

DECADAL CLIMATE PREDICTION

An Update from the Trenches

BY GERALD A. MEEHL, LISA GODDARD, GEORGE BOER, ROBERT BURGMAN, GRANT BRANSTATOR, CHRISTOPHE CASSOU, SUSANNA CORTI, GOKHAN DANABASOGLU, FRANCISCO DOBLAS-REYES, ED HAWKINS, ALICIA KARSPECK, MASAHIDE KIMOTO, ARUN KUMAR, DANIELA MATEI, JULIETTE MIGNOT, RYM MSADEK, ANTONIO NAVARRA, HOLGER POHLMANN, MICHELE RIENECKER, TONY ROSATI, EDWIN SCHNEIDER, DOUG SMITH, ROWAN SUTTON, HAIYAN TENG, GEERT JAN VAN OLDENBORGH, GABRIEL VECCHI, AND STEPHEN YEAGER

The rapidly evolving field of decadal climate prediction, using initialized climate models to produce time-evolving predictions of regional climate, is producing new results for predictions, predictability, and prediction skill.

The importance of improved information about near-term (from 1 year to several decades in advance) regional climate for many societal applications has prompted considerable research in the field of decadal climate prediction that addresses those time scales. Meehl et al. (2009a) outlined the problems and issues involved with decadal climate prediction, and reviewed some of the first results from the few model studies that had been performed up to that time. The purpose of this paper is to provide scientists and possible users of such near-term climate information with an update of research in this rapidly evolving field since the Meehl et al. (2009a) paper, including recent multimodel studies and Coupled Model Intercomparison Project phase 5 (CMIP5) results. CMIP5 includes, as part of the standard multimodel experiments to be run for analysis and comparison, a set of decadal climate prediction experiments—both hindcasts and predictions (Taylor et al. 2012). This multimodel ensemble represents a major contribution to climate science in general and decadal climate prediction in particular in allowing this coordinated set of initialized hindcasts and predictions to be made available to the international climate science community to advance our understanding of the decadal climate prediction problem.

For example, in part due to results from the initialized decadal climate predictions from CMIP5, the assessed range of near-term global warming (2016–2035) was less than the uninitialized simulations in the IPCC AR5 (Kirtman et al. 2013).

To provide an example of a group of users who could conceivably make use of decadal climate predictions, water managers must make decisions regarding water supply and infrastructure on time scales of one to several decades in advance (Barsugli et al. 2009; Means et al. 2010). Regional climate variability and change involving temperature and precipitation are important inputs for those decisions, and it is on those time scales of interest to water managers that decadal climate prediction is being applied as noted above. One question that could be asked is do current decadal climate predictions have sufficient skill or reliability on those time and space scales to be useful in helping water managers make better decisions? A corollary is whether initial conditions are playing a significant role in providing more skill than uninitialized projections.

Decadal climate prediction involves not only the skill of the predictions themselves but also the science questions that can be addressed that should lead to a better understanding of processes and modeling of

climate variability, climate sensitivity, transient climate response, and climate prediction at other time scales, such as seasonal prediction.

The “Terminology” section defines terminology, and the “Technical issues” section addresses a number of technical issues involved with decadal climate prediction such as initialization, ensemble generation, bias adjustment, and prediction verification and evaluation. The “Science issues” section includes a discussion of some of the science issues involved in decadal climate prediction studies focusing on the Indian, Atlantic, and Pacific Oceans, as well as over land areas. The “Initialized decadal climate predictions for near-term climate change” section presents near-term decadal climate predictions that have been made for time periods up to 2035, followed by conclusions in the “Conclusions and prospects for the future of decadal climate prediction” section.

TERMINOLOGY. Vice Admiral Robert Fitzroy, inspired by the loss of a ship in a violent storm in 1859, was the first person to coin a term for anticipating future weather conditions; he chose the word “forecast” (Fitzroy 1863). Later, Lewis Fry Richardson chose the word “prediction” in the title of his book, which first discussed how to produce an estimate of future weather by solving differential equations numerically (Richardson 1922).

Fitzroy and Richardson’s different choice of words to describe the process of determining future weather should be clarified in the realm of estimating the climate for the coming decades. Here the words forecast and prediction are used interchangeably.

In relation to short-term climate, a decadal climate prediction provides information about the future evolution of the statistics of regional climate from the output of a numerical model that has been initialized with observations and run with multiple ensemble members either with a single model or a multimodel ensemble on time scales of 1–30 years. A numerical weather prediction (NWP) or forecast is also generated from a numerical model that has been initialized with observations, but it attempts to track the time evolution of individual weather features typically using multimember ensembles in a probabilistic format on time scales of a week or so. Such predictions or forecasts also can take the form of a synthesis product that might include incorporation of statistical methods, statistical corrections to raw model output, or the combination of different models.

Probabilistic forecasts are thus common on all time scales. The weather man announcing scattered showers for the afternoon only predicts their statistics, not the actual places and times where these occur. The same holds for the ensemble prediction systems that provide probabilistic forecasts up to 10 days in advance and the seasonal climate forecasts now in operational use. The aim is to extend this to the 1- to 30-yr predictions discussed here. Thus, the common elements of initialization with observations and the goal of probabilistic weather and climate information can constitute a seamless or unified approach for weather and climate predictions/forecasts (Palmer et al. 2008; Hurrell et al. 2009).

An “outlook” is a summary of forecast information (e.g., probabilities of ENSO or a short text

AFFILIATIONS: MEEHL, BRANSTATOR, DANABASOGLU, KARSPECK, TENG, AND YEAGER—National Center for Atmospheric Research,* Boulder, Colorado; GODDARD—International Research Institute for Climate and Society, Palisades, New York; BOER—Canadian Centre for Climate Modeling and Analysis, Victoria, British Columbia, Canada; BURGMAN—Florida International University, Miami, Florida; CASSOU—Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique, Toulouse, France; CORTI—European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom, and Institute of Atmospheric Sciences and Climate, Italian National Research Council, Bologna, Italy; DOBLAS-REYES—Catalan Institute of Climate Sciences, Barcelona, Spain; HAWKINS AND SUTTON—National Centre for Atmospheric Science, University of Reading, Reading, United Kingdom; KIMOTO—University of Tokyo, Kashiwa, Japan; KUMAR—Climate Prediction Center, National Oceanographic and Atmospheric Administration/National Centers for Environmental Prediction, College Park, Maryland; MATEI AND POHLMANN—Max Planck Institute for Meteorology, Hamburg, Germany; MIGNOT—Institut Pierre Simon Laplace des Sciences de l’Environnement, Paris, France, and University of Bern, Bern, Switzerland; MSADEK, ROSATI,

AND VECCHI—Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey; NAVARRA—Centro Euromediterraneo sui Cambiamenti Climatici, and Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy; RIENECKER—National Aeronautics and Space Administration Goddard Space Flight Center, Greenbelt, Maryland; SCHNEIDER—George Mason University, Fairfax, Virginia, and COLA, Calverton, Maryland; SMITH—Met Office Hadley Centre, Exeter, United Kingdom; VAN OLDENBORGH—Koninklijk Nederlands Meteorologisch Instituut, De Bilt, Netherlands

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

CORRESPONDING AUTHOR: Gerald A. Meehl, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307
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/BAMS-D-12-00241.1

In final form 14 March 2013
©2014 American Meteorological Society

description), which could include subjective judgments. This is already done, for instance, for the seasonal time scale by the International Research Institute for Climate and Society (IRI) and by the National Oceanic and Atmospheric Administration (NOAA)'s Climate Prediction Center, which provides a seasonal climate outlook. The term “projection” indicates an estimate of future climate that is dependent on the externally forced climate response (e.g., the response to changes in anthropogenic greenhouse gases or aerosols) inherent in a particular emission scenario.

“Predictability” characterizes the theoretical limit of predictive skill in optimum conditions. It is the “ability to be predicted” rather than the current “ability to predict” some feature or quantity. Predictability arises from both externally forced and internally generated variability. Estimates of climate predictability are mainly, but not exclusively, based on the behavior of climate models and regions where skillful predictions on the time scales of interest may be possible as discussed in the “Predictability and forecast skill” section. Finally, the terms “hindcast” and “retrospective forecast” are perhaps more problematic. The ocean reanalysis community refers to hindcast for an atmosphere-forced ocean-only simulation. But this term is also used in the literature for initialized predictions of past cases. In the decadal climate prediction context, retrospective prediction, retrospective forecast, and hindcast are used interchangeably to refer to this type of initialized forecast of past cases.

TECHNICAL ISSUES. *Initialization and ensemble generation.* Beyond reducing the model biases themselves, one of the biggest technical challenges affecting the quality of decadal climate predictions is the initialization of the model from observations to start a decadal prediction. Modeling groups are actively exploring different techniques and methodologies for initializing decadal climate predictions. The main features of the initialization procedures used by modeling groups participating in CMIP5 are summarized in Table 1, with a list of acronyms given in Table 2. These include either partial or fully coupled assimilation of ocean and/or atmospheric observations, forcing the ocean with atmospheric observations, and, additionally, full-field or anomaly initialization. Evaluations of two of these methods, the three-dimensional initialization of the ocean with observations and the use of observed surface forcing to initialize the ocean (Matei et al. 2012c; Yeager et al. 2012; Meehl and Teng 2012; Swingedouw et al. 2012),

have shown that the latter may constitute a simple but successful alternative strategy for initialization, especially over the extratropical regions (Doblas-Reyes et al. 2011).

There are also various methods to deal with model drift away from the observed initial state. Full-field initialization brings the ocean model state close to observations and the model then drifts toward its systematic error state during the prediction and requires bias adjustment in predictions (discussed below). Anomaly initialization adds the anomalous component of the observed state to the model climatology to minimize the drift during the prediction. A comparison of the two methods shows that the former generally produces more skillful predictions on the seasonal time scale (Magnusson et al. 2012), though the latter provides more skillful predictions in hindcasts where it has been tested (Smith et al. 2012a). Hazeleger et al. (2013b, manuscript submitted to *Geophys. Res. Lett.*) find there is no significant difference between full-field and anomaly initialization in decadal prediction skill. However, the anomaly initialization method can produce mismatches between the observational anomalies and the model climatology in some regions (e.g., in sharp Gulf Stream gradient locations). A majority of modeling groups at present are leaning toward full-field initialization, but further evaluations are necessary with more models to draw definitive conclusions as to the best initialization technique.

Weather and climate predictions are well known to be sensitive to small perturbations in the initial state (e.g., Du et al. 2012), and an ensemble of initial conditions and subsequent predictions is typically generated in order to investigate this. Ensemble generation can utilize the different methodologies for perturbation, extending from ensemble Kalman filter-type assimilation methods (e.g., Karspeck et al. 2013), different start days for the initial state of the atmosphere around the time of the prediction (e.g., Yeager et al. 2012), application of breeding methods to generate optimal initial perturbations (e.g., Ham et al. 2014), or variations on those methods. This diversity of approaches results from different philosophical approaches to the initialization and ensemble generation problem.

Bias adjustment. Biases in decadal climate predictions that develop as a function of time come from a range of sources, including model drift from the observed initial state to its own preferred state, which is a product mainly of its own set of systematic errors that can happen quite rapidly (sometimes called

TABLE 1. (right) Brief summaries of initialization strategies used by (left) modeling groups for the CMIP5 decadal prediction experiments. Acronyms in this table are defined in Table 2.

Model/modeling center	Initialization method
BCC-CMI.I BCC, China	Full-field initialization. Coupled model integration with ocean <i>T</i> nudged to SODA ocean reanalysis product above 1500 m.
CanCM4 CCCma, Canada	Full-field initialization. Atmosphere: assimilate ERA-40 and Interim ECMWF Re-Analysis (ERA-Interim) with nudging to observed SSTs. Ocean: offline assimilation of SODA and GODAS subsurface ocean <i>T</i> . <i>S</i> adjusted to preserve model <i>T</i> – <i>S</i> relationship
CCSM4 NCAR, United States	Full-field initialization: (hd-1) Ocean hindcast forced with CORE2 atmospheric dataset. (da-2) Loosely coupled ocean–atmosphere Ensemble Kalman Filter (EnKF, see Table 2) analysis. Atmosphere: assimilates raw atmospheric observations, forced with observed SST dataset. Ocean: assimilates observations of subsurface <i>T</i> and <i>S</i> , forced with atmospheric EnKF analysis.
CFSv2–2011 NCEP, United States	Full-field initialization. Coupled (atmosphere, ocean, ice, land) three-dimensional variational data assimilation (3DVAR) using the NCEP CFSR. Atmosphere constrained by raw observations, ocean constrained by observed <i>T</i> and <i>S</i> , sea ice constrained by a combination of gridded and satellite-derived products. Additional nudging of ocean surface temperatures to SST reanalysis products (HADISST and Reynolds SST).
CFSv2–2011 COLA, United States	Full-field initialization. Ocean initialized using the NEMOVAR ocean reanalysis interpolated to the ocean model grid. Atmosphere, sea ice, and land initialized from CFSR reanalysis.
CMCC-CM CMCC, Italy	Full-field initialization. Atmosphere: uninitialized twentieth-century coupled model simulations. Ocean: three realizations of CMCC-INGV ocean synthesis of <i>T</i> and <i>S</i> .
CNRM-CM5 CNRM-CERFACS (France)	Full-field initialization. Coupled model integration with ocean <i>T</i> and <i>S</i> nudged to NEMOVAR ocean reanalysis product (NEMOVAR is a multivariate 3DVAR data assimilation of <i>T</i> and <i>S</i> observations into the NEMO ocean model).
EC-Earth (consortium)	Full-field initialization. Atmosphere and land: initialized from ERA-40 and ERA-Interim. Ocean: NEMOVAR ocean reanalysis product. Sea ice initial conditions come from a LIM2 run forced with the DFS4.3 atmospheric forcing.
FGOALS-g2 LASG-CESS, China FGOALS-s2 LASG-IAP, China	Full-field initialization. Atmosphere/land: none. Ocean: nudging to SST, <i>T</i> , and <i>S</i> with dynamic bias correction. Sea ice: none. Ensembles: perturbed atmosphere/ocean/land/sea ice with different initial time.
GEOS-5 NASA-GMAO, United States	Full-field initialization. Coupled assimilation with atmosphere constrained by MERRA reanalyses and observed precipitation. Ensemble OI used in ocean to assimilate observed subsurface <i>T</i> and <i>S</i> and surface height, temperature, and salinity. Ensembles: two-sided breeding method used to generate optimal initial perturbations
GFDL-CM2.1 NOAA-GFDL (United States)	Full-field initialization. Coupled assimilation using EnKF. Atmosphere: constrained by an atmospheric reanalysis. Ocean: assimilates observations of <i>T</i> , <i>S</i> , and SST
HadCM3 Met Office Hadley Centre, United Kingdom	(i1) Anomaly initialization: coupled integration with relaxation to atmospheric anomalies and HadCM3 climatology from ERA-40 and operational ECMWF analysis, and ocean anomalies and HadCM3 climatology from an offline anomaly ocean reanalysis product. (i2) Full-field initialization: as in anomaly initialization in (i1) above, but with coupled integration relaxed to full values rather than anomalies.
IPSL-CM5A-LR IPSL (France)	Anomaly initialization. Ensemble of coupled integrations with nudging toward anomalous Extended Reconstructed Sea Surface Temperature (ERSST) data (www.esrl.noaa.gov/psd/) (nudging strength: $-40 \text{ W m}^{-2} \text{ K}^{-1}$).
MIROC4h, MIROC5 MIROC, Japan	Anomaly initialization. Coupled integration using observation-based gridded ocean <i>T</i> and <i>S</i> dataset.
MPI-ESM-LR, MPI-ESM-MR MPI-M, Germany	Anomaly initialization. Coupled integration using ocean <i>T</i> and <i>S</i> fields from an offline ocean hindcast forced with NCEP atmospheric reanalysis.
MRI-CGCM3, MRI, Japan	Anomaly initialization. Coupled integration using observation-based gridded ocean <i>T</i> and <i>S</i> dataset.

TABLE 2. Definition of acronyms used in Table 1.	
3DVAR	Three-dimensional variational data assimilation
BCC-CM	Beijing Climate Center Climate Model
CanCM4	Canadian Coupled Global Climate Model version 4
CCCma	Canadian Centre for Climate Modeling and Analysis
CCSM4	Community Climate System Model version 4
CFSR	Climate Forecast System Reanalysis
CFSv2	Climate Forecast System version 4
CMCC	Centro Euro-Mediterraneo per i Cambiamenti Climatici
CMCC-CM	CMCC Climate Model
CMCC-INGV	CMCC Istituto Nazionale di Geofisica e Vulcanologia
CNRM-CERFACS	Centre National de Recherches Météorologiques-European Centre for Research and Advanced Training in Scientific Computation
CNRM-CM5	CNRM Climate Model version 5
COLA	Center for Ocean–Land–Atmosphere
CORE2	Coordinated Ocean–Ice Reference Experiments, phase 2
DFS4.3	Drakkar Forcing Set version v.3
EC-Earth	European Earth System Model
EnKF	Ensemble Kalman filter
ERA-40	40-yr ECMWF Re-Analysis
FGOALS	Flexible Global Ocean–Atmosphere–Land System model [also FGOALS g2 (s2)]
GEOS-5	Goddard Earth Observing System version 5
GFDL	Geophysical Fluid Dynamics Laboratory
GFDL-CM2.1	GFDL Climate Model version 2.1
GODAS	Global Ocean Data Assimilation System
HadCM3	Hadley Centre Climate Model version 3
HADISST	Hadley Centre Sea Ice and Sea Surface Temperature dataset
IPSL	Institute Pierre Simon Laplace
IPSL-CM5A-LR	IPSL Climate Model version 5A, Low Resolution
LASG-IAP (CESS)	State Key Laboratory Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics–Institute of Atmospheric Physics (Center of Earth System Sciences)
LIM2	Louvain-la-Neuve Sea Ice Model version 2
MERRA	Modern-Era Retrospective Analysis for Research and Applications
MIROC	Model for Interdisciplinary Research on Climate
MPI-M	Max-Planck-Institut für Meteorologie
MPI-ESM-LR/MR	MPI Earth System Model Low Resolution/Medium Resolution
MRI	Meteorological Research Institute
MRI-CGCM3	MRI Coupled Global Climate Model version 3
NASA-GMAO	National Aeronautics and Space Administration Global Modeling and Assimilation Office
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction
NEMO	Nucleus For European Modeling of the Ocean
NEMOVAR	Variational data assimilation system for NEMO
NOAA	National Oceanic and Atmospheric Administration
OI	Optimal interpolation
SODA	Simple Ocean Data Assimilation

initialization shock), inability to realistically simulate the natural modes of interannual-to-multidecadal variability, uncertain future levels of radiative forcings (such as volcanic eruptions and aerosols), and insufficient and imperfect observations. Additionally, there is uncertainty from insufficient sampling of the natural variability owing to both the short history of the hindcasts as well as the limited number of predictions. The rate and spatial pattern of bias growth may give useful information about the physical processes that lead to prediction error and may allow targeted model improvements. Nevertheless, current predictions generally must attempt to remove this bias in order to be useful in predicting small signals.

Some of the issues in adjusting prediction biases are given by WCRP (2011), which discusses the removal of the mean bias from the predictions. The result for full-field bias adjustment (see Smith et al., 2013, for discussion of bias adjustment for anomaly initialization versus full-field initialization) is equivalent to calculating a predicted climatological average for each forecast range and considering the forecast anomalies obtained by subtracting the average. The same calculation is done for the observations, and the anomalies are compared. Mean bias adjustment does not address issues such as potential trends (time dependence) in the drift/bias. Bias adjustment can be illustrated in a schematic of a set of decadal predictions (Fig. 1). In this example, which is more comparable to the case for full-field initialization, the model drifts from its observed initial states (dashed lines) toward its preferred climate state, which is closer to the uninitialized model state (gray line) with a stronger trend than is observed. Systematic model errors may be removed by subtracting the average rate of drift over all hindcasts (inset). In this example, however, removal of the mean bias produces states that remain biased low early in the period and biased high later in the period, indicating the bias adjustments are too small early on, and too large later (solid line). Replacing the average drift correction with a correction that varies over the period may be needed (van Oldenborgh et al. 2012). In other words, an “average” bias adjustment is an issue if the model drift has a substantial trend. Additional corrections may be possible for conditional biases related to different estimated magnitudes of modeled and observed variability (Goddard et al. 2012a). A further complication is that the character of the drift in initialized predictions can depend on the observing system or the particular initial state (e.g., Kumar et al. 2012; Vecchi et al. 2013), which can lead to changes in observing systems being imprinted on the bias adjustments and thus on the predictions.

Mean bias also can be removed by calculating a model climatological average for each hindcast time period, creating anomalies by subtracting that model climatology for each time period from the model hindcasts from those same time periods, and doing the same calculation for the observations (e.g., García-Serrano and Doblas-Reyes 2012, and applied in Figs. 4 and 8 below).

Assessing the required number of start dates, the number of ensemble members, and suitable adjustment methodologies to enable a reliable estimate of the bias is important in order to improve the decadal predictions. Because of issues involved with sampling model variability, trend, and conditional bias as noted above, more robust estimates of the bias adjustment are possible with more start dates for the hindcasts. This can involve start dates every year, rather than every 5 years as originally planned for CMIP5 (García-Serrano and Doblas-Reyes 2012), and this has become a recommendation as part of CMIP5. For the 30-yr predictions for CMIP5, where there are only three start dates for the hindcasts, correcting the bias presents an even greater challenge. One method is to use the year 10 bias adjustment for years 11–30, assuming most of the drift occurs by year 10 (Meehl and Teng 2012, 2014). There are additional complications for precipitation compared to temperature (Doblas-Reyes et al. 2013). For example, the more noisy spatial character of precipitation likely requires some spatial averaging (Goddard et al. 2012a) and presents even greater challenges for skillful decadal climate predictions at local scales.

A side benefit of dealing with the technical issues of removing model bias is a method to help quantify the observed transient climate response (TCR) and equilibrium climate sensitivity (ECS) to increasing CO₂. Hawkins et al. (2014) show that after all the temperature biases mentioned above are removed from model hindcasts, what is left is the model bias that is due to the transient response to increasing greenhouse gases (GHGs) compared to uninitialized runs. This is effectively the conditional bias, at least for temperature, where the bulk of that bias is in the trend. Utilizing different versions of the same set of predictions with models with a range of climate sensitivities allows a constraint on the uncertainties in the observed TCR and ECS.

Additionally, it has been shown that the bias adjustment procedure can correct for model trends that may differ from observations. For example, bias adjustments for the CMIP5 multimodel dataset can reduce a greater-than-observed decadal trend in the models, and thus improve hindcast skill in both

initialized and uninitialized simulations (Meehl and Teng 2014). Since the magnitude of the bias adjustment can be large relative to the signals predicted (Kharin et al. 2012; Kim et al. 2012), it is an important goal to reduce the systematic errors of the models to minimize the initialization shock. An analog can be found in the early days of global coupled climate models that used “flux correction” to account for egregious model systematic errors—for example, a very weak or nonexistent Atlantic meridional overturning circulation (AMOC) (Manabe and Stouffer 1988). However, even with such large model errors being corrected in that way, relevant climate change information from that generation of models was obtained as anomalies from control or reference runs (e.g., Cubasch et al. 2001). Such results have been consistent with later generations of models that were improved and consequently made less use of flux correction (Meehl et al. 2007), and none of the present generation of atmosphere–ocean general circulation models (AOGCMs) in CMIP5 (Taylor et al. 2012) use flux correction. Presumably the need for bias adjustment for decadal climate prediction will be reduced as subsequent generations of models continue to improve.

Predictability and forecast skill. Evaluating the skill of hindcasts is important for quantifying the spatial–temporal credibility of predictions as well as providing a lower bound for the predictability of the system (e.g., Goddard et al. 2012a; Wang et al. 2012). Boer et al. (2013) use hindcasts to evaluate both predictability and skill in a forecast system. Predictability may be quantified in a variety of ways including correlation or mean square error (e.g., Goddard et al. 2012a). The spread of states in prediction ensembles can also be used to determine for how long and by how much the predicted probability density function (PDF) is distinguishable from the corresponding climatological PDF, thus providing an estimate of predictability (e.g., Branstator and Teng 2010). Some investigators have used relative entropy from information theory (Kleeman 2002; Majda et al. 2005; DelSole 2004) as a comprehensive measure of predictability, though this method works best with large ensembles or large numbers of hindcasts.

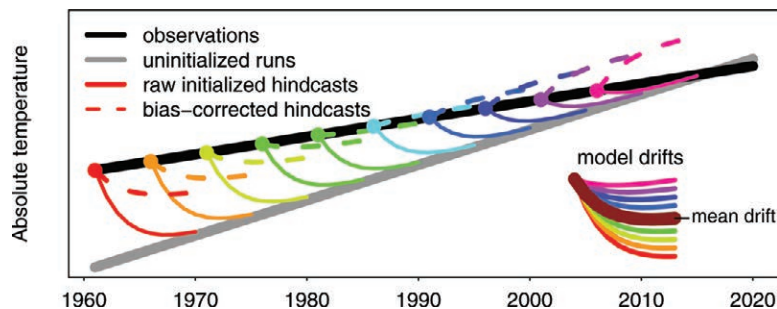


FIG. 1. Schematic figure illustrating the potential difficulty of bias adjustment when there is a long-term background trend that differs from the observed, together with short-term model drift following initialization. The drift of the raw initialized hindcasts from the initialized “observations” (solid black line) are the thin colored solid lines that approach the uninitialized model-forced simulation (solid gray line). In the inset, the mean drift (heavy solid brown line) is calculated as the average of the drifts from each initial state (colored lines). If this mean drift is removed from each hindcast, the resulting dashed colored lines drift negative early in the hindcast period and positive later in the hindcast period as a consequence of the difference in the observed and simulated trends. This is overcome by using drift corrections that depend on the initial time of the forecast rather than the average drift (Kharin et al. 2012).

Hawkins and Sutton (2009a) schematically suggested that as a prediction progresses, initial condition predictability would become relatively unimportant as the spread from past and projected changes in anthropogenic external forcing increased, with some crossover point reached at around a decade. Subsequent studies have now quantified the timing of that crossover point for some regions and variables. For example, Branstator and Teng (2010, 2012) use relative entropy to measure the predictability of upper-ocean heat content in the North Atlantic (Fig. 2). Early in the predictions there is the potential for the initial state to have a large impact, but that steadily decreases. The predictive information due to past and ongoing changes in greenhouse gases is lower at first but increases with forecast range until, after about 8 years, the crossover point is reached and there is greater potential skill provided by the external forcing. These results are in broad agreement with indirect estimates of the predictability of the relative contributions of internally generated and forced components of temperature as characterized by variance ratios in CMIP3 experiments (Boer 2011) and in measures of predictability and prediction skill for hindcasts made with the CCCma model (Boer et al. 2013).

This raises one key difference between numerical weather prediction and decadal climate prediction—namely, the much smaller number of independent initial states in the latter from which to sample in order to assess decadal climate prediction skill. As

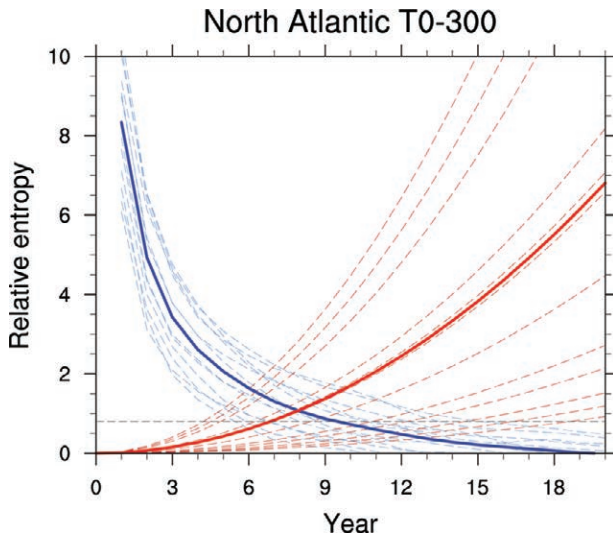


FIG. 2. Predictability of upper 300-m temperature of the North Atlantic for 12 CMIP5 models resulting from initialization (dashed blue lines) and the response to RCP4.5 forcing (dashed red lines). Here, predictability is indicated by the information contained in perfectly predicted distributions, as measured by relative entropy. The thick solid lines are multimodel averages of the relative entropy. The crossover point near year 8 for the multimodel averages is where information originating in the initial conditions (blue line) begins to be exceeded by information resulting from external forcing (red line). The horizontal dashed line indicates the 90th percentile of relative entropy from ensembles of approximately 18 random states drawn from a control run. (From Branstator and Teng 2012.)

noted above, modeling groups have used initial states starting in 1960 and run either every 5 years or every year thereafter. The question of how many independent initial states would be needed to produce statistical confidence has yet to be definitively answered. Such studies need to be performed to define the limits of a properly validated forecast system. However, as will be discussed later, measures of reliability suggest that even with the limited number of initial states available, there may be some regions that produce reliable (i.e., the probabilities of the hindcasts for a specific event match the relative observed frequencies) climate information (Corti et al. 2012), although this needs to be explored in other models (e.g., Ho et al. 2013).

Statistical methods have also been applied to quantify predictability. These include linear inverse modeling (Penland 1989; Newman 2007, 2013; Hawkins and Sutton 2009b; Teng and Branstator 2011; Hawkins et al. 2011; Zanna 2012), multivariate regression propagators (DelSole and Tippett 2009; Branstator et al. 2012), analogs (Branstator et al. 2012;

Ho et al. 2012), and coarse-grain clusters. Each has been used to study various aspects of predictability that can provide measures of possible predictive skill. For example, Branstator et al. (2012) and Branstator and Teng (2012) used two of these methods to compare the predictability properties of various comprehensive AOGCMs and found average saturation times (i.e., the time at which initialized predictions cannot be discerned from randomly generated variability) for subsurface temperature had substantial model-to-model variations, especially in the North Atlantic. In that region saturation occurred in less than 6 years in some models and after more than 15 years in others (Fig. 2), thus highlighting the extent to which estimates of North Atlantic predictability can be model dependent. It is also likely that statistical predictions can be used as a benchmark to assess the skill of dynamical predictions (Hoerling et al. 2011; Smith et al. 2012b; Ho et al. 2012; DelSole et al. 2013).

Techniques for evaluation of skill of decadal predictions can benefit from the extensive experience gained from the much longer history of seasonal predictions and their verifications. Seasonal predictions share features that are common to their decadal counterpart (Goddard et al. 2012b): a prediction format that is probabilistic, a need for prediction calibration, and a need to convey estimates of skill to the users as the skill varies strongly with region and season.

An essential measure for the probabilistic predictions is their reliability. The reliability of a probabilistic prediction, though not necessarily a measure of accuracy, refers to a comparison of prediction probability of an event against the observed frequency. For reliable probabilistic prediction, the prediction probability for an event to happen should be the same as the observed frequency of occurrence (Kumar 2007). For example, if an event such as precipitation at a particular location is predicted to be above the climatological average with a 60% chance, then, in this definition of a reliable prediction, for 60% of cases the precipitation for the verifying analysis will be above its mean value. Unreliable probabilistic predictions have severe implications on the efficacy of economic decision-making processes (Vizard et al. 2005). Decadal prediction reliability will be discussed further in the “Predictive skill over land” section.

The U.S. Climate Variability and Predictability (CLIVAR) Decadal Prediction Working Group (DPWG) recommends evaluating decadal predictions for different time and space averages (Goddard et al. 2012a). Though initialized decadal climate predictions can be run for time scales of 1–30 years

(Taylor et al. 2012), for time averages the recommendation is to evaluate prediction skill for 2–5- and 6–9-yr lead-time averages to represent interannual time scales (but not to include year 1 so as to exclude skill from the seasonal time scale) and for 2–9-yr lead-time averages to assess subdecadal-scale predictions. For spatial scales, the recommendation is to provide evaluations for at least the observational grid scale and the regional scale. The choice for the spatial extent of the regional scale is guided by a compromise between the correlation skill and spatial signal-to-noise ratios. Such an analysis suggests that at least 5° latitude by 5° longitude represents a reasonable scale for smoothing precipitation and temperature (Räisänen and Ylhäisi 2011). Statistical significance of skill depends crucially on the length of the verification time series and autocorrelation within the period, with robust estimates requiring verification over a large sample or long time series (Kumar 2009).

Decadal hindcast skill is frequently compared with uninitialized climate change simulations. Both are intended to capture the forced response to changing atmospheric composition, but only the initialized decadal hindcasts carry potential information on the time evolution of internally generated climate variability. Another approach is to attempt to remove the forced component from the hindcasts and the observations and to consider the skill of the remaining variability (van Oldenborgh et al. 2012). However, the identification of the forced component is difficult, especially in the observations, and is usually approximated by fitting to some curve.

It has been noted in earlier studies that for other climate model applications a multimodel ensemble outperforms most single model results (e.g., Reichler and Kim 2008). This characteristic seems also to apply to most initialized decadal climate hindcasts as has been shown in various ways, for example, by Chikamoto et al. (2012a), Kim et al. (2012), and Smith et al. (2012b).

SCIENCE ISSUES. To advance the science of decadal climate prediction, there have been several coordinated climate modeling exercises such as the Ensembles-Based Predictions of Climate Changes and their Impacts Project (ENSEMBLES; van der Linden and Mitchell 2009; van Oldenborgh et al. 2012; García-Serrano and Doblas-Reyes 2012) and the most recent CMIP5 (Taylor et al. 2012) as noted above. Most of the decadal climate prediction experiments in CMIP5 are hindcasts designed to assess historical predictive skill. For the near-term future decadal climate predictions, the current representative concentration

pathway (RCP) scenarios (Moss et al. 2010) used in climate model simulations of future climate change in CMIP5 do not diverge much in terms of globally averaged climate response until nearly 2035 (e.g., Meehl et al. 2012). Therefore, the CMIP5 decadal predictions used the RCP4.5 scenario. However, the observed time evolution of aerosols and short-lived species could provide some uncertainty not captured in the RCP scenarios (Moss et al. 2010) since the RCP scenarios all provide estimates of aerosol removal that could be different from recent observations (Shindell et al. 2012, though there are indications that recent observed sulfate aerosol concentrations are comparable to those in the RCP scenarios (Klimont et al., 2013). Thus, there is the possibility that near-term climate predictions in CMIP5 could have aerosols in RCP4.5 that are not necessarily consistent with recent observations in some regions.

There is also evidence that aerosol changes could influence Atlantic hurricane activity in the coming decades, with aggressive mitigation in RCP2.6 leading to increased storm frequency (Villarini and Vecchi 2012, 2013; Dunstone et al. 2013). With regards to other external forcings, some groups include a climatological solar cycle in spite of issues with its predictability, and it is acknowledged that volcanoes are inherently unpredictable so they are not included in the future predictions.

Another science issue involved with decadal climate prediction is the possible conditional nature of the prediction skill; that is, some initial states could be more predictable and thus lead to more accurate predictions than others (Griffies and Bryan 1997; Collins et al. 2006; Branstator and Teng 2010). Reliable estimates of skill conditioned on the occurrence of specific circumstances in the initial states are even more difficult to determine owing to the smaller sample for verification.

Distinguishing between skill arising from external factors and internal variability is not always possible (Solomon et al. 2011). Decadal climate variability combines elements of stochastic forcing and internally generated mechanisms such as the AMOC (Latif and Keenlyside 2011; Srokosz et al. 2012; Liu 2012). Initialized predictions that aim to capture the evolution of internal variability may be hindered by the aforementioned initialization shocks. For surface temperature, some of the near-term predictive skill arises from the warming trend associated with increases in anthropogenic greenhouse gases (Smith et al. 2010; van Oldenborgh et al. 2012) as was shown to be the case for North Atlantic upper-ocean heat content in Fig. 2.

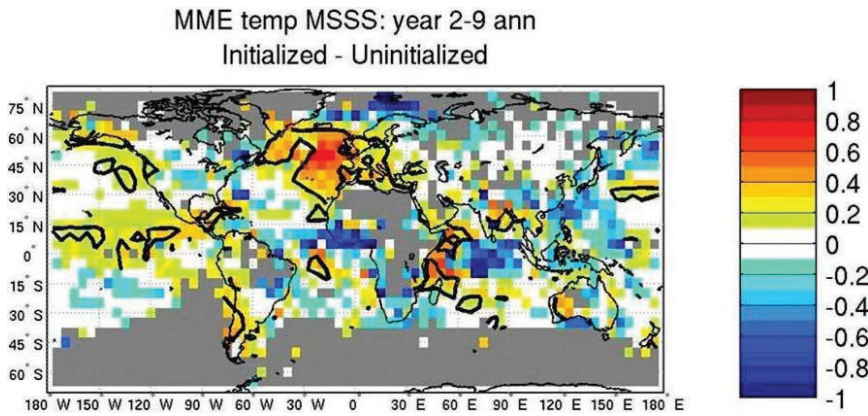


FIG. 3. Mean squared skill score (MSSS) differences for decadal temperature hindcasts from a 12-member multimodel ensemble from CMIP5, for the initialized hindcasts (“forecasts”) minus the uninitialized hindcasts (“reference”) as predictions of the observed climate. The forecast target is years 2–9 following the initialization every 5 years from 1961 to 2006 (i.e., 10 hindcasts for each model). The verification takes place at the grid scale of $5^\circ \times 5^\circ$ of the observations. Contour line indicates statistical significance that the MSSS is positive at the 95% confidence level, with areas in yellow and orange indicating enhanced predictive skill from initialization (after Goddard et al. 2012a).

To understand sources of regional predictive skill, analysis of single model results and multimodel experiments has recently pointed to possible sources of predictive skill arising from either internally generated decadal variability related to physical mechanisms, external forcing, or a combination of both (Fyfe et al. 2011). A multimodel example showing regions where there is additional prediction skill from initialization on the subdecadal time scale over and above uninitialized simulations is shown in Fig. 3. Additional skill coming from initialization has a regional pattern, with significantly improved skill compared to uninitialized simulations over areas of the North Atlantic and eastern Pacific.

Another multimodel example of predictive skill for years 6–9 at the subdecadal time scale (Fig. 4 after Doblas-Reyes et al. 2013) shows regions where overall predictive skill for temperature (with contributions from initialization and external forcing, including volcanoes) from the CMIP5 multimodel ensemble has

(2012) for years 2–5 and 6–9, and Guémas et al. (2013) for years 2–5 and 6–9. Thus, the pattern in Fig. 4 seems relatively robust for the current generation of multimodel decadal climate hindcasts, and the main question that is now being addressed is why this pattern arises (i.e., internally generated mechanisms captured by initialization or products of external forcing). We address this below for each ocean basin and then for land areas.

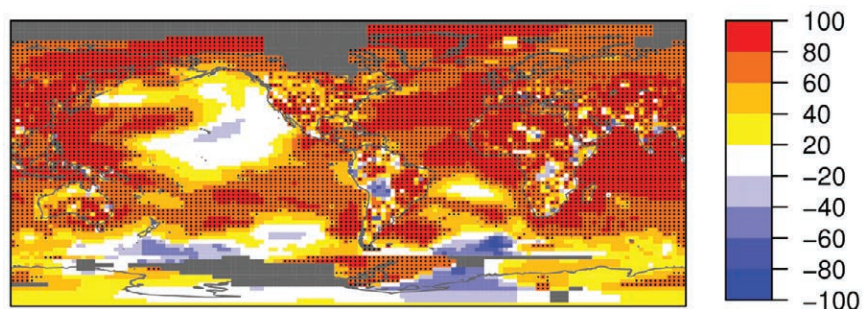


FIG. 4. Surface air temperature predictive skill (correlation with observations), predictions for years 6–9 averages based on CMIP5 multimodel ensemble mean hindcasts (see Table 1 for details). Results are from initialized hindcasts with 5-yr intervals between start dates from 1960 to 2005. Correlations are calculated by averaging all the hindcasts from all the models for each start date for the 6–9-yr hindcasts, subtracting the model climatologies to form anomalies, then taking the time series of those anomalies for each start date at each grid point and correlating corresponding gridpoint time series from the observations formulated in the same way as the model hindcast anomalies, and those correlations are plotted at each grid point to form the map. Stippling indicates significance at the 95% level, and darker red colors with stippling indicate areas with more significant predictive skill (Doblas-Reyes et al. 2013).

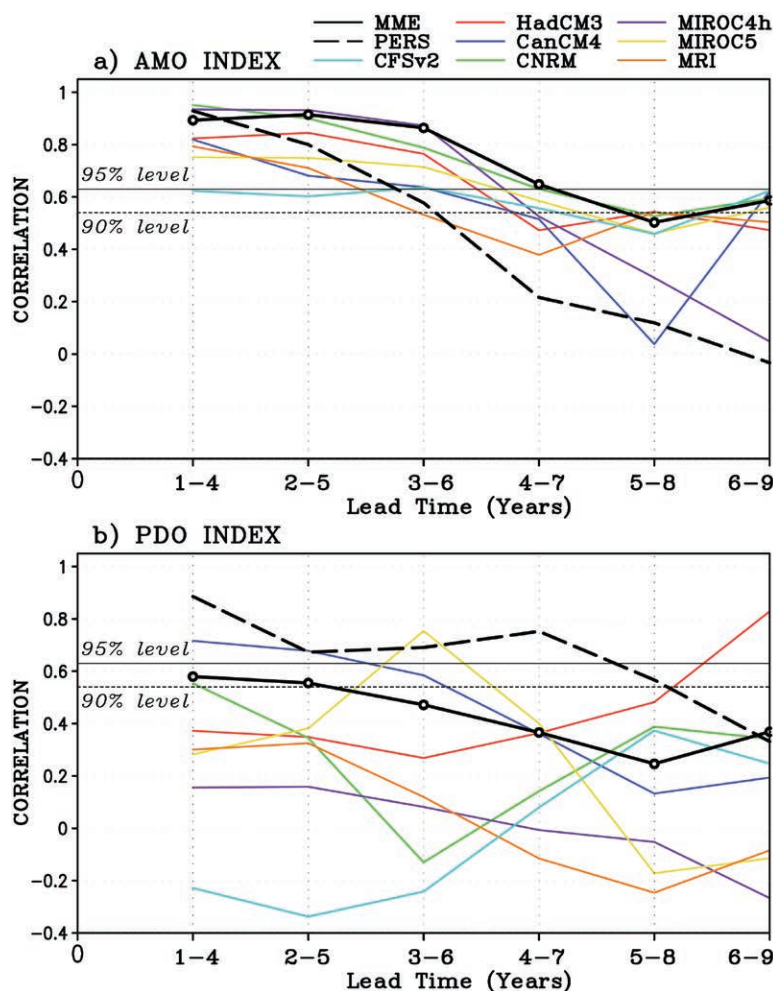
been quantified. There are indications of greater skill (relative to other areas in the initialized hindcasts, not necessarily relative to uninitialized simulations) over the North Atlantic, western Pacific, eastern tropical Pacific, and Indian Oceans (darker red colors), with less skill over parts of the North Pacific (lighter colors) as also noted by other multimodel studies for different prediction time frames (Goddard et al. 2012a; see also <http://clivar-dpwwg.iri.columbia.edu>): Kim et al. (2012) for years 2–5, Chikamoto et al. (2012a) for years 2–4 and 5–9, van Oldenborgh et al.

Predictive skill in the Indian Ocean. The Indian Ocean area repeatedly stands out as the region with the highest surface temperature prediction skill worldwide in state-of-the-art decadal climate prediction studies (e.g., Fig. 4). This skill has so far been shown to be in large part due to the externally forced trend from increasing GHGs, which has been greater than the internally generated decadal climate variability (Ho et al. 2012; Guémas et al. 2013a). Therefore, such predictive skill due to external forcing is also present in uninitialized projections (Goddard et al. 2012a) as shown by the comparison between forced and internally generated components (Boer 2011) and uninitialized and initialized historical simulations (Guémas et al. 2013). For the studies performed so far, it is likely that predictive skill for near-term climate change in the Indian Ocean region is mainly due to the externally forced response, and decadal predictions are quite reliable there (Corti et al. 2012). However, decadal variability of regional SST patterns in the Indian Ocean has been shown to modulate interannual variability across the entire Indo-Pacific region (Meehl and Arblaster 2011, 2012). Subsequent studies need to be performed to assess these aspects related to initialization in the Indian Ocean that could improve predictions of in-

terannual variability with significant regional impacts involving the Asian–Australian monsoon.

Predictive skill in the Atlantic Ocean. A number of studies find that initialization improves the predictive skill of temperature in the North Atlantic (e.g., Smith et al. 2010; van Oldenborgh et al. 2012; Pohlmann et al. 2009; Keenlyside et al. 2008; Matei et al. 2012c; Doblas-Reyes et al. 2013; Yang et al. 2013; A. Rosati et al. 2014, unpublished manuscript; R. Msadek et al. 2014, unpublished manuscript; Ho et al. 2012; Ham et al. 2014; Hazeleger et al. 2013a). Enhanced skill from initialization in a multimodel context for surface temperature at the subdecadal time scale (Fig. 4) shows improved skill in the North Atlantic region, and this is expected to be at least partially related to skillful predictions of the AMOC coming from initialization (Delworth et al. 2007; Knight et al. 2005; Swingedouw et al. 2012). The reasoning behind this is that there is likely a connection between the AMOC and surface temperature associated with the Atlantic multidecadal oscillation [AMO; sometimes

FIG. 5. Correlation skill for (a) the AMO index and (b) the PDO index, computed from seven models from the CMIP5 multimodel ensemble. Correlations are calculated from the time series obtained by averaging hindcasts and model results for various prediction times (e.g., 1–4 years, 2–5 years, etc.). The corresponding observational anomalies for each of those hindcast periods are formed into a comparable time series, and then the correlation of the two time series is calculated for each prediction time period and plotted in the two panels. High correlations indicate better predictive skill, and are generally better for prediction time periods closer to the initial time (e.g., predictions for years 1–4 are generally better than for those for years 6–9). Each model is represented by a colored line, the multimodel ensemble is the solid black line, and persistence (taking the initial state and persisting it in time) is the dashed black line. The solid and dashed horizontal lines represent significance at the 95% and 90% levels, respectively (Kim et al. 2012).



more generically referred to as Atlantic multidecadal variability (AMV)] to produce decadal climate prediction skill at time scales less than 10 years for North Atlantic SSTs associated with the AMO (Fig. 5) and, thus, for consequent atmospheric variability in the northern North Atlantic region. For example, Gastineau and Frankignoul (2012) showed that in six climate models, an intensification of the AMOC is followed by a weak sea level pressure response that resembles a negative phase of the North Atlantic Oscillation (NAO). However, current models do not yet show any skill in decadal NAO forecasts.

As noted above, accurate initialization of the AMOC is likely key for extending the predictive skill of North Atlantic SST and upper-ocean heat content up to a decade ahead beyond the skill of persistence (Latif and Keenlyside 2011; Srokosz et al. 2012; Matei et al. 2012c; Yeager et al. 2012; Robson et al. 2012a,b;

R. Msadek et al. 2014, unpublished manuscript). However, assessing skill in predicting the AMOC is difficult because of the strong seasonal cycle and the relatively short dataset for evaluation. The recent discourse between Matei et al. (2012a,b) and Vecchi et al. (2012) highlights this challenge. Matei et al. (2012a,b) reported multiyear predictive skill in the initialized decadal predictions of the AMOC when assessed against RAPID observations. However, Vecchi et al. (2012) argue that these findings are sensitive to the choice of validation metrics and the treatment of the AMOC seasonal cycle.

Consistent with observations of the NAO, Labrador Sea convection, and North Atlantic subpolar gyre strength, recent multimodel ocean analyses suggest that the AMOC at 45°N increased from the 1960s to the mid-1990s, and decreased thereafter (Pohlmann et al. 2013). Aspects of this multimodel AMOC behavior are simulated in predictions up to 5 years ahead with initialized climate models (Fig. 6a), potentially providing a physical basis for improved skill in the North Atlantic from initialization. This potential predictability is not found in models driven only by external radiative forcing changes (Fig. 6b). This analysis also shows that the multimodel hindcasts fail to predict the rise between 1991 and 1995. Meanwhile, other studies show potential skill in predicting the mid-1990s climate shift in the northern North Atlantic associated with AMOC variations (Yeager et al. 2012; Robson et al. 2012b; R. Msadek et al. 2014, unpublished manuscript). Chikamoto et al. (2012b) tie this shift in the North Atlantic to predictability of a climate shift in the western Pacific around this same time in initialized hindcasts. Consistent with earlier comparisons made by Collins et al. (2006) of AMOC potential predictability in several models, Msadek et al. (2010) found the leading mode of AMOC variability to have a potential predictability up to two decades, while Teng et al. (2011) and Persechino et al. (2012) found it saturates after only about a decade. Interestingly, these studies and many others (Branstator et al. 2012; García-Serrano and Doblas-Reyes 2012; van Oldenborgh et al. 2012; Smith et al. 2010; Matei et al. 2012c; Yang et al. 2013; A. Rosati et al. 2014, unpublished manuscript; Terray 2012; García-Serrano et al. 2012) point to upper-layer ocean temperature and SST being more predictable in the subpolar gyre of the North Atlantic than in other regions. However, this result appears to be model dependent (Kim et al. 2012; Ham et al. 2014) and is likely influenced by model biases. Additionally, if a model has significant “oscillatory” variability in AMOC, this translates into longer predictability if the oscillatory

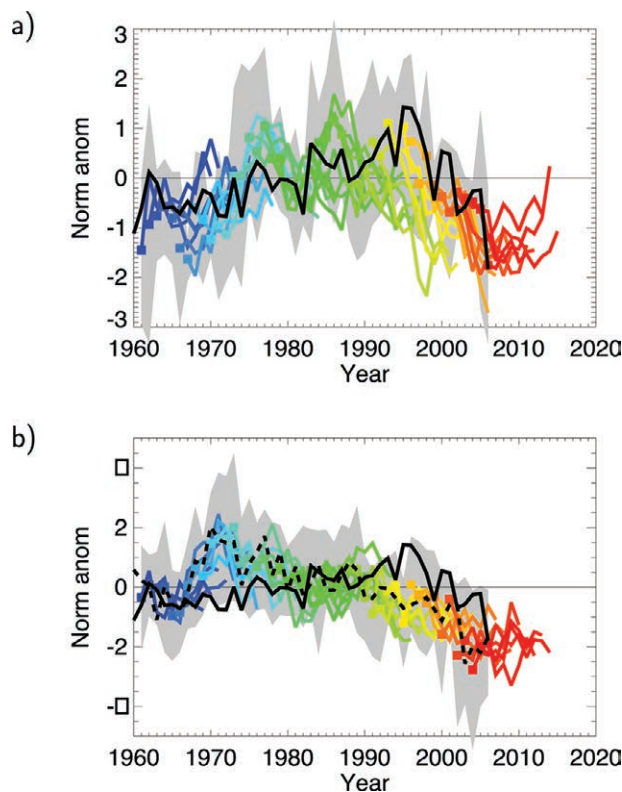


FIG. 6. Decadal predictions of the AMOC. (a) Multi-model annual mean AMOC at 45°N of retrospective predictions (colored, each color representing a different start year), together with the multimodel mean analysis (black). Squares are plotted for the first year of the retrospective predictions and the gray shading represents their 95% ensemble range. (b) As in (a) but for uninitialized retrospective predictions with HadCM3. The dashed black curve in (b) shows the ensemble mean of the transient simulations (Pohlmann et al. 2013).

behavior is realistic, with the reverse also being true. Thus, the predictability ranges in these studies can be explained to a certain extent by their AMOC variability. This has practical consequences since surface temperature is the connection through which the AMOC can impact the overlying atmosphere (e.g., Gastineau et al. 2012).

As an example, beginning in the winter of 1995/96 in the space of just a few years, sea surface temperatures in the North Atlantic subpolar gyre rose by about 1°C, and upper-ocean heat content also underwent a major step change. Yeager et al. (2012) and Robson et al. (2012a,b) conclude that the warming was primarily caused by the enhanced meridional heat transport associated with a strengthened AMOC, which was itself a response to the persistent positive phase of the NAO that occurred in the 1980s and early 1990s. The initialization with very anomalous AMOC conditions in the early 1990s is what allows the Hadley Centre's Decadal Prediction System (DePreSys) and the National Center for Atmospheric Research's Community Climate System Model version 4 (CCSM4) decadal predictions to capture the rapid warming, despite relatively poor skill at predicting surface heat flux. The unusually negative NAO index that occurred in the winter of 1995/96 contributed to the rapidity of the warming but was not the fundamental cause. In addition, the impacts of this event over land also appear somewhat predictable (Robson et al. 2013). Consistent results are being found using other models and prediction systems (e.g., R. Msadek et al. 2014, unpublished manuscript), with strong preconditioning by the NAO essential for the correct initialization of AMOC strength and subsequent skillful prediction of the anomalous heat advection from the south. While the warming was predicted by the hindcasts initialized as early as in 1991 in the National Center for Atmospheric Research (NCAR) and DePreSys prediction systems (Yeager et al. 2012; Robson et al. 2012b), only the 1995 prediction was found to yield a rise in ocean heat content comparable to observations in the Geophysical Fluid Dynamics Laboratory (GFDL) predictions (R. Msadek et al. 2013, unpublished manuscript). This suggests that a preconditioning of a few years is needed to get the anomalous enhancement of the AMOC that drives the northward heat transport but that the duration of that preconditioning might be model dependent.

Although the North Atlantic stands out in most CMIP5 models as the primary region where skill might be improved because of initialization (e.g., Fig. 4), encouraging results have also been found in the tropical Atlantic. Retrospective multiyear predictions

of North Atlantic hurricane frequency have been investigated in two climate models: DePreSys and the GFDL CM2.1 (Smith et al. 2010; Vecchi et al. 2013). High correlations that are significant relative to climatology were found from as early as 10 years in advance in both models. Key to the large retrospective correlation in the initialized predictions is capturing the observed upward shift in hurricane frequency in the mid-1990s and the resulting trend over the whole time period. In DePreSys, initialization improves the skill via remote ocean conditions in the North Atlantic subpolar gyre and tropical Pacific, which influence the tropical Atlantic through atmospheric teleconnections (Dunstone et al. 2013). In the Vecchi et al. (2013) study, the improvement in skill from initialization was related to improvements in the tropical North Atlantic. However, much of the skill in both models arose from external forcings. It has been suggested that aerosols from previous volcanic eruptions (Otterå et al. 2010) and/or anthropogenic sources (Booth et al. 2012) could play a nonnegligible role in producing additional skill in the decadal hindcasts for the Atlantic. Indeed, models suggest that anthropogenic aerosols may have depressed Atlantic hurricane activity since 1860, with a particularly prominent influence in recent decades, possibly producing decadal modulations in phase with those observed (Dunstone et al. 2013; Villarini and Vecchi 2013). The role of external forcing in decadal climate prediction therefore deserves particular attention in future studies.

Predictive skill in the Pacific. As seen in Figs. 3 and 4, prediction skill in the North Pacific is less compared with the Atlantic and Indian Oceans. This is due, in part, to the Pacific being inherently more sensitive to initial state uncertainty (Branstator et al. 2012; Branstator and Teng 2012) as well as to uncertainty in the mechanisms of internally generated decadal climate variability in the Pacific. Interannual climate variability in the Pacific is dominated by El Niño–Southern Oscillation (ENSO) and the relationship between ENSO and decadal variability in the Pacific remains a subject of debate. Some argue that the broad “ENSO like” pattern of Pacific decadal variability (PDV) related to the Pacific Decadal Oscillation (PDO) or Interdecadal Pacific Oscillation (IPO) is simply a residual pattern that results from the spatial asymmetries of ENSO and skewness in ENSO statistics. Others argue that decadal changes in the tropical Pacific mean state are forced by separate mechanisms, and may in fact influence the amplitude, frequency, and teleconnections of ENSO (Power et al. 1999a; Meehl and Hu 2006; Matei et al. 2008; Meehl et al. 2010).

A case for initialization providing additional decadal prediction skill by capturing a mechanism producing decadal variability was made by Mochizuki et al. (2010), who found predictive skill over the extratropical North Pacific related to the PDO. This was confirmed also by updated versions of the prediction system (Mochizuki et al. 2012). Mochizuki et al. (2010) and Chikamoto et al. (2012a) show that the source of skill resides in the model's ability to

follow observed subsurface temperature changes in the North Pacific. In a different set of models, Guémas et al. (2012) show that the failure in representing two major warming events that occurred around 1963 and 1968 is the primary explanation for the low predictive skill in that basin (Folland et al. 2002; Salinger et al. 2001; Power et al. 1999a). They also show, in agreement with Mochizuki et al. (2010), that the 1963 warm event stemmed from the

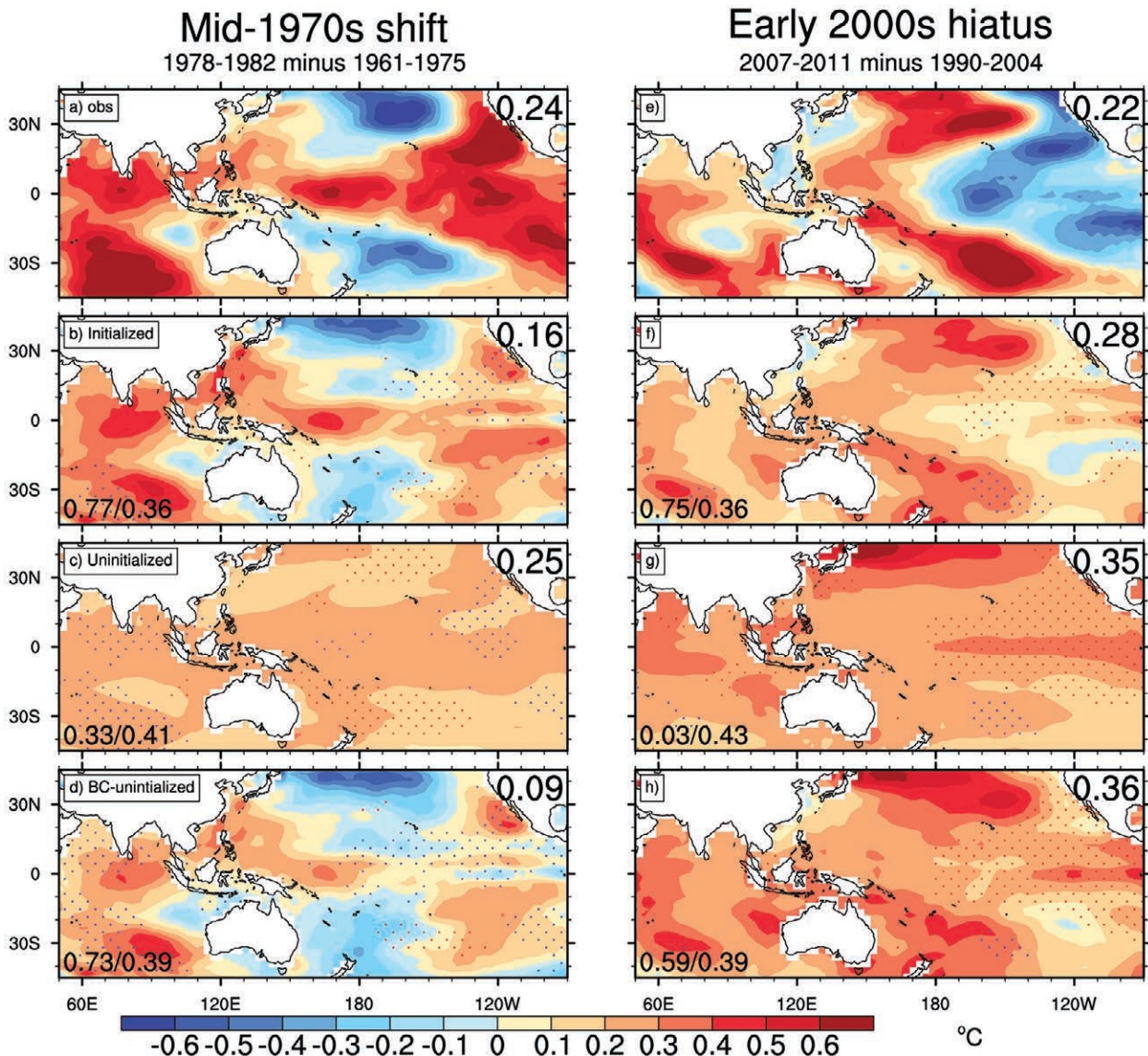


FIG. 7. Spatial distribution of 5-yr mean surface air temperature changes ($^{\circ}\text{C}$) for the CMIP5 models during the mid-1970s shift (1978–82 minus 1961–75 for hindcasts initialized in 1976) and during the early 2000s hiatus period (2007–11 minus 1990–2004 for hindcasts initialized in 2005). Area-weighted globally averaged surface air temperature anomalies are shown at the top right of each panel. Pattern correlations and RMS errors are shown at the lower left in each panel, with higher correlations and lower RMS errors indicating more skillful predictions. For the mid-1970s shift and early-2000s hiatus, the (b),(f) initialized predictions show significant pattern correlations with observations, while the (c),(g) uninitialized free-running simulations have lower values that are not significant. Bias-adjusting the uninitialized simulations (d), (h) produces improvements over the raw uninitialized (Meehl and Teng 2012, 2014).

propagation of a warm anomaly along the Kuroshio–Oyashio Extension. However, the spatial extent of observed decadal variability over the Pacific is not limited to the northern extratropics but also extends into the tropics and the Southern Hemisphere. Others have presented modeling evidence that some of the decadal variability in the South Pacific represents the low frequency and oceanic responses to prior ENSO activity that arises through the accumulation of ENSO-driven surface heat flux forcing or through the excitation of low frequency wind-driven Rossby waves (Power and Colman 2006).

Keenlyside et al. (2008), Smith et al. (2010), and van Oldenborgh et al. (2012) also show some signs of improved skill through initialization in predictions of multiyear tropical Pacific temperatures. That skill related to physical mechanisms associated with the IPO/PDO in the North Pacific and the AMO in the Atlantic is greater than for the PDO (Doblas-Reyes et al. 2013; Kim et al. 2012). However, case studies in two models (Meehl and Teng, 2012) and in the CMIP5 models (Meehl and Teng, 2014) for the mid-1970s climate shift and the early 2000s hiatus show improved skill for those large climate fluctuations of the IPO compared to the free-running uninitialized models (Meehl and Teng 2012), and this was shown to be the case for the CMIP5 multimodel ensemble as well (Meehl and Teng 2014). Results from an initialized model indicates greater skill in predicting the hiatus compared to uninitialized (Guemas et al., 2013b). Figure 7 shows that initialized hindcasts for the mid-1970s shift have greater anomaly pattern correlations with the observations in the Pacific than for the uninitialized free-running simulations (Fig. 7c). Similar results are shown for the early-2000s hiatus (initialized hindcasts in Fig. 7 have higher anomaly pattern correlations with the observations than the free-running uninitialized simulations). Meanwhile, the hindcasts are improved in the uninitialized simulations if they are bias adjusted as noted earlier.

Predictive skill over land. One of the ultimate aims of decadal climate prediction is to provide skillful and reliable predictions of societally relevant quantities over land areas where there are large human populations that could be affected by decadal climate variability. Skillful decadal predictions of North Atlantic Ocean temperatures could lead to skillful predictions of important climate impacts over land, including rainfall over the African Sahel, India, and Brazil; Atlantic hurricanes; and summer climate over Europe and America (Sutton and Hodson 2005; Zhang and Delworth 2006; Knight et al. 2005; Dunstone et al.

2011; Sutton and Dong 2012). In particular, with regards to Sahelian precipitation, it has been found that there is currently a lack of overall predictive skill for precipitation in West Africa, though there are hints of an ability to predict large shifts associated with the Sahelian drought of the 1970s and 1980s at multiannual time scales (van Oldenborgh et al. 2012; MacLeod et al. 2012; García-Serrano et al. 2013; Gaetani and Mohino 2013).

Older studies detrended the AMO with a linear trend, which implied that many of the effects are in fact the signature of the nonlinear global warming trend. When the AMO is defined relative to a nonlinear global warming trend the effects over land are much smaller. Nevertheless, there is emerging evidence of skillful predictions of temperature and precipitation over the United States and Europe following the mid-1990s warming of the subpolar gyre (Robson et al. 2013).

Similarly, skillful decadal predictions of Pacific SSTs associated with, for example, the IPO, could produce improved decadal predictions of rainfall over North and South America, Asia, Africa, and Australia (Power et al. 1999b; Deser et al. 2004; Meehl and Hu 2006; Smith et al. 2012a). The combination of Pacific and Atlantic decadal variability of SSTs may help to explain multidecadal U.S. drought frequency (McCabe et al. 2004; Schubert et al. 2004, 2009), thus implying that better decadal predictions of SSTs in both basins could produce more skillful drought outlooks over the United States.

Despite improved predictions of Atlantic and Pacific Ocean temperatures arising from initialization, improved predictions of temperature and rainfall over land appear to be less robust (Goddard et al. 2012a). Predictive skill of temperature over some land areas is less than over ocean areas (Fig. 4), and the patterns of predictive skill for precipitation over land are noisier than for temperature (comparing Figs. 4 and 8). However, there are a number of regions over land where there are indications of some skill for predictions of years 6–9 (Fig. 8)—for example, over the western and parts of the northeastern United States, areas of western and southern Africa, northern Europe and northern Asia, and southern South America. Some of these characteristics have also been seen in the projections where the same models are used for both initialized and uninitialized hindcasts. Similar results have also been found in a number of individual decadal prediction systems (Matei et al. 2012c; Bellucci et al. 2013; Müller et al. 2012). Furthermore, the frequency of extreme events is predicted with higher skill than persistence in

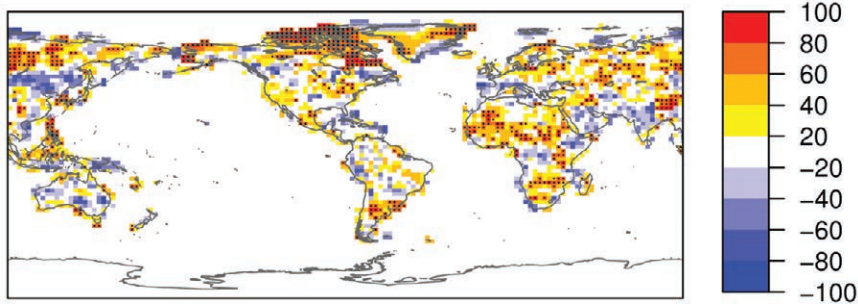


FIG. 8. Precipitation predictive skill (correlation with observations), predictions for years 6–9 averages based on CMIP5 multimodel ensemble mean hindcasts (see Table 1 for details). Results are from initialized hindcasts with 5-yr intervals between start dates from 1960 to 2005. Correlations are calculated by averaging all the hindcasts from all the models for each start date for the 6–9-yr hindcasts, subtracting the model climatologies to form anomalies, then taking the time series of those anomalies for each start date at each grid point and correlating corresponding gridpoint time series from the observations formulated in the same way as the model hindcast anomalies, and those correlations are plotted at each grid point to form the map. Stippling indicates significance at the 95% level, and darker red colors with stippling indicate areas with more significant predictive skill (Doblas-Reyes et al. 2013).

many land regions for temperature and over Europe for rainfall, although this skill arises mainly from radiative forcing beyond the first year (Eade et al. 2012; Hanlon et al. 2013).

As noted earlier, decadal climate predictions ultimately need to be assessed for their reliability, and this is particularly relevant for, at minimum, temperature and precipitation over land areas owing to near-term decadal climate impacts on water resources, agriculture, and other societally relevant applications. Policy makers need to have a good measure of the reliability of a decadal prediction to factor into the decision-making process. An attempt to quantify reliability of surface temperature decadal predictions is illustrated in an analysis of a multimodel ensemble from the European Centre for Medium-Range Weather Forecasts (ECMWF) in Fig. 9. Reliable temperature predictions for lead times up to 6–9 years are shown for global and selected regions over land (i.e., Europe and Africa) and ocean areas of the North Atlantic and Indian Ocean, while the North Pacific is less reliable than the North Atlantic and Indian Oceans. However, the forecast resolution (a measure of how much the forecast probabilities differ from the climatological probability of the event) is reduced when the forced trends are removed (Corti et al. 2012). That is, in all regions considered in Corti et al. (2012), with the exception of the Indian Ocean, the reliability is maintained after detrending, but the forecast resolution is reduced, thus reducing the Briar skill score. In

predicting regional precipitation must be performed (e.g., Doblas-Reyes et al. 2013), and also include other quantities such as winds and humidity that will influence the usefulness of decadal climate predictions.

INITIALIZED DECADAL CLIMATE PREDICTIONS FOR NEAR-TERM CLIMATE CHANGE.

Some of the first decadal climate predictions for near-term climate change were made in the pioneering work of Smith et al. (2007), Keenlyside et al. (2008), and Pohlmann et al. (2009). More recent initialized predictions of globally averaged surface air temperature agree with Smith et al. (2007) and Keenlyside et al. (2008) in predicting somewhat smaller global warming magnitude compared to uninitialized simulations out to 2020 (Fyfe et al. 2011; Mochizuki et al. 2012), and this was reflected by a similar assessment in the IPCC AR5 for the period 2016–2035 (Kirtman et al. 2013). This is attributed partly to the negative phase of the IPO in the initialized state. Similarly, Meehl and Teng (2012) analyzed 30-yr initialized predictions from a single model using two different initialization methodologies to show that for the 20-yr average of 2016–35, the initialized predictions had an average of about 15% less globally averaged warming than the free-running uninitialized projections for that same time period. Subsequently, they analyzed the CMIP5 multimodel ensemble and found a similar result (Meehl and Teng 2014). They attributed this to the recent hiatus of global warming

in addition, Ho et al. (2013) explore the spread-error ratio (or dispersion) of SST predictions in different sets of DePreSys hindcasts and find that the dispersion is very lead-time dependent, with the ensembles tending to be underdispersed on lead times up to around 2 years and overdispersed at longer lead times, revealing that the forecast system is not well calibrated.

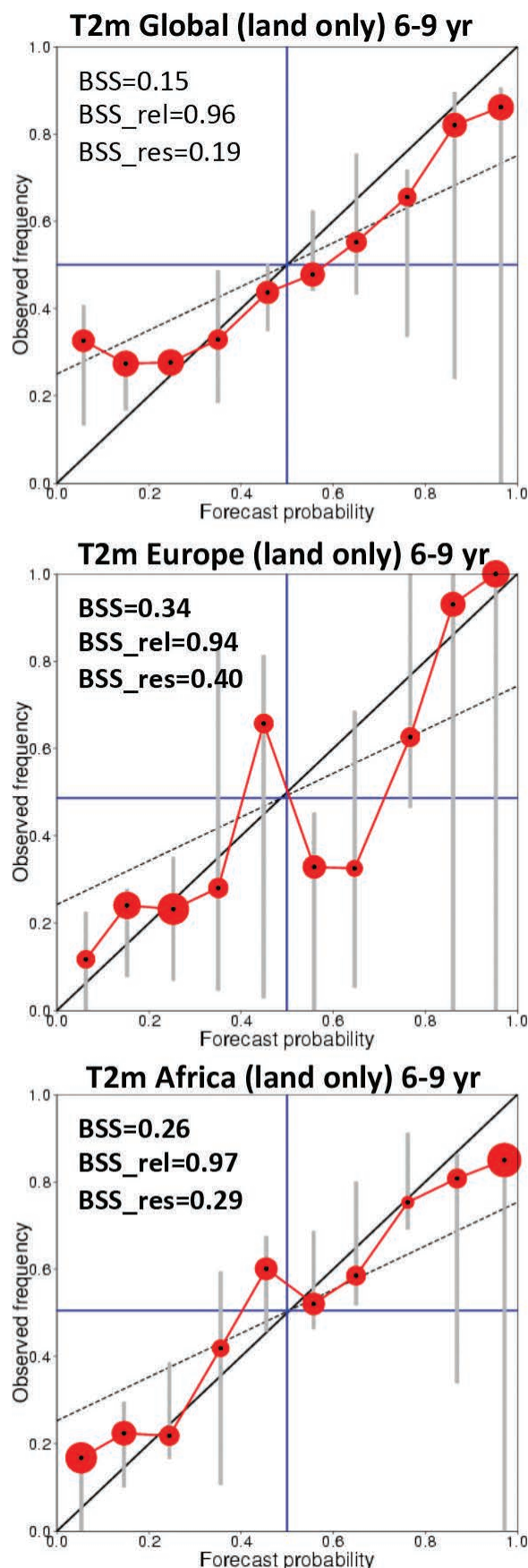
Up to now most emphasis in decadal climate prediction has been on surface temperature and understanding the possible associated mechanisms and processes that could produce predictive skill. More evaluations of skill in

associated with the negative phase of the IPO and the bias adjustment procedure that reduces a larger-than-observed warming trend in the models.

A recent exercise has been organized to perform experimental decadal climate predictions with a multimodel dataset (Smith et al. 2012b). Those predictions use nine initialized AOGCMs as well as two empirical models (Lean and Rind 2009; Ho et al. 2012). The predictions are initialized in 2011 and made for the time periods 2012–16 and 2016–20. Results for this latter period are shown in Fig. 10 and indicate somewhat less warming than indicated by uninitialized projections in most regions, consistent with the Meehl and Teng (2012, 2014) results. Based on earlier assessments of predictive skill and reliability in various regions from different initialized model datasets, including, for example, ENSEMBLES, DePreSys, and CMIP5 (e.g., Figs. 4 and 9), and taking into account the more consistent results from the larger number of models in CMIP5, these predictions are likely most reliable and therefore useful in the North Atlantic, western Pacific, and Indian Oceans and over land areas of Europe and Africa (e.g., Corti et al. 2012).

CONCLUSIONS AND PROSPECTS FOR THE FUTURE OF DECADEAL CLIMATE PREDICTION. The decadal climate prediction element of the CMIP5 experimental design provides a coordinated multimodel dataset of decadal hindcasts and predictions that extend earlier decadal prediction activities like DePreSys and ENSEMBLES. This

FIG. 9. Attributes diagrams for ECMWF multimodel decadal hindcast for surface air temperature above the (top) median for global land areas, (middle) European land areas, and (bottom) African land areas for years 6–9 of the hindcasts. The size of the bullets represents number of predictions in each probability category. The blue horizontal and vertical lines indicate the climatological frequency of the event in observations and predictions, respectively. Gray vertical bars indicate the uncertainty in the observed frequency for each probability category estimated at the 95% confidence level. The Brier skill score (BSS) and its components (reliability and resolution) are indicated at the top left corner of each panel (positive values indicate prediction skill better than climatology, and in bold when BSS indicates a reliable prediction significant at the 95% level). Black dashed line separates skillful (below the dashed line) from unskillful (above the dashed line) predictions. A flatter slope of lines connecting red bullets indicates an overconfident and thus unreliable prediction; a steeper slope indicates a more reliable prediction (Corti et al. 2012).



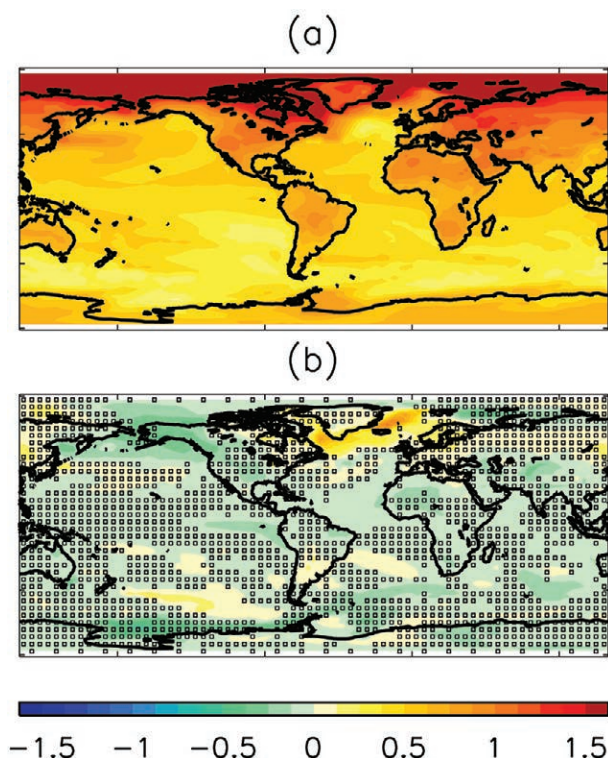


FIG. 10. (a) Initialized predictions for a multimodel average (eight models) of surface air temperature for the 5-yr period 2016–20, and (b) initialized results minus those for uninitialized simulations, with negative values indicating that the initialized predictions are cooler than the uninitialized simulations, with stippling indicating differences significant at the 90% level (Smith et al. 2012b).

information is useful for analysis of predictability and predictions on time scales from 1 to 30 years as defined by CMIP5, and also has the potential to provide insights into the workings of the climate system—for example, through identifying mechanisms in the Atlantic associated with the AMOC or in the Pacific associated with the IPO. There are advantages of initializing with observations because there is predictive skill in some areas for the first few years of an initialized prediction compared to simulations with only external forcing factors such as increasing greenhouse gases. Analyses of multimodel datasets have shown that contributions from initialization and externally forced trends in various combinations for the North Atlantic, Indian, and western Pacific make SSTs in those regions more predictable than other oceanic areas and that a multimodel average outperforms most single models for the decadal prediction problem.

Though modeling groups have applied a variety of methods for initialization, there is still no clear

indication of which method is the best. Model initialization and how it is applied to the decadal climate prediction problem is still an active research problem, particularly whether there are advantages to better initialization of the cryosphere and land surface.

Bias adjustment derives from model errors and how to account for them in the predictions, and has been addressed using a wide variety of techniques. It is likely that, with further model improvements, bias adjustment will become less of an issue. In the meantime, the bias adjustment process can be used for other purposes, such as deriving estimates of climate sensitivity of the observed climate system. As the current decadal climate predictions are verified over the next several years, it will be interesting to see if the predictions of a somewhat reduced rate of global warming, compared to free-running uninitialized projections, will be accurate.

Focused studies of predictability of specific, highly predictable patterns (DelSole et al. 2011; Yang et al. 2013), or of high predictability found in certain regions (Branstator et al. 2012; Branstator and Teng 2012) may be avenues of future research that can enhance the benefits of initialization for decadal climate predictions. Understanding the physical processes and sources of skill will be crucial for gaining confidence in forecasts. Consideration of more than oceanic conditions is another direction that predictability studies are likely to take. However, recent investigations of surface conditions over land (Jia and DelSole 2011; Teng et al. 2011) and of Arctic sea ice (Holland et al. 2010; Blanchard-Grigglesworth et al. 2011; Toyoda et al. 2011) find predictability is limited to just a few years for those quantities.

The producers of decadal climate predictions must be careful to quantify and caveat future predictions so as not to raise expectations and ensure that the output will be put to appropriate uses. To return to the example of the water resource planning community as potential users of decadal climate prediction information involving temperature and precipitation and associated evapotranspiration, there are indications that initialized predictions contain more reliable climate information for some land regions than uninitialized predictions (Corti et al. 2012). Since climate information is only one of a number of factors taken into account for water resource management decisions, water managers are already using uninitialized projections as input to their decision-making process over the next few years to several decades in the future (Means et al. 2010). Improvements to climate information, such as the initialized predictions that

show about 15% less global warming over the next few decades compared to the uninitialized projections, are of use to that community right now with regards to evapotranspiration impacts on water resources, even with the caveats that accompany them. Predictions of the time evolution of regional temperature and precipitation will likely be more reliable in the next generation of decadal climate predictions, and will be of even more use to that community. However, the reliability characteristics of hindcasts need to be explored in a wider range of models.

A useful analog as to how these types of activities evolve arises from the ENSO community (McPhaden et al. 2010). They progressed to where they had the modeling capability and physical understanding to begin to make initialized ENSO predictions in the early 1990s (Stockdale et al. 1998). However, as with decadal climate prediction, they were dealing with model systematic errors and initialization shock. As a consequence, methods were developed, such as the “two tier” approach, whereby the initialized ocean was run, and then SSTs from that simulation were used to force an atmospheric model to produce temperature and precipitation climate predictions over North America and elsewhere (Bengtsson et al. 1993). As the models continued to improve and initialization shock became less of an issue, a number of modeling groups made experimental ENSO predictions with fully coupled global climate models, and the groups began to compare their predictions informally to build credibility (McPhaden et al. 2010). This is similar to what is happening now in the decadal climate prediction community (Smith et al. 2012b). Early in the twenty-first century, ENSO prediction transitioned to an operational activity, and today ENSO predictions are used by a wide variety of stakeholders (Jin et al. 2008). It is likely that such an evolution will take place for decadal climate prediction for global climate.

Regarding upcoming decadal prediction activities, first and foremost the climate models used to make such predictions must improve. Though there is a desire to add more complexity to climate models, there is also a push to improve the representation of processes and feedbacks and use models of ever higher resolution. All these efforts should improve the model simulations, reduce the need for bias adjustments, and provide more reliable predictions. In addition, hindcasts from larger numbers of initial states and larger ensemble sizes will be of value to better define reliability of decadal climate predictions, although metrics such as the dispersion may be relatively robust with limited start dates (Ho et al.

2013). More assessments of precipitation predictions will be performed to go beyond analysis of surface temperatures in the predictions. Projects such as Seasonal-to-Decadal Climate Prediction for the Improvement of European Climate Services (SPECS) and “Mittelfristige Klimaprognose” (meaning decadal climate prediction) (MiKlip) are underway, and planning has begun to incorporate a new set of decadal climate predictions experiments into the sixth phase of CMIP (CMIP6). As noted above, future decadal climate predictions will likely use higher-resolution versions of global coupled climate models, which should result in improved predictions of regional climate variability and change through better representation of climate processes. For example, a great improvement in the representation of the quasi-persistent Euro-Atlantic flow regimes is found with increasing model resolution (Dawson et al. 2012). If the ENSO prediction experience noted above is any guide, experimental decadal predictions done now within the climate science community (e.g., Smith et al. 2012b) will likely become more formalized in operational decadal climate prediction activities and climate services within the next 5–10 years.

It remains an important question as to whether or not decadal climate predictions will end up providing useful information to a wide group of stakeholders. Indications now are that temperature, with a greater signal-to-noise ratio, shows the most promise, with precipitation being more challenging. These two quantities are typically the ones that have been addressed so far in the literature. Since sources of skill are time dependent, it is important to emphasize that for the first 5 or so years of a decadal prediction, skill could come from the initial state, and after that skill arises because of the external forcing, with some regions having potentially greater skill than others. Further quantification with other variables needs to be done and applied in reliability studies, which are just now beginning, in order to demonstrate usefulness of decadal climate predictions.

ACKNOWLEDGMENTS. The authors acknowledge the Aspen Global Change Institute (AGCI) in Aspen, Colorado, 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. The authors thank three anonymous reviewers for their constructive comments that helped considerably to improve and clarify points made in the manuscript. Portions of this study were supported by the Office of

Science (BER); U.S. Department of Energy; the National Science Foundation; the NOAA MAPP and NOAA CPO programs; the DECC/Defra Met Office Hadley Centre Climate Programme (GA01101); the European Community's Seventh Framework Programme (FP7/2007-2013) THOR and COMBINE projects; the BMBF North Atlantic II project; the NASA Modeling, Analysis, and Prediction program; the French GICC EPIDOM Project; and the Spanish RUCSS project.

REFERENCES

- Barsugli, J., C. Anderson, J. B. Smith, and J. M. Vogel, 2009: Options for improving climate modeling to assist water utility planning for climate change. Western Utilities Climate Alliance (WUCA) White Paper, 144 pp. [Available online at www.wucaonline.org/assets/pdf/actions_whitepaper_120909.pdf.]
- Bellucci, A., and Coauthors, 2013: Decadal climate predictions with a coupled OAGCM initialized with oceanic reanalyses. *Climate Dyn.*, **40**, 1483–1497, doi:10.1007/s00382-012-1468-z.
- Bengtsson, L., U. Schlese, E. Roeckner, M. Latif, T. Barnett, and N. Graham, 1993: A two-tiered approach to long-range climate forecasting. *Science*, **261**, 1026–1029.
- Blanchard-Wrigglesworth, E., C. M. Bitz, and M. M. Holland, 2011: Influence of initial conditions and climate forcing on predicting Arctic sea ice. *Geophys. Res. Lett.*, **38**, L18503, doi:10.1029/2011GL048807.
- Boer, G. J., 2011: Decadal potential predictability of twenty-first century climate. *Climate Dyn.*, **36**, 1119–1133.
- , V. Kharin, and W. J. Merryfield, 2013: Decadal predictability and forecast skill. *Climate Dyn.*, **41**, 1817–1833, doi:10.1007/s00382-013-1705-0.
- Booth, B. B. B., N. J. Dunstone, P. R. Halloran, T. Andrews, and N. Bellouin, 2012: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, **484**, 228–232.
- Branstator, G., and H. Y. Teng, 2010: Two limits of initial-value decadal predictability in a CGCM. *J. Climate*, **23**, 6292–6311.
- , and —, 2012: Potential impact of initialization on decadal predictions as assessed for CMIP5 models. *Geophys. Res. Lett.*, **39**, L12703, doi:10.1029/2012GL051974.
- , —, G. Meehl, M. Kimoto, J. Knight, M. Latif, and A. Rosati, 2012: Systematic estimates of initial value decadal predictability for six AOGCMs. *J. Climate*, **25**, 1827–1846.
- Chikamoto, Y., and Coauthors, 2012a: Predictability of a stepwise shift in Pacific climate during the late 1990s in hindcast experiments by MIROC. *J. Meteor. Soc. Japan*, **90A**, 1–21.
- , M. Kimoto, M. Watanabe, M. Ishii, and T. Mochizuki, 2012b: Relationship between the Pacific and Atlantic stepwise climate change during the 1990s. *Geophys. Res. Lett.*, **39**, L21710, doi:10.1029/2012GL053901.
- Collins, M., and Coauthors, 2006: Interannual to decadal climate predictability in the North Atlantic: A multimodel-ensemble study. *J. Climate*, **19**, 1195–1203.
- Corti, S., A. Weisheimer, T. N. Palmer, F. J. Doblas-Reyes, and L. Magnusson, 2012: Reliability of decadal predictions. *Geophys. Res. Lett.*, **39**, L21712, doi:10.1029/2012GL053354.
- Cubasch, U., and Coauthors, 2001: Projections of future climate change. *Climate Change 2001: The Scientific Basis*, J. T. Houghton, Eds., Cambridge University Press, 525–582.
- Dawson, A., T. N. Palmer, and S. Corti, 2012: Simulating regime structures in weather and climate prediction models. *Geophys. Res. Lett.*, **39**, L21805, doi:10.1029/2012GL053284.
- DelSole, T., 2004: Predictability and information theory. Part I: Measures of predictability. *J. Atmos. Sci.*, **61**, 2425–2440.
- , and M. K. Tippett, 2009: Average predictability time. Part II: Seamless diagnosis of predictability on multiple time scales. *J. Atmos. Sci.*, **66**, 1172–1187.
- , M. K. Tippett, and J. Shukla, 2011: A significant component of unforced multidecadal variability in the recent acceleration of global warming. *J. Atmos. Sci.*, **68**, 909–926.
- , L. Jia, and M. K. Tippett, 2013: Decadal prediction of observed and simulated sea surface temperatures. *Geophys. Res. Lett.*, **40**, 2773–2778, doi:10.1002/grl.50185.
- Delworth, T. L., R. Zhang, and M. E. Mann, 2007: Decadal to centennial variability of the Atlantic from observations and models. *Ocean Circulation: Mechanisms and Impacts*, *Geophys. Monogr.*, Vol. 173, Amer. Geophys. Union, 131–148.
- 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., M. A. Balmaseda, A. Weisheimer, and T. N. Palmer, 2011: Decadal climate prediction with the ECMWF coupled forecast system: Impact of ocean observations. *J. Geophys. Res.*, **116**, D19111, doi:10.1029/2010JD015394.
- , and Coauthors, 2013: Initialized near-term regional climate change prediction. *Nature Commun.*, **4**, 1715, doi:10.1038/ncomms2704.

- Du, H., F. J. Doblas-Reyes, J. García-Serrano, V. Guémas, Y. Soufflet, and B. Wouters, 2012: Sensitivity of decadal predictions to the initial atmospheric and oceanic perturbations. *Climate Dyn.*, **39**, 2013–2023, doi:10.1007/s00382-011-1285-9.
- Dunstone, N. J., D. M. Smith, and R. Eade, 2011: Multi-year predictability of the tropical Atlantic atmosphere driven by the high latitude North Atlantic Ocean. *Geophys. Res. Lett.*, **38**, L14701, doi:10.1029/2011GL047949.
- , —, B. B. Booth, L. Hermanson, and R. Eade, 2013: Anthropogenic aerosol forcing of Atlantic tropical storms. *Nature Geosci.*, **6**, 534–539.
- Eade, R., E. Hamilton, D. M. Smith, R. J. Graham, and A. A. Scaife, 2012: Forecasting the number of extreme daily events out to a decade ahead. *J. Geophys. Res.*, **117**, D21110, doi:10.1029/2012JD018015.
- Fitzroy, R., 1863: *The Weather Book: A Manual of Practical Meteorology*. Vol. 2. Longman, Green, Longman, Roberts, and Green, 464 pp.
- Folland, C. K., J. A. Renwick, M. J. Salinger, and A. B. Mullan, 2002: Relative influences of the Interdecadal Pacific Oscillation and ENSO on the South Pacific convergence zone. *Geophys. Res. Lett.*, **29**, 1643, doi:10.1029/2001GL014201.
- Fyfe, J. C., W. J. Merryfield, V. Kharin, G. J. Boer, W.-S. Lee, and K. von Salzen, 2011: Skillful predictions of decadal trends in global mean surface temperature. *Geophys. Res. Lett.*, **38**, L22801, doi:10.1029/2011GL049508.
- Gaetani, M., and E. Mohino, 2013: Decadal prediction of the Sahelian precipitation in CMIP5 simulations. *J. Climate*, **26**, 7708–7719.
- García-Serrano, J., and F. J. Doblas-Reyes, 2012: On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. *Climate Dyn.*, **39**, 2025–2040, doi:10.1007/s00382-012-1413-1.
- , —, and C. A. S. Coelho, 2012: Understanding Atlantic multi-decadal variability prediction skill. *Geophys. Res. Lett.*, **39**, L18708, doi:10.1029/2012GL053283.
- , —, R. J. Haarsma, and I. Polo, 2013: Decadal prediction of the dominant West African monsoon rainfall modes. *J. Geophys. Res.*, **118**, 5260–5279, doi:10.1002/jgrd.50465.
- Gastineau, G., and C. Frankignoul, 2012: Cold-season atmospheric response to the natural variability of the Atlantic meridional overturning circulation. *Climate Dyn.*, **39**, 37–57.
- , F. D'Andrea, and C. Frankignoul, 2012: Atmospheric response to the North Atlantic Ocean variability on seasonal to decadal time scales. *Climate Dyn.*, **40**, 2311–2330, doi:10.1007/s00382-012-1333-0.
- Goddard, L., and Coauthors, 2012a: A verification framework for interannual-to-decadal predictions experiments. *Climate Dyn.*, **40**, 245–272, doi:10.1007/s00382-012-1481-2.
- , B. Kirtman, J. Hurrell, J. Murphy, T. Stockdale, and C. Vera, 2012b: Two time scales for the price of one (almost). *Bull. Amer. Meteor. Soc.*, **93**, 621–629.
- Griffies, S. M., and K. Bryan, 1997: Predictability of North Atlantic multidecadal climate variability. *Science*, **275**, 181–184.
- Guémas, V., F. J. Doblas-Reyes, F. Lienert, Y. Soufflet, and H. Du, 2012: Identifying the causes for the low decadal climate forecast skill over the North Pacific. *J. Geophys. Res.*, **117**, D20111, doi:10.1029/2012JD018004.
- , S. Corti, J. García-Serrano, F. J. Doblas-Reyes, M. A. Balmaseda, and L. Magnusson, 2013a: The Indian Ocean: The region of highest skill worldwide in decadal climate prediction. *J. Climate*, **26**, 726–739.
- , F. J. Doblas-Reyes, I. Andreu-Burillo, and M. Asif, 2013b: Retrospective prediction of the global warming slowdown in the past decade. *Nat. Climate Change*, **3**, 649–653, doi:10.1038/nclimate1863.
- Ham, Y.-G., M. M. Rienecker, M. J. Suarez, Y. Vikhliav, B. Zhao, J. Marshak, G. Vernieres, and S. D. Schubert, 2014: Decadal prediction skill in the GEOS-5 forecast system. *Climate Dyn.*, **42**, 1–20.
- Hanlon, H. M., G. C. Hegerl, S. F. B. Tett and D. M. Smith, 2013: Can a decadal forecasting system predict temperature extreme indices? *J. Climate*, **26**, 3728–3744.
- 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.
- , J. Robson, R. Sutton, D. Smith, and N. Keenlyside, 2011: Evaluating the potential for statistical decadal predictions of SSTs with a perfect model approach. *Climate Dyn.*, **37**, 2495, doi:10.1007/s00382-011-1023-3.
- , B. Dong, J. Robson, and R. Sutton, 2014: The interpretation and use of biases in decadal climate predictions. *J. Climate*, in press, doi:10.1175/JCLI-D-13-00473.1.
- Hazeleger, W., and Coauthors, 2013a: Predicting multi-year North Atlantic Ocean variability. *J. Geophys. Res.*, **118**, 1087–1098, doi:10.1002/jgrc.20117.
- , and Coauthors, 2013b: Multi-year climate predictions using two initialization strategies. *Geophys. Res. Lett.*, **40**, 1794–1798.

- Ho, C. K., E. Hawkins, L. Shaffrey, and F. Underwood, 2012: Statistical decadal predictions for sea surface temperatures: A benchmark for dynamical GCM predictions. *Climate Dyn.*, **41**, 917–935, doi:10.1007/s00382-012-1531-9.
- , —, —, J. Broecker, L. Hermanson, J. M. Murphy, and D. M. Smith, 2013: Examining reliability of seasonal to decadal sea surface temperature forecasts: The role of ensemble dispersion. *Geophys. Res. Lett.*, **40**, 5770–5775.
- Hoerling, M., and Coauthors, 2011: On North American decadal climate for 2011–20. *J. Climate*, **24**, 4519–4528.
- Holland, M. M., D. A. Bailey, and S. Vavrus, 2010: Inherent sea ice predictability in the rapidly changing arctic environment of the Community Climate System Model, version 3. *Climate Dyn.*, **36**, 1239–1253, doi:10.1007/s00382-010-0792-4.
- Hurrell, J., G. A. Meehl, D. Bader, T. Delworth, B. Kirtman, and B. Wielicki, 2009: A unified modeling approach to climate system prediction. *Bull. Amer. Meteor. Soc.*, **90**, 1797–1817.
- Jia, L., and T. DelSole, 2011: Diagnosis of multiyear predictability on continental scales. *J. Climate*, **24**, 5108–5124.
- , and —, 2012: Multi-year predictability of temperature and precipitation in multiple climate models. *Geophys. Res. Lett.*, **39**, L17705, doi:10.1029/2012GL052778.
- Jin, E. K., and Coauthors, 2008: Current status of ENSO prediction skill in coupled ocean–atmosphere models. *Climate Dyn.*, **31**, 647–664, doi:10.1007/s00382-008-0397-3.
- Karspeck, A., S. Yeager, T. Hoar, N. Collins, K. Raeder, J. Anderson, G. Danabasoglu, and J. Tribbia, 2013: An Ensemble adjustment Kalman Filter for the ocean component of CCSM4. *J. Climate*, **26**, 7392–7413.
- Keenlyside, N., M. Latif, J. Jungclauss, L. Kornblueh, and E. Roeckner, 2008: Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature*, **453**, 84–88.
- Kharin, V. V., G. J. Boer, W. J. Merryfield, J. F. Scinocca, and W.-S. Lee, 2012: Statistical adjustment of decadal predictions in a changing climate. *Geophys. Res. Lett.*, **39**, L19705, doi:10.1029/2012GL052647.
- Kim, H.-M., P. J. Webster, and J. A. Curry, 2012: Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. *Geophys. Res. Lett.*, **39**, L10701, doi:10.1029/2012GL051644.
- Kirtman, B., and Coauthors, 2013: Near-term climate change: Projections and predictability. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker et al., Eds., Cambridge University Press, 953–1028.
- Kleeman, R., 2002: Measuring dynamical prediction utility using relative entropy. *J. Atmos. Sci.*, **59**, 2057–2072.
- Klimont Z., S. J. Smith, and J. Cofala, 2013: The last decade of global anthropogenic sulfur dioxide: 2000–2011 emissions. *Environ. Res. Lett.*, **8**, 1–6, doi:10.1088/1748-9326/8/1/014003.
- 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.
- Kumar, A., 2007: On the interpretation of skill information for seasonal climate predictions. *Mon. Wea. Rev.*, **135**, 1974–1984.
- , 2009: Finite samples and uncertainty estimates for skill measures for seasonal predictions. *Mon. Wea. Rev.*, **137**, 2622–2631.
- , M. Chen, L. Zhang, W. Wang, Y. Xue, C. Wen, L. Marx, and B. Huang, 2012: An analysis of the nonstationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) version 2. *Mon. Wea. Rev.*, **140**, 3003–3016.
- Latif, M., and N. S. Keenlyside, 2011: A perspective on decadal climate variability and predictability. *Deep-Sea Res. II*, **58**, 1880–1894.
- Lean, J. L., and D. H. Rind, 2009: How will Earth’s surface temperature change in future decades? *Geophys. Res. Lett.*, **36**, L15708, doi:10.1029/2009GL038932.
- Liu, Z., 2012: Dynamics of interdecadal climate variability: A historical perspective. *J. Climate*, **25**, 1963–1995.
- MacLeod, D. A., C. Caminade, and A. P. Morse, 2012: Useful decadal climate prediction at regional scales? A look at the ENSEMBLES stream 2 decadal hindcasts. *Environ. Res. Lett.*, **7**, 044012, doi:10.1088/1748-9326/7/4/044012.
- Magnusson, L., M. Alonso-Balmaseda, S. Corti, F. Molteni, and T. Stockdale, 2012: Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors. *Climate Dyn.*, **41**, 2393–2409, doi:10.1007/s00382-012-1599-2.
- Majda, A., R. Abramov, and M. Grote, 2005: *Information Theory and Stochastics for Multiscale Nonlinear Systems*. American Mathematical Society, 133 pp.
- Manabe, S., and R. J. Stouffer, 1988: Two stable equilibria of a coupled ocean–atmosphere model. *J. Climate*, **1**, 841–866.

- Matei, D., N. Keenlyside, M. Latif, and J. Jungclauss, 2008: Subtropical forcing of tropical Pacific climate and decadal ENSO modulation. *J. Climate*, **21**, 4691–4709.
- , J. Baehr, J. H. Jungclauss, H. Haak, W. A. Müller, and J. Marotzke, 2012a: Multiyear prediction of monthly mean Atlantic Meridional Overturning Circulation at 26.5°N. *Science*, **335**, 76–79.
- , —, —, —, —, and —, 2012b: Response to Comment on “Multiyear prediction of monthly mean Atlantic Meridional Overturning Circulation at 26.5°N.” *Science*, **338**, 604, doi:10.1126/science.1223200.
- , H. Pohlmann, J. H. Jungclauss, W. A. Müller, H. Haak, and J. Marotzke, 2012c: Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model. *J. Climate*, **25**, 8502–8523.
- McCabe, G. J., M. A. Palecki, and J. L. Betancourt, 2004: Pacific and Atlantic Ocean influences on multi-decadal drought frequency in the United States. *Proc. Natl. Acad. Sci. USA*, **101**, 4136–4141, doi:10.1073/pnas.0306738101.
- McPhaden, M. J., A. J. Busalacchi, and D. L. T. Anderson, 2010: A TOGA retrospective. *Oceanography*, **23**, 86–103.
- Means, E. M., and Coauthors, 2010: Decision support planning methods: Incorporating climate change uncertainties into water planning. Western Utilities Climate Alliance White Paper, 102 pp. [Available online at www.wucaonline.org/assets/pdf/pubs_whitepaper_012110.pdf].
- Meehl, G. A., and A. Hu, 2006: Megadroughts in the Indian monsoon region and southwest North America and a mechanism for associated multi-decadal Pacific sea surface temperature anomalies. *J. Climate*, **19**, 1605–1623.
- , and J. M. Arblaster, 2011: Decadal variability of Asian–Australian and monsoon–ENSO–TBO relationships. *J. Climate*, **24**, 4925–4940.
- , and —, 2012: Relating the strength of the tropospheric biennial oscillation (TBO) to the phase of the Interdecadal Pacific Oscillation (IPO). *Geophys. Res. Lett.*, **39**, L20716, doi:10.1029/2012GL053386.
- , and H. Teng, 2012: Case studies for initialized decadal hindcasts and predictions for the Pacific region. *Geophys. Res. Lett.*, **39**, L22705, doi:10.1029/2012GL053423.
- , and —, 2014: CMIP5 multi-model initialized decadal hindcasts for the mid-1970s shift and early-2000s hiatus and predictions for 2016–2035. *Geophys. Res. Lett.*, in press.
- , and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 747–845.
- , and Coauthors, 2009: Decadal prediction: Can it be skillful? *Bull. Amer. Meteor. Soc.*, **90**, 1467–1485.
- , A. Hu, and C. Tebaldi, 2010: Decadal prediction in the Pacific region. *J. Climate*, **23**, 2959–2973.
- , and Coauthors, 2012: Climate system response to external forcings and climate change projections in CCSM4. *J. Climate*, **25**, 3661–3683.
- Mochizuki, T., and Coauthors, 2010: Pacific decadal oscillation hindcasts relevant to near-term climate prediction. *Proc. Natl. Acad. Sci. USA*, **107**, 1833–1837.
- , and Coauthors, 2012: Decadal prediction using a recent series of MIROC global climate models. *J. Meteor. Soc. Japan*, **90A**, 373–383.
- Moss, R. H., and Coauthors, 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747–756, doi:10.1038/nature08823.
- Msadek, R., K. W. Dixon, T. L. Delworth, and W. Hurlin, 2010: Assessing the predictability of the Atlantic meridional overturning circulation and associated fingerprints. *Geophys. Res. Lett.*, **37**, L19608, doi:10.1029/2010GL044517.
- Müller, W., and Coauthors, 2012: Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology. *Geophys. Res. Lett.*, **39**, L22707, doi:10.1029/2012GL053326.
- Newman, M., 2007: Interannual to decadal predictability of tropical and North Pacific sea surface temperatures. *J. Climate*, **20**, 2333–2356.
- , 2013: An empirical benchmark for decadal forecasts of global surface temperature anomalies. *J. Climate*, **26**, 5260–5269.
- Otterå, O. H., M. Bentsen, H. Drange, and L. Suo, 2010: External forcing as a metronome for Atlantic multidecadal variability. *Nat. Geosci.*, **3**, 688–694, doi:10.1038/NGEO955.
- Palmer, T. N., F. J. Doblas-Reyes, A. Weisheimer, and M. J. Rodwell, 2008: Toward seamless prediction: Calibration of climate change projections using seasonal forecasts. *Bull. Amer. Meteor. Soc.*, **89**, 459–470.
- Penland, C., 1989: Random forcing and forecasting using principal oscillation pattern analysis. *Mon. Wea. Rev.*, **117**, 2165–2185.
- Persechini, A., J. Mignot, D. Swingedouw, S. Labetoulle, and E. Guilyardi, 2012: Decadal predictability of the Atlantic Meridional Overturning Circulation and climate in the IPSL-CM5A-LR model. *Climate Dyn.*, **40**, 2359–2380, doi:10.1007/s00382-012-1466-1.
- Pohlmann, H., J. Jungclauss, 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.

- , D. M. Smith, M. A. Balmaseda, N. S. Keenlyside, S. Masina, D. Matei, W. A. Muller, and P. Rogel, 2013: Predictability of the mid-latitude Atlantic meridional overturning circulation in a multimodel system. *Climate Dyn.*, **41**, 775–785, doi:10.1007/s00382-013-1663-6.
- Power, S., and R. Colman, 2006: Multi-year predictability in a coupled general circulation model. *Climate Dyn.*, **26**, 247–272.
- , T. Casey, C. Folland, A. Colman, and V. Mehta, 1999a: Inter-decadal modulation of the impact of ENSO on Australia. *Climate Dyn.*, **15**, 319–324.
- , F. Tseitkin, V. Mehta, S. Torok, and B. Lavery, 1999b: Decadal climate variability in Australia during the 20th century. *Int. J. Climatol.*, **19**, 169–184.
- Räisänen, J., and J. S. Ylhäisi, 2011: How much should climate model output be smoothed in space? *J. Climate*, **24**, 867–880.
- Reichler, T., and J. Kim, 2008: How well do coupled models simulate today's climate? *Bull. Amer. Meteor. Soc.*, **89**, 303–311.
- Richardson, L. F., 1922: *Weather Prediction by Numerical Process*. Cambridge University Press, 236 pp.
- Robson, J., R. Sutton, K. Lohmann, D. Smith, and M. D. Palmer, 2012a: Causes of the rapid warming of the North Atlantic Ocean in the mid-1990s. *J. Climate*, **25**, 4116–4134.
- , —, and D. Smith, 2012b: Initialized decadal predictions of the rapid warming of the North Atlantic Ocean in the mid 1990s. *Geophys. Res. Lett.*, **39**, L19713, doi:10.1029/2012GL053370.
- , —, and —, 2013: Predictable climate impacts of the decadal changes in the ocean in the 1990s. *J. Climate*, **26**, 6329–6339.
- Salinger, M. J., J. A. Renwick, and A. B. Mullan, 2001: Interdecadal Pacific Oscillation and South Pacific climate. *Int. J. Climatol.*, **21**, 1705–1721.
- Schubert, S., M. Suarez, P. J. Pegion, R. D. Koster, and J. T. Bacmeister, 2004: On the cause of the 1930s dust bowl. *Science*, **303**, 1855–1859.
- , and Coauthors, 2009: A U.S. CLIVAR project to assess and compare the responses of global climate models to drought-related SST forcing patterns: Overview and results. *J. Climate*, **22**, 5251–5272.
- Shindell, D. T., and Coauthors, 2012: Radiative forcing in the ACCMIP historical and future climate simulations. *Atmos. Chem. Phys. Discuss.*, **12**, 21 105–21 210, doi:10.5194/acpd-12-21105-2012.
- 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.
- , R. Eade, N. J. Dunstone, D. Fereday, J. M. Murphy, H. Pohlmann, and A. A. Scaife, 2010: Skilful multi-year predictions of Atlantic hurricane frequency. *Nat. Geosci.*, **3**, 846–849, doi:10.1038/ngeo1004.
- , A. A. Scaife, and B. P. Kirtman, 2012a: What is the current state of scientific knowledge with regard to seasonal and decadal forecasting? *Environ. Res. Lett.*, **7**, 015602, doi:10.1088/1748-9326/7/1/015602.
- , and Coauthors, 2012b: Real-time multi-model decadal climate predictions. *Climate Dyn.*, **41**, 2875–2888, doi:10.1007/s00382-012-1600-0.
- , R. Eade, and H. Pohlmann, 2013: A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction. *Climate Dyn.*, **41**, 3325–3338, doi:10.1007/s00382-013-1683-2.
- Solomon, A., and Coauthors, 2011: Distinguishing the roles of natural and anthropogenically forced decadal climate variability: Implications for prediction. *Bull. Amer. Meteor. Soc.*, **92**, 141–156.
- Srokosz, M., M. Baringer, H. Bryden, S. Cunningham, T. Delworth, S. Lozier, J. Marotzke, and R. Sutton, 2012: Past, present, and future changes in the Atlantic meridional overturning circulation. *Bull. Amer. Meteor. Soc.*, **93**, 1663–1676.
- Stockdale, T. N., A. J. Busalacchi, D. E. Harrison, and R. Seager, 1998: Ocean modeling for ENSO. *J. Geophys. Res.*, **103** (C7), 14 325–14 355, doi:10.1029/97JC02440.
- Sutton, R. T., and D. L. R. Hodson, 2005: Atlantic Ocean forcing of North American and European summer climate. *Science*, **309**, 115–118.
- , and B. Dong, 2012: Atlantic Ocean influence on a shift in European climate in the 1990s. *Nat. Geosci.*, **5**, 788–792, doi:10.1038/NCEO1595.
- Swingedouw, D., J. Mignot, S. Labetoulle, E. Guilyardi, and G. Madec, 2012: Initialisation and predictability of the AMOC over the last 50 years in a climate model. *Climate Dyn.*, **40**, 2381–2399, doi:10.1007/s00382-012-1516-8.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498.
- Teng, H., and G. Branstator, 2011: Initial-value predictability of prominent modes of North Pacific subsurface temperature in a CGCM. *Climate Dyn.*, **36**, 1813–1834, doi:10.1007/s00382-010-0749-7.
- , —, and G. A. Meehl, 2011: Predictability of the Atlantic overturning circulation and associated surface patterns in two CCSM3 climate change ensemble experiments. *J. Climate*, **24**, 6054–6076.
- Terray, L., 2012: Evidence for multiple drivers of North Atlantic multi-decadal climate variability. *Geophys. Res. Lett.*, **39**, L19712, doi:10.1029/2012GL053046.
- Toyoda, T., and Coauthors, 2011: Impact of the assimilation of sea ice concentration data on an atmosphere-

- ocean-sea ice coupled simulation of the Arctic Ocean climate. *Sci. Online Lett. Atmos.*, **7**, 37–40.
- van der Linden, P., and J. F. B. Mitchell, Eds., 2009: ENSEMBLES: Climate change and its impacts: Summary of research and results from the ENSEMBLES project. Met Office Hadley Centre Rep., 160 pp. [Available from Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, United Kingdom.]
- van Oldenborgh, G., F. Doblas Reyes, B. Wouters, and W. Hazeleger, 2012: Decadal prediction skill in a multi-model ensemble. *Climate Dyn.*, **38**, 1263–1280.
- Vecchi, G. A., and Coauthors, 2012: Comment on “Multiyear prediction of monthly mean Atlantic meridional overturning circulation at 26.5°N.” *Science*, **338**, 604, doi:10.1126/science.1222566.
- , and Coauthors, 2013: Multiyear predictions of North Atlantic hurricane frequency: Promise and limitations. *J. Climate*, **26**, 5337–5357.
- Villarini, G., and G. A. Vecchi, 2012: Twenty-first-century projections of North Atlantic tropical storms from CMIP5 models. *Nature Climate Change*, **2**, 604–607, doi:10.1038/nclimate1530.
- , and —, 2013: Projected increases in North Atlantic tropical cyclone intensity from CMIP5 models. *J. Climate*, **26**, 3231–3240.
- Vizard, A. L., G. A. Anderson, and D. J. Buckley, 2005: Verification and value of the Australian Bureau of Meteorology township seasonal rainfall forecasts in Australia, 1997–2005. *Meteor. Appl.*, **12**, 343–255.
- Wang, B., and Coauthors, 2012: Preliminary evaluations of FGOALS-g2 for decadal predictions. *Adv. Atmos. Sci.*, **30**, 674–683, doi:10.1007/s00376-012-2084-x.
- WCRP, 2011: Data and bias correction for decadal climate predictions. International CLIVAR Project Office Publication Series 150, 5 pp. [Available online at www.wcrp-climate.org/decadal/references/DCPP_Bias_Correction.pdf.]
- Yang, X., and Coauthors, 2012: A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system. *J. Climate*, **26**, 650–661.
- Yeager, S., A. Karspeck, G. Danabasoglu, J. Tribbia, and H. Teng, 2012: A decadal prediction case study: Late 20th century North Atlantic Ocean heat content. *J. Climate*, **25**, 5173–5189.
- Zanna, L., 2012: Forecast skill and predictability of observed Atlantic sea surface temperatures. *J. Climate*, **25**, 5047–5056.
- 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.