Clim Dyn (2013) 40:1483–1497
DOI 10.1007/s00382-012-1468-z

Decadal climate predictions with a coupled OAGCM initialized with oceanic reanalyses

A. Bellucci · S. Gualdi · S. Masina · A. Storto · E. Scoccimarro · C. Cagnazzo · P. Fogli · E. Manzini · A. Navarra

Received: 22 December 2011 / Accepted: 20 July 2012 / Published online: 5 August 2012

Abstract We investigate the effects of realistic oceanic initial conditions on a set of decadal climate predictions performed with a state-of-the-art coupled ocean-atmosphere general circulation model. The decadal predictions are performed in both retrospective (hindcast) and forecast modes. Specifically, the full set of prediction experiments consists of 3-member ensembles of 30-year simulations, starting at 5-year intervals from 1960 to 2005, using historical radiative forcing conditions for the 1960–2005 period, followed by RCP4.5 scenario settings for the 2006–2035 period. The ocean initial states are provided by ocean reanalyses differing by assimilation methods and assimilated data, but obtained with the same ocean model. The use of alternative ocean reanalyses yields the required perturbation of the full three-dimensional ocean state aimed at generating the ensemble members spread. A full-value initialization technique is adopted. The predictive skill of the system appears to be driven to large extent by trends in the radiative forcing. However, after detrending, a residual skill over specific regions of the ocean emerges in the near-term. Specifically, natural fluctuations in the North Atlantic sea-surface temperature (SST) associated with large-scale multi-decadal variability modes are predictable in the 2–5 year range. This is consistent with significant predictive skill found in the Atlantic meridional overturning circulation over a similar timescale. The dependency of forecast skill on ocean initialization is analysed, revealing a strong impact of details of ocean data assimilation products on the system predictive skill. This points to the need of reducing the large uncertainties that currently affect global ocean reanalyses, in the perspective of providing reliable near-term climate predictions.

Keywords Decadal predictions · Climate predictability · Decadal variability · Meridional overturning circulation · Ocean reanalyses

A. Bellucci (✉) · S. Gualdi · S. Masina · A. Storto · P. Fogli · A. Navarra
Centro Euro-Mediterraneo sui Cambiamenti Climatici, Viale A. Moro 44, 40127 Bologna, Italy
e-mail: alessio.bellucci@cmcc.it

S. Gualdi · S. Masina · E. Scoccimarro · A. Navarra
Istituto Nazionale di Geofisica e Vulcanologia, Bologna, Italy

C. Cagnazzo
Consiglio Nazionale delle Ricerche, Rome, Italy

E. Manzini
Max-Planck-Institut für Meteorologie, Hamburg, Germany

1 Introduction

Understanding the sources of predictability in the coupled ocean-atmosphere system over different temporal scales has always represented a primary goal in both theoretical and operational geophysical fluid dynamics. Numerical weather predictions and century-scale climate projections encompass the tails of a wide range of phenomena, involving very different sets of interactions between various components of the climate system. If the former is a pure initial value problem, whose predictive skill crucially depends on the knowledge of the initial state of the atmosphere (with the ocean evolution playing a minor role), the latter is a boundary value problem, with the expected changes in the chemical composition of the atmosphere (and in particular of those compounds that affect the radiative balance of the Earth) determining most of the predictability of the climatic system over multi-decadal and centennial scales. Seasonal and decadal

predictions lie somewhat in between. On the seasonal timescale, changes in greenhouse gases and aerosols can be assumed to play a negligible role on the climate evolution, influenced mainly by the initial state of the ocean, whereas decadal predictions appear to be a rather hybrid problem, where both initial and boundary (radiative forcing) conditions play an equally important role (Meehl et al. 2009). As such, the decadal range is crucial in that over this timescale both natural unforced fluctuations and anthropogenically driven changes are competing factors in shaping the climate variability at both global and regional scales. The slowly evolving ocean is a major driver of internal variability over decadal timescales. Low frequency variability of the thermohaline circulation, particularly active in the North Atlantic sector, as well as the large thermal inertia of the ocean, in turn leading to long-lasting upper ocean heat content anomalies, introduce decadal-scale memory in the climatic system.

Pioneering works (Griffies and Bryan 1997; Boer 2000; Pohlmann et al. 2004) based on simple model-only frameworks, investigated the predictability of the coupled ocean-atmosphere system, highlighting the central role played by ocean dynamics. Specifically, these experiments identified the North Atlantic and the Southern Ocean (and to a lesser extent the North and tropical Pacific) as regions where decadal predictability may be found. A major outcome of these efforts was to underline the relevance of the information embedded in the initial conditions of a dynamical model for predicting a significant fraction of the total internal variability. These promising results paved the way for more sophisticated approaches to decadal predictions. Recently, a number of initiatives have been undertaken aimed at providing more robust estimates of decadal predictability by constraining the initial conditions of the coupled system with the observed state of the ocean and atmosphere (Smith et al. 2007; Keenlyside et al. 2008; Pohlmann et al. 2009; Doblas-Reyes et al. 2011a). Initializing a coupled model with realistic conditions represents a significant step forward compared to previous long-term climate simulations , such as those performed for the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4), where the emphasis was on the forced response of the climate system to external (anthropogenic and natural) drivers. The awareness that a proper initialization procedure may enhance the predictive skill of a coupled model over multi-annual and longer timescales has led the climate science community to expand the former set of standard IPCC scenario simulations so as to include decadal predictions as part of the upcoming Coupled Model Intercomparison Project Phase 5 (CMIP5) effort (Meehl et al. 2009).

In this paper, the decadal predictive skill of the model developed at the Centro Euro-Mediterraneo per i Cambiamenti Climatici (CMCC-CM) is assessed. The predictability associated with the slowly evolving ocean state is investigated by initializing the model using dynamically balanced ocean reanalyses obtained through data assimilation techniques. The uncertainties associated with our knowledge of the ocean state are sampled by using different ocean reanalyses, all based on the same ocean dynamical model but different assimilation methodologies and observations. These perturbing elements (assimilated data amount and assimilation technique) are shown to produce a sufficiently large spread of the initialized simulations.

The paper is structured as follows. The dynamical model and the experimental design adopted to perform the decadal prediction experiments are described in Sect. 2. Details on ocean reanalyses used to initialize the dynamical model are given in Sect. 3. The sea surface temperature forecast skill score is discussed in Sect. 4, while forecast skill over land is discussed in Sect. 5. An evaluation of the predictive skill for specific climate variability indices is described in Sects. 6 (Atlantic SST Variability), 7 (Atlantic Meridional Overturning Circulation) and 8 (Pacific Decadal Oscillation). Summary and conclusions are reported in Sect. 9.

2 Model, experimental design and verification methods

The dynamical model used to perform the decadal prediction experiments is the global coupled general circulation model developed at the Centro Euro-Mediterraneo per i Cambiamenti Climatici (CMCC-CM; Scoccimarro et al. 2011). The atmospheric component is ECHAM5 (Roeckner et al. 2003) with a T159 horizontal resolution (corresponding to approximately 80 Km) and 31 hybrid sigma-pressure levels in the vertical with a top at 10 hPa. The ocean component is the OPA8.2 model (Madec et al. 1998) in the ORCA2 global configuration, solving primitive equations on a tripolar grid, with 2 poles in the Northern Hemisphere. The resolution is $2° \times 2°$L31 with a meridional refinement near the Equator approaching a minimum 0.5° grid spacing. The ocean model does also include the Louvain-La-Neuve (LIM) model for the dynamics and thermodynamics of sea-ice (Fichefet and Morales-Maqueda 1999).

The full set of prediction experiments consists of 3-member ensembles of 10 or 30-year simulations, initialized on the 1st of January and the 1st of November for years 1960–2005 with a 5-year spacing, yielding 20 different start-dates. CMIP5 historical radiative forcing conditions, including greenhouse gases (GHG), aerosols, ozone and solar irradiance variability, are used for the 1960–2005 period, followed by RCP4.5 scenario settings

for the 2006–2035 period. No volcanic aerosols effect is included in the model forcing. The ocean state is initialized using full-values from three different ocean reanalyses (see Sect. 3), the sea-ice is initialized through a model climatology, while the atmospheric initial states are derived from a control simulation performed with historical twentieth century radiative forcing conditions.

It is important to underline that, in the present configuration of the decadal prediction system, only one ensemble member is generated for a selected ocean reanalysis. While the use of different ocean reanalyses allows an efficient sampling of the uncertainties in the ocean initial conditions associated with ocean data assimilation schemes and observational data, here no sampling is provided of the uncertainties related to the atmospheric and land surface initial states. The combined perturbation of ocean (through an ensemble of ocean syntheses), atmosphere and land surface conditions has been recently applied in a seasonal prediction study (Zhu et al. 2012) showing the potential benefits for the overall skill of a climate prediction system.

Due to the specific design adopted for this set of experiments, the informative content associated with the *real world* dynamical status resides entirely in the ocean initial conditions (estimated through the use of ocean reanalyses). Thus, in this work we will be mainly focusing on oceanic state variables. Although predictive skill over land may eventually arise in the system through ocean-amosphere "bridges" involving teleconnection patterns (see Sect. 5), a detailed analysis of interannual-to-decadal predictability over continental areas is outside the scope of this paper.

In the following analyses we mainly use anomaly correlation as a predictive skill evaluation measure. Correlations are computed over lead-times 1, 2–5 and 6–9 years, for equally sub-sampled predictions and observations.

Because the full-value initialization approach is affected by the long-term adjustment of the system towards its own mean state, a drift removal procedure is applied to model data by subtracting the average forecast from the individual raw forecasts. This procedure allows for the decontamination of model data from spurious non-physical trends. Assuming that negative correlations are physically meaningless, a simple one-sided Student's t-test is used to verify statistical significance.

In addition to anomaly correlations, a deterministic metric based on the mean squared skill score (Murphy 1988) is used to assess whether the initialization of SST hindcasts leads to more accurate predictions with respect to an uninitialized climate simulation. For this purpose, a historical simulation performed with the CMCC-CM model following the CMIP5 protocol, using an arbitrary initialization and the same boundary conditions adopted for the predictions, is used as a reference forecast.

3 Ocean reanalyses

The three reanalyses used to initialize the ocean differ in the data assimilation method used. One of them is performed with an Optimal Interpolation (OI) scheme and the other two with a three-dimensional variational data assimilation system (3DVAR1 and 3DVAR2), which in turn varies in the specification of the background-error vertical covariances. Both the analysis methods include a three-dimensional bivariate correction of temperature and salinity fields. The OPA8.2 ocean general circulation model is used in all of the reanalyses, with the same resolution and configuration as used in the coupled model. Surface forcing fields are taken from the European Center for Medium-Range Weather Forecasts (ECMWF) ERA-40 atmospheric reanalyses (Uppala et al. 2005) until December 2001, and from January 2002 onwards from the ECMWF operational analyses and forecasts.

The OI scheme is based on the System for Ocean Forecasting and Analysis (SOFA) assimilation package (De Mey and Benkiran 2002), which was implemented for the global ocean by Bellucci et al. (2007). It consists of an OI performed on the reduced-order space of the ten dominant vertical empirical orthogonal functions (EOFs). EOFs are season-dependent, and were calculated from an assimilation-blind model hindcast, at full model horizontal resolution and without any spatial filtering; this strategy was proved to lead to the best verification skill scores (Masina et al. 2011). Horizontal correlations are assumed to be Gaussian, with a constant correlation length-scale equal to 300 Km. The OI system assimilates all the observations from the EN3v2a data set of the ENSEMBLES project (Ingleby and Hudlleston 2007). The 3DVAR reanalyses use the OceanVar variational assimilation system (Dobricic and Pinardi 2008), in its global ocean implementation (Storto et al. 2011b). Like the OI, vertical covariances are modeled by 10-mode seasonal bivariate EOFs of temperature and salinity. Horizontal correlations are obtained by applying a four-iteration first order recursive filter, with horizontally homogeneous and vertically varying correlation length-scales. Along with the in-situ observations from the EN3v2a data set, the 3DVAR reanalyses also assimilate along-track, sea-level anomaly (SLA) observations provided by AVISO (Le Traon et al. 1998) over the period 1992—onwards. Note that expendable bathythermographs (XBT) fall rates are corrected according to Wijffels et al. (2008) within the 3DVAR reanalyses. A local hydrostatic adjustment scheme is used for the SLA assimilation, by means of which the sea-level anomaly increment is covariated in vertical profiles of temperature and salinity increments (Storto et al. 2011b). The two 3DVAR reanalyses differ only in the method used to calculate the background-error covariances. In 3DVAR1,

the model variability is assumed to reproduce reasonably well the background error covariances. Consequently, the vertical EOFs are the same ones as those in the OI analysis. In 3DVAR2, vertical EOFs were computed from the differences between the members of an ensemble variational assimilation experiment and its ensemble mean, with the aim of representing the model error evolution rather than the model variability. The ensemble method (Storto et al. 2011a; see also Storto and Randriamampianina 2010, for a theoretical explanation) allows a better simulation of the true error evolution of both the data assimilation and the forecast model steps. The spread among the ensemble members was obtained by perturbing the observations, the surface forcing fields and the non-advective terms of the state variable tendencies (Palmer et al. 2009).

4 Sea surface temperature skill

In the present section, global mean and local SST forecast skill is evaluated using HadISST temperatures (Rayner et al. 2003) as a verification data set.

Figure 1 shows the time series of global mean SST anomalies (with respect to 1960–1990 baseline) in observations and ensemble mean predictions, with the latter averaged over lead times 1, 2–5 and 6–9 years. Most of the interannual variability in the observed global SST record can be traced back to El Niño-Southern Oscillation (ENSO) modulation and episodic volcanic eruptions, while the longer term variability is dictated by the upward trend associated with increasing GHGs concentrations in the atmosphere. The effect of initialization is clearly reflected in the 1-year lead time hindcasts, (correlated with equally sub-sampled SST observations at 0.91; Fig. 2), while slightly lower skill in the 2–5 years range emerges ($r = 0.86$), in contrast with the larger ($r = 0.93$) score for the 6–9 years range (Fig. 2). The poor sampling of the observed state, related to the small number of start dates, is most likely responsible for the relatively lower skill shown for the 2–5 year near term. In particular, ENSO events are dramatically under-sampled due to the particular selection of start dates, as no major ENSO episode occurs during the years selected for initialization. This leads to a lack of coherence between the interannual fluctuations in hindcast and observed global SSTs. In the longer 6–9 year range, predicted SSTs show a much reduced variance at interannual timescale, with prevailing variability related to changes in the external radiative forcing, leading in turn to the large correlation with the observed record.

SST predictability associated with specific initialization reanalyses (OI, 3DVAR1 and 3DVAR2) is contrasted with the ensemble mean in Fig. 2. The ensemble-mean SSTs show an improved skill with respect to most of the

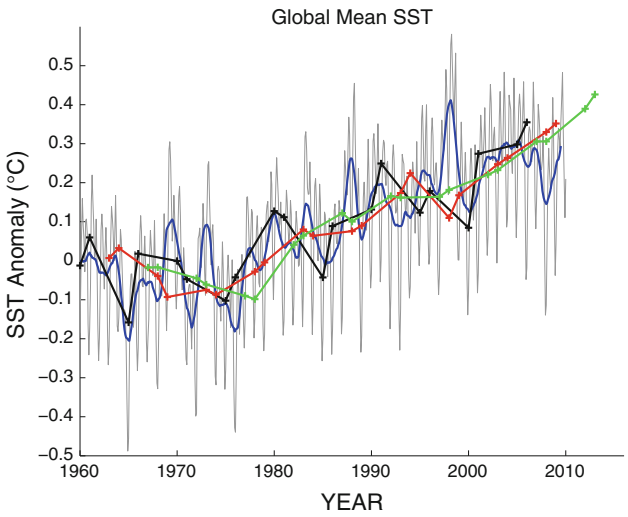

Fig. 1 Global mean SST anomalies (relative to 1960–1990) from HadISST (*grey* monthly mean; *thick blue* low-pass filtered with a 12-month moving average), and 1-year (*black*), 2–5 years (*red*) and 6–9 years (*green*) ensemble-mean predictions

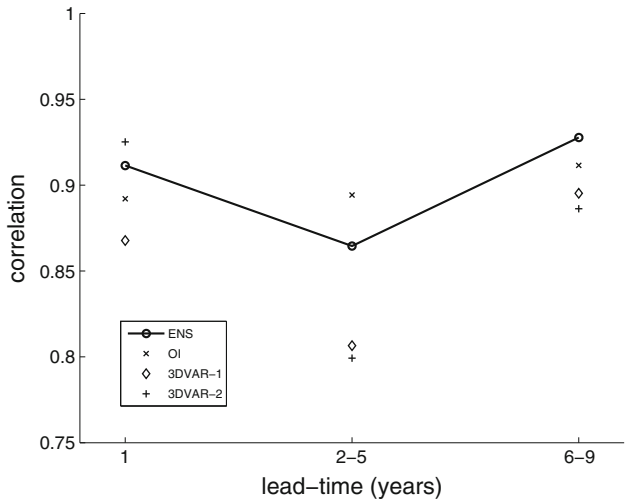

Fig. 2 Correlations between ensemble-mean (*circle*), OI- (*cross*), 3DVAR1- (*diamond*), 3DVAR2- (*plus*) initialized predictions and observed (HadISST) global mean SST, at several lead-times

individual members, without systematically outperforming the single forecasts. All correlation values shown in Fig. 2 are statistically significant at the 95 % level, according to a Student *t* test.

In order to quantify the predictive skill at sub-basins scale, the local forecast skill for sea surface temperatures is examined. The predictive skill for annual mean SST at different lead-times is evaluated via anomaly correlation coefficient (ACC) patterns, for the period 1960–2010. ACC patterns for SST at lead times 2–5 and 6–9 years, are shown in Fig. 3. According to a regular one-sided *t* test,

used to assess the statistical significance, correlations larger than 0.4 are significant at the 95 % level, for 18-points correlations (2–5 and 6–9 years). Thus, in the near-term (2–5 year lead times) SSTs display predictive skill over wide regions of the world oceans, in particular over the Indian Ocean, the extra-tropical North and South Atlantic, and over the western North Pacific. In the long term (6–9 year lead times), predictive skill is even larger than in the 2–5 year range, something that is consistent with the large score displayed by the global mean SST signal (Fig. 2).

In order to disentangle the relative contribution of the upward trend in the anthropogenic radiative forcing from the unforced variability on the overall SST skill score, ACC is re-computed after removing local SST tendencies from both observations and model results (Fig. 3, bottom panels). After detrending, ACC patterns show a much lower skill, both in the near and long term, with statistically significant residual skill in the near 2–5 year range over extensive regions of the North Atlantic basin, Southern ocean and North Pacific basins. The low skill featured by the equatorial Eastern Pacific may be possibly attributed to the afore-mentioned lack of predictive skill associated with ENSO events. This result is consistent with the low skill shown by the corresponding full-field ACC pattern (Fig. 3, top-left panel) over the same region.

In the longer 6–9 years term, local forecast skill exceeds the 95 % significance threshold over most of the North Atlantic area and for wide regions in the extra-tropical North Pacific. Interestingly, the equatorial Eastern Pacific, showing no skill in the near range, exhibits a skill "re-emergence" in the 6–9 years term. A similar feature was found in the ENSEMBLES multi-model decadal prediction system (van Oldenborgh et al. 2012).

In order to assess whether the initialization of the hindcasts leads to more accurate predictions with respect to an uninitialized climate simulation (i.e., a projection initialized with an arbitrary initial state and performed for the same time period of the predictions using identical forcing conditions) a deterministic metric, based on the mean squared skill score (MSSS) was used (Murphy 1988). The MSSS is based on the mean squared error (MSE):

$$MSE(Y,X) = \frac{1}{n}\sum_{j=1}^{n}(Y_j - X_j)^2 \quad (1)$$

where Y_j and X_j are the forecast and observed values at lead times j = 1,n years, respectively. The MSE can be calculated for both the initialized decadal hindcasts and for the uninitialized hindcasts (MSE_y and MSE_w, respectively). Then, the MSSS is defined as:

$$MSSS = 1 - \frac{MSE_y}{MSE_w} \quad (2)$$

Thus, positive MSSS values indicate that the initialization supplies added value to the prediction, with respect to a simply forced simulation. For an ideally perfect forecast, MSE_y is zero, and MSSS is 1. However, MSSS can take negative values, in case the error associated with the prediction is larger than for the uninitialized simulation. Here we diagnose the MSSS for SST using a CMIP5

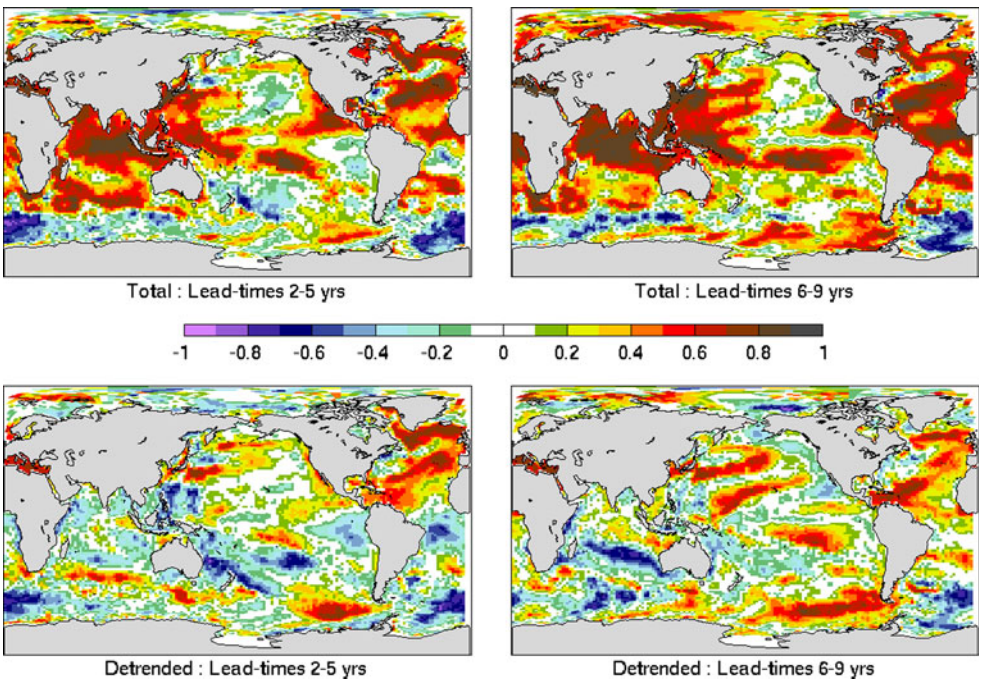

Fig. 3 Anomaly correlation coefficient of SST hindcasts for years 2–5 (*left*) and 6–9 (*right*), with (*top*) and without (*bottom*) trend. Correlations larger than 0.4 are statistically significant at the 95 % level according to a one-sided Student's *t* test

historical simulation, spanning the same period of the predictions, as a reference uninitialized forecast. All hindcasts have been bias-corrected, prior to the MSSS computation. The MSSS pattern (shown in Fig. 4) reveals that, over most of the global ocean, the skill score is positive, indicating that overall the initialization improves the model predictive skill with respect to the uninitialized control simulation. Enhanced skill is particularly pronounced over the North Atlantic basin, south-eastern Pacific, and extensive portions of the Southern Ocean. However, there are wide areas (notably, the equatorial Pacific, south-west Pacific, southern Indian Ocean as well as most of the Arctic) featuring negative MSSS, indicating no improvements associated with hindcast initialization. Comparing the MSSS pattern with the ACC maps for SSTs (Fig. 3), there are indications that the areas featuring negative anomaly correlations, roughly coincide with regions characterized by negative correlations, although a close comparison does not apply in this case as the MSSS is computed using the whole set of available lead times (1–10 years).

5 Forecast skill over continental areas

In order to evaluate the predictive skill of decadal hindcasts over continental areas, ACC patterns for 2-m temperature and precipitation over land are diagnosed, in the 2–5 and 6–9 years range. CRUTEM3 (Brohan et al. 2006) and GPCC v4 (Schneider et al. 2008) data are used as observational reference for 2-m temperature and precipitation, respectively.

Similarly to SSTs, the skill of near-surface temperature is mainly determined by the long-term trend in the prescribed boundary conditions, as can be inferred from the dramatic loss of skill found after the detrending procedure (Fig. 5). Anomaly correlations computed by including the trend (Fig. 5, top panels), display positive correlations over most of the land surface, to large extent exceeding the 95 % statistical significance threshold. The skill is negative over parts of the south american continent, particularly over Bolivia and surrounding areas. However, it must be noticed that no trends are found in observations, over this area (Fig. 3.9 in Trenberth et al. 2007).

After removing the trend (Fig. 5, bottom panels), residual skill is found over extensive continental areas. Part of this signal, and specifically the enhanced predictability found over the Mediterranean basin, Middle-East and North America bear strong similarities with the teleconnection pattern of the Atlantic Multidecadal Oscillation (Knight et al. 2005; van Oldenborgh et al. 2012), suggesting that predictability over land may arise from predictive skill over neighboring oceanic areas. The predictability associated with the natural multidecadal variability in the Atlantic sector is further analysed in Sects. 6 and 7.

In Fig. 6, the anomaly correlations for precipitation over land are displayed. Only the total skill (i.e., without removing the trend) is shown, as trends in precipitation represent a much smaller fraction of the total variability, compared with surface temperatures. Predictive skill in precipitation is generally low, with vast areas featuring negative correlations, particularly in the 2–5 years range. However, a few regions exhibit a spatially coherent pattern of statistically significant skill, most notably the Sahel and northern Scandinavia in the 6–9 years term. Aside from these specific areas, the general outlook of the ACC patterns for precipitation is extremely noisy, making the identification of the processes which are relevant for predictability particularly difficult.

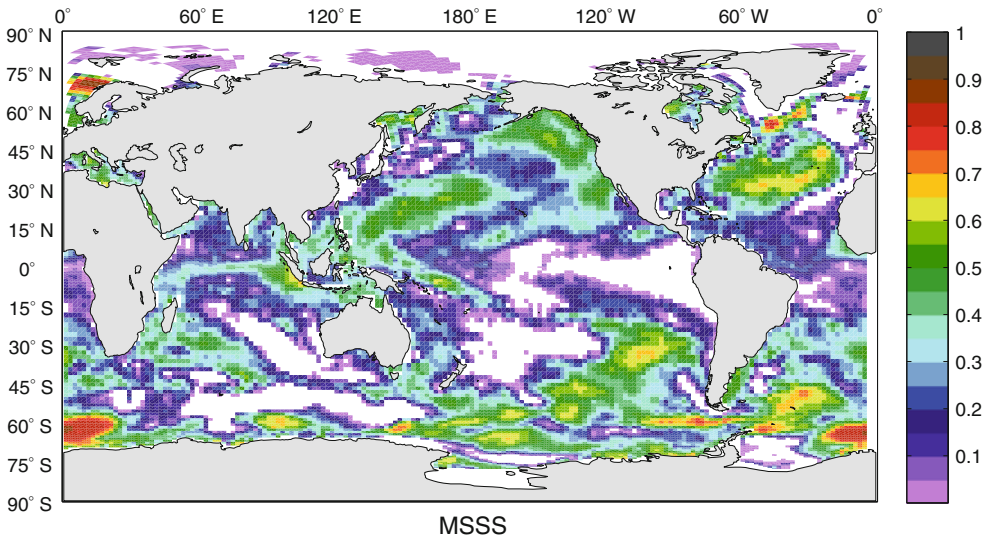

Fig. 4 Mean squared skill score (MSSS) evaluated for SSTs. A historical uninitialized simulation spanning the same period of the predictions, and identical forcing conditions has been used as a reference hindcast. Negative MSSS values are shown in *white*

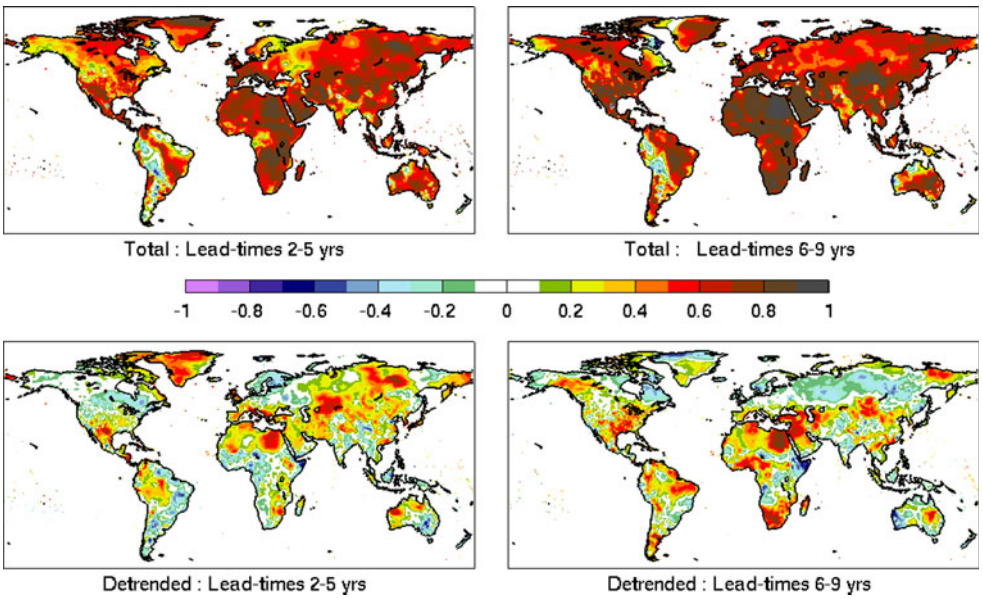

Fig. 5 Same as Fig. 3, but for 2-m temperatures over land

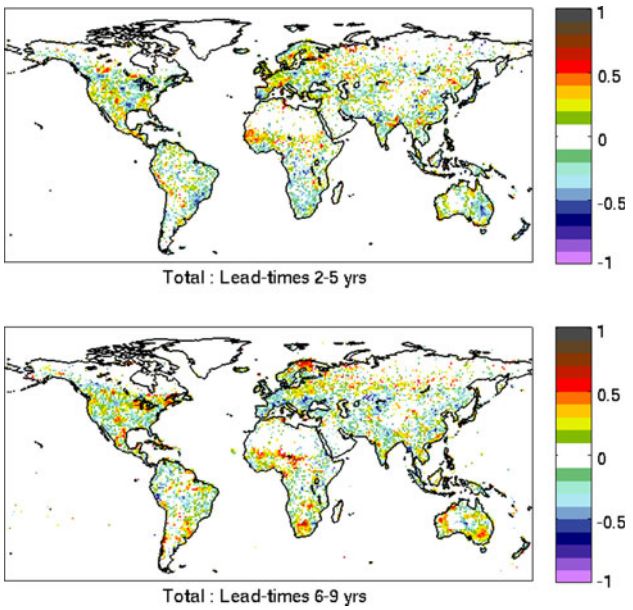

Fig. 6 Same as Fig. 3, but for precipitation over land. Only total skill is shown (trends included)

In the following sections, a more detailed analysis is carried out of the predictability associated with regional climate variability indices in the Atlantic and Pacific basins.

6 Atlantic multi-decadal variability

In the previous sections, the Atlantic sector emerged as a region featuring a significant predictive skill, even after removing local trends, indicating that natural, multi-annual fluctuations may locally enhance the predictability of the system. Figure 7 shows the time series of the Atlantic SST Dipole (hereafter AD; Keenlyside et al. 2008) index, which is here defined as the area-averaged SST difference over (40–60N,60–10W) minus (10–40S,30W–10E), for observations (HadISST) and ensemble mean hindcast/forecast simulations. The predicted AD index is shown for year 1, years 2–5 and years 6–9. Predictive skill is evident not only after 1 year ($r = 0.74$) but also over the near-term 2–5 year range ($r = 0.67$), while in the longer 6–9 year term only marginally significant skill is found ($r = 0.38$). Interestingly, the reduced skill for longer lead-times appears to be due to a delayed response of the predicted AD index, causing the AD hindcast to be out of phase with respect to observations. The above-mentioned lack of skill for the longer 6–9 year range does not allow a proper forecast (i.e., referring to the decade starting on year 2011), since year 2005 is the latest start date. Nonetheless, it is interesting to highlight the predicted persistency of the AD warm phase, after 2010, consistent with the latest part of the observed record.

Next we focus on the Atlantic Multidecadal Oscillation (AMO) index (Schelsinger and Ramankutty 1994). Several definitions of the AMO index can be found in the literature (Enfield et al. 2001). A particularly thorny issue is the discrimination of natural variability, arising from processes internal to the Atlantic sub-system, from the global anthropogenic warming signal. Trenberth and Shea (2006; hereafter TS06) suggest a possible way to remove the global warming signal from the AMO index, by subtracting the global mean SST, thus leading to a revised AMO, freed from the global warming signature (see also van Oldenborgh et al. 2009 for an alternative AMO-detrending

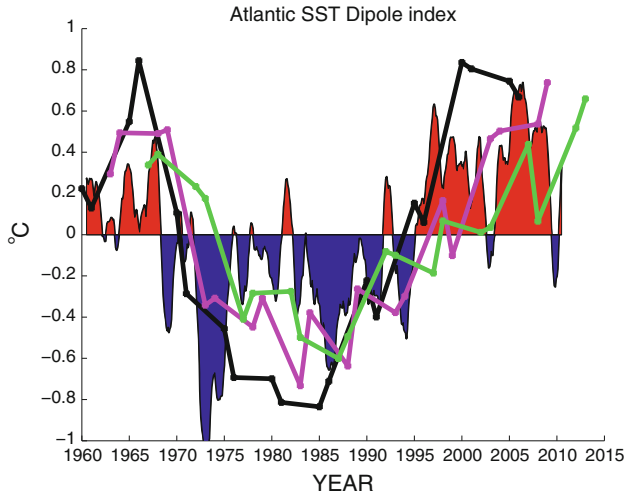

Fig. 7 Observed (*shading*) and predicted Atlantic SST dipole index for year 1 (*black*), years 2–5 (*purple*) and years 6–9 (*green*). Observed dipole index is based on monthly HadISST data, low-pass filtered with a 12-month moving average

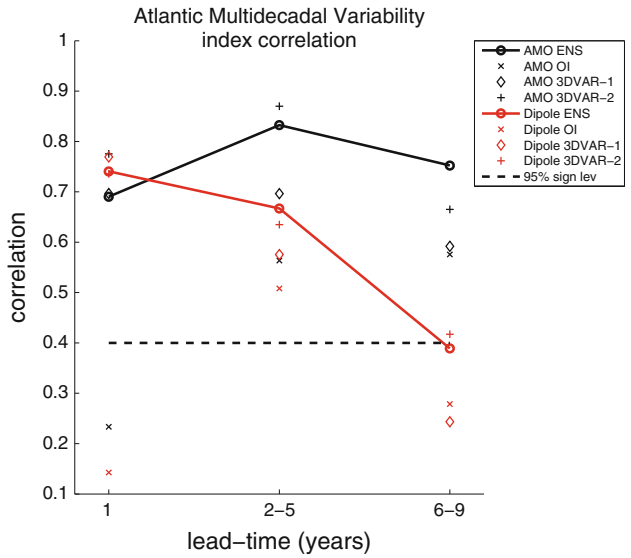

Fig. 8 Correlations between ensemble-mean (*solid line*), OI (*cross*), 3DVAR1 (*diamond*), 3DVAR2 (*plus*) predictions and observations, for Atlantic Dipole index (*red*) and Atlantic Multidecadal Oscillation index (*black*) at several lead-times. HadISST temperatures are used as reference data set. The 95 % statistical significance threshold according to a one-sided Student *t* test is also shown (*dashed*)

procedure). Correlations for the AMO index using the revised TS06 definition are shown in Fig. 8. The comparison with the AD index shows slightly lower values for the 1-year lead time, but improved skill for years 2–5 and 6–9. This result is consistent with findings of van Oldenborgh et al. (2012), who also show an improved coherence between observations and the predicted AMO index for the 2–5 year slower variations of the AMO.

Predictability in the Atlantic region is then assessed against the specific data assimilation products used for the initialization by separately evaluating the predictive skill exhibited by OI and the two 3DVAR reanalyses, for both AD and AMO indices. The analysis reveals that OI initial conditions yield a generally lower skill (systematically for AMO index) with respect to 3DVAR1 and 3DVAR2 (Fig. 8). Overall, the ensemble mean prediction displays statistically significant correlation at all lead-times, in several cases outperforming individual ensemble members. This result, which is well-known from seasonal predictions, appears to hold also for decadal predictions. The lack of skill revealed by OI with respect to 3DVAR-initialized predictions, points to some major deficiencies of the OI method in constraining the initial state of the ocean, although it is unclear whether this is related to the assimilation method or the amount of the assimilated data. This aspect is further investigated in the context of the predictability of the meridional overturning circulation (see next section).

Summarizing, the Atlantic basin is confirmed to be a region featuring large predictability, up to the 6–9 year range.

7 Atlantic meridional overturning circulation

Low frequency fluctuations associated with the strength of the slowly evolving global ocean conveyor represent a potential source of predictability in the climate system. The meridional mass transport across zonal sections in oceanic basins, estimated through the meridional overturning streamfunction, provides a dynamical proxy for the local strength of the thermohaline circulation (THC). The Atlantic component of the meridional overturning circulation is particularly effective in transferring heat meridionally, and as such plays an important role on the global energy cycle. Interest in predicting the Atlantic meridional overturning circulation is further supported by evidence of a strong link with coordinated changes in the surface temperature field of the Northern Hemisphere (Knight et al. 2005; Collins et al. 2006).

Figure 9 shows the time series of the maximum Atlantic meridional overturning circulation at 26N (hereafter MOC) for each of the ocean reanalyses used to initialize the decadal predictions. This particular latitude is selected because several observational estimates exist based on transatlantic zonal sections (Bryden et al. 2005; Cunningham et al. 2007) and direct daily measures through the RAPID monitoring system (Kanzow et al. 2008) at approximately the same latitude. A striking feature emerging from this comparison is the large uncertainty affecting both the mean state and the variability of the MOC. The root-mean-square deviations

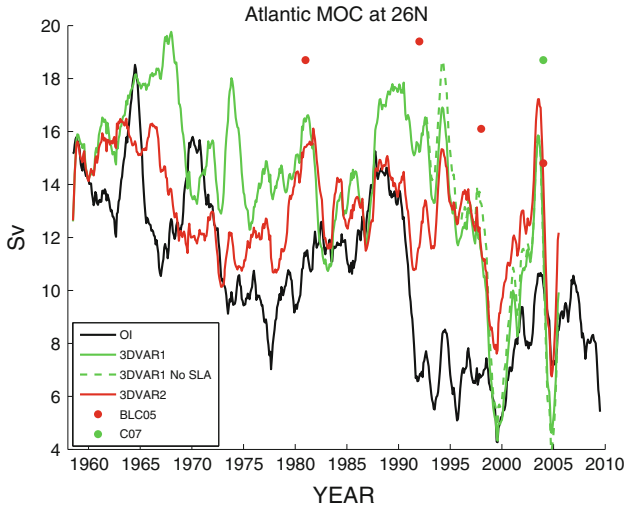

Fig. 9 Atlantic meridional overturning circulation at 26°N as represented by OI (*black*), 3DVAR1 (*green*), 3DVAR1-No SLA (*dashed green*) and 3DVAR2 (*red*) reanalyses. Observational estimates from Bryden et al. (2005; *red circles*) and Cunningham et al. (2007; *green circle*) are also shown. Units are in Sv (1 Sv = $10^6 m^3 s^{-1}$)

between the different estimates of the MOC range between 2.4 Sv (3DVAR1 and 3DVAR2) and 4.1 Sv (OI and 3DVAR1), while cross-correlations vary between 0.55 (OI and 3DVAR2) and 0.7 (3DVAR1 and 3DVAR2). The MOC amplitude in 3DVAR reanalyses is more consistent with existing direct observational estimates, compared to OI. The largest discrepancies emerge during the 1990s, with the OI undergoing a severe drop towards a weaker MOC state.

In order to disentangle the effects of the assimilated data set from the assimilation method on the MOC evolution, an additional reanalysis is shown (3DVAR1 NoSLA), where the 3DVAR1 scheme is used to assimilate in-situ temperature and salinity data, but no altimetry data. This additional reanalysis was run from 1992 onwards. The comparison with 3DVAR1 (Fig. 9) displays a relatively small enhancement in the MOC amplitude (1–2 Sv) but no substantial changes in the MOC variability. Overall, the impact of SLA assimilation on MOC appears to play a second order role, suggesting that the major differences between MOC estimates are determined by the assimilation method. This result is further corroborated by additional analyses performed on the meridional overturning mean states in the different ocean reanalyses (not shown).

Summarizing, the reanalyses show a broad agreement in reproducing a decrease in the MOC from the 1990s onward, that is partly consistent with the few available observations (although the large uncertainties affecting the MOC direct estimates should be carefully considered; see in particular, the year 2004 estimates) and reconstructions based on ocean syntheses (Pohlmann et al. 2012).

While, in principle, data assimilation is expected to contribute to a reduction of the uncertainties affecting the estimate of the ocean state, compared to model or observational estimates alone, this seems not to be the case for the MOC (Balmaseda et al. 2007). A large spread affecting the 26N Atlantic overturning in ECMWF ORAS3 and NEMOVAR reanalyses (similar to the one featured by the present set of ocean reanalyses) is reported in Balmaseda et al. (2010; see their Fig. 4). The reasons behind this behavior of the MOC are not well understood, although Balmaseda et al. found some indication of a MOC strength dependency on the specific treatment of ocean data near the coastlines.

The detected MOC discrepancies in the reanalyses highlight the large spread of the perturbations imposed to the initial state of oceanic dynamical fields when different data assimilation products are used to initialize the ocean component of the coupled model. The large uncertainties displayed by the Atlantic MOC remind us that identifying the *true* field for this specific component of the ocean circulation is not straightforward. This is a well-known problem in all the existing reanalyses of the global ocean (Balmaseda et al. 2010; Kröger et al. 2012; Muñoz et al. 2011). For this reason, in the following analysis, the ensemble members initialized with a specific ocean reanalysis will be treated as a distinct data set. Thus, MOC predictability for each sub-set will be verified against the corresponding initialization ocean data assimilation product.

In Fig. 10, the Atlantic MOC at 26N for each initialization reanalysis and the corresponding predictions averaged over lead-times 1, 2–5 and 6–9 years are shown. The comparison highlights the large sensitivity of predicted MOC to the specific ocean initialization data set. In particular, OI-initialized hindcasts exhibit a systematic upward drift in the 2–5 years transient after initialization, which is virtually absent in the 3DVAR-initialized predictions. An analysis of the MOC average forecast for each prediction set reveals that the detected upward drift in the 2–5 year lead time range is caused by an initialization *shock*, which is considerably larger in the OI set compared to 3DVAR (not shown). Overall, hindcasts initialized with 3DVAR reanalyses show a fairly good consistency with their reanalysis counterpart in the 2–5 year range, skillfully capturing the low frequency variability at multi-annual scales. In the 6–9 year range, MOC values are affected by the long term adjustment towards the model climatology (13–14 Sv). After drift removal (performed by subtracting the average forecast from each raw forecast), the residual MOC probability density functions in the initialized runs exhibit a much narrower range of values compared to ocean reanalyses (Fig. 11). A similar underestimation of MOC variance in model simulations with respect to ocean

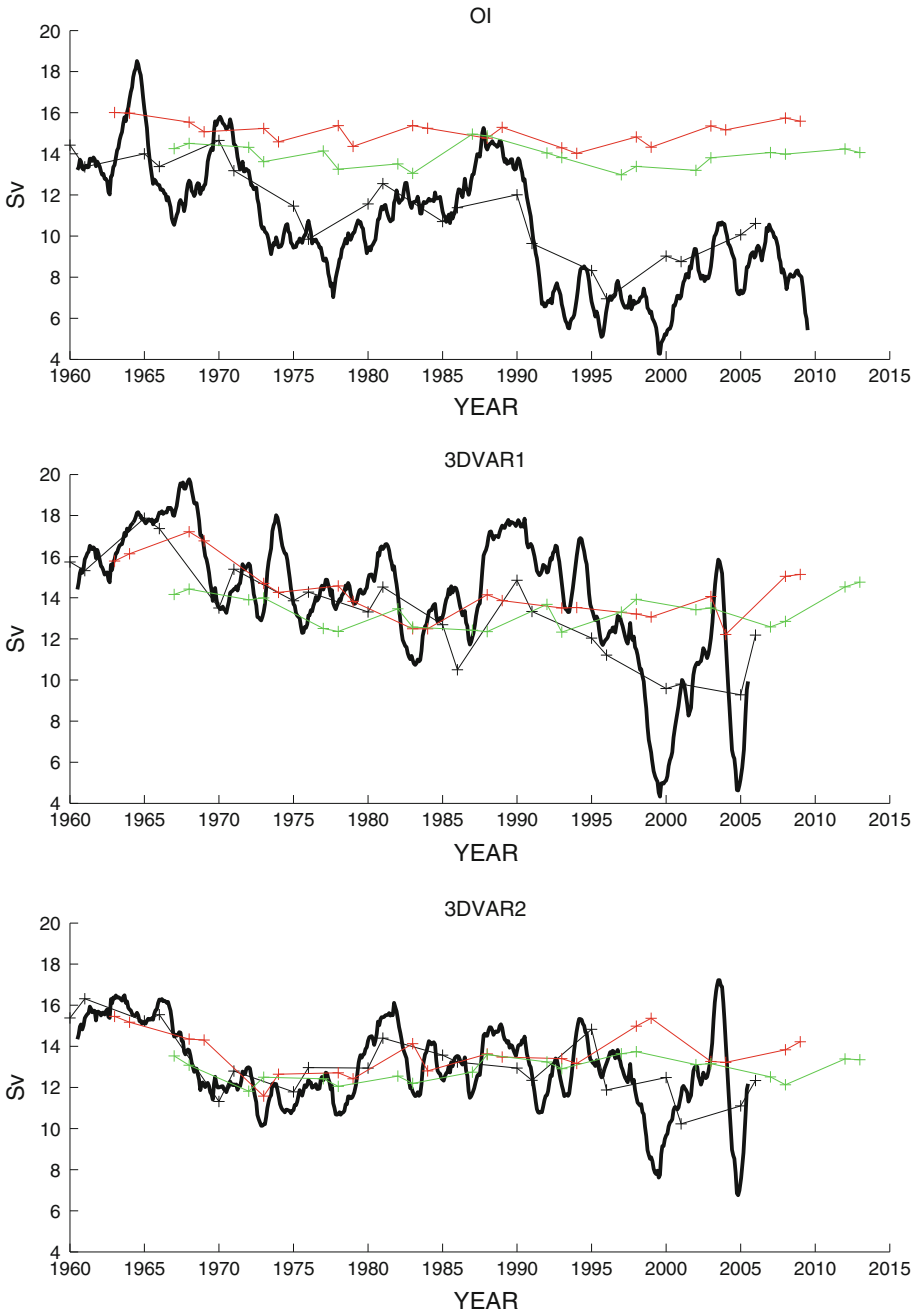

Fig. 10 Atlantic meridional overturning circulation at 26°N (Sv) in ocean reanalysis (*black solid*) and 1 year (*black plus*), 2–5 years (*red plus*) and 6–9 years (*green plus*) predictions under (*top*) OI, (*middle*) 3DVAR1 and (*bottom*) 3DVAR2 initial conditions

data assimilation products is reported in Balmaseda et al. (2007) for ECMWF operational ocean reanalyses.

In Fig. 12, the anomaly correlation between predicted and "observed" (i.e., from reanalyses) annual mean MOC at several lead times and across latitudes is shown, for each initialization set. Compared to OI, the 3DVAR-initialized predictions display an overall longer term predictability skill (up to 3–4 years, versus 1–2 years in OI) and for a wider range of latitudes. In turn, 3DVAR2, when contrasted to 3DVAR1, shows a lower skill in the northern North Atlantic (north of 50N), but larger skill south of the Equator.

Interestingly, MOC and AMO display a largely consistent level of predictability, both showing predictive skill in the 2–5 years range, particularly under 3DVAR initialization. This result is in agreement with a well-documented low frequency co-variability of the Atlantic SST field and MOC, found in both models and proxy-based observational reconstructions (Delworth and Mann 2000; Knight et al. 2005; Collins et al. 2006). In the widely accepted paradigm of the MOC being a major driver of the Atlantic SST multi-decadal variability, the initialization of the meridional overturning circulation turns out to be a crucial aspect for skilful decadal predictions in the Atlantic sector.

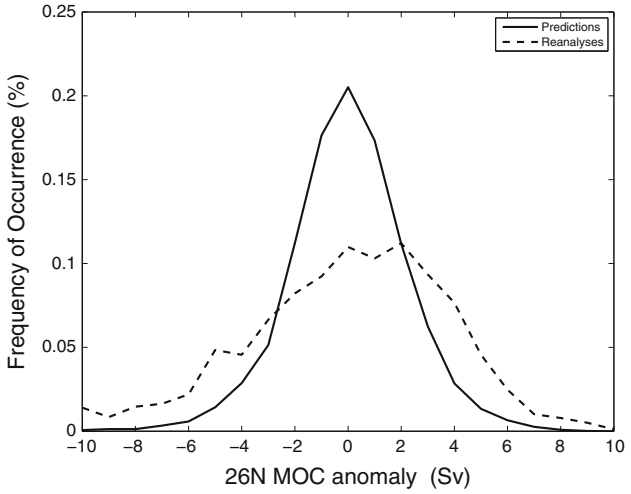

Fig. 11 Probability density function of Atlantic meridional overturning circulation anomalies at 26°N (Sv), from the full set of predictions (*solid*) and ocean renalyses (*dashed*). Model overturning anomalies were obtained by removing the average forecast from each single raw forecast

8 Pacific decadal oscillation

Predictability skill in the North Pacific region is assessed against the dominant pattern of internal variability over the Pacific area, the so-called Pacific Decadal Oscillation (PDO). PDO is a winter time climate variability mode (Mantua et al. 1997) featuring a large-scale covariance pattern, involving mid-latitude and tropical regions in the North Pacific, with fluctuations at decadal and inter-decadal time scales. Following Mantua et al. (1997) we define a PDO index based on the leading Empirical Orthogonal Function/Principal Component (EOF1/PC1) of winter-mean (January–February–March;JFM) SST, north of 20N. In the present analysis, only January-initialized predictions were used, as these have been integrated for 30 years, unlike November-initialized simulations which were mainly integrated for 10 years only. Since 10 years are not sufficient to constrain a decadal-scale variability signal, only the longer 30-years subset of simulations was used for detecting the PDO signature. It should be emphasized that PDO does not always emerge as the dominant EOF in the simulated North Pacific SSTs, but rather as the second variability mode (EOF2). For those specific cases, the corresponding PC2 was used as a proxy of the PDO index.

Hindcast/forecast of the PDO index for different start dates is shown in Fig. 13, for ensemble mean SSTs. A simple visual inspection reveals a rapid loss of coherency between the observed and predicted PDO signal. This is further confirmed by correlation between observed and ensemble-mean PDO index at several lead-times (Fig. 14), showing significant correlation (around 0.9) only for 1-year lead time, followed by an abrupt decay below statistically significant levels for longer lead times. When predictive skill is evaluated against different initialization data sets (Fig. 14) large differences are found between OI and 3DVAR initialization experiments. In particular, OI shows the lowest skill scores (never reaching statistical significance), while 3DVAR2 displays a rather counter-intuitive behavior, with the highest score in the long 6–9 years term. The ensemble-mean procedure enhances the predictive skill for the 1-year and, partly, for the 2–5 year lead times, but not in the longer time range. The overall poor predictive capability associated with the PDO in the multi-annual

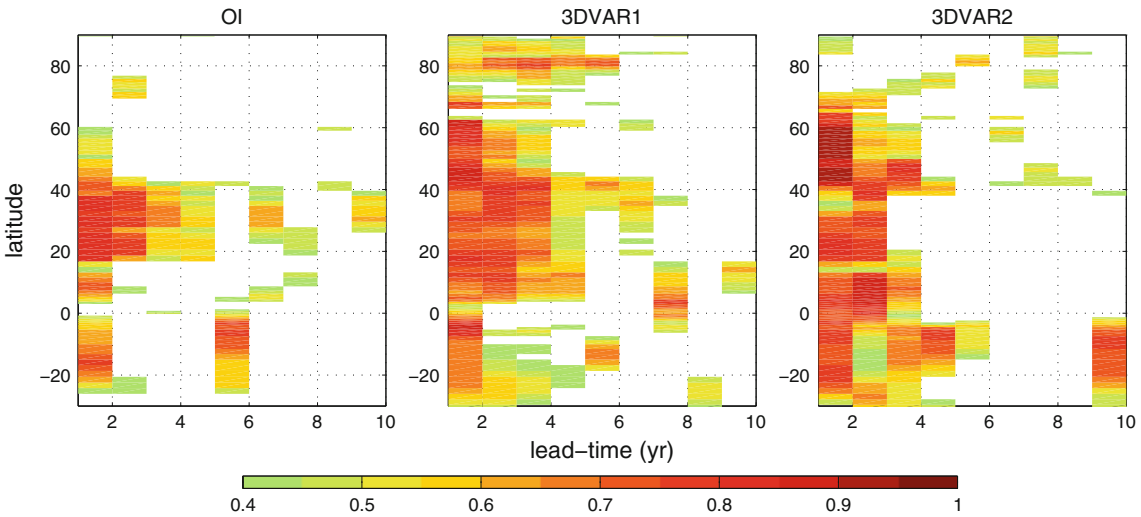

Fig. 12 Anomaly correlation of the Atlantic meridional overturning circulation for predictions initialized with OI (*left*), 3DVAR1 (*middle*) and 3DVAR2 (*right*) reanalysis, at different lead-times (*horizontal axis*) and latitudes (*vertical axis*). Each prediction set is verified against the corresponding initialization reanalysis, which is used as truth. Only statistically significant values at the 95 % level according to a one-sided Student *t* test are shown

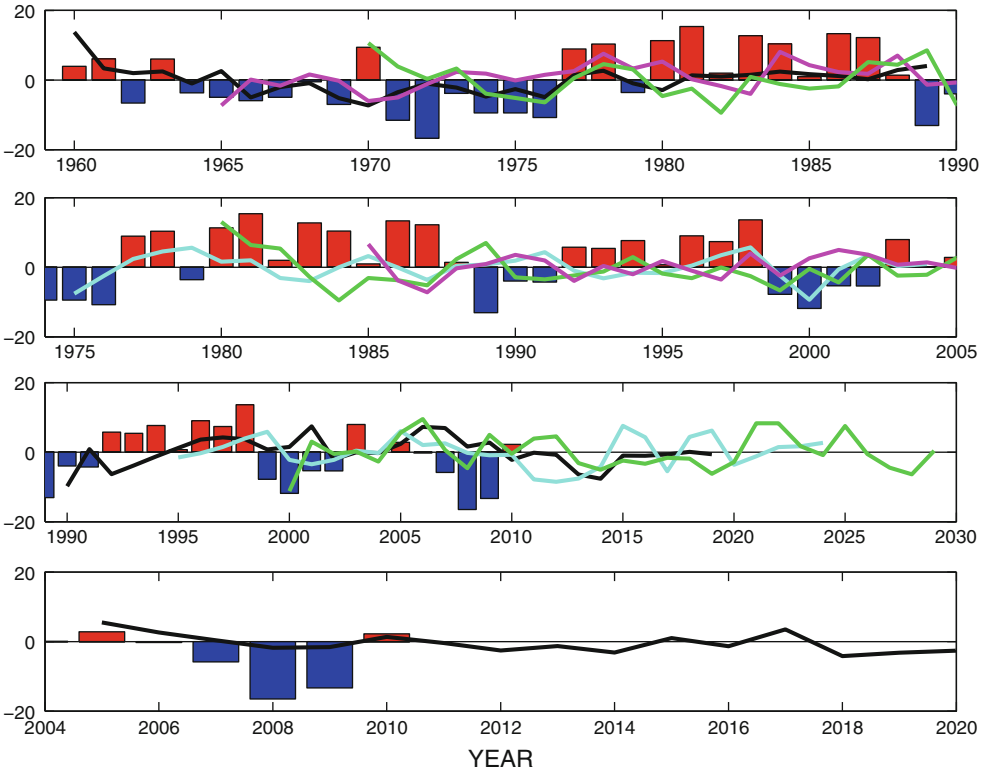

Fig. 13 Time series of PDO index in observations (HadISST; *bars*) and hindcast/forecast (*colour curves*) computed as the leading principal component of winter mean (JFM) SST in the Pacific basin, north of 20N. Predicted PDO is evaluated from ensemble-mean SSTs. Different panels refer to different temporal segments of the hindcast/forecast period

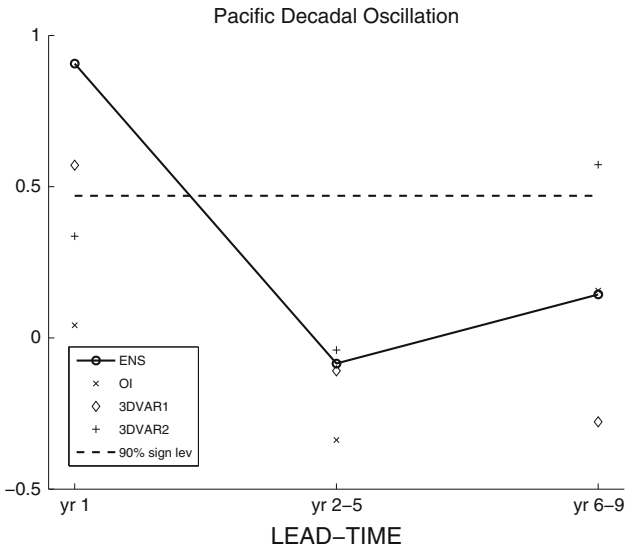

Fig. 14 Correlations between ensemble-mean (*circle*), OI- (*cross*), 3DVAR1- (*diamond*), 3DVAR2- (*plus*) initialized predictions and observed (HadISST) PDO index, at several lead-times.The 90 % statistical significance threshold according to a one-sided Student *t* test is also shown (*dashed*)

range is consistent with the relatively low skill score found in the ACC patterns over wide areas of the North Pacific sector (Figs. 3, 4). Mochizuki et al. (2010) show a similarly low SST predictability in the extratropical North Pacific, but a sensibly larger skill when vertically averaged temperatures over the top 300 m are considered (their Fig. 2). The rapid loss of skill featured by the PDO index for lead times longer than 1 year has also been found in an ongoing analysis of CMIP5 models (Kim 2011, personal communication).

9 Summary and conclusions

The skill of a set of CMIP5 near-term predictions performed with a state-of-the-art ocean-atmosphere general circulation model (OAGCM) has been analysed. In particular, the predictive skill dependency on ocean initialization has been examined by using three distinct ocean reanalyses as initialization data set, differing in the data assimilation methodology and the amount of the assimilated observations. A large fraction of SST skill is accounted for by the global upward trend in surface temperatures, while contributions to near-term predictability from ENSO are virtually absent due to the poor sampling of the historical record (few starting dates which accidentally miss most of the major ENSO episodes in the 1960–2005 period). Since no volcanic activity is included in the model forcing, any residual skill (i.e. computed after removing the linear trend from both the predictions and the observations) must be mainly associated with initialization, although some spurious skill caused by the non-linearity of the actual trend may still exist (see van Oldenborgh et al.

2012 for a detailed discussion on this issue). Statistically significant predictability after detrending is found over extensive regions of the North Atlantic both in the near (2–5 year) and long (6–9 year) term. Some skill re-emergence in the equatorial Pacific, together with enhanced predictability in the northern North Pacific is also found in the 6–9 years range. The ACC patterns for SSTs from the present set of predictions reveal strong analogies with a similar analysis made for the ENSEMBLES multi-model set (van Oldenborgh et al. 2012; Doblas-Reyes et al. 2011b) showing that, to a large extent, there is consistency in the predictive capabilities displayed by a wide range of climate models. The skill score patterns for detrended SSTs provide a spatial mapping of predictability associated with the unforced component of climate variability driven by natural fluctuations, which can be ascribed solely to the internal dynamics of the climate system. The emerging picture corroborates a feature revealed in model-only potential predictability studies (Boer 2000) as well as in realistically initialized decadal predictions efforts (Pohlmann et al. 2009; Keenlyside et al. 2008; van Oldenborgh et al. 2011). Both approaches indicate the North Atlantic as a region exhibiting enhanced predictability at multi-annual time scales. Consistent with this finding, indices of Atlantic multidecadal variability, including an interhemispheric Atlantic SST dipole index, and a detrended version of the AMO index, are skilfully predicted in the 2–5 year range, and (marginally, for the dipole index) in the 6–9 year range, by model hindcasts. According to a widely accepted paradigm based on model evidence and theoretical speculation (Knight et al. 2005), AMO appears to be related to coherent changes in the Atlantic THC. In this study, the Atlantic MOC shows a degree of predictability up to 3–4 years over a wide range of latitudes. Thus an overall self-consistent picture emerges from this study, connecting the predictive skill of (detrended) North Atlantic SSTs with predictability of internal variability modes involving large-scale coherent SST and THC changes in the Atlantic sector.

A substantial asymmetry appears when contrasting the predictive skill in the Atlantic area with the North Pacific, the latter (evaluated through the PDO index) showing no skill beyond 1 year lead time. This result may point to model deficiencies in representing the essential processes underlying decadal fluctuations in the extra-tropical North Pacific. To the extent that PDO is an ENSO-driven process (Newman et al. 2003), misrepresentations of tropical-extratropical interactions may negatively affect the internal variability over the North Pacific region in the coupled model (Cherchi et al. 2012).

The envisaged predictive capabilities beyond the anthropogenic trend, involving the slow dynamics of unforced components of climate variability, support the usefulness of model initialization practices aimed at performing reliable decadal predictions (within the predictability limits that are inherent to the specific dynamical model in use; Branstator and Teng 2010). Appropriately constraining the initial state of the dynamical model becomes crucial in limiting the error growth associated with the chaotic nature of the climate system. Here, an ensemble of three ocean reanalyses, based on different data assimilation techniques and assimilated datasets, was used with the purpose of initializing 3-member ensemble hindcasts/forecasts and thus providing the initial state perturbations that are needed for generating an ensemble members spread. Ranking climate predictions on the basis of the specific initialization reanalysis reveals a strong predictive skill dependency on the ocean data assimilation product used to constrain the oceanic initial conditions. In particular hints of an improved predictive skill associated with 3DVAR1 and 3DVAR2 initializations, compared to OI, have been found. On the other hand, no significant differences were found when contrasting the two 3DVAR implementations. This result suggests that the details regarding the background error parameterization have little effect on the simulated evolution of the climate state, both methods providing *equivalent* (as fas as predictability is concerned) representations of the initial state of the ocean. Bearing in mind that OI only assimilates in-situ temperature and salinity profiles, while 3DVAR reanalyses additionally assimilate sea-level anomaly data, the large divergences between 3DVAR and OI may, in principle, reflect differences both in the applied assimilation methodology and in the amount of assimilated data. However, specifically for the Atlantic MOC, some indications have been found of a minor influence of SLA assimilation, with respect to the use of OI versus a 3DVAR assimilation technique. On the methodological side, the vertically increasing background error horizontal correlation lenght-scale featured by 3DVAR, against a constant one in OI, might improve the representation of deep ocean circulation features (and therefore, of the MOC) by extending the radius of influence of the observations in the deeper oceanic layers.

Viewing this study in a wider perspective, the large uncertainties affecting the currently available oceanic reanalyses (Carton and Santorelli 2008; Munoz et al. 2011) are likely to produce a similarly large spread in climate predictions (Kröger et al. 2012). While observable fields such as SST and SLA are well constrained through synoptic satellite-based observing systems, the severe undersampling, in space and time, of sub-surface ocean (which has only recently started to be alleviated by the ARGO profiling floats programme) may be a major reason for the large spread affecting dynamically connected fields such as MOC, ocean heat transport and heat content in ocean

reanalyses. In particular, the scarcity of ocean observational data may strengthen the dependency of oceanic reanalyses on the adopted dynamical model and the details of the data assimilation method.

In order to provide skilful near-term climate predictions in the foreseeable future, a substantial reduction of the uncertainties that currently impoverish our knowledge of the oceanic state is mandatory. This requires a major effort involving a wide range of expertise, including ocean data assimilation and observational communities.

Acknowledgments The authors gratefully acknowledge the support from the EU FP7 COMBINE Project (Grant Agreement number 226520) and the Italian Ministry of Education, University and Research and Ministry for Environment, Land and Sea through the Project GEMINA. Stimulating discussions with Jurgen Kröger, Rein Haarsma and Andrea Borrelli, and insightful comments on the manuscript by Panos Athanasiadis are also acknowledged.

References

Balmaseda MA, Smith GC, Haines K, Anderson D, Palmer TN, Vidard A (2007) Historical reconstruction of the Atlantic meridional overturning circulation from the ECMWF operational ocean reanalyses. Geophys Res Lett 34:L23615

Balmaseda M, Mogensen K, Molteni F, Weaver AT (2010) The NEMOVAR-COMBINE ocean reanalyses. COMBINE Tech Rep 1

Branstator G, Teng H (2010) Two limits of initial-value decadal predictability in a CGCM. J Clim 23:6292–6311

Bellucci A, Masina S, Di Pietro P, Navarra A (2007) Using temperature salinity relations in a global ocean implementation of a multivariate data assimilation scheme. Mon Wea Rev 135:3785–3807

Boer GJ (2000) A study of atmosphere-ocean predictability on long timescales. Clim Dyn 16:469–477

Brohan P, Kennedy JJ, 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. doi:10.1029/2005JD006548

Bryden HL, Longworth HR, Cunningham SA (2005) Slowing of the Atlantic meridional overturning circulation at 25 N. Nature 438:655–657

Carton J, Santorelli A (2008) Global decadal upper-ocean heat content as viewed in nine analyses. J Clim 21:6015–6035

Cherchi A, Masina S, Navarra A (2012) Tropical Pacific-North Pacific teleconnection in a coupled GCM: remote and local effects. Int J Climatol (in press). doi:10.1002/joc.2379

Collins M et al (2006) Interannual to decadal climate predictability in the North Atlantic: a multimodel-ensemble study. J Clim 19:1195–1203

Cunningham SA et al (2007) Temporal variability of the Atlantic meridional overturning circulation at 26.5N. Science 317: 938–941

Delworth TL, Mann ME (2000) Observed and simulated multidecadal variability in the Northern Hemisphere. Clim Dyn 16:661–676

De Mey P, Benkiran M (2002) A multi-variate reduced-order optimal interpolation method and its application to the Mediterranean basin-scale circulation. In: Pinardi N, Woods JD (eds) Ocean forecasting: conceptual basis and applications. Springer, Berlin, pp 281–306

Doblas-Reyes FJ, Balmaseda MA, Weisheimer A, Palmer TN, (2011a) Decadal climate prediction with the ECMWF coupled forecast system: impact of ocean observations. J Geophys Res 116. doi:10.1029/2010JD015394

Doblas-Reyes FJ , van Oldenborgh GJ, Garcia-Serrano J, Pohlmann H, Scaife AA, Smith D (2011b) CMIP5 near-term climate prediction. CLIVAR Exch 56 16(2):8–11

Dobricic S, Pinardi N (2008) An oceanographic three-dimensional assimilation scheme. Ocean Model 22:89–105

Enfield DB, Mestas-Nunez AM, Trimble PJ (2001) The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S. Geophys Res Lett 28:2077–2080

Fichefet T, Morales-Maqueda MA (1999) Modeling the inuence of snow accumulation and snow-ice formation on the seasonal cycle of the Antarctic sea-ice cover. Clim Dyn 15:251–268

Griffies SM, Bryan K (1997) A predictability study of simulated North Atlantic multidecadal variability. Clim Dyn 13:459–487

Ingleby B, Huddleston M (2007) Quality control of ocean temperature and salinity profiles. Historical and real-time data. J Mar Syst 65:158–175

Kanzow T, Hirschi JJ-M, Meinen CS, Rayner D, Cunningham SA, Marotzke J, Johns WE, Bryden HL, Beal LM, Baringer MO (2008) A prototype system of observing the Atlantic meridional overturning circulation—scientific basis, measurement and risk mitigation strategies, and first results. J Oper Oceanogr 1:19–28

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

Knight JR, Allan R, Folland CK, Vellinga M, Mann ME (2005) A signature of persistent natural thermohaline circulation cycle in observed climate. Geophys Res Lett 32:L20708

Kröger J, Müller W, von Storch J-S (2012) Impact of different ocean reanalyses on decadal climate prediction. Clim Dyn. doi:10.1007/s00382-012-1310-7

Le Traon P, Nadal F, Ducet N (1998) An improved mapping method of multisatellite altimeter data. J Atmos Ocean Technol 15:522–534

Madec G, Delecluse P, Imbard M, Lvy C (1998) OPA 8.1 ocean general circulation model reference manual. Note du Pole de Modlisation 11, Institut Pierre-Simon Laplace, 91 p

Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis R (1997) A Pacific interdecadal climate oscillation with impacts on salmon production. Bull Am Meteor Soc 78:1069–1079

Masina S, Di Pietro P, Storto A, Navarra A (2011) Global ocean reanalyses for climate applications. Dyn Atm Oceans 52:341–366

Meehl GA, Goddard L, Murphy J, Stouffer RJ, Boer G, Danabasoglu G, Dixon K, Giorgetta MA, Greene AM, 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 Meteor Soc 90:1467–1485

Mochizuki T et al (2010) Pacific decadal oscillation hindcasts relevant to near-term climate prediction. PNAS 5:1833–1837

Muñoz E, Kirtman B, Weijer W (2011) Varied representation of the Atlantic meridional overturning across multidecadal ocean reanalyses. Deep-Sea Res II 58:1848–1857

Murphy AH (1988) Skill scores based on the mean square error and their relationships to the correlation coefficient. Mon Wea Rev 116:2417–2424

Newman M, Compo GP, Alexander MA (2003) ENSO-forced variability of the pacic decadal oscillation. J Clim 16:3853–3857

Palmer TN, Buizza R, Doblas-Reyes F, Jung T, Leutbecher M, Shutts GJ, Stenheimer M, Weisheimer A (2009) Stochastic parametrization and model uncertainty. Tech Rep No 598, ECMWF Tech Memo

Pohlmann H, Botzet M, Latif M, Roesch A, Wild M, Tschuck P (2004) Estimating the decadal predictability of a coupled AOGCM. J Clim 17:4463–4472

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

Pohlmann H, Smith D, Balmaseda M, Keenlyside N, Masina S, Matei D, Müller W, Rogel P, da Costa E (2012) Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system. Clim Dyn (under review)

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(D14):4407. doi:10.1029/2002JD002670

Roeckner E et al (2003) The atmospheric general circulation model ECHAM5. Part I: model description. MPI Rep 349, p 127

Schlesinger ME, Ramankutty N (1994) An oscillation in the global climate system of period 65–70 years. Nature 367:723–726

Schneider U, Fuchs T, Meyer-Christoffer A, Rudolf B (2008) Global precipitation analysis products of the GPCC. Tech. Rep., Global precipitation climatology centre (GPCC), Deutscher Wetterdienst, Offenbach, Germany

Scoccimarro E, Gualdi S, Bellucci A, Sanna A, Fogli PG, Manzini E, Vichi M, Oddo P, Navarra A (2011) Effects of tropical cyclones on ocean heat transport in a high resolution coupled general circulation model. J Clim 24:4368–4384

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

Storto A, Randriamampianina R (2010) Ensemble variational assimilation for the representation of background error covariances in a high latitude regional model. J Geophy Res 115:D17204

Storto A, Masina S, Dobricic S (2011a) Ensemble spread-based assessment of observation impact: application to a global ocean analysis system. Q J R Meteorol Soc (submitted)

Storto A, Dobricic S, Masina S, Di Pietro A (2011b) Assimilating along-track altimetric observations through local hydrostatic adjustment in a global ocean variational assimilation system. Mon Wea Rev 139:738–754

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

Trenberth KE et al (2007) Observations: surface and atmospheric climate change. 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

Uppala S et al (2005) The ERA-40 reanalysis. Q J R Meteorol Soc 131:2961–3012

van Oldenborgh GJ, te Raa LA, Dijkstra HA, Philip SY (2009) Frequency- or amplitude-dependent effects of the Atlantic meridional overturning on the tropical Pacific Ocean. Ocean Sci 5:293–301

van Oldenborgh GJ, Doblas-Reyes FJ, Wouters B, Hazeleger W (2012) Decadal prediction skill in a multi-model ensemble. Clim Dyn 38:1263–1280. doi:10.1007/s00382-012-1313-4

Wijffels SE, Willis J, Domingues CM, Barker P, White NJ, Gronell A, Ridgway K, Church JA (2008) Changing expendable bathythermograph fall rates and their impact on estimates of thermosteric sea level rise. J Clim 21:5657–5672

Zhu J, Huang B, Marx L, Kinter JL III, Balmaseda MA, Zhang R-H, Hu Z-Z (2012) Ensemble ENSO hindcasts initialized from multiple ocean analyses. Geophys Res Lett 39:L09602