

Anthropogenic Warming of the Oceans: Observations and Model Results

David W. Pierce*

Tim P. Barnett

*Climate Research Division, Scripps Institution of Oceanography
La Jolla, California*

Krishna M. AchutaRao

Peter J. Gleckler

*Lawrence Livermore National Laboratory
Livermore, California*

Jonathan M. Gregory

*University of Reading
Reading, U.K.*

Warren M. Washington

*National Center for Atmospheric Research
Boulder, Colorado*

6 October 2005 (Version 2)

Submitted to *J. Climate*

*Corresponding Author. Address:

Climate Research Division
Mail Stop 0224
Scripps Institution of Oceanography
voice: (858) 534-8276
fax: (858) 534-8561
dpierce@ucsd.edu

Abstract

Observations show the oceans have warmed over the past 40 years, with appreciable regional variation and more warming at the surface than at depth. Comparing the observations with results from two coupled ocean-atmosphere climate models (PCM and HadCM3) that include anthropogenic forcing shows remarkable agreement between the observed and model-estimated warming. In this comparison the models were sampled at the same locations as gridded yearly observed data. In the top 100 m of the water column the warming is well separated from natural variability, including both variability arising from internal instabilities of the coupled ocean-atmosphere climate system and that arising from volcanism and solar fluctuations. Between 125 and 200 m the agreement is not significant, but then increases again below, and remains significant down to 600 m. Analysis of PCM's heat budget indicates the warming is driven by an increase in net surface heat flux that reaches 0.7 watts m^{-2} by the 1990s; the downward longwave flux increases by 3.7 watts m^{-2} , which is not fully compensated by an increase in the upward longwave flux of 2.2 watts m^{-2} . Latent and net solar heat flux each decrease by about 0.6 watts m^{-2} . The changes in the individual longwave components are distinguishable from the pre-industrial mean by the 1920s, but due to cancellation of components, changes in the net surface heat flux do not become well separated from zero until the 1960s. Changes in advection can also play an important role in local ocean warming due to anthropogenic forcing, depending on the location. The observed sampling of ocean temperature is highly variable in space and time, but sufficient to detect the anthropogenic warming signal in all basins, at least in the surface layers, by the 1980s.

1. Introduction

Observations show the climate is warming. Nearly a decade's worth of rigorous detection and attribution studies (e.g. IPCC 2001, IDAG 2005), focused mainly on atmospheric temperature, show the warming is outside the range of natural variability. Natural variability includes the effects of both fluctuations that arise within the coupled ocean-atmosphere system due to oscillations or instabilities ("internal" variability), and effects imposed on the climate from natural sources external to the oceans and atmospheres, such as volcanism and solar fluctuations ("external" variability). The warming's effects can be seen qualitatively in accelerating melting of glaciers (e.g. Thompson 2000, 2003), disruption of Arctic ecosystems (e.g. Tynan and DeMaster (1997)), and bleaching of the world's coral reefs (Knowlton 2001), to name a few examples.

The ocean's density and heat capacity are orders of magnitude greater than the atmosphere's; it is the thermal flywheel that stabilizes the climate system. Indeed, 84% of observed climate warming of the Earth system (oceans, atmosphere, cryosphere, continents) over the past 50 years has gone to heat the oceans; only 3% has gone to heat the atmosphere (Levitus et al. 2005). The global ocean is therefore the logical place to study the nature of observed warming.

Detection and attribution studies of ocean warming using climate models were pioneered by Barnett et al. (2001) using the Parallel Climate Model version 1 (PCM), and Levitus et al. (2001) using the Geophysical Fluid Dynamics Laboratory (GFDL R30) model, who compared simulated ocean warming to the Levitus et al. (2000) observed ocean temperature data set. Both found a good match between modeled and observed increase in ocean heat content, but only if anthropogenic forcing was included. The conclusion was that observed ocean warming is outside the expected range of natural variability, and attributable to anthropogenic forcing. Subsequently the issue was examined by Reichert et al. (2002) using the European Center/Hamburg model (ECHAM4/OPYC3); Gent and Danabasoglu (2004) with the Community Climate System

Model version 2 (CCSM2); and Gregory et al. (2004) with the Hadley Centre coupled climate model version 3 (HadCM3). All found good agreement between model-predicted and observed ocean warming, although sampling issues, somewhat weak model variability, and the way the signal could be more easily detected in the top part of the water column than at greater depths were generally noted. None of these studies examined in detail the vertical structure of the warming signal. Hansen et al. (2002) and Sun and Hansen (2003), although not formal detection and attribution studies, examined the anthropogenically forced increase in ocean heat content in the Goddard Institute for Space Studies atmosphere model (GISS SI2000) coupled to various ocean models (including HYCOM), and showed globally-averaged vertical profiles of anthropogenic ocean warming. Gregory (2000) showed zonal-mean ocean temperature increases in an anthropogenically forced model experiment, and analyzed the role of ocean vertical heat transport (particularly at high latitudes) in producing the surface climate response.

More recently, Barnett et al. (2005; B05 hereafter) revisited this issue with a new version of the observed, gridded ocean temperature data set (Levitus et al. 2005). Compared to Barnett et al. (2001), B05 also included a basin-scale analysis of the temporal evolution of vertical penetration of heat into the oceans; an estimate the role of time-averaged surface heat flux and advection in contributing to the temperature change; detection with estimated historic solar and volcanic forcing taken into account; results from a different climate model (HadCM3, Gordon et al. 2000); and a more restrictive sampling scheme that better mimics observations (more on this below). The conclusion of B05 was that human activities are largely responsible for the observed ocean warming over the last 40 years. Natural variability, either internal to the coupled ocean-atmosphere system or external (solar/volcanic), was incapable of explaining the ocean observations.

Because of their brevity, Barnett et al. (2001) and B05 did not fully address a number of questions attending their main conclusion. Our objective is to more fully explore these questions and determine their potential impact on the detection and attribution results for ocean warming. Accordingly, this paper is organized as follows. In section 2 we briefly

describe the data and models used. In section 3 we examine the models' natural variability and compare it to observations; a reasonably close correspondence is needed if natural variability is to be accepted or ruled out as a potential cause of observed ocean warming. In section 4 we illustrate a model-based "fingerprint" of ocean warming and compare it to both observations and model estimates of natural internal and external variability. In section 5 we illustrate how changes in net surface heat flux and advection together determine changes in local heat storage. In section 6 we examine the effects sampling variability has on the detection and attribution results. Finally, in section 7 we present a summary of our findings and conclusions.

2. Model and Observational Data Used

a. Models

We use data from two O-AGCMS: the Parallel Coupled Model (PCM; Washington et al. 2000) and HadCM3 (Gordon et al. 2000). These two models have been developed independently for decades, so they have little in common except their purpose. Neither uses flux adjustments. Complete descriptions of both these models and their simulations are available in the above references, so they will only be summarized briefly here.

The atmospheric component of PCM is the CCM3 atmospheric general circulation model (Kiehl et al. 1998), a spectral model used here at T42 resolution (equivalent to about 280 by 280 km grid spacing), with 18 vertical layers in a hybrid sigma coordinate scheme. A land surface model with soil moisture, vegetation types, and a simplified runoff scheme is included. The ocean component is the Parallel Ocean Program (POP; Smith et al. 1992, Dukowicz and Smith 1994), with a horizontal resolution of 384 by 288 points on a displaced-pole grid (roughly $2/3^\circ$ resolution), with 32 vertical levels clustered near the surface to improve the representation of surface mixing processes. A dynamic-thermodynamic sea-ice model based on Zhang and Hibler (1997) is included, with an elastic-viscous-plastic rheology for computational efficiency (Hunke and Dukowicz 1997). The ice model is formulated on its own grid, which has a total of 320 by

640 grid points, with a physical grid spacing of roughly 27 km.

To determine the effect of anthropogenic forcing on climate, we used a 12 member ensemble (listed in Fig. 6) forced with estimated historical emissions of greenhouse gasses, ozone, and the direct effect of sulfate aerosols (Washington et al. 2000; Dai et al. 2004). These runs did not include the effects of black carbon, solar fluctuations, or volcanoes. We compared this to a 520 year control run with greenhouse gas and sulfate aerosol concentrations fixed at 1870 values, which provides an estimate of natural internal climate variability. For isolating the effects of natural variability external to the coupled ocean-atmosphere system, we used a four member ensemble driven only by estimated solar fluctuations and volcanic forcing over the past 50 years (Meehl et al. 2003).

HadCM3 is based on the UKMO unified forecast and climate model. The atmospheric resolution is 2.5° by 3.75° , with 19 vertical levels in a hybrid vertical coordinate. It represents effects of minor trace gases in addition to carbon dioxide, water vapor, and ozone. The direct and some indirect effects of aerosols are included (Gordon et al. 2000).

The ocean component of HadCM3 uses a 1.25° by 1.25° horizontal resolution, with 20 vertical levels concentrated in the upper part of the water column to better resolve surface processes. This ocean model shows considerably improved heat budgets compared to earlier, coarser resolution versions. A simple thermodynamic sea-ice model is included, with parameterizations of ice drift and leads. A four member ensemble was used (listed in Fig. 7) that incorporated effects of both anthropogenic and natural external (solar, volcanic) forcing. This differs from the PCM anthropogenically forced runs, which did not include solar and volcanic forcing. A 700 year control run with no external or anthropogenic forcing provided an estimate of natural internal variability. Tett et al. (2002) and Johns et al. (2003) analyze anthropogenic runs using this model.

b. Observations

We use the latest version of the National Oceanographic Data Center's gridded historical ocean temperature data set (Levitus et al. 2005), available at http://www.nodc.noaa.gov/OC5/DATA_ANALYSIS/heat_intro.html (accessed 17 Mar

2005). Using this data, Levitus et al. (2005) estimated the world ocean heat content (surface to 3000 m) increased 14.5×10^{22} J between 1955 and 1998.

The data is on a 1° by 1° grid, and has been infilled (i.e., interpolated) to fill gaps where no observations exist. Yearly data is provided on 16 standard depth levels in the top 700 m.

Locations where the ocean is sampled vary considerably from year to year. To help take this into account, two additional fields are supplied with the data set. The *gp* mask shows, for each grid box, the number of observations taken within ~ 450 km of that grid box at the same depth. (This is a change from the *gp* field in Levitus et al. (2000), where a radius of ~ 900 km was used.) The *dd* mask shows the number of observations taken within the grid box itself. Both masks vary in latitude, longitude, depth, and time. For most results shown here, only locations the *dd* mask indicates as sampled are included, for both models and observations. This is a departure from Barnett et al. (2001) and Reichert et al. (2002), who used the *gp* mask; this change is an attempt to be more conservative, by avoiding using any infilled data. It also motivates changing from total heat content, used in Barnett et al. (2001), to volume averaged temperature, since the total amount of heat depends on the volume sampled in any particular year. Gregory et al. (2004) (their Fig. 2) illustrate how total estimated heat content can vary substantially based on assumptions made to infill data. (Gent and Danabasoglu 2004 used volume averaged temperature as well, but did so to normalize by varying basin sizes and avoid use of heat content units.) In section 6a we compare the results of using *dd* and *gp* masks.

Pentads (5-year averaged data) are provided on the same horizontal grid as yearly data, but extended to 3000 m using 28 depth levels. We focus on yearly data in this work, since we use the *dd* mask to sample the model in the same way as the observations, and potentially having a single sample represent temperature anomalies for five years becomes uncomfortably long, even at depth. The issue of seasonal sampling of the observations, and how it might influence the annual anomalies, is discussed in section 6d.

3. Natural Variability: Models vs. Observations

It is important to estimate the likelihood that observed ocean warming arises from natural climate variability, both internal and external. To do this one must compare observed ocean warming to fluctuations expected without anthropogenic forcings. The latter requires estimates from models; it then becomes important to examine whether the models' natural variability is realistically strong. We do this by comparing control model results to the detrended observations. The detrended observations also include natural external variability (solar and volcanic forcing) and any non-linear effects of anthropogenic forcing, unlike model control runs. To address this, we also compare detrended observations to detrended anthropogenically forced model runs. (The separate effect of external forcing in the absence of anthropogenic forcing is shown for PCM in Section 4.)

Perhaps the first and most obvious way to compare variability is simply through maps of standard deviation of annual sea surface temperature (SST) anomalies; this is shown in Fig. 1 for observations (1945-2004) and the long model control runs. Both the observed and model values are detrended by removing the least squares best fit line (although there is very little trend in the model control run surface temperatures). After detrending, solar and volcanic variability (along with any non-linear response to anthropogenic forcing) remains in the observations, but is absent from the model control runs. Even so, there is some tendency for the models to overestimate SST variability. For example, PCM has too much North Pacific variability, as the Pacific Decadal Oscillation in PCM (standard deviation 1.1°C) is stronger than observed (0.74°C). HadCM3 has more North Atlantic variability than observed. Both models have roughly the observed strength in the central tropical Pacific (associated with ENSO), although PCM misses the high variance along the west coast of South America and HadCM3 extends the variability too far into the tropical Pacific warm pool.

The standard deviation of annually averaged values is useful to examine, but strong interannual variability might mask low-frequency differences. Time series and spectra of PCM and HadCM3 temperatures (0-700 m) from the control runs (internal variability

alone) are already compared to detrended observations in B05, and so not repeated here. However, it is unclear if anthropogenic forcing has any effect on the amplitude or frequency of natural modes of internal climate variability, such as ENSO, the Pacific Decadal Oscillation, or the North Atlantic Oscillation. Accordingly, in Fig. 2 we show spectra of globally averaged temperature (0-700 m, *dd* sampled points only) in observations and anthropogenically-forced model runs, both detrended by removing the least squares best fit line to remove the linear part of any anthropogenic signal. (Prior to detrending the model values have control model drift removed as described in Appendix A; Gent and Danabasoglu 2004 illustrate and discuss drifts that can be found in a control run, albeit for a different model than used here.)

The models and observations have similar power in interannual frequencies (> 0.15 cy yr^{-1}), but PCM becomes deficient at longer timescales. This is less a problem with HadCM3, which has more low-frequency variability than PCM, and matches the observed spectrum rather well. A more complete spectral analysis is shown in Appendix B, where the same data used to make Figure 2 is analyzed by basin and depth. The conclusion is low-frequency spectral power in the detrended observations generally falls within the 90% confidence limits of the PCM model ensembles, while HadCM3 tends to have too much variability above 100 m and too little below 400 m. This would have the tendency of making any anthropogenic warming of the surface ocean (where the signal is concentrated) harder to detect than it should be.

Gregory et al. (2004) compared observed low frequency temperature variability to the HadCM3 control run (natural internal variability only) by computing the standard deviation of volume averaged, 5-yr running mean temperature at each level from the detrended data (their Fig. 4). They found a peak in observed variability at 400 m that the model did not reproduce. The detailed analysis of standard deviations in Appendix B shows that the PCM control run exhibits the same behavior, but also that the observed 400 m “peak in variability” arises not because ocean variability is unusually strong at 400 m, but because it is unusually spatially coherent. This reduces cancellation between

uncorrelated points seen to a greater extent above and below 400 m when volume-average temperature is calculated. The models' average standard deviation of temperature by basin is rather similar to the observations, as shown in Fig. 3. The black dots are observations; the range from HadCM3's control run is shown as the gray shaded region, while PCM's control run is shown as the crosshatched region. Both models have somewhat weak internal variability in the South Atlantic below 250 m and in the Indian Ocean around 100 m. In other basins and depths the observed curve generally falls within the range of HadCM3 ensembles, while PCM tends to show overly strong variability.

The analyses discussed so far use detrended data, yet any anthropogenic response might be a near-linear trend over the time period. It is therefore useful to show a simple comparison of the linear trends (1960-1999) found in the observations, control model runs, and anthropogenic model runs. This is shown in Fig. 4 for PCM and Fig. 5 for HadCM3; all data has been sampled with the *dd* mask. The results vary by model and basin, but generally the observed 40-yr trends fall within the trends seen in the anthropogenically forced runs, and outside the trends seen in the control run ensemble members in the upper 100 m of the water column. This is confirmed by a K-S test of the trends from the control and anthropogenic ensemble members (results shown on the plot).

In summary, we have performed a number of comparisons of the observed data with the model control and anthropogenic runs to evaluate the models' ability to reproduce climate variability. Comparing the spectra of the detrended observations to the detrended anthropogenically forced runs shows PCM's level of variability is in general agreement with observations; HadCM3 tends to have too much variability above 100 m, and too little below 400 m. Comparing low-frequency standard deviations of temperature to the observations shows the model control runs bracket the observed values, but lack large-scale coherence in temperature fluctuations around 400 m in the southern hemisphere western Pacific. However, B05 (their Fig. 2) show no detectable warming signal at that location and depth, so the implication of this model shortcoming for the detection issue is moot. Comparing simple linear trends in the observations to the control

and anthropogenic model runs shows the observed trends are consistent with those found in the anthropogenic runs, and well separated from those found in the control runs in the upper part of the water column. None of these analyses indicate the models have systematically weak variability in the top 100 m (where the ocean warming signal is greatest). We conclude that the models provide a good estimate of natural internal and external variability for detection and attribution studies.

4. Ocean Warming Fingerprint

The time series for yearly volume average temperature anomaly (0-700 m) are shown for PCM in Fig. 6 and HadCM3 in Fig. 7. Left panels show results from the anthropogenically forced ensemble members, right panels show same-length chunks from the control run. Anomalies are calculated relative to the base period of 1957-1990 to match the treatment of the observations (Levitus et al. 2005). For brevity, in the rest of this work we will refer to the yearly, volume averaged temperature anomaly as the “temperature”. The values in Figs. 6 and 7 are taken from sampled points only, as indicated by the *dd* mask. The period shown is 1955-1999, in deference to PCM runs, which end in 1999. (Levitus et al. (2005), their Fig. 1, shows a similar graph for observed heat content that extends through 2003.)

Both PCM and HadCM3 show a great deal of temporal variability, which differs significantly between ensemble members. For instance, PCM realizations B05.03 and B06.08 show near-constant temperatures over the last 12 years of record, while B06.28, B06.85, and B06.97 show cooling in the final years of the record. Most ensembles show multi-year periods where temperature is cooler than a previously attained value. Such variability can be seen in observations as well, although the decadal swing from 1970 to 1980 is larger than anything seen in the PCM ensemble members (but comparable to swings seen in HadCM3’s *abw1* and *abw2*). The consistent strong trends seen in the observations and anthropogenically forced runs of both models are not found in the control runs, which lack anthropogenic forcing (right panels of Figs. 6 and 7).

The temperature series by region for PCM are shown in Fig. 8. Individual ensemble

members (gray lines) show a great deal of variability, both year-to-year and between realizations. The observations (thick solid line) are consistent with individual ensemble members in the amplitude of year-to-year variability and in the overall trend by basin. Taking the average over ensembles (thick dashed line) removes much of the higher frequency variability not of interest here, but still worth showing to illustrate the model generates basin-wide variability consistent with observations. The anthropogenic forcing increases gradually and monotonically during this time period, so for the rest of the analysis we average by decade to reduce effects of high-frequency internal natural climate variability.

a. Definition of Fingerprint

We desire to define a signal that concisely expresses both spatial and temporal ocean warming found in the model runs with anthropogenic forcing. Given such a fingerprint, we can examine how likely it is to find this pattern in the observations, unforced control run (natural internal variability), and runs with natural external forcing (solar fluctuations and volcanoes).

We first remove the influence of model drift by differencing temperature in the forced ensemble runs from the simultaneous low-frequency fitted temperature in the control run (details in Appendix A). We then define an ensemble common signal (ECS) for an ensemble of n members as follows. For each member we concatenate, by ocean basin, the 4 decadal temperature values to form $C(n, \hat{t})$. Decades are taken as 1960-1969, etc. Here \hat{t} can take on 24 values (4 decades \times 6 basins; the x axis in Fig. 9). We do this independently for each of the 16 standard depth levels in the top 700 m (the maximum depth at which yearly data is given in the observed data set). To get the final ECS, we use standard principal component analysis (PCA) to decompose C :

$$C(n, \hat{t}) = \sum_m E(n, m)P(m, \hat{t}) \quad (1)$$

where m is the mode number, E is the empirical orthogonal function, and P is the principal component. We take the leading principal component, $P(1, \hat{t})$, as the ECS.

The ECS can be formed from any set of ensemble members, such as the PCM or HadCM3 ensemble members alone. Here, we define our ensemble common model fingerprint of ocean warming (or simply “fingerprint”) as the ECS of the 12 PCM ensemble members taken together with the 4 HadCM3 ensemble members. It thus represents the common warming signal in all the ensembles available from the two models. The fingerprint explains 86% of the variance in the surface layers, dropping to 30% of variance at 700 m. As tests we tried both volume weighting C before PCA analysis, and simply averaging the ensemble members to obtain the fingerprint; neither made much difference. Equally weighting the mean of the 12 PCM ensembles and the mean of the 4 HadCM3 ensembles to form the fingerprint changed it somewhat in the bottom 3 levels (500-700 m), but had little effect higher in the water column, where the signal is concentrated.

The ensemble common model fingerprint at 50 m is shown in Fig. 9 (dashed line with squares). Also illustrated is the same quantity calculated from observations (thick line with dots), and the 16 individual ensemble members that go into making the model fingerprint (thin gray lines). The observed warming matches the model fingerprint quite well. Consistent differences between oceans can be seen; for example, warming has been greater in the North than South Atlantic.

The amplitude of the fingerprint decreases with depth as shown in Fig. 10. Peak-to-peak values drop from 0.3 C near the surface to 0.05 C at 600 m. Again, differences between basins are apparent; the signal drops less in the North and South Atlantic than in the South Pacific and South Indian oceans.

The ensemble common model fingerprint is compared to the observations, and to the models’ anthropogenically forced and control run ensemble members, in Fig. 11. The figure is in the same format used in B05, although here the entire world ocean fingerprint is used instead of individual basins. Briefly, “signal strength” is defined as:

$$S = (F \cdot T) / \|F\| \quad (2)$$

where F is the ensemble common model fingerprint at a given level and T is the target

being compared to, taken at the same level. The signal strength is similar to a correlation, but without normalization by T (equivalently, it is the correlation between F and T multiplied by the standard deviation of T). It retains information on both the agreement in temporal evolution of the two signals and the strength of T . Retention of signal strength is an important point; we would deem a model unsuccessful at reproducing the target signal if it predicted radically too weak or strong a signal, even with a perfect correlation.

Over most of the water column there is good agreement between the observed signal strength and the anthropogenically forced ensemble members, in both PCM and HadCM3. Moreover, in the upper part of the water column there is a clear separation between the observations and the natural internal variability found in the control runs (crosshatched region). Consistent with B05, we conclude that the observed warming cannot be explained by natural internal variability, but can be explained by anthropogenic forcing. The results of section 3 should be kept in mind when forming this conclusion; i.e., it is important to assess a model's natural internal variability, and compare it to the available observations, before the model estimate of natural internal variability can sensibly be compared to the observed warming signal. Here we have done this comparison and found the models' estimation of natural internal variability to be in accord with observations.

The triangles in the left panel of Fig. 11 show the signal strength for the ECS of PCM runs forced by estimated solar and volcanic forcing alone, i.e., natural external variability (triangles). The signal strength in these runs falls within the range of natural internal variability, and distinctly outside the envelope of the anthropogenically forced runs. The conclusion is that natural external variability cannot explain the observed ocean warming. In sum, natural internal variability and solar/volcanic fluctuations are too weak to explain the observed ocean warming signal, but the warming is consistent with that expected from anthropogenic forcing.

The correlation between the ensemble common model fingerprint and the observations is shown in Fig 12. The ensemble common model fingerprint explains 80-90% of the variance of the observed signal over the top 75 m. The correlation drops to a minimum at

depths between 125 and 150 m, possibly because of variability in the thermocline. It then increases again below, such that the model fingerprint is significantly correlated with the observations and explains $\sim 35\%$ of the observed variance between 250 and 600 m.

Once the global fingerprint concatenated by ocean has been calculated, it can be split apart again to perform detection individually by basin. This is illustrated in B05, and so will not be repeated here, although we will note that adding an additional 7 PCM anthropogenically forced realizations to the 5 used in B05 made little difference (not shown).

5. Change in Heat Flux Components

The ocean warming in various basins is accomplished by an increase in net surface heat flux (NSHF) due to anthropogenic forcing, the regional details of which do not appear to have been shown before. Several previous studies have shown globally averaged radiative forcing at various atmospheric levels (e.g. Hansen and et al. (2002), Johns et al. (2003)), while Pierce (2004) showed global surface fluxes. Sun and Hansen (2003) have an interesting analysis of NSHF and ocean heat storage, as well as changes in ocean heat transport in an anthropogenically forced simulation, but do not separate out the contribution of the various surface heat flux component fields. In this section we show the regional anthropogenic change in surface heat flux components and oceanic heat storage estimated by PCM. (It is not possible to show the same quantities from HadCM3 as not all the surface heat flux components were saved.)

a. Net Surface Heat Flux

Figure 13 shows PCM's change in surface heat flux components relative to the 1880-1919 mean (the South Pacific, which is similar to the North Pacific, is omitted). All components have the sign convention that positive values heat the ocean, including LW-surf, the longwave flux emitted from the surface (which is then always negative).

Globally averaged (lower right panel), the change in NSHF relative to pre-industrial conditions reaches about 0.71 ± 0.05 watts m^{-2} by the 1990s. Willis et al. (2004),

analyzing satellite altimetric height data with observed in-situ temperature profiles, estimated a global ocean warming rate of 0.86 ± 0.12 watts m^{-2} from 1993 to 2003, which suggests the model NSHF might be slightly low. The increase in downward longwave (LW) flux from the atmosphere (3.7 watts m^{-2}) is not completely compensated for by an increase in upward LW (2.2 watts m^{-2}) from the surface, resulting in a net LW increase by about 1.5 watts m^{-2} . Incoming shortwave radiation has a negative trend, due to the effects of increasing sulfate aerosols and, possibly, changes in cloud cover. The trend in latent heat flux is also negative, suggesting increased evaporation. The sensible heat flux increases by about 0.6 watts m^{-2} .

The anthropogenically-forced changes in individual heat flux components (particularly LW) are distinguishable from the pre-industrial mean by the 1920s. However, due to near-cancellation of the competing effects, NSHF (and thus the ocean warming) does not become well separated from the pre-industrial mean until the 1960s. This multi-decade lag between the effects of anthropogenic forcing on the atmosphere and on NSHF (and hence surface temperatures) does not seem to have been noted before. It should be noted that the anthropogenically-forced PCM runs used here do not include solar and volcanic forcing, which might alter the times by which the runs can be distinguished.

Individual oceans each have their own signature of how flux components evolve. For example, in the North Indian ocean NSHF is near zero until the 1990s, due to evaporative cooling and a strong decrease in incoming shortwave radiation (c.f. Gent and Danabasoglu 2004, who noted the Indian ocean's lack of monotonically increasing warming in CCSM2 was an "interesting exception" to the other basins). There is also considerable variability between ensemble members, as indicated by the whiskerplots. For example, the North Indian ocean shows ensemble members with a negative NSHF change by the 1990s. Not every region is obliged to exactly reproduce the globally-averaged trends. This variability prevents the same calculation from being sensibly performed for HadCM3, since the heat flux components were only saved for one ensemble member.

b. Advection and Local Heat Storage

The rate of change of heat storage in a basin depends on NSHF and net heat advection into the basin (cf. Gregory (2000), who examine the effects of vertical ocean heat transport on surface temperature change in an anthropogenically forced model run). With the anthropogenically forced PCM ensemble members it is possible to calculate local heat storage's rate of change and compare it to NSHF, thereby estimating the role changes in net advection play in producing ocean warming.

Figure 14 shows time series (averaged by decade) of NSHF and changes in local heat storage. The contribution of advection is then calculated as a residual between the other two. Values are shown relative to the average over 1880-1919, since we are interested in examining anthropogenically forced changes. To facilitate comparison, changes in heat are normalized by surface area to yield an equivalent surface heat flux in watts m^{-2} . The entire depth of the water column is used (without any masking), so the effect of vertical heat transfers is not included.

Advection's role in accomplishing the anthropogenic ocean temperature change can be minimal or important, depending on the region. For example, changes in net advective warming play little role in North Atlantic anthropogenic warming, but significantly warm the North Indian ocean, and cool the South Pacific. The differences between ensemble members (not shown) can be reasonably large.

The trends are also interesting, albeit noisy. The South Pacific shows a steadily decreasing contribution to ocean warming from advection, which partly offsets the NSHF to result in reduced warming. The South Indian ocean, by contrast, shows increasing warming due to advection, and so warms more than would be expected due to NSHF alone.

The conclusion is that advection plays an important role in local ocean warming due to anthropogenic forcing. It would be interesting to explore how the change in advective heating is accomplished; for example, by changes in mean or eddy circulation, or by changes in the temperature difference between water flowing into and out of the basin.

However, this is outside the scope of the present work.

6. Sampling Variability and Model Uncertainty

The fraction of ocean volume sampled to 700 m for various oceans, as given by the *dd* mask, is shown in Fig. 15. Northern hemisphere oceans (solid symbols) are noticeably better sampled than southern hemisphere oceans (open symbols). Even in the northern oceans, the actual locations sampled can vary wildly from year to year. In this section we explore how sampling might potentially affect the detection and attribution results.

Gregory et al. (2004), in their analysis of the possible effects of sampling and data infill on the observed data set, indicated that caution on this point is warranted, since estimates of ocean heat content can vary significantly depending on how infill data is calculated. Additionally, Sun and Hansen (2003) specifically noted that surface heat fluxes estimated from air-sea temperature differences “do not seem to be consistent with” the large decadal-timescale swing in ocean heat content found in the observational data set, and the discrepancy might be due to incomplete sampling. Although we avoided using heat content and sampled only at the same places as the observations partly for these reasons, sampling is clearly an important issue that deserves consideration.

a. *dd* Versus *gp* Masking

We have already noted the difference between the observed data’s *dd* mask (sampled points only; used here and in B05) and the *gp* mask (points within ~450 km of a sample, used by Barnett et al. 2001 and Reichert et al. (2002)). The difference using these masks makes to the detection result (Fig. 11), by level, is shown in Fig. 16. There is little difference between results using either mask. More detailed results with detection by basin (not shown) give differences slightly bigger than seen in Fig. 16, but still not enough to affect the detection and attribution results. Reichert et al. (2002) compared using the *gp* mask to the full infilled data set from Levitus et al. (2001) and similarly found it made little difference to the conclusions in a formal detection and attribution study. This does not necessarily imply there is little difference between the sampled and infilled data itself;

rather, the signal detection strategy employed here (with the concatenated signal applied to Eq. 2) is designed to maximize the signal to noise ratio as much as possible, thereby avoiding some problems with using infill data. Further details on the impact of infilling are given in AchutaRao et al. (2005).

b. Accuracy of Observed Estimates

We have focused on volume averaged yearly temperature anomaly in various oceans. How accurately is this known, given the incomplete sampling?

It is straightforward to estimate this from the model. We use PCM, as there are more ensemble members to work with than with HadCM3 (12 vs. 4). For each year's *dd* sampling mask, we compare the ensemble member's true basin averaged temperature to the value estimated after masking. We then average into decades so the effect of changes in sampling over the years can be more easily seen. The 90% confidence intervals for observational error in estimated volume averaged temperature are shown in Fig. 17, using sampling masks from the 1960s (gray swath) and 1990s (crosshatched region). The intervals are centered around zero; however the purpose of constructing such intervals is to compare them to the signal. Values in Fig. 17 are therefore offset by the decade's mean temperature anomaly with respect to the pre-industrial control run, for the 1960s (white squares) and 1990s (black dots).

For example, in the North Atlantic (top left panel) there has been warming of about 0.1°C by the 1960s. However the sampling masks used in that era could have misestimated the basin average temperature by up to $\pm 0.1^{\circ}\text{C}$, making detection of the warming problematical by that time. By the 1990s the warming is greater and sampling better, leading to a clear separation of measured temperature anomalies from zero at all depths. Examination of the other decades (not shown) indicates the confidence intervals for all basins except the South Atlantic separate from zero (at least in the surface levels) starting in the 1970s; the South Atlantic then becomes separated from zero in the 1980s. The overall conclusion is that, given the evolving sampling density and strength of the signal, sampling is sufficient to confidently detect the anthropogenic warming signal in all

oceans by the 1980s. Detection would be possible even earlier were a formal mechanism for climate change detection used, such as the fingerprint technique in B05, rather than simple decadal differences.

c. Model Heat Uptake Differences

Different climate models show different different ocean heat uptakes (Sokolov et al. 2003). How might this affect the detection and attribution results? The Climate Model Intercomparison Project (CMIP; Meehl et al. 2000) has accumulated a database of coupled climate runs with 1% yearly increasing CO₂ forcing that can be used to examine this question.

The ocean warming experienced after 80 years for a selection of models is shown in Table 1. The difference between the minimum and maximum warming found by basin (expressed here as a ratio) can be large; for example, the model with most warming showed the North Indian ocean increasing in temperature almost 8 times more than the model with least.

A crude estimate of the effect this model uncertainty has on the detection and attribution results can be made as follows. If $R(b)$ is the maximum/minimum ratio from Table 1 for basin b , and $F(t, b)$ is the ensemble common model warming fingerprint (a function of decade t and basin), then we construct estimated signals for the “minimum” and “maximum” models as follows:

$$S_{min}(t, b) = F(t, b) \left(1 - \frac{R - 1}{R + 1}\right) \quad (3)$$

$$S_{max}(t, b) = F(t, b) \left(1 + \frac{R - 1}{R + 1}\right). \quad (4)$$

F is at the midpoint of S_{min} and S_{max} , so the constructed signals are centered on the ensemble common model fingerprint, and at every point $S_{max}/S_{min} \equiv R$. The constructed signals for the 50 m level are shown in Fig. 18. We can then treat these constructed signals as any other signal, using Eq. 2.

The results are shown in Fig. 19. Observations are shown as circles, the standard ensemble common model fingerprint as squares, and the minimum/maximum constructed

signals as the black and white triangles, respectively. The crosshatched region is the 90% confidence interval of the unforced control run from HadCM3, used here to be conservative as it is a wider region than found in PCM (see Fig. 16). Even with the minimum model signal, the forced result (black triangles) falls outside the 90% confidence interval of the unforced control in the upper part of the water column, indicating that natural internal variability would be rejected as an explanation of the warming. Still, the standard model fingerprint is a considerably better match to observations than either the minimum or maximum constructed signals. If various climate models exhibit the same difference in ocean heat uptake during the historical era as they show in the 1% CO₂ increase CMIP runs, this might allow identification of models whose simulations agree poorly with observations. This could then be taken into account when considering their projections of future climate change.

In summary, model uncertainty in heat uptake is significant, but not large enough to affect the detection results presented above and in B05.

d. Seasonality of the Observed Sampling

There is no difficulty forming annually averaged temperature anomalies from the models, since complete monthly data coverage is available. For the observations, however, samples are made at specific times, and there is unlikely to be uniform seasonal coverage in any particular region. To address whether this could affect our analysis, we consider three questions. 1) Does the observed ocean temperature trend vary by season? If not, then even strongly seasonally biased annual anomalies would have little effect on our analysis of the low-frequency signal. 2) Is the seasonal temperature cycle sufficiently well known to correctly compute anomalies from the value expected at the observed time? This is a prerequisite for accurately computing the annual anomalies. 3) Is the ocean warming fingerprint present only in the upper part of the water column affected by the seasonal cycle? If not, then there is little reason to believe that any seasonal bias is affecting our conclusions. These questions will be considered in turn.

1) TEMPERATURE TREND BY SEASON

The best place to examine the seasonality of observed ocean temperature trends is at the surface, since the amplitude of the seasonal cycle decreases with depth. This also allows the long record of SST to be used. The linear trend in SST anomalies, by season, over the most recent 100 years (1905-2004) computed from the Kaplan et al. (1998) extended data set (downloaded from <http://www.cdc.noaa.gov/>) is shown in Fig. 20. The seasonal trends differ in details, but there is broad agreement on the larger scales. The pattern correlations between the various seasons range from 0.58 to 0.92, with a mean of 0.76; all are significant at the 99% level. Similar agreement is seen using the most recent 80- and 50-yr periods (not shown). The lack of large systematic differences in low-frequency trend by season suggests that even if seasonally biased sampling were affecting the annual anomaly, it would have minimal impact on the low-frequency signal analyzed here.

2) REPRESENTATION OF ANNUAL CYCLE

The seasonal cycle of ocean temperatures in the *World Ocean Atlas 2001* (WOA2001; Stephens et al. 2002), an earlier version of the temperature data set used here, was examined by Antonov et al. (2004). They found the significant increase in data in WOA2001 compared to an even earlier (1994) version of the data set made little difference to the depicted seasonal cycle. The implication is that there is sufficient data in WOA2001 to accurately model the seasonal cycle. In fact, they show that 99% of seasonal changes in ocean heat content, in both hemispheres, arise from the annual harmonic alone (their Table 1). Evidently, fine temporal sampling is not required to form an accurate annual cycle on basin scales. The annually averaged version of the data set we use here is based on even more data than used in WOA2001 (Levitus et al. 2005). It therefore would be expected to have an even more accurate depiction of the annual cycle, which would allow annual anomalies to be correctly computed even from irregularly sampled data.

3) DETECTION WITH DEPTH

If seasonally biased observations were somehow mimicking the models' anthropogenic warming fingerprint, the correlation between the ensemble common model

fingerprint and the observed signal would become insignificant at depths little influenced by the annual cycle. Figure 12 shows this is not the case. The correlation of the model anthropogenic warming fingerprint with the observations is statistically significant, and nearly uniform, from 250 to 600 m. At such depths, temperatures are little influenced by the seasonal cycle (Gleckler et al. 2005). The fact that the signal is detectable at depth is a strong indication that it does not arise from seasonal biases in the observations.

In summary, multidecadal trends in SST do not exhibit large differences by season; previous work suggests that the data set used here has a good representation of the seasonal cycle; and the model anthropogenic warming fingerprint is significantly correlated with the observations at depths little influenced by the seasonal cycle (250-600 m). We therefore conclude that the results shown here are not influenced by possible seasonal sampling biases in the observations.

7. Conclusions

The purpose of this work has been to explore a number of issues related to detection and attribution of ocean warming, including an evaluation of natural climate variability (internal and external) in the coupled climate models, the possible influence of sampling on the results, an estimate of the strength of model heat uptake uncertainty (compared to signal strength), the role of the surface heat flux components and advection in accomplishing the ocean warming, and the possible effects of seasonally biased observations. We used two independent coupled climate models, PCM and HadCM3 (neither of which uses flux correction) and compared them to the Levitus et al. (2005) yearly observed, gridded ocean temperature data set (0-700 m). With few exceptions, we sampled the models in accordance with the observation's *dd* sampling mask, which shows when and where an individual $1^\circ \times 1^\circ$ grid box contains an observation. Our main conclusions are as follows.

- PCM and HadCM3 show natural variability (internal and external) at levels similar to observed at time scales shorter than 0.15 cy yr^{-1} , except for having overly weak variability in basin-averaged temperatures between 300 and 500 m. This latter is not due

to weak variability at a given point, but rather because observations show more coherence in variability at this depth (resulting in lack of cancellation when the basin-average temperature is computed) than the models. This increase in coherent variability at depth in the observations is not a global phenomena, but rather localized to the southern hemisphere western tropical Pacific. Overall, the models provide a good measure of natural climate variability for anthropogenic detection studies.

- The ensemble common model fingerprint of ocean warming due to anthropogenic forcing agrees well with observed ocean warming in the top 100 m of the water column, explaining about 80-90% of the observed variance (globally and decade averaged). The warming is outside the envelope expected from natural variability, both internal and external (solar and volcanic fluctuations) to the climate system. Between 250 m and 600 m the correlation is weaker but still statistically significant. The conclusion that the warming is outside the range expected from natural variability depends on the estimate of natural variability employed. Both models' estimates of natural variability are consistent with the available observations. Since the observed ocean warming is outside this range of natural variability, we conclude the ocean is warming due to anthropogenic causes.

- The increase in ocean heat content in PCM is driven by an 0.7 watts m^{-2} increase in net surface heat flux over the world oceans since 1960. Downward longwave radiation increases by 3.7 watts m^{-2} , which is not completely compensated by an increase in upward longwave radiation of 2.2 watts m^{-2} . Net incoming shortwave and latent heat have negative trends. The anthropogenically forced changes in the longwave components become distinguishable from pre-industrial values by the 1920s, but near cancellation of terms prevents the net heat flux from becoming well separated from the pre-industrial mean until the 1960s, a gap of some decades.

- In PCM (where data is available), changes in advection play an important part in the modeled anthropogenic change in ocean temperature in certain regions. For example, in the South Indian ocean advection increases the overall warming, while it decreases the warming in the South Pacific.

- Model estimates suggest the observed sampling is sufficient to estimate basin-averaged temperatures to about ± 0.1 °C. This is highly dependent on the basin and depth, however. Comparing this uncertainty to the model-projected strength of anthropogenic ocean warming, the historical coverage is sufficient to reliably detect ocean warming in all oceans.

- CMIP2 results suggest ocean basin heat uptake in various climate models can vary by a wide margin (a factor of 2 to almost 8, depending on the region). Synthetically constructed ocean warming fingerprints corresponding to the “maximum” and “minimum” heat uptake ocean models suggests that even in the case of minimum heat uptake, the anthropogenically forced model signal falls outside the expected range of natural internal climate variability in the surface layers.

- Analysis of ocean temperature trends by season, representation of the seasonal cycle, and detectability with depth indicate our results are unaffected by possible seasonal sampling biases in the observations.

Our conclusion that observed ocean warming arises from anthropogenic sources agrees with earlier authors such as Barnett et al. (2001), Levitus et al. (2001), Reichert et al. (2002), Gent and Danabasoglu (2004), and Gregory et al. (2004). The additional information presented here and in Barnett et al. (2005) on the vertical structure of the signal explains why it is more detectable in the upper part of the water column, as noted in Gent and Danabasoglu (2004). Additionally, illustrating the evolution of net surface heat flux and the role advection plays sheds more light on the processes driving the ocean warming signal. The analysis of detectability and sampling issues as a function of depth shows the limitations of the historical observations and imply how they might be usefully extended in the future. Lastly, it should be noted that the two models used here, although producing good climate simulations overall, likely do not sample the full spread of model differences found in global climate models in use today; a fuller analysis that takes more model results into account would be worthwhile. As models with higher resolution or improved physical parameterizations become available, the kind of analysis shown here

could reveal ever more detailed information about how anthropogenic forcing is likely to affect our planet.

8. Acknowledgements

This work is a contribution from the International Detection and Attribution Group funded by the NOAA and DOE through NOAA's CCDD program. We gratefully acknowledge DOE support through grants DE-FG03-01ER63255 to the Scripps Institution of Oceanography and DOE-W-7405-ENG-48 to PCMDI at LLNL. Work at the Hadley Centre was supported by the UK Department for Environment, Food and Rural Affairs under contract PECD 7/12/37 and by the Government Meteorological Research and Development Programme. We especially wish to thank Syd Levitus for his generosity in making his updated ocean data set available, and to colleagues at NCAR and the Hadley Centre for performing and making available the model runs used in this work. Kaplan SST data provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA, from their web site. This work benefited from the comments of anonymous reviewers, for which we thank them.

9. Appendix A: Removing Model Drift

Our objective is to remove the model drift from the detection results. In Barnett et al. (2001), we fit a second-order trend to the control run and took the difference between the anthropogenically forced ensembles and fitted trend as the signal. The difference between the unforced control run and the fitted trend was used as the noise estimate. This gets more difficult when using the “dd” sampling mask, since the sampling mask itself changes significantly from year to year.

Let:

$s(\mathbf{x}, t)$ be the sampling mask (1 if grid box was sampled, 0 otherwise)

$c(\mathbf{x}, t)$ be the unforced control run model data

$e(\mathbf{x}, t, n)$ be the anthropogenically forced ensemble run model data

$v(\mathbf{x})$ be the volume of the grid box

where \mathbf{x} is the location, t is the year (1955-1999), and n is the ensemble number.

Then, the volume averaged temperature for the control run (\mathcal{T}^c) and an ensemble member (\mathcal{T}^e) are time series defined as:

$$\mathcal{T}^c(t) = \frac{\sum_{\mathbf{x}} c \cdot s \cdot v}{\sum_{\mathbf{x}} s \cdot v} \quad (5)$$

and

$$\mathcal{T}^e(t, n) = \frac{\sum_{\mathbf{x}} e(n) \cdot s \cdot v}{\sum_{\mathbf{x}} s \cdot v} \quad (6)$$

If no low-frequency fitting to the control run were desired, the signal for any particular ensemble member n could be defined as $\mathcal{T}^e(t, n) - \mathcal{T}^c(t)$. The problem with this comes when estimating the noise. If a procedure analogous to that used in Barnett et al. (2001) were employed, the noise would be: $\mathcal{T}^e(t) - \mathcal{T}^c(t) \equiv 0$. So that is not useful.

We instead define γ , the time series of the control run sampled with *one fixed mask*,

where the fixed mask is taken at year τ :

$$\gamma(t, \tau) = \frac{\sum_{\mathbf{x}} c(\mathbf{x}, t) \cdot s(\mathbf{x}, \tau) \cdot v}{\sum_{\mathbf{x}} s(\mathbf{x}, \tau) \cdot v} \quad (7)$$

The advantage of this is $\partial\gamma/\partial t$ depends only on the way $c(\mathbf{x}, t)$ depends on t . By contrast, $d\mathcal{T}^c(t)/dt$ depends on the way both $c(\mathbf{x}, t)$ and $s(\mathbf{x}, t)$ depend on t . This is not helpful, since changes in s with time are completely non-physical; they might be driven by budgets for observing programs, for example. We want to get at the way $c(\mathbf{x}, t)$ changes in time, as it is the model drift to be eliminated.

Since the time dependence of γ only depends on a quantity with a physically sensible time variation ($c(\mathbf{x}, t)$) we can reasonably fit a second order trend to it over time (t) to produce $\gamma^*(t, \tau)$.

Now consider only the diagonal elements of the γ^* array:

$$\Gamma^*(t, \tau) = \begin{cases} \gamma^*(t, \tau) & \text{if } t = \tau \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

The two time indices (t and τ) are then redundant, so we drop the τ and refer to this simply as $\Gamma^*(t)$. Γ^* represents the second order trend fitted volume-averaged temperature (i.e., retaining only model drift) seen at time t , using the correct sampling mask for that time.

We then estimate the signal for an ensemble as:

$$\mathcal{T}^e(t, n) - \Gamma^*(t) \quad (9)$$

and the noise as:

$$\mathcal{T}^c(t) - \Gamma^*(t) \quad (10)$$

This estimates the signal as the difference between what is seen at each year and what would be expected to be seen if only model drift were acting at that same year, using the correct sampling mask for the year in question.

A subtle point is that the final quantity re-introduces the temporal dependence on the sampling mask. I.e., $d/dt(\mathcal{T}^c(t) - \Gamma^*(t))$ depends on ds/dt again (since $\Gamma^*(t)$ depends on

$s(t)$). The trend fitting removes the *effect of model drift*, not the *effect of sampling mask changes*. Another point worth noting is that after this procedure, the temperatures are anomalies with respect to the control run, which represents pre-industrial conditions. We must then compute the anomalies with respect to the observed era (1957-1990) to match the treatment of the observations (Levitus et al. 2005).

10. Appendix B: Variability in Models and Observations

1) SPECTRA BY BASIN

Figure 2 shows temperatures globally averaged over all depths, but our primary interest is how ocean warming varies regionally (in the North and South Atlantic, Indian, and Pacific oceans) and with depth (cf. Lysne and Deser 2002 for a regional model-data comparison of variability in the Pacific). The data used to generate the global, depth-averaged spectrum in Fig. 2 is broken apart by basin and depth in Fig. 21 for the observations, Fig. 22 for PCM, and Fig. 23 for HadCM3. Only frequencies less than 0.2 cy/yr are shown since we are interested in examining the slow ocean warming over decades (in Section 4 temperatures are averaged by decade before forming the ocean warming fingerprint). The low-frequency power at depth should be considered uncertain due to poor sampling in early decades.

Places where observed values fall outside the model ensemble range are shown in the panels marked “outside” in Figs. 22 and 23. Solid circles indicate the observed power falls above the 90% confidence limit of the ensembles; hollow circles indicate where it falls below. A circle (hollow or solid) will be seen by chance 10% of the time; the actual percentage of locations with circles is noted in the plot title for each basin. For the purposes of detecting anthropogenic ocean warming, solid circles indicate the model is deficient in variability, which could lead to a false detection of anthropogenic influences on the oceans; hollow circles indicate the model has too much variability, which could lead to the false rejection of anthropogenic forcing as the cause of ocean warming.

The spectral power density seen in PCM is close to the values seen in observations in

all basins except the North Atlantic, where PCM tends to exhibit too much variability below 75 m. In all other basins the fraction of observed values falling outside the model ensemble range is near that expected from chance. In HadCM3, there is a tendency to have weak variability at depth, and overly strong variability in the surface layers, especially in the South Atlantic, South Pacific, and North Indian basins.

2) ANALYSIS OF STANDARD DEVIATION

Gregory et al. (2004) compared observed low frequency temperature variability to the HadCM3 control run (natural internal variability only) by computing the standard deviation of volume averaged, 5-yr running mean temperature at each level from the detrended data (their Fig. 4). We will refer to this quantity as “SD(Avg(T))”; it is shown for observed temperatures (detrended by removing the best fit line) as the broken line with triangles in Fig. 24. There is a peak in observed variability at 400 m, which neither model control run reproduced (Fig. 25, right panel).

The subsurface peak in observed SD(Avg(T)) gives the impression that were a ship to occupy a station, there would be greater temperature variability measured at 400 m than above or below. This is not true, as can be seen in the global average of individual standard deviations at each point, “Avg(SD(T))”. Generally $SD(Avg(T)) \ll Avg(SD(T))$ because significant cancellation occurs when the volume average temperature is taken. How much cancellation depends on the spatial coherence of temperature fluctuations, so an indication of coherence can be seen by comparing SD(Avg(T)) to Avg(SD(T)).

Figure 24 shows both quantities. The two curves have been normalized by their maximum values so they can be easily compared. They differ at the surface, where there is less cancellation (i.e., more spatial coherence) in ocean temperature than below. Presumably this is due to the atmosphere’s imposition of its typical length scales on the surface ocean. More importantly, Avg(SD(T)) shows no peak in variability at 400 m. This implies that the observed 400 m “peak in variability” arises not because ocean variability is unusually strong at 400 m, but because it is unusually spatially coherent. This reduces cancellation between uncorrelated points seen to a greater extent above and below 400 m

when volume-average temperature is calculated.

Figure 25 compares observations to the control model runs for both methods of calculating variability. Data have been detrended as described in section 3. The left panel shows $\text{Avg}(\text{SD}(T))$; by this measure the HadCM3 control run is generally consistent with observations, while the PCM control run has somewhat too strong internal variability between 20 and 600 m. The right panel shows $\text{SD}(\text{Avg}(T))$; as in Gregory et al. (2004), the main discrepancy is that both models miss the observed maximum at 400 m. Comparing the panels shows the models' failure to reproduce the peak in $\text{SD}(\text{Avg}(T))$ at 400 m happens mainly because they have less spatial coherence at that depth, not because variability at individual points is too weak. Further analysis (not shown) indicates the observed spatial coherence driving the peak at 400 m is not a global phenomena, but rather is confined roughly to the northwest half of a box from 105°E to 130°W , and 4°S to 42°S . As a general characterization, Fig. 25 shows that the model control runs bracket the observed pointwise variability at all depths, but have less spatial coherence than observed (but still realistic pointwise variability) around 400 m.

References

- AchutaRao, K. M., P. J. Gleckler, B. D. Santer, K. E. Taylor, D. W. Pierce, T. P. Barnett, 2005: Variability of Ocean Heat Uptake: Reconciling Observations and Models. *JGR Oceans* (submitted).
- Antonov, J. I., S. Levitus, T. P. Boyer, 2004: Climatological annual cycle of ocean heat content. *Geophys. Res. Lett.*, v. **31** L04304 doi:10.1029/2003GL018851.
- Barnett, T. P., D. W. Pierce, R. Schnur, 2001: Detection of anthropogenic climate change in the world's oceans. *Science*, v. **292** p. 270-274.
- Barnett, T. P., D. W. Pierce, K. AchutaRao, P. Gleckler, B. Santer, J. Gregory, W. Washington, 2005: Penetration of human-induced warming into the world's oceans. *Science*, v. **309** p. 284-287.
- Dai, A., W. M. Washington, G. A. Meehl, T. W. Bettge, W. G. Strand, 2004: The ACPI climate change simulations. *Climatic Change*, v. **62** p. 29-43.
- Dukowicz, J. K., R. D. Smith, 1994: Implicit free-surface method for the Bryan-Cox-Semtner ocean model. *J. Geophys. Res.*, v. **99** p. 7991-8014.
- Gent, P. R., G. Danabasoglu, 2004: Heat uptake and the thermohaline circulation in the Community Climate System Model, version 2. *J. Climate*, v. **17** p. 4058-4069.
- Gleckler, P. J., K. R. Sperber, K. AchutaRao, 2005: The annual cycle of global heat content: observed and simulated. *J. Geophys. Res.*, submitted.
- Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. Mitchell, 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. *Climate Dynamics*, v. **16** p. 147-168.
- Gregory, J. M., 2000: Vertical heat transports in the ocean and their effect on time-dependent climate change. *Climate Dynamics*, v. **16** p. 501-515.
- Gregory, J. M., H. T. Banks, P. A. Stott, J. A. Lowe, M. D. Palmer, 2004: Simulated and observed decadal variability in ocean heat content. *Geophys. Res. Lett.*, v. **31** doi:10.1029/2004GL020258.

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

Hunke, E. C., J. K. Dukowicz, 1997: An elastic-viscous-plastic model for sea ice dynamics. *J. Phys. Oceanogr.*, v. **27** p. 1849-67.

IDAG (International ad hoc Detection, Attribution Group), 2005: Detecting and Attributing External Influences on the Climate System: A Review of Recent Advances. *J. Climate*, in press.

IPCC, 2001: Climate Change 2001: Synthesis Report. Contribution of working groups I, II, and III to the third assessment report of the intergovernmental panel on climate change. Cambridge University Press, 397 pp.

Johns, T. C., J. M. Gregory, W. J. Ingram, C. E. Johnson, A. Jones, J. A. Lowe, J. F. Mitchell, D. L. Roberts, D. M. Sexton, D. S. Stevenson, S. F. Tett, M. J. Woodage, 2003: Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios. *Climate Dynamics*, v. **20** p. 583-612.

Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, B. Rajagopalan, 1998: Analysis of global sea surface temperature 1856-1991. *J. Geophys. Res.*, v. **C9** p. 18567-18589.

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, D. J. Williamson, P. J. Rasch, 1998: The National Center for Atmospheric Research Community Climate Model: CCM3. *J. Climate*, v. **11** p. 1131-1149.

Knowlton, N., 2001: The future of coral reefs. *PNAS*, v. **98** p. 5419-5425.

Levitus, S., J. I. Antonov, T. P. Boyer, 2000: Warming of the world ocean. *Science*, v. **287** p. 2225-2229.

Levitus, S., J. I. Antonov, J. L. Wang, T. L. Delworth, K. W. Dixon, A. J. Broccoli, 2001: Anthropogenic warming of Earth's climate system. *Science*, v. **292** p. 267-70.

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

Lysne, J., C. Deser, 2002: Wind-driven thermocline variability in the Pacific: A

model-data comparison. *J. Climate*, v. **15** p. 829-845.

Meehl, G. A., G. J. Boer, C. Covey, M. Latif, R. J. Stouffer, 2000: The Coupled Model Intercomparison Project (CMIP). *Bull. Amer. Met. Soc.*, v. **81** p. 313-318.

Meehl, G. A., W. M. Washington, T. M. Wigley, J. M. Arblaster, A. Dai, 2003: Solar and greenhouse gas forcing and climate response in the twentieth century. *J. Climate*, v. **16** p. 426-444.

Pierce, D. W., 2004: Beyond the means: Validating climate models with higher-order statistics. *Computing in Science and Engineering*, v. **6** Sep/Oct 22-29.

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

Smith, R. D., J. K. Dukowicz, R. C. Malone, 1992: Parallel ocean general circulation modeling. *Physica D*, v. **60** p. 38-61.

Sokolov, A. P., C. E. Forest, P. H. Stone, 2003: Comparing Oceanic Heat Uptake in AOGCM Transient Climate Change Experiments. *J. Climate*, v. **16** p. 1573-0.

Stephens, C., J. I. Antonov, T. P. Boyer, M. E. Conkright, R. Locarini, T. D. O'Brien, 2002: World Ocean Atlas 2001 Volume 1: Temperature. NOAA Atlas NESDIS 49, U.S. Govt. Printing Office, Washington, D. C., 167 pp.

Sun, S., J. E. Hansen, 2003: Climate simulations for 1951-2050 with a coupled atmosphere-ocean model. *J. Climate*, v. **16** p. 2807-2826.

Tett, S. F., G. S. Jones, P. A. Stott, D. C. Hill, J. F. Mitchell, M. R. Allen, W. J. Ingram, T. C. Johns, C. E. Johnson, A. Jones, D. L. Roberts, D. M. Sexton, M. J. Woodage, 2002: Estimation of natural and anthropogenic contributions to twentieth century climate change. *J. Geophys. Res.*, v. **107** 4306 doi:10.1029/2000JD000028.

Thompson, L. G., 2000: Ice core evidence for climate change in the Tropics: implications for our future. *Quaternary Science Reviews*, v. **19** p. 19-35.

Thompson, L. G., E. Mosley-Thompson, M. E. Davis, P. N. Lin, K. Henderson, T. A. Mashiotta, 2003: Tropical glacier and ice core evidence of climate change on annual to millennial time scales. *Climatic Change*, v. **59** p. 137-155.

Tynan, C. T., D. P. DeMaster, 1997: Observations and predictions of Arctic climatic change: Potential effects on marine mammals. *Arctic*, v. **50** p. 308-322.

Washington, W. M., J. W. Weatherly, G. A. Meehl, A. J. Semtner, T. W. Bettge, A. P. Craig, W. G. Strand, J. Arblaster, V. B. Wayland, R. James, Y. Zhang, 2000: Parallel Climate Model (PCM) control and transient simulations. *Climate Dynamics*, v. **16** p. 755-74.

Willis, J. K., D. Roemmich, B. Cornuelle, 2004: Interannual variability in upper ocean heat content, temperature, and thermosteric expansion on global scales. *J. Geophys. Res.*, v. **109** C12036 doi:10.1029/2003JC002260.

Zhang, J., W. D. Hibler, 1997: On an efficient numerical method for modeling sea ice dynamics. *J. Geophys. Res.*, v. **102** p. 8691-702.

Figure Captions

Figure 1. Standard deviation (C) of detrended annually averaged sea surface temperature anomalies. Top: Observed, 1945-2004. Middle: PCM control run (520 years). Bottom: HadCM3 control run (700 years). Contour interval is 0.2 C; values greater than 0.6 C are shaded. Model results are from the control run with no anthropogenic, solar, or volcanic variability.

Figure 2. Spectra of yearly volume averaged temperature anomaly, 0-700 m, at *dd* sampled points only, 1955-1999. Thick line: Observed. Grey shading: envelope from four HadCM3 anthropogenically forced runs. Stippled: envelope from 12 PCM anthropogenically forced runs. All series are detrended by removing the best-fit line before computing the spectrum. The cross in the upper right shows the 90% confidence interval and bandwidth centered on 0.3 cy yr^{-1} for the observations.

Figure 3. Basin-averaged standard deviation of temperature by depth (at *dd* sampled points only), from yearly values averaged into non-overlapping 5-year pentads and detrended by removing the best fit line. Dots: observed. Shaded: range in HadCM3 control run. Crosshatched: range in PCM control run.

Figure 4. Histograms of the linear trend (C/decade) in volume average temperature from the PCM control runs (black) and anthropogenically forced ensemble members (gray), compared to the observations ('X'). Using a K-S test, '****' indicates the probability the gray and black values are drawn from the same underlying distribution is less than 0.01; '***', less than 0.05; '**', less than 0.10.

Figure 5. As in Fig. 4, but for HadCM3.

Figure 6. Time series of yearly global volume average temperature anomaly (0-700 m), at *dd* sampled points only. Left panel: from observations (bottom, thick line) and 12 PCM ensemble members with anthropogenic forcing estimated over the historical period. Right panel: from the PCM control run with natural internal climate variability alone. Curves are offset by 0.12 C; ensemble member case IDs are given along the right edge.

Figure 7. Same as Fig. 6, but for HadCM3.

Figure 8. Time series of yearly volume averaged temperature by ocean, 0-700 m, at *dd* sampled points only. Thick solid line: observed. Thick dashed line: PCM ensemble average. Thin lines: individual PCM ensemble members.

Figure 9. The ocean-decade concatenated “fingerprint” of ocean warming at 50 m depth. Dashed heavy line with squares: ensemble common model fingerprint generated using both PCM and HadCM3. Solid heavy line with dots: observed. Thin lines: individual model ensemble members. The X axis is concatenated decade by ocean; the Y axis is volume averaged temperature (at *dd* mask sampled points).

Figure 10. Ensemble common model fingerprint of ocean warming, by depth, computed from the 12 PCM and 4 HadCM3 anthropogenically forced ensemble members taken together. The *x* axis is concatenated decadal values by basin; the first *x* point therefore reflects the volume averaged temperature anomaly in the North Atlantic (basin labels are across the top) in the 1960s, the second *x* point the 1970s, etc. The highly characteristic sawtooth shape of the fingerprint in the surface layers indicates all basins experienced distinct warming from the 1960s to the 1990s. Curves are offset by 0.1 C. Only data at *dd* mask sampled points are included.

Figure 11. Strength of the ensemble common model fingerprint of ocean warming found in the observations (circles), anthropogenically forced model runs (90% confidence interval shown by shaded region), and control run with internal natural variability only (90% confidence interval around zero shown by crosshatched region). Left: for PCM. Right: for HadCM3. PCM also had a separate run with only solar and volcanic forcing; the strength of the fingerprint found in the ensemble common signal of these runs is shown by the triangles in the left panel. Only data at *dd* mask sampled points was included.

Figure 12. Correlation between the ensemble common model fingerprint of ocean warming and the observations. Values outside the shaded region are statistically significant at the 95% level. Only data at *dd* mask sampled points was included.

Figure 13. Components of the surface heat flux, by ocean (and world: lower right) from the PCM anthropogenically forced runs. Values are anomalies relative to the first 40

years, averaged by decade. For all components, positive values act to warm the surface. The whiskerplots show the mean, interquartile range, and minimum/maximum of the ensembles.

Figure 14. Time series of net surface heat flux (NSHF), changes in local heat storage, and heating due to advection from the PCM anthropogenically forced ensembles. Values are decadal averages relative to the 1880-1919 average.

Figure 15. Fraction of ocean volume sampled to 700 m, based on the $1^\circ \times 1^\circ$ yearly *dd* mask from Levitus et al. (2005).

Figure 16. Strength of ensemble common model warming fingerprint (C) in the observations (circles), unforced control runs (90% confidence interval around zero shown by crosshatching), and anthropogenically forced ensemble members (shaded region). Left column: using *dd* mask (sampled points only), for PCM (top) and HadCM3 (bottom). Right column: using *gp* mask (points within 450 km of samples included), for PCM (top) and HadCM3 (bottom).

Figure 17. Uncertainty in observed volume averaged temperature by basin, estimated using PCM. Grey shaded region: 90% confidence interval of the error, using *dd* sampling locations of the 1960s. Crosshatched region: using *dd* sampling locations of the 1990s. So the uncertainty can be compared to the signal, values are offset by the anthropogenically forced warming seen in the 1960s (white squares) and 1990s (black dots) relative to the pre-industrial era.

Figure 18. Thick line with squares: ensemble common model ocean warming fingerprint. Dashed line with triangles: the constructed “minimum” model signal, based on models with the minimal ocean heat uptake (by basin) in the CMIP2 data set. Thin line with circles: the constructed “maximum” model signal. Only data at *dd* mask sampled points was included. See text for details.

Figure 19. Strength of global ocean warming signal as a function of depth. Squares: ensemble common model fingerprint. Circles: observations. Black triangles: for the constructed “minimum” heat uptake model. White triangles: for the “maximum” heat

uptake model. Crosshatched region: 90% confidence interval for HadCM3 unforced control model run. Only data at *dd* mask sampled points was included.

Figure 20. Linear trend in SST anomalies (C/century) over the period 1905-2004, by season. Values greater than 0.25 C/century are shaded. Contour interval is 0.25 C. Data have been spatially filtered with a boxcar of width 3000 km.

Figure 21. As in Fig. 2, but shown as a function of basin and depth for observations only. Plotted are the the logs of the spectral power density; values > -4 are shaded. Contour interval is 0.5.

Figure 22. As in Fig. 21, but for PCM. The panels marked “outside” indicate locations where the observed spectral power (Fig. 21) falls outside the 90% confidence interval of the PCM ensembles’ spectra; solid circles indicate where the observed spectral density exceeds the model range, and hollow circles indicate where the observed spectral density is smaller than the minimum model value. See text for details.

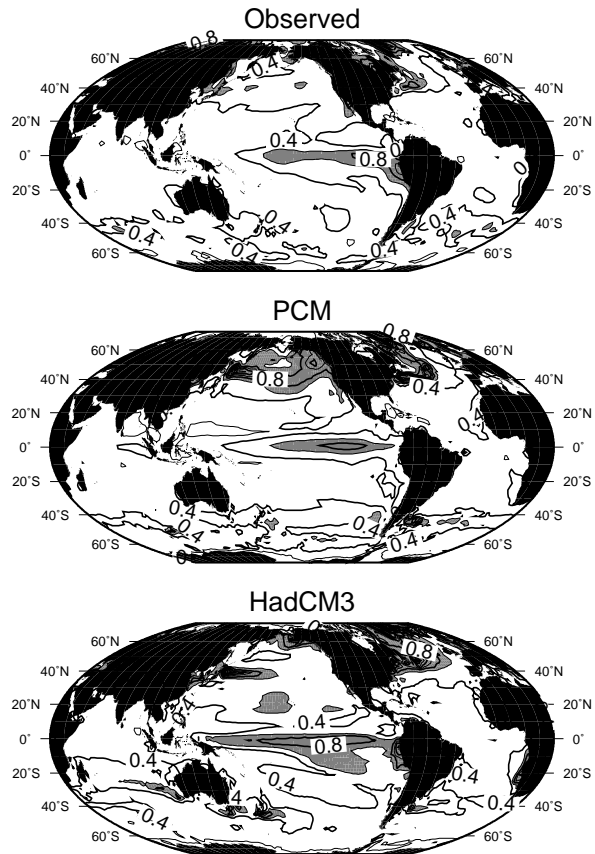
Figure 23. As in Fig. 22, but for HadCM3.

Figure 24. Standard deviation of detrended 5-yr running mean globally averaged temperature as a function of depth (at *dd* sampled points only). Both curves are from observations, but differ in the way they were computed. Black triangles: standard deviation of globally averaged temperature. White squares: global average of the standard deviation of temperature at each point. Curves are normalized by their maximum value.

Figure 25. Standard deviation of detrended (removal of best fit line) 5-yr running mean globally averaged temperature as a function of depth, comparing observations (black dots), HadCM3 control run (gray shaded region shows the ensemble range), and PCM control run (crosshatched region shows ensemble range). Left: computed as the global average of the standard deviation of temperature at each point. Right: computed as the standard deviation of globally averaged temperature. All data is at *dd* sampled points only.

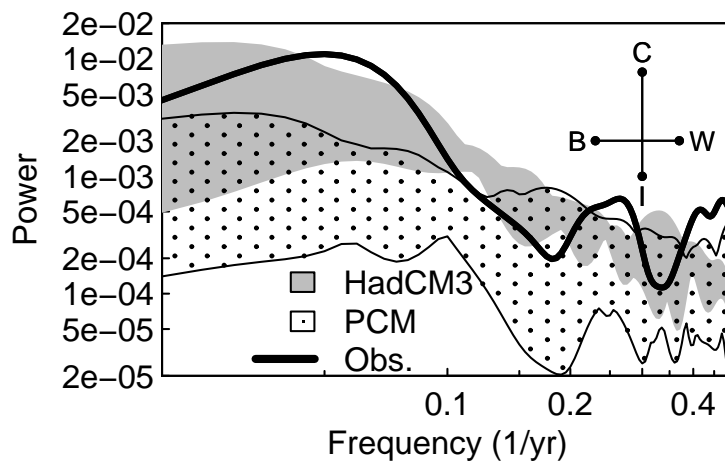
Table Captions

Table 1. Volume averaged increase in temperature (C) by ocean, for various climate models, after 80 years of 1% per year increase in atmospheric CO₂ concentration. See AchutaRao et al. (2005) for details.



Wed Sep 7 14:20:05 2005 gtc.surf_sd_v2.pl

Figure 1: Standard deviation (C) of detrended annually averaged sea surface temperature anomalies. Top: Observed, 1945-2004. Middle: PCM control run (520 years). Bottom: HadCM3 control run (700 years). Contour interval is 0.2 C; values greater than 0.6 C are shaded. Model results are from the control run with no anthropogenic, solar, or volcanic variability.



_2004/yearly/temp/compare_spec_with_models2.R Mon May 16 12:28:31 2005

Figure 2: Spectra of yearly volume averaged temperature anomaly, 0-700 m, at *dd* sampled points only, 1955-1999. Thick line: Observed. Grey shading: envelope from four HadCM3 anthropogenically forced runs. Stippled: envelope from 12 PCM anthropogenically forced runs. All series are detrended by removing the best-fit line before computing the spectrum. The cross in the upper right shows the 90% confidence interval and bandwidth centered on 0.3 cy yr^{-1} for the observations.

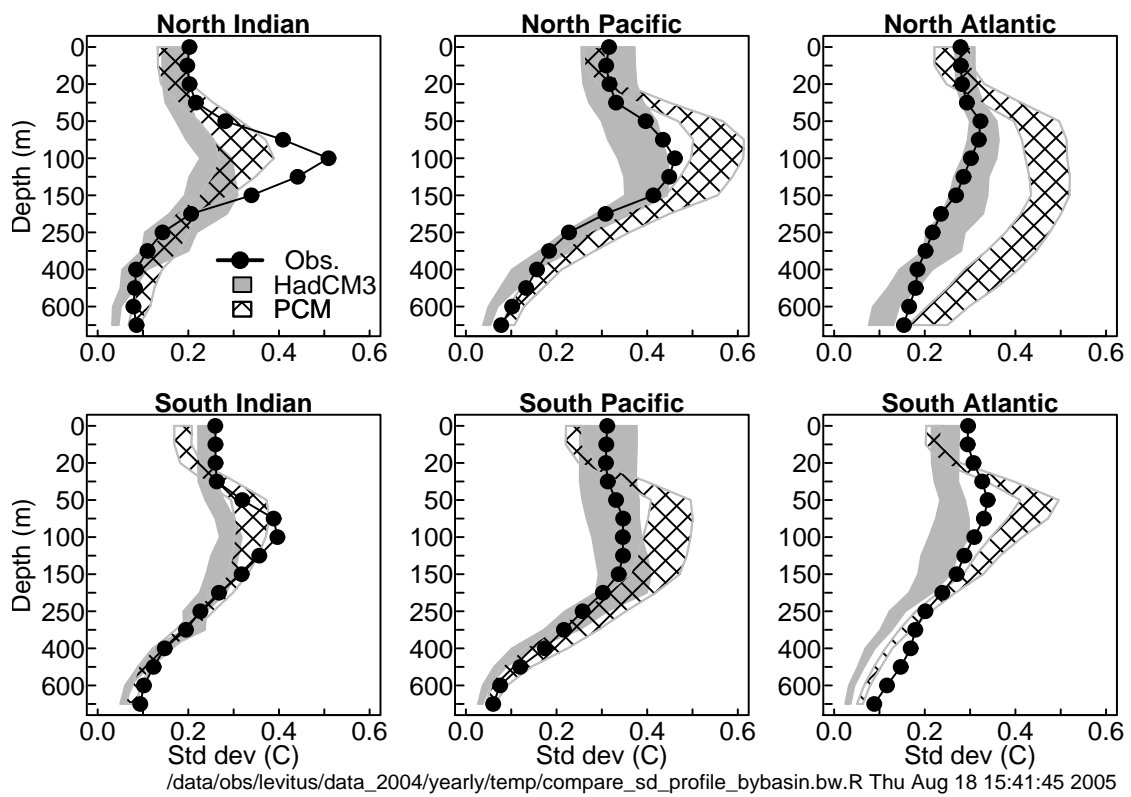


Figure 3: Basin-averaged standard deviation of temperature by depth (at *dd* sampled points only), from yearly values averaged into non-overlapping 5-year pentads and detrended by removing the best fit line. Dots: observed. Shaded: range in HadCM3 control run. Crosshatched: range in PCM control run.

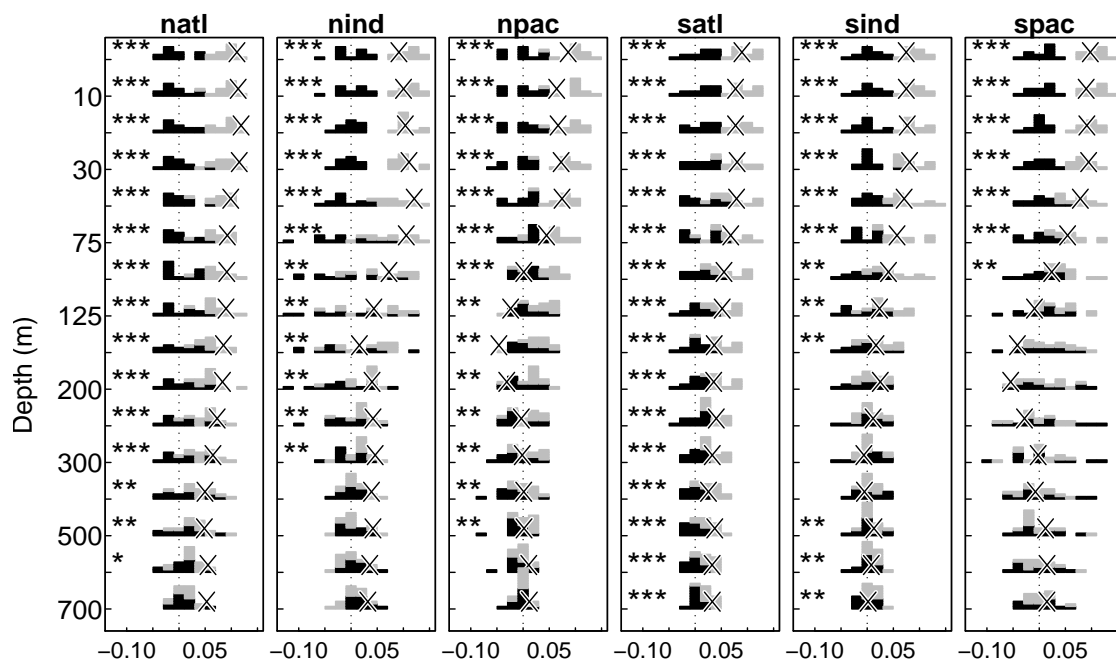


Figure 4: Histograms of the linear trend (C/decade) in volume average temperature from the PCM control runs (black) and anthropogenically forced ensemble members (gray), compared to the observations ('X'). Using a K-S test, '***' indicates the probability the gray and black values are drawn from the same underlying distribution is less than 0.01; '**', less than 0.05; '*', less than 0.10.

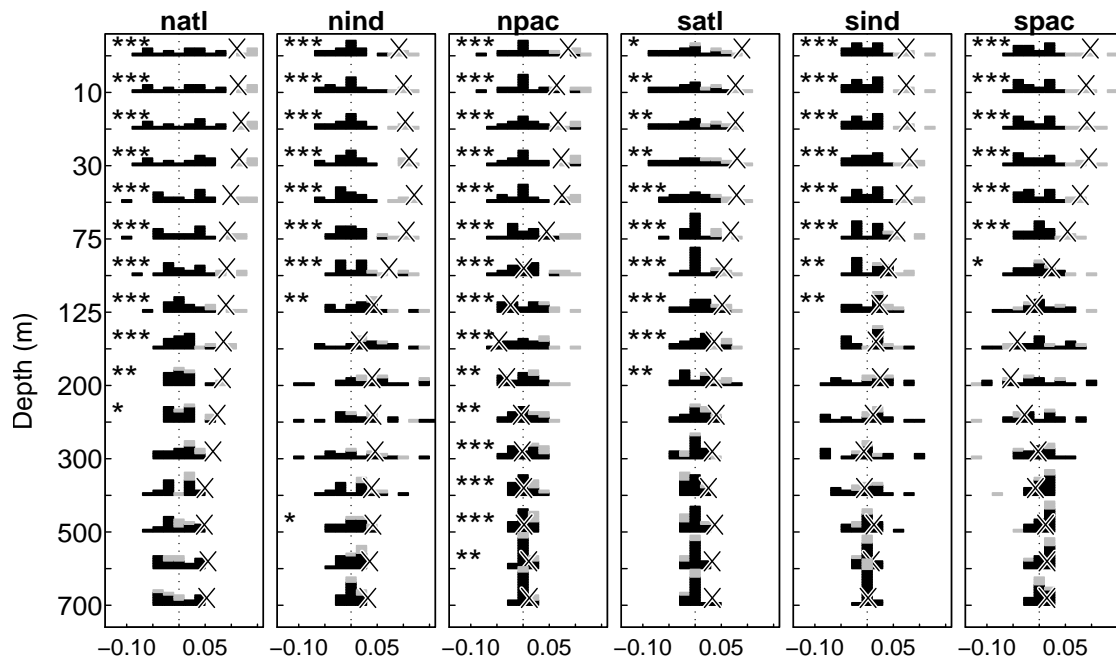
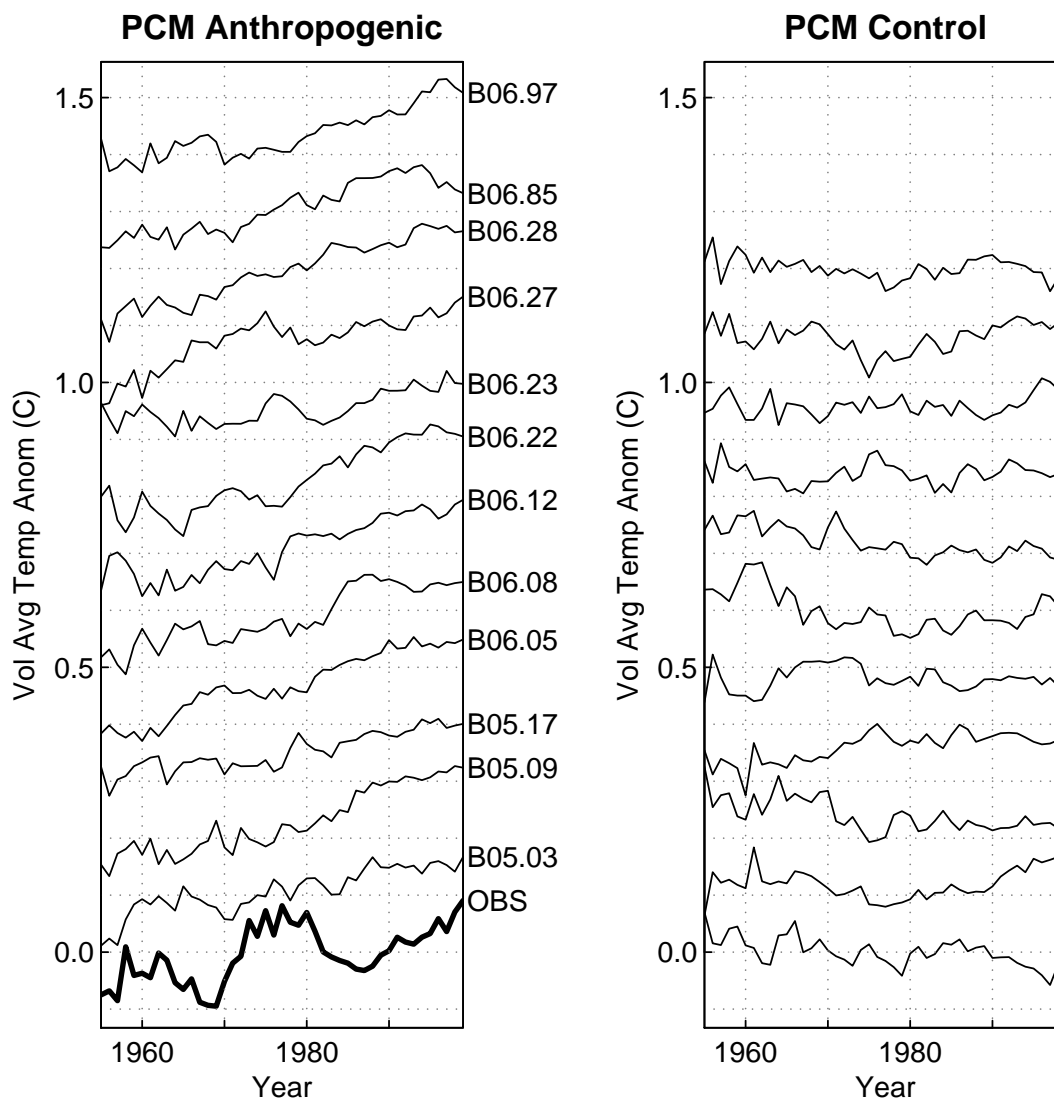
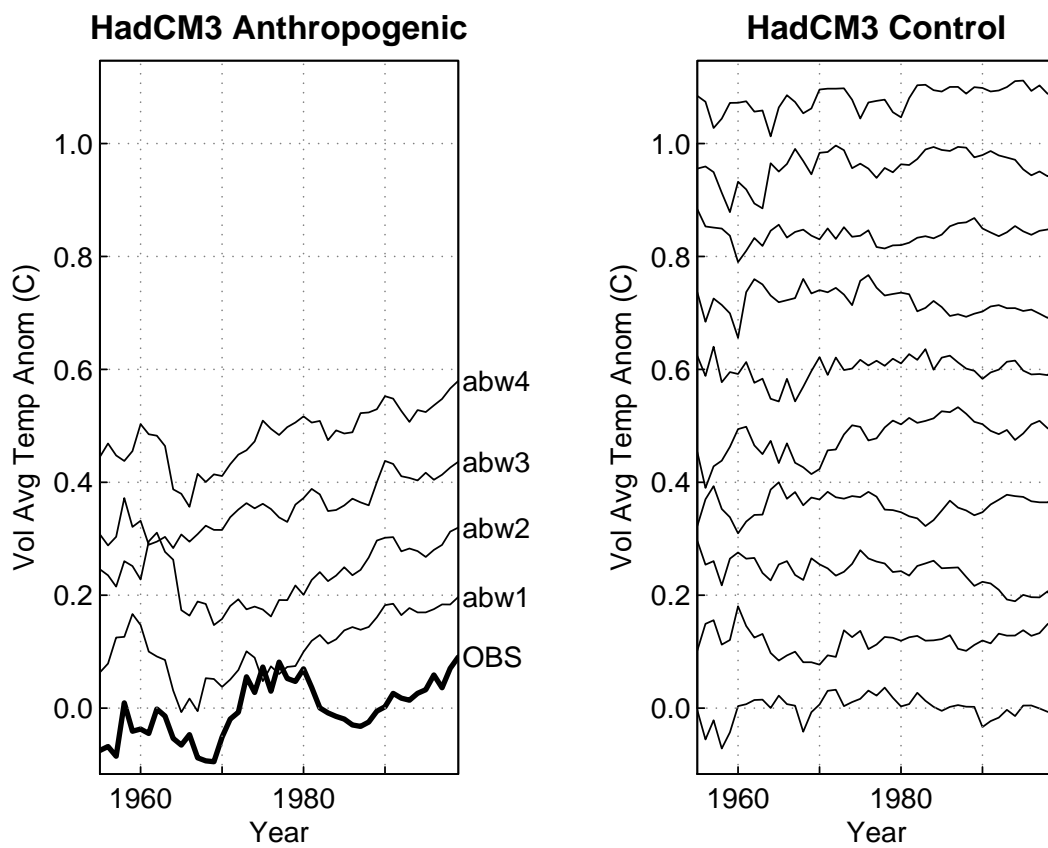


Figure 5: As in Fig. 4, but for HadCM3.



/data/obs/levitus/data_2004/yearly/temp/plot_pcm_10pan_v2.R Tue Sep 20 13:47:08 2005

Figure 6: Time series of yearly global volume average temperature anomaly (0-700 m), at *dd* sampled points only. Left panel: from observations (bottom, thick line) and 12 PCM ensemble members with anthropogenic forcing estimated over the historical period. Right panel: from the PCM control run with natural internal climate variability alone. Curves are offset by 0.12 C; ensemble member case IDs are given along the right edge.



/data/obs/levitus/data_2004/yearly/temp/plot_had_10pan_v2.R Sat Sep 3 10:51:02 2005

Figure 7: Same as Fig. 6, but for HadCM3.

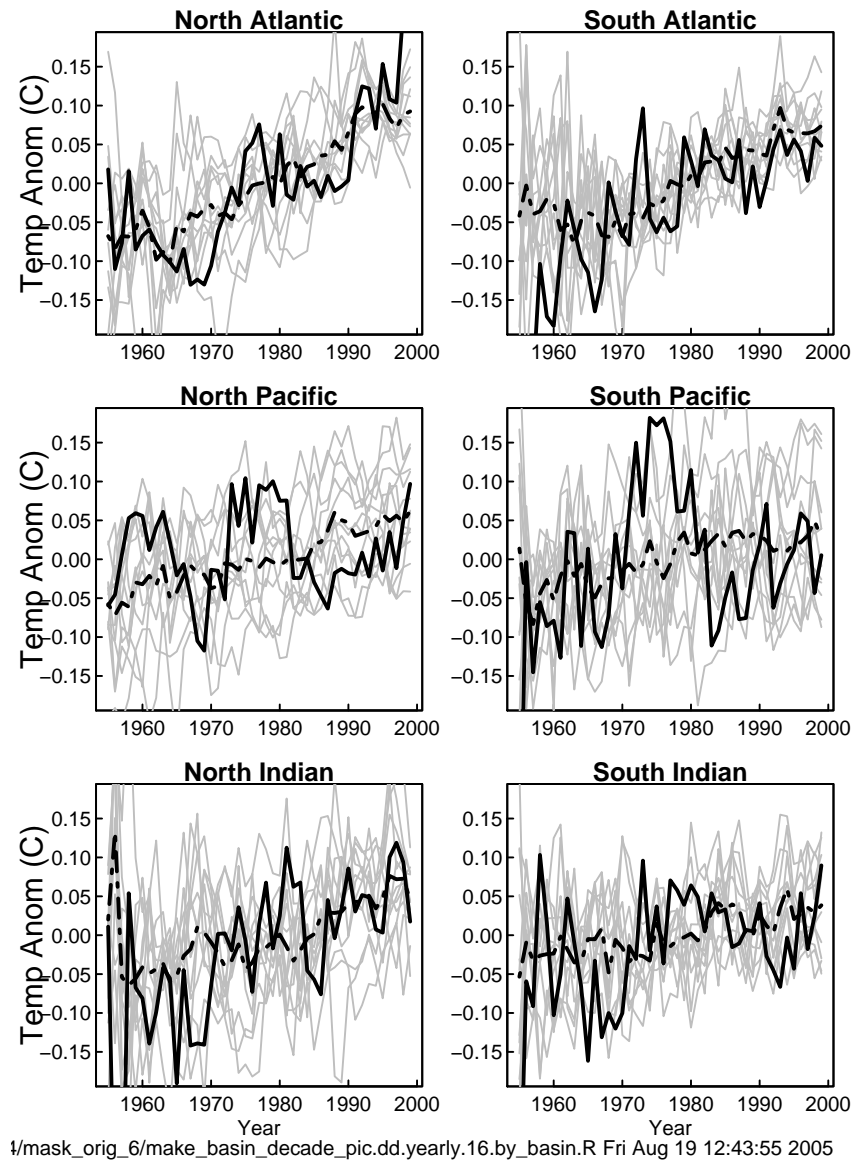
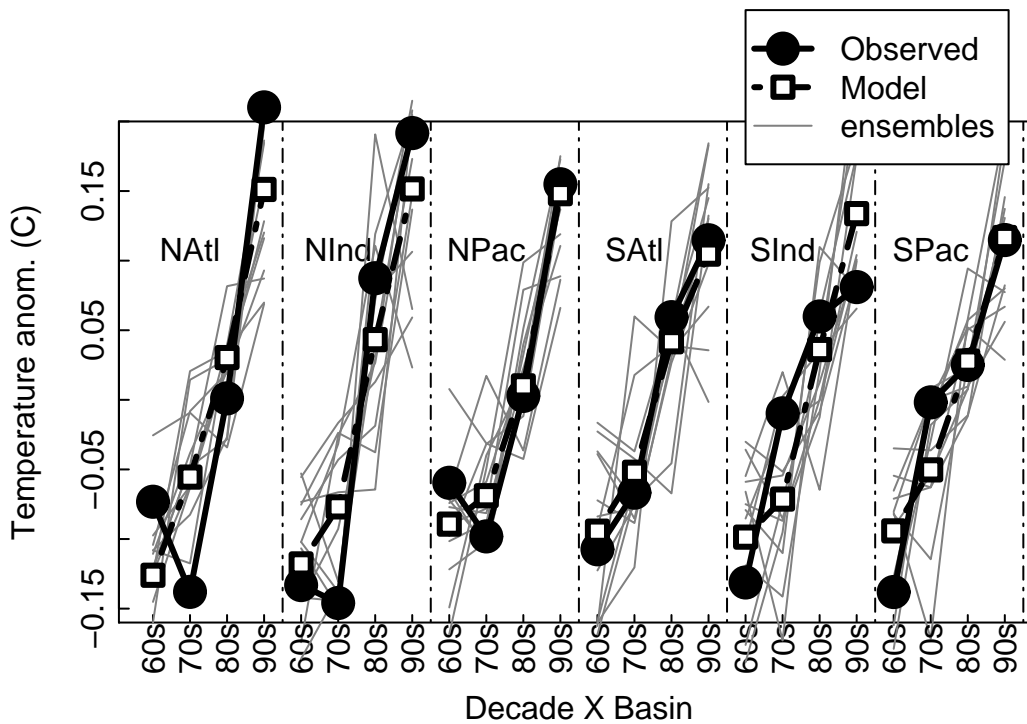
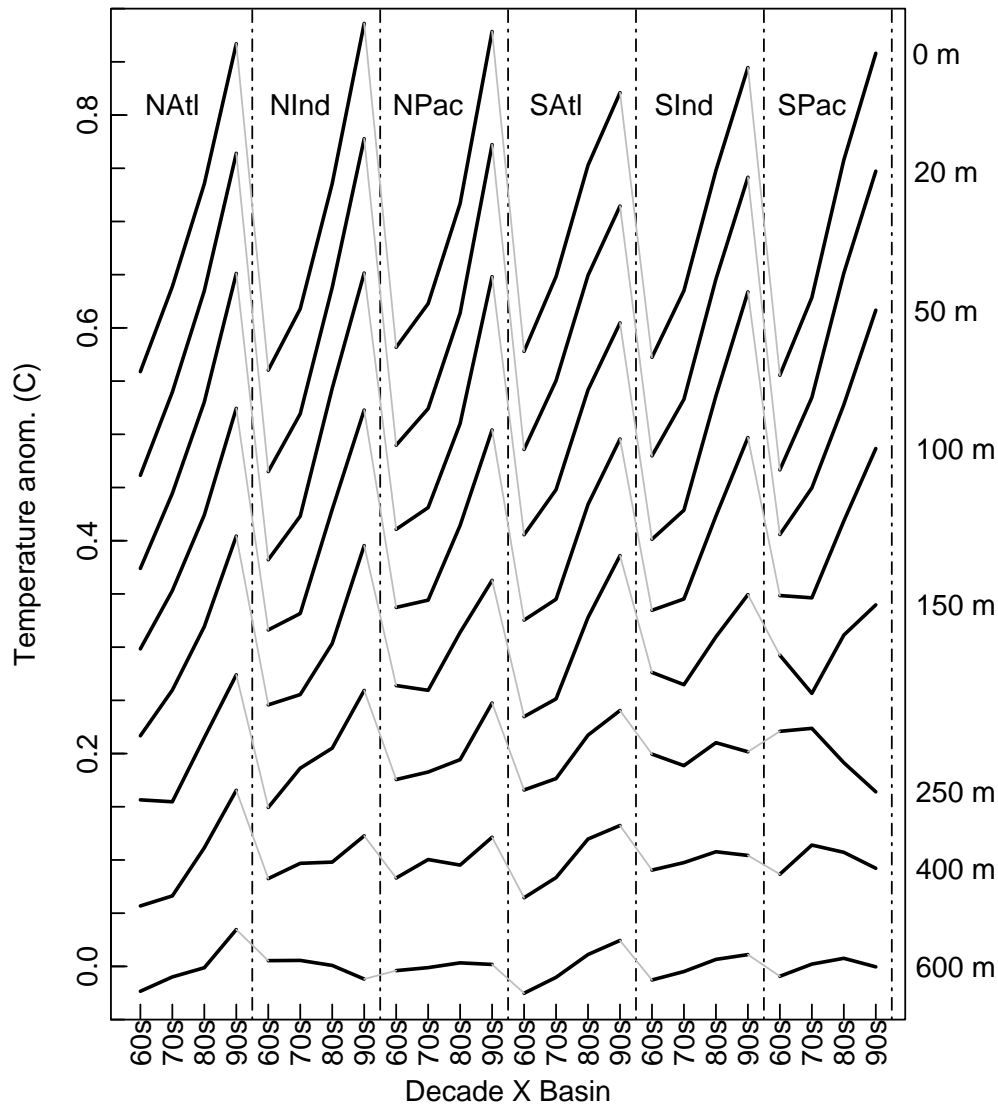


Figure 8: Time series of yearly volume averaged temperature by ocean, 0-700 m, at *dd* sampled points only. Thick solid line: observed. Thick dashed line: PCM ensemble average. Thin lines: individual PCM ensemble members.



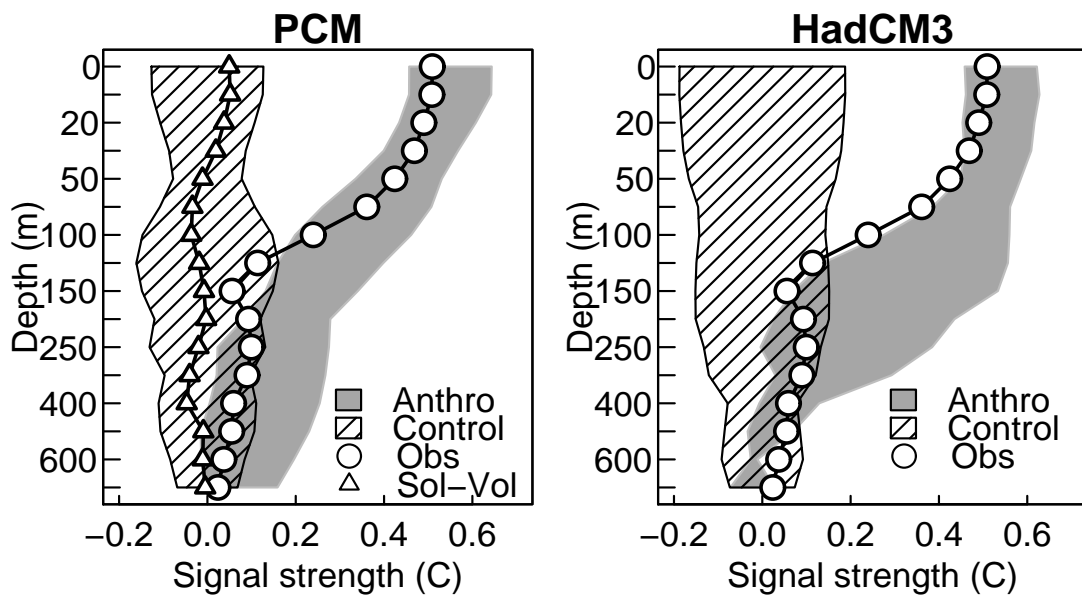
/data/pcm/hc_lev_2004/mask_orig_6/plot_fingerprint_v2.R Mon Aug 29 15:23:22 2005

Figure 9: The ocean-decade concatenated “fingerprint” of ocean warming at 50 m depth. Dashed heavy line with squares: ensemble common model fingerprint generated using both PCM and HadCM3. Solid heavy line with dots: observed. Thin lines: individual model ensemble members. The X axis is concatenated decade by ocean; the Y axis is volume averaged temperature (at *dd* mask sampled points).



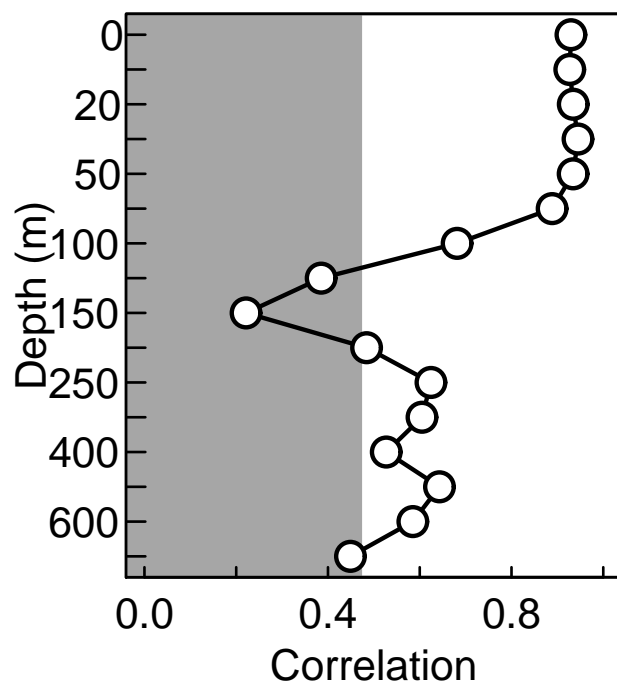
/data/pcm/hc_lev_2004/mask_orig_6/plot_fingerprint_stack_v2.R Tue Aug 30 14:37:24 2005

Figure 10: Ensemble common model fingerprint of ocean warming, by depth, computed from the 12 PCM and 4 HadCM3 anthropogenically forced ensemble members taken together. The x axis is concatenated decadal values by basin; the first x point therefore reflects the volume averaged temperature anomaly in the North Atlantic (basin labels are across the top) in the 1960s, the second x point the 1970s, etc. The highly characteristic sawtooth shape of the fingerprint in the surface layers indicates all basins experienced distinct warming from the 1960s to the 1990s. Curves are offset by 0.1 C. Only data at dd mask sampled points are included.



data/pcm/hc_lev_2004/mask_orig_6/taylor_diagram_replacement_v2.fourth_try.R Sat Sep 3 16:49:09 2005

Figure 11: Strength of the ensemble common model fingerprint of ocean warming found in the observations (circles), anthropogenically forced model runs (90% confidence interval shown by shaded region), and control run with internal natural variability only (90% confidence interval around zero shown by crosshatched region). Left: for PCM. Right: for HadCM3. PCM also had a separate run with only solar and volcanic forcing; the strength of the fingerprint found in the ensemble common signal of these runs is shown by the triangles in the left panel. Only data at *dd* mask sampled points was included.



nt_v2.fourth_try.CORR.R Thu Sep 1 13:33:48 2005

Figure 12: Correlation between the ensemble common model fingerprint of ocean warming and the observations. Values outside the shaded region are statistically significant at the 95% level. Only data at *dd* mask sampled points was included.

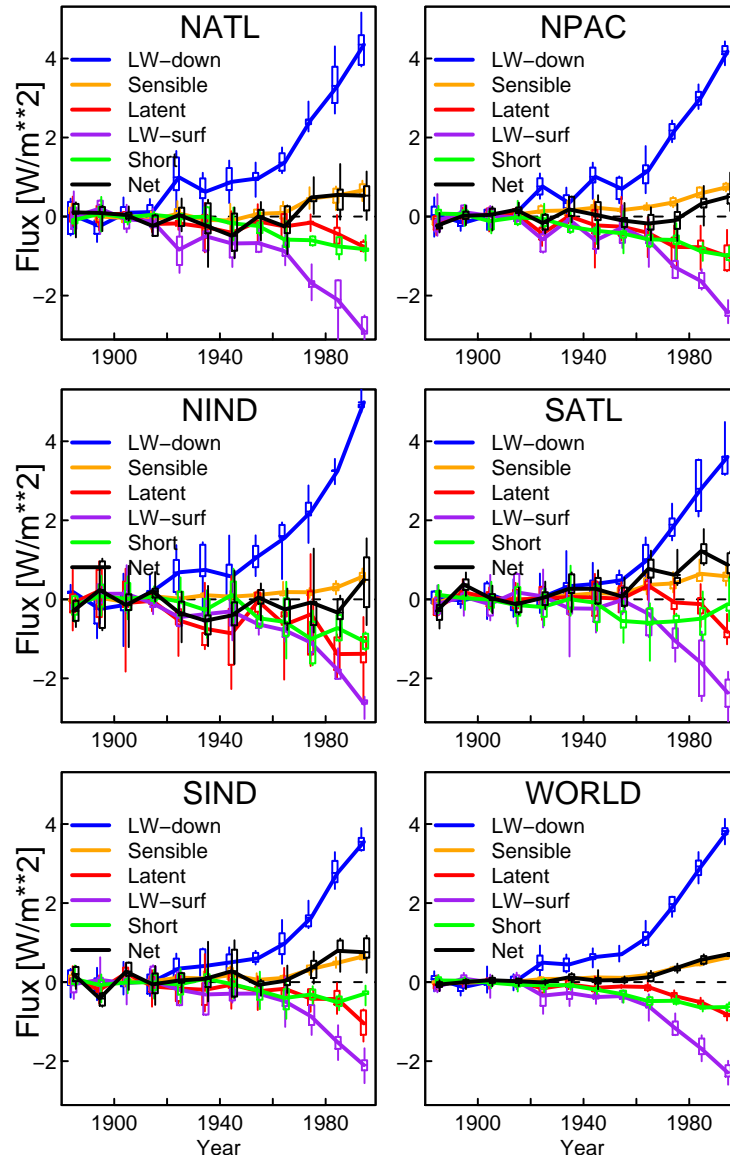


Figure 13: Components of the surface heat flux, by ocean (and world: lower right) from the PCM anthropogenically forced runs. Values are anomalies relative to the first 40 years, averaged by decade. For all components, positive values act to warm the surface. The whiskerplots show the mean, interquartile range, and minimum/maximum of the ensembles.

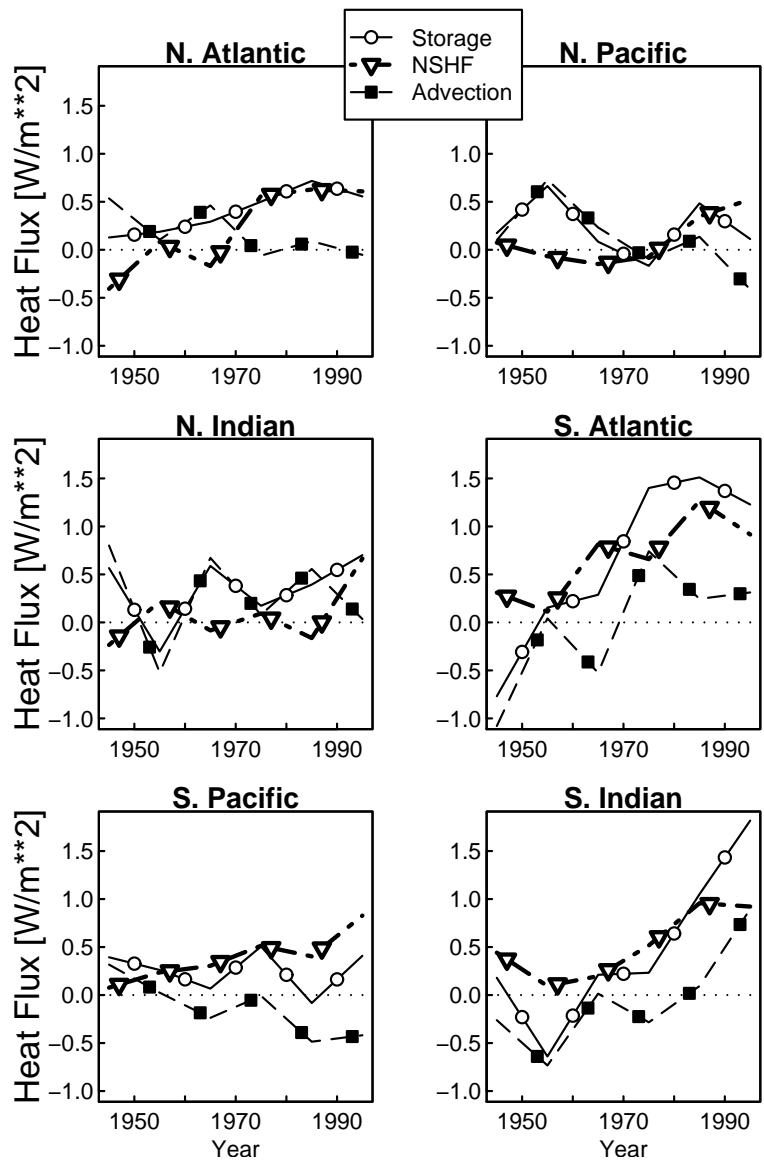


Figure 14: Time series of net surface heat flux (NSHF), changes in local heat storage, and heating due to advection from the PCM anthropogenically forced ensembles. Values are decadal averages relative to the 1880-1919 average.

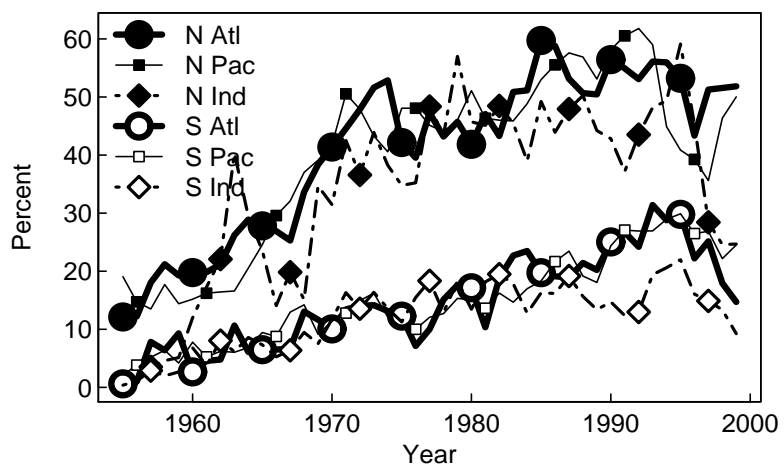
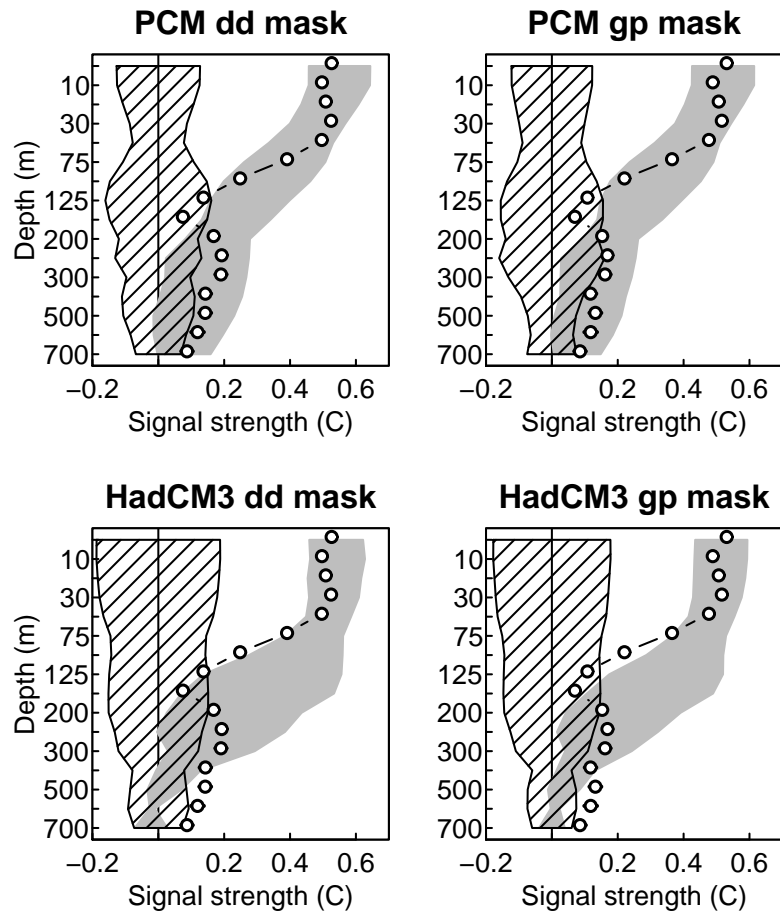


Figure 15: Fraction of ocean volume sampled to 700 m, based on the $1^\circ \times 1^\circ$ yearly *dd* mask from Levitus et al. (2005).



/data/pcm/hc_lev_2004/mask_orig_6/plot_dd_vs_gp.R Sat Sep 3 16:53:21 2005

Figure 16: Strength of ensemble common model warming fingerprint (C) in the observations (circles), unforced control runs (90% confidence interval around zero shown by crosshatching), and anthropogenically forced ensemble members (shaded region). Left column: using *dd* mask (sampled points only), for PCM (top) and HadCM3 (bottom). Right column: using *gp* mask (points within 450 km of samples included), for PCM (top) and HadCM3 (bottom).

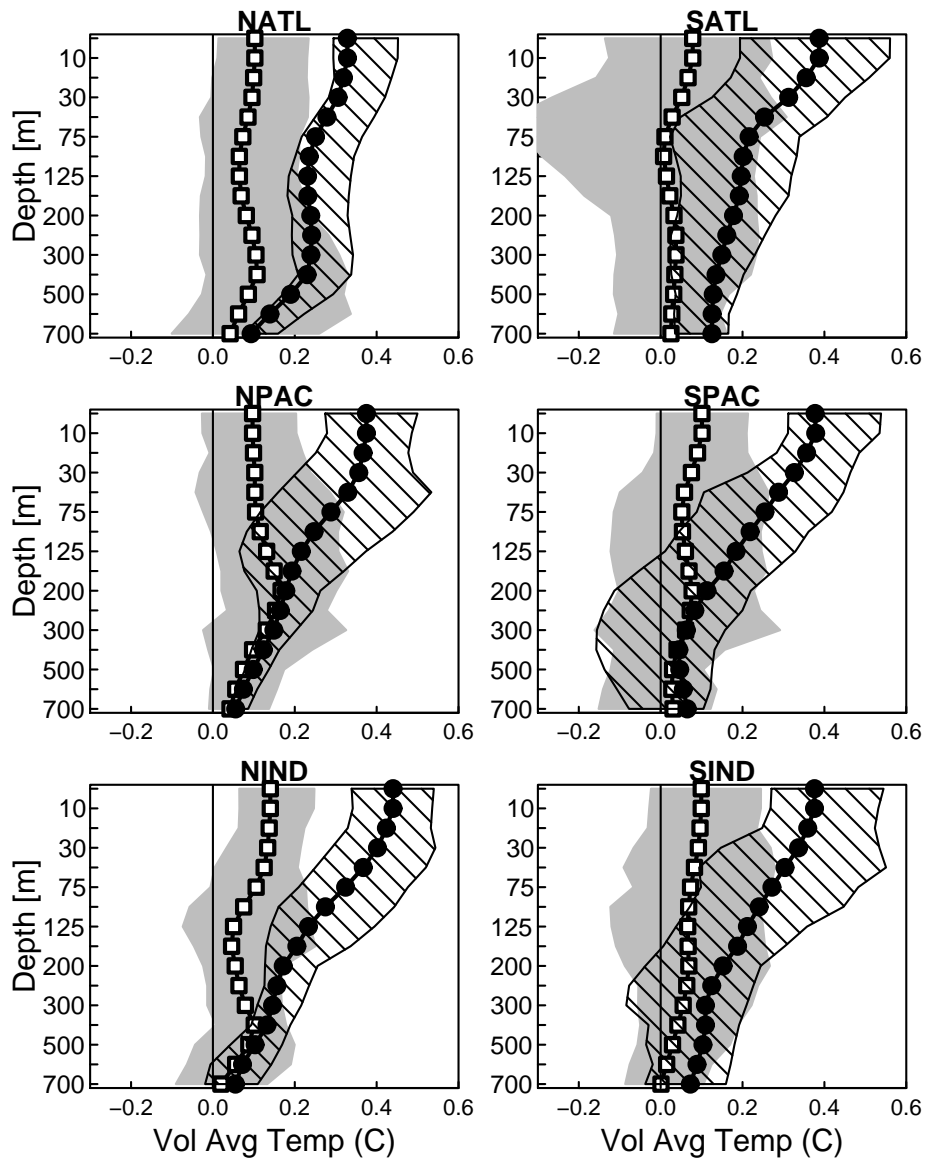
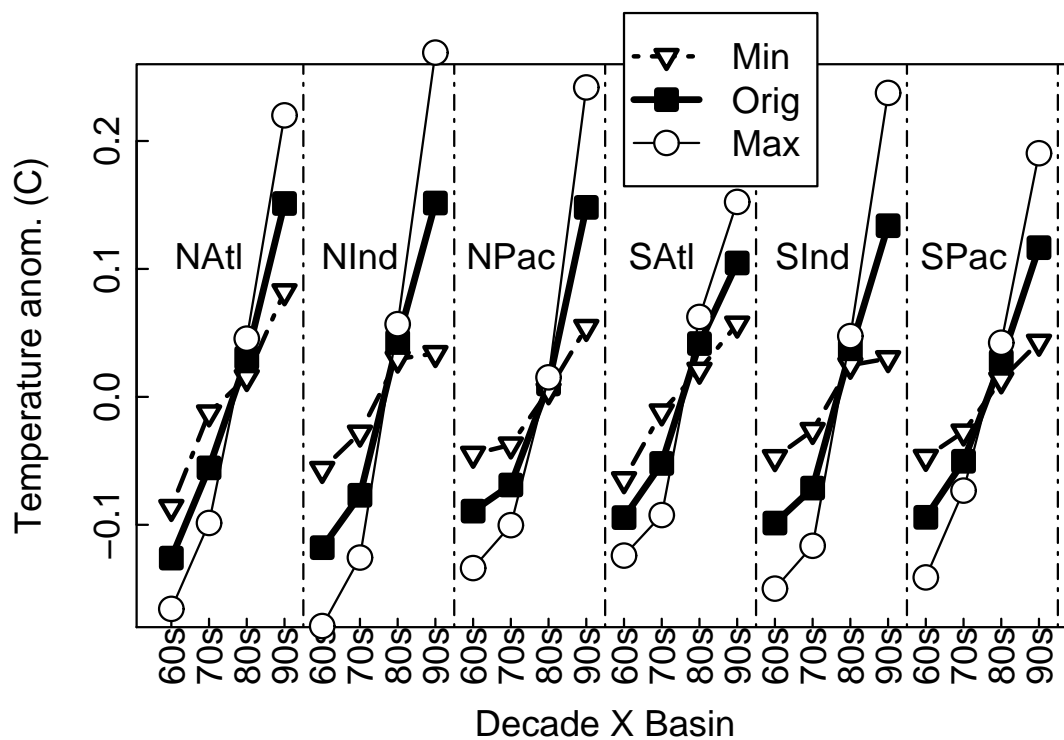
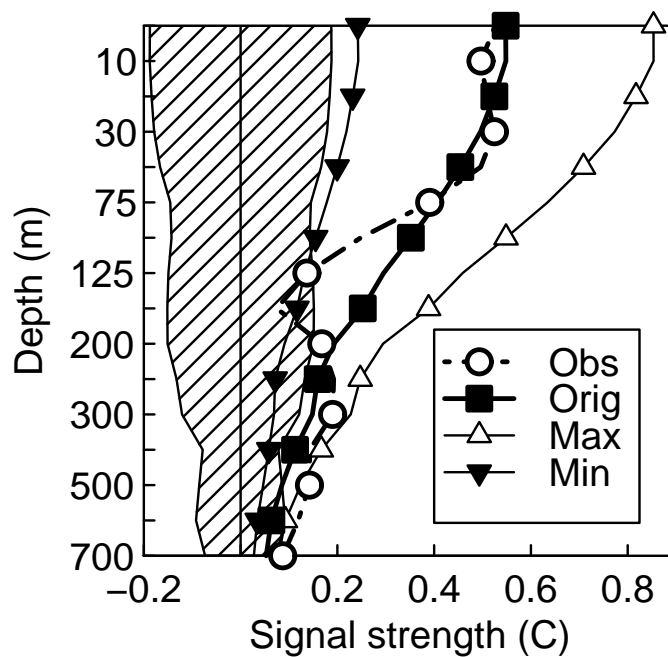


Figure 17: Uncertainty in observed volume averaged temperature by basin, estimated using PCM. Grey shaded region: 90% confidence interval of the error, using *dd* sampling locations of the 1960s. Crosshatched region: using *dd* sampling locations of the 1990s. So the uncertainty can be compared to the signal, values are offset by the anthropogenically forced warming seen in the 1960s (white squares) and 1990s (black dots) relative to the pre-industrial era.



/data/pcm/hc_leve_2004/mask_orig_6/str_minmax_fingerprints.level_5.R Tue Aug 23 15:19:03 2005

Figure 18: Thick line with squares: ensemble common model ocean warming fingerprint. Dashed line with triangles: the constructed “minimum” model signal, based on models with the minimal ocean heat uptake (by basin) in the CMIP2 data set. Thin line with circles: the constructed “maximum” model signal. Only data at *dd* mask sampled points was included. See text for details.



mask_orig_6/str_minmax_fingerprints.R Tue Aug 23 15:17:14 2005

Figure 19: Strength of global ocean warming signal as a function of depth. Squares: ensemble common model fingerprint. Circles: observations. Black triangles: for the constructed “minimum” heat uptake model. White triangles: for the “maximum” heat uptake model. Crosshatched region: 90% confidence interval for HadCM3 unforced control model run. Only data at *dd* mask sampled points was included.

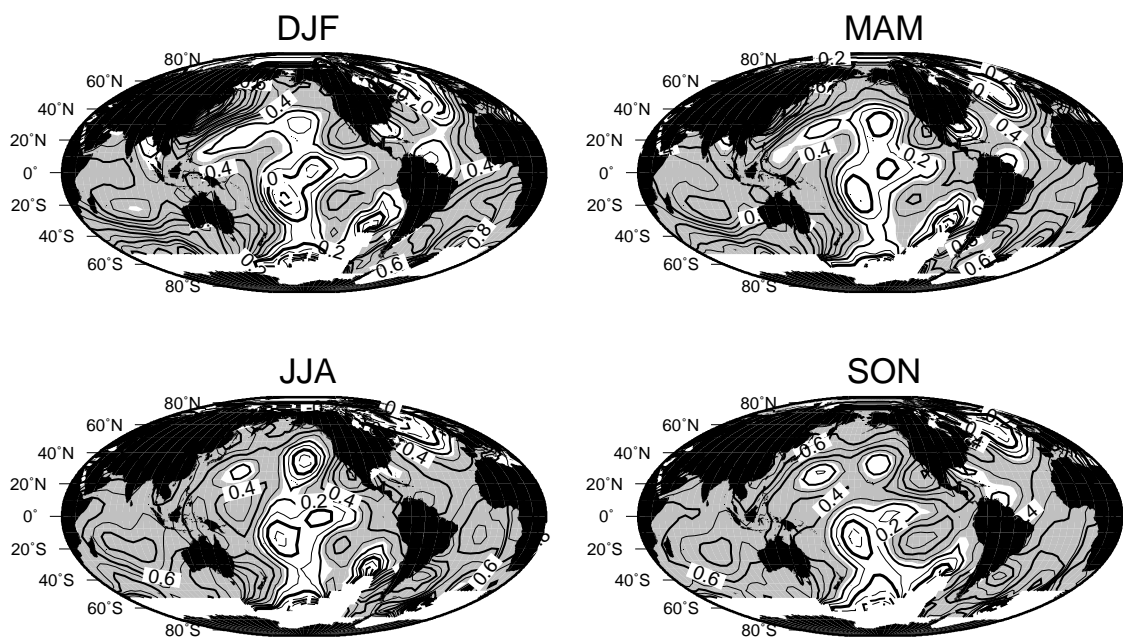


Figure 20: Linear trend in SST anomalies (C/century) over the period 1905-2004, by season. Values greater than 0.25 C/century are shaded. Contour interval is 0.25 C. Data have been spatially filtered with a boxcar of width 3000 km.

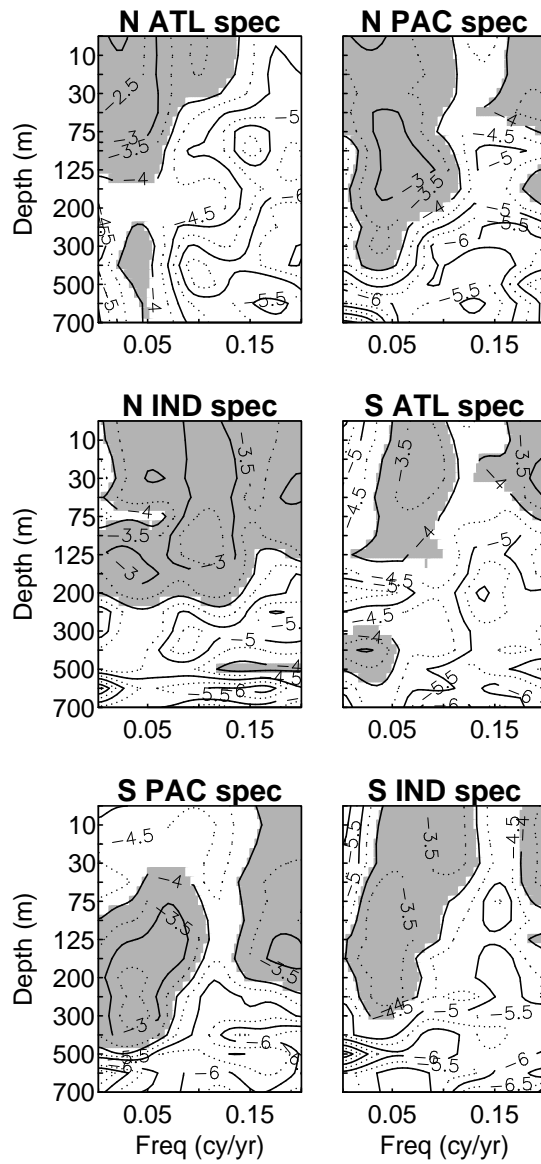


Figure 21: As in Fig. 2, but shown as a function of basin and depth for observations only. Plotted are the the logs of the spectral power density; values > -4 are shaded. Contour interval is 0.5.

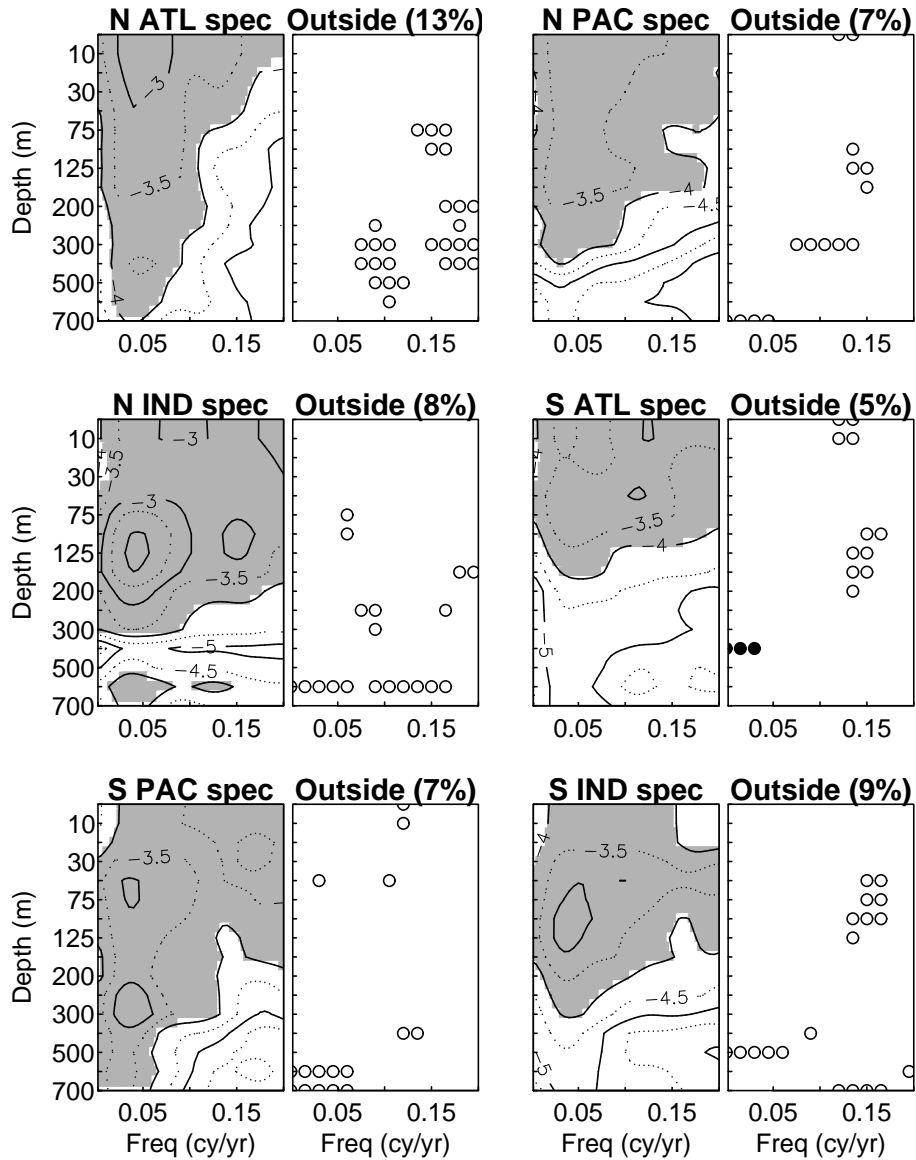


Figure 22: As in Fig. 21, but for PCM. The panels marked “outside” indicate locations where the observed spectral power (Fig. 21) falls outside the 90% confidence interval of the PCM ensembles’ spectra; solid circles indicate where the observed spectral density exceeds the model range, and hollow circles indicate where the observed spectral density is smaller than the minimum model value. See text for details.

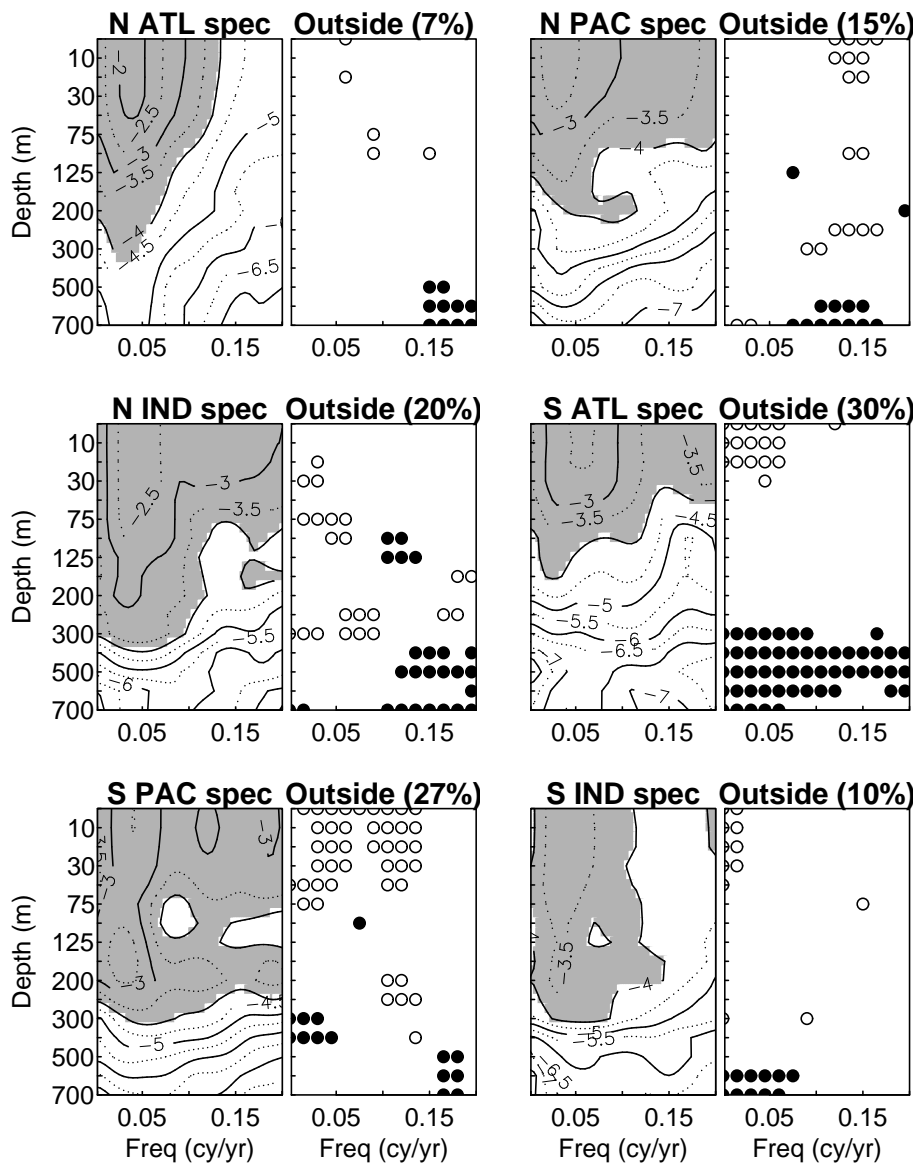
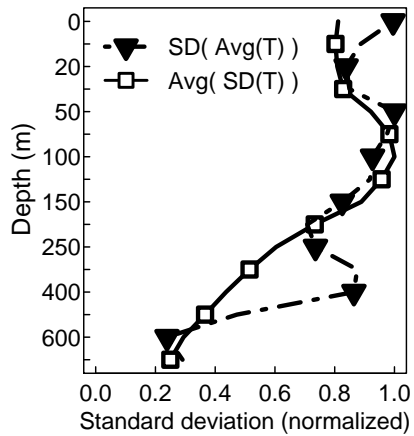
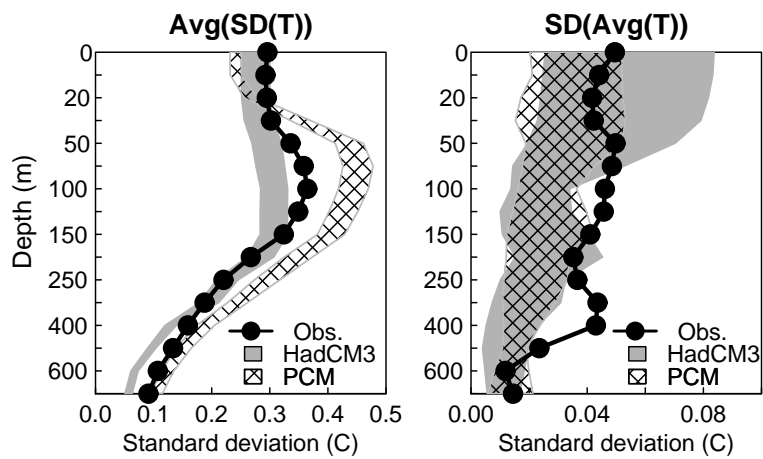


Figure 23: As in Fig. 22, but for HadCM3.



/sd_of_avg_t_vs_avg_of_sd_t_world.R Wed May 11 11:25:49 2005

Figure 24: Standard deviation of detrended 5-yr running mean globally averaged temperature as a function of depth (at *dd* sampled points only). Both curves are from observations, but differ in the way they were computed. Black triangles: standard deviation of globally averaged temperature. White squares: global average of the standard deviation of temperature at each point. Curves are normalized by their maximum value.



ata/obs/levitus/data_2004/yearly/temp/sd_of_avg_t_vs_avg_of_sd_t_world_v2.R Thu Aug 18 15:11:15 2005

Figure 25: Standard deviation of detrended (removal of best fit line) 5-yr running mean globally averaged temperature as a function of depth, comparing observations (black dots), HadCM3 control run (gray shaded region shows the ensemble range), and PCM control run (crosshatched region shows ensemble range). Left: computed as the global average of the standard deviation of temperature at each point. Right: computed as the standard deviation of globally averaged temperature. All data is at *dd* sampled points only.

Model	N Atl	S Atl	N Pac	S Pac	N Ind	S Ind
csiro-c97a	0.531	0.377	0.339	0.270	0.521	0.316
bcm-c01a	0.322	0.252	0.114	0.067	0.066	0.100
hadcm3-c97a	0.498	0.183	0.215	0.224	0.215	0.218
mri-c02a	0.365	0.253	0.239	0.199	0.422	0.241
pcm-c99a	0.277	0.163	0.133	0.129	0.121	0.123
cccma-c01a	0.505	0.436	0.294	0.296	0.274	0.315
gfdl-c96a	0.506	0.319	0.280	0.232	0.319	0.200
md-c99a	0.427	0.227	0.192	0.161	0.199	0.156
Maximum	0.531	0.436	0.339	0.296	0.521	0.316
Minimum	0.277	0.163	0.114	0.067	0.066	0.100
Max/Min	1.919	2.675	2.989	4.449	7.926	3.163

Table 1: Volume averaged increase in temperature (C) by ocean, for various climate models, after 80 years of 1% per year increase in atmospheric CO₂ concentration. See AchutaRao et al. (2005) for details.