Clim Dyn
DOI 10.1007/s00382-012-1600-0

Real-time multi-model decadal climate predictions

Doug M. Smith · Adam A. Scaife · George J. Boer · Mihaela Caian · Francisco J. Doblas-Reyes · Virginie Guemas · Ed Hawkins · Wilco Hazeleger · Leon Hermanson · Chun Kit Ho · Masayoshi Ishii · Viatcheslav Kharin · Masahide Kimoto · Ben Kirtman · Judith Lean · Daniela Matei · William J. Merryfield · Wolfgang A. Müller · Holger Pohlmann · Anthony Rosati · Bert Wouters · Klaus Wyser

Received: 27 July 2012 / Accepted: 12 November 2012

Abstract We present the first climate prediction of the coming decade made with multiple models, initialized with prior observations. This prediction accrues from an international activity to exchange decadal predictions in near real-time, in order to assess differences and similarities, provide a consensus view to prevent over-confidence in forecasts from any single model, and establish current collective capability. We stress that the forecast is experimental, since the skill of the multi-model system is as yet unknown. Nevertheless, the forecast systems used here are based on models that have undergone rigorous evaluation and individually have been evaluated for forecast skill. Moreover, it is important to publish forecasts to enable open evaluation, and to provide a focus on climate change in the coming decade. Initialized forecasts of the year 2011 agree well with observations, with a pattern correlation of 0.62 compared to 0.31 for uninitialized projections. In particular, the forecast correctly predicted La Niña in the Pacific, and warm conditions in the north Atlantic and USA. A similar pattern is predicted for 2012 but with a weaker La Niña. Indices of Atlantic multi-decadal

Electronic supplementary material The online version of this article (doi:10.1007/s00382-012-1600-0) contains supplementary material, which is available to authorized users.

D. M. Smith (✉) · A. A. Scaife · L. Hermanson
Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK
e-mail: doug.smith@metoffice.gov.uk

G. J. Boer · V. Kharin · W. J. Merryfield
Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, BC, Canada

M. Caian · K. Wyser
Rossby Centre, Swedish Meteorological and Hydrological Institute, 60176 Norrköping, Sweden

F. J. Doblas-Reyes · V. Guemas
Institut Català de Ciències del Clima, Carrer del Doctor Trueta, 203, 08005 Barcelona, Spain

E. Hawkins · C. K. Ho
NCAS-Climate, Department of Meteorology, University of Reading, Reading RG6 6BB, UK

W. Hazeleger · B. Wouters
Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands

W. Hazeleger
Wageningen University, Wageningen, The Netherlands

M. Ishii
Meteorological Research Institute, Japan Meteorological Agency, Tsukuba 305-0052, Japan

M. Kimoto
Atmosphere and Ocean Research Institute, University of Tokyo, Kashiwa 277-8568, Japan

B. Kirtman
RSMAS/MPO, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149, USA

J. Lean
Space Science Division, Naval Research Laboratory, Washington, DC 20375, USA

D. Matei · W. A. Müller · H. Pohlmann
Max-Planck-Institut für Meteorologie, Bundesstraße 53, 20146 Hamburg, Germany

A. Rosati
Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, NJ, USA

Published online: 05 December 2012

variability and Pacific decadal variability show no signal beyond climatology after 2015, while temperature in the Niño3 region is predicted to warm slightly by about 0.5 °C over the coming decade. However, uncertainties are large for individual years and initialization has little impact beyond the first 4 years in most regions. Relative to uninitialized forecasts, initialized forecasts are significantly warmer in the north Atlantic sub-polar gyre and cooler in the north Pacific throughout the decade. They are also significantly cooler in the global average and over most land and ocean regions out to several years ahead. However, in the absence of volcanic eruptions, global temperature is predicted to continue to rise, with each year from 2013 onwards having a 50 % chance of exceeding the current observed record. Verification of these forecasts will provide an important opportunity to test the performance of models and our understanding and knowledge of the drivers of climate change.

Keywords Decadal climate prediction · Multi-model ensemble · Forecast

1 Introduction

It is very likely that climate has already changed in response to human activities, with much larger changes expected by the end of this century if greenhouse gas concentrations continue to rise (IPCC 2007). However, many sectors of society now require climate forecasts of the coming decade in order to make decisions on how to respond to a changing climate, for example, for land management and crop productivity (Mendelsohn et al. 2007), energy usage, tourism and public health (Khasnis and Nettleman 2005). To address this need there is a growing international effort to develop decadal climate predictions (e.g. Meehl et al. 2009).

Climate in the coming decade is likely to be influenced both by external forcing factors, including greenhouse gases, anthropogenic aerosols, volcanic aerosols and changes in solar irradiance, and also by natural internal variability. External forcing factors are included in simulations and projections of centennial-scale climate change using dynamical climate models (IPCC 2007), and are an important source of decadal predictability (e.g. Hoerling et al. 2011). However, uncertainties would be narrowed by additionally predicting internal variability. This requires the current state of the climate system to be taken into account, and the development of decadal predictions has therefore focused on additionally initializing dynamical climate models with observations (e.g. Smith et al. 2007; Keenlyside et al. 2008; Pohlmann et al. 2009; Mochizuki et al. 2010). In addition to predicting some aspects of internal variability, initialization may also improve the skill of near term climate predictions by correcting the model's response to previous external forcing factors. Decadal forecasts have also been developed using empirical approaches (Lean and Rind 2009; Hawkins et al. 2011; Ho et al. 2012).

The growing need for decadal climate predictions is recognized by the inclusion of a protocol for historical tests in the latest climate model inter-comparison project (CMIP5, Taylor et al. 2012) for informing the upcoming IPCC fifth assessment report. These historical tests consist of a number of retrospective forecasts (hereafter referred to as hindcasts) with climate models initialized with observations that would have been available at the start of each hindcast. By comparing the hindcasts with subsequent observations it is possible to estimate the likely skill of actual forecasts (e.g. Kim et al. 2012; van Oldenborgh et al. 2012).

The latest start date of the core CMIP5 experiments is 2005 (Taylor et al. 2012); these simulations are therefore of limited utility for assessing climate in the coming decade. However, having developed the capability to perform the CMIP5 hindcasts many forecasting centers are now also producing decadal forecasts in near real-time. Recognizing this, the 15th session of the WMO Commission for Climatology (http://www.wmo.int/pages/prog/wcp/ccl/cclxv/index.php) recommended action to start the coordination and exchange of decadal forecasts. At this stage decadal forecasts are regarded as experimental, and our effort is primarily a research exercise aimed at assessing differences and similarities between the forecasts and identifying a consensus view. Nevertheless, there is a need to publish forecasts so that they can be evaluated openly. Assessing discrepancies between forecasts and subsequent observations can reveal weaknesses in initialization strategies, model simulations of internal variability, model responses to external forcing, and uncertainties in future forcing factors, all of which are invaluable for improving future forecasts. Conversely, recognizing agreement between forecasts and observations helps build confidence in future forecasts and longer range climate projections. Here we document the activity to exchange decadal forecasts in near real-time, and present the first multi-model decadal forecast. We also assess the impact of initialization on this forecast, and compare to empirical forecasts. The paper is organized as follows. In Sect. 2 we describe our approach, and provide further details of the different prediction systems in the Supplementary Information. Forecast results are presented in Sect. 3, with a summary and conclusions provided in Sect. 4.

2 Approach

To facilitate an on going exchange and assessment of decadal climate forecasts, the UK Hadley Centre has inaugurated a "Decadal Exchange", in which decadal forecasts

Table 1 Summary of decadal forecasts (see Supplementary Information for further details of prediction systems)

Name	Institute	Method D: Dynamical E: Empirical	References	Start month	Resolution[a] Atmosphere (A) Ocean (O)	Ensemble size		
						2011[b]	2012[b]	Uninitialized
CCCMA	Canadian Centre For Climate Modelling and Analysis	D: Full field initialization	Merryfield et al. (2012), Fyfe et al. (2011), Boer et al. (2012)	Jan	A: 2.8 × 2.8 L35 O: 0.94 × 1.4 L40	10	10	10
GFDL	Geophysical Fluid Dynamics Laboratory, USA	D: Full field initialization	Zhang and Rosati (2010), Zhang et al. (2007)	Jan	A: 2.5 × 2.0 L24 O:1.0 × 1.0[c] L50		10[d]	10[d]
IC3/KNMI	Institut Català de Ciències del Clima, Spain and Royal Netherlands Meteorological Institute	D: Full field initialization	Hazeleger et al. (2010, 2012), Du et al. (2012)	Nov	A: T159 L62 O: 1 × 1[c] L42	10	10	3
MIROC5	University of Tokyo, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan	D: Anomaly initialization	Tatebe et al. (2012), Chikamoto et al. (2012)	July, Oct, Jan	A:1.41 × 1.41 L40 O:1.41 × 0.80 L50	9	9	12[e]
MOHC	Met Office Hadley Centre, UK	D: Anomaly initialization	Smith et al. (2007, 2010)	Sep	A: 3.75 × 2.5 L19 O: 1.25 × 1.25 L20	10	10	10
MPI	Max Planck Institute, Germany	D: Anomaly initialization	Matei et al. (2012), Müller et al. (2012)	Jan	A: T63 L47 O: 1.5 × 1.5 L40	10	10	3
MRI	Meteorological Research Institute, Japan	D: Anomaly initialization	Yukimoto et al. (2012)	July, Oct, Jan	A:1.125 × 1.125 L48 O: 1.0 × 0.46 L51	9	9	4[e]
RSMAS	Rosenstiel School of Marine and Atmospheric Science, University of Miami, USA	D: Full field initialization	Kirtman and Min (2009), Collins et al. (2006)	Jan	A: 2.5 × 2.5 O: 1 × 1	3	3	
SMHI	Swedish Meteorological and Hydrological Institute	D: Anomaly initialization	Hazeleger et al. (2010, 2012)	Sep	A: T159 L 62 O: 1 × 1[c] L42	3	7	3
NRL	Naval Research Laboratory, USA	E: Multiple linear regression	Lean and Rind (2008, 2009), Kopp and Lean (2011)	Jan	A: 5 × 5		1	
Reading (AR1)	University of Reading, UK	E: Trend plus auto-regression	Ho et al. (2012)	Jan	O: 1 × 1	1	1	
Reading (CA)	University of Reading, UK	E:Trend plus constructed analogue	Ho et al. (2012)	Jan	O: 1.0 × 1.0 (Atlantic only)	1	1	

[a] Degrees longitude × latitude, or spectral (T159 ≈ 1.125°), L vertical levels

[b] The first calendar year of the forecast

[c] Reducing to 1/3° latitude at the equator

[d] Only ensemble mean data were provided

[e] Ensemble consists of simulations with RCP2.6, RCP4.5, RCP6.0 and RCP8.5

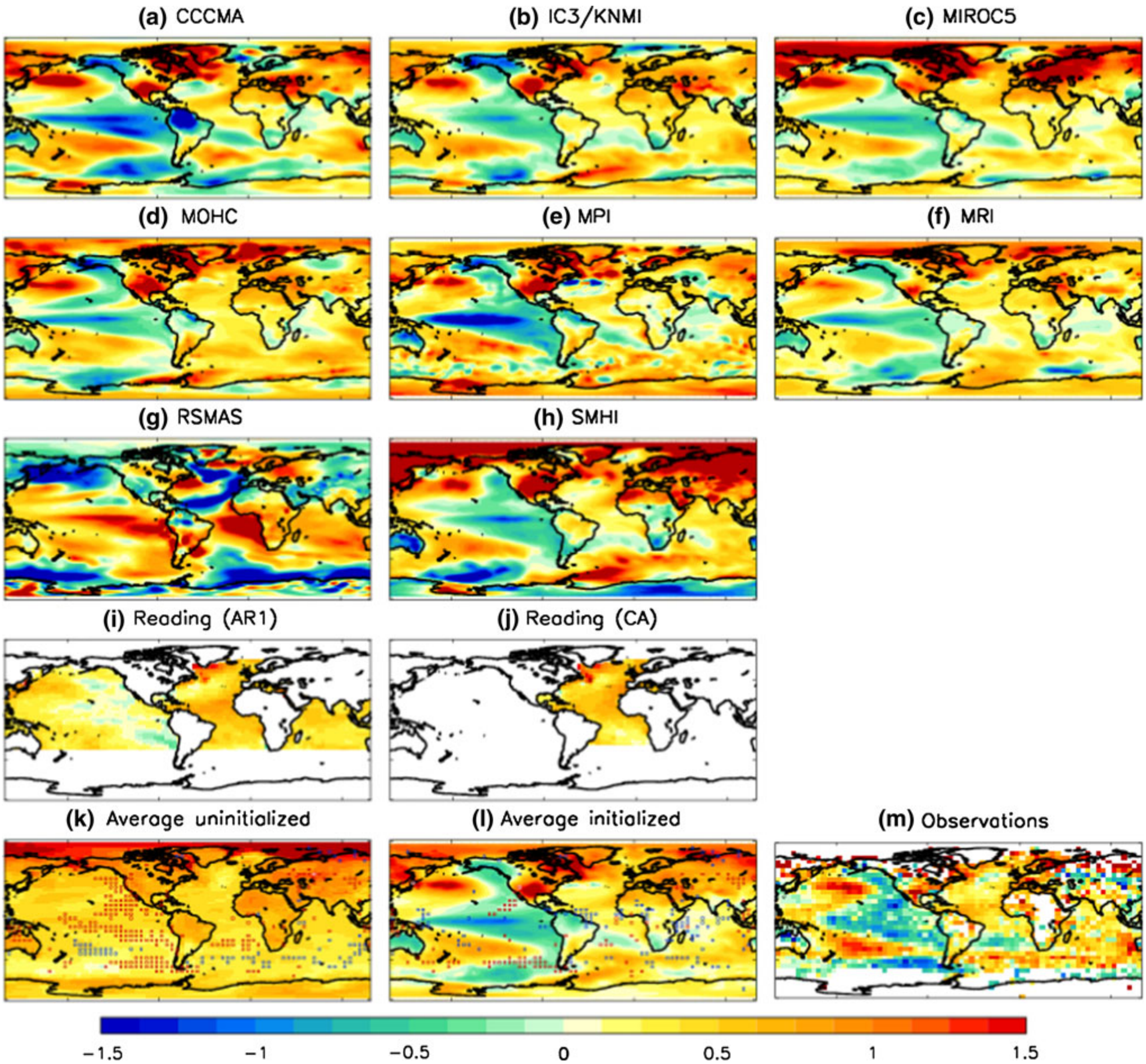

Fig. 1 Forecast and observed temperature anomalies for 2011. **a**–**j** Ensemble mean forecast from each prediction system, showing the first calendar year of forecasts starting between 1st September 2010 and 1st January 2011. Average forecasts (**k**, **l**) are for those systems for which uninitialized projections are available (see Table 1 for further details). The stippling in these indicates where the 5–95 % forecast confidence range (diagnosed from the spread of the individual ensemble members) is warmer (*red/white*) or cooler (*blue/black*) than the observations. Observations (**m**) are taken from HadCRUT3 (Brohan et al. 2006). All anomalies are degrees centigrade relative to the average of the period 1971–2000

provided by participating international scientists are updated and exchanged annually. To date two exchanges have taken place, and are summarised in Table 1 with further details of each prediction system provided in the Supplementary Information. The first (2011) exchange consists of forecasts from eight dynamical climate models starting between 1st July 2010 and 1st January 2011, plus two empirical forecasts. The second (2012) exchange consists of nine dynamical model forecasts starting between 1st July 2011 and 1st January 2012, plus three empirical forecasts. Each dynamical model forecast consists of ensembles of between 3 and 10 members, each with slightly different initial conditions. In this way the effects of unpredictable noise and, to some extent, observational uncertainty, are sampled. The multi-model ensemble additionally samples modelling uncertainties, and the overall ensemble spread provides an estimate of forecast uncertainties. Future external forcing factors were prescribed according to the

Table 2 Skill of forecasts of 2011

Name	Pattern correlation		Global mean bias[b] (°C)	
	Initialized	Uninitialized	Initialized	Uninitialized
CCCMA	0.54	0.01	−0.04	0.30
IC3/KNMI	0.55	0.29	0.02	0.29
MIROC5	0.57	0.44	0.06	0.19
MOHC	0.47	0.22	0.06	0.38
MPI	0.54	0.13	−0.06	0.38
MRI	0.55	0.08	−0.04	−0.04
RSMAS	−0.10		0.04	
SMHI	0.53	0.17	0.15	0.27
Dynamical ensemble mean[a]	0.62	0.31	0.02	0.25
Reading (AR1)	0.48		0.19	

[a] Averaged over those systems for which uninitialized forecasts are available (see Table 1 for further details)

[b] The area weighted mean of forecast minus observation at each grid point

RCP4.5 pathway of the CMIP5 protocol (Meinshausen et al. 2011). Our ensemble does not currently sample uncertainties in future emissions. With the possible exception of sulphate aerosols, emission uncertainties are thought to have little impact for the coming decade.

Currently, global fields of annual mean temperature anomalies for each calendar year of the forecast and each ensemble member are exchanged, although this may be extended to other variables, including rainfall, in future. All anomalies are computed to be relative to the period 1971–2000, with each forecasting centre applying adjustments to correct for model biases before exchanging the data. The procedure for bias adjustment depends on the initialization approach, and interested readers are referred to ICPO 2011 for full details.

In order to assess the impact of initialization, eight uninitialized dynamical model projections of the effects of external forcing factors have also been exchanged. These are obtained from simulations starting around 1850 with initial conditions taken from simulations of pre-industrial climate, so that any internal variability would not be expected to be in phase with reality. Such simulations are conventionally analysed in terms of anomalies from a fixed period (e.g. IPCC 2007). However, here we remove biases in these simulations in exactly the same way as the initialized forecasts (ICPO 2011), so that the impact of initialization may be diagnosed as the difference between initialized and uninitialized simulations.

3 Results

Users of climate forecasts ideally require accompanying skill and reliability estimates (e.g. Goddard et al. 2012). These metrics are typically determined by performing hindcast experiments, and have been assessed to some extent for most of the prediction systems included here (see citations in Table 1). In general, decadal hindcasts of surface temperature show high skill at predicting the warming trend due to external radiative forcing, with modest improvements through initialization especially in the north Atlantic and to some extent the tropical and north Pacific (e.g. Smith et al. 2010; van Oldenborgh et al. 2012; Kim et al. 2012; Chikamoto et al. 2012). However, we have not yet assessed the expected skill of the multi-model forecast developed under the Decadal Exchange activity, which we present here, and stress that our results are experimental at this stage. We also note that hindcasts do not necessarily provide an accurate estimate of forecast skill. For example, hindcasts may underestimate the skill of current forecasts, which benefit from greatly improved observations of the sub-surface ocean provided by the Argo array, but may overestimate forecast skill due to unintentional use of observations that would not have been available in a real forecast situation.

Maps of temperature anomalies for the first calendar year of the 2011 exchange are presented in Fig. 1 for each prediction system, the initialized and uninitialized multi-model means (averaged over those systems for which uninitialized forecasts are available, see Table 1), together with verifying observations. Temperature anomalies for 2011 are dominated by La Niña conditions in the Pacific (with a tongue of cool temperatures in the tropical Pacific and a horseshoe pattern of warm temperatures to the north, west and south), a cool Australia, warm high latitudes and USA, and a warm north Atlantic sub-polar gyre and tropical Atlantic. With the exception of RSMAS, all the forecasts capture this pattern well (with pattern correlations around 0.5 for each system, increasing to 0.62 for the multi-model mean, Table 2) and much better than the uninitialized projections (multi-model pattern correlation of 0.31). Furthermore, for all systems the bias is less than or equal to that of their uninitialized counterpart (Table 2), even though the uninitialized projections have been bias-corrected in the same way. Indeed, the multi-model bias is reduced from 0.25 to 0.02 °C through initialization. By definition, the observations are expected to lie outside of the 5–95 % confidence interval diagnosed from the forecast ensemble spread in 10 % of grid points. This actually occurs in 18 % of grid points for the uninitialized forecast, and 14 % for the initialized forecast (stippling in Fig. 1 k, l), suggesting that both could be slightly over-confident

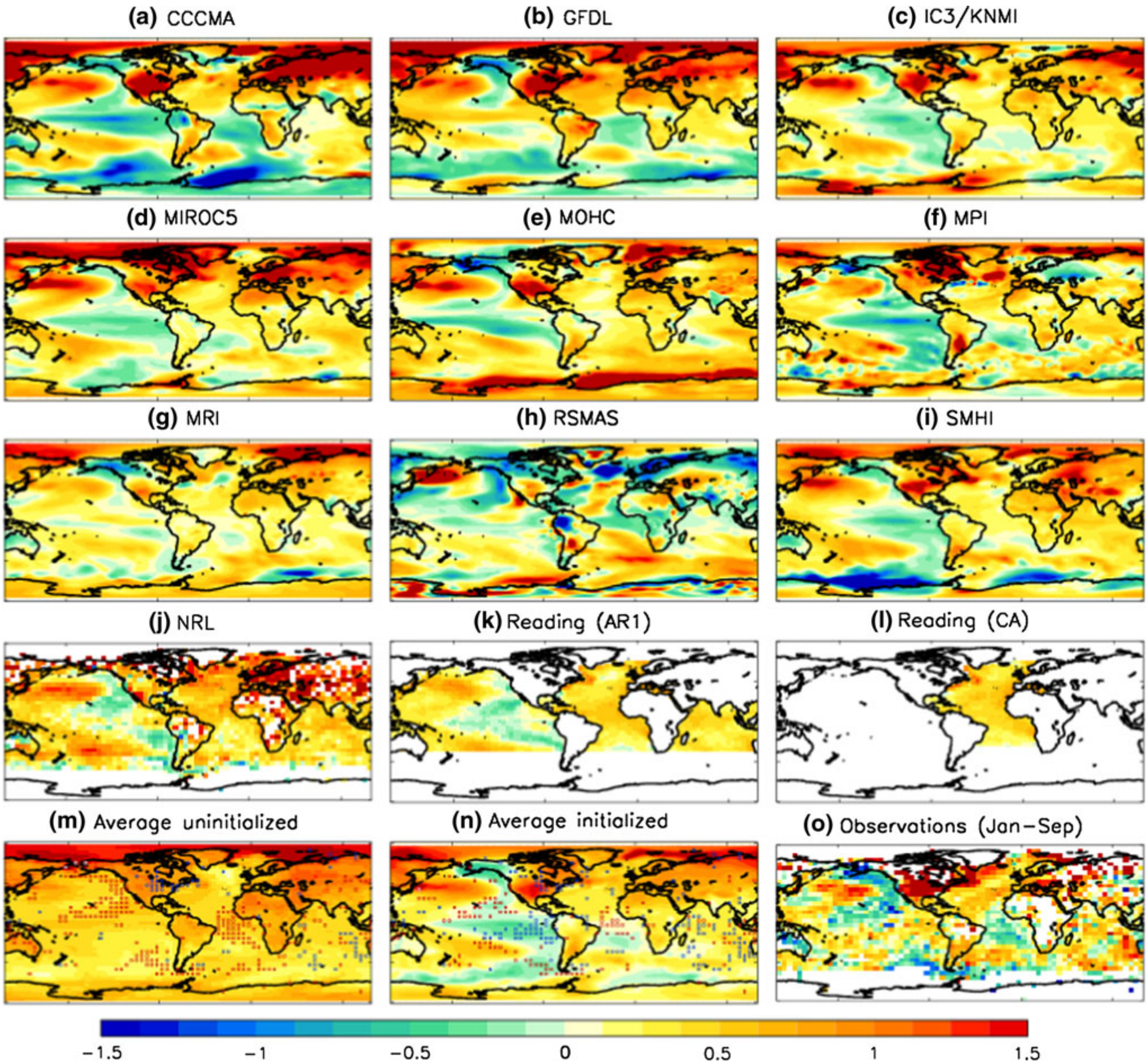

Fig. 2 As Fig. 1 but for 2012, showing the first calendar year of forecasts starting between 1st September 2011 and 1st January 2012

although we have not assessed the statistical significance of this.

Forecast temperatures for the first calendar year of the 2012 exchange (Fig. 2) show a similar pattern to 2011, although La Niña is weaker and the tropical Atlantic dipole stronger relative to 2011. Observations for the whole year are not yet available, but the initial nine months of 2012 are in reasonable agreement with the multi-model forecast (pattern correlation 0.58).

Forecasts for the 5-year periods 2012–2016 and 2016–2020 are presented in Figs. 3 and 4 for each system, and in Fig. 5 in terms of the mean and confidence intervals diagnosed from those systems for which individual ensemble members are available (Table 1). In theory, ensemble members could be weighted depending on their skill in hindcast experiments. However, evidence from seasonal forecasts suggests that in the majority of cases using equal weights for each member performs as well as more complex schemes (e.g. DelSole et al. 2012). We therefore assume that all ensemble members are equally likely. For both periods, the multi-model forecast indicates with at least a 90 % probability that temperatures will be warmer than the 1971–2000 mean for nearly all land regions (the main exception being Alaska for 2012–2016), the tropical Atlantic, Indian and western Pacific oceans (Fig. 5b, e). Predictions of the Southern Ocean are

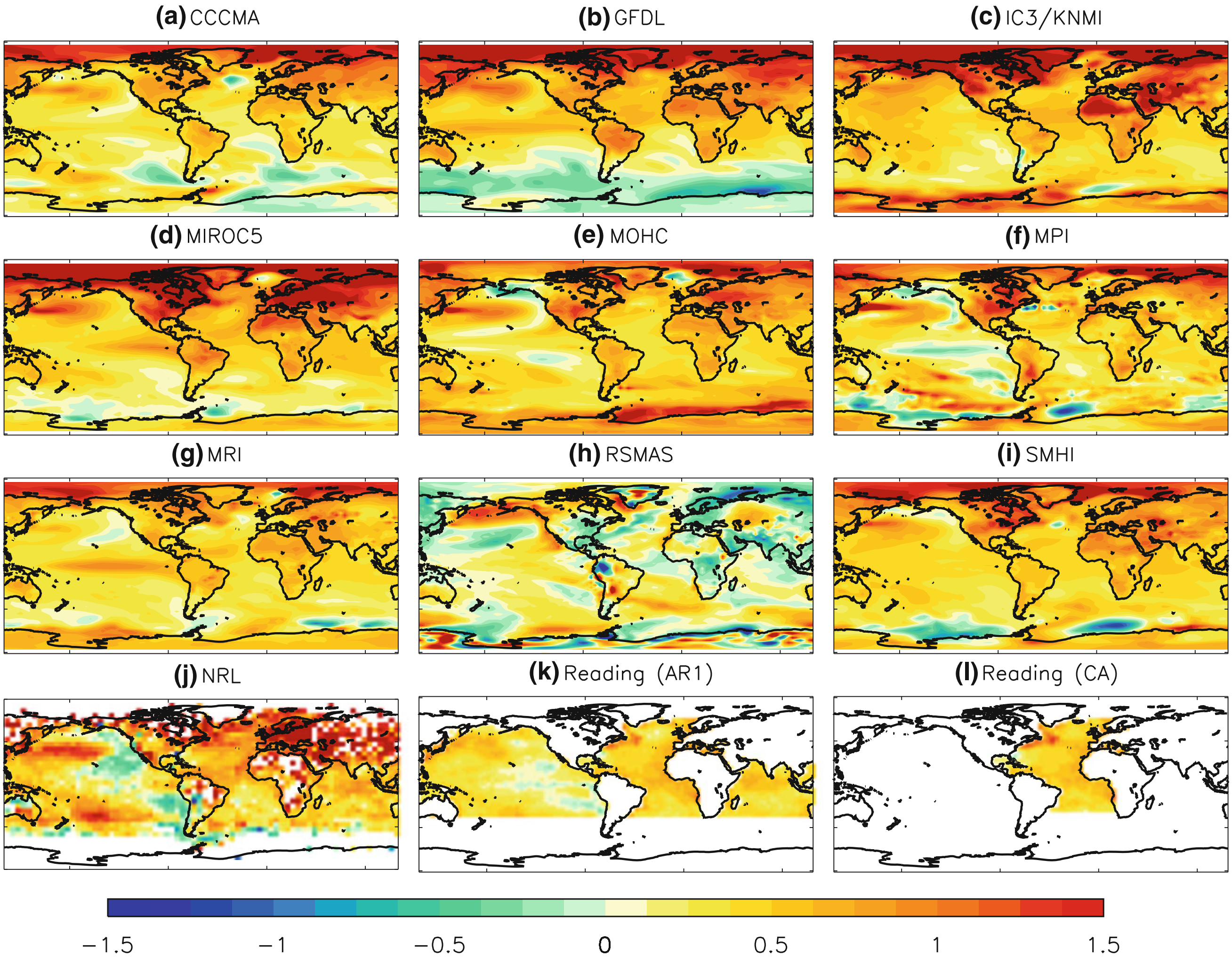

Fig. 3 Forecast temperature anomalies (as Fig. 2) for the 5-year period 2012–2016

particularly uncertain, with some systems (e.g. GFDL and RSMAS) forecasting cool anomalies even for the period 2016–2020.

The multi-model mean forecast for both periods (Fig. 5a, d) shows many similarities with long-term projections of climate change in response to increasing greenhouse gases (Meehl et al. 2007), with more warming of the land than ocean, and the largest warming over the Arctic. However, a warming minimum over the Atlantic sub-polar gyre (Meehl et al. 2007) is not evident. Indeed, the initialized forecasts are significantly warmer than the uninitialized ones in this region (Figs. 6, 7). Further investigation is required to determine whether this difference is related to differences in the Atlantic meridional overturning circulation, as suggested from analysis of coupled climate models (Knight et al. 2005; Delworth et al. 2007), and whether there are related differences in climate impacts such as rainfall over the Sahel, Amazon, USA and Europe, or Atlantic hurricane activity (Sutton and Hodson 2005; Knight et al. 2006; Zhang and Delworth 2006; Smith et al. 2010; Dunstone et al. 2011). This is outside the scope of this preliminary forecast exchange, which is limited to surface temperature only, but motivates future activity of the Decadal Exchange.

Outside the sub-polar gyre, initialization of the models results in cooler forecasts (relative to uninitialized counterparts) in almost all regions (except the Kuroshio extension), with a significant impact in many regions for 2012–2016 but mainly in the high latitude northern Pacific for 2016–2020 (Figs. 6, 7). Indeed, ensemble mean initialized forecasts of globally averaged temperature are significantly cooler than uninitialized ones until 2015 (Fig. 8a; Table 3), with a concomitant narrowing of the uncertainty range (red shading in Fig. 8a). This is consistent with the recent hiatus in global warming (e.g. Easterling and Wehner 2009). However, dynamical models and empirical approaches both predict that global mean temperature will continue to rise. Although uncertainties in

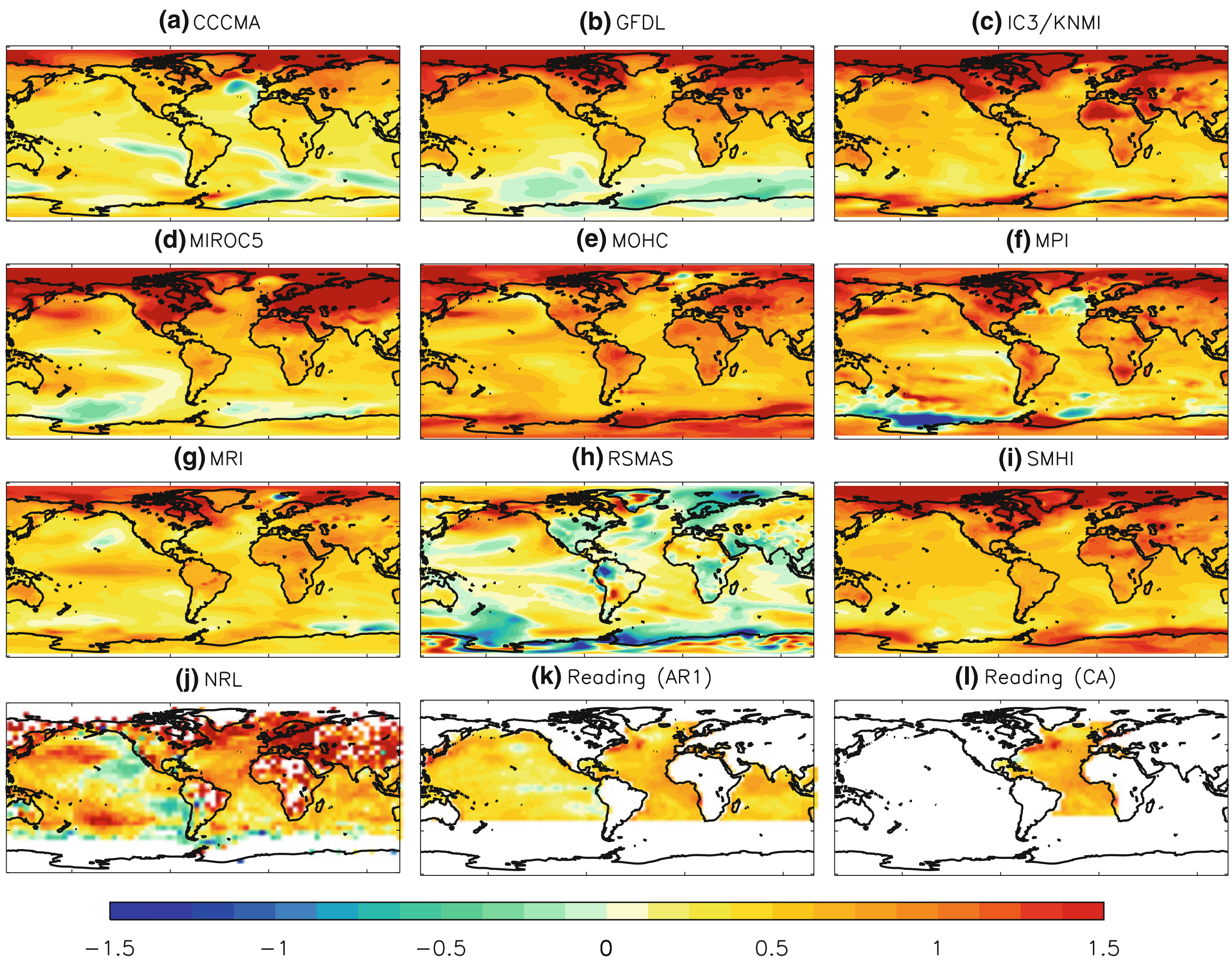

Fig. 4 Forecast temperature anomalies (as Fig. 3) for the 5-year period 2016–2020

predictions of an individual year are large beyond the first forecast year, dynamical models predict at least a 50 % chance of each year from 2013 onwards exceeding the current observed record of 0.4 °C above the 1971–2000 average (Table 3). This is similar to the forecast published by Smith et al. (2007) in which years from 2010 onwards were predicted to have a 50 % chance of exceeding the warmest year on record. If the recent hiatus in global warming is natural internal variability (Katsman and van Oldenborgh 2011; Meehl et al. 2011), current climate models (Knight et al. 2009) and the forecasts presented here suggest that it is unlikely to continue for many more years. Assessment of our forecast climate with actual observations, as they become available, will help determine whether other factors that are not well represented in current models prove to be important. These include enhanced anthropogenic aerosol emissions from coal burning, contributions from small volcanic eruptions, changes in solar irradiance, and trends in stratospheric water vapor (Lean and Rind 2009; Kaufmann et al. 2011; Solomon et al. 2010, 2011; Hansen et al. 2011). Either way, the ongoing assessment of global temperature forecasts and observations in the indefinite future promises to provide an important test of the forecasts and of our understanding and knowledge of the drivers of global warming.

Forecasts of other selected climate indices are presented in Fig. 8b–d. Dynamical models predict a slight warming of about 0.5 °C for the Niño3 region over the decade, with indices of Atlantic multidecadal variability (AMV) and Pacific decadal variability (PDV) showing no signal beyond their climatological distributions after 2015. The NRL empirical method (which inputs only future anthropogenic and solar influences) suggests that the AMV and PDV will continue in warm and cold phases, respectively, throughout the decade. However, forecasts for an individual year are uncertain, and these signals may be difficult to

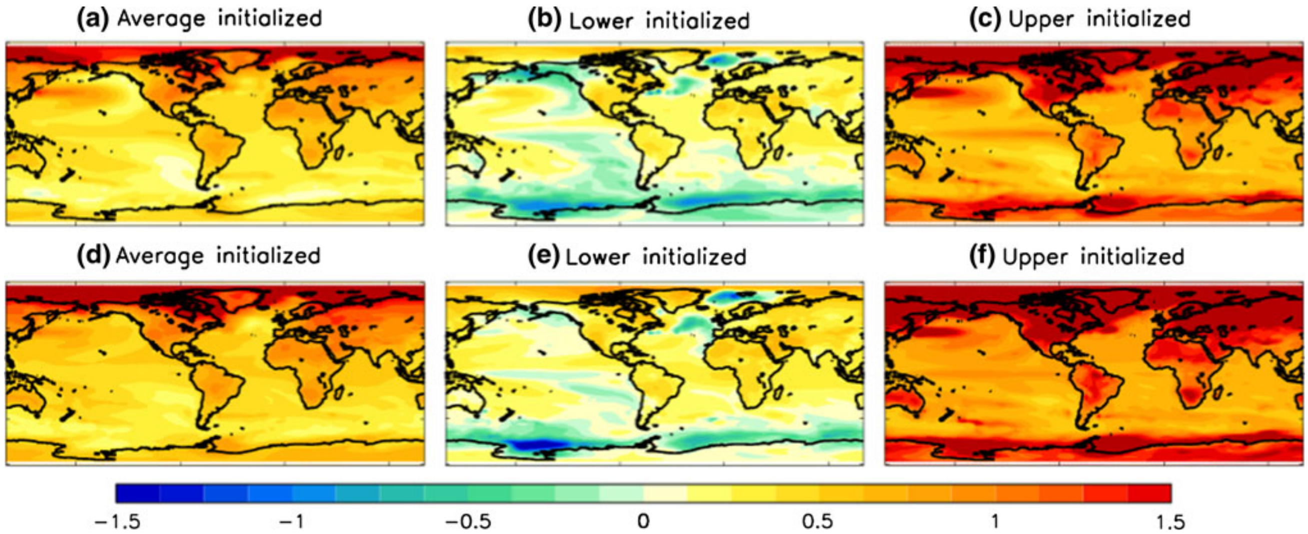

Fig. 5 Average initialized forecasts for the 5-year periods 2012–2016 (**a**–**c**) and 2016–2020 (**d**–**f**). The average, *lower* and *upper* values are diagnosed from the spread of the individual ensemble members (except GFDL since only ensemble means were provided), such that there is a 10 % chance of the observations being cooler than the *lower* (**b**, **e**), and a 10 % chance of the observations being warmer than the *upper* (**c**, **f**). Note that the actual anomaly patterns in the *lower* and *upper* maps are unlikely to occur since extreme fluctuations would not be expected at all locations simultaneously. The *upper* and *lower* confidence limits are diagnosed from the spread of the individual ensemble members assuming they are normally distributed and all members are equally likely

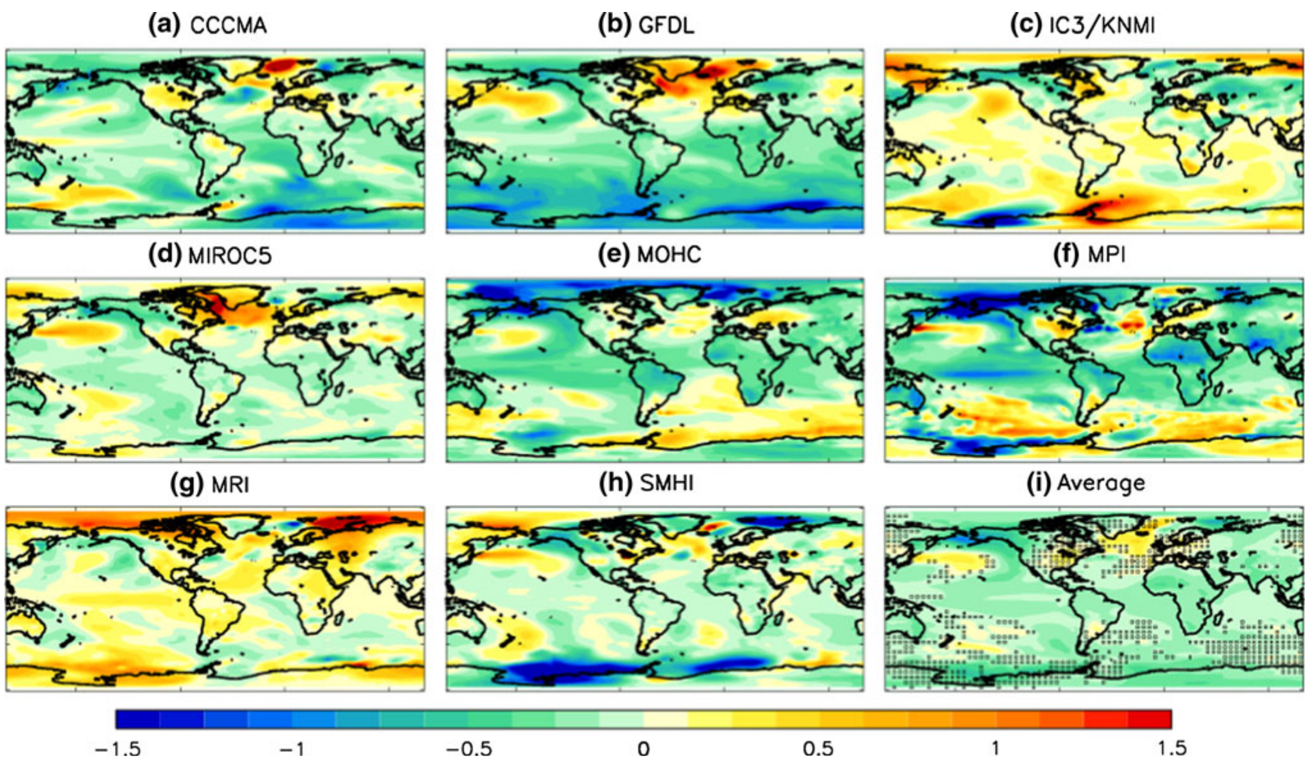

Fig. 6 Impact of initialization (initialized minus uninitialized ensemble means) on forecasts of the period 2012–2016. Unstippled regions in **i** indicate a 90 % or higher probability that differences between the initialized and uninitialized ensemble means did not occur by chance (based on a 2 tailed *t* test of differences between the two ensemble means assuming the ensembles are normally distributed)

distinguish amongst large inter-annual variability. Initialization has little impact on these indices beyond the first 3 years.

The lack of impact of initialization on AMV forecasts is somewhat surprising given that other studies show improved skill through initialization in the north Atlantic

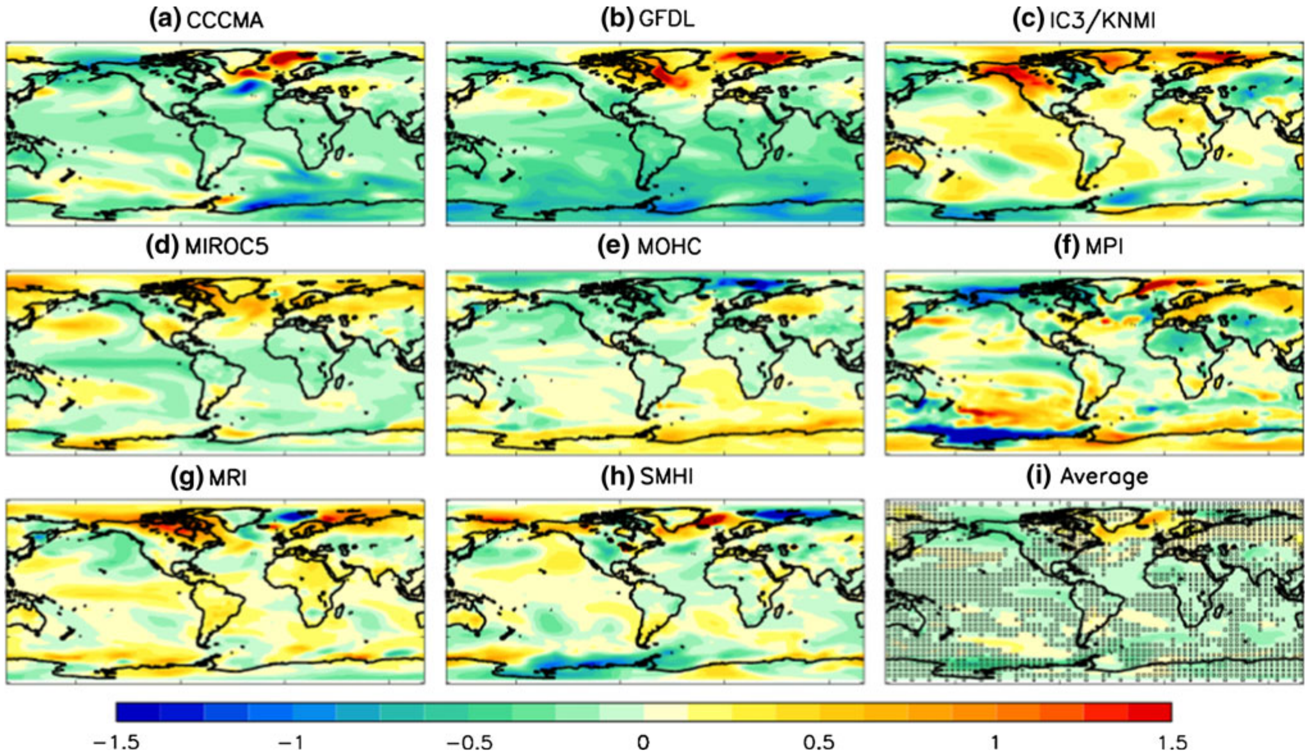

Fig. 7 Impact of initialization (as Fig. 6) on forecasts of the period 2016–2020

(Keenlyside et al. 2008; Pohlmann et al. 2009; Smith et al. 2010; van Oldenborgh et al. 2012; García-Serrano and Doblas-Reyes 2012; Matei et al. 2012). We investigated this ambiguity further by computing for each grid point the number of years in which the initialized ensemble mean is significantly different to the uninitialized ensemble mean (Fig. 9a). This does indeed show a maximum, of 7–10 years, in the north Atlantic but it is confined to the sub-polar gyre region (consistent with Fig. 7) and is therefore not included in the AMV index used here (which is based on latitudes south of 60°N).

To further investigate the impact of initialization, we construct an observed estimate of decorrelation timescale, computed as the number of years for which absolute lagged correlations of detrended sea surface temperatures are statistically significant. This can be thought of as the number of years for which an observation at a particular location would influence a forecast at that location based on statistical regression relationships. It is an imperfect estimate because although detrending will remove most of the influences of slowly increasing concentrations of greenhouse gases, it will not remove all of the impacts of higher frequency external forcing from volcanic eruptions, solar variability and anthropogenic aerosols. Nevertheless, this exercise also highlights the sub-polar gyre as a region with relatively large (10 years or longer) decorrelation times (Fig. 9b). Interestingly, this decorrelation time shows other similarities with the actual impact of initialization (Fig. 9a), with relatively large values in the tropical Atlantic and Indian Ocean. In many places, the decorrelation time is longer than the impact time, perhaps suggesting a potential for greater impact of initialization to be achieved in the future through a combination of improved initialization techniques, better models and more observations. However, the impact time is larger than the decorrelation time in some regions, notably the eastern tropical and northern Pacific. Further investigation of these similarities and differences promise to be instructive, and remain for future work.

4 Summary and conclusions

The growing need for decadal climate predictions is recognized by the inclusion of a protocol for historical tests in the latest coupled model inter-comparison project (CMIP5, Taylor et al. 2012), which will inform the upcoming IPCC fifth assessment report. The focus of those experiments is the historical period in order to assess the expected skill of decadal forecasts. However, many forecasting centers that have developed the capability to perform the CMIP5 historical tests are also making actual decadal climate forecasts in near real-time. These forecasts have been collated in a "Decadal Exchange" as a research exercise, aimed at

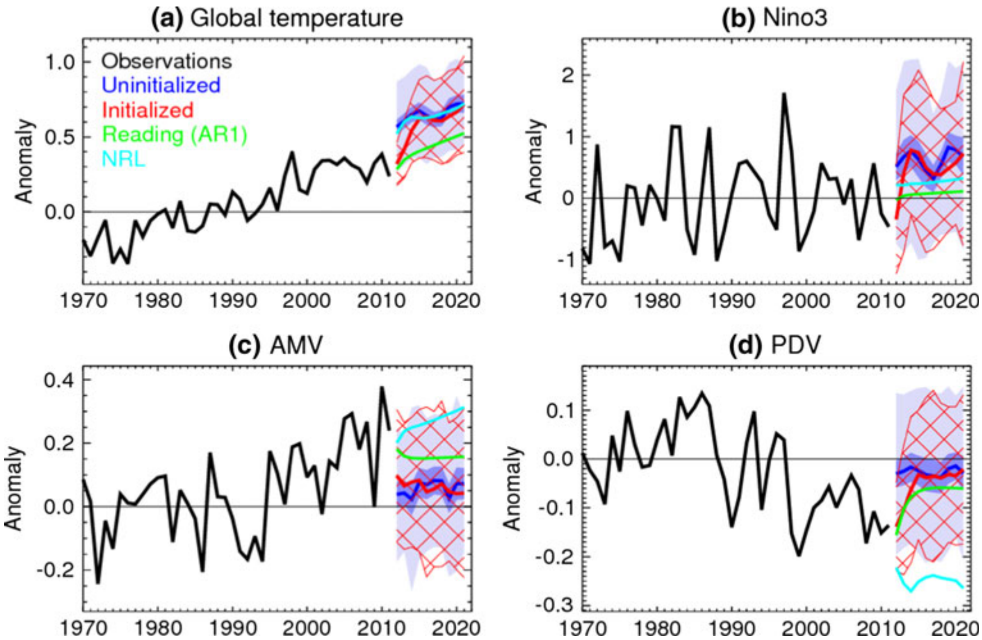

Fig. 8 Forecasts of selected climate indices. **a** Globally averaged temperature. **b** Temperature averaged over the Niño3 region of the tropical Pacific (150–90°W, 5°S–5°N). **c** Atlantic multi-decadal variability (AMV), computed as the north Atlantic (80°–0°W, 0°–60°N) minus globally averaged temperature between 60°S and 60°N (Trenberth and Shea 2006). **d** Pacific decadal variability (PDV), computed as the projection of the forecasts onto the first EOF of detrended sea surface temperature in the north Pacific (110°E–100°W, 20°–90°N) from HADISST (Rayner et al. 2003) over the period 1900–2001. The *dark blue curves* and *light blue shading* show the mean and 5–95 % confidence interval diagnosed from the individual members of the uninitialized dynamical model ensemble (see Table 1). The *red curves* and hatching show the equivalent initialized forecasts starting between 1st September 2011 and 1st January 2012. The *dark blue shading* shows the 5–95 % confidence interval where differences between initialized and uninitialized ensemble means could have occurred by chance. Note that GFDL forecasts are omitted from these plots because only ensemble means were provided. All anomalies are annual means relative to the average of the period 1971–2000. Observations (*black curves*) are taken from HadCRUT3 (Brohan et al. 2006)

Table 3 Annual globally averaged temperature anomaly (as in Fig. 8a) from dynamical model and empirical forecasts starting between 1st September 2011 and 1st January 2012

Years	Initialized			Uninitialized			NRL	Reading (AR1)
	Mean	Lower	Upper	Mean	Lower	Upper		
2012	0.32	0.18	0.46	**0.57**	0.27	0.87	**0.52**	0.28
2013	**0.42**	0.22	0.61	**0.61**	0.31	0.90	**0.60**	0.35
2014	**0.54**	0.33	0.76	**0.65**	0.35	0.95	**0.64**	0.38
2015	**0.62**	0.36	0.88	**0.67**	0.36	0.98	**0.63**	**0.40**
2016	**0.63**	0.36	0.91	**0.64**	0.32	0.96	**0.63**	**0.43**
2017	**0.61**	0.34	0.88	**0.62**	0.36	0.87	**0.64**	**0.45**
2018	**0.61**	0.32	0.90	**0.63**	0.37	0.89	**0.66**	**0.47**
2019	**0.64**	0.33	0.94	**0.70**	0.41	1.00	**0.67**	**0.49**
2020	**0.67**	0.38	0.96	**0.72**	0.42	1.02	**0.69**	**0.51**
2021	**0.72**	0.40	1.04	**0.72**	0.45	0.99	**0.72**	**0.53**

The upper and lower values represent the 5–95 % confidence interval diagnosed from the dynamical models ensemble spread. Anomalies are degrees centigrade relative to the average of the period 1971–2000. Forecast mean values highlighted in bold indicate a greater than 50 % chance of exceeding the warmest observed value in the HadCRUT3 dataset (of 0.40 °C occurring in 1998). Initialized mean values that are significantly cooler than the uninitialized means (at the 5 % level based on a one sided *t* test) are underlined

assessing and understanding differences and similarities between the forecasts, identifying a consensus view in order to prevent over-confidence in a single model, and establishing current collective capability.

Two exchanges have taken place so far: of forecasts starting on or before 1st January 2011, and on or before 1st January 2012. Each exchange consists of forecasts from up to 9 dynamical climate models and 3 empirical techniques. A potentially important aspect of decadal climate model predictions is that they are initialized with observations of the current state of the climate system. In order to assess the impact of initialization, uninitialized simulations of the effects of external forcing factors have also been exchanged. Details of the forecasts are presented here, both to document this activity, and to provide a focus on the coming decade. We anticipate that this activity will be ongoing, with participating groups continuing to exchange annual climate forecasts, followed by validation and assessment to inform subsequent forecasts.

Forecast temperature anomalies for 2011 agree well with observations (with an initialized multi-model mean

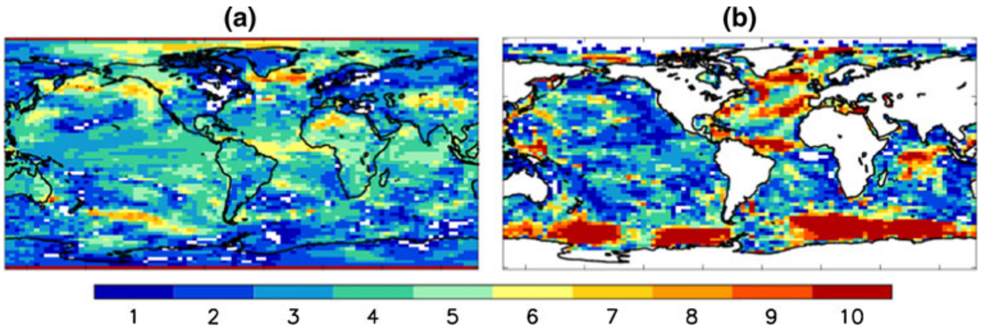

Fig. 9 Timescales of initialization impact. **a** The number of years for which there is at least a 90 % chance that the initialized and uninitialized ensemble means are different. The plot shows the average of the 2011 and 2012 exchanges. **b** Observed decorrelation time, computed as the number of years for which there is at least a 90 % chance that lagged correlations of detrended sea surface temperatures did not occur randomly. This was computed from annual mean HadISST data (Rayner et al. 2003) over the period 1891–2011

pattern correlation of 0.62, compared to 0.31 for uninitialized simulations). The initialized forecast captures the La Niña temperature pattern in the Pacific, together with a warm north Atlantic sub-polar gyre and tropical Atlantic and warm conditions in the southern Atlantic and Indian Oceans. The forecast also correctly captures cool conditions over Australia and Alaska, and very warm conditions over the USA. Observed cool conditions in central Eurasia were not predicted in the ensemble mean, but were captured by the ensemble spread. Forecast temperatures for 2012 show a similar pattern, with a continuing cool Alaska and warm USA, but with a reduced La Niña and strengthened tropical Atlantic dipole.

Considering the 5-year periods 2012–2016 and 2016–2020 and assuming no future volcanic eruptions, the dynamical models predict with 90 % probability that temperatures will be warmer than the 1971–2000 mean for nearly all land regions (the main exception being Alaska for 2012–2016), the tropical Atlantic, Indian and western Pacific oceans. Forecasts for the Southern Ocean are uncertain, with some models predicting cool anomalies for the entire decade.

Dynamical model predictions of AMV and PDV show no signal beyond climatological distributions after 2015. However, the NRL empirical method suggests that AMV and PDV will continue in positive and negative phases, respectively, throughout the decade. Dynamical models also predict Niño3 to warm slightly, by around 0.5 °C, over the decade. Uncertainties for an individual year are large, however, and these signals may be obscured by large inter-annual variability.

In most regions, and for indices of global temperature, Niño3, AMV and PDV, initialization has little impact beyond the first 4 years. However, initialization does have a prolonged impact in the north Atlantic sub-polar gyre and high latitude north Pacific, with the initialized forecast significantly warmer and cooler, respectively, than the uninitialized ensemble mean throughout the decade. Outside the sub-polar gyre, initialization cools the forecast almost everywhere, especially for 2012–2016. This includes most land regions, apart from the USA and Western Europe. Although uncertainties in forecasts for each individual year are large, significant cooling of the ensemble mean through initialization is expected to alter the predicted probabilities of extreme events (Hamilton et al. 2012; Eade et al. 2012), although this has not been quantified since only annual mean data have been analysed.

The cooling impact of initialization is consistent with the recent hiatus in global warming (e.g. Easterling and Wehner 2009), since initialized forecasts start from anomalously cool conditions relative to the expected warming from greenhouse gases. Indeed, initialized forecasts of globally averaged temperature are significantly cooler than uninitialized ones until 2015. However, in the absence of significant volcanic eruptions, global mean temperature is predicted to continue to rise. Assuming all ensemble members to be equally likely, the dynamical models predict at least a 50 % chance of each year from 2013 onwards exceeding the current observed record.

In future the exchange of decadal forecasts will be extended to include more models, and other variables, including rainfall. We reiterate that the forecasts presented here are experimental. Decadal climate prediction is immature, and uncertainties in future forcings, model responses to forcings, or initialization shocks could easily cause large errors in forecasts. Nevertheless, we believe it is important to publish such forecasts so that they may be evaluated openly as verifying observations become available. Such evaluation is likely to provide important insights into the performance of climate models and our understanding and knowledge of the drivers of climate change as climate science progresses.

Acknowledgments This work was supported by the joint DECC/Defra Met Office Hadley Centre Climate Programme (GA01101), and the EU FP7 THOR and COMBINE projects. MPI work was supported by the BMBF Nordatlantik and MiKlip projects. Ben Kirtman was supported by NOAA grants NA10OAR4320143 and NA10OAR4310203. IC3 work was supported by the EU-funded QWeCI (FP7-ENV-2009-1-243964) and CLIM-RUN projects (FP7-ENV-2010-265192), the MICINN-funded RUCSS (CGL2010-20657) projects and the Catalan Government. The IC3 authors thankfully acknowledge the computer resources, technical expertise and assistance provided by the Red Española de Supercomputación (RES), the European Centre for Medium-Range Weather Forecasts (ECMWF) under the special project SPESICCF, and Muhammad Asif for his invaluable support in running the experiments. NASA supported J. Lean. MIROC work was supported by the KAKUSHIN program of the Ministry of Education, Culture, Sports, Science, and Technology, Japan. The Earth Simulator of JAMSTEC was employed to perform MIROC experiments. MRI work was supported by the Japan Meteorological Agency research program and partly by the KAKUSHIN Program.

References

Boer GJ, Kharin VV, Merryfield WJ (2012) Decadal predictability and forecast skill. Clim Dyn (submitted)

Brohan P, Kennedy J, Harris I, Tett SFB, Jones PD (2006) Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. J Geophys Res 111: D12106

Chikamoto Y, Kimoto M, Ishii M, Mochizuki T, Sakamoto TT, Tatebe H, Komuro Y, Watanabe M, Nozawa T, Shiogama H, Mori M, Yasunaka S, Imada Y (2012) An overview of decadal climate predictability in a multi-model ensemble by climate model MIROC. Clim Dyn. doi:10.1007/s00382-012-1351-y

Collins WD et al (2006) The community climate system model version 3 (CCSM3). J Clim 19:2122–2143

DelSole T, Yang X, Tippett MK (2012) Is unequal weighting significantly better than equal weighting for multi-model forecasting? Q J R Meteorol Soc. doi:10.1002/qj.1961

Delworth TL, Zhang R, Mann ME (2007) Decadal to centennial variability of the Atlantic from observations and models. In: Ocean circulation: mechanisms and impacts. Geophysical monograph series 173. American Geophysical Union, Washington, DC, pp 131–148

Du H, Doblas-Reyes FJ, García-Serrano J, Guemas V, Soufflet Y, Wouters B (2012) Sensitivity of decadal predictions to the initial atmospheric and oceanic perturbations. Clim Dyn. doi:10.1007/s00382-011-1285-9

Dunstone NJ, Smith DM, Eade R (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

Eade R, Hamilton E, Smith DM, Graham RJ, Scaife AA (2012) Forecasting the number of extreme daily events out to a decade ahead. J Geophys Res 117:D21110. doi:10.1029/2012JD018015

Easterling DR, Wehner MF (2009) Is the climate warming or cooling? Geophys Res Lett 36:L08706

Fyfe JC, Merryfield WJ, Kharin V, Boer GJ, Lee W-S, von Salzen K (2011) Skillful predictions of decadal trends in global mean surface temperature. Geophys Res Lett 38:L22801. doi:10.1029/2011GL049508

García-Serrano J, Doblas-Reyes FJ (2012) On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. Clim Dyn. doi:10.1007/s00382-012-1413-1

Goddard L, Kumar A, Solomon A, Smith D, Boer G, Gonzalez P, Kharin V, Merryfield W, Deser C, Mason S, Kirtman B, Msadek R, Sutton R, Hawkins E, Fricker T, Hegerl G, Ferro C, Stephenson D, Meehl GA, Stockdale T, Burgman R, Greene A, Kushnir Y, Newman M, Carton J, Fukumori I, Delworth T (2012) A verification framework for interannual-to-decadal predictions experiments. Clim Dyn. doi:10.1007/s00382-012-1481-2

Hamilton E, Eade R, Graham RJ, Scaife AA, Smith DM, Maidens A, MacLachlan C (2012) Forecasting the number of extreme daily events on seasonal timescales. J Geophys Res 117:D03114. doi:10.1029/2011JD016541

Hansen J, Sato M, Kharecha P, von Schuckmann K (2011) Earth's energy imbalance and implications. Atmos Chem Phys 11: 13421–13449. doi:10.5194/acp-11-13421-2011

Hawkins E, Robson J, Sutton R, Smith D, Keenlyside N (2011) Evaluating the potential for statistical decadal predictions of sea surface temperatures with a perfect model approach. Clim Dyn 37:2459–2509. doi:10.1007/s00382-011-1023-3

Hazeleger W et al (2010) EC-earth: a seamless earth-system prediction approach in action. Bull Am Meteorol Soc 91: 1357–1363. doi:10.1175/2010BAMS2877.1

Hazeleger W et al (2012) EC-earth V2.2: description and validation of a new seamless earth system prediction model. Clim Dyn. doi:10.1007/s00382-011-1228-5

Ho CH, Hawkins E, Shaffrey L, Underwood F (2012) Statistical decadal predictions for sea surface temperatures: a benchmark for dynamical GCM predictions. Clim Dyn. doi:10.1007/s00382-012-1531-9

Hoerling M et al (2011) On North American decadal climate for 2011–20. J Clim 24:4519–4528. doi:10.1175/2011JCLI4137.1

ICPO (International CLIVAR Project Office) (2011) Data and bias correction for decadal climate predictions. International CLIVAR Project Office, CLIVAR Publication Series No. 150, 6 pp. Available from http://eprints.soton.ac.uk/171975/1/150_Bias_Correction.pdf

IPCC (2007) Climate change 2007: the physical science basis. In: Solomon S et al (eds) Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Katsman CA, van Oldenborgh GJ (2011) Tracing the upper ocean's 'missing heat'. Geophys Res Lett 38:L14610

Kaufmann RK, Kauppib H, Mann ML, Stock JH (2011) Reconciling anthropogenic climate change with observed temperature 1998–2008. Proc Natl Acad Sci USA 108:11790–11793

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

Khasnis A, Nettleman MD (2005) Global warming and infectious disease. Arch Med Res 36:689–696. doi:10.1016/y.aremed.2005.03.041

Kim H-M, Webster PJ, Curry JA (2012) Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. Geophys Res Lett 39:L10701. doi:10.1029/2012GL051644

Kirtman BP, Min D (2009) Multi-model ensemble ENSO prediction with CCSM and CFS. Mon Weather Rev. doi:10.1175/2009MWR2672.1

Knight JR, Allan RJ, Folland CK, Vellinga M, Mann ME (2005) A signature of persistent natural thermohaline circulation cycles in observed climate. Geophys Res Lett 32:L20708. doi:10.1029/2005GL024233

Knight JR, Folland CK, Scaife AA (2006) Climatic impacts of the Atlantic Multidecadal Oscillation. Geophys Res Lett 33:L17706. doi:10.1029/2006GL026242

Knight JR, Kennedy J, Folland C, Harris G, Jones GS, Palmer M, Parker D, Scaife A, Stott P (2009) Do global temperature trends over the last decade falsify climate predictions? Bull Am Meteorol Soc 90:S1–S196

Kopp G, Lean JL (2011) A new low value of total solar irradiance: evidence and climate significance. Geophys Res Lett 38:L01706. doi:10.1029/2010GL045777

Lean J, Rind D (2008) How natural and anthropogenic influences alter global and regional surface temperatures: 1889 to 2006. Geophys Res Lett 35:L18701. doi:10.1029/2008GL034864

Lean JL, Rind DH (2009) How will earth's surface temperature change in future decades? Geophys Res Lett 36:L15708. doi:10.1029/2009GL038932

Matei D, Pohlmann H, Jungclaus J, Müller WA, Haak H, Marotzke J (2012) Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model. J Clim. doi:10.1175/JCLI-D-11-00633.1

Meehl GA, Stocker TF, Collins W, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao ZC (2007) Global climate projections. In: Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Meehl GA, Goddard L, Murphy J, Stouffer RJ, Boer G, Danabasoglu G, Dixon K, Giorgetta MA, Greene A, Hawkins E, Hegerl G, Karoly D, Keenlyside N, Kimoto M, Kirtman B, Navarra A, Pulwarty R, Smith D, Stammer D, Stockdale T (2009) Decadal prediction: can it be skillful? Bull Am Meteorol Soc 90: 1467–1485

Meehl GA, Arblaster JM, Fasullo JT, Hu A, Trenberth KE (2011) Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. Nat Clim Chang 1:360–364. doi:10.1038/NCLIMATE1229

Meinshausen M, Smith SJ, Calvin K, Daniel JS, Kainuma MLT, Lamarque JF, Matsumoto K, Montzka SA, Raper SCB, Riahi K, Thomson A, Velders GJM, van Vuuren DPP (2011) The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. Clim Chang 109(1–2):213–241

Mendelsohn R, Basist A, Dinar A, Kurukulasuriya P, Williams C (2007) What explains agricultural performance: climate normals or climate variance? Clim Chang 81:85–99. doi:10.1007/s10584-006-9186-3

Merryfield WJ, Lee W-S, Boer GJ, Kharin VV, Scinocca JF, Flato GM, Ajayamohan RS, Fyfe JC, Tang Y, Polavarapu S (2012) The Canadian seasonal to interannual prediction system. Part I: models and Initialization. Mon Weather Rev (submitted)

Mochizuki T, Ishiia M, Kimoto M, Chikamoto Y, Watanabe M, Nozawa T, Sakamoto TT, Shiogama H, Awaji T, Sugiura N, Toyoda T, Yasunaka S, Tatebe H, Mori M (2010) Pacific decadal oscillation hindcasts relevant to near-term climate prediction. Proc Natl Acad Sci 107:1833–1837

Müller WA, Baehr J, Haak H, Jungclaus JJ, Kröger J, Matei D, Notz D, Pohlmann H, von Storch JS, Marotzke J (2012) Forecast skill of multi-year seasonal means with the MPI decadal prediction system. Geophys Res Lett. doi:10.1029/2012GL053326

Pohlmann H, Jungclaus J, Köhl A, Stammer D, Marotzke J (2009) Initializing decadal climate predictions with the GECCO oceanic synthesis: effects on the North Atlantic. J Clim 22:3926–3938

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J Geophys Res 108:4407. doi:10.1029/2002JD002670

Smith DM, Cusack S, Colman AW, Folland CK, Harris GR, Murphy JM (2007) Improved surface temperature prediction for the coming decade from a global climate model. Science 317:796–799. doi:10.1126/science.1139540

Smith DM, Eade R, Dunstone NJ, Fereday D, Murphy JM, Pohlmann H, Scaife AA (2010) Skilful multi-year predictions of Atlantic hurricane frequency. Nat Geosci. doi:10.1038/ngeo1004

Solomon S et al (2010) Contributions of stratospheric water vapor to decadal changes in the rate of global warming. Science 327:1219–1223

Solomon S et al (2011) The persistently variable 'background' stratospheric aerosol layer and global climate change. Science 333:866–870

Sutton RT, Hodson DLR (2005) Atlantic Ocean forcing of North American and European summer climate. Science 309:115–118

Tatebe H, Ishii M, Mochizuki T, Chikamoto Y, Sakamoto TT, Komuro Y, Mori M, Yasunaka S, Watanabe M, Ogochi K, Suzuki T, Nishimura T, Kimoto M (2012) Initialization of the climate model MIROC for decadal prediction with hydographic data assimilation. J Meteorol Soc Jpn (Special issue on the recent development on climate models and future climate projections) 90A:275–294

Taylor KE, Stouffer RJ, Meehl GA (2012) An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc 92:485–498. doi:10.1175/BAMS-D-11-00094.1

Trenberth KE, Shea DJ (2006) Atlantic hurricanes and natural variability in 2005. Geophys Res Lett 33:L12704

van Oldenborgh GJ, Doblas-Reyes FJ, Wouters B, Hazeleger W (2012) Skill in the trend and internal variability in a multi-model decadal prediction ensemble. Clim Dyn 38(7):1263–1280. doi:10.1007/s00382-012-1313-4

Yukimoto S, Adachi Y, Hosaka M, Sakami T, Yoshimura H, Hirabara M, Tanaka TY, Shindo E, Tsujino H, Deushi M, Mizuta R, Yabu S, Obata A, Nakano H, Koshiro T, Ose T, Kitoh A (2012) A new global climate model of Meteorological Research Institute: MRI-CGCM3—model description and basic performance. J Meteorol Soc Jpn 90A:23–64. doi:10.2151/jmsj.2012-A02

Zhang R, Delworth TL (2006) Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. Geophys Res Lett 33:L17712. doi:10.1029/2006GL026267

Zhang S, Rosati A (2010) An inflated ensemble filter for ocean data assimilation with a biased coupled GCM. Mon Weather Rev 138:3905–3931

Zhang S, Harrison MJ, Rosati A, Wittenberg A (2007) System design and evaluation of coupled ensemble data assimilation for global oceanic studies. Mon Weather Rev 135:3541–3564