DECADAL PREDICTION

Can It Be Skillful?

BY GERALD A. MEEHL, LISA GODDARD, JAMES MURPHY, RONALD J. STOUFFER, GEORGE BOER, GOKHAN DANABASOGLU, KEITH DIXON, MARCO A. GIORGETTA, ARTHUR M. GREENE, ED HAWKINS, GABRIELE HEGERL, DAVID KAROLY, NOEL KEENLYSIDE, MASAHIDE KIMOTO, BEN KIRTMAN, ANTONIO NAVARRA, ROGER PULWARTY, DOUG SMITH, DETLEF STAMMER, AND TIMOTHY STOCKDALE

A new field called "decadal prediction" will use initialized climate models to produce time-evolving predictions of regional climate that will bridge ENSO forecasting and future climate change projections.

NEED FOR DECADAL PREDICTIONS. Prolonged drought in the American Southwest, increased hurricane activity in the tropical Atlantic since the late 1990s, changing fisheries regimes, extreme events, like the 2003 European heat wave, and the need to adapt to time-evolving climate change and increasing temperatures have raised concern among policy and decision makers about climate change in the near term, that is, out to 10–30 yr, referred to as the "decadal" time scale. Impacts resulting from these conditions have significant social, economic, and environmental implications and are consistent with the climate simulations of the twentieth-century and projections of the twenty-first-century climate of some models (Seager et al. 2007; Knutson and Tuleya 2004; Meehl et al. 2007). Some aspects of observed changes have been attributed to naturally occurring decadal variability (Goldenberg et al. 2001; McCabe et al. 2004; Zhang and Delworth 2006; Meehl et al. 2009a). Anthropogenically forced climate change, intrinsic climate variability, and natural external forcings (e.g., major volcanic eruptions or possibly the solar cycle) act together to produce the time-evolving climate. Given no future information on the third, the first two must thus be addressed to provide the best information on climate shifts over the coming several decades.

AFFILIATIONS: MEEHL AND DANABASOGLU—NCAR,* Boulder, Colorado; GODDARD AND GREENE—IRI, New York, New York; MURPHY AND SMITH—Hadley Centre, Exeter, United Kingdom; STOUFFER AND DIXON—GFDL, Princeton, New Jersey; BOER—Environment Canada, Gatineau, Quebec, Canada; GIORGETTA—MPI, Saarbrüken, Germany; HAWKINS—NCAS-Climate, University of Reading, Reading, United Kingdom; HEGERL—University of Edinburgh, Edinburgh, United Kingdom; KAROLY—University of Melbourne, Melbourne, Victoria, Australia; KEENLYSIDE—IFM-GEOMAR, Kiel, Germany; KIMOTO—University of Tokyo, Tokyo, Japan; KIRTMAN—University of Miami, Miami, Florida; NAVARRA—INGV, Rome, Italy; PULWARTY—NOAA, Boulder, Colorado; STAMMER—University of Hamburg, Hamburg, Germany; STOCKDALE—ECMWF, Reading, United Kingdom

*The National Center for Atmospheric Research is sponsored by the National Science Foundation.

CORRESPONDING AUTHOR: Gerald A. Meehl, 1850 Table Mesa Dr., NCAR, Boulder, CO 80305
E-mail: meehl@ncar.ucar.edu

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/2009BAMS2778.1

In final form 27 March 2009

The prospect of decadal prediction and its recognized importance has led, in part, to the initiation, in several countries, of climate services intended to bridge the gap between the seasonal-to-interannual (SI) climate information provided by the National Meteorological and Hydrological Services and the broad-scale, longer-duration horizon information considered by the Intergovernmental Panel on Climate Change (IPCC) assessments. In the United States, the National Oceanic and Atmospheric Administration (NOAA), in partnership with other agencies, is discussing formation of a National Climate Service that would, among other things, serve the near-term climate change information needs of the Regional Integrated Sciences and Assessments (RISAs; see www.climate.noaa.gov/cpo_pa/risa), the Regional Climate Centers (RCCs; see www.ncdc.noaa.gov/oa/climate/regionalclimatecenters.html), and the newly established National Integrated Drought Information System (NIDIS; see www.drought.gov). In Germany, the Climate Service Centre (CSC), funded by the German ministry for education and research [Bundesministerium für Bildung and Forschung (BMBF)] for an initial period of about 5 yr, will start (likely early in 2009) a program for climate prediction over the next few decades (information online at www.clisap.de). In the United Kingdom, the U.K. Climate Impacts Programme (UKCP09; see http://ukcp09.defra.gov.uk/), established in 1997, provides climate model projections of twenty-first-century climate for use in national assessments of climate impacts and adaptation strategies. UKCP has published new probabilistic scenarios based on ensembles of climate model projections for a series of 30-yr periods covering from 2010–39 to 2070–99. In Italy, the Euromediterranean Center for Climate Change (www.cmcc.it) has been established with the mission to develop Earth system models for climate scenarios. It is focusing on the near-term period (2010–40) with high-resolution global models, using an approach that includes realistic initial conditions and emission scenarios. In September of 2009, the Third World Climate Conference established a Global Framework for Climate Services to initiate international cooperation in the provision of climate change information to stakeholders.

In addition, partnerships developed through boundary organizations such as the RISAs, the International Research Institute for Climate and Society (IRI), and others will need to be engaged to test and evaluate the benefits and limits of decadal-scale knowledge in appropriate decision-making environments.

The ability to provide meaningful decadal predictions using dynamical models has yet to be firmly established, but pioneering efforts at initialized coupled ocean–atmosphere 10-yr predictions have begun (Smith et al. 2007; Keenlyside et al. 2008; Pohlmann et al. 2009). Later in this paper we describe an initiative by the climate science community [as part of the Coupled Model Intercomparison Project phase 5 (CMIP5)] to carefully examine the ability of dynamical models to simulate and predict decadal variability, to test the benefits and limitations of different initialization schemes, and ultimately to quantify the potential contributions of decadal climate outlooks over and above the projections typically considered in previous IPCC assessments, which have focused mostly on the forced response (i.e., the response of the climate system to external forcings, such as anthropogenic greenhouse gases). In addition, decadal prediction will likely involve higher-resolution climate models for better simulation of both regional climate and climate extremes because the coupled models available today are barely capable of representing regional events that require a higher resolution, such as precipitation extremes (e.g., Kimoto et al. 2005) or tropical cyclones (Gualdi et al. 2008). Model initialization could potentially yield higher skill just by assimilating persistent anomalies (e.g., in upper-ocean heat content) even if they lack the ability to accurately simulate internal variability. Initialization may also enable more realistic simulation of the slow oceanic changes associated with decadal variability. Results from CMIP5 would be relevant to more coordinated or "seamless" future climate predictions where a number of time scales could be predicted, including the decadal, using different versions of the same model (e.g., Shukla et al. 2009; Hurrell et al. 2009; Shapiro et al. 2009, manuscript submitted to *Bull. Amer. Meteor. Soc.*).

BACKGROUND. The current practice for providing climate change information over the next several decades is to look at those time periods in ensemble averages of forced climate change simulations using various future emission scenarios that typically are run to 2100 (Fig. 1). Using this technique, it can be seen that some regional climate change information on decadal time scales already can be obtained mainly from two sources: 1) climate change commitment (e.g., Wetherald et al. 2001; Meehl et al. 2005; Wigley 2005), and 2) the forcing from increasing greenhouse gases (e.g., Lee et al. 2006; Stott and Kettleborough 2002). Climate change commitment arises because at any point in time the slower-warming oceans are lagging behind the land areas. Thus, the oceans provide thermal inertia for the climate system. The time scale of this lag for the upper ocean is decades, and for the deep ocean it is 1,000 yr or more. This implies that even if greenhouse gas concentrations were stabilized today, the climate system would continue to warm at a rate of about 0.1°C decade^{-1} for the next several decades for a total of about 0.6°C after 100 yr (Meehl et al. 2007). There also would be additional climate change due to further anticipated increases in greenhouse gases.

The pattern and magnitude of surface temperature change is similar for three different emission scenarios for the period of 2011–30 (Fig. 1) because the climate system response is comparable over the next few decades no matter which scenario is followed (Meehl et al. 2007). Only in the second half of the twenty-first century does the climate response depend significantly on which emission pathway is followed. Therefore, barring a large volcanic eruption that could cool the system for a few years, a decadal prediction system already has some potential built-in skill simply from climate change commitment and forcing (Lee et al. 2006). However, useful predictions of the forced response would still depend on the signal-to-noise ratio of the quantity being predicted, and the model predictions themselves would likely contain significant uncertainties, even for the next 10–30 yr (e.g., Hawkins and Sutton 2009a).

Climate change projections from coupled models do exhibit decadal variability; however, they have been started from randomly selected preindustrial states, so the inherent variability in projections is not synchronized with observations. Furthermore, such variability is typically averaged out using multimember ensembles from individual models, or multimodel ensembles, so that only an estimate of the forced response remains (Meehl et al. 2007). Climate variability results in a range of possible outcomes spanned by the models about the mean forced response. For the next few decades, however, the actual time-evolving climate is of interest. Presumably, better climate change information could be obtained if the models could track the time evolution of the inherent decadal variability in combination with the forced response. Even an ability to capture no more than the decay of existing "anomalies" toward the mean forced response would be an improvement over traditional climate projections, where the change over the next 20–30 yr relative to the recent past typically takes no account of whether that recent past has been "above" or "below" what might have been expected. Initializing climate predictions, testing performance over recent decades, and understanding the present state of the climate are all linked. Initialized predictions should better quantify the uncertainty range in the near future by taking into account internal variability and the mean forced response.

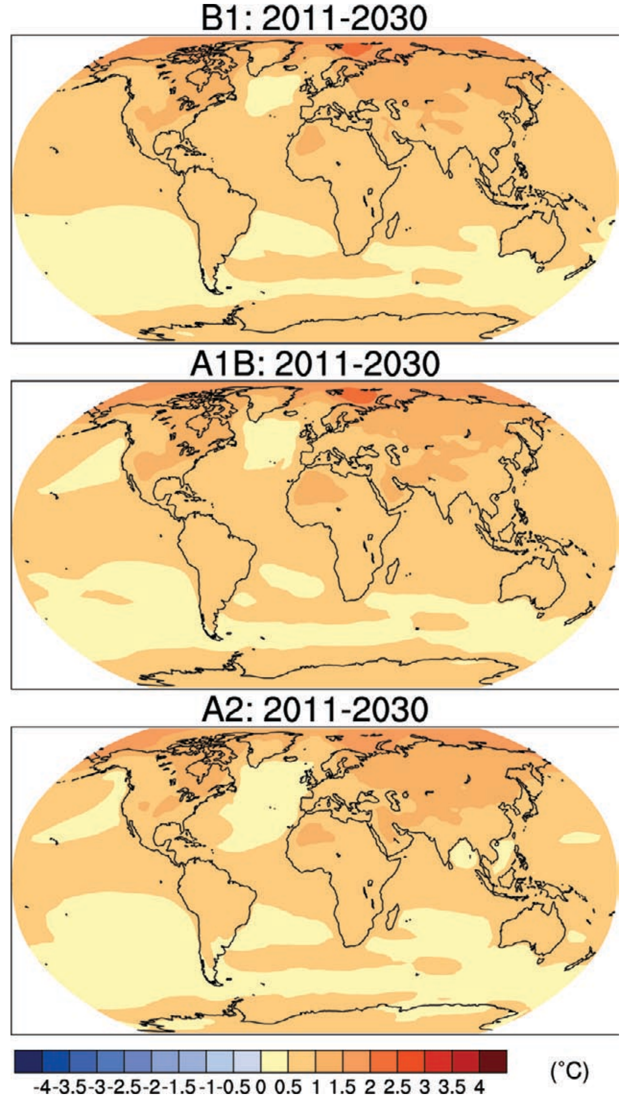


Fig. 1. Near-term surface air temperature anomalies from CMIP3 multimodel projections, 2011–30 minus 1980–99 (°C), for the (top) low, (middle) medium, and (bottom) high emission scenarios from IPCC AR4 (Figure: from *Climate change 2007: The Physical Science Basis*. Working Group I Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Fig. 10.8, Cambridge University Press).

Projections of how anticipated changes in greenhouse gases and aerosols will influence climate over time scales of several decades to centuries (dec–cen) can be considered primarily as "boundary condition problems" (Fig. 2). Such model-based projections seek to describe climate trends, not the details of individual

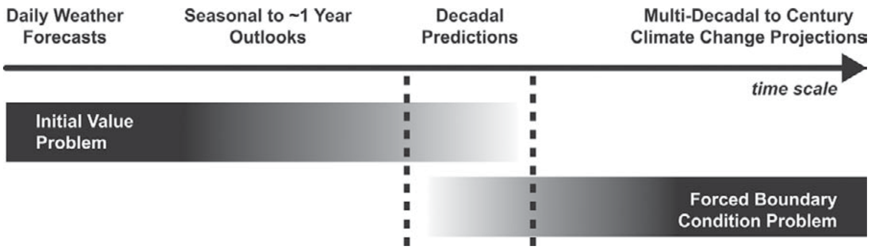


Fig. 2. Schematic illustrating progression from initial value problems with daily weather forecasts at one end, and multidecadal to century projections as a forced boundary condition problem at the other, with seasonal and decadal prediction in between.

days, seasons, or years. In contrast, daily weather forecasts and shorter-term SI climate predictions [e.g., El Niño–Southern Oscillation (ENSO) forecasts] can be thought of as "initial value problems," for which detailed knowledge of the observed current conditions are crucially needed to define the starting point (the initial conditions). Lorenz (1963) demonstrated how, even if one possessed a hypothetically perfect numerical model representing all of the physical processes completely and without error, unavoidable uncertainties in the initial conditions will invariably grow and contaminate the numerical simulation of transient weather systems. This sensitivity to initial conditions (sometimes referred to as the "butterfly effect") limits to about 2 weeks the time period over which even a perfect model could yield skillful weather forecasts. When considering El Niño, a quasi-oscillatory phenomenon that evolves more slowly than synoptic weather systems, skillful numerical forecasts of monthly mean or seasonal mean conditions (Shukla 1984) can be made with a lead time of 6–12 months (Kirtman et al. 2002). For example, at 8 months multimodel correlation coefficients for Niño-3.4 are approximately 0.75, and then they drop to 0.6 at 10 months, and then 0.5 at 12 months. However, predictability varies on decadal time scales (e.g., Tang et al. 2008), and the ultimate predictability limits are not well established.

For many climate variables, decision makers are interested in the 10–30-yr time horizon (e.g., Pulwarty 2003), a time period that is characterized by a forced climate change signal that is often weaker than or comparable to the magnitude of internally generated climate variations. If skillful decadal climate predictions are to be realized, the time scale for which initial conditions are shown to impact the predictions will need to be extended by roughly an order of magnitude beyond today's El Niño forecasts. That is, decadal prediction involves having some predictable signal in the initial state that has been ignored in traditional dec–cen climate change simulations.

In the decadal time range, at the confluence between dec–cen and SI, there may be a "sweet spot" for an enhanced signal-to-noise ratio of climate change information. The relative uncertainty in global-mean, decadal-mean surface air temperature predictions initially decreases with lead time as the predictions transition from initial state dependence to the forced response out to about 40 yr (Fig. 3). At longer lead times the emissions scenario uncertainty generally becomes dominant (Hawkins and Sutton 2009a).

Even if uncertainty is low in the decadal range relative to other periods, there remains the question of the signal-to-noise ratio, namely, the extent to which predictable regional variations could rise above noise from uncertainties in the forced response, and also from unpredictable aspects of internal variability, on those time and space scales (Barnett et al. 2008). On continental scales, the observed response to external

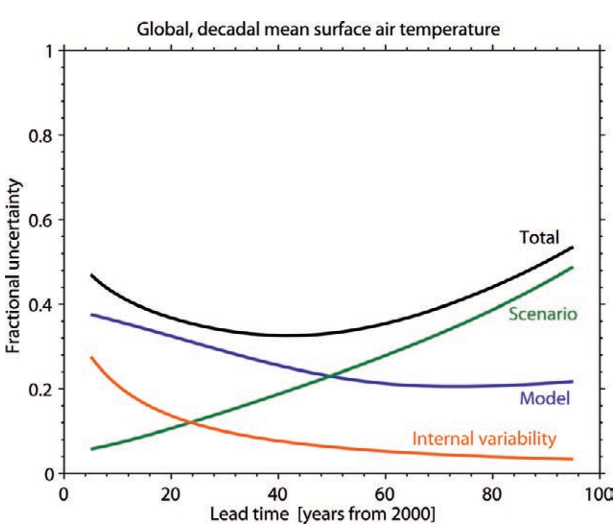


Fig. 3. The relative importance of different sources of uncertainty in IPCC GCM projections of decadal-mean global-mean surface air temperature in the twenty-first century is shown by the fractional uncertainty (i.e., the prediction uncertainy divided by the expected mean change, relative to the 1971–2000 mean). Model uncertainty is the dominant source of uncertainty for lead times up to 50 yr, with internal variability being important for the first decade or so. Scenario uncertainty becomes important at multidecadal lead times (from Hawkins and Sutton 2009a).

forcing has clearly emerged from decadal climate variability (Hegerl et al. 2007). However, on spatial scales smaller than the subcontinental scale, it takes several decades for the forced signal to emerge (Karoly and Wu 2005; Knutson et al. 1999). The situation becomes more difficult for other climate variables, such as precipitation, where presently even large-scale forced changes are only marginally separable from internal climate variability (e.g., Zhang et al. 2007; Min et al. 2008).

Thus, some unresolved questions remain regarding not only how to conduct decadal predictions, but also regarding the quality and usefulness of the results. As we stand at the threshold of a new area of research, there are a variety of science questions that need to be addressed, and we turn to those next.

OBSERVED AND MODELED DECADAL PHENOMENA THAT COULD POTENTIALLY CONTRIBUTE TO PREDICTION SKILL.

CMIP3 models can already simulate the magnitude of observed decadal surface temperature variability over land. The potential for skillful decadal

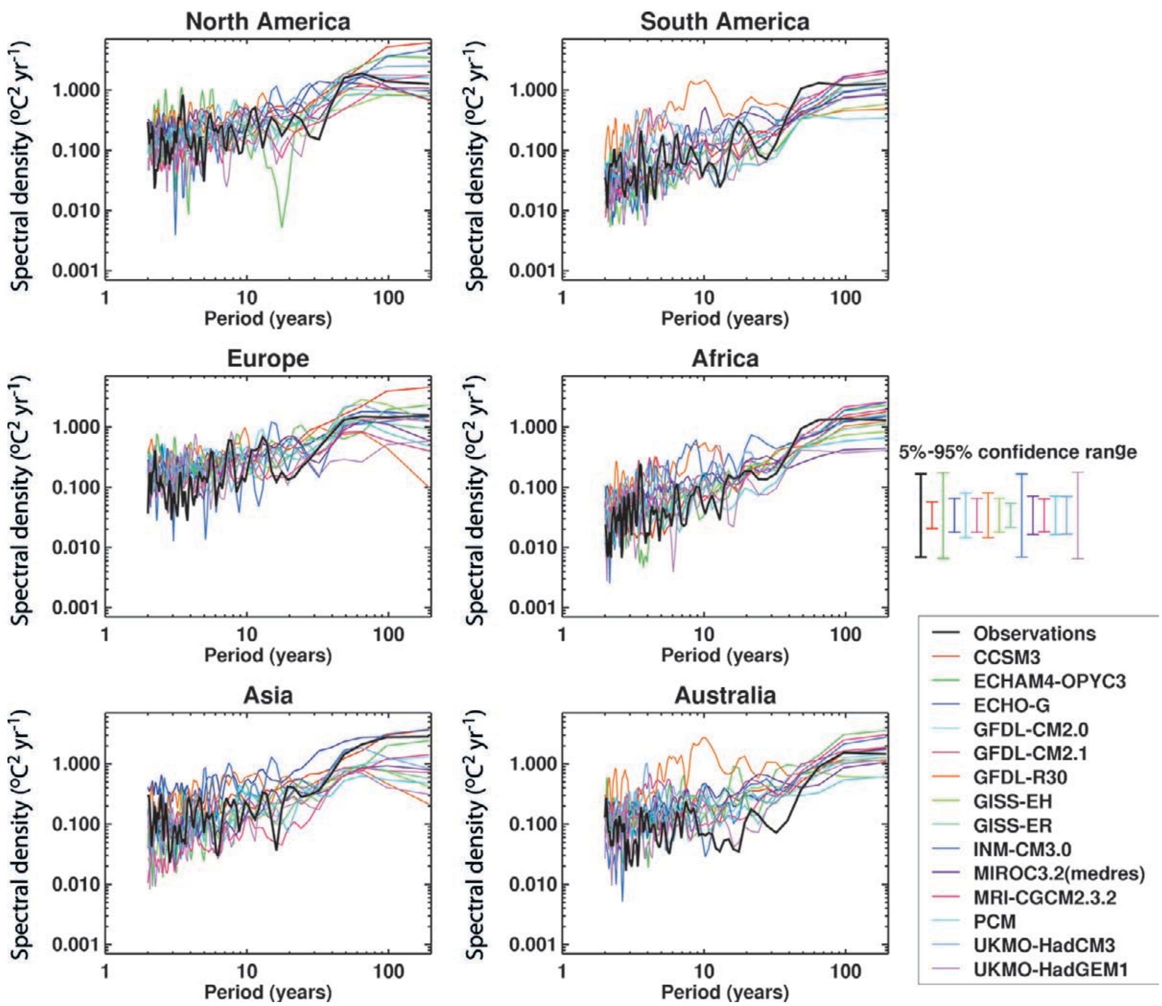


Fig. 4. Comparison of variability as a function of time scale for continental mean temperature for continental regions (°C² yr⁻¹) from the observed record [Hadley Centre Climatic Research Unit gridded surface temperature data set (HadCRUT3)] and from CMIP3 AOGCM simulations assessed for the IPCC AR4. Models include both anthropogenic and natural forcings. All power spectra are estimated using a Tukey–Hanning filter of width 97 yr. The model spectra displayed are the averages of the individual spectra estimated from individual ensemble members. Most models simulate variability on decadal time scales and longer, which is consistent with observations. Exceptions where models are inconsistent with the observations (at the 10% significance level) are two models over South America, five models over Asia, and two models over Australia (Figure: from *Climate Change 2007: The Physical Science Basis*. Working Group I Contribution to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Fig. 9.8, Cambridge University Press).

predictions depends largely on whether models simulate sufficient decadal climate variability both in terms of magnitude as well as structure. In that regard, the temperature variability of coupled climate models over global and continental space scales has been shown to be realistic, even on time scales of multiple decades (Fig. 4; Hegerl et al. 2007). Note that it is essential for comparisons between model-simulated and observed variability to compare data that contain both the response to external forcing and internal climate variability.

For this purpose, twentieth-century simulations are particularly useful, because they should have similar contributions from external forcing as the observations. A comparison of reconstructed past temperature and models forced with appropriate external forcing in paleoclimate simulations also confirms that the current generation of climate models appear to be able to simulate low-frequency temperature variability on spatial scales of continents and larger (Jansen et al. 2007), as well as on subcontinental scales (e.g., Karoly and Wu 2005; Hegerl et al. 2007). Models also simulate many mechanisms of climate variability (e.g., improved simulations of El Niño) and show similar patterns of coupled variability as observed (Randall et al. 2007).

It is presently less well understood to what extent teleconnections arising from modes of variability are well simulated. Also, there is some evidence that the precipitation variance in many latitude bands is underestimated by climate models on average by about a factor of 2 (Zhang et al. 2007). However, firm conclusions cannot be drawn because the sparse spatial coverage of observed precipitation data can lead to an overestimate of precipitation variability.

Examples of decadal time-scale phenomena that could increase decadal prediction skill. Pacific. In the Pacific Ocean region, there are several candidate sources of decadal variability that, if captured in a decadal prediction, could contribute to enhanced prediction skill certainly over some regions of the Pacific Ocean and perhaps could extend to other regions over land in the Pacific rim countries. One is externally forced by the 11-yr solar cycle (e.g., van Loon and Shea 1999; van Loon and Labitzke 1998; van Loon et al. 2007), with the amplitude of tropical Pacific SST variability about half that of ENSO (Meehl et al. 2008).

A number of mechanisms have been proposed to explain this response. Two of the most likely involve either stratospheric ozone being affected by solar forcing, with concomitant changes in temperature and winds starting in the stratosphere and extending into the troposphere (e.g., Shindell et al. 1999), with

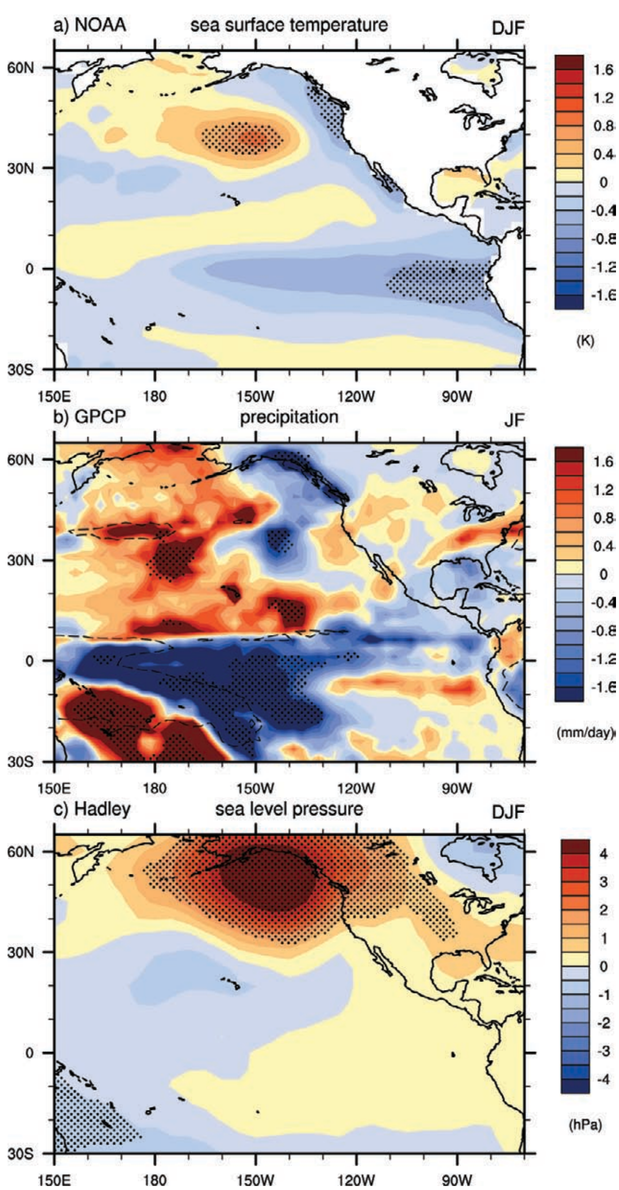
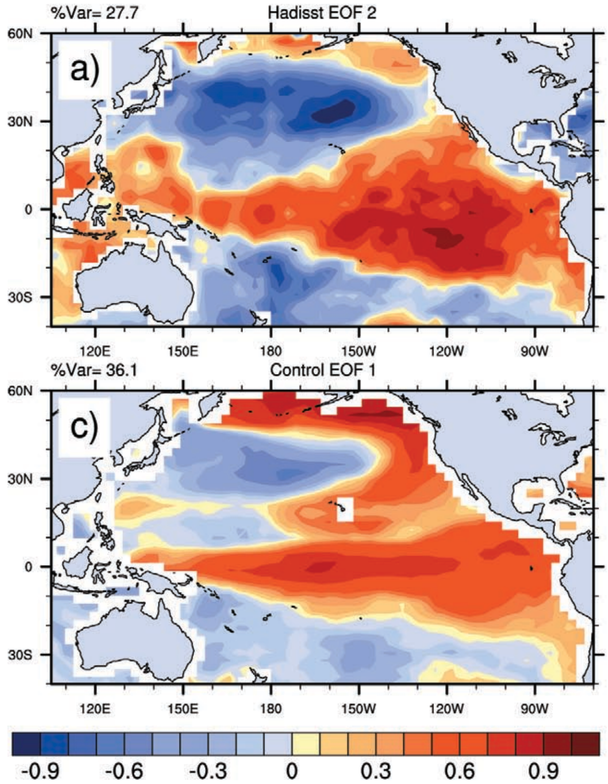


Fig. 5. (a) The average anomalies of sea surface temperature in 11 solar peak years (°C), computed relative to all other years, Dec–Feb, from the NOAA Extended Reconstructed Sea Surface Temperature dataset; (b) the average tropical rainfall anomalies [Global Precipitation Climatology Project (GPCP) gridded precipitation dataset] in the solar peak years starting in the late 1970s (mm day^{-1}), Jan–Feb, in comparison to all other years. Dashed line is the 6 mm day^{-1} contour from the long-term mean climatology; (c) same as (a), but for the average anomalies of sea level pressure (Hadley Centre sea level pressure dataset) in 11 solar peaks (hPa), Dec–Feb. Shading indicates significance at or above the 95% level, indicating the relative magnitude of the anomalies compared to the noise (Meehl et al. 2008).



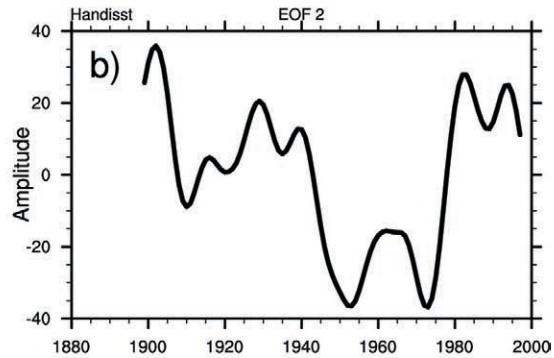


Fig. 6. (a) The second EOF (the first EOF is the trend) of 13-yr low-pass-filtered non-detrended observed SSTs for the period of 1890–2006, (b) PC time series for second EOF, (c) the first EOF of 13-yr low-pass-filtered SSTs from a 300-yr period of an unforced model control run (Meehl et al. 2009a). Units for panels (a) and (c) are arbitrary, PC time series is in °C.

strengthened tropical precipitation (Balachandran et al. 1999), or a direct coupled ocean–atmosphere response to solar forcing that also enhances tropical precipitation that produces a La Niña–like pattern of SSTs in the Pacific with anomalously cold surface temperatures in the equatorial eastern Pacific as shown in Fig. 5 (e.g., van Loon et al. 2007). These mechanisms appear to work in the same sense and add together to produce an enhancement of the climatological mean precipitation in the Pacific region, stronger Hadley and Walker circulations, intensified trade winds and upwelling, and cooler equatorial Pacific SSTs (Meehl et al. 2008, 2009a). Thus, by taking into account a climatological 11-yr solar forcing, there could be some enhanced predictive skill for the tropical Pacific and associated teleconnections to midlatitude continental climate (Fig. 5; Meehl et al. 2008).

Another source of Pacific decadal variability that is related to internally generated variability is referred to as either the Pacific decadal oscillation (PDO; e.g., Mantua et al. 1997) or the North Pacific Index (NPI; Deser et al. 2004). Both usually denote decadal variability in the North Pacific. The Interdecadal Pacific Oscillation (IPO; e.g., Power et al. 1999) has similar patterns of variability in the North Pacific, but also encompasses SST patterns across the entire Pacific Ocean region, with an explained low-frequency variance of tropical Pacific SSTs of about 10% (Power and Colman 2006). The PDO and IPO are usually characterized by a low-pass-filtered SST EOF pattern that has an "El Niño–like" character, with SST anomalies of one sign in the tropical central and eastern Pacific and northeastern and southeastern midlatitude Pacific, and opposite sign anomalies in the northwest and southwest Pacific (Fig. 6a), with decadal-to-multidecadal time scales of variability (Fig. 6b). The PDO and NPI focus on the part of this pattern in the North Pacific, and the IPO takes into account the entire Pacific-wide pattern. In the observations, there are a variety of other forcing mechanisms that could contribute to this pattern during the twentieth century, but long control runs with global coupled models show that this pattern on the decadal time scale is internally generated by the models (Fig. 6c). This suggests that the similar low-frequency pattern in the observations (pattern correlation of +0.63 between the observations in Fig. 6a and the model pattern in Fig. 6c; note that the observed pattern is noisy due to small number of samples) is internally generated as well. However, this type of variability has been connected to extremes such that the Pacific decadal variability index correlates at 0.8 with the number of unusually warm or cold daily temperatures over parts of North America in the Northern Hemisphere's cold season (e.g., Kenyon and Hegerl 2008), and up to 0.6 with an index of intense precipitation.

Though it has been postulated that the PDO/NPI/IPO are simply products of amplitude modulation of interannual El Niño and La Niña events (e.g., Jin 2001), possibly related to changes in the tropical Pacific background state (Imada and Kimoto 2009), arguments have been made that these decadal phenomena may not be a product of dynamical ocean–atmosphere coupling because models with only a mixed layer ocean produce the PDO pattern (Pierce et al. 2001), or, conversely, they may have deterministic mechanisms that are separate from ENSO and thus may be predictable beyond the ENSO time scale (e.g., White and Cayan 1998; Meehl and Hu 2006; Power and Colman 2006).

There is also the issue of whether and/or how much external forcing has affected the time evolution of the PDO/IPO as depicted in Fig. 6a. For example, a significant shift occurred in the mid-1970s when the tropical Pacific SSTs transitioned from a relatively cool state to relatively warm conditions (e.g., Trenberth and Hurrell 1994). Though this shift could have been entirely natural, there is evidence that the transition had a partial contribution from changes in external forcing from increases in anthropogenic greenhouse gases (Meehl et al. 2009b). Thus, predictive skill on the decadal time scale for the Pacific is most likely to be achieved by taking into account the interactions of external forcing (both anthropogenic and natural) and internally generated decadal variability (Mochizuki et al. 2009, manuscript submitted to *Proc. Natl. Acad. Sci.*).

Atlantic. Many coupled general circulation models (CGCMs) used in climate studies show multidecadal oscillations in their Atlantic meridional overturning circulations (AMOCs; e.g., Delworth et al. 1993; Dong and Sutton 2005; Danabasoglu 2008). These AMOC fluctuations are mostly irregular, and their periods range from 20 to more than 100 yr among models. Other ocean fields also exhibit similar variability, including sea surface temperatures (SSTs) and northward heat transport.

Observational studies based on instrumental and proxy data also show distinct multidecadal variability in SSTs with a broad hemispheric pattern in the Atlantic Ocean and with periods of about 40–70 yr (e.g., Kushnir 1994; Delworth and Mann 2000). This multidecadal variability is sometimes referred to as the Atlantic Multidecadal Oscillation (AMO) or Atlantic multidecadal variability (AMV). It has been shown to have climate impacts in regions outside the North Atlantic (Knight et al. 2006), such as those associated with multidecadal variations of the North American and western European summertime climate (Sutton and Hodson 2005), and Northern Hemisphere–averaged surface temperature (Zhang et al. 2007). There is a broad resemblance between model-simulated and observed multidecadal SST variability patterns in the North Atlantic that is usually associated with the AMOC (e.g., Delworth et al. 1993).

The variability of AMOC, and possibly the associated climate changes, may be predictable on decadal or longer time scales (Griffies and Bryan 1997), implying potential predictability for seasonal hurricane activity in the North Atlantic. The presence of such multidecadal intrinsic variability also complicates climate studies investigating anthropogenic effects. Some recent modeling studies, however, suggest stronger ties of this AMOC variability with the North Atlantic Oscillation (NAO; e.g., Eden and Jung 2001; Dai et al. 2005; Dong and Sutton 2005; Danabasoglu 2008). The driving mechanism(s) of this AMOC oscillation, as well as whether it represents an atmosphere–ocean coupled mode or an ocean-alone mode, remain largely unresolved, showing differences among various climate models (Latif et al. 2006).

SCIENCE AND DATA ISSUES. *Initialization.* Initializing climate models offers the potential to predict internal variability in addition to externally forced climate change on decadal time scales and is thought to be at the heart of the decadal predictability/prediction problem. Although idealized model experiments show considerable promise for predicting internal variability, particularly in the North Atlantic (Collins et al. 2006), there are technical obstacles that must be overcome if such potential predictability is to be achieved in reality. A fundamental problem is that climate models are unable to simulate the observed climate perfectly. When initialized with observations, models therefore drift toward their preferred imperfect climatology, leading to biases in the forecasts. It is standard practice to remove such biases from SI forecasts by an a posteriori empirical correction computed from a series of hindcasts (Stockdale 1997). This strategy is potentially less applicable for decadal prediction, because the smaller magnitude of the predictable signal is more likely to be masked by inaccuracies in the bias correction computed from the comparatively short period, and because nonlinearities will inevitably grow with the length of the experiments. An alternative approach, known as "anomaly initialization" (Schneider et al. 1999), has therefore been tried (Barnett et al. 2004; Pierce et al. 2004; Smith et al. 2007; Keenlyside et al. 2008; Pohlmann et al. 2009). In this approach, models are initialized with observed anomalies added to the model climate, and the mean model climate state is subtracted to obtain forecast anomalies. The model climate is usually obtained from transient simulations of the twentieth century, making this approach relatively expensive. These two approaches for dealing with model drift have been tested to some degree in SI forecasts. For example, ENSO hindcasts, made using the NOAA Coupled Forecast System, have been found to give similar skill using either approach. However, their relative merits have yet to be quantified on decadal time scales.

Historically the subsurface ocean has been very sparsely observed, and some of the data appear to be significantly biased (Domingues et al. 2008; Ishii and Kimoto 2009), making the development and testing of ocean initialization schemes difficult. A simple approach that avoids the difficulties with historical subsurface ocean observations is to initialize models by assimilating only sea surface temperatures (Keenlyside et al. 2008), relying on ocean transport processes in the model to initialize the subsurface ocean indirectly. At the National Center for Atmospheric Research (NCAR) and Max-Planck Institut (MPI), an alternative approach is being tested in which subsurface ocean temperature and salinity can be diagnosed from an ocean model forced by atmospheric reanalysis data based on observations, and then nudged into a coupled model to produce initial conditions for forecasts (P. Gent and D. Matei 2009, personal communication). However, the direct use of subsurface ocean observations would be expected to improve forecast skill. Several reanalyses of historical ocean observations have been constructed and are being evaluated through the Climate Variability and Predictability (CLIVAR) Global Synthesis and Observations Panel (GSOP) intercomparison project to help understand the differences among these products, and to provide insight into why they may disagree. Temperature and salinity fields from two of these have already been used to initialize models for decadal forecasts (Smith et al. 2007; Pohlmann et al. 2009), and there is evidence that analyzed currents can also be included in the initialization (Kirtman and Min 2009; G. Danabasoglu and J. Tribbia 2009, personal communication). In this way, modeling groups without data assimilation schemes can perform initialized climate predictions. Ultimately, however, fully coupled data assimilation schemes that take advantage of covariances between ocean and atmosphere variables to generate an optimal estimate of the climate system would seem to potentially offer the most forecast skill, and they are being developed by some groups (Sugiura et al. 2008; A. Rosati et al. 2009, personal communication).

A significant issue related to initialization is the treatment of sea ice and the ocean conditions under the sea ice. Most ocean synthesis products have strong climatological restoration poleward of 60°N and 60°S and do not assimilate observations there. Dense water formation there may have a strong climate impact, and the initial state in those regions must somehow be made coherent with the ocean initialization elsewhere.

Studies of historical periods are important in order to assess the likely skill of forecasts over a range of different climate states. Recent and planned improvements to the observational network promise significant improvements in future forecast skill. Perhaps the most important of these is the recent deployment of a global array of profiling floats by the Argo program (see www.argo.ucsd.edu/). These provide many more measurements of both temperature and salinity over the upper 2 km of the world's oceans than were available historically, potentially offering a step change in our ability to initialize ocean heat and density anomalies.

In addition to ocean temperature and salinity, initialization of other aspects of surface climate, notably sea ice, snow cover, frozen soil, and soil moisture, may have the potential to contribute to predictive skill beyond the seasonal time scale. Initialization of these variables has not been attempted in decadal prediction studies to date, although the process of ocean initialization [and of atmospheric initialization in the case of Smith et al. (2007)] may allow some aspects of the observed anomalous patterns to be captured in the initial conditions. Additionally, the technique used in the Global Soil Wetness Project (GSWP), whereby atmospheric forcing is used to initialize soil moisture, could be applied to the decadal prediction problem. Explicit initialization could also be investigated, for example, by using measurements of soil moisture from the planned Soil Moisture and Ocean Salinity (SMOS) satellite, or by initializing sea ice thickness with observations from the planned CryoSat-2 satellite.

Generation of ensemble forecasts. The decadal prediction problem requires strategies for sampling the spread of possible outcomes consistent with initial state uncertainties. These can be represented by using ensembles of coupled model forecasts distinguished by perturbations to the applied initial conditions.

Smith et al. (2007) created four member hindcast ensembles involving the use of assimilated analyses of ocean and atmospheric observations taken from consecutive days immediately preceding the hindcast start date. These were then combined with ensembles started from previous seasons to improve the sampling of uncertainties in low-frequency variability. Keenlyside et al. (2008) performed three-member hindcast ensembles, with each member started from simulations of twentieth-century climate into which analyses of observed SST anomalies had been assimiliated. These techniques are sufficient to generate a spread of possible outcomes. For example, confidence limits on hindcasts of annually averaged global temperature, diagnosed from the ensemble spread, can give a reasonably good indication of actual hindcast errors (see Smith et al. 2007, and supporting online material therein). However, it is likely that more sophisticated approaches could achieve a better characterization of the growth of forecast uncertainties associated with initial state errors.

One approach could be to identify a set of perturbations that optimally capture the fastest-growing forecast errors, following methods such as breeding vectors (Toth and Kalnay 1997; Vikhliaev et al. 2007) or singular vectors (Molteni et al. 1996). Such techniques are commonly used in ensemble weather forecasting and are now being applied to longer-term climate predictions (e.g., Kleeman et al. 2003). An example of an optimal perturbation is shown in Fig. 7 for the Atlantic domain in the third climate configuration of the Met Office (HadCM3) coupled model (Hawkins and Sutton 2009b), showing that the far North Atlantic is the most sensitive region to small anomalies in this model, and is thus the optimal region both for perturbations to sample forecast uncertainties and for targeted observations to help constrain its predictions. An alternative option could be the use of ensemble assimilation methods such as the ensemble Kalman filter (Evensen 1994), in which analyses of observations are created by using the forecast model and observations to update an ensemble of previous analyses, accounting for analysis, model, and observational errors. In prediction systems which assimilate analyses of observations created offline, another alternative could be to perturb the analyses consistent with their errors, noting that these would arise both from the observations themselves, and from the analysis methods used to convert them into spatially complete fields. Given that different approaches all have potential strengths and limitations, it remains an open research question to identify the best methods for representing initial state uncertainties in decadal predictions.

Modeling errors are known to be an important source of uncertainty in predictions of internal climate variability on seasonal time scales (e.g., Hagedorn et al. 2005) and the response to externally forced climate change on multidecadal time scales (e.g., Meehl et al. 2007). Given that uncertainties arising from forced climate change are likely to contribute significantly to the total uncertainty in predictions for the next few decades (Fig. 3; Hawkins and Sutton 2009a), it is important that ensemble forecasts are constructed to sample model as well as initial state uncertainties. The multimodel approach of constructing ensembles from different available GCMs has been shown to provide improved estimates of uncertainty in seasonal forecasts compared to single-model ensembles using only perturbed initial conditions (Hagedorn et al. 2005). This method has improved attribution results, for example, for precipitation, thus suggesting increased skill (e.g., Zhang et al. 2007), and has been used extensively to provide quantitative uncertainty estimates in multidecadal climate change projections

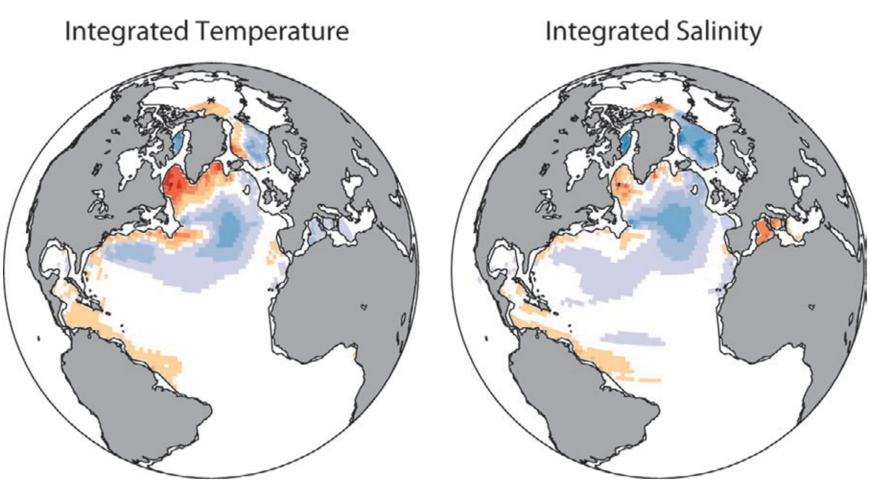


Fig. 7. An optimal perturbation for the Atlantic domain from the HadCM3 model, using a linear inverse modeling approach (from Hawkins and Sutton 2009b). The panels show integrated (left) temperature (in K) and (right) salinity (in PSU) multiplied by five from the surface to a depth of 1,800 m. The colored regions indicate where the ocean is sensitive to small anomalies, and are thus the optimal regions for initial condition perturbations and for targeted observations to improve forecast skill. The color scale is the same in both panels and is arbitrary. White regions represent small anomalies of either sign.

(e.g., Tebaldi and Knutti 2007; Meehl et al. 2007). A second approach that is based on systematic perturbation of uncertain parameters in a single model has also been studied in the context of long-term climate prediction (e.g., Murphy et al. 2007). A third method consists of applying random rather than sustained perturbations to the model physics, through the introduction of terms designed to represent stochastic aspects of the parameterization of subgrid-scale processes. These stochastic–dynamic parameterization schemes have been applied to the seasonal forecast problem (e.g., Berner et al. 2008) and also are proposed for longer-term climate predictions (T. N. Palmer et al. 2009). The European Union ENSEMBLES project has undertaken an initial comparison of these three methods of sampling modeling uncertainties in seasonal and annual hindcasts (Doblas-Reyes et al. 2009), and this study is currently being extended to a set of decadal hindcasts initialized during the period of 1960–2005.

Predictability and predictions. Both internal variability and the forced response are important sources of potential predictability in global-scale projections (Fig. 3). At a regional level their relative importance varies significantly (Boer 2009a), with the forced response largest over parts of the tropical oceans, and the internal variability contribution larger over the middle- and high-latitude oceans. Given that the predictable component of climate anomalies on annual to decadal time scales in any location may typically be modest compared to the unpredictable component, it is important that a large dataset of hindcasts is built up to provide robust estimates of skill. Additionally, decadal variability could change in a future warmer climate, thus there is the potential that decadal potential predictability of the internally generated component could decrease (Boer 2009b).

Hindcast studies performed over a limited number of past decades will inevitably give results that depend on the characteristics of observed decadal variability during the relevant period. For example, Smith et al. (2007) found that initializing their climate model with observed ocean anomalies (compared with parallel simulations without initialization) gave a particularly large improvement in regional skill over the Indian and Australasian sectors of the southern oceans, based on hindcasts started from dates covering 1982–2001. This was largely because of the disparity between observed upper-ocean temperature anomalies and those simulated in the uninitialized hindcasts, which happened to be large in this region over these particular decades. However, the largest disparities could occur in other parts of the world during alternative periods, so there is no guarantee that regional variations in skill diagnosed from past cases will be a robust guide to future performance. Pohlmann et al. (2009) come to similar conclusions.

For projecting a decade or two ahead, the role of uncertainties in anthropogenic emissions of forcing agents is likely to be relatively small in general (e.g., Fig. 3), although there could be exceptions in regions where uncertainties in the forcing resulting from spatially heterogeneous agents, such as tropospheric aerosols (Schulz et al. 2006), are largest. Another issue concerns the impact of future variations in natural forcing, such as the solar cycle and explosive volcanic eruptions (Shiogama et al. 2009, manuscript submitted to *SOLA*). In a real forecasting situation the solar cycle and eruptions cannot be predicted, so there is a strong case for assuming no past knowledge of those forcings in hindcast studies (a "no cheating" strategy). The hindcast studies published to date (see "Examples of decadal prediction" section) all follow this principle. If in practice there is no eruption during a forecast period, then a subset of hindcasts based on the no-cheating strategy will give some guidance on likely forecast skill. However, if there is an eruption during the forecast period, then past skill statistics based on a no-cheating strategy will be less informative, because any hindcasts for which the observed verification data were affected by a post-initialization eruption will give a misleadingly pessimistic estimate of skill. On the other hand, hindcast skill estimates assuming prior knowledge of solar activity and past eruptions could be too optimistic, because a source of large forced anomalies [the response to which is likely to be relatively predictable (e.g., Soden et al. 2002)] would be present in the hindcast dataset, but not in the forecast. Another factor is that past knowledge of solar variability and eruptions is now assumed in most historical climate change simulations (see Table 10.1 of Meehl et al. 2007), so there is a case for following the same strategy in initialized decadal hindcasts from a resource perspective, because this allows existing historical climate simulations (at least for modeling groups possessing these) to be used as a "no initialization" baseline for the assessment of hindcast skill.

Examples of decadal predictions. There were three recent efforts at decadal prediction, all with the following similar strategy: Initialize a global climate model from observations and reanalyses and run it forward 10 yr, while accounting for changes in external forcing (natural and anthropogenic). In the first

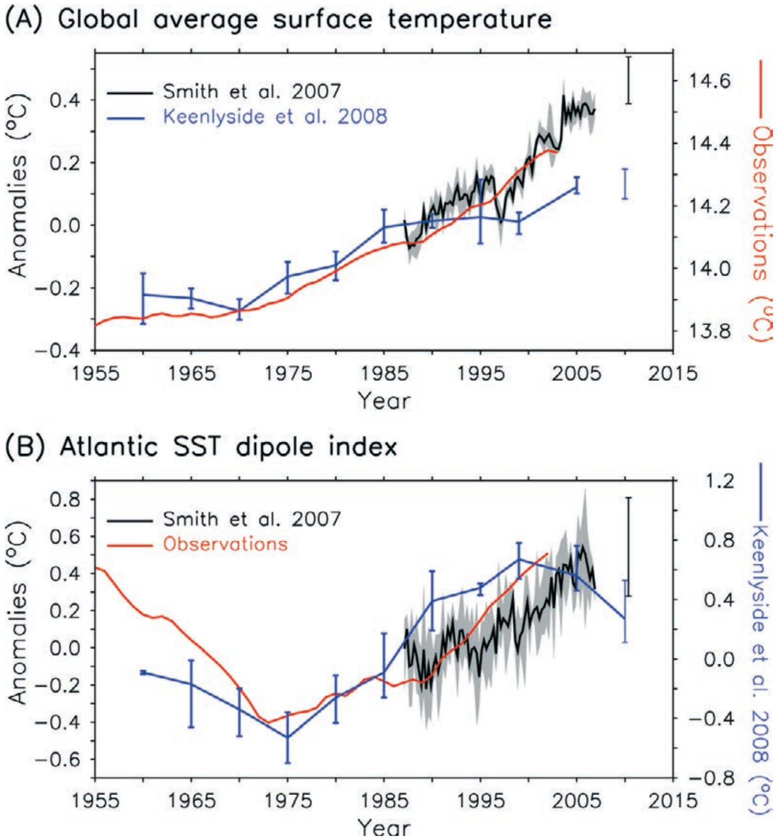


FIG. 8. Decadal prediction examples. Observed and hindcast values of (a) 10-yr mean global mean surface temperature and (b) an Atlantic SST dipole index. The latter is a proxy for MOC fluctations and is defined as the average SST difference for 60°–10°W, 40°–60°N minus 50°–0°W, 40°–60°S. Hindcasts begin in 1982 (1955) in Smith et al. (2007) and Keenlyside et al. (2008), with a four (three) member forecast every season (5 yr); shading (error) indicates the ensemble range. The error bars centered on 2010 represent actual forecasts for the period of 2005–15. Hindcasts for Smith et al. (2007) and Keenlyside et al. (2008) are adjusted to have the observed means over the 1979–2001 (1955–2005) period. Note the different axis used in (b) for Keenlyside et al. (2008). Observations are from HadISST 1.1 and HadCRU3.

work Smith et al. (2007) showed that global-mean temperature could be predicted out to a decade in advance (Fig. 8a), with more skill than obtained when only external radiative forcing (boundary condition) changes are accounted for (Fig. 8). Beyond the first year, this skill enhancement resulted mainly from initialization of the upper-ocean heat content. There was also skill enhancement in predictions of multiyear averages of surface temperature in some regions, including the Indian Ocean and parts of the Southern Ocean.

In the second study Keenlyside et al. (2008) demonstrated that SST variations associated with the Atlantic MOC could be predicted a decade in advance, but because of an overly strong MOC signal, their strength was overestimated (Fig. 8b). There was skill in predicting 10-yr mean surface temperature variations over parts of the North Atlantic sector, including Europe and North America, and the tropical Pacific, greater than that obtained from the specification of external radiative forcing alone. Ten-year-averaged global surface temperature variations were also predictable (Fig. 8a), but with marginally less skill than that obtained from radiative forcing only.

In both studies forecasts were made for the next 10 yr (Fig. 8b), and in both cases natural internal variability was found to temporarily offset anthropogenic global warming. The offset was largest in Keenlyside et al. (2008), whose results suggest a temporary lull in global warming for the next decade; however, the simplicity of the scheme employed needs to be kept in mind. The results of both studies highlight the impact of internal variability on the evolution of surface temperature, both globally and regionally, over the next decade and warrant further investigation.

The third study, Pohlmann et al. (2009), showed predictive skill through the initialization up to the decadal time scale, particularly over the North Atlantic. Viewed over all time scales analyzed here (annual, 5-yr mean, and 10-yr mean), greater skill for the North Atlantic SST is obtained in the hindcast experiments than either damped persistence or a trend forecast. The hindcast Atlantic Meridional Overturning Circulation follows closely that of the German contribution to the Estimating the Circulation and Climate of the Ocean (GECCO) oceanic synthesis used in the initialization. Hindcasts of global-mean temperature do not obtain greater skill than either damped persistence or a trend forecast, resulting from the SST errors in the GECCO synthesis, outside the North Atlantic. An ensemble of forecast experiments is performed subsequently over the period of 2002–11. North Atlantic SST from the forecast experiment agrees well with observations until the year 2007 and is higher than those simulated without the oceanic initialization, averaged over the forecast period. The results confirm that in decadal climate predictions, both the initial and the boundary conditions must be accounted for.

DECADAL PREDICTION EVALUATION.

An important advantage of decadal over centennial predictions is that the likely skill can potentially be quantified in tests of past cases, or hindcast experiments, as has been mentioned previously. In these, the model is used to "forecast" a historical period, but only using data that would have been available prior to this period (though this is not quite an independent test because these data generally are used to develop, test, and tune the model). The accuracy of the model can then be assessed by comparing with what actually happened. A large set of such hindcasts is typically made in order to obtain a robust estimate of the likely skill and reliability of actual forecasts.

There are many different measures of skill, although no one measure can capture all aspects of a forecast quality. For experimental forecasts with limited numbers of hindcasts, such as in the case of decadal predictions, estimates of forecast quality encounter many limitations, and care is needed when interpreting the results. For example, the correlation between forecast and observed anomalies can be a useful and easily interpretable measure of the ability to predict the phase of natural cycles such as ENSO. However, on decadal time scales very high anomaly correlations can be achieved simply by predicting the warming trend in response to increased greenhouse gases, giving little guidance on any ability to predict natural internal variability beyond the forced response. Even if one were to investigate the relative skill of two different forecasts, such as comparing initialized forecasts with radiatively forced forecasts, improved forecast quality resulting from the initial conditions does not necessarily indicate that the initial conditions provide predictability of natural decadal variability. The improvement may just have come about by the initial conditions better quantifying the ocean's thermal state, and thus with bias correcting the radiatively forced projections. The rate of change between two points in time, or over two distinct periods, may provide additional information on the contribution of ocean dynamics to low-frequency climate variability and change.

Beyond these simple diagnostics, detection and attribution techniques that use ideas from signal processing (Hegerl et al. 2007) may help to separate the influence of forcing and initial conditions in the presence of climate variability, though this would likely require multiple large ensembles (one with boundary forcing, and one with both initialized and boundary forcing) to adequately sample the signal-to-noise ratio. Assessing the statistical significance of any differences is also an important aspect of such comparisons, and care is needed to ensure that uncertainties arising from finite ensemble sizes, finite hindcast sets, and correlation of errors are properly accounted for.

Some examples of different skill measures are provided in recent examples of decadal predictions noted in the previous section. Smith et al. (2007) used root-mean-square error of their decadal hindcasts as their skill measure and showed that global average anomalies of annual-mean surface temperature were predicted with significantly higher skill by initialized forecasts rather than by uninitialized, radiatively forced forecasts. Keenlyside et al. (2008) used correlations of time series of initialized hindcasts with observations and climate model projections with radiative forcing changes only for several different average surface temperature series. For the global-mean surface temperature, both the initialized hindcasts and the climate model projections show very high correlations with observations resulting from the large trend in global mean temperature over the period considered. In fact, the correlation of the twentieth-century radiatively forced projections with observations is greater than that of the hindcasts, but only marginally at the 5% significance level.

One complication when measuring skill from hindcasts is that the coverage of subsurface ocean observations has recently dramatically improved with the deployment of Argo floats. Actual forecasts benefiting from Argo data are therefore potentially significantly more skillful than hindcasts based on very sparse historical observations, though the problem of data inhomogeneity is likely to be especially serious in the Southern Ocean. Experiments are therefore currently underway to assess the impact of Argo data on decadal forecasts. Another aspect also under investigation is whether skill depends on the initial state. For example, are forecasts initialized from an extreme phase of natural internal variability more skillful than those on average? This is the situation with seasonal-to-interannual climate forecasts dependant on ENSO, and it is expected to also be the case with decadal predictions (Griffies and Bryan 1997).

Confidence in decadal forecasts requires not only an assessment of model performance in hindcasts but also an understanding of the physical mechanisms giving rise to any predicted changes in climate. Additionally, an essential component of evaluating decadal predictions is to determine the effect of model systematic errors on the predictions, both in the simulation of mean climate and coupled processes that contribute to decadal time-scale variability. There is an urgent need not only to quantify model bias, but to reduce those biases. This will be an important aspect of the research activities involved with decadal prediction in CMIP5 (described below).

Ultimately, not just the quality but also the value of decadal forecasts should be quantified in terms of the societal or economic value of the prediction information to climate-related decisions or impacts studies. However, such estimates are meaningless if derived outside the context of the actual decision setting. Even in regions where it might be useful, there will be forecasts of opportunity when the information carries value together with other environmental and social indicators. The quality of the prediction information must first be assessed before prototype information can be developed and tested for value. In these cases, the consistency and probabilistic quality of the information must be measured. Of greatest interest to decision makers is the risk or likelihood of adverse or beneficial thresholds that affect management triggers.

CMIP5 COORDINATED DECADAL PREDICTABILITY/PREDICTION EXPERIMENT.

The new CMIP5 protocol for coordinated climate change experiments to be performed over the next 5 yr includes an experimental design that focuses on decadal predictability and prediction. The goal is to provide a research framework for exploring the question of how predictable climate is from one to three decades in advance, and how skillful decadal predictions out to about the year 2035 might be. The detailed requirements for the project are described by Taylor et al. (2008; see also www.pcmdi.llnl.gov/). CMIP5 emerged from extensive discussions in and beyond the CLIVAR and Working Group on Coupled Modeling (WGCM) communities, and builds on the decadal prediction protocols of the European ENSEMBLES project. Only a brief overview is given here.

There are two core experiments that are considered essential to a meaningful decadal predictability/prediction exercise, and there are a number of tier-1 experiments that add additional insight into the science questions involved with decadal prediction (Fig. 9). The first core experiment is to make a series of 10-yr hindcasts with initial observed climate states every 5 yr, starting near 1960. How to create the initial climate states is left to the discretion of the modeling groups because, as noted above, how best to initialize models is one of the central unanswered questions involved with decadal prediction. These 10-yr hindcasts should allow estimates of both the theoretical limits of decadal predictability and our present ability to make decadal predictions, accounting for both the regional decadal phenomena discussed earlier, and the climate change commitment from previous increases of GHGs. The minimum ensemble size from any given starting point is 3 members, although 10 or more ensemble members are desirable.

The second core experiment extends the integrations starting from 1960, 1980, and 2005 to 30 yr, and explores predictability and prediction over time scales thought to be more influenced by external forcing from increasing GHGs. Depending on how the initial conditions are prepared, the experimental design for the 30-yr integrations does not necessarily require long control runs of the coupled model, and thus opens the door for a wider class of models to be used in short-term climate prediction. In both core experiments, volcanic aerosol and solar cycle variability is prescribed during each integration using actual values for the past, and assuming a climatological 11-yr solar cycle and no eruptions in the future. These forcings allow an assessment of the predictability and prediction of the internal variability of the climate system, and a clean comparison with the standard CMIP5 twentieth-century runs. They allow an estimate of the skill of decadal predictions when the forcing is known, which for the future means an estimate conditional on no major volcanic eruptions.

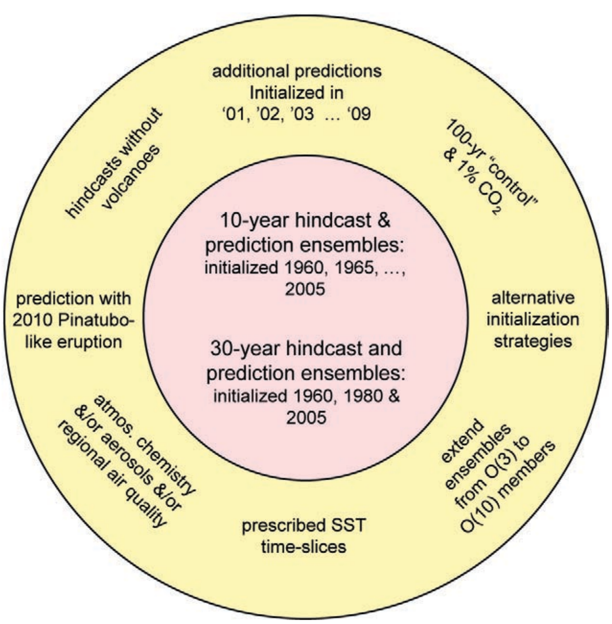


FIG. 9. Schmatic of decadal predictability/prediction experiments as part of CMIP5 (from Taylor et al. 2008).

The tier-1 integrations include simulations that start from initial climate states representing each of the years in this century when the ocean data coverage is much better than in previous years, in particular due to the Argo float data. There is also the option to perform high atmospheric resolution time slice experiments where the historical SSTs are either derived from observations or models. Further runs can study the impact of volcanoes, and others can include interactive atmospheric chemistry to investigate the impact of various short-lived species and pollutants on the predictions.

It is intended that this CMIP5 activity will not only set up a framework for coordinated multimodel experiments to address various science questions involved with decadal predictability/prediction, including the effect of model simulation errors on decadal prediction skill, but also provide the foundation for the simulations to be assessed as part of the IPCC Fifth Assessment Report (AR5). Decadal prediction is very much a research question at this early stage. Therefore, results from decadal prediction experiments must be carefully evaluated in the AR5 process so that results from CMIP5 are not misused.

CONCLUSIONS. Decadal prediction, a new field of study, focuses on time-evolving regional climate conditions over the next 10–30 yr, which is a time period of interest to infrastructure planners, water resource managers, and others. The decadal time scale offers a critical bridge for informing adaptation strategies as climate varies and changes. However, because decadal prediction is so new, there are a number of outstanding scientific and technical questions that need to be addressed. One of the chief challenges is how to initialize the modeled climate system. Because decadal prediction lies between seasonal/interannual forecasting and longer-term climate change projections, there is some knowledge from El Niño forecasting that can be applied to decadal prediction, and climate change commitment and forcing changes also provide some information as to how skillful decadal predictions might be. One of the interesting science questions involves whether an initial climate state, particularly an initial observed ocean state, can capture the proposed mechanisms that could contribute to enhanced regional prediction skill (e.g., AMOC and AMO in the Atlantic, PDO/IPO in the Pacific). Because an accurate observed initial climate state is thought to be important for decadal prediction skill, it is important to maintain a comprehensive global climate observing system, with particular emphasis on the ocean (e.g., Trenberth 2008). The use of observations to evaluate model biases, the effect of model systematic errors on prediction skill, and how to reduce those biases, are major challenges for decadal prediction.

There are questions regarding how to evaluate decadal prediction skill, what form decadal information would take, and the role such information would play in applications. An experimental framework to address decadal predictability/prediction over the next 5 yr has been incorporated into the coordinated climate change experiments of CMIP5. Some of the results of these experiments will be assessed for the IPCC AR5, in addition to guiding research activity in decadal prediction to at least 2013.

Finally, related to the question in the title, there is the issue of how skillful a decadal forecast needs to be before it is actually used. Though a definitive answer is not yet known, there are some examples where this question has been addressed with respect to the use of climate predictions for certain applications (e.g., Changnon and Vonnhame 1986; Changnon 1992; Changnon et al. 1995). As the science of decadal prediction is developed, the skill of such predictions related to their usefulness and application must also be evaluated.

ACKNOWLEDGMENTS. The authors acknowledge the Aspen Global Change Institute (AGCI) in Aspen, Colorado, and Director John Katzenberger, for hosting the workshop that laid the foundations for this paper, and particularly all the attendees who contributed to the workshop discussions that led to this paper. Funding for the AGCI workshop was provided by NASA, NSF, NOAA, DOE, and AIMES. We thank Laurent Terray and one anonymous reviewer for constructive comments. Portions of this study were supported by the Office of Science (BER), U.S. Department of Energy, Cooperative Agreement DE-FC2-97ER62402, and the National Science Foundation.

REFERENCES

Balachandran, N., D. Rind, P. Lonergan, and D. Shindell, 1999: Effects of solar cycle variability on the lower stratosphere and the troposphere. *J. Geophys. Res.,* **104,** 27 321–27 339.

Barnett, T., R. Malone, W. Pennell, D. Stammer, A. Semtner, and W. Washington, 2004: The effects of climate change on water resources in the west: Introduction and overview. *Climatic Change,* **62,** 1–11.

——, and Coauthors, 2008: Human-induced changes in the hydrology of the western United States. *Science,* **319,** 1080–1083.

Berner, J., F. J. Doblas-Reyes, T. N. Palmer, G. J. Shutts, and A. Weisheimer, 2008: Impact of a quasi-stochastic cellular automaton backscatter scheme on the systematic error and seasonal prediction skill of a global climate model. *Philos. Trans. Roy. Soc. London,* **A366,** 2561–2579.

Boer, G. J., 2009a: Decadal potential predictability of 21st century climate. *Climate Dyn.,* in press.

——, 2009b: Changes in interannual variability and decadal potential predictability under global warming. *J. Climate,* **22,** 3098–3109.

Changnon, S. A., 1992: Contents of climate predictions desired by agricultural decision makers. *J. Appl. Meteor.,* **31,** 1488–1491.

——, and D. R. Vonnhame, 1986: Use of climate predictions to decide a water management problem. *J. Amer. Water Resour. Assoc.,* **22,** 649–652.

——, J. M. Changnon, and D. Changnon, 1995: Uses and applications of climate forecasts for power utilities. *Bull. Amer. Meteor. Soc.,* **76,** 711–720.

Collins, M., and Coauthors, 2006: Interannual to decadal climate predictability in the North Atlantic: A multi-model ensemble study. *J. Climate,* **19,** 1195–1203.

Dai, A., A. Hu, G. A. Meehl, W. M. Washington, and W. G. Strand, 2005: Atlantic thermohaline circulation in a coupled general circulation model: Unforced variations versus forced changes. *J. Climate,* **18,** 3270–3293.

Danabasoglu, G., 2008: On multi-decadal variability of the Atlantic meridional overturning circulation in the Community Climate System Model version 3 (CCSM3). *J. Climate,* **21,** 5524–5544.

Delworth, T. L., and M. E. Mann, 2000: Observed and simulated multidecadal 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.

Deser, C., A. S. Phillips, and J. W. Hurrell, 2004: Pacific Interdecadal climate variability: Linkages between the tropics and the north Pacific during boreal winter since 1900. *J. Climate,* **17,** 3109–3124.

Doblas-Reyes, F. J., and Coauthors, 2009: Addressing model uncertainty in seasonal and annual dynamical ensemble forecasts. *Quart. J. Roy. Meteor. Soc.,* in press, doi:1002/qj.464.

Domingues, C. M., J. A. Church, N. J. White, P. J. Gleckler, S. E. Wijffels, P. M. Barker, and J. R. Dunn, 2008: Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature,* **453,** 1090–1093, doi:10.1038/nature07080.

Dong, B., and R. T. Sutton, 2005: Mechanism of interdecadal thermohaline circulation variability in a coupled ocean–atmosphere GCM. *J. Climate,* **18,** 1117–1135.

Eden, C., and T. Jung, 2001: North Atlantic interdecadal variability: Oceanic response to the North Atlantic Oscillation (1865–1997). *J. Climate,* **14,** 676–691.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 10 143–10 162.

Goldenberg, S., C. Landsea, A. Mestas-Nuñuz, and W. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science,* **293,** 474–479.

Griffies, S. M., and K. Bryan, 1997: Predictability of North Atlantic multidecadal climate variability. *Science,* **275,** 181–184.

Gualdi, S., E. Scoccimarro, and A. Navarra, 2008: Changes in tropical cyclone activity due to global warming: Results from a high-resolution coupled general circulation model. *J. Climate,* **21,** 5204–5228.

Hagedorn, R., F. J. Doblas-Reyes, and T. N. Palmer, 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. *Tellus,* **57A,** 219–233.

Hawkins, E., and R. Sutton, 2009a: The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.,* **90,** 1095–1107.

——, and ——, 2009b: Decadal predictability of the Atlantic Ocean in a coupled GCM: Forecast skill and optimal perturbations using linear inverse modeling. *J. Climate,* **22,** 3960–3978.

Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. Global climate projections. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 663–745.

Hurrell, J., G. A. Meehl, D. Bader, T. Delworth, B. Kirtman, H.-L. Pan, and B. Wielicki, 2009: A unified modeling approach to climate system prediction. *Bull. Amer. Meteor. Soc.,* in press.

Imada, Y., and M. Kimoto, 2009: ENSO amplitude modulation related to Pacific decadal variability. *Geophys. Res. Lett.,* **36,** L03706, doi:10.1029/2008GL036421.

Ishii, M., and M. Kimoto, 2009: Reevaluation of historical ocean heat content variations with an XBT depth bias correction. *J. Oceanogr.,* **65,** 287–299.

Jansen, E., and Coauthors, 2007: Paleoclimate. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 433–498.

Jin, F. F., 2001: Low frequency modes of tropical ocean dynamics. *J. Climate,* **14,** 3874–3881.

Karoly, D. J., and Q. Wu, 2005: Detection of regional surface temperature trends. *J. Climate,* **18,** 4337–4343.

Keenlyside, N., M. Latif, J. Jungclaus, L. Kornblueh, and E. Roeckner, 2008: Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature,* **453,** 84–88.

Kenyon, J., and G. C. Hegerl, 2008: Influence of modes of climate variability on global temperature extremes. *J. Climate,* **21,** 3872–3889.

Kimoto, M., N. Yasutomi, C. Yokoyama, and S. Emori, 2005: Projected changes in precipitation characteristics around Japan under the global warming. *SOLA,* **1,** 85–88, doi:10.2151/sola.2005-023.

Kirtman, B. P., and D. Min, 2009: Multi-model ensemble ENSO prediction with CCSM and CFS. *Mon. Wea. Rev.,* **137,** 2908–2930.

——, J. Shukla, M. Balmaseda, N. Graham, C. Penland, Y. Xue, and S. Zebiak, 2002: Current status of ENSO forecast skill. A report to the Climate Variability and Predictability (CLIVAR) Numerical Experimentation Group (NEG), CLIVAR Working Group on Seasonal to Interannual Prediction, 31 pp. [Available online at www.clivar.org/publications/wg_reports/wgsip/nino3/report.htm.]

Kleeman, R., Y. Tang, and A. M. Moore, 2003: The calculation of climatically relevant singular vectors in the presence of weather noise as applied to the ENSO problem. *J. Atmos. Sci.,* **60,** 2856–2868.

Knight, J. R., C. K. Folland, and A. A. Scaife, 2006: Climate impacts of the Atlantic Multidecadal Oscillation. *Geophy. Res. Lett.,* **33,** L17706, doi:10.1029/2006GL026242.

Knutson, T. R., and R. E. Tuleya, 2004: Impact of CO_2-induced warming on simulated hurricane intensity and precipitation: Sensitivity to the choice of climate model and convective parameterization. *J. Climate,* **17,** 3477–3495.

——, T. L. Delworth, K. W. Dixon, and R. J. Stouffer, 1999: Model assessment of regional surface temperature trends (1949-1997). *J. Geophys. Res.,* **104,** 30 981–30 996.

Kushnir, Y., 1994: Interdecadal variations in North Atlantic sea surface temperature and associated atmospheric conditions. *J. Climate,* **7,** 141–157.

Latif, M., M. Collins, H. Pohlmann, and N. Keenlyside, 2006: A review of predictability studies of Atlantic sector climate on decadal time scales. *J. Climate,* **19,** 5971–5987.

Lee, T. C. K., F. W. Zwiers, X. Zhang, and M. Tsao, 2006: Evidence of decadal climate prediction skill resulting from changes in anthropogenic forcing. *J. Climate,* **19,** 5305–5318.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.,* **20,** 131–140.

Mantua, N. M., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.,* **78,** 1069–1079.

McCabe, G., M. Palecki, and J. Betancourt, 2004: Pacific and Atlantic influences on multidecadal drought frequency in the US. *Proc. Natl. Acad. Sci.,* **101,** 4136–4141.

Meehl, G. A., and A. Hu, 2006: Megadroughts in the Indian monsoon region and southwest North America and a mechanism for associated multidecadal Pacific sea surface temperature anomalies. *J. Climate,* **19,** 1605–1623.

——, W. M. Washington, W. D. Collins, J. M. Arblaster, A. Hu, L. E. Buja, W. G. Strand, and H. Teng, 2005: How much more global warming and sea level rise? *Science,* **307,** 1769–1772.

——, and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 747–845.

——, J. M. Arblaster, G. Branstator, and H. Van Loon, 2008: A coupled air–sea response mechanism to solar forcing in the Pacific region. *J. Climate,* **21,** 2883–2897.

——, A. Hu, and B. D. Santer, 2009a: The mid-1970s climate shift in the Pacific and the relative roles of forced versus inherent decadal variability. *J. Climate,* **22,** 780–792.

——, J. M. Arblaster, K. Matthes, F. Sassi, and H. van Loon, 2009b: Amplifying the Pacific climate system response to a small 11 year solar cycle forcing. *Science,* **325,** 1114–1118.

Min, S.-K., X. Zhang, and F. W. Zwiers, 2008: Human induced Arctic moistening. *Science,* **320,** 518–520, doi:10.1126/science.1153468.

Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.,* **122,** 73–119.

Murphy, J. M., B. B. B. Booth, M. Collins, G. R. Harris, D. M. H. Sexton, and M. J. Webb, 2007: A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. *Philos. Trans. Roy. Soc. London,* **A365,** 1993–2028.

Palmer, T. N., F. J. Doblas-Reyes, A. Weisheimer, G. J. Shutts, J. Berner, and J. M. Murphy, cited 2009: Towards the Probabilistic Earth-System Model. [Available online at http://aps.arxiv.org/abs/0812.1074.]

Pierce, D. W., T. P. Barnett, N. Schneider, R. Saravanan, D. Dommenget, and M. Latif, 2001: The role of ocean dynamics in producing decadal climate variability in the North Pacific. *Climate Dyn.,* **18,** 51–70, doi:10.1007/s003820100158.

——, ——, R. Tokmakian, A. Semtner, M. Maltrud, J. Lysne, and A. Craig, 2004: The ACPI project, element 1: Initializing a coupled climate model from observed initial conditions. *Climate Change,* **62,** 13–28.

Pohlmann, H., J. H. Jungclaus, A. Köhl, D. Stammer, and J. Marotzke, 2009: Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic. *J. Climate,* **22,** 3926–3938.

Power, S., and R. Colman, 2006: Multi-year predictability in a coupled general circulation model. *Climate Dyn.,* **26,** 247–272, doi:10.1007/s00382-005-0055-y.

——, T. Casey, C. Folland, A. Colman, and V. Mehta, 1999: Interdecadal modulation of the impact of ENSO on Australia. *Climate Dyn.,* **15,** 319–324.

Pulwarty, R., 2003: Climate and water in the West: Science, information and decisionmaking. *Water Resour.,* **124,** 4–12.

Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis,* S. Solomon, et al., Eds., Cambridge University Press, 589–662.

Schneider, E. K., B. Huang, Z. Zhu, D. G. DeWitt, J. L. Kinter, B. Kirtman, and J. Shukla, 1999: Ocean data assimilation, initialization and predictions of ENSO with a coupled GCM. *Mon. Wea. Rev.,* **127,** 1187–1207.

Schulz, M., and Coauthors, 2006: Radiative forcing by aerosols as derived from the AeroCom present-day and pre-industrial simulations. *Atmos. Chem. Phys. Discuss.,* **6,** 5095–5136.

Seager, R., and Coauthors, 2007: Model projections of an imminent transition to a more arid climate in southwestern North America. *Science,* **316,** 1181–1184.

Shindell, D., D. Rind, N. Balachandran, J. Lean, and J. Longergan, 1999: Solar cycle variability, ozone, and climate. *Science,* **284,** 305–308.

Shukla, J., 1984: Predictability of time averages: Part II. The influence of boundary forcings. *Problems and Prospects in Long and Medium Range Weather Forecasting,* D. M. Burridge and E. Kallen, Eds., Springer-Verlag, 155–206.

——, R. Hagedorn, B. Hoskins, J. Kinter, J. Marotzke, M. Miller, T. Palmer, and J. Slingo, 2009: Strategies: Revolution in climate prediction is both necessary and possible: A declaration at the World Modeling Summit for Climate Prediction. *Bull. Amer. Meteor. Soc.,* **90,** 175–178.

Smith, D., S. Cusack, A. Colman, C. Folland, G. Harris, and J. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global circulation model. *Science,* **317,** 796–799.

Soden, B. J., R. T. Wetherald, G. L. Stenchikov, and A. Robock, 2002: Global cooling after the eruption of Mount Pinatubo: A test of climate feedback by water vapour. *Science,* **296,** 727–730.

Stockdale, T. N., 1997: Coupled ocean–atmosphere forecasts in the presence of climate drift. *Mon. Wea. Rev.,* **125,** 809–818.

Stott, P. A., and J. A. Kettleborough, 2002: Origins and estimates of uncertainty in predictions of 21st century temperature rise. *Nature,* **416,** 723–726.

Sugiura, N., T. Awaji, S. Masuda, T. Mochizuki, T. Toyoda, T. Miyama, H. Igarashi, and Y. Ishikawa, 2008: Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. *J. Geophys. Res.,* **113,** C10017, doi:10.1029/2008JC004741.

Sutton, R. W., and D. L. R. Hodson, 2005: Atlantic Ocean forcing of North American and European summer climate. *Science,* **309,** 115–118.

Tang, Y., Z. Deng, X. Zhou, Y. Cheng, and D. Chen, 2008: Interdecadal variation of ENSO predictability in multiple models. *J. Climate,* **21,** 4811–4833.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2008: A summary of the CMIP5 experimental design. CLIVAR, 30 pp. [Available online at www.clivar.org/organization/wgcm/references/Taylor_CMIP5.pdf.]

Tebaldi, C., and R. Knutti, 2007: The use of multi-model ensembles in probabilistic climate projections. *Philos. Trans. Roy. Soc. London,* **A365,** 2053–2076.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.,* **125,** 3297–3319.

Trenberth, K. E., 2008: Observational needs for climate prediction and adaptation. *WMO Bull.,* **57,** 17–21.

——, and J. W. Hurrell, 1994: Decadal atmosphere-ocean variation in the Pacific. *Climate Dyn.,* **9,** 303–319.

van Loon, H., and K. Labitzke, 1998: The global range of the stratospheric decadal wave. *J. Climate,* **11,** 1529–1537.

——, and D. J. Shea, 1999: A probable signal of the 11-year solar cycle in the troposphere of the Northern Hemisphere. *Geophys. Res. Lett.,* **26,** 2893–2896.

——, G. A. Meehl, and D. J. Shea, 2007: Coupled air–sea response to solar forcing in the Pacific region during northern winter. *J. Geophys. Res.,* **112,** D02108, doi:10.1029/2006JD007378.

Vikhliaev, Y., B. Kirtman, and P. Schopf, 2007: Decadal North Pacific bred vectors in a coupled GCM. *J. Climate,* **20,** 5744–5764.

Wetherald, R. T., R. J. Stouffer, and K. W. Dixon, 2001: Committed warming and its implications for climate change. *Geophys. Res. Lett.,* **28,** 1535–1538.

White, W., and D. Cayan, 1998: Quasi-periodicity and global symmetries in interdecadal upper ocean temperature variability, *J. Geophys. Res.,* **103** (C10), 21 335–21 354.

Wigley, T. M. L., 2005: The climate change commitment. *Science,* **307,** 1766–1769, doi:10.1126/science.1103934.

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.

Zhang, X., F. W. Zwiers, G. C. Hegerl, F. H. Lambert, N. P. Gillett, S. Solomon, P. A. Stott, T. Nozawa, 2007: Detection of human influence on twentieth-century precipitation trends. *Nature,* **448,** 461–465, doi:10.1038/nature06025.