CrossMark

# Comparison of full field and anomaly initialisation for decadal climate prediction: towards an optimal consistency between the ocean and sea-ice anomaly initialisation state

**Danila Volpi[1,2] · Virginie Guemas[1,3] · Francisco J. Doblas-Reyes[1,4]**

**Abstract** Decadal prediction exploits sources of predictability from both the internal variability through the initialisation of the climate model from observational estimates, and the external radiative forcings. When a model is initialised with the observed state at the initial time step (Full Field Initialisation—FFI), the forecast run drifts towards the biased model climate. Distinguishing between the climate signal to be predicted and the model drift is a challenging task, because the application of a-posteriori bias correction has the risk of removing part of the variability signal. The anomaly initialisation (AI) technique aims at addressing the drift issue by answering the following question: if the model is allowed to start close to its own attractor (i.e. its biased world), but the phase of the simulated variability is constrained toward the contemporaneous observed one at the initialisation time, does the prediction skill improve? The relative merits of the FFI and AI techniques applied respectively to the ocean component and the ocean and sea ice components simultaneously in the EC-Earth global coupled model are assessed. For both strategies the initialised hindcasts show better skill than historical simulations for the ocean heat content and AMOC along the first two forecast years, for sea ice and PDO along the first forecast year, while for AMO the improvements are statistically significant for the first two forecast years. The AI in the ocean and sea ice components significantly improves the skill of the Arctic sea surface temperature over the FFI.

## 1 Introduction

Society needs suitable strategies for adaptation to climate change in the short to medium term (Trenberth 2008). Decadal prediction aims at responding to such a need by providing interannual climate information. It bridges seasonal forecasts and climate change projections as it focuses on intermediate time scales of several years to a few decades (e.g. Oldenborgh et al. 2011). The added value of decadal climate forecasts over climate-change projections is that the former are initialised to take into account the best knowledge of the state of the climate system at the start of the prediction (Doblas-Reyes et al. 2013a). This allows to exploit the climate predictability arising from internal variability. Due to the relative novelty of the climate prediction field and the lack of knowledge about the best method to initialise the decadal predictions, several alternatives are currently being explored. In broad terms, the initialisation is performed using either the best estimate of the observed climate system (full field initialisation, FFI (Pohlmann et al. 2009)) or using the assimilation of observed anomaly variables onto an estimate of the model mean climate

✉ Danila Volpi
danila.volpi@bsc.es

[1] Barcelona Supercomputing Center, Barcelona, Spain

[2] Department of Mathematics and Statistics, University of Reading, Reading, UK

[3] Centre National de Recherches Météorologiques/Groupe d'Etude de l'Atmosphère Météorologique, Météo-France,CNRS, Toulouse, France

[4] ICREA, Pg. Lluis Companys 23, 08010 Barcelona, Spain

_Springer

(anomaly initialisation, AI (Smith et al. 2007)). Each of these initial states has its respective benefits and drawbacks: on one hand, when the model is initialised with FFI the forecast run drifts towards the model climate because of its climate biases which develop on various timescales (Magnusson et al. 2012; Doblas-Reyes et al. 2010). The FFI is commonly performed also for seasonal climate predictions as has the advantages of starting direclty from the observed state, beeing easy to implement and does not have issues of consistency with the climate of the model, as it could happen with AI. However, at such timescales the climate prediction drift, defined as the evolution of the difference between the model and the observed climatology,[1] does not stand as a major issue, whereas it does for decadal forecasts. Sanchez-Gomez et al. (2015) provide a physical and dynamical description of the drifts in the CNRM-CM5 coupled model in decadal hindcasts. They found an initial shock in the Pacific that causes an excitation of ENSO warm events. This also affects the atmospheric drift over a large part of the planet through teleconnection.

Thus at decadal time scale, the climate signal may be polluted by the drift and making a distiction between the two is a challenging task; the application of a-posteriori bias correction has the risk of removing part of the variability signal. On the other hand, AI relies on the hypothesis that the model variability is independent of the model mean state, and therefore only the phase of the simulated variability is constrained towards the contemporaneous observed one. If such hypothesis is valid, it follows that even with its biases, a model could predict correctly the superimposed observed variability. Beyond the fact that this hypothesis has not been proven, its limitation is that if the biases include a spatial shift, the anomalies will indeed not be located at the correct place with respect to the climatology and therefore the amplitude of the variability would differ when the mean state differs. Moreover, since the initial state of AI is an artificial state, there is the risk of generating inconsistency between the different variables and model components.

Hazeleger et al. (2013) compared the two initialisation techniques in the context of the CMIP5[2] exercise, using EC-Earth v2.3 (Hazeleger et al. 2010). The FFI was applied by replacing the model state at the initialisation time with an estimate of the full field observed initial state (from reanalysis). Analogously, the AI was performed by replacing the model state of ocean horizontal velocities, temperature and salinity with a state composed by their respective observed anomalies at the initial time added to the model climatology. The same technique was applied to the sea ice component. The results showed that even though the hindcasts initialised from AI had a smaller drift than FFI, a drift was still present and the AI did not lead to any robust increase in skill. The FFI improved the skill in the North Atlantic over both AI and the historical run, but for the global two metres temperature and for the sea surface temperature in the Pacific the historical run performed the best. However, the fact that only some of the ocean variables were anomaly initialised (and not the full ocean state) could have lead to some inconsistencies because some variables were starting from the model attractor and others from the observed attractor.

Smith et al. (2013) made a similar comparison using the HadCM3 (Gordon et al. 2000) model, and a different assimilation approach called nudging: two sets of hindcasts were initialised from two different assimilation runs. One of these assimilation runs consisted in relaxing the atmosphere winds, potential temperature and surface pressure, plus the ocean temperature and salinity towards their respective observational values. In the second assimilation run the same variables as above were relaxed towards the respective observed anomalies, added to the model climatologies. The added value brought by the nudging is the consistency between the variables within a single model component and between model components. However, the model is initialised with a state which is close to the ideal one (i.e. the observed phase of climate variability), but it is not exactly the observed one. The results for multi-annual predictions did not show any statistically significant difference between the two methods. On seasonal timescales the FFI provides more skilful regional predictions than the AI, for both temperature and precipitation.

In this work, we assess the multi-year prediction skill of hindcasts initialised with the same assimilation techniques as in Hazeleger et al. (2013), namely the FFI and the AI replacement. In addition, we attempt at having an ocean and sea ice states as consistent as possible, by initialising all the prognostic variables in the ocean and sea ice components and applying thorough corrections at their interface. We address the following questions:

1. Does this AI method applied to all the prognostic variables prevent the initial drift? How does this translate into a potentially better prediction skill than FFI?
2. What is the added-value of using a sea ice initial state consistent with the ocean initial state?

Section 2 describes the initialisation procedures, the model in use and the hindcasts set-up. The skill comparison between the FFI and the AI hindcasts are presented in

---

[1] We compute the climatology as the average of the predictions over the start dates and it evolves with the forecast time.

[2] Coupled Model Intercomparison Project Phase 5 sponsored by WCRP: http://cmip-pcmdi.llnl.gov/cmip5/index.html.

Sect. 3, while the discussion and the conclusions are drawn in Sect. 4.

## 2 Methodology

### 2.1 Climate model

The model in use is the coupled general circulation model EC-Earth version 2.3 (Hazeleger et al. 2010). The atmospheric component is based on the European Centre for Medium-Range Weather Forecasts integrated forecasting system (IFS), with 62 vertical levels and a TL159 horizontal resolution. The ocean component is the NEMO model (Madec 2008; Ethe et al. 2006), with ORCA1 configuration (about 1 degree with enhanced tropical resolution) and 42 vertical levels. The sea-ice component is LIM2 (Fichefet and Maqueda 1997; Goosse and Fichefet 1999) directly embedded into NEMO. The atmospheric and ocean components are coupled via OASIS3 (Valcke 2006). Information on the atmospheric chemistry and the dynamic vegetation are prescribed from observations. The atmospheric top is at 5 hPa, so only the lower stratosphere is resolved. EC-Earth v2.3 develops a cold sea surface temperature bias in the Tropical Pacific and the North Atlantic, with peaks of $-4$ to $-8\,^{\circ}$C, and a warm bias in the Southern Ocean with peaks of 2–4 °C. The biases maps of figure S1 of the Supplementary Material support the results found in Hazeleger et al. (2013): the regional pattern is similar, with the strongest cold bias in the tropical regions and in the North Atlantic subpolar gyre, that for its slow evolution is attributed to oceanic processes. A warm bias is found in the Southern Ocean, that is associated with atmospheric processes for its quick development. Moreover, Du et al. (2012) showed that the model underestimates precipitation and drifts towards a less dry condition after the first forecast year. In terms of ocean heat content, they also show that the model develops a warm bias in the mixed layer during the boreal winter and a cold bias in the boreal summer. A substantial bias is also seen in the AMOC, consistently with the findings of this study.

### 2.2 Reference simulations: the NOINI and the FFI hindcasts

The benchmark hindcasts of this work are the FFI experiment of Du et al. (2012) that is part of the CMIP5 exercise, and an uninitialised model experiment. In the FFI experiment, all the variables from each model component are initialised. The atmosphere and land surface initial conditions are taken from the ERA-40 reanalysis (Uppala et al. 2005) for start dates before 1989 and ERA-Interim (Dee et al. 2011) afterwards. The ocean initial conditions are taken

from NEMOVAR-ORAS4 (Mogensen et al. 2012) which is forced with ERA40 and ERA-Interim, while the sea-ice initial conditions are produced with a simulation using NEMO v2.0 coupled to LIM2 driven by DFS4.3 ocean forcing data (Brodeau et al. 2009). The DFS4.3 forcing data are derived from ERA40 data with tropical surface air humidity, Arctic sea surface temperature and global wind field corrections based on high-quality observations. Thus, we consider that the oceanic and sea-ice initial conditions are fairly coherent.

The observed volcanic and anthropogenic aerosol load and greenhouse gas concentration are prescribed using observed values up to 2005. After that date the RCP4.5 scenario was used, as well as a background solar irradiance level and a constant background volcanic aerosol load. Every 2 years between 1960 and 2004, on the 1st of November, a set of 5 simulations (referred to as members here) were started and run for 5 years. The choice of having one start date every 2 years is a good compromise between the computational cost and having enough data to compute robust statistics. These 5 members start from atmosphere initial perturbations based on singular vectors (Magnusson et al. 2008), which are added at the initial time to all the prognostic variables except for humidity (Du et al. 2012).

Each of the ocean member is initialised from a different 3D-Var ensemble member of NEMOVAR-ORA_S4 ocean re-analysis.

The uninitialised experiment, called NOINI, is a 3-member hindcast derived from long historical simulations up to 2005, and simulations following the representative concentration pathways 4.5 (RCP4.5) after 2006 produced in the framework of CMIP5. In the NOINI experiment, the internal variability is not in phase with the observed variability since each member has been initialised in 1850 from a different date of a pre-industrial control simulation. However, based on recent literature (Otterå et al. 2010; Swingedouw et al. 2015), part of the internal variability could be phased through their response to volcanic eruptions. Anthropogenic aerosols have also been proposed to influence decadal variability (Booth et al. 2013). The NOINI experiment as well as all the experiments implemented in this study, employs the same external radiative forcings as described for the FFI.

### 2.3 The anomaly initialised simulations

To answer the questions listed in the Introduction, two experiments with AI have been implemented:

- assuming that most of the memory is held by the ocean initial states (Old and Haines 2006), an ocean AI (referred as O-AI thereafter) configuration has been implemented. In this case the ocean model component

has been anomaly initialised and the atmosphere and sea ice components have been full field initialised

- to account for the importance of a consistent ocean and sea ice initial state, an ocean and sea ice AI (referred to as OSI-AI thereafter) configuration has also been implemented. The ocean and sea ice model components have been anomaly initialised and the atmosphere component has been full field initialised.

The initial state is obtained by replacing the model anomalies with the observed ones. The term "model anomalies" refers to the difference between the contemporaneous state and its long-term average, defined over a predefined period of time, that in this study has been chosen to be 1971–2000.

The O-AI experiment has the atmosphere and the sea ice components initialised as in the FFI experiment, while the ocean is anomaly initialised. The reference data for initialising the ocean is the five-member NEMOVAR-ORAS4 ocean reanalysis (Mogensen et al. 2012; Balmaseda et al. 2012).

In the OSI-AI experiment the AI is implemented for the ocean and sea ice model components. The reference data used for the sea ice initialisation is the 5-member sea ice reconstruction proposed by Guemas et al. (2014), that covers the period 1958–2006. This reconstruction, called HistDfsNudg, has been produced with LIM2, the sea ice model included in EC-Earth 2.3, forced with ERA-Interim and nudged toward NEMOVAR-ORAS4. It is biased with respect to the HadISST sea-ice extent, but it is still the best estimate of sea-ice cover (in particular is the best estimate of sea-ice thickness available) that could be used to initialise LIM2, as it has the same grid resolution and it provides estimates for all the sea-ice model variables. The OSI-AI experiment makes use of sea ice initial perturbations, by using different members for the sea ice initial conditions. The hindcasts of the O-AI and OSI-AI experiments are respectively composed by a set of 5-member, 5-year long simulations, starting every 2 years on the 1st of November as the FFI experiment.

## 2.4 Ensuring the physical consistency

The sea-ice variables are sensitive to the geographical position where the anomalies are placed, especially because most of them are bounded. Moreover, the variables from the sea-ice model component and the ocean need to be consistent in the OSI-AI initialisation. The sea-ice and the snow thickness cannot be negative nor the ice concentration; the SST must be at the freezing point ($-1.865\,°$C) in presence of sea ice. However, correcting the SST by cooling it down could result in having a column of water with temperature cooler at the surface than the levels below. This would trigger ocean convection and cause the formation of

deep water mass. The consequence would be a potential strong alteration of the thermohaline circulation. In order to avoid this the aim has been to create a quasi-stable profile. The simplest way for that would have been to cool down also all the levels below the SST which are warmer than the first level. However, this assumes that the temperature cools down with depth which is the case in most places of the ocean, but not in the Arctic, where the temperature cools down with depth close to the surface and it warms up in the bottom layers (and its salinity increases). To minimise the alteration of the temperature profile but at the same time avoid the creation of ocean convection, only the subsurface water (arbitrary chosen down to 25 metres) has been cooled down to the freezing temperature in the presence of sea ice. Finally, where the ice thickness and the ice concentration are zero, when the temperature at the surface was below the freezing point, it has been warmed up to 0.01 °C above the freezing point.

## 2.5 Skill assessment

The metrics that we used to evaluate the skill of the hindcasts experiments are the anomaly correlation (AC) and the Root Mean Square Error (RMSE) as a function of the forecast time $f$ applied to the ensemble mean forecast anomalies. The forecast anomalies (not to be confused with the model anomalies defined in Sect. 2.3) are calculated by subtracting the forecast climatology to each hindcast. The forecast climatology at each grid point depends on the forecast time. It is estimated by averaging the hindcast variable across the starting dates and the members using only hindcast values for which observations are available at the corresponding dates. This data selection process is referred to as per-pair (García-Serrano and Doblas-Reyes 2012). The implementation of the per-pair method guarantees the use of all the observational data available and at the same time a consistent estimation of the model and reference climatologies. Since the climatology is calculated as a function of the forecast time, the anomalies vary depending on the forecast time they are associated with, as there is a sliding verification window. The anomaly correlation is then defined as:
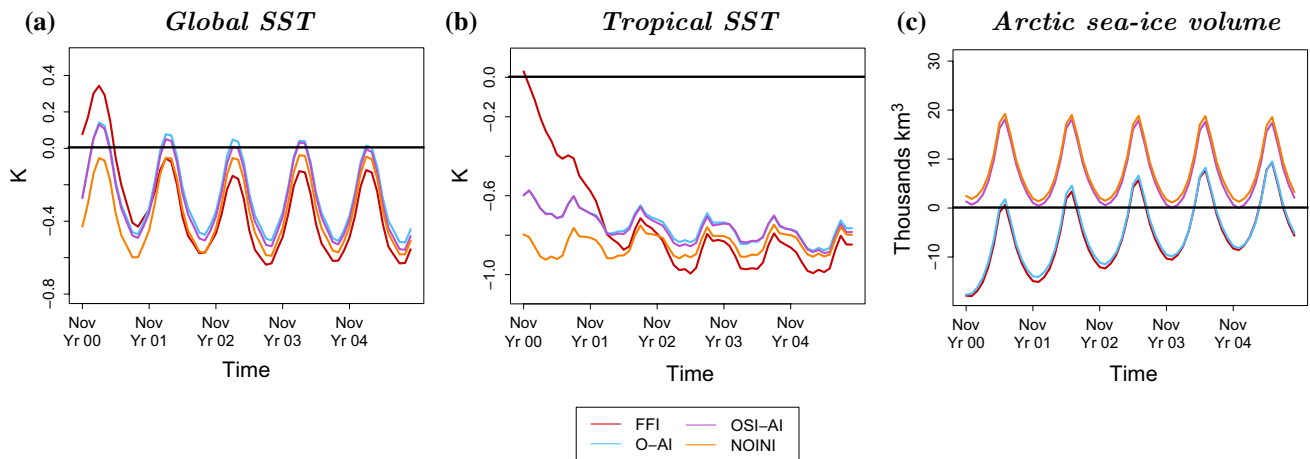
$$AC(f) = \frac{\sum_{d=1}^{n} [x_{d,f}]'[o_{d,f}]'}{\sqrt{\sum_{d=1}^{n} [x_{d,f}]'^2 \sum_{d=1}^{n} [o_{d,f}]'^2}} \qquad (1)$$

The root mean square error is given by:

$$RMSE(f) = \sqrt{\frac{\sum_{d=1}^{n} [[x_{d,f}]' - [o_{d,f}]']^2}{n}} \qquad (2)$$

In Eq. 1, $[x_{d,f}]'$ indicates the hindcast per-pair ensemble mean anomalies (for example ensemble mean anomalies of

## Bias and drift



**(a)** *Global SST*   **(b)** *Tropical SST*   **(c)** *Arctic sea-ice volume*

FFI ——   OSI–AI ——
O–AI ——   NOINI ——

**Fig. 1** Drift of **a** global mean SST averaged between 60 °S and 65 °N, **b** tropical SST averaged between 30 °S and 30 °N and **c** Arctic sea-ice volume respectively, as a function of the forecast time. FFI is represented in *red*, O-AI in *blue*, OSI-AI in *purple* and NOINI in *orange*.

See more details about these experiments in Sect. 2. The temperatures drift (**a**, **b**) are computed against ERSST data, while the sea-ice volume against the HistDfsNudg reconstruction which was also used to implement the AI technique

global mean temperature) as a function of the forecast time $f$ and the start date $d$, and $n$ is the number of start dates. Similarly $[o_{d,f}]'$ indicates the observed anomalies as a function of $f$ and $d$. Note that $'$ indicates the forecast or observed per-pair anomalies.

The confidence interval is calculated with a t-distribution for the AC, and with a $\chi^2$ distribution for the RMSE. The dependence between the hindcasts is accounted for in the computation of the confidence interval using Von Storch and Zwiers (2001) formula to estimate an effective number of independent data as a function of the autocorrelation (Trenberth 1984). The skill scores are computed after applying a one-year running mean on the forecast months in order to filter out seasonal climate variability and focus on interannual prediction skill.

## 3 Results

### 3.1 Forecast biases and drift

Figure 1 shows the bias of the experiments along the forecast time for global SST (60 °S–65 °N), tropical SST (30 °S –30 °N) and sea-ice volume in the Arctic. The evolution of the bias along the forecast time defines the drift. Regional maps of the bias for the different experiments are shown in figure S1 of the Supplementary Material. While for temperature the reference[3] data is ERSST (Smith et al. 2008), for

sea-ice volume the HistDfsNudg reconstruction was used, which was also used as a source of sea ice initial conditions in the AI technique. The model drifts until it reaches a stabilized state, i.e. stable biases, which we can see in NOINI. NOINI does not drift and stands as the reference that the other experiments will reach sooner or later depending on the variable and region. The FFI global SST climatology (red line in Fig. 1a) is in a warmer than observed state at the beginning of the forecast, and then it drifts towards a cooler than observed climate from the second forecast year. This is known as an overshoot (Hazeleger et al. 2013). The warm bias during half of the first forecast year is due to the warm bias in the Antarctic Circumpolar Current which is already settled during forecast months 2–4 (figure S1 of the Supplementary Material). The cold biases in the tropical regions and in the North Atlantic subpolar gyre take longer to develop (figure S1 of the Supplementary Material). The NOINI experiment (orange line) does not drift, it oscillates within its annual cycle. The two AI experiments (blue line O-AI, purple line OSI-AI) also have a warm overshoot followed by a drift toward a cold bias, but it is reduced with respect to FFI. The drift of the FFI experiment (red line in Fig. 1b) in the tropical SST is even more pronounced than the global SST drift, although no overshoot is seen in this case. The two AI experiments have similar behaviour with a drift which is smaller compared to FFI (blue line O-AI, purple line OSI-AI). The AI techniques allow for reducing the SST drift but this reduction is not entirely successful since some drift is still present. Figure 1c shows that regarding the sea-ice component the FFI (in red) and the O-AI (in blue) experiments have the same drift since they

---
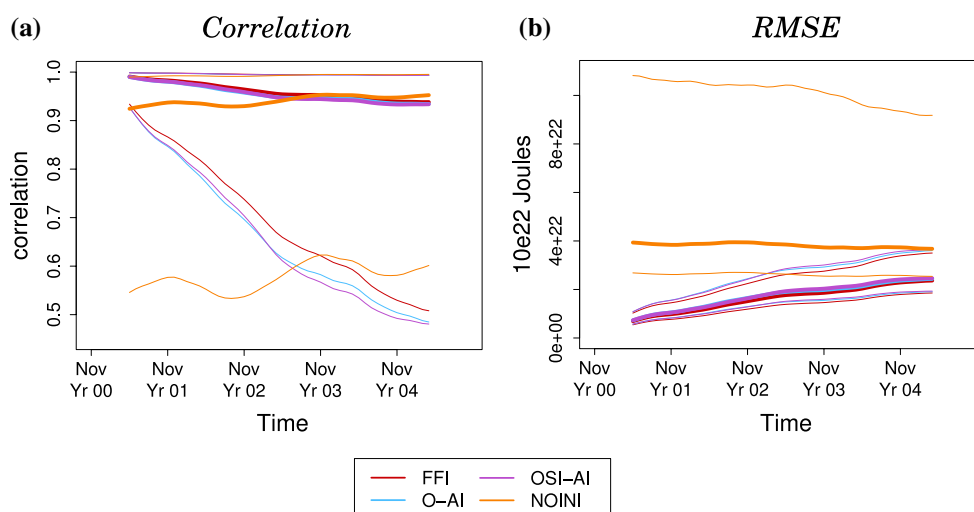
## Ocean heat content



**Fig. 2** Correlation and root mean square error for the total global ocean heat content with respect to NEMOVAR-ORAS4. *Red* for FFI, *blue* for O-AI, purple for OSI-AI and orange for NOINI. The *thin lines* represent the 95 % confidence interval obtained with a t-distri- bution for the correlation and a $\chi^2$ distribution for the RMSE. The dependence between the hindcasts is accounted for in the computa- tion of the confidence interval using Von Storch and Zwiers (2001) formula

are initialised with the same technique. The OSI-AI (in purple) seems very efficient in suppressing the forecast drift. When using PIOMAS (Zhang and Rothrock 2003) as a validation dataset the amplitude of the bias is larger (not shown).

The following sections will assess whether the reduction of the drift with the AI technique allows for a better prediction skill.

### 3.2 Predicting the thermohaline circulation

The PREDICATE project[4] aimed at assessing the level of predictability of decadal fluctuations in the Atlantic-European climate and concluded that there was a substantial influence of the ocean on the climate predictability at interannual to multidecadal timescales. Several studies illustrated the crucial role of the thermohaline circulation and the variations in Atlantic SST on the European climate (Gastineau and Frankignoul 2014; Msadek et al. 2011; Sutton and Hodson 2003; Mignot and Frankignoul 2004). Yeager et al. (2012) and Robson et al. (2012) linked changes in the North Atlantic subolar gyre heat content with changes in the AMOC. Predictability studies concluded that the Atlantic Meridional Overturning Circulation (AMOC) is sensitive to changes in the oceanic initial conditions (Msadek et al. 2013; Persechino et al. 2012; Collins et al. 2006) and the thermohaline circulation is potentially

predictable at least a decade in advance for the Atlantic. Other works as the one from Matei et al. (2012) also show skilful predictions of the sea surface temperature in the Atlantic that can be attributed to the initialisation of the AMOC.

We focus first on the global ocean heat content and the AMOC, the latter measured in terms of averaged northward transport of mass at the surface and southward transport in depth in the $40-55\,^{\circ}$N band of the Atlantic, vertically integrated in the top 1–2 km depth (Atlantic Meridional overturning stream function).

Figure 2a shows the anomaly correlation (calculated as in Eq. 1) of the global ocean heat content plotted along the forecast time. The thin lines in the plot represent the 95 % confidence interval obtained with a t-distribution. If, for example, the confidence intervals of the O-AI and NOINI experiments were not overlapping, the correlations of these two experiments would be significantly different. Moreover, if the confidence interval of the correlation of a specific experiment crosses the 0 line, then the correlation of this experiment is not significant. In general, in this work it is challenging to obtain significant results due to the low number of start dates available (23) associated with the short length of the verification period (from 1960 to 2004 i.e. when we have reliable observations for validation), and due to the small ensemble size. Figure 2a shows no obvious difference in correlation between the three initialised experiments (FFI -red-, O-AI -blue- and OSI-AI -purple-): the three of them perform better than NOINI (orange) for the first three forecast years. However, this difference is

---

[4] PREDICATE: http://www.ugamp.nerc.ac.uk/predicate/.

## 300m ocean heat content correlation for forecast year 1
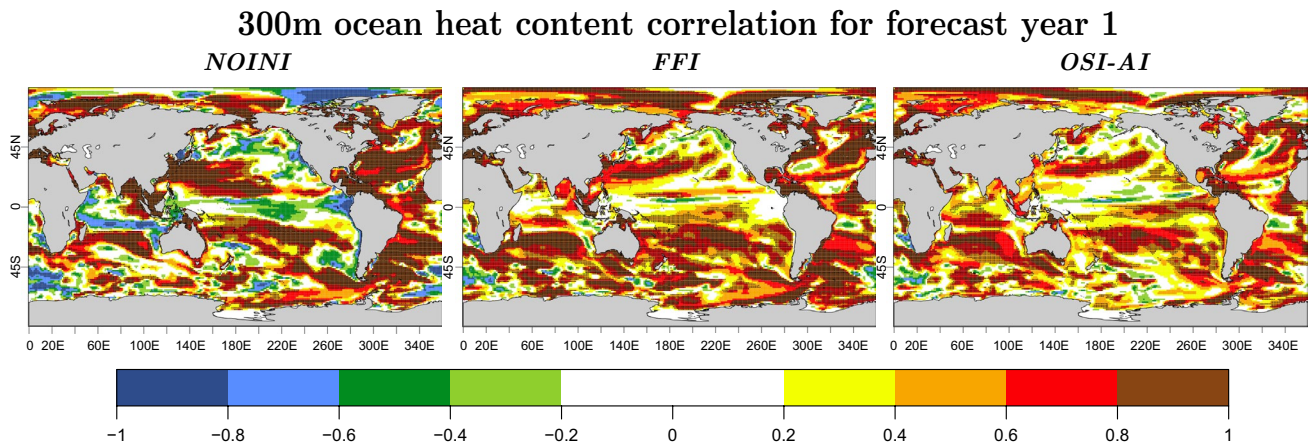


**Fig. 3** Correlation of the top 300 m depth ocean heat content, respectively for NOINI, FFI and OSI-AI experiments. The reference data are taken from NEMOVAR-ORAS4. *Dots* are used for the 95 % confidence significant correlations, where a two-sided test is applied

## AMOC
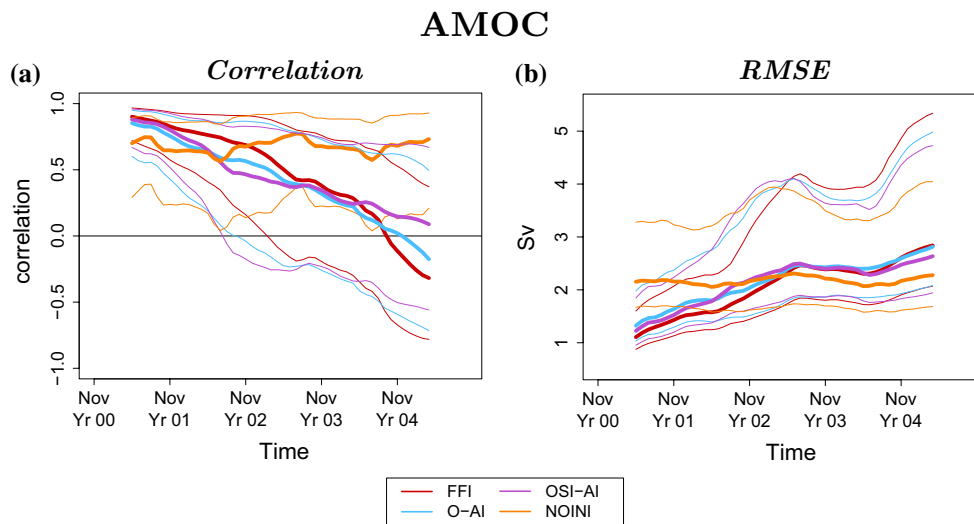


**Fig. 4** Correlation and root mean square error for the Atlantic meridional overturning stream function averaged in the $40 - 55\,°N$ band and 1–2 km depth with respect to NEMOVAR-ORAS4 as a function of the forecast time. *Red* for FFI, *blue* for O-AI, *purple* for OSI-AI and orange for NOINI. The *thin lines* represent the 95 % confidence interval obtained with a t-distribution for the correlation and a $\chi^2$ distribution for the RMSE. The dependence between the hindcasts is accounted for in the computation of the confidence interval using Von Storch and Zwiers (2001) formula

not significant for the correlation score. The reduction in RMSE with initialisation is significant for the first 2 forecast years (Fig. 2b). The skill of the three initialised experiments degrades with the forecast time toward the skill of NOINI which is nearly constant. Figure 3 shows the ocean heat content correlation maps averaged in the top 300 m depth respectively for NOINI, FFI and OSI-AI experiments. The initialised experiments improve the regions of negative skill shown in NOINI in the Arctic, Indian Ocean, part of the tropical and North Pacific. As it will be shown also for SST, the major difference between FFI and OSI-AI is seen in the tropical Pacific, where FFI performs better.

In the North Atlantic subpolar gyre there is a relatively low skill in the region corresponding to the Gulf stream, where NOINI performs better.

The skill of the initialised experiments (O-AI, OSI-AI and FFI) for the AMOC (Fig. 4a) drops below the NOINI skill after the first forecast year and correlations become non significantly different from 0 in the second forecast year. The RMSE and ACC show consistent skill results for the comparison between the NOINI and the 3 initialized experiments (Fig. 4b). Figure S2 of the Supplementary Material shows that the drift of AMOC is reduced in the AI experiments. Thus the initialised experiments show a
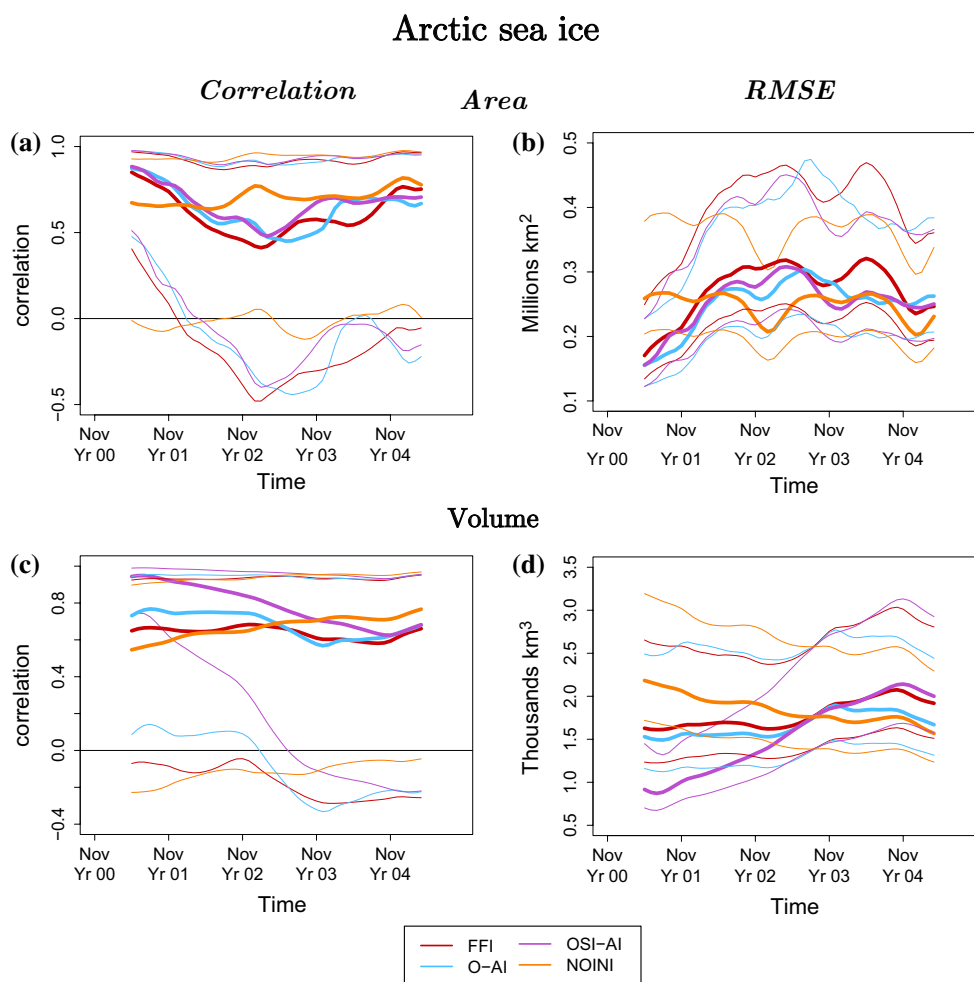
## Arctic sea ice



**Fig. 5** Correlation and RMSE of Arctic sea-ice area (**a–b**) and sea-ice volume (**c–d**). The reference data is the HistDfsNudg reconstruction. *Red* indicates FFI, *blue* O-AI, *purple* OSI-AI and *orange* NOINI. The *thin lines* represent the 95 % confidence interval as in the previous figures

similar loss of skill but a different magnitude of drift and therefore the latter is likely not to be the cause of such loss of skill.

### 3.3 Predicting the Arctic sea ice

Figure 5 illustrates the skill of the four experiments in predicting the Arctic sea-ice conditions. The variables shown are the area (first row in Fig. 5), and volume (second row in Fig. 5). When looking at these results one should keep in mind that the sea ice product used to initialise FFI is less coherent with the ocean state than the one used to initialise the AI experiment. The correlation of the sea-ice area (Fig. 5a) shows that the three initialised experiments improve their skill over NOINI during the first and a half forecast year. However, none of the differences are statistically significant, and the confidence intervals of the initialised experiments overlap the 0 line after the first and a half forecast year. There is a recovery of skill in Fig. 5a after the

second forecast year that could possibly come from the re-emergence of the SST signal that could have been stored in the ocean heat content (Chevallier and Salas-Mélia 2012) or the sea-ice thickness (Blanchard-Wrigglesworth et al. 2011). A similar behaviour was found in the study of Du et al. (2012). The OSI-AI clearly improves the skill in Arctic sea-ice volume (purple line Fig. 5c) with respect to the other experiments for approximately the first 3 forecast years. This improvement are partly explained by sea ice thickness and partly by sea ice extent (related to the SST shown in Sect. 3.5) improvements. The improvements in sea ice volume may be linked to the suppression of the drift shown in Fig. 1c. O-AI (blue line in Fig. 5c) and FFI (red line in Fig. 5c) also exhibit a slightly better skill than NOINI respectively during the first two and a half years and during the first forecast year.

A good prediction of the thickness/volume needs a good initialisation of the energy content in the sea ice system, for which a proper initialisation of the sea-ice component

AMO
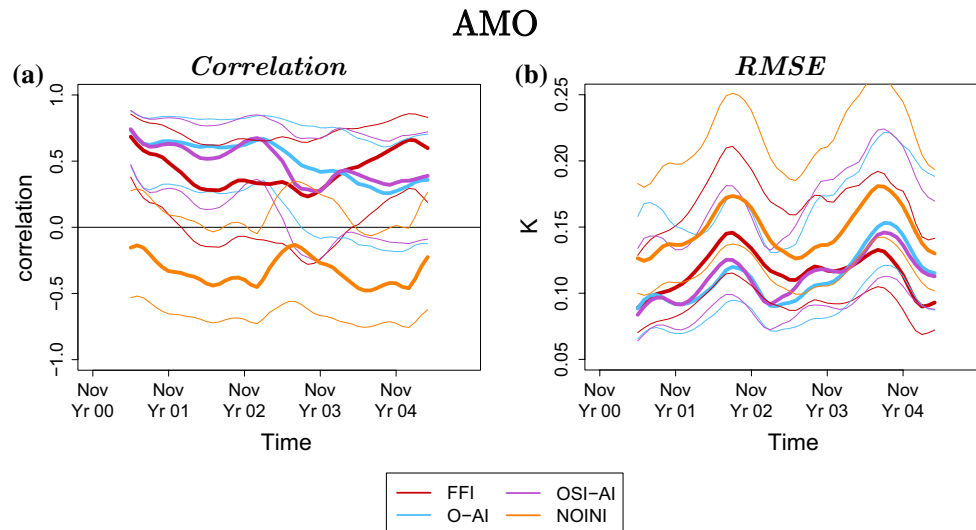


**Fig. 6** Skill of AMO, calculated as the difference between the mean SST anomalies in the North Atlantic and the global mean SST anomalies between 60 °S and 60 °N as a function of the forecast time. **a** cor-relation and **b** RMSE with respect to the ERSST data. *Red* for FFI, *blue* for O-AI, purple for OSI-AI and orange for NOINI. The *thin lines* represent the 95 % confidence interval as in the previous figures

and the improved consistency between the ocean and sea-ice variables given by the OSI-AI seems to play a key role. Analogous results are found in the RMSE (Fig. 5d): the OSI-AI experiment has the lowest error for the first 3 forecast years, followed by the O-AI, the FFI and the NOINI experiments. The results of the RMSE (Fig. 5b, d) are consistent with what is seen for the AC. The positive impact of the initialisation is more obvious in terms of RMSE than with AC, because the AC is a measure of skill that is not sensitive to errors in the linear trend along the start dates as a simple estimate of the response to climate change (Goddard et al. 2012). The results of the difference between the initialised experiments are qualitatively similar when using PIOMAS as validation data (not shown), although the differences between the experiments are reduced.

### 3.4 Impact on the main modes of climate variability

The Atlantic multidecadal oscillation (AMO, also called Atantic multidecadal variability -AMV-) is a pattern of multidecadal climate variability consisting in alternating phases of warm and cold sea surface temperature (SST) over the North Atlantic (Deser et al. 2010). It is thought to be the surface fingerprint of the thermohaline circulation (Kerr 2000; Knight et al. 2005). Its oscillation period ranges between 60 and 80 years (Schlesinger and Ramankutty 1994; Kushnir 1994). The AMO index is calculated as the difference between the mean SST anomalies in the North Atlantic and the global mean SST anomalies between 60 °S and 60 °N following the definition of Trenberth and Shea (2006). When looking at the AMO skill (Fig. 6a), the added value of the initialisation is clear. Consistently

to what shown in the multi-model ensemble study of Dob-las-Reyes et al. (2013a) NOINI is not able to predict the AMO (the correlation with the ERSST is negative but not-significant). The three initialised experiments have a posi-tive and significant skill, with a correlation with the ref-erence data of around 0.8 at the beginning of the forecast period. A similar impact of initialisation is also shown in Doblas-Reyes et al. (2013a). The correlation of the O-AI and OSI-AI experiments remains above 0.5 until the end of the third year, while the beneficial effect of initialisation lasts only for the first forecast year in the FFI experiment. The improvements in the O-AI and OSI-AI experiments compared to the NOINI one are significant for the first 2 forecast years (the confidence intervals do not overlap the NOINI one), while the improvements of FFI are significant only for the first forecast year. The positive impact of the initialisation is confirmed by the RMSE plot (Fig. 6b).

To further explore the performance of the different AMO predictions of the FFI and the OSI-AI experiments, Fig. 7 shows the AMO time series for the two experiments, both plotted against the ERSST data. The most interesting fea-ture in this comparison is the increasing trend of the last 20 years. Both experiments tend to underestimate the mini-mum in the first half of the 90s, especially with the hind-casts started in 1990 (light blue). However, the OSI-AI experiment captures properly the maximum in 1995 (dark pink Fig. 7b), which is overestimated and shifted in time in the FFI hindcast (Fig. 7a). Also the increasing trend hap-pening between the year 2002 and 2008 is not captured by FFI, while is partly captured by OSI-AI (start dates in red, orange and yellow). While the minimum in the 60 s is shifted in time, the maximum in 1965 is well captured by
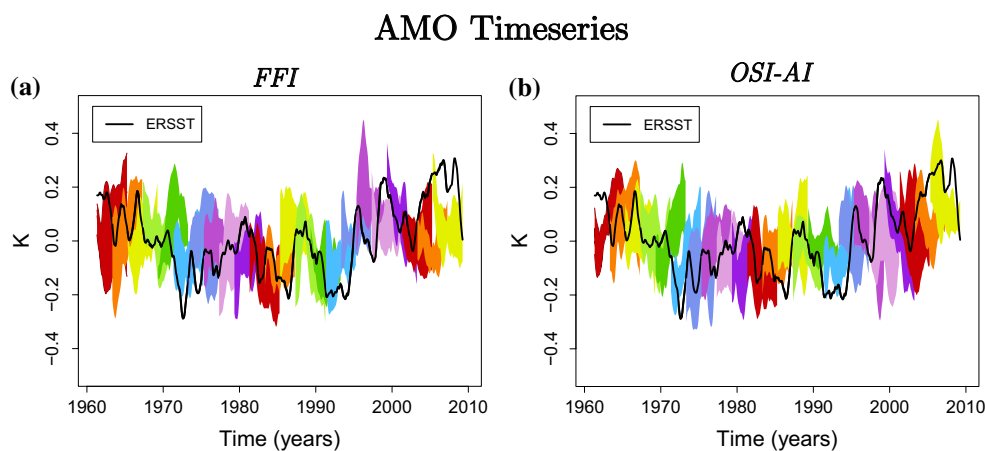
## AMO Timeseries



**Fig. 7** Comparison of AMO time series **a** FFI and **b** OSI-AI. In *black* is plotted the ERSST time series. The *colours* indicate the spread of the ensemble for the different start dates

## PDO
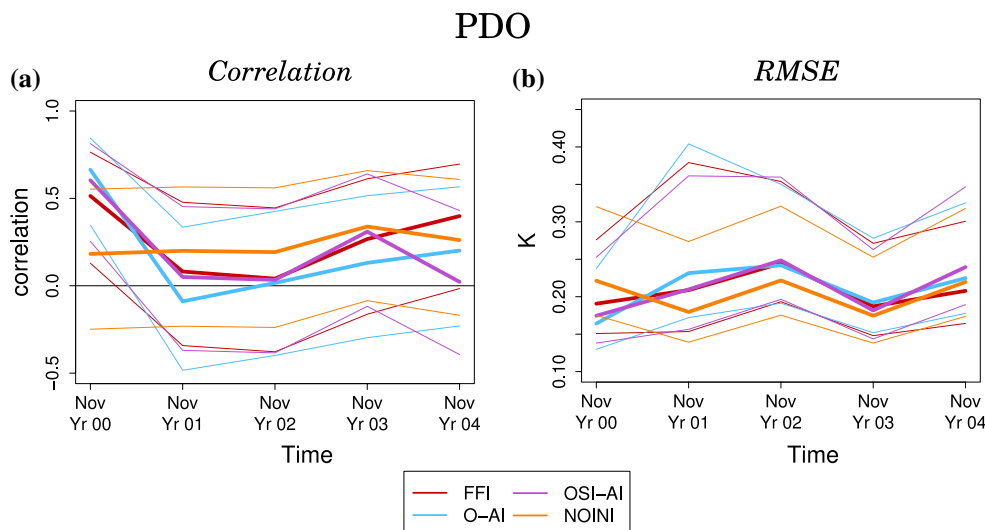


**Fig. 8** PDO calculated as the leading principal component of the Pacific annual SST variability in the domain $20 - 65\,°$N as a function of the forecast year. **a** correlation and **b** RMSE with respect to the ERSST data. *Red* for FFI, *blue* for O-AI, *purple* for OSI-AI and

*orange* for NOINI. The *thin lines* represent the 95 % confidence interval as in the previous figures. Note that annual mean values of SST were used in the calculation

both experiments (orange hindcasts). Finally in the 70s, the light and dark blue hindcasts fit better the observation with OSI-AI. However these results are limited by the small ensemble size (5 members).

We focus next on the most long-lived sea surface temperature mode in the Pacific, the Pacific Decadal Oscillation (PDO). The PDO is defined as the leading principal component of the Pacific annual SST variability calculated in the domain $20 - 65\,°$N. The PDO substantially impacts the North Pacific and North American climates, with phases persisting for 20–30 years (Anderson et al. 2009). It has also been linked to variations in surface air temperature, snowpack, precipitation and marine ecosystems (Mantua

et al. 1997). Previous studies have shown that the PDO interannual predictability is low (Newman 2007) and that the prediction skill is limited to about one year (Lienert and Doblas-Reyes 2013). The three initialisation methods tend to improve the skill of the PDO over NOINI, for the first forecast year (Fig. 8a). Such improvement is also shown in the RMSE plot (Fig. 8b) for the first year of forecast. None of the differences in forecast skill are significant though.

### 3.5 Temperature and precipitation

To have an overall view of the regional improvements and degradations of the OSI-AI experiment over the FFI and

## SST RMSE Ratio for forecast year 1
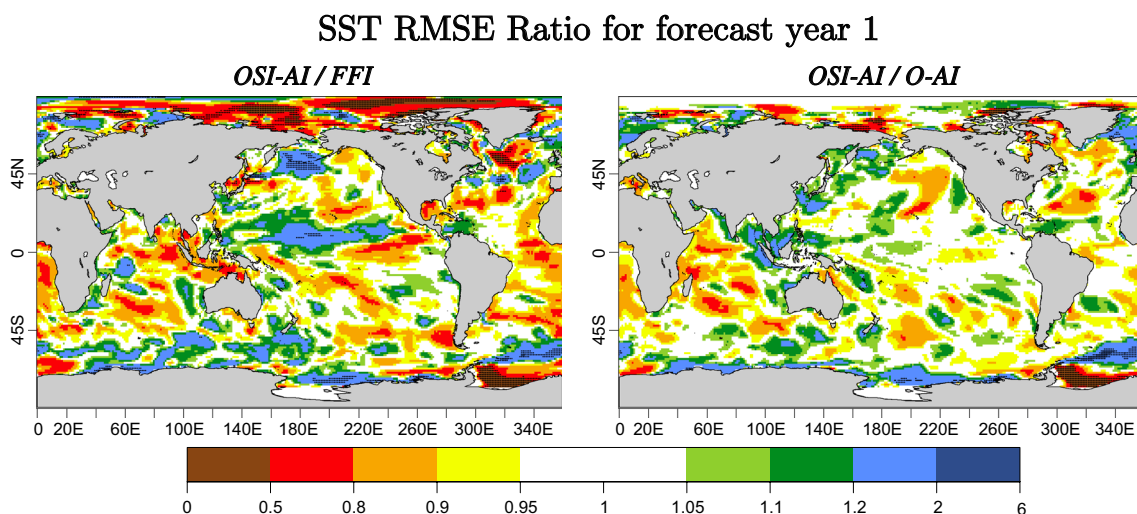


**Fig. 9** Ratio of sea surface temperature RMSE maps for the first forecast year, calculated against ERSST data: the first panel is the ratio between OSI-AI/FFI, second panel between OSI-AI/O-AI. When the ratio is smaller than 1 (*red*, *yellow* areas) the OSI-AI experiment has smaller RMSE, i.e. a better skill of the prediction. Vice versa, when the ratio is larger than 1 (regions in *blue*) the skill is degraded. The *black dots* over the *colours* indicate where the RMSE ratio is 95 % significantly higher or smaller than 1 according to a Fisher test. *Maps* show less than 5 % of the areas found as significant

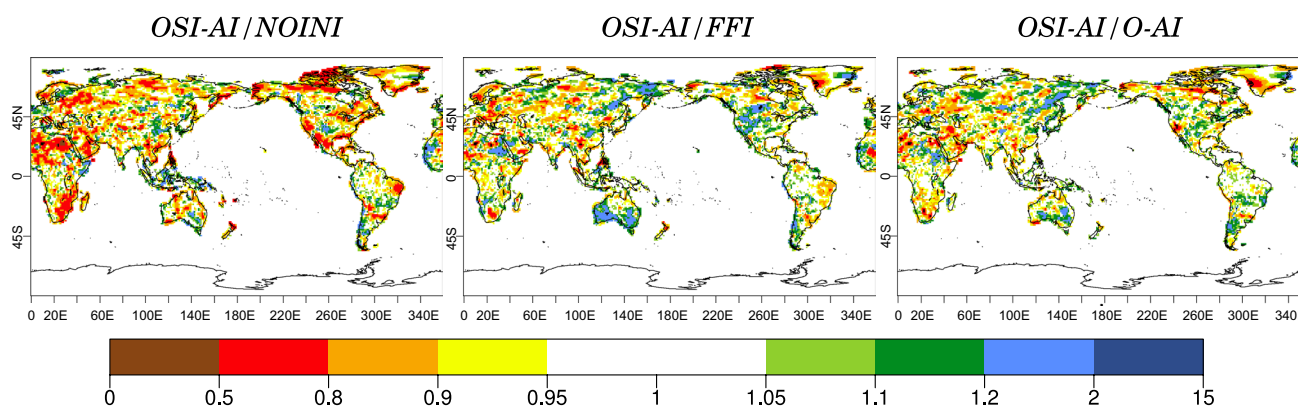## Near surface temperature RMSE ratio for forecast year 1
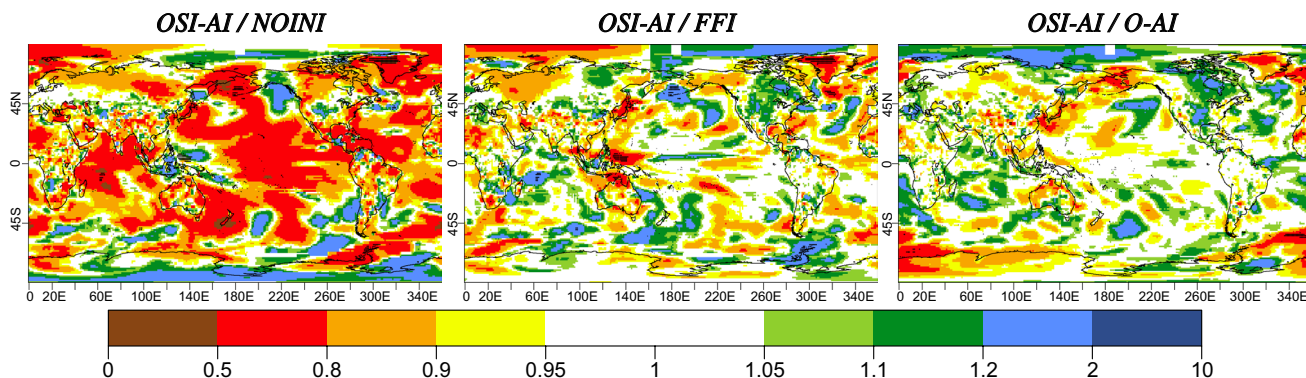


**Fig. 10** Near surface temperature RMSE ratio maps for the first forecast year, calculated against GHCN-ERSST-GISS merged data: *first column* OSI-AI/NOINI, *second column* OSI-AI/FFI, *third column* OSI-AI/O-AI

O-AI ones, Fig. 9 shows the ratio between the RMSE of SST during the first forecast year, respectively for the OSI-AI over the FFI experiment in the first panel, and the OSI-AI over the O-AI experiment in the second panel. When the RMSE ratio is smaller than 1 the OSI-AI has a better SST skill over the other experiments, vice versa it degrades the skill when the RMSE ratio is larger than 1. The SST skill is improved by OSI-AI over FFI (first panel Fig. 9) at the sea-ice in the Arctic region, the Labrador Sea and the Weddell Sea. To better illustrate the skill in the Arctic region figure S4 of the Supplementary Material shows the polar stereographic version of Fig. 9. However the skill is degraded

in the Bering Sea, part of the North and the Tropical Pacific and the Ross Sea in the Antarctic. The regions of statistically significant improvements represent less than the 5 % of the total grid boxes. The second panel of Fig. 9 shows that OSI-AI improves the SST prediction over the O-AI experiment in some regions of the Arctic and in the Weddell Sea. Some degradations in skill occur in some regions of the Antarctic (Fig. 9 second panel) and the North Atlantic.

The most societally relevant climate variables over land are the near surface temperature and precipitation. Figure 10 shows the RMSE ratio for near surface temperature between OSI-AI and NOINI (first column), OSI-AI and FFI (second

## Precipitation RMSE ratio for forecast year 1



**Fig. 11** Land precipitation RMSE ratio maps for the first forecast year, calculated against the GPCC data (Rudolf et al. 2010): **a** OSI-AI/NOINI, **b** OSI-AI/FFI and **c** OSI-AI/O-AI

column), OSI-AI and O-AI (third column), for the first forecast year, calculated against a combination of land surface temperatures from the National Centres for Environmental Prediction (NCEP) GHCN/CAMS dataset (Fan and Dool 2008), SST from the National Climatic Data Center (NCDC) ERSST v3b dataset (Smith et al. 2008) and, north of 60 °N, the GISTEMP dataset with 1200 km decorrelation scale (Hansen et al. 2010). The first panel of Fig. 10 shows a gain in skill of the OSI-AI experiment over NOINI in the Pacific, the Central America and Greenland. The corresponding maps of correlation difference show that such improvements are statistically significant in terms of correlation (Figure S5 of the Supplementary Material). The differences with the other initialised experiments (second and third panels of Fig. 10 ) are not statistically significant.

The skill maps of precipitation (Fig. 11) are very patchy and not significant improvements are shown. The AI seems not to be efficient in the ENSO region and the correspondent teleconnections; this could be caused by a possible spatial shift of the observed anomalies with respect to the position where the model would simulate the corresponding anomalies. Bellucci et al. (2014) found a similar loss of skill of the AI multi-model ensemble in the tropical Pacific. They suggested the inconsistency between the observed anomalies with the underlying background state of the model as a possible cause. Similarly in the comparison between OSI-AI and O-AI (Fig. 11 third panel) there are some losses of skill over North America, Australia and Asia. No significant improvements are detected.

## 4 Summary and conclusions

In this work, two anomaly initialisation (AI) techniques have been implemented and compared with the full field initialisation (FFI) method and a reference simulation for

which no effort has been put into initialising the model variability with the phase of the observed variability referred to as an 'uninitialized experiment' in what follows. One anomaly initialised experiment is performed initialising the ocean with AI of all the oceanic prognostic variables. The other consists in an ocean and sea ice AI of all prognostic variables with consistency corrections between the ocean and sea ice components. The full field initialised experiment has the three model components initialised with the observed full fields. Such comparison aimed at investigating whether removing the drift through AI and optimizing the consistency between the ocean and the sea-ice state has a positive impact on the prediction skill.

Previous studies such as Bellucci et al. (2014), Smith et al. (2013), Hazeleger et al. (2013) have found small differences in the comparison between AI and FFI. Although differences in skill between our various experiments are rather weak, they are of the same order of magnitude as the progress in seasonal forecast skill over the last 10 years (Doblas-Reyes et al. 2013b; Vitart 2014). In general, in this work it is challenging to obtain statistically significant results due to the low number of start dates available and the small ensemble size. In the full field initialised experiment, the climate state drifts toward the state of the uninitialised experiment, reaching the equilibrium with different timescales depending on the location and variable. The AI technique allows for a strong reduction of such drift. All the variables considered in this study have shown improvements in forecast skill in the initialised hindcasts compared to the uninitialised ones. In particular the ocean heat content shows a better skill up to 3 years ahead and the Atlantic Meridional Overturning stream function up to 2 years ahead, but the difference between the various initialisation techniques are minor. Initialising the ocean and sea ice components in a consistent way through the AI technique

improves the skill of Arctic sea-ice volume from 45 to 10 % (respectively at the beginning and at the end of the forecast) with respect to the full field initialised predictions. The improvement with respect to the ocean only AI experiment varies between 30 to 8 %, and the improvements over the uninitialised experiment varies between 72 % to a degradation of the 10 % at the end of the forecast. The ocean and sea ice AI implies a better initialisation of the sea-ice volume that also translate into near surface temperature skill in the Arctic. All the initialised experiments have a better skill in predicting the Arctic sea-ice area over the uninitialised experiment, but the skill difference between the initialised experiments are small.

The skill in the Atlantic Multidecadal Oscillation (AMO) using the AI technique is significantly better than in the uninitialised experiment along the first 2 forecast years, and it is better also than the full field hindcast skill, although the difference in skill is not significant. Although the improvements are not statistically significant, the skill of the anomaly initialised experiments in hindcasting the Pacific Decadal Oscillation (PDO) index seems to be better when using the AI than the FFI technique during the first forecast year, the latter being already better than the uninitialised experiment for the first forecast year. In a nutshell, removing the drift through AI and improving the consistency between the ocean and the sea-ice state increase the skill compared to the full field or the ocean only anomaly initialised experiments. On the other hand, an important area where the skill is degraded by AI is the tropical Pacific, supporting the findings of previous studies (Bellucci et al. 2014; Smith et al. 2013). This degradation also affects the skill of precipitation in some of the ENSO teleconnection regions (North America, Western Asia and Australia). Such result might be related to the limitation of the AI hypothesis mentioned in the Introduction: the model bias in the tropical Pacific could include a spatial shift that would cause the initial observed anomalies no to be placed where the model would simulate the corresponding anomalies. Some additional work has explored the possibility of refining further the AI technique to aim at a larger skill. The ideas that has been implemented aim, on one hand, at providing the most suitable initialisation for the density variable which plays a crucial role in the ocean circulation (Volpi et al. 2016). On the other hand, the observed anomaly has been scaled in order to take into account the different amplitudes of the observed versus the model variability. Further investigation should focus on overcoming the issue of the spatial shift and implementing the AI in the ocean, sea ice and atmosphere components to suppress the remaining drift and optimize the consistency between the different model components.

## References

Anderson DLT, Doblas-Reyes FJ, Balmaseda M, Weisheimer A (2009) Decadal variability: processes, predictability and prediction. ECMWF Technical Memorandum 591, http://www.ecmwf.int/publications/library/do/references/show?id=89132

Balmaseda MA, Mogensen KS, Weaver AT (2012) Evaluation of the ecmwf ocean reanalysis oras4. Quart J Roy Meteor Soc. doi:10.1002/qj.2063

Bellucci A, Haarsma R, Gualdi S, Athanasiadis PJ, Caian M, Cassou C, Fernandez E, Germe A, Jungclaus J, Kroger J, Matei D, Muller W, Pohlmann H, Salas-Mélia D, Sanchez E, Smith D, Terray L, Wyser K, Yang S (2014) An assessment of a multi-model ensemble of decadal climate predictions. Clim Dyn 44:2787–2806. doi:10.1007/s00382-014-2164-y

Blanchard-Wrigglesworth E, Bitz CM, Holland MM (2011) Influence of initial conditions and climate forcing on predicting arctic sea ice. Geophys Res Lett 38(L18503): doi:10.1029/2011GL048807

Booth J, Wang S, Polvani L (2013) Midlatitude storms in a moister world: lessons from idealized baroclinic life cycle experiments. Clim Dyn 41:787–802. doi:10.1007/s00382-012-1472-3

Brodeau L, Barnier B, Treguier A, Penduff T, Gulev S (2009) An era40-based atmospheric forcing for global ocean circulation models. Ocean Model 31:88–104. doi:10.1016/j.ocemod.2009.10.005

Chevallier M, Salas-Mélia D (2012) The role of sea ice thickness distribution in the arctic sea ice potential predictability: a diagnostic approach with a coupled gcm. J Clim 25:3025–3038

Collins M, Booth BBB, Harris GR, Murphy JM, Sexton DMH, Webb M (2006) Towards quantifying uncertainty in transient climate change. Clim Dyn 365:1957–1970

Dee DP, Uppala SM, Simmons AJ, Berrisford P, PP, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Holm EV, Isaksen L, Kållbergc P, Kähler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thpaut JN, Vitart F, (2011) The era-interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteor Soc 137:553–597

Deser C, Alexander MA, Xie SP, Phillips AS (2010) Sea surface temperature variability: patterns and mechanisms. Ann Rev Mar Sci 2:115–143. doi:10.1146/annurev-marine-120408-151453

Doblas-Reyes F, Andreu-Burillo I, Chikamoto Y, García-Serrano J, Guèmas V, Kimoto M, Mochizuki T, Rodrigues L, van Oldenborgh G (2013a) Initialized near-term regional climate change prediction. Nat Commun 4(1715): doi:10.1038/ncomms2704

Doblas-Reyes FJ, Weisheimer A, Palmer TN, Murphy JM, Smith D (2010) Forecast quality assessment of the ensembles seasonal-to-decadal stream 2 hindcasts. ECMWF Technical Memorandum 621, http://www.ecmwf.int/publications/library/do/references/show?id=89771

Doblas-Reyes FJ, García-Serrano J, Lienert F, Biescas AP, Rodrigues LRL (2013b) Seasonal climate predictability and forecasting: status and prospects. WIREs Clim Change 4:245–268. doi:10.1002/wcc.217

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

Ethe C, Aumont O, Foujols MA, Levy M (2006) Nemo reference manual, tracer component: Nemo-top. preliminary version. Note du Pole de modlisation, Institut Pierre-Simon Laplace (IPSL) France 28:1288–1619

Fan Y, van den Dool H (2008) A global monthly land surface air temperature analysis for 1948-present. Geophys Res Lett 113(D01103): doi:10.1029/2007JD008470

Fichefet T, Maqueda MAM (1997) Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics. J Geophys Res 102:12609–12646

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

Gastineau G, Frankignoul C (2014) Influence of the north atlantic sst variability on the atmospheric circulation during the twentieth century. J Clim 28:1396–1416. doi:10.1175/JCLI-D-14-00424.1

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 (2012) A verification framework for interannual-to-decadal predictions experiments. Clim Dyn. doi:10.1007/s00382-012-1481-2

Goosse H, Fichefet T (1999) Importance of ice-ocean interactions for the global ocean circulation: a model study. J Geophys Res 104:13337–23355

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mithell JFB, Wood RA (2000) The simulation of sst, sea ice extents and ocean heat and ocean heat transports in a version of the hadley centre coupled model without flux adjustments. Clim Dyn 16:147–168

Guemas V, Doblas-Reyes F, Mogensen K, Keely S, Tang Y (2014) Ensemble of sea ice initial conditions for interannual climate predictions. Clim Dyn 43:2813–2829. doi:10.1007/s00382-014-2095-7

Hansen J, Ruedy R, Sato M, Lo K (2010) Global surface temperature change. Rev Geophys 48(RG4004): doi:10.1029/2010RG000345

Hazeleger W, Severijns C, Semmler T, Stefânescu S, Yang S, Wang X, Wyser K, Dutra E, Baldasano JM, Bintanja R, Bougeault P, Caballero R, Ekman AML, Christensen JH, van den Hurk B, Jimenez P, Jones C, Kållberg P, Koenigk T, McGrath R, Miranda P, van Noije T, Palmer T, Parodi JA, Schmith T, Selten F, Storelvmo T, Sterl A, Tapamo H, Vancoppenolle M, Viterbo P, Willân U (2010) Ec-earth: a seamless earth-system prediction approach in action. Bull Am Meteorol Soc 91(10):1357–1363. doi:10.1175/2010BAMS2877.1

Hazeleger W, Guemas V, Wouters B, Corti S, Andreu-Burillo I, Doblas-Reyes FJ, Wyser K, Caian M (2013) Multiyear climate predictions using two initialization strategies. Geophys Res Lett 40(9):1794–1798. doi:10.1002/grl.50355

Kerr RA (2000) A north atlantic climate pacemaker for the centuries. Science 288(5473):1984–1985. doi:10.1126/science.288.5473.1984

Knight JR, Allan RJ, Folland CK, Vellinga M, Mann ME (2005) A signature of persistent natural thermohaline circulation cycles in observed climate. Geophys Res Lett 32(20): doi:10.1029/2005GL024233

Kushnir Y (1994) Interdecadal variations in the north atlantic sea surface temperature and associated atmospheric conditions. J Clim 7:141–157

Lienert F, Doblas-Reyes F (2013) Decadal prediction of interannual tropical and north pacific sea surface temperature. J Geophys Res 118:5913–5922. doi:10.1002/jgrd.50469

Madec G (2008) Nemo ocean engine. Note du Pole de modlisation, Institut Pierre-Simon Laplace (IPSL) France 27:12881619

Magnusson L, Leutbecher M, Kallen E (2008) Comparison between singular vectors and breeding vectors as initial perturbations for the ecmwf ensemble prediction system. Mon Wea Rev 134:4092–4104

Magnusson L, Alonso-Balmaseda M, Corti S, Molteni F, Stockdale T (2012) Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors. ECMWF Technical Memorandum 676. http://www.ecmwf.int/publications/library/do/references/show?id=90506

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

Matei D, Pohlmann H, Jungclaus J, Müller 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. doi:10.1175/JCLI-D-11-00633.1

Mignot J, Frankignoul C (2004) Interannual to interdecadal variability of sea surface salinity in the atlantic and its link to the atmosphere in a coupled model. J Geophys Res 109(C04005): doi:10.1029/2003JC002005

Mogensen KS, Balmaseda MA, Weaver A (2012) The nemovar ocean data assimilation as implemented in the ecmwf ocean analysis for system4. ECMWF Technical Memorandum 657 (in preparation)

Msadek R, Frankignoul C, Li LZX (2011) Mechanisms of the atmospheric response to north atlantic multidecadal variability: a model study. Clim Dyn 36:1255–1276. doi:10.1007/s00382-010-0958-0

Msadek R, Johns W, Yeager S, Danabasoglu G, Delworth T, Rosati A (2013) The atlantic meridional heat transport at 26.5 n and its relationship with the moc in the rapid array and the gfdl and ncar coupled models. J Clim 26:4335–4356

Newman M (2007) Interannual to decadal predictability of tropical and north pacific sea surface temperatures. J Clim 20:2333–2356. doi:10.1175/JCLI4165.1

Old C, Haines K (2006) North atlantic subtropical mode waters and ocean memory in hadcm3. J Clim 19:1126–1148

Otterà OH, Bentsen M, Drange H, Suo L (2010) External forcing as a metronome for atlantic multidecadal variability. Nat Geosci 3:688–694

Persechino A, Mignot J, Swingedouw D, Labetoulle S, Guilyardi E (2012) Decadal predictability of the atlantic meridional overturning circulation and climate in the ipsl-cm5a-lr model. Clim Dyn 40(9):2359–2380. doi:10.1007/s00382-012-1466-1

Pohlmann H, Jungclaus J, Köhl A, Stammer D, Marotzke J (2009) Initializing decadal climate predictions with the gecco oceanic synthesis: Effects on the north atlantic. J Clim 22:3926–3938

Robson J, Sutton R, Lohmann K, Smith D, Palmer MD (2012) Causes of the rapid warming of the north atlantic ocean in the mid-1990s. J Clim 25: doi:10.1175/JCLI-D-11-00443.1

Rudolf B, Becker A, Schneider U, Meyer-Christoffer A, Ziese M (2010) The new gpcc full data reanalysis version 5 providing high-quality gridded monthly precipitation data for the global land-surface is public available since december 2010. GPCC Status Report. http://www.dwd.de/bvbw/generator/DWDWWW/Content/Oeffentlichkeit/KU/KU4/KU42/en/Reports__Publications/GPCC__status__report__2010,templateId=raw,property=publicationFile.pdf/GPCC_status_report_2010.pdf

Sanchez-Gomez E, Cassou C, Ruprich-Robert Y, Fernandez E, Terray L (2015) Drift dynamics in a coupled model initialized for

decadal forecasts. Clim Dyn 46(5):1819–1840. doi:10.1007/s00382-015-2678-y

Schlesinger ME, Ramankutty N (1994) An oscillation in the gloabl climate system of period 65–70 years. Nature 367:723–726. doi:10.1038/367723a0

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

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

Smith T, Reynolds R, Peterson T, Lawrimore J (2008) Improvements to noaa's historical merged land-ocean surface temperature analysis (1880–2006). J Clim 21:2283–2296

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

Swingedouw D, Ortega P, Mignot J, Guilyardi E, Masson-Delmotte V, Butler PG, Khodri M, Séférian R (2015) Bidecadal north atlantic ocean circulation variability controlled by timing of volcanic eruptions. Nature Communications 6(6545): doi:10.1038/ncomms7545

Trenberth KE (1984) Some effects of finite sample size and persistence on meteorological statistics. part i: Autocorrelations. Mon Wea Rev 112:2359–2368. doi:10.1175/1520-0493(1984)112,2359:SEOFSS.2.0.CO;2

Trenberth KE (2008) Observational needs for climate prediction and adaptation. WMO Bull 57:17–21

Trenberth KE, Shea DJ (2006) Atlantic hurricanes and natural variability in 2005. Geophys Res Lett 33(L12704): doi:10.1029/2006GL026894

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, Bechtold VDC, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg LVD, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hlm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, Mcnally AP, Mahfouf J, Morcrette J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The era-40 reanalysis. Q J R Meteor Soc 131:2961–3012

Valcke S (2006) Oasis3 user guide. PRISM Support Initiative Report 3:64

Van Oldenborgh GJ, Doblas-Reyes FJ, Wouters B, Hazeleger W (2011) Decadal prediction skill in a multi-model ensemble. Clim Dyn 38:1263–1280. doi:10.1007/s00382-012-1313-4

Vitart F (2014) Evolution of ecmwf sub-seasonal forecast skill scores. Q J R Meteorol Soc 140(683):1889–1899. doi:10.1002/qj.2256

Volpi D, Guemas V, Doblas-Reyes FJ, Hawkins E, Nichols N (2016) Decadal climate prediction with a refined anomaly initialisation approach. Clim Dyn. doi:10.1007/s00382-016-3176-6

Von Storch H, Zwiers F (2001) Statistical analysis in climate research. Cambridge University Press, Cambridge

Yeager S, Karspeck A, Danabasoglu G, Tribbia J, Teng H (2012) A decadal prediction case study: Late twentieth-century north atlantic ocean heat content. J Clim 25: doi:10.1175/JCLI-D-11-00595.1

Zhang J, Rothrock DA (2003) Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates. Mon Weather Rev 131:845–861