Supporting Online Material for:

Penetration of Human-Induced Warming into the World's Oceans

T. P. Barnett*[1], D. W. Pierce[1] , K. M. AchutaRao[2], P. J. Gleckler[2], B. D. Santer[2], J. M. Gregory[3], W. M. Washington[4]

[1] Climate Research Division, Scripps Institution of Oceanography, 0224, La Jolla, CA 92037. [2] PCMDI/LLNL, P.O. Box 808, Livermore, CA 94550. [3] UK Met Office Hadley Centre and University of Reading, Reading, U.K. [4] NCAR, P.O. Box 3000, Boulder, CO 80307. * To whom correspondence should be addressed. E-mail tbarnett@ucsd.edu

Materials and Methods

1. Ocean warming fingerprint

We start with the yearly averaged temperature anomaly, volume averaged (over a basin) at observed locations only using the "dd" mask from the observations (*S1*). This is done independently for each of 16 vertical levels, from the surface to 700 m (we checked all the way to 3000 m, but the results showed little statistically significant signal below 700 m for most basins). The anomaly is with respect to the simultaneous value of the long control run fitted with a second order polynomial, to remove model drift. We then form decadal averages (the 1960s through the 1990s) to reduce the effects of annual and decadal climate variability not of interest here, and form anomalies over this period to match the treatment of the observations. Let the result be called T_v. The decadal averaging also helps eliminates the large scale spatial noise (*S2*), which the model will not reproduce since it did not see the exact forcing the real ocean did.

For each of the five anthropogenically forced ensemble members (started from different times in the long control run) we form a so-called "concatenated" signal (*S3*). This is a vector with 24 entries: one set of 4 decadal entries for each of the 6 basins (North and South Atlantic, Pacific, and Indian). It is laid out so that the first four entries are the 4 values of T_v for the North Atlantic, the second four entries are the 4 values of T_v for the North Indian, and so on for the rest of the basins. This concatenated signal combines both temporal and geographical information in a compact way.

To get the final signal, we must extract information common to the concatenated signals from all ensemble members. To do this we first formed a two-dimensional array (in X and T), where each ensemble member contributed one X entry, and the time-basin concatenated signal extended along the T axis. We then performed a principal component analysis (PCA) on this array. PCA is a filtering device to better isolate the signal that is coherent between the ensemble members. The leading principal component has the

concatenated signal for all six basins; we split it back apart by basin to form the "ensemble common signal" (ECS) for each basin.

2. Definition of signal strength

The observed signal strength P_o is given by P_o = (ECS · observations) / || ECS ||. This number was compared to P_{nv}, the estimate of signal strength from the natural internal variability runs. Again, these computations were only carried out only where data existed in the observed data set. The distribution of the multiple signal estimates from the control run allows statistical confidence statements to be made on the observed versus model signal strength.

The bars on the observations represent the +/- two standard deviation uncertainty limits due to sampling. They were estimated by a bootstrap method that chose, in random order, the sequence of yearly observed sampling masks to apply to the PCM data. This was repeated 1000 times for each of the five ensemble members. Subsequent processing was the same as for the unperturbed case. We found that the inter-ensemble difference between the 5 realizations was much larger than the intra-ensemble sampling uncertainty. In short, the results of this study are insensitive to sampling uncertainty associated with the irregular space-time distribution of ocean temperature observations.

3. Analysis of control run

We began with 520 years of the PCM control run and 700 years of the HadCM3 control run. We took these to represent the models' natural climate state subject to internal forcing only, i.e. natural internal variability. The control runs were divided into non-overlapping chunks the same length as the observations, then analyzed the same way as the observed data. The resulting distributions of signal strength estimates for each ocean and level formed a probability distribution and hence confidence limits on the expected magnitude of natural variability.

Supporting text

The time series and spectra of the control run chunks are compared to observations in Fig. S1. The spectra indicate that model levels of internal natural variability are similar to observed, except for a large observed event that occurred in the late 70s and early 80s. Previous studies have discussed the possibility this event is related to sampling fluctuations (*S4*), as it appears inconsistent with inferred surface fluxes and so may be due to an undersampled horizontal redistribution of heat (*S5*). In any event, with only 40 years of observations, decadal variability is poorly estimated in the observations, and century scale or longer variations are essentially not sampled at all. Thus we urge caution in concluding that the levels of variability in the models are the same as observed, but they appear reasonably close insofar as can be judged from the available data. The separation between the strength of the ocean warming signal in the unforced model control run and observations can be seen in Fig. S2 (same as Fig. 2 of the main text but normalized by the standard deviation of the control run). Observed values are typically 3 or more standard deviations greater than zero in the upper part of the water

column, which suggests that the observed ocean warming signal is strong enough to be confidently detected even if the model's natural variability is underestimated.

Figures

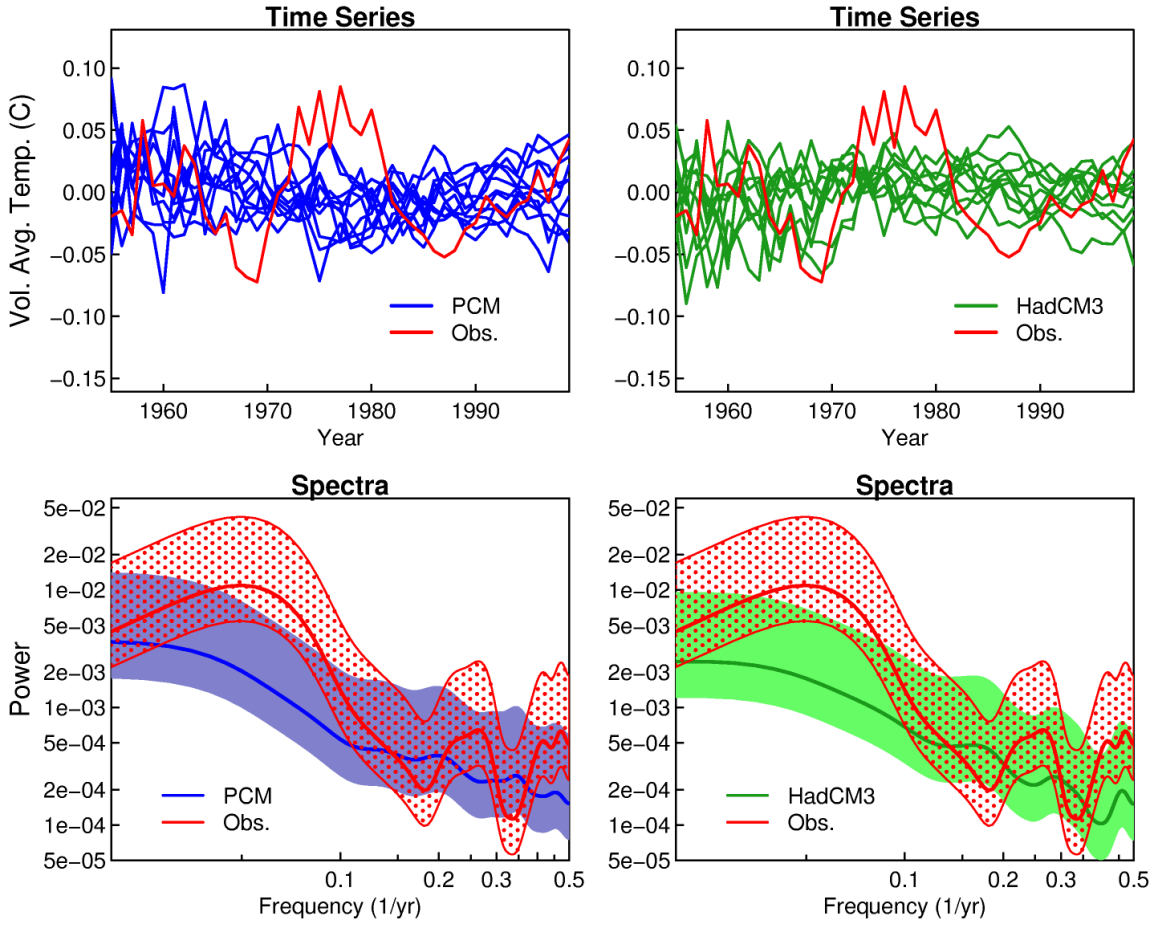

Fig. S1. Comparison of natural internal climate variability in models and observations. Top row: Time series of yearly global ocean temperature in the upper 700 m at sampled points from observations 1955-1999 (detrended; red line), compared to unforced control runs from PCM (left) and HadCM3 (right). Lower row: Power spectra of observed and model data shown in upper row. The observed spectrum and 90% confidence interval are shown in red; the ensemble mean model spectrum and 90% confidence interval are shown in blue for PCM (left) and green for HadCM3 (right). A similar analysis by depth for HadCM3 is shown in (*S4*), and by basin and depth in (*S6*).

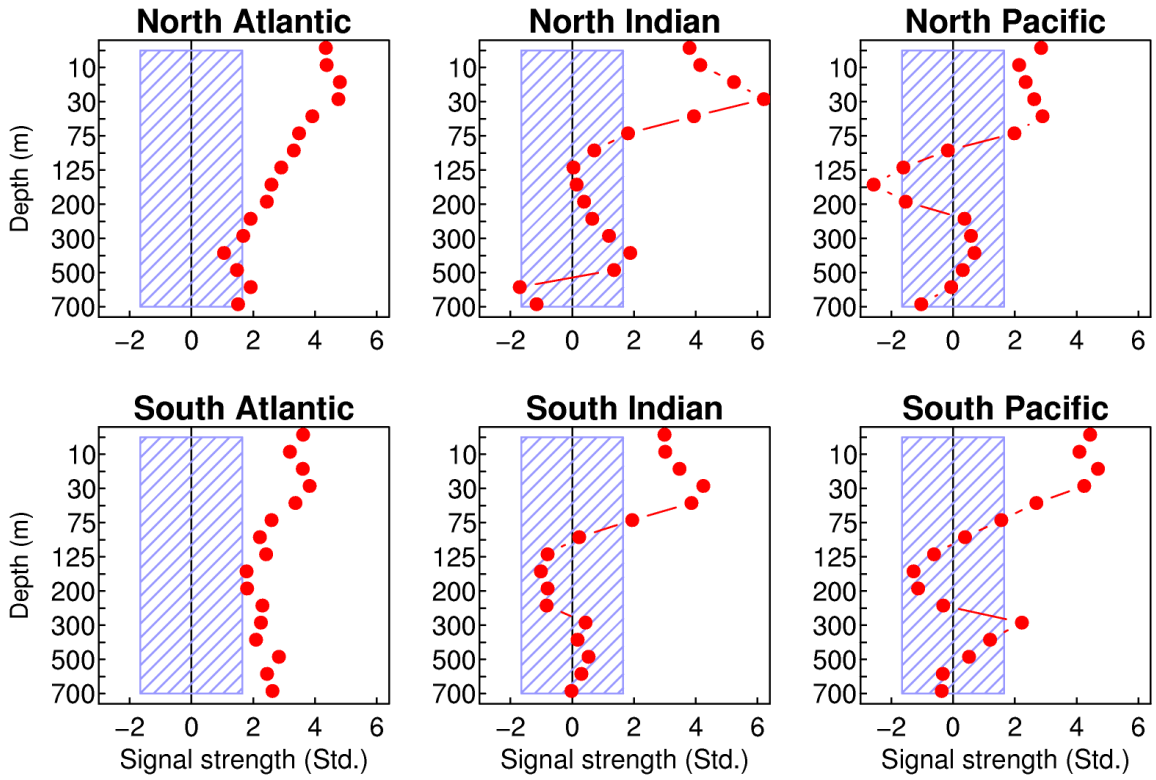

Fig. S2. Same as Fig. 2 in the main text, but with values normalized by the standard deviation of the ocean warming signal strength in the unforced PCM control run. The blue hatched region shows the 90% confidence interval from the control run, representing the range of signal strengths expected to arise from natural internal climate variability. Red dots show the signal strength in the observations.

Supporting Material References and Notes

S1. S. Levitus, J. Antonov, T. Boyer, *Geophys. Res. Lett.*, **32**, L02604, doi:10.1029/2004GL021592 (2005).

S2. J. K. Willis, D. Roemmich, and B. Cornuelle, *J. Geophys. Res.*, **109**, C12036, doi:10.1029/2003/JC002260 (2004).

S3. Barnett, T. P., G. Hegerl, T. Knutson, S. Tett, *J. Geophys. Res.*, **105** (D12), 15,525 (2000).

S4. J. M. Gregory, H.T. Banks, P. A. Stott, J. A. Lowe, M. D. Palmer, *Geophys. Res. Lett*. **31** L15312, 10.1029/2004GL020258 (2004).

S5. S. Sun, J. E. Hansen, *J. Clim*. **16**, 2807 (2003).

S6. D. W. Pierce, T. P. Barnett, K. M. AchutaRao, P. J. Gleckler, B. D. Santer, J. M. Gregory, W. M. Washington, in prep (2005).