

The Atlantic Multidecadal Oscillation Inferred from the Forced Climate Response in Coupled General Circulation Models

JEFF R. KNIGHT

Met Office Hadley Centre, Exeter, United Kingdom

(Manuscript received 13 May 2008, in final form 11 September 2008)

ABSTRACT

Instrumental sea surface temperature records in the North Atlantic Ocean are characterized by large multidecadal variability known as the Atlantic multidecadal oscillation (AMO). The lack of strong oscillatory forcing of the climate system at multidecadal time scales and the results of long unforced climate simulations have led to the widespread, although not ubiquitous, view that the AMO is an internal mode of climate variability. Here, a more objective examination of this hypothesis is performed using simulations with natural and anthropogenic forcings from the Coupled Model Intercomparison Project phase 3 (CMIP3) database. Ensemble means derived from these data allow an estimate of the response of models to forcings, as averaging leads to cancellation of the internal variability between ensemble members. In general, the means of individual model ensembles appear to be inconsistent with observed temperatures, although small ensemble sizes result in uncertainty in this conclusion. Combining the ensembles from different models creates a multimodel ensemble of sufficient size to allow for a good estimate of the forced response. This shows that the variability in observed North Atlantic temperatures possess a clearly distinct signature to the climate response expected from forcings. The reliability of this finding is confirmed by sampling those models with low decadal internal variability and by comparing simulated and observed trends. In contrast to the inconsistency with the ensemble mean, the observations are consistent with the spread of responses in the ensemble members, suggesting they can be accounted for by the combined effects of forcings and internal variability. In the most recent period, the results suggest that the North Atlantic is warming faster than expected, and that the AMO entered a positive phase in the 1990s. The differences found between observed and ensemble mean temperatures could arise through errors in the observational data, errors in the models' response to forcings or in the forcings themselves, or as a result of genuine internal variability. Each of these possibilities is discussed, and it is concluded that internal variability within the natural climate system is the most likely origin of the differences. Finally, the estimate of internal variability obtained using the model-derived ensemble mean is proposed as a new way of defining the AMO, which has important advantages over previous definitions.

1. Introduction

The Atlantic multidecadal oscillation (AMO) has become recognized in recent years as an important feature of the variability of observed climate. First identified in the instrumental climate record, the AMO is the alternation between periods of several decades with cool and warm sea surface temperature (SST) anomalies throughout the North Atlantic Ocean (Kerr 2000). Recognition of the AMO emerged from attempts to understand prominent land-based climate variability, in the form of significant multidecadal drying in the Afri-

can Sahel. Folland et al. (1986) linked Sahel drought to changes in an interhemispheric pattern of SST differences. Key to the timing of the recognition of significant multidecadal variability in the North Atlantic was the emergence of more accurate global datasets based upon carefully quality controlled and homogenized instrumental data (Folland et al. 1984). Subsequently, further analysis of these data clarified the character of North Atlantic variability (Kushnir 1994; Mann and Park 1994; Delworth and Mann 2000; Folland et al. 1999; Parker et al. 2007) and showed it had an approximately 65–70-yr period that was also manifest in global mean temperature (Schlesinger and Ramankutty 1994). The length of globally widespread instrumental records is less than 150 yr, or approximately two AMO periods; therefore, it is not possible to deduce from these data whether

Corresponding author address: Jeff R. Knight, Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB, United Kingdom.
E-mail: jeff.knight@metoffice.gov.uk

the observed AMO is indeed a long-lived oscillation, or simply a set of fluctuations confined to the observed period. Evidence from palaeoclimate data and climate models has, therefore, been advanced to assess this issue. Data from collections of multiple paleoproxies (Mann et al. 1995, 1998) and tree-ring data alone (Gray et al. 2004) indicate that the AMO was likely present for at least the last 400–500 yr. Furthermore, control simulations of some coupled ocean–atmosphere general circulation models, which have constant levels of external forcing agents, show significant multidecadal variability in the North Atlantic arising from internal mechanisms related to the oceanic meridional overturning circulation (Delworth et al. 1993; Timmermann et al. 1998; Delworth and Mann 2000; Latif et al. 2004; Jungclauss et al. 2005). The patterns, amplitudes, and periods of this variability differ considerably between models; however, a realistic simulation of observed AMO characteristics has been demonstrated with the third climate configuration of the Met Office Unified Model (HadCM3) (Knight et al. 2005).

The AMO has been shown to be relevant to the multidecadal climate variability of numerous regional climates (Knight et al. 2006; Baines and Folland 2007). As well as the aforementioned Sahel rainfall, in the tropics, the AMO has been shown to influence northeast Brazilian rainfall (Folland et al. 2001), the Indian (Goswami et al. 2006; Zhang and Delworth 2006) and East Asian (Lu et al. 2006) monsoons, and it may influence El Niño–Southern Oscillation (ENSO) (Dong et al. 2006). A link with the occurrence of major Atlantic hurricanes, via AMO-related changes in vertical wind shear in the main hurricane development region, has been suggested by observational (Goldenberg et al. 2001) and modeling (Knight et al. 2006; Zhang and Delworth 2006) evidence. This has coincided with interest in whether the increase in Atlantic hurricane activity in recent decades is attributable to natural variability or to anthropogenic climate change (Emanuel 2005). The AMO is also thought to affect midlatitude climate, including North American summer rainfall (Enfield et al. 2001; Sutton and Hodson 2005) and European climate (Sutton and Hodson 2005; Knight et al. 2006). In addition to its implications for regional climates, the AMO is also important for understanding observed large-scale climate, such as the Hadley and Walker circulations (Baines and Folland 2007), and the departure from a uniform warming trend in twentieth-century Northern Hemisphere mean temperatures (Zhang et al. 2007). Quantifying the role of the AMO in the instrumental record is, therefore, essential for making confident climate projections. This is even more so for projections of regional climate and for decadal time scales (Smith et al. 2007).

The variability of the North Atlantic SST record beyond interannual time scales (see Fig. 1) is usually interpreted as the superposition of a multidecadal oscillation and secular warming. The term “oscillation” in this context refers to the assumption that the variability has a characteristic frequency range and is part of a long series with stationary mean and variance. This is a more liberal definition than is used in physics and follows similar usage elsewhere in climatology (e.g., the North Atlantic Oscillation). The simple conceptual division between secular warming and an oscillation is motivated by the recognition of a dominant contribution to climate warming during the instrumental era from greenhouse gases (GHGs); a priori we expect the response of climate to a smooth, monotonic increase in GHG concentrations to be a similarly steady warming. The multidecadal fluctuations in the observational record, therefore, appear to have a different origin to the longer-term change, justifying this partitioning of the climate record. Various analytical methods have been used to perform this separation in the North Atlantic, with the aim of identifying the multidecadal component as the AMO. The simplest method is to remove a least squares linear fit to the North Atlantic mean SST time series (Enfield et al. 2001; Knight et al. 2005; Sutton and Hodson 2005; Zhang et al. 2007). This has the advantage of simplicity and is straightforward to communicate and reproduce, but it also has the important drawback that the rate of anthropogenic warming likely varied considerably through time. Nonlinearity potentially distorts the shape of the AMO time series in the most recent part of the record. In an attempt to improve on linear detrending, Trenberth and Shea (2006) define an AMO index as the difference between North Atlantic and global mean temperatures. Their aim is to isolate variability peculiar to the North Atlantic from the globally ubiquitous climate warming signal. Variability generated in the North Atlantic, however, is likely to have a signature in global mean temperature (Knight et al. 2005; Zhang et al. 2007), so subtracting global mean temperature from North Atlantic temperature is likely to lead to the partial cancellation of the AMO signal. Mann and Emanuel (2006) go further to include global mean temperature as a predictor of seasonal tropical North Atlantic temperatures in a multiple regression model, which also includes a climate model–derived aerosol term. The AMO component is identified as the residual of the model. Again, this approach risks fitting the temperature with a global mean term that already contains an AMO signal, falsely diminishing the size of the AMO residual. It is also not clear to what extent the aerosol effect is already present in the global mean. An alternative approach to the problem of separating the

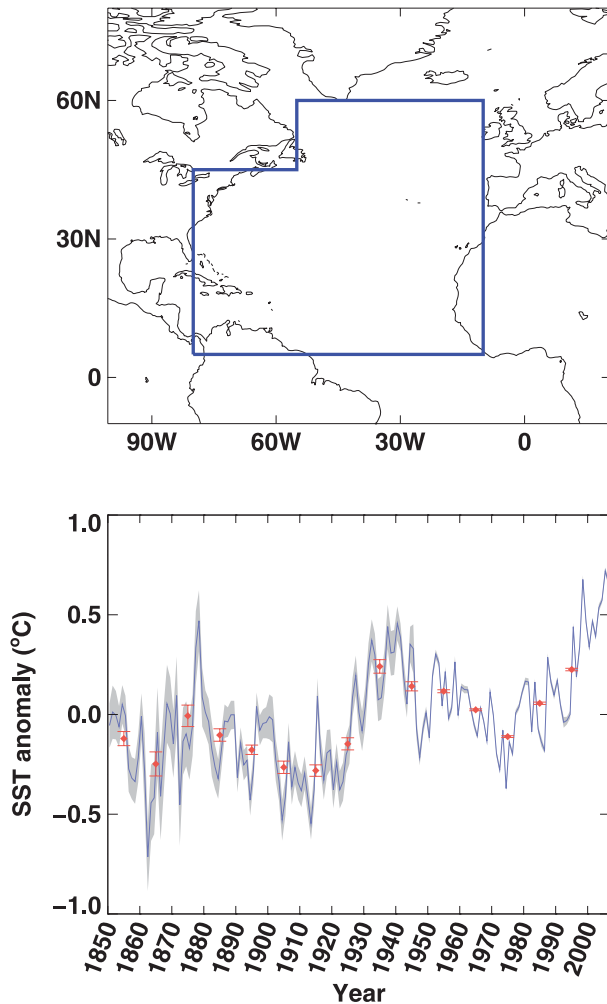


FIG. 1. Definition of the North Atlantic SST index with (top) area used and (bottom) time series of observed Atlantic SST averaged over the area in the top panel (blue) with 95% uncertainty limits defined by the gray shading. Decadal mean SST values for decades 1850–59, 1860–69, etc. are shown by red diamonds, with their 95% uncertainties represented by the red bars. The data are anomalized with respect to 1900–99.

AMO from the nonlinear secular component was taken by Parker et al. (2007), who identified the third principal component of an analysis of decadal filtered global SST data from 1891–2005 with the AMO. Their analysis also produces a first principal component (PC) time series that is almost identical to a similarly filtered series of global mean SST. This suggests that the sought-after separation is performed relatively efficiently by principal component analysis (PCA). Nevertheless, it is difficult to argue a priori that this should necessarily be the case, as PCs are statistical constructs and are constrained to be orthogonal to each other. As a result, there remains uncertainty about the detailed accuracy of the partitioning of variance between the global warming and AMO PCs.

All of the above estimates of the AMO are subject to uncertainty, as a result of not only the inherent uncertainty in the data but also through the arbitrariness of the chosen climate change component subtracted from North Atlantic SSTs. These surrogates—whether linear trend, global mean temperature, or first principal component—are employed on the grounds of plausibility and ease of computation from the observational record. Yet, there is no observational benchmark against which their suitability can be gauged. This is because the observational record cannot provide the response of North Atlantic SST to climate forcings needed to estimate the AMO directly. In this study, results from climate model simulations forced by estimates of past natural and anthropogenic forcing agents are used to provide this estimate. The success of climate models in simulating past global mean (Stott et al. 2000, 2006; Broccoli et al. 2003; Meehl et al. 2004; Hegerl et al. 2007) and regional (Knutson et al. 2006) temperatures suggests that they are sufficiently accurate for this purpose. The use of models does not completely eliminate uncertainty, as estimates of the simulated responses to twentieth-century forcings are potentially limited by model error and ensemble size. Nevertheless, it does remove the uncertainty in using a proxy measure of the forced signal, because this signal can be directly obtained from the models. It is also a completely independent source of physical information to the observed SST. Time-dependent variations in the strength of volcanic and anthropogenic aerosol cooling set against the background of increasing greenhouse-gas warming could have the appearance of a multidecadal oscillation. Simulations using all (natural and anthropogenic) forcings allow this component to be explicitly removed from the SST residual, which is, therefore, attributable to internal variability in the Atlantic climate system. The aim of this study is to establish to what extent the observed multidecadal variability known as the AMO is a result of external climate forcings or internal variability. To do this, results from the twentieth-century model ensembles from phase 3 of the Coupled Model Intercomparison Project (CMIP3) are used. These data provide a sufficiently large collection of realizations of twentieth-century climate to enable a useful estimate of the North Atlantic Ocean response to climate forcings. The use of multimodel data also helps to guard against obtaining model dependent results. Kravtsov and Spangale (2008) have shown that there are important differences between observed surface temperatures in the Atlantic region (including adjacent land areas) and those from a multimodel ensemble mean of surface temperatures of a set of these simulations. Here, a statistical comparison is performed to establish to what extent the model estimated forced

response is distinct from observations, focusing on North Atlantic SST as the most direct representation of AMO variability.

2. Data and methods

The simulation data for this study are from the World Climate Research Programme (WCRP) Working Group on Coupled Modeling (WGCM) CMIP3 multimodel dataset. CMIP3 data were obtained from the archive held by the Program for Climate Model Diagnosis and Intercomparison (PCMDI) at the Lawrence Livermore National Laboratory (LLNL) (available online at <https://esg.llnl.gov:8443/index.jsp>). Data were contributed by many climate modeling centers to this archive, which was a key primary resource drawn upon by the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) (Solomon et al. 2007). The analysis here primarily uses results for the “twentieth-century climate” (20C3M) simulations, which simulate the period from the mid-to-late-nineteenth century to the year 2000. For some of the models, a single 20C3M simulation is available; however, for others there are initial-value ensembles of various sizes. The forcing agents specified in the 20C3M experiments also vary between models, ranging from only the main anthropogenic factors (GHGs and tropospheric sulfate aerosol) to more comprehensive sets (including natural solar variability and volcanic aerosol amounts, tropospheric and stratospheric ozone changes, and land use change). For a reasonable comparison with observed historical temperatures, it is necessary to use those simulations that include the main natural and anthropogenic forcings (see, e.g., Stott et al. 2000). As a result, a subset of the CMIP3 20C3M data from simulations or ensembles using (as a minimum) solar, volcanic, GHG, and aerosol forcings are used here. The data for the Met Office Hadley Centre models [third climate configuration of the Met Office Unified Model (UKMO HadCM3) and Hadley Centre Global Environmental Model version 1 (UKMO HadGEM1)] from the CMIP3 database were supplemented by extra ensemble members available locally.

Monthly surface temperature (output variable name *ts*) fields were retrieved from the PCMDI CMIP3 archive, along with land–sea mask data. The *ts* data are a combination of marine SST and land surface temperature, so by masking for marine areas the SST can be retrieved. This is done in preference to using the SST fields in the archive (output variable name *tos*) which are available for fewer of the simulations. Masked monthly surface temperature data were aggregated into decadal means, and an area-weighted average over the domain in

Fig. 1 was calculated. This region covers most of the low- and midlatitude North Atlantic Ocean while avoiding areas where sea ice is likely to be present to preserve the equivalence of surface temperature and SST. Nevertheless, the possible influence of *ts* values over regions of sea ice on the inferred SST index was tested by repeating the analysis in the following sections on indices derived from the *tos* data directly. Significant results very similar to those derived using *ts* data are found for the *tos* indices, so only the *ts* results will be presented. The simulations’ North Atlantic SST indices are the basis for comparisons between the model data and historical observations of North Atlantic SST. An observational index for the area in Fig. 1 was derived for 1870–2000 as an area-weighted decadal average of data from the second Met Office Hadley Centre Sea Surface Temperature Dataset (HadSST2) (Rayner et al. 2006). HadSST2 is a gridded dataset that includes missing data points rather than using statistical methods (Rayner et al. 2003) to in-fill grid boxes with insufficient individual observations. It is provided with uncertainty estimates for various diagnostic quantities, such as global and hemispheric mean SSTs (available online at <http://www.hadobs.org>), and in principle these can be derived for any area mean. For this work, the 95% uncertainty limits from the combination of measurement, sampling, coverage, and bias uncertainties have been derived for the SST index used.

The basic comparisons of observed and simulated SST using the 20C3M data are extended to examine trends in the most recent period by using data from the CMIP3 IPCC Special Report on Emissions Scenarios (SRES) A1B projections for 2000–07 and further data from HadSST2. The use of projection data is valid to represent this period, because the scenario calculations are initialized from 20C3M calculations, and greenhouse gas and aerosol amounts have not markedly diverged from scenario assumptions. In addition, stratospheric volcanic aerosol has remained low throughout this period, and a large fraction of the 2000–07 warming is “committed warming” and so dependent only on climate forcing before the year 2000.

3. Results

a. Decadal means in individual model ensembles

Comparisons of simulated and observed North Atlantic area–mean SST time series for the range of models meeting the forcing and data availability criteria are shown in Fig. 2. For each model, the decadal responses in the ensemble members show a spread arising from the uncorrelated internal decadal variability. The

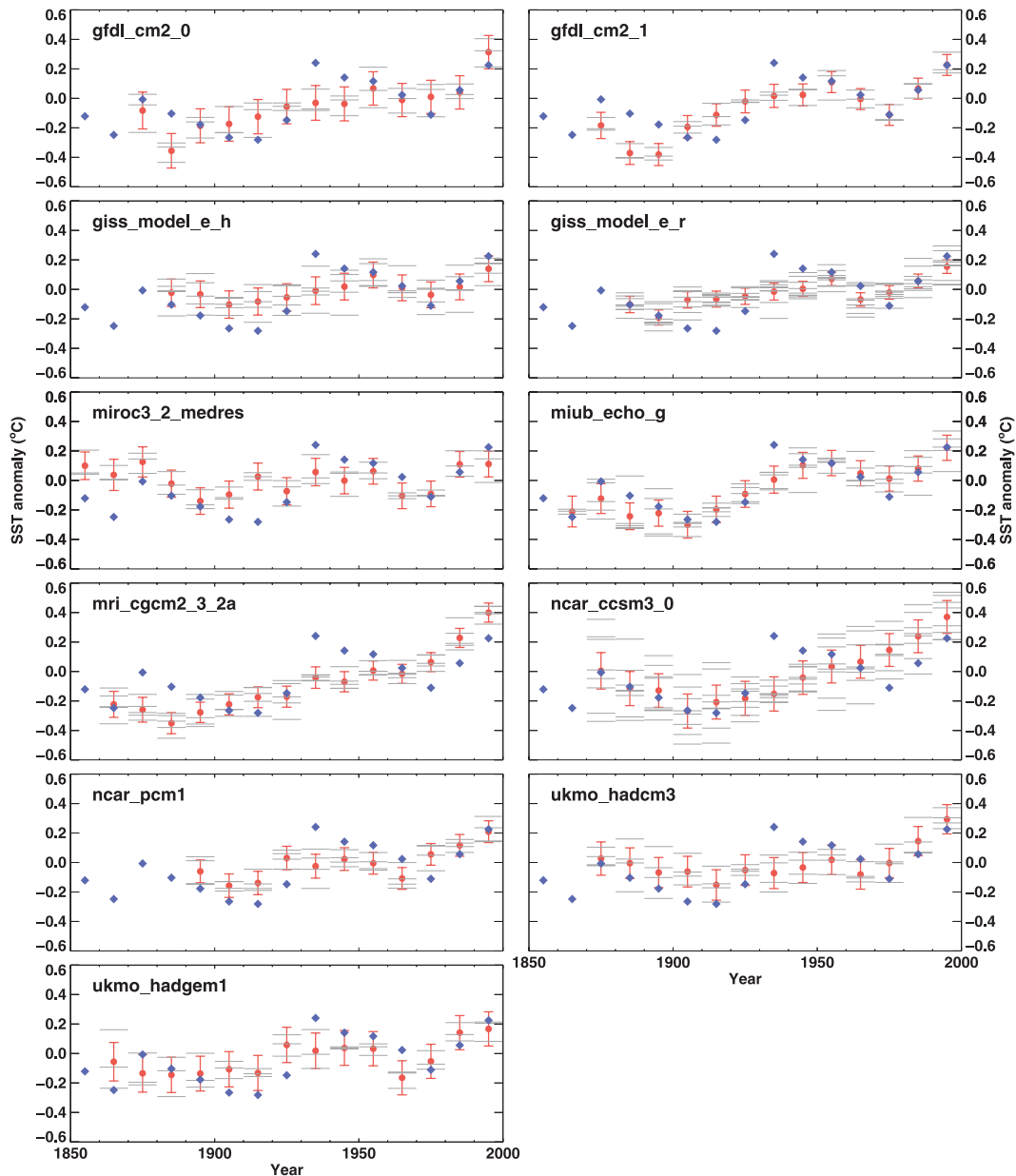


FIG. 2. Decadal mean North Atlantic SST index anomalies for a range of CMIP3 model ensembles. Values for individual ensemble members (gray bars) are derived as area averages over the domain shown in Fig. 1, and decadal anomalies for 1850–59, 1860–69, etc. are calculated with respect to the period 1900–99. Ensemble means are plotted as red circles. Decadal mean SST anomalies from HadSST2 (blue diamonds; see Fig. 1) are also calculated with respect to the period 1900–99. Estimates of the 95% uncertainty in the ensemble mean are produced with a standard deviation derived from the data in all decades of each ensemble. These are combined with the 95% uncertainty limits for the observed decadal mean SST index anomalies. The resulting overall uncertainties (red bars), therefore, allow for the interpretation that the observations (blue diamonds) are inconsistent with the ensemble mean (at the 95% level) if they lie outside the bars. The models are identified by the names by which they are referred to in the CMIP3 database (see URL in the text).

decadal ensemble mean is estimated as the average of the decadal temperature anomalies in each ensemble member. Because the temperature for each member contains a common signal from the forcings and a ran-

dom component as a result of internal variability, this averaging is likely to lead to the cancellation of the variability and a better estimate of the forced response than any single member alone. Nevertheless, without a

very large ensemble, the ensemble mean remains an uncertain estimate of this forced response. The CMIP3 ensembles analyzed here have only between 3 and 9 members each, so this uncertainty is not negligible. A simple approach to estimate the uncertainty is to apply a *t* test to each decade for each model; however, this method generates very broad 95% uncertainty limits (to 1°C or more), as the small sample size additionally makes the estimate of the standard deviation required for the test very uncertain. In this case, many of the models' ensemble means show statistical consistency with observations simply because their estimates are so poorly constrained (not shown). In addition, the size of the uncertainty in any given model can vary unrealistically from decade to decade based on the chance spread of a small number of members. A preferable approach is to obtain a more accurate estimate of the standard deviation of internal variability by pooling the deviations from the ensemble mean for all the decades together. This produces a smaller uncertainty that is uniform in each decade, although it is necessary to make the assumption that the magnitude of internal decadal variability does not change in response to the forcings. A priori there is no reason to assume that this sensitivity exists, and there is little evidence for it in observations or models (Hegerl et al. 2007). Even if the internal variance does change slightly, the error in the standard deviation this will introduce is still likely to be small compared to the uncertainty in the standard deviation estimated in an individual decade. For these reasons, this assumption is made in the calculation of uncertainties in Fig. 2. In addition to the uncertainty in the estimate of the ensemble mean, the observational uncertainty in the observed decadal mean SST is added in quadrature (Lanzante 2005). The resulting uncertainty range in Fig. 2 no longer reflects the range of possible values of the ensemble mean, but it shows the 95% limits of significant differences between the ensemble mean and observations. This allows consistency to be directly attributed wherever the observational curve lies within the plotted uncertainty envelope. In general, the observational uncertainty estimates are smaller than the uncertainties in the simulated ensemble mean, so they act to broaden the uncertainty range only slightly.

The comparison of mean North Atlantic SST between the Geophysical Fluid Dynamics Laboratory Climate Model version 2.0 (GFDL CM2.0, hereafter CM2.0) and observations is illustrative of many features seen in the comparisons using other models. The overall combined model and data uncertainty is over 0.2°C, most of which comes from uncertainty in the model mean as a result of the small ensemble size (three members). This

uncertainty is quite large compared to the time variation of the mean—only the means for the decades 1880–89 and 1990–99 lie further from the twentieth-century average (zero anomaly) than this. The relatively large size of the uncertainties suggests that the ability of the ensemble to provide a well-constrained estimate of the forced response in the model and thus a useful comparison with observations may be limited. Nevertheless, there are five decades for which the observed North Atlantic SST is statistically inconsistent with the ensemble mean despite the small ensemble size: 1880–89, 1910–19, 1930–39, 1940–49, and 1970–79. This is far in excess of the number of apparent inconsistencies expected by chance at the 95% confidence level used; if the observations were entirely consistent with the estimated ensemble mean, then for 13 independent decades there would be much less than a 1% chance that they would lie outside the uncertainty limits in five or more decades. Unsurprisingly, these five decades correspond to periods around the peaks and troughs of the observed AMO (the local minima and maxima of the decadal variability in Fig. 1), and are in the correct sense, so that decades corresponding to periods conventionally considered AMO negative are below the ensemble mean value, and vice versa.

The results from other models, such as the GFDL CM2.1 (GFDL CM2.1, hereafter CM2.1), show similar features but also some differences. For example, while the decade 1880–89 is, again, significantly warmer than represented by the model ensemble, this is also true of the adjacent decades 1870–79 and 1890–99. In addition, 1910–19 is joined by 1920–29 as significantly cooler than in the ensemble. On the other hand, 1970–79 is consistent with the CM2.1 ensemble where it is not for CM2.0. The ensemble spread in the CM2.1 model seems to be less than in CM2.0, leading to narrower uncertainty ranges on the ensemble mean. This clearly favors the identification of more discrepancies between the model and the observations, of which there are seven in CM2.1. The possibility that so many apparently significant differences arise by chance is even more improbable than for the CM2.0 case. Careful inspection of all the models in Fig. 2 shows that the decadal spread of the modeled temperature varies considerably between models [compare, e.g., National Center for Atmospheric Research (NCAR) Community Climate System Model, version 3 (CCSM3; eight members) with Goddard Institute for Space Studies Model E-R (GISS-ER; nine members)]. All the models show warming over the twentieth century, as would be expected from the imposition of increases in atmospheric greenhouse gases. Some show a more secular increase (e.g., Meteorological Research Institute Coupled GCM, version 2.3.2a (MRI

CGCM2.3.2a) than others, which possess a period of cooling (e.g., UKMO HadGEM1). This suggests that there is some uncertainty in the simulation of North Atlantic climate over the past 100–150 yr. Nevertheless, each model simulation has four or more decades when the observed decadal temperature is statistically inconsistent with the decadal ensemble mean, and the decades that are most frequently outside the limits of the ensemble mean estimates are those at the extreme phases of the AMO. Indeed, the 1930–39 decade is significantly warmer than the means of all the model ensembles, and 1910–19 is cooler in 9 out of 11 cases. This suggests that none of the models is able to reproduce the decadal variability of observed North Atlantic SST solely as a response to the imposed forcing. That said, the aforementioned width of the uncertainty ranges as a result of the small ensemble sizes available may be acting to prevent a more clear-cut identification of the differences. For example, the observed decadal mean for 1970–79 is lower than the means of every individual model ensemble, but it is only significantly so in 7 of these 11 cases. To further clarify whether the AMO is part of the forced response in models, it is necessary to obtain larger ensemble sizes. This is done here by amalgamating the various models' ensembles into a multimodel ensemble, and this is the subject of the next section.

b. Multimodel ensemble

A multimodel ensemble is created from the decadal anomalies of all members of all model ensembles without selection or weighting (Fig. 2). The decadal observations are compared with the statistics of this multimodel ensemble in Fig. 3. The multimodel ensemble contains 52 members after 1890, so it is much larger than the largest (nine members) individual model ensemble and even more so than the typical ensemble size (three to four members). The uncertainty in the multimodel ensemble mean is not as small as would be expected for a 52-member ensemble of simulations with a single model, as there is also a spread in the models' ensemble means, but it is still considerably smaller than for any of the individual models. The uncertainty range shown in Fig. 3 is, again, a combination of model and observational uncertainties, with the model uncertainty being the larger component. For the later decades of the twentieth century, the combined uncertainty range is somewhat less than 0.1°C , spreading to about 0.1°C in the early twentieth century. It is wider in the nineteenth century, when fewer ensemble members contribute to the multimodel ensemble. The first decade with data for a large fraction of the 52 members is 1870–79, which has 37, so 1850–59 and 1860–69 are disregarded here.

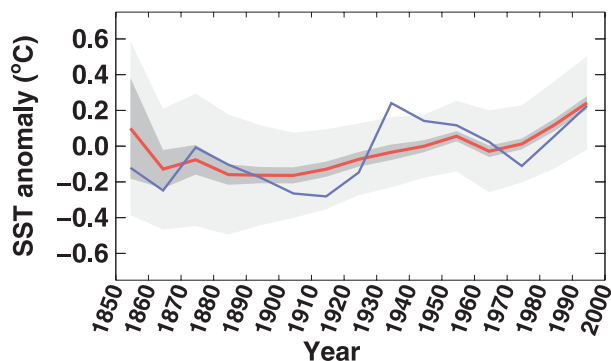


FIG. 3. Comparison of decadal multimodel ensemble mean (red curve) with the observed decadal North Atlantic SST index (blue curve). As in Fig. 2 but the 95% uncertainty range of the estimated multimodel decadal mean is combined with observational uncertainty to produce an overall consistency range (dark gray shading). In addition, an estimate of the 95% range of responses in the multimodel ensemble is combined with observational uncertainty (light gray shading).

Overall, the smaller size of the uncertainty in the multimodel ensemble mean suggests that a clearer separation of the estimated forced response from observations is possible.

The multimodel ensemble mean appears to dip slightly in the decade 1880–89, which could be associated with volcanic forcing in the models linked with the eruption of Krakatau in 1883. It subsequently remains stable before beginning to rise in the early part of the twentieth century. Rising global mean temperature during this period has been attributed to increases in solar irradiance and greenhouse gases (Tett et al. 2002). It then falls temporarily in the decade 1960–69, which could be a response to anthropogenic aerosol increases or cooling from the Mount Agung eruption of 1963. The multimodel ensemble mean then rises more rapidly in the last three decades of the twentieth century, as greenhouse gases increase further. The shape of the curve is very similar to estimates of total forced change in global mean temperature (Hegerl et al. 2007). In contrast, the instrumentally derived decadal North Atlantic temperature is dominated by the peaks and troughs of the AMO superimposed on a general warming trend. The uncertainties show that the observations and the multimodel ensemble mean are inconsistent at the 95% level in 9 of the 13 decades considered. These are 1900–09, 1910–19, 1920–29 (all cooler in observations), 1930–39, 1940–49, 1950–59, 1960–69 (all warmer), 1970–79, and 1980–89 (both cooler). In fact, the only decades that show statistical consistency are before 1900 and the decade 1990–99. Of course, unless the multimodel ensemble mean is a perfect reproduction of observations, then the frequency of inconsistency should increase as

the uncertainty limits decrease. The differences here, however, are not random decadal noise, as they are coherent between decades and are a considerable fraction (several tenths of a degree Celsius) of the forced variability. This implies that the observed average North Atlantic SST is not well reproduced by the multimodel mean. Equivalent to this, it appears the AMO is inconsistent with the models' estimate of the response to forcings.

Also plotted in Fig. 3 is an estimate of the 95% range of North Atlantic SSTs in the multimodel ensemble members (as opposed to the range of the multimodel ensemble mean discussed above). For any given member, this is a function of both the forced response and the particular realization of internal variability that this member contains. The observations are consistent with this range of possible area-mean SSTs in every decade except 1930–39, which is the furthest from the multimodel ensemble mean. The inconsistency in this decade does not imply inconsistency overall, however, as there is about a 50% chance that one or more of the 13 decades considered will lie outside the 95% limits. The agreement between the observations and the model range suggests that although the observed variability does not appear as a forced response in the models, its magnitude is consistent with their internal variability. This is despite a lack of realistic AMO behavior in most of the models.

c. Sensitivity to model selection

The internal variability of each model used in the multimodel ensemble can be estimated using the distribution of differences between decadal SSTs in the ensemble members and the ensemble means in every decade. Table 1 shows the decadal standard deviations for each model included in Fig. 2 plus an estimated standard deviation of the internal variability in observations. This is derived from the set of decadal differences between the observed SST and the multimodel ensemble mean, which is taken to be an estimate of the external (forced) component of the observed response. For all but one model, the internal variability is less than in observations, often considerably so. To formalize the extent of this shortfall, an F test was performed on the ratio of each model standard deviation with that from observations, taking into account the number of realizations of the variability used in formulating each. It is found that for consistency with the estimated observed standard deviation (0.13°C) at the 95% confidence level, a model requires a standard deviation of about 0.07°C (the actual level varies slightly with the number of ensemble members and decades in each calculation). Five of the 11 models are found to have internal variability in

TABLE 1. Standard deviation of decadal SST anomalies from the ensemble mean for each model included in the analysis and for HadSST observations. Boldface type indicates that the model is consistent with estimated observed values (see text).

Model name	Ensemble size	Standard deviation (°C)
GFDL CM2.0	3	0.079
GFDL CM2.1	3	0.049
GISS-EH	5	0.087
GISS-ER	9	0.065
MIROC3.2(medres) ^a	3	0.061
MIUB ECHO-G ^b	5	0.085
MRI CGCM2.3.2a	5	0.065
NCAR CCSM3	8	0.148
NCAR PCM1 ^c	4	0.063
UKMO HadCM3	4	0.086
UKMO HadGEM1	3	0.081
HadSST	—	0.128

^a Model for Interdisciplinary Research on Climate 3.2, medium-resolution version.

^b Meteorological Institute of the University of Bonn ECHAM and the global Hamburg Ocean Primitive Equation.

^c NCAR Parallel Climate Model.

decadal North Atlantic SST that is significantly lower than in observations. The models that are consistent (highlighted in bold in Table 1) also tend to be less variable (except NCAR CCSM3), but the uncertainty in estimating the observed internal standard deviation from just 15 decades means that it is not possible to discount chance as the source of any one model's shortfall.

The widespread lack of decadal internal variability in the models suggests that the multimodel uncertainty ranges in Fig. 3 may be too narrow, as the uncertainty in the multimodel ensemble mean is a function of the spread in the model ensembles. This could undermine, to some extent, the finding that the observations are inconsistent with the multimodel ensemble mean. To test this, a separate multimodel ensemble is constructed from the six models that have statistically consistent estimates of internal variability with that derived from observations. The resulting multimodel ensemble comprises about half the number of models and about half the number of members (28) as the original ensemble. The shape of the multimodel ensemble mean (Fig. 4) is essentially unchanged by the selection of models, but its uncertainty range is typically about 50% larger. Both factors of the typical model spread being larger in the new ensemble and the smaller overall ensemble size could contribute to this increase. A measure of the effect of the change in spread is obtained from examining the uncertainty in the mean of an ensemble of the five least variable models, which is of similar overall size (24 members). This shows uncertainty ranges that are about the same size as the full ensemble, suggesting the effect

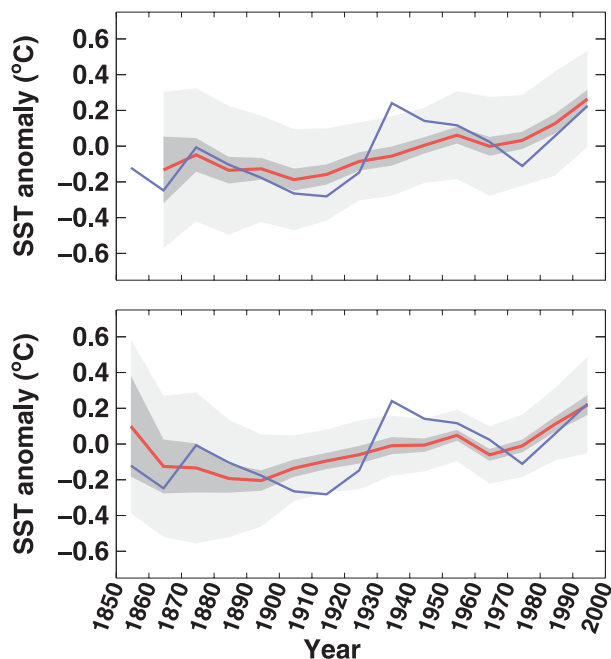


FIG. 4. Sensitivity of the statistical comparison of observations with the multimodel ensemble to the internal variability of models in the multimodel ensemble. As in Fig. 3 but for (top) models with variability consistent with observations (models are listed in bold in Table 1) separated from (bottom) those with inconsistently low variance at the 95% confidence limit.

of reducing ensemble size is about the same as the effect of selecting the models with lower variability. The broader uncertainties in the higher variability multimodel ensemble inevitably make for more stringent statistical comparisons than the full ensemble. Despite this, there is still inconsistency between the multimodel ensemble mean and the observations in all the decades found in Fig. 3 except for 1960–69. The finding that the AMO is inconsistent with the forced response in models holds, as the extreme decades of the AMO are still statistically distinct from the multimodel ensemble mean. The range of SST values represented by the models (including model forced responses and internal variability) is slightly broader in the ensemble of higher variability models (although still not broad enough to include the decade 1930–39). In contrast, the lower variability ensemble has a narrower 95% uncertainty range and additionally excludes the decade 1910–20. Although not highly significant, this suggests that these models probably do not have sufficient internal variability to reproduce the observed AMO.

d. Analysis of trends

A minor drawback of the type of analyses presented above arises from the potential differences between the

shapes of the responses to imposed forcings estimated by the various model ensembles. Although the ensemble means all show similar features (Fig. 2), they are not identical, so any difference in the total amount of twentieth-century warming would lead to opposite offsets at the beginning and end of the anomalized series. In turn, the spread of the multimodel ensemble could be inflated as a result of this uncertainty in the forced response. This may mask the contribution of decadal-scale variability (of more interest here than the centennial change) to the multimodel mean. To overcome this, decadal trends in the models are compared with those from observations, as these trends are a clearer reflection of the forcings within the trend period than the temperature anomalies themselves. Best-fit linear trends for a range of interval lengths from 10 yr to half the length of the data series were computed using annual North Atlantic mean SST values in each multimodel ensemble member and these were averaged to produce the ensemble mean trend. The use of a range of trend periods ensures that the conclusions drawn from the analysis are robust to the choice of a particular interval length. Since trend values tend to decrease with increasing interval length, linear temperature changes (i.e., trend multiplied by length) are plotted in Fig. 5. For most of the simulated period, the changes are positive, which is consistent with the general warming seen in Figs. 2 and 3. This warming is concentrated in two periods, with positive changes of several tenths of a degree Celsius across a wide range of decadal time scales from the late nineteenth to the mid-twentieth century and in the late twentieth century. Periods of cooling or stable temperatures are found in the mid-to-late-nineteenth century and centered about 1960, but the amount of cooling is generally weaker than the warming seen in other periods. These features correspond to those identified in global mean temperature (Trenberth et al. 2007), and the global-scale warming has been attributed to increases in solar forcing and greenhouse gases and cooling as a result of volcanic forcing and sulfate aerosols (Hegerl et al. 2007). In contrast to the periods with warming, the magnitude of North Atlantic cooling tends to decrease rapidly at longer trend intervals. This suggests an influence of short-lived phenomena, such as the volcanic eruptions of Krakatau in 1883 and Agung in 1963.

The observed temperature changes are also plotted in Fig. 5. Again, there are periods of warming and cooling, but in each case the magnitude of changes is greater than for the multimodel ensemble mean. Warming of more than 0.6°C centered about 1925 is found that is approximately invariant across a range of time scales. At first sight, this is reminiscent of the first phase of

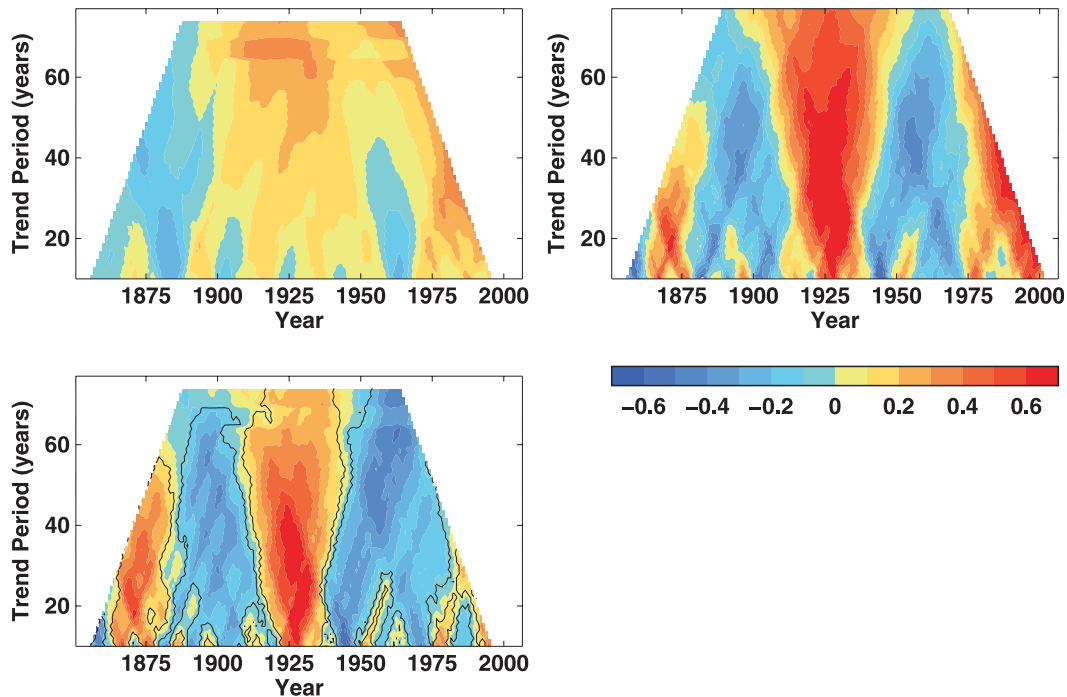


FIG. 5. Comparison of ensemble mean trends in the multimodel ensemble and observations. Trends multiplied by trend period ($^{\circ}\text{C}$) are plotted as a function of the midpoint year of the trend interval and the length of trend interval for (upper left) the multimodel ensemble mean, (upper right) HadSST observations, and (lower left) the difference. Multimodel ensemble mean changes were derived as the average of the changes computed in all the individual members. The mean value for each interval length and central year is computed separately, averaging all simulations for which the length of the data series allows the computation of a trend. Here, 95% confidence limits on the difference are shown as the solid contours.

warming in the multimodel ensemble, but it is dissimilar in being more than twice as large and occurring over a much shorter period; 30-yr changes, for example, are positive between about 1913 and 1940, whereas they are positive between about 1890 and 1952 for the models' ensemble mean. Similar length periods of cooling bracket the warming, centered on approximately 1895 and 1955. Again, these features are consistent over a wide range of choices of interval length. The magnitude of the changes in these phases is similar to each other but about half that of the warming phase. Nevertheless, this is still about double the cooling changes seen in the multimodel mean. The latter cooling period, at least, does not seem to coincide with the 1963 volcanic eruption as well as in the multimodel ensemble, as the observed cooling is already peaking by this time. There is additionally a warming period in about 1870 and, like the multimodel trends, a manifestation of warming in recent decades. The peak of this recent change is more than 0.6°C , which appears larger than that for the mean of the models. This is not a fair comparison, however, as the observational data are plotted up to 2007, whereas the model data stop in 1999.

The difference between the observed and multimodel changes (computed where data overlap) bears a strong resemblance to the observed temperature changes, with very similar phases of warming and cooling and similar magnitudes. A small reduction in the peak warming rate is found on account of the models being able to produce a small amount of mid-twentieth-century warming. Overall, the differences are so stark that they are significant at the 95% confidence level for almost all times and intervals for which changes were computed. The analysis essentially confirms the findings from the multimodel ensemble of decadal means that the models are inconsistent with the historical North Atlantic SSTs. This is also borne out by comparisons with the individual model ensembles (not shown), none of which produce the magnitude or the characteristics of the observed trends.

e. Recent warming

As mentioned above, the difference in changes at the very end of the comparison period (up to 1999) in Fig. 5 suggests that simulated warming is slower than observed. For example, the 30-yr change for 1970–99 is about 0.3°C in the multimodel ensemble mean, but it is

about 0.5°C in observations. This anomaly occurs only at the very end of the series, however, so it is not clear whether this is a robust feature. More up-to-date data from models are required to examine the simulation of the recent past. Although the CMIP3 20C3M set of past simulations only extend to 2000, it is possible to get a model estimate of forced warming in the period since 2000 by supplementing the 20C3M data with data from climate projections using forcings scenarios. Several factors contribute to the feasibility of this approach. First, a number of the CMIP3 scenario simulations are performed as continuations of specific 20C3M simulations from the year 2000, so it is possible to produce a continuous ensemble that spans the pre- and post-2000 periods. Second, much of the temperature change in the 2000–07 period examined is likely to be “committed warming,” determined by increases in greenhouse gases in the 20C3M period. Third, the anthropogenic forcings projected by the scenarios have been quite similar to those observed in the recent period (Rahmstorf et al. 2007), and there has been no volcanic eruption with a significant climatic effect. SRES A1B scenario data (denoted as sresa1b in the CMIP3 archive) are used on the basis that more data are available for this scenario than others, and that the forcings diverge only slightly by 2007. A multimodel ensemble of combined model realizations of North Atlantic SST up to 2007 is produced for each member of the model ensembles in Fig. 2 for which a sresa1b continuation is available. This results in a 25-member ensemble, although the representation of models in this ensemble is more uneven than in the larger 20C3M ensemble. For example, the models GISS-ER and NCAR CCSM3 account for 12 of the 25 members.

A comparison of the ensemble mean of this multimodel ensemble with observed decadal mean SSTs is shown in Fig. 6. During the 20C3M period (before 2000), the ensemble mean has similar values and shape to that of the full ensemble in Fig. 3. The ensemble mean corresponding to the sresa1b period is plotted as the average over eight years (2000–07) rather than a decade, as are the equivalent data from observations. This shows, as would be expected from simulations of recent global trends (Hegerl et al. 2007), accelerating warming between the decade 1990–99 and the partial decade 2000–07. There is confirmation of the tendency seen in Fig. 2 for observed SSTs to be lower than the multimodel ensemble mean in the decades 1970–79 and 1980–89 and to be similar in the decade 1990–99. Despite the rapid warming in the simulations, the observed 2000–07 mean is slightly higher than the ensemble average. This difference lies on the 95% uncertainty limit, suggesting that the best estimate of the forced response

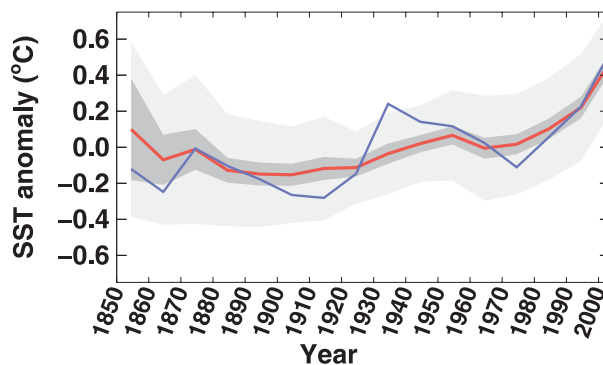


FIG. 6. As in Fig. 3 but for a multimodel ensemble of combined CMIP3 20C3M and sresa1b simulations. The last point corresponds to the mean SST anomaly over the period 2000–07 rather than a whole decade. The uncertainty limits on the comparison account for this.

provided by the models is unlikely to explain the full extent of recent North Atlantic warmth. The full range of responses in individual ensemble members in the last eight years is consistent with observations, however, implying that observed SSTs could result from a combination of forcing and internal variability.

Although this analysis leads to the conclusion that the most recent observed SST appears to be too high to be explained by forcings alone, there are implicit uncertainties that cannot be ignored as safely as in the twentieth-century-only case. As already mentioned, the relatively small size of the ensemble allows some models to have a more dominant influence on the multimodel ensemble mean. In particular, the NCAR CCSM3 model is the model with the most members (seven) and shows the highest ensemble mean warming relative to the twentieth-century mean. This may act to positively bias the ensemble mean, and so artificially diminish the amount of additional warmth that would otherwise be attributed to the AMO. Correcting the overpopulation of the ensemble by particular models by, for example, reducing the number of members it contributes, just acts to reduce overall ensemble size and therefore increases uncertainty. Systematic differences between the amount of warming in the multimodel ensemble members will be most pronounced at the 2000–07 endpoint, as anomalies are defined over the 1900–99 period. These difficulties once again suggest that comparison of trends averaged across the ensemble may be more reliable than anomalies. Figure 7 shows linear changes in the ensemble of extended realizations and their differences with observed changes. Simulated changes are similar to the previous analysis (Fig. 5) except for the increased emphasis on strong warming in recent times. While there is a hint that observed changes in the last two to three

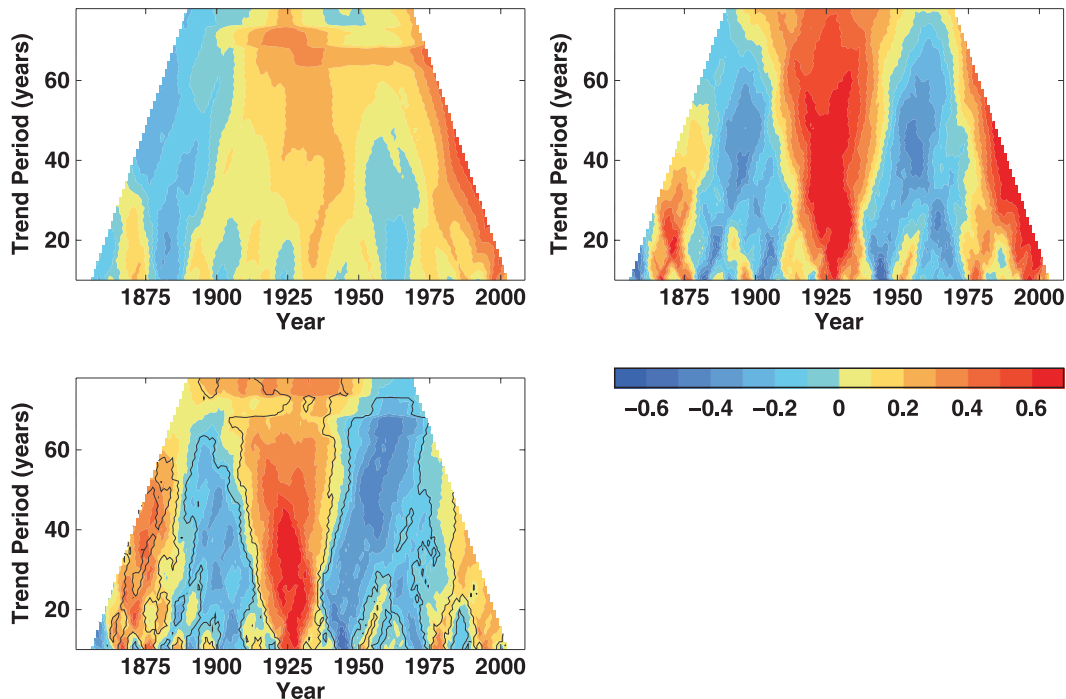


FIG. 7. As in Fig. 5, but for the extended period up to 2007, as provided by the composited 20C3M and sresa1b simulations.

decades are greater than simulated changes in Fig. 5, this becomes clearer when the scenario simulation data are used to extend the analysis. Observed changes up to the present day, measured over all periods up to 40 yr (1968–2007), are statistically significantly higher at the 95% confidence level. This includes changes over periods shorter than 20 yr, suggesting that enhanced warming is going on right up to the present time. Having said this, the magnitude of the changes (up to about 0.3°C) does not rival those seen in earlier warming phases, such as that in the 1920s (with values of more than 0.6°C), and is less than recent changes in the models (about 0.4°C). The implication is that recent Atlantic warming is only partly internal variability, and that the scale of this variability is more modest than at times in the past. Nevertheless, the results show that in common with most of the twentieth century, there are likely to have been changes in North Atlantic SST in recent decades that are unrelated to forcings.

4. Discussion and conclusions

The results presented here show that for area-mean North Atlantic SST, the ensemble mean of a large multimodel ensemble of climate simulations is inconsistent with observations for most of the twentieth century. Interpreting the multimodel mean as an estimate of the

response of climate to natural and anthropogenic forcings suggests the observed SST is inconsistent with the forced response in climate models. The differences in the shape of the simulated and observed time series are large, with the models showing changes similar to those in global mean temperature, while the observations exhibit large multidecadal fluctuations. In addition, the differences are robust to the lack of decadal internal variability in some of the models that compose the multimodel ensemble. The results are confirmed by an analysis of the multimodel ensemble of trends over a range of multidecadal periods, which avoids anomalization from an arbitrary climatological period. There are four possibilities that may individually or jointly explain these findings: (i) there are errors in the observed North Atlantic SST record, (ii) there are errors in the models' responses to forcings, (iii) there are errors in the forcing data used to drive the models, and/or (iv) the differences are the manifestation of internal variability in North Atlantic climate.

The 95% limits of observational uncertainty, comprising estimates of measurement, sampling, coverage, and bias uncertainties, are provided as part of the HadSST2 dataset (Rayner et al. 2006) and have been included in this analysis. In addition, the dataset is corrected for known historical biases. Historically, the North Atlantic has been relatively well observed, so

data uncertainty tends to be small compared to that from the spread of model realizations. In principle, however, there could be unaccounted time-varying biases that narrow the differences between models and observations. Nevertheless, the chance of this substantially bridging the gap seems very remote, considering the range of independent data [e.g., night marine air temperature (Rayner et al. 2003) and island and coastal station records] that corroborate the marked multidecadal fluctuations seen in the SST records.

Models can respond incorrectly to forcings as a result of, for example, incorrect feedbacks. A familiar example is the range of sensitivities of different models to future levels of CO₂. In the twentieth century, the magnitude of forcings is much smaller, but there are presumably still some differences in response between the models. The effectiveness of models to simulate global mean temperature in the twentieth century, however, argues that this is not greatly limiting. At the scale of the North Atlantic Ocean, there may be larger differences in responses, however, as differences in regional responses could be “averaged out” in the global mean. Nevertheless, within current understanding the North Atlantic is not anticipated to respond fundamentally differently to past forcings than other ocean basins, so the observed SST history is unexpected. For example, there is no reason to suggest that highly nonlinear responses to the forcings, such as the forced excitation of an internal mode of North Atlantic variability, are absent in the models. Rather, at multidecadal time scales, the forcings appear most likely to produce responses that are essentially linear everywhere (Hegerl et al. 2007).

In terms of errors in forcings, although the history of carbon dioxide and other well-mixed greenhouse gas concentrations is relatively well known, changes in other forcing agents are less so—examples include volcanic aerosols, solar variability, and tropospheric aerosol concentrations (Forster et al. 2007). Two periods of explosive volcanic activity occurred between 1850 and 2000: 1883–1912 and 1963–91. The intervening period was relatively quiescent. Coupled with the tendency of volcanic stratospheric aerosols to cool climate, this suggests the potential for a sequence of cool and warm periods like that in observations. Although represented in the models, if the forcing from these eruptions were higher than is specified, the additional cooling may explain the difference with observations. A problem with this idea is that the first anomalously cool period in observations occurs between about 1900 and 1930, which is too long after the first volcanic period to be consistent with the time scale of volcanic effects of a few years (see, e.g., Jones et al. 2003). In the second period, the largest eruption is that of Mount Pinatubo in 1991, but the

decade 1990–99 shows good consistency between the ensemble and observed SSTs. In addition to problems of timing, volcanic aerosol rapidly spreads in the zonal direction in the stratosphere. This makes it unclear why the simulated North Atlantic should have a deficient response to this forcing when there is no apparent deficiency in the global mean response (Stott et al. 2006).

Total solar irradiance (TSI) may well have increased in the early part of the twentieth century (Forster et al. 2007), and so it is potentially a candidate for explaining the rise in North Atlantic SST at this time. Reconstructions also show, however, that TSI has likely remained stable in the latter part of the twentieth century, so it cannot explain the ongoing variability. In addition, suggested TSI changes appear to be a weak forcing globally (Forster et al. 2007), although Meehl et al. (2003) suggest the possibility of regional subtropical feedbacks in response to combined solar and greenhouse forcing. It is unclear, however, how much of an effect this mechanism might have throughout the North Atlantic. Uncertainties in tropospheric aerosols are a potentially better explanation, as anthropogenic sulfate aerosols have sources close to the North Atlantic (in Europe and North America) and are known to cool climate. In addition, sulphur dioxide emissions in this part of the world are known to have increased before about 1965 and declined thereafter, possibly contributing to the phases of cooling and subsequent warming observed in the North Atlantic at about this time (Fig. 1). The models included in this study represent, as a minimum, the direct effect of sulfate aerosols, so to a first approximation this cooling should be accounted for in the multimodel ensemble. Aerosol forcing estimates are uncertain, however, so the tendency of the models not to reproduce the coolness of the 1970s and 1980s could arise if the aerosol forcing is too weak. Against that, the models tend to reproduce the global cooling during this period quite well, and the ensembles produced with models having more complex aerosol treatments do not show systematically different responses. As for TSI changes, aerosol trends cannot explain the changes throughout the record; the rapid warming of the North Atlantic in the early twentieth century occurred during the early stages of growth of aerosol emissions. It remains possible that with multiple changes to the specification of forcings the observations could be reproduced, but the probability of the necessary differences between the true forcings and those used in existing models rapidly diminishes with the number of factors considered.

With the likelihood that the errors in the data, models, and forcings cannot account for observed North

Atlantic SSTs, the only remaining possibility is that the differences represent genuine internal variability at multidecadal time scales. This paradigm has been widely used as a hypothesis to explain historical North Atlantic SST and has given rise to the AMO terminology. The results presented here support the existence of an internally generated AMO and strengthen the findings of Knight et al. (2005), who demonstrate that an AMO with a similar pattern, amplitude, and period to that observed can be produced in a climate model simulation without external forcings.

Estimation of the component of regional SST change associated with the forced response using the multimodel ensemble is an advance on previous estimates using, for example, global mean temperature, as the models are a completely independent source of physical information. The method presented here removes the unquantifiable uncertainty associated with the arbitrariness of such methods. In its place is uncertainty over the accuracy of simulated responses, but this is at least amenable to investigation by, for example, making multimodel ensembles comprising different sets of models. At the very least, an index of the AMO derived by subtracting the CMIP3 multimodel ensemble mean can be described as the best estimate available at this time; such an index is shown in Fig. 8. For the period 1850–1999 this is derived from the full ensemble used in Fig. 3 to make use of the greatest amount of data. A value for 2000–07 is added using the smaller ensemble of extended simulations. The AMO index clearly shows the familiar phases of twentieth-century variability and also suggests part of an AMO cycle in the nineteenth century, although this has lower confidence in part as a result of larger ensemble spread at this time.

The AMO produced by this new definition retains many of the properties of series computed by linear detrending (e.g., Knight et al. 2005). It has the same sequence of phases with very similar zero-crossing times. The overall amplitude is reduced as a result of the increased structure in the model-derived forced SST but only marginally. The coolness of the periods 1900–25 and 1965–95 is not quite as intense, but the warm period of 1925–65 has higher peak anomalies, although these are not as sustained. An alternative approach to linear detrending is the subtraction of global mean temperature (Trenberth and Shea 2006). This gives an AMO sequence that appears to be less variable than the AMO in Fig. 8, which is consistent with the concern that the AMO is partly cancelled by an AMO-related component in the global mean. Further, Trenberth and Shea (2006) claim that their analysis shows that the apparent positive AMO in the late nineteenth century found by linear detrending

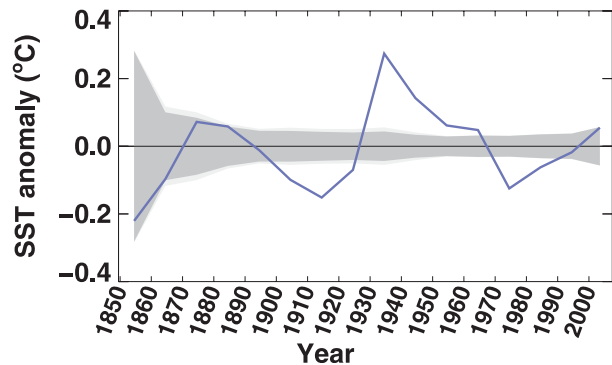


FIG. 8. A new definition of the AMO. The blue curve shows the AMO estimated as the difference between the area-average North Atlantic SST in observations and a multimodel ensemble mean. For the period up to 1999 inclusive, the differences are derived from the decadal analysis in Fig. 3. For 2000–07, the 8-yr mean difference from the extended analysis in Fig. 6 is used. The dark gray shading shows the 95% uncertainty limits of a significance test of the difference accounting for ensemble spread. The light gray shading shows the limits of a similar test for the combined effects of ensemble spread and observational SST uncertainty.

is an artifact of the detrending methodology. Using a different definition of the AMO, this analysis hints that this feature may be real, although confidence in it is less than for twentieth-century AMO phases. Mann and Emanuel (2006) use scaled global mean temperature and model-derived aerosol terms as a statistical model to fit observed seasonal SST in the tropical North Atlantic. This analysis is reminiscent of attribution analyses using global mean temperature in place of greenhouse gas increases. The result is an AMO estimate with very little multidecadal amplitude. Overfitting is clearly a risk with such a model when there are so few effective degrees of freedom in the low-frequency SST time series. Moreover, the results found here, which account for all forcings in a self-consistent way, demonstrate that there is a considerable residual AMO after forcings are accounted for. A much closer approximation to the AMO in Fig. 8 is that derived by Parker et al. (2007) as the third principal component of low-frequency global SSTs. This has a very similar shape after about 1925. Before this there is a similar cool phase, but its duration is shorter than in the version of the AMO presented here. Their AMO series is given in standardized units; however, inspection of the associated projection onto global SSTs implies an AMO amplitude that is quite similar to that found here. Although the principal component analysis is not guaranteed a priori to isolate the AMO, the comparison with the new AMO shows that it does so fairly well. Kravtsov and Spannagle (2008) estimate the AMO as the leading principal component of the difference between global surface temperatures in observations and the mean of a

multimodel ensemble of CMIP3 runs with a somewhat different composition to the one used here. Their index is similar to that in Fig. 8 before about 1950, but thereafter it has systematically lower values. Most likely the inclusion of simulations without natural forcings in this ensemble overemphasizes anthropogenic warming in the latter period, which in fact had considerable volcanic cooling.

In the present day, the AMO defined using the multimodel ensemble forcings estimate is positive, but it has a small anomaly (about $+0.06^{\circ}\text{C}$) that is on the limit of significance at the 95% level; this smaller value of the AMO in the present day is different from that obtained from the AMO estimated using linear detrending, which does not account for the enhanced anthropogenic warming rate in recent decades. Trenberth and Shea (2006) obtain a similar result by subtracting global mean temperature despite having different AMO values earlier in the series. Parker et al. (2007) also show a marginally positive AMO in the early twenty-first century. These results give a reasonably firm indication that the AMO became positive in the 1990s. The relatively small size of the present-day anomaly compared to values in the mid-twentieth century may partly explain why Sahel rainfall, which has been linked to the AMO (Folland et al. 1986; Zhang and Delworth 2006), has not recovered to the levels measured in this period. The AMO derived here shows a slower rate of increase in recent times to that seen in earlier AMO warming episodes. This might suggest that the AMO is now behaving differently than in the past. On the other hand, the difference is consistent with variability in AMO characteristics in successive cycles seen in very long model simulations of the AMO (Knight et al. 2005). It is not clear whether the recent increase in the AMO will culminate in a modest maximum, or whether the AMO will continue to grow to the scale of the mid-twentieth-century positive event. As the AMO has been linked to variations in the strength of the meridional overturning circulation in the Atlantic Ocean (Knight et al. 2005), decadal forecasts based on detailed initialization of the ocean (Smith et al. 2007) may give a perspective on this. It is not certain, however, that these forecasts are yet sufficiently skillful in predicting the oceanic changes that are likely to determine the detailed evolution of the future AMO. In any case, an examination of past intervals between AMO transitions (Enfield and Cid-Serrano 2006) suggests that the current positive phase of the AMO will last for one to three decades into the future. This implies the occurrence of a wide range of climatic anomalies (see, e.g., Knight et al. 2006; Zhang and Delworth 2006) in addition to those expected from anthropogenic climate change.

Acknowledgments. The author would like to thank A. Scaife and C. Folland for their useful discussions in the development of this work, which was supported by the joint DECC and MoD Integrated Climate Programme—Contracts GA01101 (DECC) and CBC2B0417 (MoD), Annex C5. The modeling groups the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP's Working Group on Coupled Modelling (WGCM) are acknowledged for their roles in making available the WCRP CMIP3 multimodel dataset. Support of this dataset is provided by the Office of Science of the U.S. Department of Energy.

REFERENCES

- Baines, P. G., and C. K. Folland, 2007: Evidence for a rapid global climate shift across the late 1960s. *J. Climate*, **20**, 2721–2744.
- Broccoli, A. J., K. W. Dixon, T. L. Delworth, T. R. Knutson, R. J. Stouffer, and F. Zeng, 2003: Twentieth-century temperature and precipitation trends in ensemble climate simulations including natural and anthropogenic forcing. *J. Geophys. Res.*, **108**, 4798, doi:10.1029/2003JD003812.
- Delworth, T. L., and M. E. Mann, 2000: Observed and simulated multi-decadal variability in the Northern Hemisphere. *Climate Dyn.*, **16**, 661–676.
- , S. Manabe, and R. J. Stouffer, 1993: Interdecadal variations of the thermohaline circulation in a coupled ocean–atmosphere model. *J. Climate*, **6**, 1993–2011.
- Dong, B., R. T. Sutton, and A. A. Scaife, 2006: Multidecadal modulation of El Niño–Southern Oscillation (ENSO) variance by Atlantic Ocean sea surface temperatures. *Geophys. Res. Lett.*, **33**, L08705, doi:10.1029/2006GL025766.
- Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436**, 686–688.
- Enfield, D. B., and L. Cid-Serrano, 2006: Projecting the risk of future climate shifts. *Int. J. Climatol.*, **26**, 885–895.
- , A. M. Mestas-Núñez, and P. J. Trimble, 2001: The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. *Geophys. Res. Lett.*, **28**, 2077–2080.
- Folland, C. K., D. E. Parker, and F. Kates, 1984: Worldwide marine temperature fluctuations 1856–1981. *Nature*, **310**, 670–673.
- , —, and T. N. Palmer, 1986: Sahel rainfall and worldwide sea temperatures, 1901–85. *Nature*, **320**, 602–607.
- , —, A. W. Colman, and R. Washington, 1999: Large-scale modes of ocean surface temperature since the late nineteenth century. *Beyond El Niño: Decadal and Interdecadal Climate Variability*, A. Navarra, Ed., Springer-Verlag, 73–102.
- , A. W. Colman, D. P. Rowell, and M. K. Davey, 2001: Predictability of northeast Brazil rainfall and real-time forecast skill, 1987–98. *J. Climate*, **14**, 1937–1958.
- Forster, P., and Coauthors, 2007: Changes in atmospheric constituents and in radiative forcing. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 129–234.
- Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Núñez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**, 474–479.
- Goswami, B. N., M. S. Madhusoodanan, C. P. Neema, and D. Sengupta, 2006: A physical mechanism for North Atlantic SST influence on the Indian summer monsoon. *Geophys. Res. Lett.*, **33**, L02706, doi:10.1029/2005GL024803.

- Gray, S. T., L. J. Graumlich, J. L. Betancourt, and G. T. Pederson, 2004: A tree-ring-based reconstruction of the Atlantic multidecadal oscillation since 1567 A.D. *Geophys. Res. Lett.*, **31**, L12205, doi:10.1029/2004GL019932.
- Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 663–745.
- Jones, P. D., A. Moberg, T. J. Osborn, and K. R. Briffa, 2003: Surface climate responses to explosive volcanic eruptions seen in long European temperature records and mid-to-high latitude tree-ring density around the Northern Hemisphere. *Volcanism and the Earth's Atmosphere, Geophys. Monogr.*, Vol. 139, Amer. Geophys. Union, 239–254.
- Jungclaus, J. H., H. Haak, M. Latif, and U. Mikolajewicz, 2005: Arctic–North Atlantic interactions and multidecadal variability of the meridional overturning circulation. *J. Climate*, **18**, 4013–4031.
- Kerr, R. A., 2000: A North Atlantic climate pacemaker for the centuries. *Science*, **288**, 1984–1985.
- Knight, J. R., R. J. Allan, C. K. Folland, M. Vellinga, and M. E. Mann, 2005: A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophys. Res. Lett.*, **32**, L20708, doi:10.1029/2005GL024233.
- , C. K. Folland, and A. A. Scaife, 2006: Climate impacts of the Atlantic Multidecadal Oscillation. *Geophys. Res. Lett.*, **33**, L17706, doi:10.1029/2006GL026242.
- Knutson, T. R., and Coauthors, 2006: Assessment of twentieth-century regional surface temperature trends using the GFDL CM2 coupled models. *J. Climate*, **19**, 1624–1651.
- Kravtsov, S., and C. Spanngale, 2008: Multidecadal climate variability in observed and modeled surface temperatures. *J. Climate*, **21**, 1104–1121.
- Kushnir, Y., 1994: Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Climate*, **7**, 141–157.
- Lanzante, J. R., 2005: A cautionary note on the use of error bars. *J. Climate*, **18**, 3699–3703.
- Latif, M., and Coauthors, 2004: Reconstructing, monitoring, and predicting multidecadal-scale changes in the North Atlantic thermohaline circulation with sea surface temperature. *J. Climate*, **17**, 1605–1614.
- Lu, R., B. Dong, and H. Ding, 2006: Impact of the Atlantic Multidecadal Oscillation on the Asian summer monsoon. *Geophys. Res. Lett.*, **33**, L24701, doi:10.1029/2006GL027655.
- Mann, M. E., and J. Park, 1994: Global-scale modes of surface temperature variability on interannual to century time scales. *J. Geophys. Res.*, **99**, 25 819–25 833.
- , and K. A. Emanuel, 2006: Atlantic hurricane trends linked to climate change. *Eos, Trans. Amer. Geophys. Union*, **87**, 233–244.
- , J. Park, and R. S. Bradley, 1995: Global interdecadal and century-scale climate oscillations during the past five centuries. *Nature*, **378**, 266–270.
- , R. S. Bradley, and M. K. Hughes, 1998: Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, **392**, 779–787.
- Meehl, G. A., W. M. Washington, T. M. L. Wigley, J. M. Arblaster, and A. Dai, 2003: Solar and greenhouse gas forcing and climate response in the twentieth century. *J. Climate*, **16**, 426–444.
- , —, C. M. Ammann, J. M. Arblaster, T. M. L. Wigley, and C. Tebaldi, 2004: Combinations of natural and anthropogenic forcings in twentieth-century climate. *J. Climate*, **17**, 3721–3727.
- Parker, D., C. Folland, A. Scaife, J. Knight, A. Colman, P. Baines, and B. Dong, 2007: Decadal to multidecadal variability and the climate change background. *J. Geophys. Res.*, **112**, D18115, doi:10.1029/2007JD008411.
- Rahmstorf, S., A. Cazenave, J. A. Church, J. E. Hansen, R. F. Keeling, D. E. Parker, and R. C. J. Somerville, 2007: Recent climate observations compared to projections. *Science*, **316**, 709.
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.
- , P. Brohan, D. E. Parker, C. K. Folland, J. J. Kennedy, M. Vanicek, T. Ansell, and S. F. B. Tett, 2006: Improved analyses of changes and uncertainties in marine temperature measured in situ since the mid-nineteenth century: The HadSST2 dataset. *J. Climate*, **19**, 446–469.
- Schlesinger, M. E., and N. Ramankutty, 1994: An oscillation in the global climate system of period 65–70 years. *Nature*, **367**, 723–726.
- Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, and J. M. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global climate model. *Science*, **317**, 796–799.
- Solomon, S., D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, H. L. Miller Jr., and Z. Chen, Eds., 2007: *Climate Change 2007: The Physical Science Basis*. Cambridge University Press, 996 pp.
- Stott, P. A., S. F. B. Tett, G. S. Jones, M. R. Allen, J. F. B. Mitchell, and G. J. Jenkins, 2000: External control of twentieth century temperature by natural and anthropogenic causes. *Science*, **290**, 2133–2137.
- , G. S. Jones, P. Thorne, J. Lowe, C. Durnam, and T. Johns, 2006: Transient climate simulations with the HadGEM1 climate model: Causes of past warming and future climate change commitment. *J. Climate*, **19**, 2763–2782.
- Sutton, R. T., and D. L. Hodson, 2005: Atlantic Ocean forcing of North American and European summer climate. *Science*, **309**, 115–118.
- Tett, S. F. B., and Coauthors, 2002: Estimation of natural and anthropogenic contributions to twentieth century temperature change. *J. Geophys. Res.*, **107**, 4306, doi:10.1029/2000JD000028.
- Timmermann, A., M. Latif, R. Voss, and A. Grötzner, 1998: Northern Hemispheric interdecadal variability: A coupled air–sea mode. *J. Climate*, **11**, 1906–1931.
- Trenberth, K. E., and D. J. Shea, 2006: Atlantic hurricanes and natural variability in 2005. *Geophys. Res. Lett.*, **33**, L12704, doi:10.1029/2006GL026894.
- , and Coauthors, 2007: Observations: Surface and atmospheric climate change. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 235–336.
- Zhang, R., and T. L. Delworth, 2006: Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophys. Res. Lett.*, **33**, L17712, doi:10.1029/2006GL026267.
- , —, and I. M. Held, 2007: Can the Atlantic Ocean drive the observed multidecadal variability in Northern Hemisphere mean temperature? *Geophys. Res. Lett.*, **34**, L02709, doi:10.1029/2006GL028683.