**SI Text**

**1. Observational Data**

Observational data for $W_o$, the column-integrated atmospheric moisture content over oceans, were provided by Remote Sensing Systems (RSS) (Santa Rosa, CA) (1, 2). All analyses reported on here rely on version 6 of the SSM/I-derived $W_o$ data set produced by RSS. Data were available as monthly means on a 2.5° × 2.5° latitude/longitude grid and span the period July 1987 through December 2006.

We used version 2 of the NOAA Extended Reconstructed Sea Surface Temperature data set (ERSST) (3) for the SST variability calculations shown in Fig. 3*C*. ERSST data were available from January 1854 to December 2005 in the form of monthly means on a regular 2° × 2° latitude/longitude grid. Reconstruction of high-frequency SST anomalies involved use of empirically derived spatial modes of variability to interpolate observations in times of sparse coverage. Further details of the ERSST data set are available online (4).

**2. Modeling Groups Contributing to IPCC Database**

At the time this research was conducted, 15 modeling groups had performed a wide range of simulations in support of the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR4). Climate data from these simulations were made available to the scientific community through the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison (PCMDI). Five modeling groups provided $W$ results for at least two different model configurations. Results from a total of 22 different climate models were analyzed.

We considered two sets of simulations here: preindustrial control runs and 20CEN experiments with historical changes in a number of different anthropogenic and natural forcings. In IPCC terminology, these integrations are referred to as “picntrl” and “20c3m,” respectively.

Official designations of the 15 modeling groups that supplied *W* data are listed below (with model acronyms in brackets):

1. Bjerknes Center for Climate Research, Norway [BCCR-BCM2.0].

2. Canadian Centre for Climate Modeling and Analysis, Canada [CCCma-CGCM3.1(T47) and CCCma-CGCM3.1(T63)].

3. National Center for Atmospheric Research, U.S.A. [CCSM3 and PCM].

4. Météo-France/Centre National de Recherches Météorologiques, France [CNRM-CM3].

5. Commonwealth Scientific and Industrial Research Organization (CSIRO) Atmospheric Research, Australia [CSIRO-Mk3.0].

6. Max-Planck Institute for Meteorology, Germany [ECHAM5/MPI-OM].

7. Meteorological Institute of the University of Bonn, Meteorological Research Institute of the Korean Meteorological Agency, and Model and Data group, Germany/Korea [MIUB/ECHO-G].

8. Institute for Atmospheric Physics, China [FGOALS-g1.0].

9. Geophysical Fluid Dynamics Laboratory, U.S.A. [GFDL-CM2.0 and GFDL-CM2.1].

10. Goddard Institute for Space Studies, U.S.A. [GISS-AOM, GISS-EH, and GISS-ER].

11. Institute for Numerical Mathematics, Russia [INM-CM3.0].

12. Institute Pierre Simon Laplace, France [IPSL-CM4].

13. Center for Climate System Research, National Institute for Environmental Studies, and Frontier Research Center for Global Change, Japan [MIROC-CGCM3.2(medres) and MIROC-CGCM3.2(hires)].

14. Meteorological Research Institute, Japan [MRI-CGCM2.3.2].

15. United Kingdom Meteorological Office, Hadley Centre for Climate Prediction and Research, U.K. [UKMO-HadCM3 and UKMO-HadGEM1].

**3. Forcings Used in 20CEN Runs**

Details of the natural and anthropogenic forcings used by different modeling groups in their IPCC 20CEN simulations are given in SI Table 2. This table was compiled using information that participating modeling centers provided to PCMDI (5). All model acronyms used in the table are defined in the previous section.

A total of 11 different forcings are listed in SI Table 2. A letter "Y" denotes inclusion of a specific forcing. As used here, "inclusion" signifies the specification of time-varying forcings, with changes on interannual and longer time scales. Forcings that were varied over the seasonal cycle only, or not at all, are identified with a dash. A question mark indicates a case where there is uncertainty regarding inclusion of the forcing.

Results in SI Table 2 are stratified by inclusion or omission of volcanic forcing (ALL and ANTHRO, respectively). Ten of the 12 ALL models explicitly incorporated volcanic aerosols. Two models, MRI-CGCM2.3.2 and MIUB/ECHO-G, represented volcanic effects in a more indirect manner, using estimated volcanic forcing data from refs. 6 and 7, respectively, to adjust the solar irradiance at the top of the model atmosphere. As noted in the main text, the ALL vs. ANTHRO partitioning also separates models with "total" external forcing (natural plus anthropogenic) from models with primarily anthropogenic forcing.

Although all 15 modeling groups used very similar changes in well mixed GHGs, the changes in other forcings were not prescribed as part of the experimental design. In practice, each group used different combinations of 20th century forcings and often used different data sets for specifying individual forcings. End dates for the 20c3m experiment varied among groups and ranged from 1999 to 2003.

## 4. Fingerprint Analysis

**4.1 Regridding and Masking of Data.** Model results were available on different grids (SI Table 3). To calculate fingerprints from the averages of the ALL and ANTHRO model 20CEN runs, and to obtain "pooled" noise estimates from the ALL and ANTHRO model control integrations, we regridded 20CEN and control run *W* data from all 22 models to a common 10° × 10° latitude/longitude grid. Regridding to a relatively coarse-resolution grid reduces the spatial dimensionality of the input data sets, which is of benefit in the estimation of EOFs used in the fingerprint analysis. Because changes in *W* tend to be smoothly varying (Fig. 4 *A*–*D*), regridding does not lead to appreciable loss of information on the spatial structure of the leading signal or noise modes.

Each model has a "mask," $\vec{M}(j)$, of the ocean fraction on the original model grid. The arrow denotes a vector in *p*-dimensional space, where *p* is the total number of model gridpoints, and *j* is the index over the number of ALL or ANTHRO models. Because observed *W* data were available over ocean only, each model's land *W* values had to be appropriately masked out in the regridding process, i.e., any land gridpoints within a given 10° × 10° "target" grid cell were excluded from the calculation of the ocean *W* value for the target grid cell.

For each model, we calculated the ocean fraction at every target grid cell. Global-mean values of these fractions are generally different across models, reflecting differences in the original land/sea masks. Observed $W_o$ data and their associated ocean fraction were also transformed to the same 10° × 10° target grid.

**4.2 Definition of Fingerprint.** Let $\vec{S}(t,i,j)$ represent annual-mean $W_o$ data at time $t$ from the $i$th realization of the $j$th model's 20CEN experiment. Data are expressed as anomalies relative to the smoothed initial state (1900–1909) of the experiment. Here, the total time in years is $N_t = 100$ (because all 20CEN experiments cover the common period 1900–1999), and the total number of model gridpoints is $P = 291$ (after regridding to the common 10° × 10° latitude/longitude grid and masking out land points). The ALL and ANTHRO averages, $\vec{S}_{ALL}(t)$ and $\vec{S}_{ANTHRO}(t)$, were calculated by first averaging over an individual model's 20CEN realizations (where multiple realizations were available; see SI Table 3) and then averaging over models. Because the individual model land/sea masks are not identical after regridding, the number of models contributing to the multimodel averages varies near coastlines and in the vicinity of islands.

Finally, we calculated EOFs of the $\vec{S}_{ALL}(t)$ and $\vec{S}_{ANTHRO}(t)$ data sets. The fingerprints $\vec{F}_{ALL}$ and $\vec{F}_{ANTHRO}$ are simply the first EOFs of each data set, which explain a substantial fraction of the overall variance (89% and 76%, respectively, in the $\vec{S}_{ALL}(t)$ and $\vec{S}_{ANTHRO}(t)$ moisture data; see Fig. 4 *A* and *B*) and primarily capture the large increase in $W_o$ over the 20th century (not shown).

In calculating EOFs of $\vec{S}_{ALL}(t)$ and $\vec{S}_{ANTHRO}(t)$, we had to account for intermodel differences in $\vec{M}(j)$, the regridded ocean fraction. We did this in the following way. First, the regridded $\vec{M}(j)$ values were set to zero at any grid cell with <1% ocean coverage. We then computed $\vec{M}_{ALL}$ and $\vec{M}_{ANTHRO}$, the geometrical means of the ocean fraction for the ALL and ANTHRO models. Use of the geometrical mean excludes areas in which any model has zero ocean fraction. Next, we calculated the combined geometrical mean ocean fraction, $\vec{M}_{COMB}$, which is the geometrical mean of $\vec{M}_{ALL}$, $\vec{M}_{ANTHRO}$, and the observed ocean fraction, $\vec{M}_{OBS}$. Use of $\vec{M}_{COMB}$ ensures that all EOF calculations (and all calculations in the subsequent determination of detection time) are performed on a common grid, with a common land/sea mask. Appropriate weights are carried throughout the EOF analysis. For each grid cell, the weight is the product of the combined geometrical mean ocean fraction and the grid cell's area weight.

**4.3 Calculation of Concatenated Noise Data Sets.** As noted in the main text, optimal fingerprint techniques typically require two different data sets for estimating the background noise of natural internal variability. One data set is required for optimizing the fingerprint, and the second is used for estimating the statistical significance of results. Here, we generated two noise data sets by concatenating *W* data from individual control runs. We did this separately for the ALL and ANTHRO control runs.

For example, for the 12 ALL model control runs, we regridded annual-mean *W* data to the same target $10° \times 10°$ grid used for fingerprint estimation, formed anomalies for each control run (relative to its overall time mean), and then concatenated these anomalies to form the noise data set $\vec{C}_{ALL}(t)$. The index *t* denotes the single concatenated time dimension. There are a total of 4,440 years of ALL model control run data. The 10 ANTHRO control runs were processed in a similar manner to form $\vec{C}_{ANTHRO}(t)$, with a total time dimension of 4,418 years. In calculating EOFs of $\vec{C}_{ALL}(t)$ and $\vec{C}_{ANTHRO}(t)$, we used the geometrical mean ocean fraction masks ($\vec{M}_{ALL}$ and $\vec{M}_{ANTHRO}$) that were described in the previous section. The first EOFs of $\vec{C}_{ALL}(t)$ and $\vec{C}_{ANTHRO}(t)$ explain 28% and 44% of the total variance of the concatenated control run data, respectively (see Fig. 4 *C* and *D*).

In view of substantial intermodel differences in the length of the available control runs (SI Table 3), we also generated $\vec{C}_{ALL}(t)$ and $\vec{C}_{ANTHRO}(t)$ data sets using equal lengths of control run (the first 100 years) from each model, yielding 1,200 years of data for the ALL concatenated control runs and 1,000 years for the ANTHRO controls. The EOFs estimated from these much shorter data sets are highly similar to those obtained using all available control run data. This indicates that in our specific example, the use of information from models with longer control runs does not distort the EOF structure.

Although the partitioning of 20CEN results into ALL and ANTHRO model groups is logical for the purposes of fingerprint estimation (see main text and SI Table 2), it is less meaningful for the control runs because these have no volcanic forcing. Other ways could have been devised for separating the control runs into the two data sets required for optimal fingerprinting. Here, we use the ALL vs. ANTHRO partitioning of the noise data for the sake of consistency with our fingerprint definition approach.

**4.4 Method for Estimation of Detection Time.** We begin with regridded annual-mean observational data, $\vec{O}(t)$ (from SSM/I), and the concatenated noise data from the ALL and ANTHRO model control integrations, $\vec{C}_{ALL}(t)$ and $\vec{C}_{ANTHRO}(t)$. Observed data are expressed as anomalies relative to climatological annual means over the entire SSM/I period (1988–2006); control run anomalies are defined as described above.

Two forms of detection time are computed: nonoptimized (“raw”) and optimized. We consider the raw case first and assume for illustrative purposes that both the fingerprint and the noise have been obtained from ALL model output. To define raw detection times, $\vec{O}(t)$ and $\vec{C}_{ALL}(t)$ are projected onto the fingerprint $\vec{F}_{ALL}$, yielding (respectively) a test statistic time series $Z(t)$ and a “signal-free” time series $N(t)$. We fit least-squares linear trends of increasing length $L$ to $Z(t)$ and then compare these with the standard error of the distribution of nonoverlapping $L$-length trends in $N(t)$. Detection is stipulated to occur when the trend in $Z(t)$ exceeds and remains above the 5% significance level. The test is one-tailed, and we assume a Gaussian distribution of trends in $N(t)$.

The start date for fitting linear trends to $Z(t)$ is 1988, the first complete year of the SSM/I data. We use a minimum trend length of 10 years, so the earliest possible detection time is in 1997. To explore the sensitivity of our results to the choice of fingerprint and noise data sets, we calculate detection times for all four possible combinations of the fingerprints $\vec{F}_{ALL}$ and $\vec{F}_{ANTHRO}$ and the noise data sets $\vec{C}_{ALL}(t)$ and $\vec{C}_{ANTHRO}(t)$ (see Table 1).

Optimized detection times are determined similarly but involve projection of $\vec{O}(t)$ and $\vec{C}_{ALL}(t)$ onto $\vec{F}^{*}_{ALL}$, a version of the fingerprint that has been rotated away from high noise directions (here, and subsequently, the asterisk denotes an optimized version of the fingerprint). This rotation is performed in the subspace of the first $m$ EOFs of $\vec{C}_{ALL}(t)$, where $m$ is the so-called "truncation dimension." We examined the sensitivity of optimized detection times by using three different choices of $m$ (5, 10, and 15). To avoid the introduction of "artificial skill," the same noise data set is never used for both optimizing the fingerprint and estimating the "signal free" time series $N(t)$ (8). This is why we require two noise data sets, $\vec{C}_{ALL}(t)$ and $\vec{C}_{ANTHRO}(t)$. Full details of the detection method are given elsewhere (8).

Given the short observational record lengths, we use only the spatial properties of signal and noise in rotating $\vec{F}^{*}_{ALL}$ and $\vec{F}^{*}_{ANTHRO}$. Other detection work involving longer data sets with more temporal structure has used both spatial and temporal information for fingerprint optimization (9).

Because residual control run drift was not subtracted in the formation of the $\vec{C}_{ALL}(t)$ and $\vec{C}_{ANTHRO}(t)$ data sets, there are several large "jumps" in the $N(t)$ time series at the transitions between individual control runs (SI Figs. 11 and 12). The most obvious example is the jump between the end of the GFDL-CM2.1 control run and the beginning of the GISS-EH control (SI Fig. 11). For the purposes of estimating detection time, such discontinuities inflate the standard deviation of the sampling distribution of the $L$-length trends that we fit to $N(t)$, and hence make it more difficult to obtain a statistically significant trend in the signal time series $Z(t)$. Our significance testing procedure is therefore conservative.

**4.5 Estimation of Detection Time: An Example.** SI Fig. 8 provides a specific example of how we estimate detection time. In this example, $Z(t)$ is the time series of coefficients for the projection of the SSM/I $W_o$ data onto the raw and optimized ANTHRO fingerprints, $\vec{F}_{ANTHRO}$ and $\vec{F}^{*}_{ANTHRO}$. $N(t)$ and $N^*(t)$ are the projections of $\vec{C}_{ANTHRO}(t)$, the concatenated ANTHRO model control run data, onto $\vec{F}_{ANTHRO}$ and $\vec{F}^{*}_{ANTHRO}$, respectively.

Trends in $Z(t)$ are displayed as a function of increasing trend length $L$ (SI Fig. 8*A*). In the "raw fingerprint" case, the largest $Z(t)$ trend is for the 11-year period ending in 1998. This is probably due to the influence of the large El Niño in 1997/1998 on observed $W_o$ data. For $L > 12$ years, trends in $Z(t)$ increase and then fluctuate around an asymptotic value. Trends in the projection of the SSM/I data onto the optimized fingerprint reach a similar asymptotic value but show a smoother initial increase, which suggests that optimization is successfully rotating the fingerprint away from high ENSO noise directions (see SI Fig. 7) and thus damping the influence of ENSO on $Z(t)$ trends.

As the trend interval $L$ increases, there is a reduction in the standard error of the sampling distributions of trends in $N(t)$ (SI Fig. 8*B*). This decrease may be due not only to a decrease in noise amplitude with increasing trend length, but also to dissimilarity between the fingerprint pattern and the patterns of low-frequency noise in $\vec{C}_{ANTHRO}(t)$. SI Fig. 8 *A* and *B* clearly illustrate that the increase in the S/N plotted in SI Fig. 8*C* arises primarily from the decrease in the standard error of the noise trends with longer trend interval $L$. In the optimized case, the growth in S/N is also due to an increase in the signal with increasing $L$.

**4.6 Analysis with Mean Removed.** In the "mean-removed" case, the spatial means of $\breve{O}(t)$, $\breve{S}_{ALL}(t)$, $\breve{S}_{ANTHRO}(t)$, $\breve{C}_{ALL}(t)$, and $\breve{C}_{ANTHRO}(t)$ were removed (at each gridpoint and at each time, $t$) before calculation of fingerprints, noise EOFs, and detection times.

## 5. Details of Other Statistical Analyses

**5.1 Calculation of Temporal Standard Deviations.** All temporal standard deviations estimated from the $W_o$ data shown in Fig. 3 were calculated by using linearly detrended data. This was done because some of the model simulations examined here (and the SSM/I and ERSST data) have large trends in atmospheric moisture or SST, which inflate the temporal variance.

**5.2 Calculation of Confidence Intervals for Linear Trends.** The error bars on the SSM/I $W_o$ trend in Fig. 3*B* are the "adjusted" $2\sigma$ confidence intervals for $b$, the slope parameter of the estimated least-squares linear trend in the observed data (10). The adjustment for temporal autocorrelation assumes a lag-1 autocorrelation structure of the trend residuals, $e(t)$. The lag-1 autocorrelation coefficient of $e(t)$ is used to compute an effective sample size, $n_e$, and to adjust $s_b$, the standard error of $b$. Strong temporal autocorrelation of $e(t)$ results in $n_e << n$ (the actual number of time samples) and inflates $s_b$. For the monthly-mean $W_o$ data analyzed here, $n_e$ is typically an order of magnitude smaller than $n$.

**5.3 Digital Filtering.** For display purposes, the modeled and observed $\langle W_o \rangle$ data in Fig. 1 were smoothed using a digital filter (11) with a window width of $K = 21$ months, corresponding to a half-power point of 25 months (see SI Fig. 13). This damps variability on interannual and ENSO time scales, while information on the atmospheric moisture response to volcanic forcing is largely preserved. The overall linear trend was subtracted before filtering and reinserted after filtering. Data loss was avoided by "reflecting" $(K - 1)/2$ points at the beginning and end of the time series. The same filter was used in the variability calculations shown in Fig. 3*B*. The $\langle W_o \rangle$ data from the PCM and

MIROC3.2(medres) runs in Fig. 5 were digitally filtered with a window width of $K = 145$ months, which corresponds to a half-power point of 119 months.


1. Wentz FJ, Schabel M (2000) *Nature* 403:414–416.

2. Wentz FJ, Ricciardulli L, Hilburn K, Mears C (2007) *Science* 317:233–235.

3. Smith TM, Reynolds RW (2004) *J Clim* 17:2466–2477.

4. National Climatic Data Center (2007) *Sea Surface Temperature (SST) Datasets (Gridded Data)*. Available at http://www.ncdc.noaa.gov/oa/climate/research/sst/sst.html#grid.

5. Program for Climate Model Diagnosis and Intercomparison (2007) *About the WCRP CMIP3 Multi-Model Dataset Archive at PCMDI*. Available at http://www-pcmdi.llnl.gov/ipcc/about_ipcc.php.

6. Sato M, Hansen JE, McCormick MP, Pollack JB (1993) *J Geophys Res* 98:22987–22994.

7. Crowley TJ (2000) *Science* 289:270–277.

8. Santer BD, Mikolajewicz U, Brüggemann W, Hasselmann K, Höck H, Maier-Reimer E, Wigley TML (1995) *J Geophys Res* 100:10693–10725.

9. Stott PA, Tett SFB, Jones GS, Allen MR, Mitchell JFB, Jenkins GJ (2000) *Science* 290:2133–2137.

10. Santer BD, Wigley TML, Boyle JS, Gaffen DJ, Hnilo JJ, Nychka D, Parker DE, Taylor KE (2000) *J Geophys Res* 105:7337–7356.

11. Lynch P, Huang X-Y (1992) *Mon Weather Rev* 120:1019–1034.