Forced and unforced ocean temperature changes in Atlantic and Pacific tropical cyclogenesis regions

B. D. Santer[a,b], T. M. L. Wigley[c], P. J. Gleckler[a], C. Bonfils[d], M. F. Wehner[e], K. AchutaRao[a], T. P. Barnett[f], J. S. Boyle[a], W. Brüggemann[g], M. Fiorino[a], N. Gillett[h], J. E. Hansen[i], P. D. Jones[h], S. A. Klein[a], G. A. Meehl[c], S. C. B. Raper[j], R. W. Reynolds[k], K. E. Taylor[a], and W. M. Washington[c]

[a]Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA 94550; [c]National Center for Atmospheric Research, Boulder, CO 80307; [d]University of California, Merced, CA 95344; [e]Lawrence Berkeley National Laboratory, Berkeley, CA 94720; [f]Scripps Institution of Oceanography, La Jolla, CA 92037; [g]Institut für Unternehmensforschung, Universität Hamburg, 22765 Hamburg, Germany; [h]Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, United Kingdom; [i]National Aeronautics and Space Administration/Goddard Institute for Space Studies, New York, NY 10025; [j]Centre for Air Transport and the Environment, Manchester Metropolitan University, Manchester M1 5GD, United Kingdom; and [k]National Oceanic and Atmospheric Administration/National Climatic Data Center, Asheville, NC 28801

Edited by Isaac M. Held, National Oceanic and Atmospheric Administration, Princeton, NJ, and approved July 24, 2006 (received for review April 7, 2006)

Previous research has identified links between changes in sea surface temperature (SST) and hurricane intensity. We use climate models to study the possible causes of SST changes in Atlantic and Pacific tropical cyclogenesis regions. The observed SST increases in these regions range from 0.32°C to 0.67°C over the 20th century. The 22 climate models examined here suggest that century-timescale SST changes of this magnitude cannot be explained solely by unforced variability of the climate system. We employ model simulations of natural internal variability to make probabilistic estimates of the contribution of external forcing to observed SST changes. For the period 1906–2005, we find an 84% chance that external forcing explains at least 67% of observed SST increases in the two tropical cyclogenesis regions. Model ''20th-century'' simulations, with external forcing by combined anthropogenic and natural factors, are generally capable of replicating observed SST increases. In experiments in which forcing factors are varied individually rather than jointly, human-caused changes in greenhouse gases are the main driver of the 20th-century SST increases in both tropical cyclogenesis regions.

Hurricane activity is influenced by a variety of physical factors, such as sea surface temperatures (SSTs), wind shear, moisture availability, and atmospheric stability (1). Theory, observations, and modeling provide evidence of a direct link between changes in SSTs and hurricane intensity (2–6). One recent investigation found that secular SST changes in the Atlantic and Pacific tropical cyclogenesis regions (ACR and PCR) were highly correlated with a measure of hurricane intensity based on maximum wind speeds (6). This research raises an important question: What are the causes of past SST changes in areas where hurricanes develop?

This question is timely in view of the unprecedented level of activity during the 2005 Atlantic hurricane season (7) and evidence that a recent increase in the number of category 4 and 5 hurricanes is largely SST-driven (8, 9). There are, however, conflicting estimates of the relative contributions of internal climate variability and external forcing to observed SST changes. Some analyses suggest that 20th century SST changes in the ACR can be fully explained by internal variability of the climate system (10). In contrast, detection and attribution studies find a substantial anthropogenic component in observed increases in upper ocean heat content (11–13). Such work has examined the behavior of ocean heat content averaged over large ocean basins, while our investigation focuses on elucidating the causes of SST changes in the much smaller ACR and PCR.[l]

Previous research has relied on observational data to assess the relative contributions of internal noise and external forcing to SST changes in tropical cyclogenesis regions (7, 14). Partitioning of signal and noise components is difficult to achieve with observations alone. In the real world, human-induced changes in external climate forcings are superimposed on (and may even modulate) natural internal climate variability. We do not have a control experiment without anthropogenic forcings, which could be used to isolate and quantify climate noise. Such systematic experimentation can be performed only with numerical models of the climate system.

Model and Observational Data

Here, we use 22 different climate models to estimate the magnitude of SST changes arising from internally generated variability and external forcing. Our focus is on SST changes in the ACR and PCR on timescales of the last 20–100 years. We analyze both control simulations with no forcing changes and 20th-century (20CEN) experiments with estimated historical changes in external forcings (15). 20CEN forcings were not standardized across different modeling groups (see *Supporting Text*, which is published as supporting information on the PNAS web site). The 20CEN results therefore reflect differences and uncertainties in the applied forcings and in the physics and parameterizations of the models themselves. The most comprehensive experiments include changes in combined natural external forcings (solar irradiance and volcanic dust loadings in the atmosphere) and in a wide variety of anthropogenic influences (such as well mixed greenhouse gases, ozone, sulfate and black carbon aerosols, and land surface properties). All simulations were performed with coupled atmosphere–ocean General Circulation Models, in which SST changes are predicted.

Model SSTs are compared with the Extended Reconstructed SST data set (ERSST) of the National Oceanic and Atmospheric Administration (NOAA) (16) and the Hadley Centre Sea Ice and SST data set (HadISST) (17). The aim of these comparisons is to determine whether observed SST changes in the ACR and PCR can be explained by internally generated variability estimated from control simulations, and to evaluate how successfully the 20CEN runs capture important features of the observed SST behavior in these two tropical cyclogenesis regions. Use of both ERSST and HadISST data provides information on the sensitivity of our results to structural uncertainties in the observations (15, 18).

Observed and Modeled SST Time Series

We consider the observations first. In the smoothed ERSST and HadISST data (19), SSTs in the ACR were at record levels during

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: ACR, Atlantic tropical cyclogenesis region; 20CEN, 20th-century; ERSST, Extended Reconstructed Sea Surface Temperature data set; HadISST, Hadley Centre Sea Ice and Sea Surface Temperature data set; IPCC, Intergovernmental Panel on Climate Change; NOAA, National Oceanic and Atmospheric Administration; PCM, Parallel Climate Model; PCR, Pacific tropical cyclogenesis region; SST, sea surface temperature.

[b]To whom correspondence should be addressed. E-mail: santer1@llnl.gov.

[l]The ACR and PCR used here are identical to those defined in ref. 6. Gridded, monthly mean model and observational SST data were spatially averaged over 6°N–18°N, 60°W–20°W (ACR) and over 5°N–15°N, 180°E–130°E (PCR).

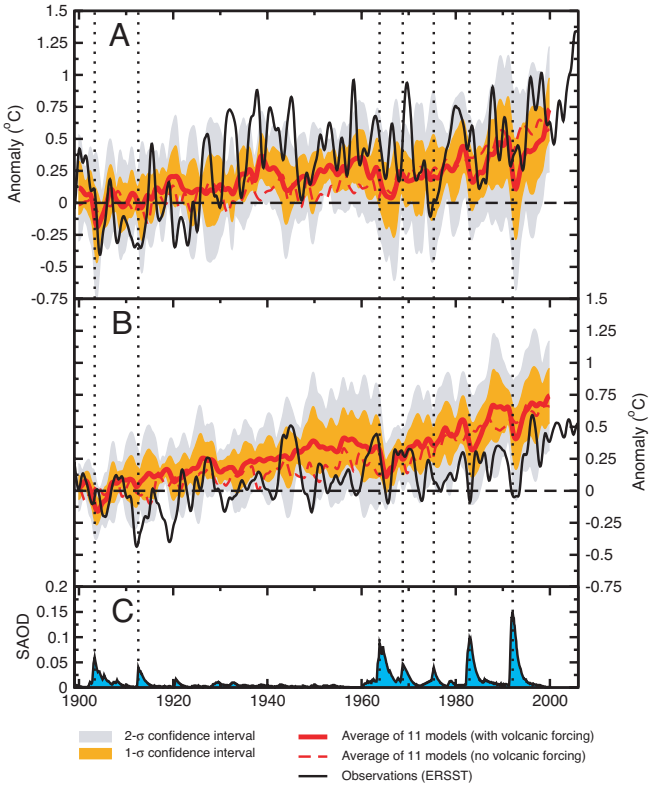

Fig. 1. Modeled and observed SST changes in tropical cyclogenesis regions and observed changes in stratospheric aerosol optical depth (SAOD). Time series of monthly mean, spatially averaged SST anomalies are for the ACR (*A*) and PCR (*B*). Observational results are from the NOAA ERSST data set (16). Results for a second observational data set (17) are very similar (see Fig. 6). Model data from simulations of 20CEN climate change are partitioned into two groups, with and without volcanic forcing (V and No-V). All model data were low-pass filtered (with window width $W = 21$ months; see Fig. 7, which is published as supporting information on the PNAS web site) before formation of V and No-V averages (19). ERSST data were smoothed with the same filter. The yellow and gray envelopes are the 1σ and 2σ confidence intervals for the V averages, calculated with the smoothed data at each time *t*. Because most 20CEN experiments end in December 1999, V and No-V averages are calculated only until that month. ERSST data are shown through December 2005. All SST anomalies were defined relative to climatological monthly means over 1900 through 1909. This reference period was chosen for visual display purposes only, and it has no impact on subsequent trend analyses or variability estimates. The amplitudes of the observed and simulated SST variability are not directly comparable, because the latter was damped by averaging over different realizations and models. (*C*) Estimate of the SAOD (21). Dotted vertical lines denote the times of maximum SAOD during major volcanic eruptions.

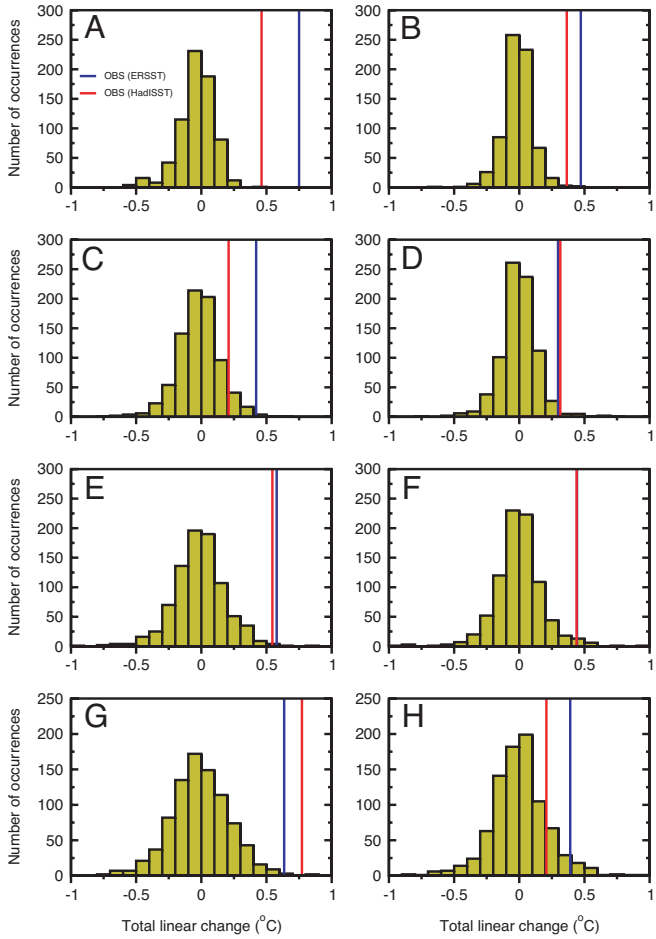

Fig. 2. Comparison between observed and simulated SST changes in the ACR (*A*, *C*, *E*, and *G*) and PCR (*B*, *D*, *F*, and *H*). Results are expressed as the total linear change, $b \times n$, where *b* is the slope parameter of the least-squares linear trend (in °C/month) and *n* is the total number of months. Trend comparisons are made on four different timescales: 100 years (*A* and *B*), 50 years (*C* and *D*), 30 years (*E* and *F*), and 20 years (*G* and *H*). Observed ACR and PCR SST trends from HadISST and ERSST were calculated over the periods 1906–2005, 1956–2005, 1976–2005, and 1986–2005. Sampling distributions of unforced SST trends on 100-, 50-, 30-, and 20-year timescales were computed as described in the main text. For visual display purposes, unforced SST trends were fitted to segments of the ACR and PCR anomaly time series that overlapped by all but 10 years. For the century-timescale results, this procedure yields 698 unforced SST trends for each tropical cyclogenesis region. Unforced trends are plotted in the form of histograms. Very similar (but less smooth) histograms are obtained if trends are fitted to nonoverlapping segments of control run SST data. Red and blue vertical lines indicate observed SST trends in the HadISST and ERSST data, respectively.

the 2005 Atlantic hurricane season (Fig. 1*A* and Fig. 6, which is published as supporting information on the PNAS web site).[m] The 2005 SST anomaly was smaller in the PCR and not unprecedented (Figs. 1*B* and 6*B*). Observed SSTs in both tropical cyclogenesis regions have increased over the 20th century, with total linear changes in HadISST and ERSST data of 0.41°C and 0.67°C in the ACR (respectively) and 0.32°C and 0.38°C in the PCR. Differences between observational data sets primarily reflect the different procedures used by the NOAA and Hadley Centre groups to infill missing SST data (16, 17).

Variability on sub- to multidecadal timescales is superimposed on these overall increases in observed SSTs (Fig. 1). Commonly discussed sources of this variability are the El Niño/Southern Oscillation and the Atlantic Multidecadal Oscillation (7, 20). In the ERSST and HadISST data, part of this variability is in phase with fluctuations in the optical depth of stratospheric aerosols produced by massive volcanic eruptions (21) (Figs. 1 and 6). This result is consistent with the identification of volcanic effects (albeit at much larger spatial scales) in many different climate variables (22–24). The relationship between SST variability and stratospheric aerosol optical depth is clearer in the PCR than in the ACR, particularly for the eruption of Mt. Pinatubo in June

[m]For visual display, the modeled and observed SST data in Figs. 1 and 6 were smoothed by using a digital filter (19) with a window width *W* of 21 months, corresponding to a half-power point of 25 months. This smoothing damps variability on interannual and El Niño/Southern Oscillation timescales, while information on the SST response to volcanic forcing is largely preserved. The overall linear trend was subtracted before filtering and was reinserted after filtering. Data loss was avoided by ''reflecting'' $(W - 1)/2$ points at the beginning and end of the time series. To estimate modeled and observed variability on decadal and longer timescales (Fig. 4*C*), we applied the same digital filter to the detrended SST anomaly data and set $W = 145$ months, yielding a half-power point at 119 months. The response functions for both choices of *W* are shown in Fig. 7.

Table 1. Statistics for observed and simulated SST trends in the ACR and PCR

Region	Dataset	Length, yr	Period	b_{OBS}	s_{CTL}	b_{OBS}/s_{CTL}	p_1	p_2
ACR	HadISST	100	1906–2005	0.046	0.015	3.158	0.000***	0.024**
ACR	HadISST	50	1956–2005	0.042	0.035	1.204	0.074*	0.177
ACR	HadISST	30	1976–2005	0.182	0.064	2.838	0.010***	0.014**
ACR	HadISST	20	1986–2005	0.386	0.117	3.291	0.002***	0.005***
ACR	ERSST	100	1906–2005	0.075	0.015	5.125	0.000***	0.000***
ACR	ERSST	50	1956–2005	0.084	0.035	2.414	0.006***	0.029**
ACR	ERSST	30	1976–2005	0.193	0.064	3.014	0.007***	0.010***
ACR	ERSST	20	1986–2005	0.318	0.117	2.710	0.005***	0.014**
PCR	HadISST	100	1906–2005	0.036	0.012	3.061	0.000***	0.012**
PCR	HadISST	50	1956–2005	0.063	0.030	2.124	0.011**	0.034**
PCR	HadISST	30	1976–2005	0.147	0.065	2.243	0.010***	0.028**
PCR	HadISST	20	1986–2005	0.104	0.105	0.987	0.140	0.255
PCR	ERSST	100	1906–2005	0.047	0.012	3.959	0.000***	0.012**
PCR	ERSST	50	1956–2005	0.059	0.030	2.004	0.017**	0.046**
PCR	ERSST	30	1976–2005	0.148	0.065	2.261	0.010***	0.028**
PCR	ERSST	20	1986–2005	0.195	0.105	1.855	0.038***	0.063*

The observed SST trends in the ACR and PCR (b_{OBS}; °C/decade) are given for the last 100, 50, 30, and 20 years of the HadISST and ERSST data sets. The standard deviation of the model-based sampling distribution of unforced SST trends (s_{CTL}; °C/decade) was calculated from nonoverlapping segments of the control run SST time series (see Fig. 2 and main text). The ratio b_{OBS}/s_{CTL} is a simple measure of signal to noise. The p values are estimates of the probability that b_{OBS} could be caused by (model-simulated) natural internal variability alone. Probabilities are based on tests against actual and absolute values of unforced SST trends (p_1 and p_2, respectively). All trends were calculated with monthly mean, spatially averaged anomaly data. Trend significance at the 10%, 5%, and 1% levels is indicated by one, two, or three asterisks, respectively.

1991 (Figs. 1 and 6). Regional differences in the observed SST changes after volcanic eruptions are expected, partly because of spatial differences in climate noise (25).

Eleven of the 22 historical forcing experiments included some representation of volcanic effects on climate (see *Supporting Text* and Tables 2 and 3, which are published as supporting information on the PNAS web site). We therefore partitioned the 20CEN results in Fig. 1 into two sets, with and without volcanic forcing (V and No-V, respectively).[n] The pronounced differences between the V and No-V averages during major eruptions support the observational evidence of volcanically induced cooling of SSTs in both tropical cyclogenesis regions.

Comparison of Observed and Unforced SST Trends

To assess whether observed ACR and PCR trends could be due to climate noise alone, we used information from 22 model control runs to generate sampling distributions of the unforced SST trends in these regions (Fig. 2). For each control run, least-squares linear trends were fitted to successive nonoverlapping segments of the ACR and PCR anomaly time series (Figs. 8 and 9, which are published as supporting information on the PNAS web site). Results from the 22 models were combined to obtain "multimodel" sampling distributions of unforced SST trends. This was done for timescales of 100, 50, 30, and 20 years, yielding trend sample sizes of $N_t = 84$, 175, 287, and 444, respectively (see *Supporting Text*). Observed SST trends in both tropical cyclogenesis regions were calculated with the last 100, 50, 30, and 20 years of the HadISST and ERSST data sets.

We then estimated p values by comparing the observed SST trend, b_{OBS}, with both actual and absolute values of $b_{CTL}(i)$, $i = 1$, N_t, the unforced SST trends from the multimodel sampling distributions. In 29 of 32 cases (2 cyclogenesis regions × 2 observational data sets × 4 trend lengths × 2 different methods of estimating p values), the null hypothesis that observed SST trends could be explained by natural internal variability (as simulated in current climate models) is rejected at the 10% level or better (Table 1). Our finding that observed SST trends in the ACR and PCR are significantly larger than model-based estimates of unforced SST variability is therefore relatively insensitive to observational uncertainty, the timescale over which trends are calculated, and the details of our significance testing strategy.

The p values partly obscure the expected relationship between the timescale of SST changes and the relative sizes of observed and unforced trends (Fig. 2). Because the amplitude of unforced variability decreases with an increase in the temporal averaging period, a slowly evolving greenhouse-gas-induced warming signal should be more easily discernible at longer than at shorter timescales (26). Such relationships are more clearly revealed by using the signal-to-noise ratio b_{OBS}/s_{CTL}, where s_{CTL} is the standard deviation of the sampling distribution of unforced SST trends (Table 1). While b_{OBS} trends over the past 100 years are at least 3.2–5.1 times larger than s_{CTL}, observed SST changes over the past 20 to 30 years typically have smaller ratios of b_{OBS}/s_{CTL}, particularly in the PCR, where they vary from 1.0 to 2.3. In the ACR, however, partly because of the unusual warmth of 2005 in the observational record (Fig. 1), even b_{OBS} trends over the past 20–30 years are 2.7–3.3 times larger than values of s_{CTL}.

Contribution of External Forcing to Observed SST Trends

The results from Figs. 1 and 2 can be used to estimate the contribution of external forcing to observed SST trends in the ACR and PCR. We do this in two different ways: (*i*) by comparing observed trends with model-based estimates of unforced trends, and (*ii*) by comparing the forced SST changes in the 20CEN experiments with observations. The first approach has the advantage that the "spread" of model-based noise estimates arises solely from structural differences in the models (e.g., in terms of physics, parameterizations, resolution, and spin-up procedures), whereas the second approach uses experiments that convolve differences in model structure and the applied external forcings.

[n]Ensembles of the 20CEN simulations were performed with 13 of the 22 models analyzed here (see *Supporting Text*). Each ensemble contains multiple realizations of the same experiment, differing only in their initial conditions, but with identical changes in external forcings. This procedure yields many different realizations of the noise that is superimposed on the climate "signal" (the response to the imposed forcing changes). Averaging over multiple realizations reduces noise and facilitates signal estimation. Here, we calculated averages over V and No-V 20CEN runs. In each case, $\overline{X}$ is the arithmetic mean of the ensemble means (for the models for which ensembles are available) and of individual realizations, i.e., $\overline{X} = (1/N_m)\sum_{j=1}^{N_m} \overline{X}_j$, where N_m is the total number of V or No-V models (11 here), and $\overline{X}_j$ is the ensemble mean signal (or individual realization) of the jth model. This weighting avoids undue emphasis on results from a single model with a large number of realizations.

In the first approach, we assume that an observed SST trend, b_{OBS}, can be decomposed into b_{EXT}, the true (but unknown) slope of the SST trend in response to external forcing, and b_{INT}, the slope of the SST trend arising from a (random) realization of natural internal variability. The percentage contribution of external forcing to the observed trend can be estimated by $F_1 = 100[(b_{OBS} - b_{INT})/b_{OBS}]$. In the real world, b_{INT} may be either positive (contributing to the observed warming) or negative (offsetting some portion of externally forced warming). Assuming that the model-based estimates of internal variability are reasonable estimates of the true amplitude of internal noise, and that the sampling distribution of this unforced trend component (derived from control run data) is Gaussian with zero mean and standard deviation s_{CTL}, the 68% confidence interval for b_{INT} is $(-s_{CTL}, +s_{CTL})$, which can be easily transformed into a corresponding confidence interval for F_1. This procedure yields F_1 values in the range $100 - D$ to $100 + D$, where $D = 100\ (s_{CTL}/b_{OBS})$. There is therefore a 16% chance that the signal percentage is less than $100 - D$, and a 16% chance that the signal percentage exceeds $100 + D$.

In the second approach, we assume that the 20CEN runs provide reliable estimates of b_{EXT}. As in the case of b_{INT}, a 68% confidence interval can be specified for b_{EXT}, i.e., $(\overline{b_V} - s_V, \overline{b_V} + s_V)$, where $\overline{b_V}$ is the model-average SST trend in the subset of 20CEN runs with volcanic forcing, and s_V is the intermodel standard deviation of SST trends in the V models. Under this assumption, the percentage contribution of external forcing to b_{OBS} is estimated by $F_2 = 100\ (\overline{b_V}/b_{OBS})$, and the $\pm 1\sigma$ range of $\overline{b_V}$ yields the error bars on the F_2 results in Fig. 3.[o]

Consider the results for the century-timescale observed trends. Values of F_1 are symmetrical around 100% (Fig. 3). Based on the multimodel sampling distributions of unforced SST trends (and on one-tailed tests), there is an 84% chance that the externally forced component of observed SST increases in the ACR and PCR is at least 67%, and an 84% chance that this component is no greater than 133%. For central values of $\overline{b_V}$, F_2 yields a larger range (55–184%) for this externally forced component. The F_2 error bars overlap with the F_1 ranges, demonstrating consistency in the signal-to-noise partitioning obtained with the two methods. This implies that our finding of $\overline{b_V} > b_{OBS}$ in the PCR is not inconsistent with an "offsetting" of an externally forced warming by a century-timescale natural cooling trend. Clearly, model error (in both the applied 20CEN forcings and the model responses) may also be important in explaining why $\overline{b_V} > b_{OBS}$.

Model Performance in Simulating Means, Variability, and Trends

The p values and F_1 results in the previous sections are only as reliable as the model-based estimates of climate noise on which they are based. The p values in Table 1 could be spuriously low (and the F_1 values in Fig. 3 spuriously high) if there were a systematic underestimate of internally generated variability in the models used here. We tried to guard against this possibility by using a large number of models to estimate s_{CTL}.

Although we lack sufficiently long observational records to evaluate model estimates of century-timescale variability, the data are of adequate length for assessing simulated SST variability on subdecadal to decadal timescales. We use the 20CEN simulations to compare modeled and observed means, variability and trends. A discussion of model performance in simulating the climatological seasonal cycles of ACR and PCR SSTs is given in the *Supporting Text* (see Fig. 10, which is published as supporting information on the PNAS web site).

[o] F_1 is calculated with observed trends over 1906–2005, 1956–2005, etc., whereas F_2 is based on b_{OBS} and $\overline{b_V}$ trends over 1900–1999 only. This is because most of the 20CEN experiments end in 1999, thus hampering direct comparisons with the full observational record.

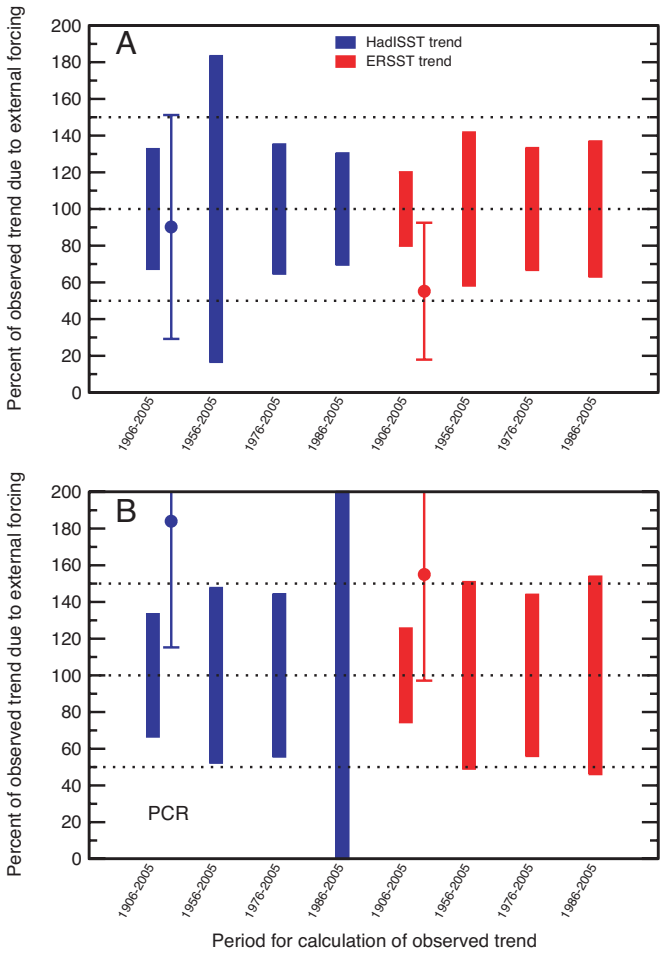

Fig. 3. Estimates of the percentage contribution of external forcing to observed SST changes in the ACR (*A*) and PCR (*B*). Results are for F_1 (solid bars) and F_2 (circles and thin error bars). For definitions of F_1 and F_2, refer to main text. In computing F_1, model estimates of s_{CTL} were obtained from histograms similar to those shown in Fig. 2, but based on trends fitted to nonoverlapping rather than overlapping segments of SST time series.

Most models systematically underestimate the climatological annual-mean SST in the ACR and PCR (Fig. 4*A*). There is no evidence of such a systematic underestimate in the temporal standard deviation of unfiltered SST anomalies, which is dominated by variability on interannual and El Niño/Southern Oscillation timescales (Fig. 4*B*). In the ACR (PCR), roughly one-third (two-thirds) of the 60 20CEN realizations overestimate observed SST variability. These variance differences are not statistically significant.

The model results in Fig. 4 *A* and *B* show apparent relationships between SST behavior in the ACR and PCR. SST biases in one tropical cyclogenesis region tend to be correlated with biases in the other region (Fig. 4*A*). There is an even stronger linear relationship (across models) between the amplitude of the high-frequency variability in the ACR and PCR (Fig. 4*B*). The apparent correlation of biases in geographically disparate regions may reflect common underlying causes, such as model errors in the large-scale mean state and in the amplitude of tropically coherent modes of variability.

Model performance in simulating variability on decadal and longer timescales is of most interest here, because this constitutes the background noise against which any slowly evolving forced signal must be detected (Fig. 4*C*). SST data were low-pass filtered to isolate variability on these timescales (see *Supporting Text*). In the

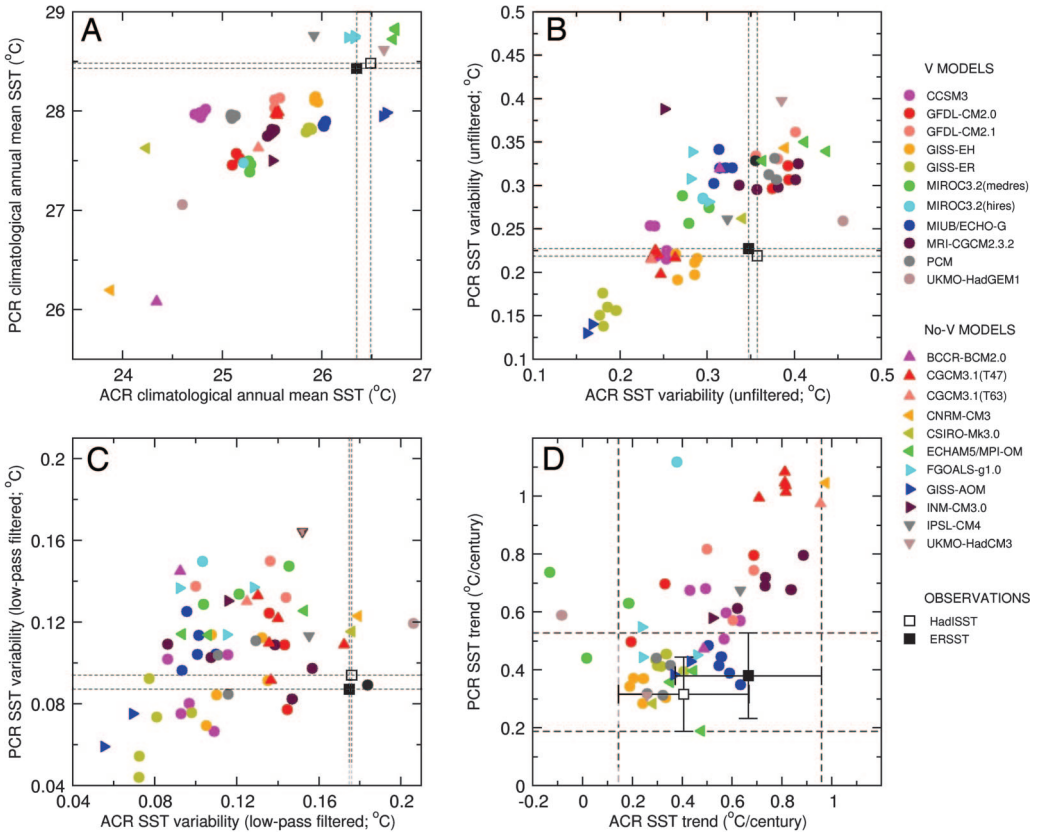

Fig. 4. Comparison of basic statistical properties of simulated and observed SSTs in the ACR and PCR. Results are for climatological annual means (*A*), temporal standard deviations of unfiltered (*B*) and filtered (*C*) anomaly data, and least-squares linear trends over 1900–1999 (*D*). For each statistic, ACR and PCR results are displayed in the form of scatter plots. Model results are individual 20CEN realizations and are partitioned into V and No-V models (colored circles and triangles, respectively). Observations are from ERSST and HadISST. All calculations involve monthly mean, spatially averaged anomaly data for the period January 1900 through December 1999. For anomaly definition and sources of data, refer to Fig. 1. The dashed horizontal and vertical lines in *A–C* are at the locations of the ERSST and HadISST values, and they facilitate visual comparison of the modeled and observed results. The black crosses centered on the observed trends in *D* are the 2σ trend confidence intervals, adjusted for temporal autocorrelation effects (see *Supporting Text*). The dashed lines in *D* denote the upper and lower limits of these confidence intervals.

ACR, the standard deviations of the filtered SST data are systematically lower in models than in observations, pointing to possible biases in model low-frequency variability.[P] Only 5 of the 22 models have 20CEN realizations with standard deviations close to or exceeding observed values. In the PCR, 21 of 22 models produce 20CEN realizations with greater than observed low-frequency SST variability. The implications of these results are discussed below.

Compared with Fig. 4 *A* and *B*, Fig. 4*C* displays much larger differences between the individual realizations of any given model's results. For example, the Parallel Climate Model (PCM) of the National Center for Atmospheric Research (27) has one 20CEN realization with low-frequency SST variability that is very similar to observed values (in both the ACR and PCR), whereas two other realizations have substantially lower ACR variability than either HadISST or ERSST. This difference illustrates that a large ensemble size (or long control run) is necessary to obtain reliable model estimates of low-frequency SST variability. It also suggests that it may be difficult to obtain a reliable observational estimate of internally generated low-frequency SST variability from the relatively short data records available.

These large differences between the temporal variance of individual realizations are also relevant to comparisons of modeled and observed trends (Fig. 4*D*). In the ACR and PCR, 20 and 13 (respectively) of the 22 models have at least one realization of the 20th century SST trend that lies within the statistical confidence intervals of the observed results. There is no evidence of a systematic model deficiency in simulating the magnitude of 20CEN SST trends in the ACR. In the PCR, nearly half of the simulated SST trends exceed the 2σ confidence interval for the observed trends.

Single-Forcing Experiments

Although our work points toward a pronounced influence of external forcing on SST changes in ACR and PCR, it does not separate and quantify the relative contributions of anthropogenic factors and natural external forcing. Separation is difficult without "single-forcing" experiments, in which key climate forcings are varied individually (rather than jointly, as in the 20CEN experiments).

Single-forcing experiments performed with PCM (27) indicate that increases in well mixed greenhouse gases are the main driver of century-timescale increases in ACR and PCR SSTs (Fig. 5). PCM's greenhouse-gas induced warming is partly offset by the cooling effects of anthropogenic sulfate aerosol particles, thus supporting observational findings in ref. 14, while solar, volcanic, and ozone forcing make much smaller contributions to the simulated SST changes over the 20th century.

Conclusions

Current model estimates of internal climate variability cannot explain observed SST increases in either the ACR or the PCR. This conclusion is insensitive to existing uncertainties in model physics and parameterizations, to observational uncertainty, and to the details of the procedure used to compare SST trends in observations and model control runs. It is also reasonably robust to the choice of time period used to estimate historical SST trends.

Our confidence in this conclusion would be undermined if models substantially underestimated the amplitude of natural internal climate variability. On decadal timescales, most current models underestimate SST variability in the ACR and overestimate variability in the PCR. It is possible that biases of similar magnitude may also apply on the multidecadal and century timescales consid-

[P]Missing or incorrectly specified forcings also influence the model-versus-observed variability differences shown in Fig. 4*C*. For example, the observed decadal variability in ACR and PCR SSTs receives a contribution from volcanic forcing (see Figs. 1 and 6), which is neglected in the No-V group of models. This missing forcing must contribute to the No-V models' underestimate of observed SST variability in the ACR.

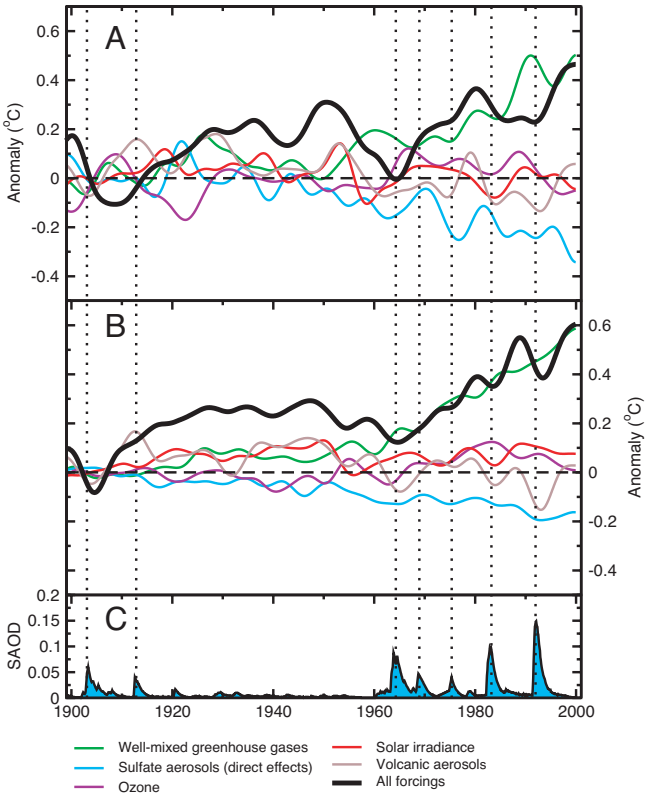

Fig. 5. Contribution of different external forcings to SST changes in tropical cyclogenesis regions. (*A* and *B*) Results are for the ACR (*A*) and PCR (*B*) and are from a 20CEN run and single-forcing experiments performed with the PCM (27). Each result is the low-pass filtered average of a four-member ensemble, with window width $W = 145$ months. For anomaly definition, refer to Fig. 1. Stratospheric aerosol optical depth (21) is also shown (*C*).

ered in Fig. 2. Even if they did, however, it is unlikely that climate noise could fully explain the large observed SST trend in the ACR over the last 100 years. This trend is at least 3–5 times larger (depending on the choice of observational data set) than s_{CTL}, the standard deviation of the model-based sampling distribution of unforced SST trends (see Table 1). Our estimates of s_{CTL} are conservative because they incorporate residual (and unphysical) climate drift. To achieve nonsignificant results (based on one-tailed tests and a 5% significance level) for the observed ACR trends over 1906–2005, the models used here would on average have to underestimate century-timescale SST variability in the ACR by a factor 2 (for the HadISST data) or a factor of >3 (for the ERSST data). Model average errors in decadal-timescale SST variability are of order 50%, not a factor of 2 or 3.[q]

In the PCR, the evidence against an internal variability explanation is even stronger. The model overestimate of the PCR low-frequency SST variability implies that the observed PCR trends (which are already highly significant over 1906–2005) are even less likely to be due to internal variability.

These results, together with other observational and modeling studies (7, 14, 28) contradict claims that internal climate noise accounts for all of the observed variability in tropical Atlantic SSTs (10). We find a large externally forced component of SST change in the ACR and PCR. On the basis of our F_1 results for the period 1906–2005, there is an 84% chance that external forcing explains at least 67% of observed SST increases in the ACR and PCR. In both regions, model simulations with external forcing by combined natural and anthropogenic effects are broadly consistent with observed SST increases. The PCM experiments suggest that forcing by well mixed greenhouse gases has been the dominant influence on century-timescale SST increases. We also find clear evidence of a volcanic influence on observed SST variability in the ACR and PCR.

Hurricanes are complex phenomena. Although changes in ocean surface temperatures may be a key influence on hurricane intensity (6, 8, 9), SSTs are only one of a variety of factors that control hurricane formation and evolution (1, 9, 29). Detailed analyses of changes in other large-scale conditions that affect tropical cyclogenesis (such as wind shear and vertical stability) are required to obtain a more complete understanding of how hurricane activity has changed and may continue to change in a warming world. Our research illustrates that models can be of considerable benefit in understanding the causes of such changes.

[q]The temporal standard deviation of the observed low-pass filtered ACR SST data, s_{filt}(OBS), is ≈0.18°C for both the HadISST and ERSST data (see Fig. 4*C*). Model-average values of this quantity, s_{filt}(MOD), are 0.12°C and 0.13°C for the V and No-V 20CEN runs.

We acknowledge the international modeling groups for providing their data for analysis, the Joint Scientific Committee/Climate Variability and Predictability Working Group on Coupled Modeling and their Coupled Model Intercomparison Project (CMIP) and Climate Simulation Panel for organizing the model data analysis activity, and the Intergovernmental Panel on Climate Change (IPCC) WG1 Technical Support Unit for technical support. We also thank Isaac Held and two anonymous reviewers for their constructive comments. The IPCC Data Archive at Lawrence Livermore National Laboratory is supported by the Office of Science, U.S. Department of Energy. HadISST data were provided by John Kennedy at the Hadley Centre for Climate Prediction and Research (Exeter, U.K.). Work at Lawrence Livermore National Laboratory was performed under the auspices of the U.S. Department of Energy, Environmental Sciences Division, Contract W-7405-ENG-48. A portion of this study was supported by the U.S. Department of Energy, Office of Biological and Environmental Research, as part of its Climate Change Prediction Program.

1. Gray WM (1968) *Mon Weather Rev* 96:669–700.
2. Emanuel KA (1987) *Nature* 326:483–485.
3. Holland GJ (1997) *J Atmos Sci* 54:2519–2541.
4. Raper SCB (1993) in *Climate and Sea Level Change: Observations, Projections and Implications*, eds. Warrick RA, Barrow EM, Wigley TML (Cambridge Univ Press, Cambridge, UK), pp 192–212.
5. Knutson TR, Tuleya RE (2004) *J Clim* 17:3477–3493.
6. Emanuel KA (2005) *Nature* 436:686–688.
7. Trenberth KE, Shea DJ (2006) *Geophys Res Lett* 33:L12704, 10.1029/2006GL026894.
8. Webster PJ, Holland GJ, Curry JA, Chang, H-R (2005) *Science* 309:1844–1846.
9. Hoyos CD, Agudelo PA, Webster PJ, Curry JA (2006) *Science* 312:94–97.
10. Goldenberg SB, Landsea CW, Mestas-Nuñez AM, Gray WM (2001) *Science* 293:474–479.
11. Barnett TP, Pierce DW, Schnur R (2001) *Science* 292:270–274.
12. Levitus S, Antonov JI, Wang J, Delworth TL, Dixon KW, Broccoli AJ (2001) *Science* 292:267–270.
13. Barnett TP, Pierce DW, AchutaRao KM, Gleckler PJ, Santer BD, Gregory JM, Washington WM (2005) *Science* 309:284–287.
14. Mann ME, Emanuel KA (2006) *EOS, Trans Am Geophys Union* 87:233, 238, 241.
15. Santer BD, Wigley TML, Mears C, Wentz FJ, Klein SA, Seidel DJ, Taylor KE, Thorne PW, Wehner MF, Gleckler, PJ, *et al.* (2005) *Science* 309:1551–1556.
16. Smith TM, Reynolds RW (2004) *J Clim* 17:2466–2477.
17. Rayner NA, Brohan P, Parker DE, Folland CK, Kennedy JJ, Vanicek M, Ansell T, Tett, SFB (2006) *J Clim* 19:446–469.
18. Thorne PW, Parker DE, Christy JR, Mears CA (2005) *Bull Am Met Soc* 86:1437–1442.
19. Lynch P, Huang, X-Y (1992) *Mon Weather Rev* 120:1019–1034.
20. Knight JR, Allan RJ, Folland CK, Vellinga M, Mann ME (2005) *Geophys Res Lett* 32:L20708, 10.1029/2005GL024233.
21. Sato M, Hansen JE, McCormick MP, Pollack JB (1993) *J Geophys Res* 98:22987–22994.
22. Gillett NP, Weaver AJ, Zwiers FW, Wehner MF (2004) *Geophys Res Lett* 31, 10.1029/2004GL020044.
23. Robock A (2000) *Rev Geophys* 38:191–219.
24. Gleckler PJ, Wigley TML, Santer BD, Gregory JM, AchutaRao KM, Taylor KE (2006) *Nature* 439:675.
25. Wigley TML (2000) *Geophys Res Lett* 27:4101–4104.
26. Santer BD, Taylor KE, Wigley TML, Johns TC, Jones PD, Karoly DJ, Mitchell JFB, Oort AH, Penner JE, Ramaswamy V, *et al.* (1996) *Nature* 382:39–46.
27. Washington WM, Weatherly JW, Meehl GA, Semtner AJ, Bettge TW, Craig AP, Strand WG, Arblaster J, Wayland VB, James R, *et al.* (2000) *Clim Dyn* 16:755–774.
28. Knutson TR, Delworth TL, Dixon KW, Held IM, Lu J, Ramaswamy V, Schwarzkopf MD, Stenchikov G, Stouffer RJ (2006) *J Clim* 10:1624–1651.
29. Bengtsson L, Botzet M, Esch M (1996) *Tellus* 48A:57–73.