parallel climate model (PCM) control and transient simulations, *Clim. Dyn.*, *16*, 755–774.

————————————

K. M. AchutaRao, P. J. Gleckler, B. D. Santer, and K. E. Taylor, Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, L-103, P.O. Box 808, Livermore, CA 94550, USA. (achutarao1@llnl.gov)

T. P. Barnett and D. W. Pierce, Scripps Institution of Oceanography, 9500 Gilman Drive, M.S. 0224, La Jolla, CA 92093-0224, USA.

T. M. L. Wigley, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000, USA.