Decadal predictions with the HiGEM high resolution global coupled climate model: description and basic evaluation

Clim Dyn
DOI 10.1007/s00382-016-3075-x

Decadal predictions with the HiGEM high resolution global coupled climate model: description and basic evaluation

L. C. Shaffrey[1] · D. Hodson[1] · J. Robson[1] · D. P. Stevens[2] · E. Hawkins[1] · I. Polo[1] · I. Stevens[2] · R. T. Sutton[1] · G. Lister[1] · A. Iwi[3] · D. Smith[4] · A. Stephens[3]

Received: 10 August 2015 / Accepted: 6 March 2016

Abstract This paper describes the development and basic evaluation of decadal predictions produced using the HiGEM coupled climate model. HiGEM is a higher resolution version of the HadGEM1 Met Office Unified Model. The horizontal resolution in HiGEM has been increased to $1.25° \times 0.83°$ in longitude and latitude for the atmosphere, and $1/3° \times 1/3°$ globally for the ocean. The HiGEM decadal predictions are initialised using an anomaly assimilation scheme that relaxes anomalies of ocean temperature and salinity to observed anomalies. 10 year hindcasts are produced for 10 start dates (1960, 1965,..., 2000, 2005). To determine the relative contributions to prediction skill from initial conditions and external forcing, the HiGEM decadal predictions are compared to uninitialised HiGEM transient experiments. The HiGEM decadal predictions have substantial skill for predictions of annual mean surface air temperature and 100 m upper ocean temperature. For lead times up to 10 years, anomaly correlations (ACC) over large areas of the North Atlantic Ocean, the Western Pacific Ocean and the Indian Ocean exceed values of 0.6. Initialisation of the HiGEM decadal predictions significantly increases skill over regions of the Atlantic Ocean, the Maritime Continent and regions of the subtropical North and South Pacific Ocean. In particular, HiGEM produces skillful predictions of the North Atlantic subpolar gyre for up to 4 years lead time (with ACC > 0.7), which are significantly larger than the uninitialised HiGEM transient experiments.

Keywords Decadal prediction · Climate variability · High-resolution climate modelling

1 Introduction

Developing skillful and statistically reliable climate predictions on seasonal to decadal timescales is one of the grand challenges of climate science. Skillful seasonal to decadal predictions would have substantial socioeconomic benefits, informing investment across a wide range of economic sectors.

Over the past few years, substantial international effort has been spent on developing decadal prediction systems. It has been shown that decadal predictions have significant skill, particularly for surface temperature, over most of the globe (Smith et al. 2007; Doblas-Reyes et al. 2011; Kim et al. 2012; Matei et al. 2012; Oldenborgh et al. 2012; Hanlon et al. 2013). A substantial component of the skill of decadal predictions arises from capturing the warming trend of temperatures from changes in external forcing such as greenhouse gases and aerosols, and from changes in temperature induced by volcanic eruptions. However, initialising decadal prediction systems leads to a significant increase in skill over the North Atlantic, the Indian Ocean and Eastern Pacific (Pohlmann et al. 2009; Smith et al. 2010; Mochizuki 2012; Doblas-Reyes et al. 2013).

✉ L. C. Shaffrey
L.C.Shaffrey@reading.ac.uk

[1] National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading RG6 6BB, UK

[2] Centre for Ocean and Atmospheric Sciences, School of Mathematics, University of East Anglia, Norwich, UK

[3] British Atmospheric Data Centre, Rutherford Appleton Laboratory, Chilton, UK

[4] Met Office Hadley Centre for Climate Prediction and Research, Exeter, UK

Published online: 15 April 2016

There are a number of challenges in developing skillful seasonal to decadal predictions. These include difficulties in initialising the prediction system, due to the choice of assimilation scheme and the sparse and inhomogeneous observations of the climate system. Another critical challenge is that climate models are biased and may inadequately represent some important physical processes. This can impact on the evolution of climate predictions.

One key question is whether the skill and statistical reliability of decadal prediction systems can be improved by using climate models with an improved representation of the climate system. One way to improve the representation of the climate in climate models is to increase their resolution (Shaffrey et al. 2009; Jung et al. 2012). The climate models used in the CMIP5 decadal prediction experiment typically have resolutions of 100–300 km in the atmosphere and 50–150 km in the ocean. Increasing the resolution of climate models improves the representation of Northern Hemisphere stationary waves and ENSO (the El Nino Southern Oscillation; Shaffrey et al. 2009), the Tropical Pacific ocean (Roberts et al. 2009), the extratropical response to ENSO (Dawson et al. 2013), the Southeast Pacific stratocumulus regions (Toniazzo et al. 2010) and anticyclonic blocking (Scaife et al. 2011).

The more relevant question for decadal prediction is whether higher resolution has an impact on the representation of decadal variability in climate models. Hodson and Sutton (2012) and Kirtman et al. (2012) discussed how increases in resolution leads to changes in the representation of decadal variability in higher resolution climate model simulations. Higher resolution has also been identified as important for the representation of specific aspects of decadal climate variability, for example variations in the Agulhas current (Biastoch et al. 2008) and high latitude ocean biases (Menary et al. 2015). However, increased resolution should not be seen as a panacea for climate model biases e.g. Patricola et al. (2012) found that increased resolution did not improve the representation of SST biases in the Tropical Atlantic.

The question of whether using a higher-resolution climate model with a better representation of climate can lead to improvements in prediction skill is addressed in this study by performing decadal predictions using the higher resolution HiGEM coupled climate model (Shaffrey et al. 2009). The aims of the study are to:

(i) Provide a description of the HiGEM decadal prediction system.
(ii) Investigate the extent to which skillful predictions can be produced on interannual to decadal timescales.
(iii) Assess whether using a higher resolution climate model with an improved representation of the climate system leads to more skillful seasonal to decadal predictions.

In Sect. 2 the HiGEM decadal prediction system and experimental design are described. In Sect. 3 the skill of the HiGEM decadal predictions is evaluated and conclusions are given in Sect. 4.

2 Experimental design, model description and initialisation

2.1 Model description

The HiGEM high resolution coupled climate model (Shaffrey et al. 2009) is based on the HadGEM1 climate configuration of the Met Office Unified Model (Johns et al. 2006). In HiGEM, the horizontal resolution of the atmosphere has been increased from $1.875° \times 1.25°$ in longitude and latitude in HadGEM1 to $1.25° \times 0.83°$ in longitude and latitude (approximately 90 km in the mid-latitudes). In the ocean, the horizontal resolution is increased from $1° \times 1°$ ($1/3° \times 1°$ in the Tropics) to $1/3° \times 1/3°$ globally (approximately 30 km). The ocean resolution in HiGEM is considered to be an eddy-permitting resolution, which allows oceanic mesoscale eddies to be represented but not fully resolved. The vertical resolution of HiGEM is the same as that of HadGEM1, i.e. 38 levels in the atmosphere and 40 levels in the ocean. The horizontal resolution of the climate models used in CMIP5 is typically 100–300 km in the atmosphere and 50–150 km in the ocean. The horizontal resolution of HiGEM is therefore higher than the typical resolutions used in the CMIP5 climate models.

The physics parametrisations remain very similar to those used in HadGEM1. The main differences are that the time-step is reduced in the ocean to 15 min and in the atmosphere to 20 min. In the ocean, the Gent-McWilliams eddy parametrisation is not used since the partially resolved ocean eddies are capable of providing the eddy component of the heat transport (for more details see Shaffrey et al. 2009). Increasing the resolution in HiGEM generally leads to a reduction in SST biases (Shaffrey et al. 2009). In particular, there is a reduction in SST biases in the Tropical Pacific (Roberts et al. 2009), although in some regions (e.g. the Southern Ocean) SST biases increase. The configuration of HiGEM used here remains the same as that used in the study of Shaffrey et al. (2009).

2.2 Experimental design

The experimental design is based upon the protocol used for the CMIP5 decadal prediction experiment (IPCC AR5, 2013). 10-year ensemble hindcasts with four members are performed for start dates every 5 years from 1 Nov 1960 to 1 Nov 2005 (i.e. 1960, 1965, 1975,...,2000, 2005).

The methodology for initialising HiGEM is described in Sect. 2.2.2.

The skill that arises from initialising HiGEM (initial condition predictability) versus the skill that arises from changes in external forcing (boundary condition predictability) can be assessed by comparing the HiGEM decadal predictions with uninitialised historical climate experiments driven by external forcing only (NOASSIM transient experiments).

2.2.1 *Historical NOASSIM transient experiments*

A four member ensemble of historical NOASSIM experiments have been performed with HiGEM using the CMIP5 RCP Historical scenario from 1 Jan 1957 to 30 Dec 2005 and with the CMIP5 RCP4.5 scenario from 1 Jan 2006 to 30 Dec 2015. In the historical NOASSIM experiments, observed values of time-varying well-mixed greenhouse gases, emissions of aerosols (SO2, black carbon and biomass burning), the incoming solar radiation, volcanic forcing and ozone are prescribed. An annual cycle of land surface parameters is used in HiGEM, but no long-term trends in land surface parameters are prescribed to reflect land use changes.

Initial conditions for the historical NOASSIM HiGEM experiments are taken from four different consecutive days at the end of a 65-year HiGEM experiment with constant late 1950s external forcing. The forcing is derived from averaged 1955–1960 CMIP3 historical well-mixed greenhouse gases, emissions of aerosols (SO2, black carbon and biomass burning), incoming solar radiation, volcanic forcing and ozone. Although the late 1950s external forcing experiment used to generate the initial conditions for the NOASSIM HiGEM experiment is only 65 years in length, the long-term drifts in ocean temperatures below 500 m are small (not shown).

2.2.2 *Prediction initialisation and anomaly assimilation*

The HiGEM decadal predictions are initialised using an anomaly assimilation approach similar to that used in Smith et al. (2007). An assimilation experiment is performed where anomalies of potential temperature and salinity throughout the depth of the ocean are strongly relaxed back to the observed anomalies.

For potential temperature, T, the conservation equation becomes

$$\frac{\partial T}{\partial t} + \nabla.(\mathbf{v}T) = F^T - \frac{T' - T'_{obs}}{\tau}, \quad (1)$$

where $\mathbf{v}$ is the three-dimensional velocity field, F^T represents subgrid-scale processes, T' is the model anomaly of potential temperature and T'_{obs} is the observed anomaly

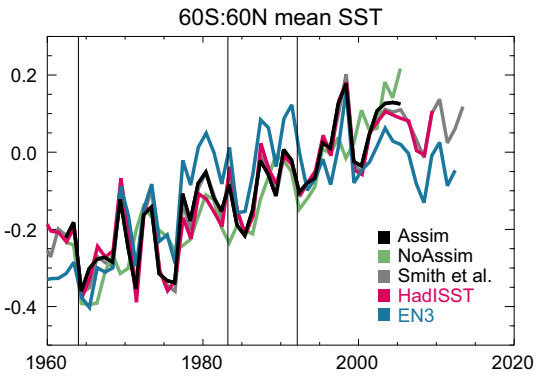

Fig. 1 Time-series of globally averaged SST anomalies (60S–60N, with respect to 1985–2006) from the HiGEM assimilation experiment (*black*) HadISST SST dataset (*red*), the EN3 ocean observation dataset (*light blue*), the analysis of Smith and Murphy (2007) (*grey*) and the ensemble mean of the four NOASSIM HiGEM transient experiments (*green*). Units: K. *Thin vertical lines mark* the timings of major volcanic eruptions: Agung (1963/4), El Chichon (1982) and Pinatubo (1991)

of potential temperature. The global relaxation timescale, τ, is chosen as 15 days. A similar equation with the same global value of τ is used for the relaxation of salinity. In the original HadCM3-based anomaly assimilation scheme a 6-h relaxation timescale was chosen (Smith et al. 2007). However, it was found in initial experiments with the eddy-permitting HiGEM climate model that such a short relaxation timescale overly constrained the ocean eddy field.

Model anomalies are determined by removing a seasonally varying 30-year climatology taken from the present day control integration described in Shaffrey et al. (2009) which uses 1990 external forcing. The observed ocean potential temperatures and salinities are taken from the ocean analysis of Smith and Murphy (2007). The observed anomalies are determined by removing by a seasonally varying 1980–2005 climatology. The periods chosen for the model and observed climatologies were used as they reflect periods of similar external forcing.

The HiGEM assimilation experiment is performed from Jan 1957 to Dec 2005 using the ocean relaxation as described above and the same Historical CMIP5 RCP forcing as used in the transient NOASSIM experiments (Sect. 2.2.1). Figure 1 shows the time-series of global SST anomalies (60S to 60N) from the HadISST dataset, the observational analysis of Smith and Murphy (2007), the ensemble mean of four transient NOASSIM experiments and from the assimilation experiment. There is very good agreement between the time-series of the HiGEM assimilation experiment and the observations. This indicates that the anomaly assimilation scheme in HiGEM is performing as expected.

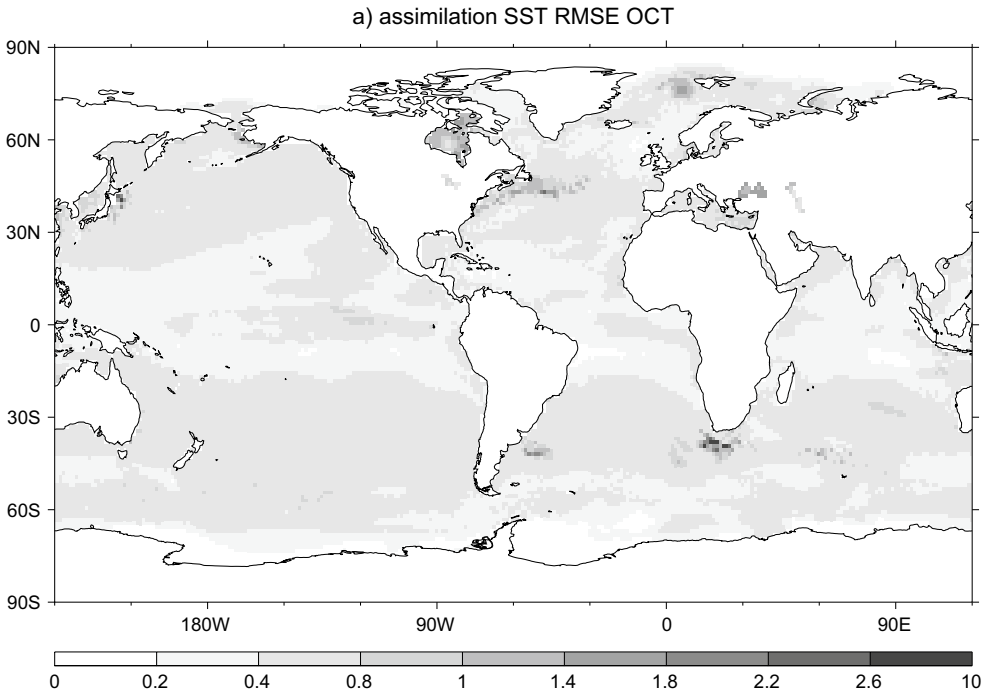

Fig. 2 RMS differences in October anomalous SST from the HiGEM assimilation experiment and ocean analysis SSTs from Smith and Murphy (2007). Units: K

The ensemble mean of the HiGEM NOASSIM experiments generally captures the observed long-term warming from 1960 to 2010. The NOASSIM experiments are also able to capture the periods of global cooling associated with volcanic eruptions in the mid 1960s (Agung), 1982 (El Chichon) and 1991 (Pinatubo).

After the assimilation experiment was performed, it was found that that an incorrect sulphate aerosol forcing had been used (where twice as much sulphate aerosol had been emitted compared to the RCP Historical scenario). As the ocean temperature and salinities are heavily constrained by the relaxation this does not strongly degrade the ability of the assimilation experiment to replicate the observed ocean anomalies (e.g. see Fig. 1).

The initial conditions used in the hindcast set considered here were created by performing a series of additional 1-year assimilation experiments with corrected sulphate aerosols emissions. These additional 1 year experiments begin 1 year before each start of the CMIP5 start date (e.g. 1 Nov 1964 for the 1 Nov 1965 start date) using the initial conditions from the original assimilation experiment. The differences in ocean temperatures and salinities between the corrected and uncorrected experiments are small since they are constrained by the anomaly assimilation. However, the additional year ensures that the initial condition for the decadal predictions have the correct sulphate aerosol loadings. This additional level of complexity in generating the initial conditions is not desirable, but unfortunately it was not possible to redo the entire assimilation experiment with corrected sulphate aerosol emissions due to the computational expense of running the high-resolution HiGEM model. The correct aerosol forcing was prescribed in the NOASSIM and HiGEM decadal prediction experiments.

Additional information about the performance of the anomaly assimilation scheme in HiGEM is provided in Fig. 2. Figure 2 shows spatial maps of the RMS (root-mean square) differences between the October anomalous SSTs from the HiGEM assimilation experiment with corrected aerosol emissions and the ocean analysis of Smith and Murphy (2007). RMS differences are typically less than 1 K except in areas of high SST variability (for example, the Gulf Stream). This again suggests that the anomaly assimilation scheme in HiGEM is performing as expected.

2.3 Evaluating predictions and prediction biases

The presence of biases can significantly influence and complicate the estimation of the skill of a decadal prediction system (Robson 2010; Kharin et al. 2012; Goddard et al. 2013). Figure 3 shows that there are lead-time dependent prediction biases in the HiGEM decadal predictions for SST anomalies. SST anomaly biases are generally small (within 0.75 K). There are systematic cold biases in the North Pacific and warm biases in the

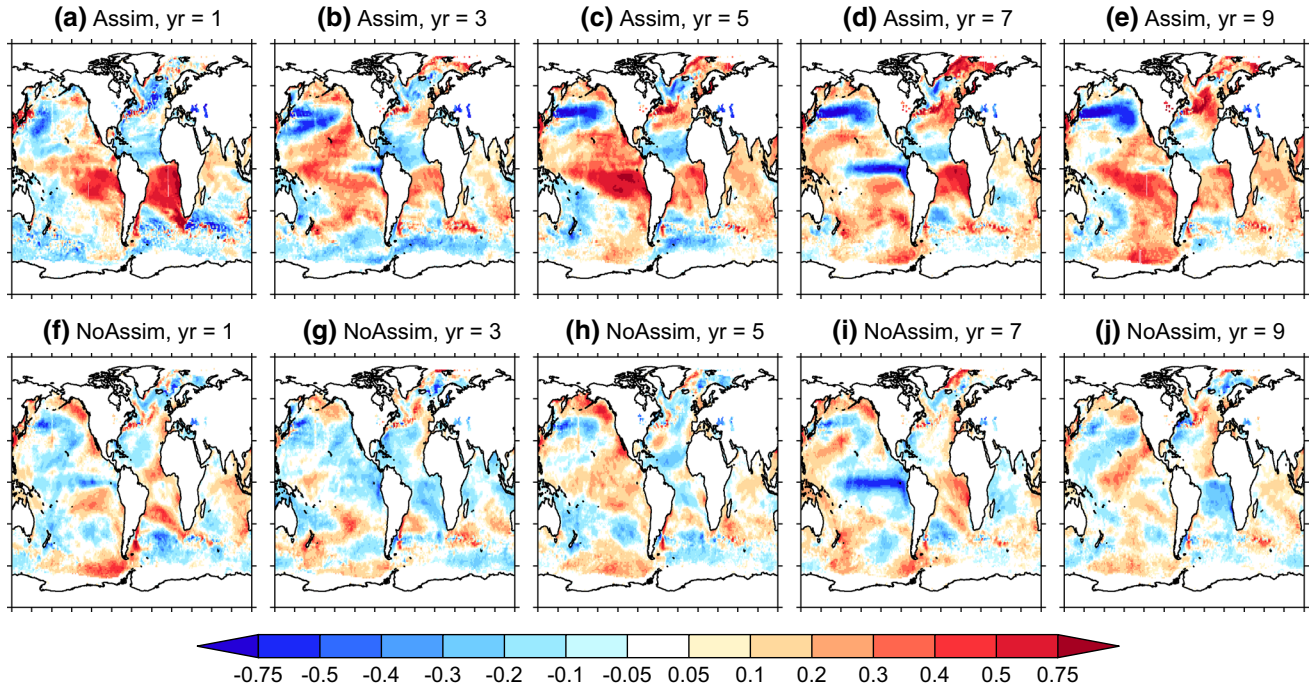

Fig. 3 Annual mean SST anomaly biases from **a**–**e** the HiGEM Decadal Predictions evaluated using the analysed SSTs of Smith and Murphy (2007) for lead times of 1, 3, 5, 7 and 9 years and **f**–**j** the same but for the NOASSIM HiGEM transient experiments. Ensemble mean SST anomaly biases are averaged across each of 10 start dates. Units K

Southeastern Pacific and Southeastern Atlantic, which may be related to climatological SST biases in uninitialised control HiGEM experiments (Shaffrey et al. 2009). There is also some indication that the prediction biases may not be well sampled in some regions. For example, the sign of the prediction bias varies from year to year in the Tropical Pacific.

Given that lead-time dependent biases exist in the HiGEM decadal predictions, the anomaly correlation skill score (ACC) is primarily used to evaluate skill as it is inherently insensitive to mean bias corrections (MBC). Other skill scores are sensitive to the exact definition of bias removal. This includes the root-mean square error (RMSE; Robson 2010) and the mean squared skill score (MSSS). This sensitivity is due to all aspects of bias removal including the period over which bias is calculated, the climatologies used to calculate the anomalies, and also the definition of the bias to be removed (e.g. the bias due to forcing errors, sampling errors, or the 'true' model bias e.g. Hawkins et al. 2014). Additional evaluation of the HiGEM hindcasts using the MSSS skill score, with analysis of the sensitivity of MSSS to bias removal, are provided in the Appendix.

To understand the impact of the ocean initialisation on skill we compare the ACC of the HiGEM decadal predictions, with the ACC from the HiGEM NOASSIM transient experiments. We test the significance of the ACC difference in the 2D spatial maps similarly to Smith et al. (2013). For the purposes of significance testing, we create synthetic transient NOASSIM members by adding random errors to the ensemble-mean of the NOASSIM transient prediction. The random errors are generated by block-bootstrapping the prediction errors (i.e. prediction anomalies minus observed anomalies) of all HiGEM decadal prediction start dates in order to create an ensemble mean error at each grid-point independently. A block length of 5 years was used to preserve the multi-annual variability. The ensemble mean error is then used to perturb the NOASSIM transient ensemble mean in the ACC calculation. The resampling of the NOASSIM transient ACC is performed 3000 times to build a probability distribution function of differences in ACC. The differences are found to be significant if they are outside the 5–95 % percentile of the resampled NOASSIM distribution.

We also apply a simple lead-time dependent correction Δ_l from the hindcasts to enable a better visual comparison with observations in Figs. 7 and 11. This mean bias correction is computed as:

$$\Delta_l = \frac{1}{YN} \sum_{yn} (X_{ynl} - O_{yl}) \tag{2}$$

where X_{ynl} is the nth ensemble member hindcast initialised from the yth start date at the lth lead, N is the number of

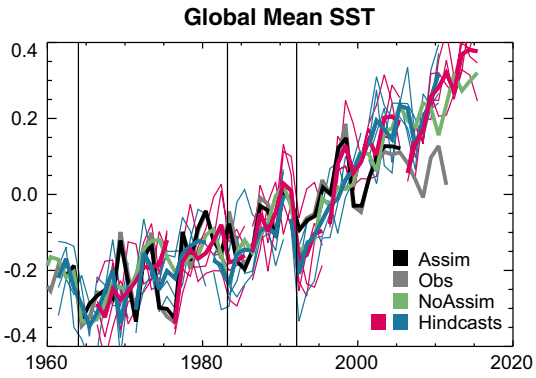

Fig. 4 Time-series of annual mean globally averaged SST anomalies (60S–60N; with respect to 1985–2006) from (*black*) the HiGEM assimilation experiment, (*grey*) the ocean analysis of Smith and Murphy (2007), (*green*) the ensemble mean of the four NOASSIM HiGEM transient experiments and (*thick red and blue*) HiGEM decadal predictions. Alternate start dates are shown *red* and *blue*. Ensemble mean predictions are indicated by *thick lines* and individual members by *thin lines*. Units: K

ensemble members, Y is the number of start dates, and O_{yl} is the corresponding observed value at the same time point.

3 An evaluation of the HiGEM decadal predictions

In this section, an evaluation of the skill of the HiGEM decadal predictions is presented. Section 3.1 focuses on evaluating prediction skill from a global perspective. Sections 3.2 and 3.3 focus on evaluating skill in the Atlantic and Pacific Oceans respectively.

3.1 Surface air temperature and upper ocean heat content

Figure 4 shows the time-series of observed global SST anomalies from Smith and Murphy (2007), the ensemble mean of the HiGEM NOASSIM transient experiments and the ensemble means of the HiGEM decadal predictions. As mentioned in Sect. 2, the ensemble mean of the HiGEM NOASSIM experiment is capable of capturing the observed warming from 1960 to 2005. Similarly the HiGEM decadal predictions can also capture the long-term warming trend and the periods of global cooling associated with volcanic eruptions. Both the HiGEM decadal predictions and the HiGEM NOASSIM transient experiments overestimate the very recent temperature trend from 2005 to 2014. The overestimation of the recent temperature trend is well documented behaviour of many climate model simulations (e.g. IPCC AR5, 2013).

3.1.1 Surface air temperature

Figure 5a–d show spatial maps of ACC for annual mean SAT (surface air temperature) for the HiGEM decadal predictions for different lead times. There is substantial skill in the HiGEM decadal predictions in predicting SAT across the different lead times (1-year, 2–3 years, 4–6 years and 7–10 years ahead). For 1-year lead times, anomaly correlations over large areas of the North Atlantic Ocean, the Western Pacific Ocean and the Indian Ocean exceed values of 0.6. For longer lead times, the skill over the Tropical Pacific decreases, but the skill of the HiGEM decadal predictions predictions increases over North America, Eurasia and Australia. The increases in skill arises from (i) the use of longer averaging periods thereby increasing the signal to noise ratio and (ii) from capturing the trend in SAT due to changes in external forcing. There is a notable lack of skill over the Eastern Pacific Ocean (see Sect. 3.3.1 for further details).

As discussed in Smith et al. (2007), a substantial proportion of the skill of decadal predictions for SAT arises since climate models are capable of reproducing the observed long-term warming trend when driven with the observed external forcing. To determine the relative contributions to prediction skill from initial conditions and external forcing, the HiGEM decadal predictions can be compared to the HiGEM NOASSIM transient experiments.

Figure 5e–h show spatial maps of the differences in the ACC between the HiGEM decadal predictions and the HiGEM NOASSIM experiments for annual mean SAT. Figure 5e–h generally show positive values indicating that the initialisation of the HiGEM decadal predictions increases prediction skill. For 1-year lead times, initialisation significantly increases skill over regions of the Atlantic Ocean, the Indian Ocean, the Maritime Continent and regions of the subtropical North and South Pacific Ocean. A significant increase in skill from initialisation can be seen over regions of the Atlantic and the subtropical North and South Pacific in years 2–3. Although there is substantial skill over the Atlantic and Pacific oceans, initialistion does not lead to a similar increase in skill over many land regions. However, there is a substantial and significant increase in skill from initialisation of the HiGEM decadal predictions over regions of the Atlantic Ocean for years 4–6 and years 7–10. The skill of the HiGEM decadal predictions in the Atlantic and Pacific Oceans will be considered in more detail in Sects. 3.2 and 3.3.

The levels of skill for SAT seen in the HiGEM decadal predictions appear to be qualitatively comparable to that seen in other decadal prediction systems (e.g. Smith et al. 2013; Chikamoto et al. 2013). One question raised in the introduction is whether a higher resolution coupled climate model with a better representation of climate is

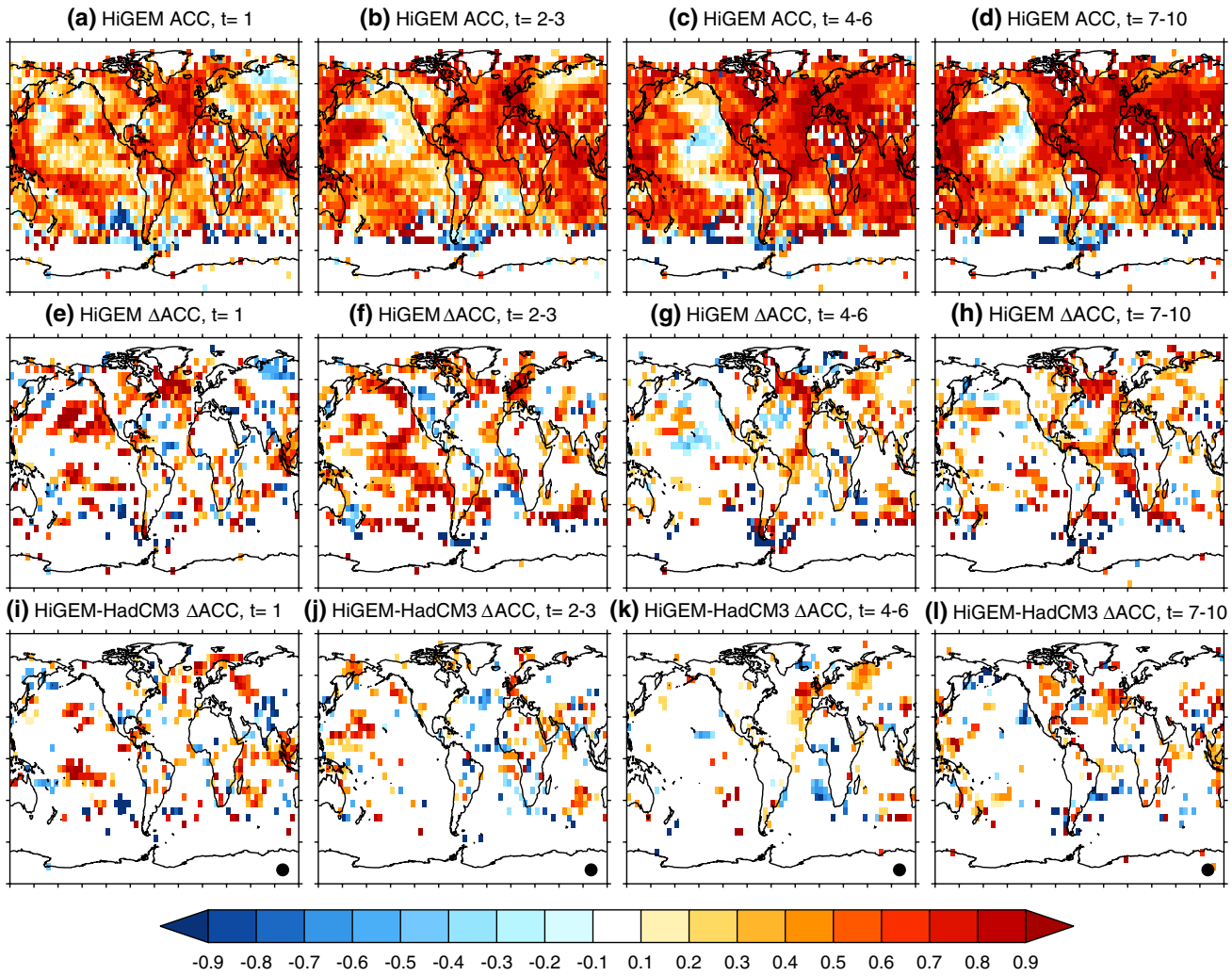

Fig. 5 Anomaly correlations of annual mean Surface Air Temperature from the HiGEM Decadal Predictions for lead times of **a** 1 year, **b** 2–3, **c** 4–6 and **d** 7–10 year. **e**–**h** Differences in anomaly correlations between the HiGEM Decadal Predictions and the HiGEM NOASSIM transient experiments calculated using a Fisher transform. **i**–**l** Differences in anomaly correlations between the HiGEM Decadal Predictions and the CMIP5 DePreSys anomaly assimilation decadal predictions (Smith et al. 2013) calculated using a Fisher transform. Only gridpoints where correlations are outside of the 5–95 % confidence levels are shown. The decadal predictions are evaluated on the native grid of the HadCRUT4 dataset. A *black circle* in the corner of plots, **i**–**l** indicates field significance (i.e that greater than 10 % of gridpoint are significant)

able to produce more skillful decadal predictions. A more quantitative comparison is presented in Fig. 5i–l, which shows the difference in ACC for the annual mean SAT predictions in the HiGEM decadal predictions minus the CMIP5 DePreSys decadal predictions that are based on the lower resolution HadCM3 model. The DePreSys decadal predictions are taken from the CMIP5 anomaly assimilation hindcast set described in Smith et al. (2013) that used the same historical forcings. Furthermore, the predictions are evaluated for the same start dates and for the same number of ensemble members. At lead times of 1 year, the HiGEM decadal predictions are significantly more skillful than DePreSys in parts of the North Atlantic, the Indian Ocean and the subtropical North and South Pacific, though less skillful over the Indian subcontinent. At longer lead times (i.e. years 2–3, 4–6 and 7–10), HiGEM appears to be significantly more skillful than DePreSys over the Eastern North Atlantic. This may reflect the improved representation of the North Hemisphere stationary wave pattern found in the HiGEM model compared to lower resolutions climate models such as HadCM3 and the other CMIP3 models (Woollings 2010; Catto et al. 2011) and CMIP5 models (Zappa et al. 2013).

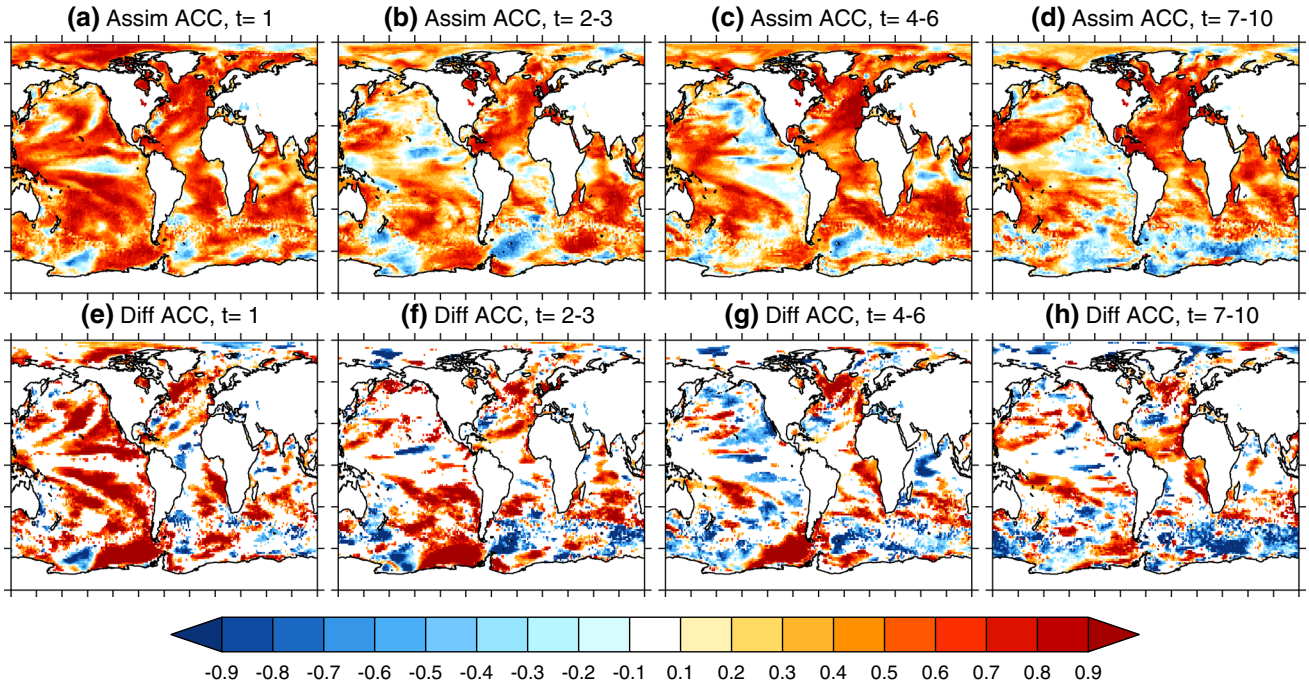

Fig. 6 Anomaly correlations of annual mean upper 100 m ocean temperatures from the HiGEM Decadal Predictions for lead times of **a** 1 year, **b** 2–3, **c** 4–6 and **d** 7–10 year. **e–h** Differences in anomaly correlations between the HiGEM Decadal Predictions and the HiGEM NOASSIM transient experiments calculated using a Fisher transform. Only gridpoints where correlations are outside of the 5–95 % confidence levels are shown. The HiGEM decadal predictions are evaluated on the native grid of the ocean analysis of Smith and Murphy (2007)

3.1.2 Upper 100 m ocean temperature

Another consideration is whether the skill in SAT can also be seen in the heat content of the upper ocean. Figure 6a–d show spatial maps of anomaly correlations for annual mean upper 100 m ocean temperature for the HiGEM decadal predictions. Again there is substantial skill in the HiGEM decadal predictions in predicting upper ocean temperature. The regions of substantial skill for upper ocean temperature generally correspond with those seen for SAT. Figure 5e–h show spatial maps of the differences in the anomaly correlations between the HiGEM decadal predictions and the HiGEM NOASSIM experiments for upper ocean temperature. The ACC skill for the upper 500 m ocean temperature was also investigated (not shown) and found to be similar to that for the upper 100 m ocean temperatures. Although there is general agreement between the skill of SAT and upper ocean temperature predictions, it is apparent that initialisation makes a larger contribution to the skill of the HiGEM decadal predictions for upper ocean temperature than for SAT.

3.2 The Atlantic multidecadal oscillation and the Atlantic meridional overturning circulation

In the previous section it was shown that HiGEM decadal predictions have substantial skill for SAT and upper 100 m ocean temperature in the North Atlantic on multi-annual timescales. It was also shown that the initialisation of the HiGEM decadal predictions significantly contributes to the prediction skill. In this section, three SST indices are used to investigate the origin of the skill of the HiGEM decadal predictions in more detail. These three indices are (i) an index of the Atlantic Multidecadal Oscillation (AMO: SSTs averaged between 0N–60N and 75W–7.5W), ii) an index of the North Atlantic Subpolar Gyre (SPG: SSTs averaged between 50N–65N and 75W–7.5W) and iii) an index of Tropical Atlantic SSTs (TA: averaged between 0N–20N and 75W–7.5W; Sutton and Hodson 2003). The AMO index is based on that from Sutton and Hodson (2005) but without any filtering applied, so that the AMO index used here includes interannual as well as decadal variability. In Sect. 3.2.2, the ability of the HiGEM decadal predictions to capture the evolution the Atlantic Meridional Overturning Circulation at 27N and 45N is evaluated.

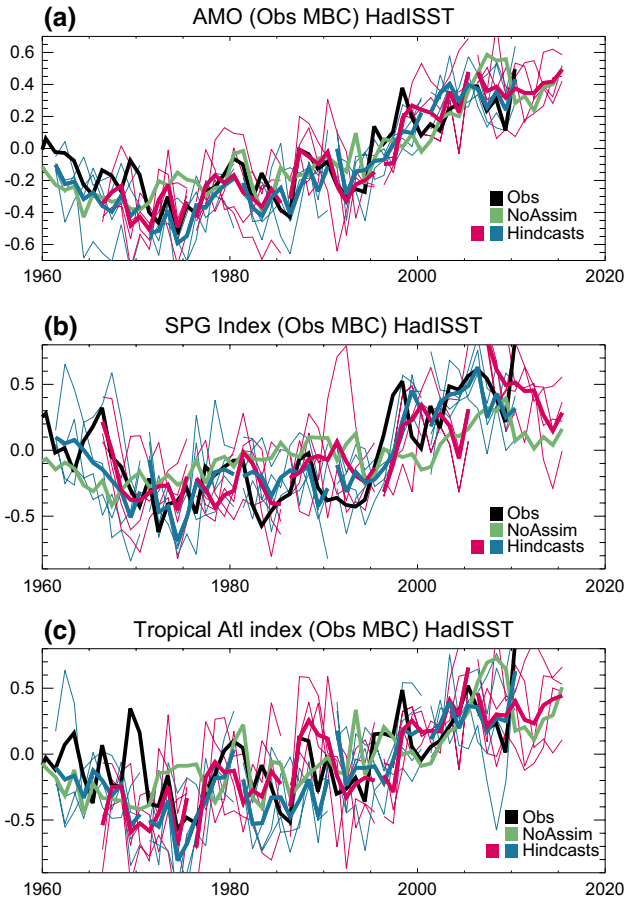

Fig. 7 Time-series of anomalies in **a** the annual mean Atlantic Multidecadal Oscillation index (SSTs averaged 0N–60N and 75W–7.5W), **b** the annual mean North Atlantic Subpolar Gyre index (SSTs averaged between 50N–65N and 75W–7.5W)) and **c** the annual mean Tropical Atlantic index (SSTs averaged between 0N–20N and 75W–7.5W). The *black line* is the observations (HadISST), the *green line* is the ensemble mean of the HiGEM NOASSIM transient experiments, *the thick red* and *blue line* sare the ensemble means of the HiGEM decadal predictions and the *thin red* and *blue lines* are the individual predictions. Units on the *y*-axis are Kelvin. The Hindcasts have been lead-time dependent corrected with respect to the observations (see Sect. 2.3 for more details)

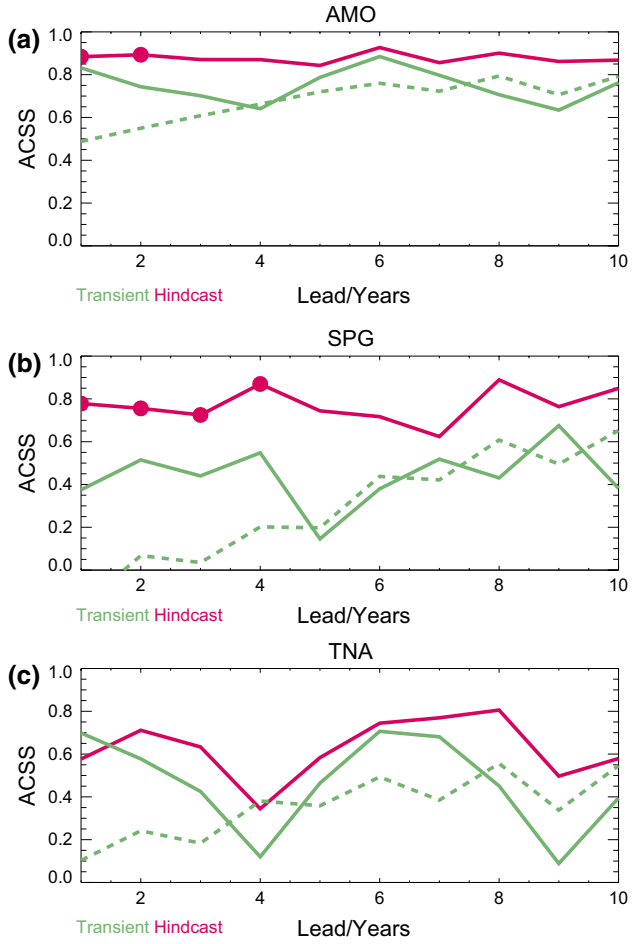

Fig. 8 Anomaly correlation as a function of lead time for the HiGEM decadal predictions (*red line*), the HiGEM NOASSIM transient experiments sampled for the same periods as the decadal predictions (*green line*) and HiGEM NOASSIM transient experiments sampled for 21 start dates (1960, 1962, 1964,...2000; *green dashed line*). Anomaly correlations are shown for the a) AMO, **b** SPG and **c** TA SST indices, as defined in Fig. 7. *Red circles* indicate where the skill of the HiGEM decadal predictions is significantly larger than the HiGEM NOASSIM experiments when sampled at 21 start dates at the 90 % level

3.2.1 The Atlantic multidecadal oscillation

Figure 7 shows time-series of the AMO, SPG and TA SST indices from the HadISST observations, the HiGEM NOASSIM transient experiments and the HiGEM decadal predictions. The HiGEM decadal predictions are capable of capturing the long term evolution of the AMO index, and in particular the observed cooling from 1960 to 1970, and warming from 1990 to 2010. The HiGEM NOASSIM experiments do not capture the cooling during the 1960s to the same extent as the HiGEM decadal predictions.

The HiGEM decadal predictions are also able to capture some of the rapid changes observed in the SPG, for example the rapid cooling in mid-1960s and the rapid warming in the mid-1990s. In contrast, the HiGEM decadal predictions are not able to capture much of the interannual variation in the TA SSTs, although both the HiGEM decadal predictions and the HiGEM NOASSIM experiments are capable of capturing the long-term warming trend.

Figure 8 shows the ACC for the AMO, SPG and TA SST indices as a function of lead time for the HiGEM decadal predictions. To assess the contribution of

initialisation to prediction skill, the ACC from the HiGEM NOASSIM experiments is also shown. To assess the sampling uncertainty from using only 10 evenly spaced start dates, the ACC is shown for the HiGEM NOASSIM experiments when sampled for same time periods as the HiGEM decadal predictions and also shown when sampled for 21 start dates. The differences between the two different sampling strategies for the HiGEM NOASSIM experiments suggest that sampling uncertainties can be substantial (see also Garcia-Serrano et al. 2014; Mignot et al. 2015).

Figure 8a shows that both the HiGEM decadal predictions can produce predictions of the AMO on multi-annual timescales with values of ACC greater than 0.8. For lead times of 1 and 2 years, the HiGEM decadal predictions are significantly more skillful at the 90 % significance level than the HiGEM NOASSIM transient experiments when sampled for 21 start dates.

Figure 8b shows the anomaly correlations for the SPG index. The HiGEM decadal predictions are also able to produce very skillful predictions of the SPG index on multi-annual timescales with values of ACC greater than 0.8. The skill in the HiGEM decadal predictions is significantly larger for years 1–4 at the 90 % significance level than that of the HiGEM NOASSIM experiments when sampled at 21 start dates. This suggests that the initialisation of the HiGEM decadal predictions leads to substantial and significant prediction skill for SST in the North Atlantic Subpolar gyre on multi-annual timescales (but is not significantly more skillful when sampled using only 10 start dates). Figure 8 also suggests that the skill of the HiGEM decadal predictions for the AMO mostly arises from capturing the observed evolution of the North Atlantic subpolar gyre. In contrast, the skill in both the HiGEM decadal predictions and the HiGEM NOASSIM experiments appears to be more modest for Tropical Atlantic SST.

3.2.2 *The Atlantic Meridional Overturning Circulation*

It is also of interest to understand whether the HiGEM decadal prediction system has any skill in capturing the evolution of the AMOC (Atlantic Meridional Overturning Circulation). Figure 9 shows the time-series of AMOC at 45N from the assimilation experiment, the ensemble mean of the HiGEM NOASSIM transient experiments and the HiGEM decadal predictions. Since there are no direct observations, the evolution of the AMOC at 45N from the assimilation experiment is taken as a proxy (in a manner similar to Pohlmann et al. 2013).

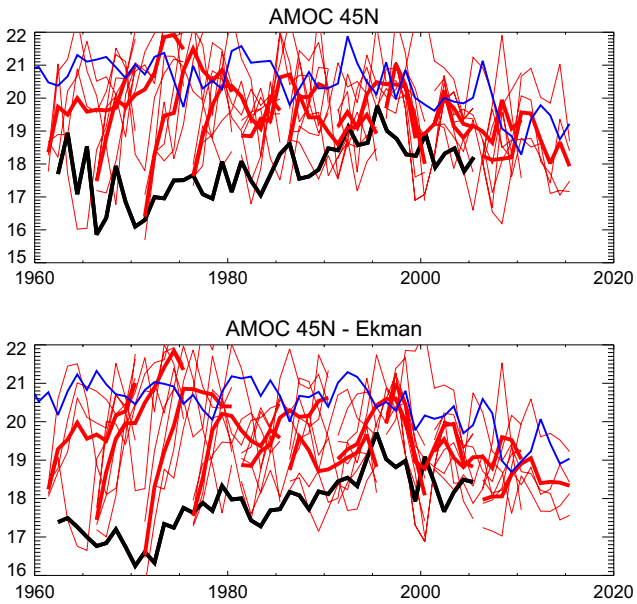

Fig. 9 Time-series of **a** the annual Atlantic Meridional Overturning Circulation (AMOC) at 45N and **b** the same but with the Ekman variability removed. The Ekman variability is removed by first regressing the anomalous AMOC at 45N onto the anomalous latitudinal windstress at 45N averaged between 100W and 0W (τ_x'). $AMOC'_{45N} = \beta\tau_x' + \epsilon$. β is then used to remove the Ekman variability from $AMOC_{45N}$ using $AMOC_{No_Ek} = AMOC_{45N} - \beta\tau_x'$. The *black line* is the assimilation experiment, the *blue line* is is the ensemble mean of the HiGEM NOASSIM transient experiments, the *thick red line* is the ensemble mean of the HiGEM decadal predictions and the *thin red lines* are the individual predictions. Units on the *y*-axis are in Sverdrups

It can be seen from 1960 to 1980 that the HiGEM decadal predictions have substantial problems with forecasting the AMOC at 45N. The AMOC at 45N in HiGEM typically has values of 20Sv, as indicated by the time-series of the HiGEM NOASSIM experiments. In contrast, the HiGEM decadal predictions from 1960 to 1980 are initialised with ocean states that give rise to substantially weaker AMOC values at 45N. The initial evolution of the HiGEM decadal predictions during 1960 to 1980 is to increase the strength of the AMOC to values more consistent with the HiGEM's climatology.

The problems with the assimilation of the AMOC may be due to the sparseness of ocean observations from 1960 to 1980, but it may also arise from the details of the anomaly assimilation scheme (e.g. the choice of climatology or relaxation timescale) or from problems with the HiGEM climate model. As shown in Sect. 3.2.1, the issues with initialisation of AMOC in the HiGEM decadal predictions

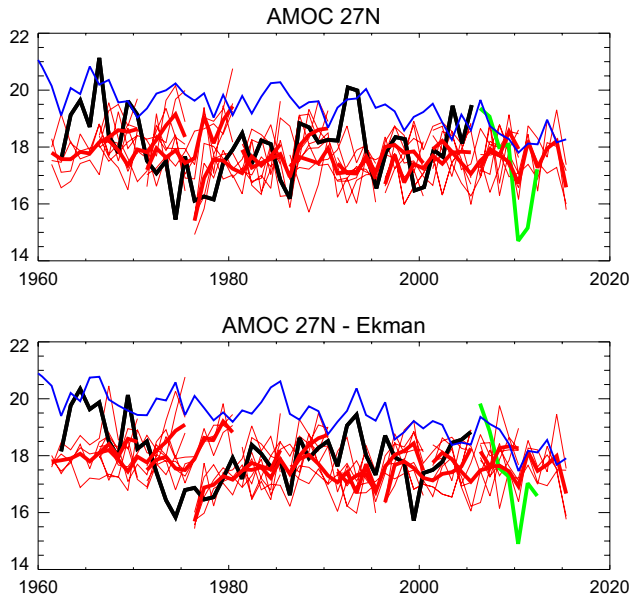

Fig. 10 Time-series of **a** the annual Atlantic Meridional Overturning Circulation at 27N and **b** the same but with the Ekman variability removed (see Fig. 9) . The *black line* is the assimilation experiment, the *blue line* is is the ensemble mean of the HiGEM NOASSIM transient experiments, the *thick red line* is the ensemble mean of the HiGEM decadal predictions and the *thin red lines* are the individual predictions. The *green line* denotes observed AMOC values from the RAPID array. Units on the *y*-axis are in Sverdrups

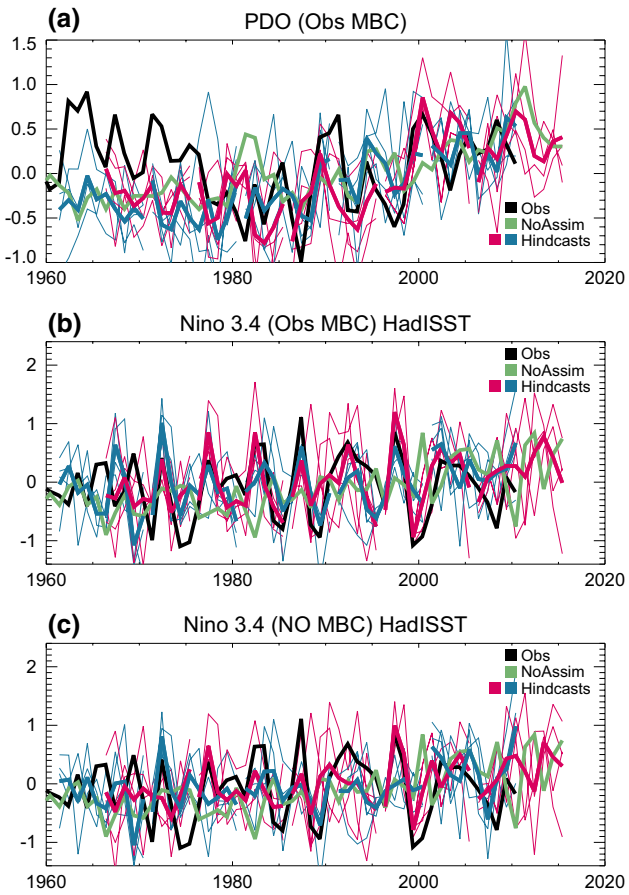

Fig. 11 Time-series of anomalies in **a** the annual mean Pacific Decadal Oscillation index (SSTs averaged 150–180W, 30–42N which is the maximum in spatial patterns of the pacific-wide PDO), **b** the annual mean Nino3.4 index (SSTs averaged 120–170W, 5S–5N) and **c** the annual mean Nino3.4 index without a lead-time dependent correction (SSTs averaged between 120–170W, 5S–5N). The *black line* is the observations (HadISST), the *green line* is the ensemble mean of the HiGEM NOASSIM transient experiments, the *thick red and blue lines* are the ensemble means of the HiGEM decadal predictions and the *thin red* and *blue lines* are the individual predictions. Units on the *y*-axis are Kelvin. In **a** and **b**, the Hindcasts have been lead-time dependent corrected with respect to the observations (see Sect. 2.3 for more details)

do not seem to substantially reduce the skill of the HiGEM decadal predictions in capturing SST in the North Atlantic subpolar gyre.

The problems with the drift in the AMOC at 45N appear to be strongest in the higher latitude North Atlantic. Figure 10 shows the time-series of AMOC at 27N from the assimilation experiment, the ensemble mean of the HiGEM NOASSIM experiments and the HiGEM decadal predictions. It can be seen from Fig. 10 that the HiGEM decadal predictions are capable of producing AMOC values that are similar in magnitude to the assimilation experiment and observations from the RAPID array. It is also apparent from 10 that there is little skill in the HiGEM decadal predictions for AMOC at 27N. It has been previously suggested that since the Ekman component of the AMOC is associated with the less predictable fluctuations in the atmosphere, removing the Ekman component may reveal the more predictable fluctuations associated with the ocean (Hermanson et al. 2014). Figures 9 and 10 indicate that removing the Ekman component makes little difference to the skill of the HiGEM decadal predictions for forecasting the evolution of the AMOC. Future research will examine whether the prediction skill increases for the recent period when there are more ocean observations available from the Argo datasets and the RAPID array to initialise and evaluate the decadal predictions.

3.3 The Pacific decadal oscillation and the El Nino southern oscillation

In this section, two SST indices are used to investigate the HiGEM decadal predictions for the Indo-Pacific region in

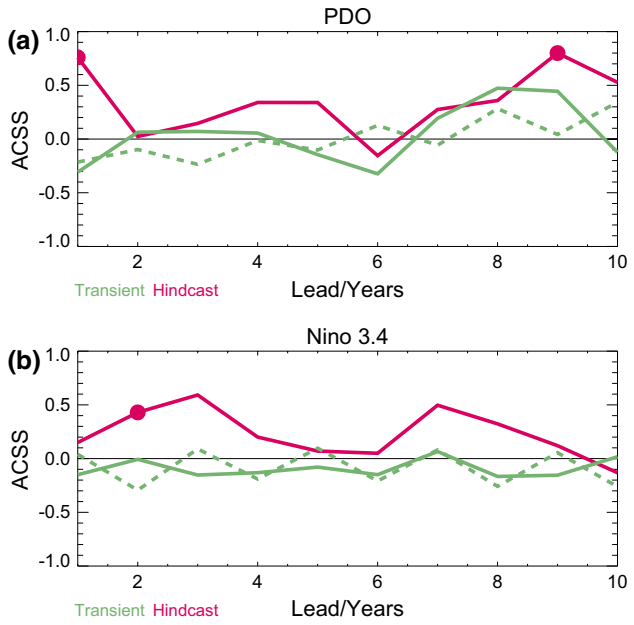

Fig. 12 Anomaly correlation as a function of lead time for the HiGEM decadal predictions (*red line*), the HiGEM NOASSIM transient experiments sampled for the same periods as the decadal predictions (*green line*) and HiGEM NOASSIM transient experiments sampled for 21 start dates dates (1960, 1962, 1964,...,2000; *green dashed line*). Anomaly correlations are shown for **a** PDO, **b** Nino3.4, as defined in Fig. 11. *Red circles* indicate where the skill of the HiGEM decadal predictions is significantly larger than the HiGEM NOASSIM experiments when sampled at 21 start dates at the 90 % level

more detail. These two indices are (i) an SST index associated with the Pacific Decadal Oscillation (PDO: averaged between 30N–42N and 150W–180W; Dawson et al. 2012) and (ii) the Nino 3.4 index (Nino 3.4: SSTs averaged between 5S-5N and 170W-120W).

3.3.1 The Pacific Decadal Oscillation

Figure 11 shows time-series of mean bias corrected anomalies in the annual mean Pacific Decadal Oscillation index, the annual mean Nino 3.4 index and the annual mean Nino 3.4 index without a mean bias correction. Figure 11a indicates that the HiGEM decadal predictions and the NOASSIM experiments are capable of capturing the recent long-term warming in the PDO index observed from 1980 to present. Some of the HiGEM decadal predictions are also capable of capturing some of the rapid changes in the PDO index (for example, the rapid warming in 1999 and 2000). However, it is also evident from Fig. 11 that the HiGEM decadal predictions have difficulty in capturing the long-term cooling observed in the PDO index observed from 1960 to 1980. This contributes to the limited ACC skill seen in Eastern Pacific SAT in Fig. 5.

Figure 12 shows the ACC for the mean bias corrected annual PDO index and the uncorrected annual Nino 3.4 index as a function of lead time. At 1 year lead time, there is modest skill (ACC is approximately 0.7) for the PDO in the HiGEM decadal predictions. This modest level of skill is nevertheless significantly greater than that of the HiGEM NOASSIM transient ensemble experiments. This suggests that the initialisation of the HiGEM decadal predictions substantially and significantly increases 1 year lead predictive skill. Figure 12 also shows some skill at lead times of year 9 for the PDO index. This may be due to sampling issues given that there is no apparent physical explanation.

3.3.2 The El Nino Southern Oscillation

Figure 11b shows the time-series of the mean bias corrected anomalies of the annual mean Nino 3.4 index. As mentioned in Sect. 2.3, the mean prediction bias may be under-sampled in the Tropical Pacific (Fig. 3). In particular, there appears to be a large cold bias in the Tropical Pacific in years 6 and 7. Given that it is very unlikely for there to be a physical explanation, it is therefore likely that the biases are due to under-sampling. The large years 6 and 7 biases manifest themselves in the mean bias corrected Nino 3.4 time-series as an artificial El Nino in each of the HiGEM decadal predictions.

Given the possible sampling issues with the mean bias correction, the uncorrected Nino 3.4 time-series is also shown in Fig. 11c. The uncorrected time-series does not suffer from the artifacts introduced by the mean bias correction. The HiGEM decadal predictions appear to be able to capture specific El Nino events (e.g. 1977/78 and 1997/98), however this does not translate into any significant skill in the annual mean Nino 3.4 index across the whole hindcast set (Fig. 12b).

In summary the skill in the Pacific is much more modest than that seen in the Atlantic Ocean. This is consistent with the results from other decadal prediction systems (e.g. Smith et al. 2013). However, there is modest but significant skill at 1-year lead time for the PDO. Future research will focus on understanding the mechanisms that might give rise to this skill in the PDO.

4 Conclusions and Discussion

The paper has described and evaluated a new decadal prediction system based on the HiGEM coupled climate model. The main conclusions of this study are:

- The HiGEM decadal predictions have substantial skill for predictions of annual SAT and 100 m upper ocean temperature. For lead times up to 10 years, anomaly correlations over large areas of the North Atlantic Ocean, the Western Pacific Ocean and the Indian Ocean exceed values of 0.6. Initialisation of the HiGEM decadal predictions significantly increases skill over regions of the Atlantic Ocean, the Maritime Continent and regions of the subtropical North and South Pacific Ocean. However, initialisation does not lead to a similar increase in skill over many land regions.
- The HiGEM decadal predictions are modestly but significantly more skillful than decadal predictions from the CMIP5 DePreSys system. At lead times of 1 year, the HiGEM decadal predictions are significantly more skillful in the Indian Ocean and the subtropical North and South Pacific. At longer lead times (i.e. years 2–3, 4–6 and 7–10), HiGEM appears to be significantly more skillful than CMIP5 DePreSys over the Eastern North Atlantic and to the west of the British Isles. This provides evidence that the skill of decadal predictions can be increased by using climate models with an improved representation of the climate system.
- The HiGEM decadal predictions can produce skillful predictions of the Atlantic Multidecadal Oscillation on multi-annual timescales. Most of the skill of the HiGEM decadal predictions arises in the North Atlantic subpolar gyre. The initialisation of the HiGEM decadal predictions results in skillful predictions for up to four years lead time (with ACC> 0.7). The skill of the HiGEM decadal predictions are significantly larger than the uninitialised HiGEM NOASSIM transient experiments.

This study has demonstrated that the HiGEM decadal predictions are capable of producing skillful multi-annual predictions. However, there is a need to better understand the physical processes that give rise to the long term predictability in the climate system (e.g. Robson et al. 2012). This will be a key focus of future work, particularly in the North Atlantic subpolar gyre where initialisation appears to result in the greatest gains in predictive skill.

These results have also highlighted the difficulty in initialising the high latitude Atlantic Meridional Overturning Circulation in the HiGEM decadal predictions. This difficulty seems to be particularly pronounced for the earlier period of the hindcast set (1960–1980). Additional future research directions include performing decadal predictions from the very recent period when there are substantially more ocean observations available to initialise and evaluate the predictions. In particular there will be a focus on performing predictions from the last decade when the Argo floats provide a step change in our understanding of the subsurface ocean.

Acknowledgments The authors would like to thank the two anonymous reviewers for providing constructive comments that improved the manuscript. The authors would also like to acknowledge the support and resources provided to the HiGEM DP team by the British Atmospheric Data Centre and by NCAS Computational Modelling Services. The authors would also like to acknowledge the Edinburgh Parallel Computing Centre and HECToR for HPC resources and the considerable support provided to the HiGEM DP team. We would also like to acknowledge the support of the National Centre for Atmospheric Science and the Natural Environment Research Council which provides funding for LS, JR, DH, EH, and RS. Doug Smith was supported by the joint DECC/Defra Met Office Hadley Centre Climate Programme (GA01101) and the EU FP7 SPECS project. Data from the RAPID-WATCH MOC monitoring project are funded by the Natural Environment Research Council and are freely available from www.rapid.ac.uk/rapidmoc.

Appendix: MSSS for HiGEM hindcasts and the impact of lead-time dependent prediction bias

A different perspective of the skill in the HiGEM CMIP5 decadal predictions is shown in Fig. 13. Figure 13 shows the mean squared skill score (MSSS; as recommended by Goddard et al. 2013) for SAT. The MSSS was calculated after a time-dependent correction of the mean-bias was performed (i.e. the average of the HiGEM prediction error for each lead time was removed from the hindcasts as recommended by WCRP; note that, as in Goddard et al. (2013), we do not use cross-validation to calculate the bias correction). Comparing to the anomaly correlation skill score (see Fig. 3) many similarities are apparent, including the improvement in the North Atlantic at most lead times and in the tropical Pacific and Atlantic.

Figure 13 (middle row) also shows the MSSS for the raw HiGEM hindcasts, that is without bias corrections applied, and highlights that the MSSS is sensitive to the treatment of bias. The bottom row of Fig. 13 shows the difference in MSSS between the full bias corrected hindcasts (top row) and the raw hindcasts (middle). It is clear from Fig. 13 that

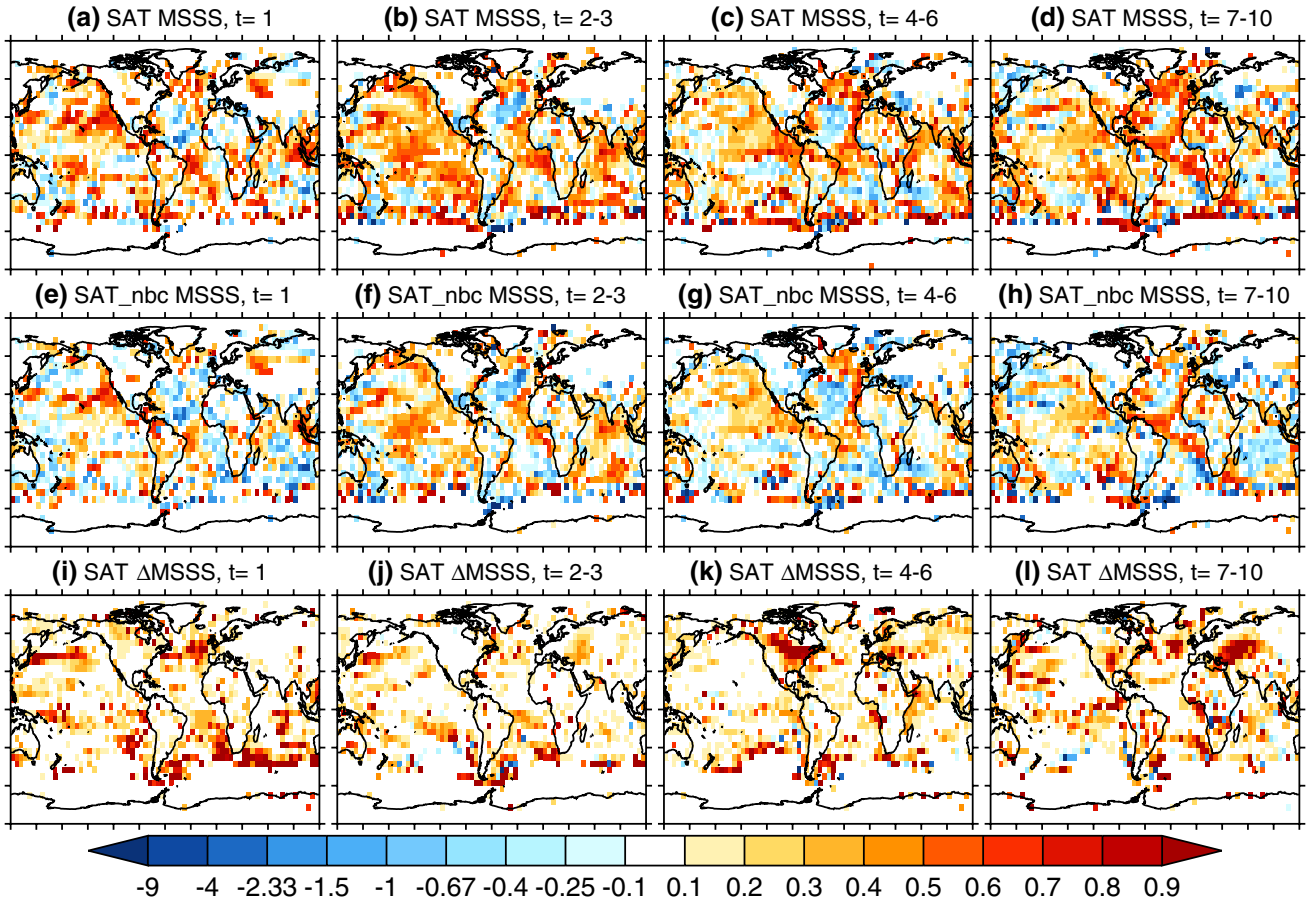

Fig. 13 SAT skill when using the Mean Squared Skill Score (MSSS) for the HiGEM decadal predictions. **a** shows the MSSS for averages of year 1 predictions when calculated using the NoAssim transient runs as the reference prediction. A full bias correction was applied to the hindcasts, but no lead-time dependent correction is applied to the NoAssim Transient runs. Positive values denote an improvement of the initialised hindcasts and MSSS was validated against HadCRUT4. **b–d** are the same as **a** but for averages of years 2-3, 4-6, and 7–10. **e–h** are the same as **a–d** but using the raw HiGEM predictions, that is without bias corrections. **i** The difference in MSSS in year 1 which is directly attributable to the bias correction (i.e. panel **a**–panel **e**). **j–l** The same as **i** but for averages of years 2–3, 4–6, and 7–10. *Colours* are only shown where there is data, and where the MSSS is significant at the $p \leq 0.05$ % based on a Monte-Carlo estimation (see main text for details)

there is a substantial increase in MSSS for the HiGEM decadal predictions when the bias correction is applied. This is particularly true for the Pacific at short lead times where the bias is large (Fig. 3), the North Atlantic, and most areas where there is an improvement in skill over land (e.g. North America in years 4-6 and 6-10, see panels k and l).

Although the bias correction improves the MSSS skill significantly over the majority of the globe, it is important to note that even when there is no bias correction the HiGEM decadal hindcasts skill improves SAT predictions over substantial areas of the globe, especially in the North Atlantic. Therefore, the analysis of the MSSS gives further confidence that the initialisation of HiGEM successfully improves predictions. However, the results hint that reducing model bias is important for the further improvement of decadal climate predictions.

References

Biastoch A, Boning CW, Lutjeharms JRE (2008) Agulhas leakage dynamics affects decadal variability in Atlantic overturning circulation. Nature 456:489–492

Catto J, Shaffrey LC, Hodges KI (2011) Northern hemisphere extratropical cyclones in a warming climate in the HiGEM high resolution climate model. J Clim 24:5336–5352

Chikamoto Y et al (2013) An overview of decadal climate predictability in a multi-model ensemble by climate model MIROC. Clim Dyn 40:1201–1222

Dawson A, Matthews AJ, Stevens DP (2012) Rossby wave dynamics of the North Pacific extra-tropical response to El Nio: importance of the basic state in coupled GCMs. Clim Dyn 37:391–405

Dawson A, Matthews AJ, Stevens DP, Roberts MJ, Vidale PL (2013) Importance of oceanic resolution and mean state on the extra-tropical response to El Nino in a matrix of coupled models. Clim Dyn 41:1439–1452. doi:10.1007/s00382-012-1518-6

Doblas-Reyes FJ, Balmaseda MA, Weisheimer A, Palmer TN (2011) Decadal climate prediction with the ECMWF coupled forecast system: Impact of ocean observations. J Geophys Res Atmos 116:D19111

Doblas-Reyes FJ et al (2013) Initialized near-term regional climate change prediction. Nat Commun 4:1715

Garcia-Serrano J, Guemas V, Doblas-Reyes FJ (2014) Added-value from initialization in predictions of Atlantic multi-decadal variability. Clim Dyn 44:2539–2555

Goddard L, Kumar A, Solomon A, Smith D, Boer G, Gonzalez P, Kharin V, Merryfield W, Deser C, Mason SJ, Kirtman BP, Msadek R, Sutton R, Hawkins E, Fricker T, Hegerl G, Ferro CAT, Stephenson DB, Meehl GA, Stockdale T, Burgman R, Greene AM, Kushnir Y, Newman M, Carton J, Fukumori I, Delworth T (2013) A verification framework for interannual-to-decadal predictions experiments. Clim Dyn 40:245–272

Hanlon HM, Hegerl GC, Tett SFB, Smith DM (2013) Can a decadal forecasting system predict temperature extreme indices? J Clim 26:3728–3744. doi:10.1175/JCLI-D-12-00512.1

Hawkins E Buwen, Dong Jon Robson, Sutton Rowan, Smith Doug (2014) The interpretation and use of biases in decadal climate predictions. J Clim 27:2931–2947. doi:10.1175/JCLI-D-13-00473.1

Hermanson L, Dunstone N, Haines K, Robson J, Smith D, Sutton R (2014) A novel transport assimilation method for the Atlantic meridional overturning circulation at 26 N. Q J R Meteorol Soc 140:2563–2572. doi:10.1002/qj.2321

Hodson DLR, Sutton RT (2012) The impact of resolution on the adjustment and decadal variability of the Atlantic Meridional Overturning Circulation in a coupled climate model. Clim Dyn 39:3057–3073. doi:10.1007/s00382-012-1309-0

Johns TC, Durman CF, Banks HT, Roberts MJ, McLaren AJ, Ridley JK, Senior CA, Williams KD, Jones A, Rickard GJ, Cusack S, Ingram WJ, Crucifix M, Sexton DMH, Joshi MM, Dong B-W, Spencer H, Hill RSR, Gregory JM, Keen AB, Pardaens AK, Lowe JA, Bodas-Salcedo A, Stark S, Searl Y (2006) The New Hadley centre climate model (HadGEM1): evaluation of coupled simulations. J Clim 19:1327–1353. doi:10.1175/JCLI3712.1

Jung T, Miller MJ, Palmer TN, Towers P, Wedi N, Achuthavarier D, Adams JD, Altshuler EL, Cash BA, Kinter JL III, Marx L, Stan C, Hodges KI (2012) High-resolution global climate simulations with the ECMWF model in project athena: experimental design, model climate and seasonal forecast skill. J Clim 25:3155–3172. doi:10.1175/JCLI-D-11-00265.1

Kharin VV, Boer GJ, Merryfield WJ, Scinocca JF, Lee W-S (2012) Statistical adjustment of decadal predictions in a changing climate. Geophys Res Lett 39:L19705. doi:10.1029/2012GL052647

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

Kirtman BP et al (2012) Impact of ocean model resolution on CCSM climate simulations. Clim Dyn 39:1303–1328

Matei D, Pohlmann H, Jungclaus J, Muller W, Haak H, Marotzke J (2012) Two tales of initializing decadal climate prediction experiments with the ECHAM5/ MPI-OM model. J Clim 25:8502–8523

Menary MB, Hodson DLR, Robson JI, Sutton RT, Wood RA, Hunt JA (2015) Exploring the impact of CMIP5 model biases on the simulation of North Atlantic decadal variability. Geophys Res Lett 42:5926–5934. doi:10.1002/2015GL064360

Mignot J, Garca-Serrano J, Swingedouw D, Germe A, Nguyen S, Ortega P, Guilyardi E, Ray S (2015) Decadal prediction skill in the ocean with surface nudging in the IPSL-CM5A-LR climate model. Clim Dyn. doi:10.1007/s00382-015-2898-1

Mochizuki T et al (2012) Decadal prediction using a recent series of MIROC global climate models. J Meteorol Soc Jpn 90A:373–383

Patricola CM, Li M, Xu Z, Chang P, Saravanan R, Hsieh J-S (2012) An investigation of tropical Atlantic bias in a high-resolution coupled regional climate model. Clim Dyn 39:2443–2463. doi:10.1007/s00382-012-1320-5

Pohlmann H, Jungclaus JH, Kohl 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

Pohlmann H, Smith DM, Balmaseda MA, Keenlyside NS, Masina S, Matei D, Müller WA, Rogel P (2013) Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system. Clim Dyn 41:775–785. doi:10.1007/s00382-013-1663-6

Roberts MJ, Clayton A, Demory M-E, Donners J, Vidale PL, Norton W, Shaffrey L, Stevens DP, Stevens I, Wood RA, Slingo J (2009) Impact of resolution on the tropical Pacific circulation in a matrix of coupled models. J Clim 22:2541–2556

Robson J (2010) Understanding the performance of a decadal prediction system. PhD thesis, University of Reading

Robson JR, Sutton KL, Smith D, Palmer MD (2012) Causes of the rapid warming of the North Atlantic in the 1990s. J Clim 25:4116–4134. doi:10.1175/jcli-d-11-00443.1

Scaife AA, Copsey D, Gordon C, Harris C, Hinton T, Keeley S, O'Neill A, Roberts M, Williams K (2011) Improved Atlantic winter blocking in a climate model. Geophys Res Lett 38:L23703. doi:10.1029/2011GL049573

Shaffrey LC et al (2009) UK-HiGEM: the new UK high resolution global environment model. Model description and basic evaluation. J Clim 22:1861–1896

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, Murphy JM (2007) An objective ocean temperature and salinity analysis using covariances from a global climate model. J Geophys Res 112:C02022

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 3:846–849

Smith DM, Eade R, Pohlmann H (2013) A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction. Clim Dyn 41:3325–3338. doi:10.1007/s00382-013-1683-2

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

Sutton RT, Hodson DLR (2003) Influence of the ocean on North Atlantic climate variability 1871–1999. J Clim 16:3296–3313

Toniazzo T, Mechoso CR, Shaffrey LC, Slingo JM (2010) Upper-ocean heat budget and ocean eddy transport in the south-east Pacific in a high-resolution coupled model. Clim Dyn 35:1309–1329

Woollings T (2010) Dynamical influences on European climate: an uncertain future. Phil Trans A 368:3733–3756

van Oldenborgh GJ, Doblas-Reyes FJ, Wouters B, Hazeleger W (2012) Decadal prediction skill in a multi-model ensemble. Clim Dyn 38:1263–1280

Zappa G, Shaffrey LC, Hodges KI (2013) The ability of CMIP5 models to simulate north Atlantic extratropical cyclones. J Clim 26:5379–5396