Trends, variability and predictive skill of the ocean heat content in North Atlantic: an analysis with the EC-Earth3 model

Teresa Carmo-Costa[1] · **Roberto Bilbao**[2] · **Pablo Ortega**[2] · **Ana Teles-Machado**[3] · **Emanuel Dutra**[3]

Abstract
This study investigates linear trends, variability and predictive skill of the upper ocean heat content (OHC) in the North Atlantic basin. This is a region where strong decadal variability superimposes the externally forced trends, introducing important differences in the local warming rates and leading in the case of the Central Subpolar North Atlantic to an overall long-term cooling. Our analysis aims to better understand these regional differences, by investigating how internal and forced variability contribute to local trends, exploring also their role on the local prediction skill. The analysis combines the study of three ocean reanalyses to document the uncertainties related to observations with two sets of CMIP6 experiments performed with the global coupled climate model EC-Earth3: a historical ensemble to characterise the forced signals, and a retrospective decadal prediction system to additionally characterise the contributions from internal climate variability. Our results show that internal variability is essential to understand the spatial pattern of North Atlantic OHC trends, contributing decisively to the local trends and providing high levels of predictive skill in the Eastern Subpolar North Atlantic and the Irminger and Iceland Seas, and to a lesser extent in the Labrador Sea. Skill and trends in other areas like the Subtropical North Atlantic, or the Gulf Stream Extension are mostly externally forced. Large observational and modeling uncertainties affect the trends and interannual variability in the Central Subpolar North Atlantic, the only region exhibiting a cooling during the study period, uncertainties that might explain the very poor local predictive skill.

Keywords Ocean heat content · Long-term trends · North Atlantic Ocean · Decadal prediction · Climate modelling

1 Introduction

Observations show that approximately 93% of the energy entering the climate system since the 1950s has been stored by the oceans (Levitus et al. 2012; Lyman et al. 2010; Johnson and Lyman 2020; Schuckmann et al. 2020). As a result, the upper layers of the ocean have warmed with distinct spatio-temporal features, shaped by the different warming rates that the ocean basins of the world have experienced. Several studies (e.g. Chen and Tung 2014; Wang et al. 2018; Zanna et al. 2019) have shown that the ocean heat content

(OHC) in the Indian basin has remained approximately constant from 1950 until the early 2000s, when a warming trend started to develop, while the Pacific and Atlantic ocean basins exhibited notable multi-decadal variability during the whole period, with a superimposed long-term warming trend. Regional differences have also emerged in some of the basins, like in the Atlantic ocean, where in recent decades the central Subpolar North Atlantic (CSPNA) region has cooled while the rest of the basin has warmed, a feature that has been labeled in recent literature as the "cold blob" when referring to the most recent decadal trends (e.g. Yeager et al. 2016), and as the "warming hole" when referring to externally forced centennial-scale changes (e.g. Drijfhout et al. 2012; Rahmstorf et al. 2015; Keil et al. 2020). This paper will investigate the origin of these local trends from a climate prediction perspective, delving on other aspects like their linearity.

While, at the global scale, the observed long-term linear warming trend can be largely explained by the increasing greenhouse gas concentrations in the atmosphere and

✉ Teresa Carmo-Costa
tmcosta@fc.ul.pt

1 Instituto Dom Luiz-Faculdade de Ciências da Universidade de Lisboa, Edifício C8, Campo Grande, 1749-016 Lisbon, Portugal

2 Barcelona Supercomputing Center, Barcelona, Spain

3 Instituto Dom Luiz-Instituto Português do Mar e da Atmosfera, Lisbon, Portugal

the counterbalancing effect of anthropogenic aerosols (e.g. Bilbao et al. 2019), at the regional scale internal climate variability can strongly contribute to the trends, and other external factors can induce non-linearities in the forced signals (Borchert et al. 2021). On decadal time-scales, ocean dynamics play an important role in shaping these regional trends (Chen and Tung 2014). Several mechanisms have been proposed to explain the recent cooling trend in the CSPNA, including the influence of anthropogenic factors (Chemke et al. 2020). The oceanic processes proposed to explain the cooling include (i) a slowdown in the meridional heat transport, triggered by a weakening of the Atlantic Meridional Overturning Circulation (AMOC; Drijfhout et al. 2012; Robson et al. 2016), (ii) a change in heat advection by the horizontal gyre circulation (Piecuch et al. 2017), and (iii) a shift in the Gulf Stream current (Ruiz-Barradas et al. 2018). More generally, changes in the ocean circulation, and in particular the AMOC, have been related in many modeling studies to intrinsic basin-scale decadal variability in the North Atlantic ocean (Knight et al. 2005; Buckley and Marshall 2016; Ortega et al. 2015; Gastineau et al. 2018), usually referred to as Atlantic Multidecadal Variability (AMV; Schlesinger and Ramankutty 1994). This implies that the recent OHC trends in the basin could be internally-driven. However, there is also modeling (Cheng et al. 2013) and indirect paleoclimatic evidence (i.e. proxies; Rahmstorf et al. 2015; Thornalley et al. 2018) supporting that the AMOC might have been weakening since the beginning of the industrial era following the increase in greenhouse gas (GHG) concentrations, which would support an externally-forced origin of the OHC cooling trends.

Both externally forced and internally-generated climate variability can be jointly exploited in climate models to produce skillful near-term predictions of the upper ocean temperatures (Meehl et al. 2009). In these predictions, typically know as decadal predictions because they extend for up to 10 years, radiative forcing changes are prescribed as boundary conditions and internal climate variability is initialised by bringing the model close to the observed state. The predictive skill is not geographically nor temporally uniform, as some regions are naturally more predictable than others, either because they are more sensitive to the external radiative forcings or because they are under the influence of more predictable modes of internal climate variability. Two prominent examples of predictable regions are the Tropical Pacific, where El Niño–Southern Oscillation (ENSO) dominates the seasonal to interannual ocean variability and prediction skill (e.g. Barnston 1994; Ham et al. 2019), and the North Atlantic, where the aforementioned AMV—thought to be at least partially linked to slow changes in both gyre and ocean circulations—provides high levels of decadal prediction skill for the OHC (e.g. Doblas-Reyes et al. 2013; Yang et al. 2013;

Yeager and Robson 2017a; Frajka-Williams et al. 2017). In the particular case of the North Atlantic, not all areas are equally predictable. Most decadal prediction systems show high levels of skill for the upper ocean temperatures at Subpolar Gyre latitudes (Matei et al. 2012; Robson et al. 2018; Mignot et al. 2016; Yeager et al. 2012, 2018; Bilbao et al. 2020). On these regions slowly-varying ocean dynamics exerts a prominent influence, but observations also suggest long predictive capacity due to long thermal memory, and therefore persistence (Buckley et al. 2019). Another North Atlantic region also linked to the large-scale ocean circulation is the Gulf Stream Extension (GSE), but it shows comparatively less persistence for the upper ocean temperatures than the Subpolar Gyre (Buckley et al. 2019). Furthermore, the GSE exhibits lower skill in many prediction systems (e.g. Matei et al. 2012; Mignot et al. 2016; Yeager et al. 2018; Bilbao et al. 2020), which suggests that the region might be inherently less predictable. However, the Gulf Stream region is also known to suffer from important systematic biases in the typical non-eddying model resolutions used for climate prediction (i.e., 1° × 1°), which might hamper predictive skill (Hewitt et al. 2017). It is also unclear if higher prediction skill is achievable in other regions like the Subtropical North Atlantic, in which decadal variability from internal ocean dynamics is less prominent but the external forcings might exert a stronger influence (Frankignoul et al. 2017). We also note that some regional decadal variations may be forced by changes in natural radiative forcings, like major volcanic eruptions (Hermanson et al. 2020; Mann et al. 2021; Borchert et al. 2021), which might additionally trigger a delayed dynamical ocean response (Swingedouw et al. 2015).

This study focuses on the upper OHC variations in the North Atlantic basin in recent decades, addressing some of the factors behind the different local OHC variability and predictability. We also explore the observational uncertainties, since many regions of the North Atlantic Ocean are still unmonitored or have remained undersampled until recently (Abraham et al. 2013). The following questions will be addressed in the study:

- Are past upper OHC changes consistent in ocean reanalyses, both at the basin-scale and regionally? If not, which are the regions where the recent OHC changes are more uncertain?
- Can climate models skilfully predict the robust upper OHC changes from reanalyses? And how much of the skill derives from the predictability of long-term linear trends?
- To what extent it is realistic to interpret linear trends as externally forced signals? Are there regions in which this assumption is reasonable?

- Which aspects of the regional trends, variability and predictability are driven by external forcings, and which ones are due to internal climate variability?

To answer these questions we examine the outputs of three ocean reanalyses as well as two sets of experiments (historical simulations and decadal predictions) with the EC-Earth3 atmosphere-ocean general circulation model (AOGCM) version that contributed to the CMIP6 exercise (Eyring et al. 2016; Döscher et al. 2021).

The paper is structured as follows: Sect. 2 describes the data, the model and all methodological aspects considered in the analysis. Section 3 presents the results and is subdivided in three parts: (i) a description of upper OHC trends and variability in reanalyses, (ii) an evaluation of the forecast skill for the upper OHC in the North Atlantic, and (iii) an assessment of the upper OHC trends and their contribution to the total skill. Finally, Sect. 4 summarizes the main results and discusses them in light of previous studies.

2 Data and methods

2.1 Model and experimental setup description

EC-Earth is an European community Earth-System Model. The simulations analysed in this study were performed with the coupled AOGCM configuration of EC-Earth, using the version 3.3, the same one contributing to CMIP6. Its atmospheric component is the Integrated Forecast System (IFS; Hazeleger et al. 2010) from the European Centre for Medium-Range Weather Forecasts (ECMWF), based on its cycle cy36r4, and has a T255 horizontal resolution (grid size approximately 80km) and 91 vertical levels. The ocean component is the version 3.6 of the Nucleus for European Modelling of the Ocean (NEMO; Madec 2016), which is itself composed of the ocean model OPA (Ocean PArallelise) and the Louvain-La-Neuve sea ice model (LIM3; Rousset et al. 2015), both run with an ORCA1 horizontal resolution (ca. 1° nominal resolution) and 75 vertical levels. The atmospheric and oceanic components are coupled through OASIS (Craig et al. 2017). The vegetation fields are prescribed and have been derived from an historical simulation performed with EC-Earth3-Veg, a different model configuration that in addition includes interactive vegetation as represented by the LPJ-GUESS model (Smith et al. 2014).

In this study we analyse the historical and the Decadal Climate Prediction Project (DCPP; Boer et al. 2016) hindcast simulations performed with EC-Earth3 at the Barcelona Supercomputing Center (BSC), all contributing to CMIP6 (Bilbao et al. 2020).

The historical simulations follow the CMIP6 protocol (Eyring et al. 2016) and are driven with prescribed historical forcings of GHG concentrations, volcanic and anthropogenic aerosol concentrations, and solar variability. For the main analyses in this study an ensemble of 10 historical simulations covering the period 1960–2014 is used. This ensemble size allows us to establish fair comparisons with the decadal prediction system, for which a total of 10 members have been produced. The historical ensemble corresponds to the members r(2,7,12,17-22,24)i1p1f1. These members were chosen because they exhibit active Labrador Sea convection during the study period, and are therefore deemed to be more realistic than the members with suppressed convection (Bilbao et al. 2020). This set of simulations will be referred to as HIST hereinafter. A larger ensemble of 23 members (r(1-4,6,7,9-25)i1p1f1), which includes the 10 members of HIST, has also been used, but for a different purpose: to address the sensitivity of the results to the number or historical simulations, as previous work (Milinski et al. 2020) suggests that 20–30 members are needed to accurately capture the forced signals of sea surface temperature in the North Atlantic.

The EC-Earth3 DCPP hindcast is described in detail in Bilbao et al. (2020). This experiment consists of an ensemble of retrospective forecasts initialised every year on the first of November from 1960 to 2018, each composed of 10 members and following a full field initialisation method. The atmospheric initial conditions use ERA-40 reanalysis (Uppala et al. 2005) for the period 1960–1978 and ERA-Interim reanalysis (Dee et al. 2011) for the period 1979–2018. The ocean and sea ice initial conditions come from a forced ocean-sea ice reconstruction with NEMO-LIM, where the model is nudged towards three-dimensional ocean temperature and salinity from the ORAS4 reanalysis (Mogensen et al. 2012) and driven by surface fluxes from the Drakkar Forcing Set (DFS5.2; Brodeau et al. 2010). This set of simulations corresponds to the DCPP members r1i1p1f1-r10i1p1f1 and will be referred to as PRED henceforth.

All simulations are available at the CMIP6-Earth System Grid Federation (ESGF) portal (https://esgf-data.dkrz.de/search/cmip6-dkrz/) and can also be downloaded from the corresponding BSC-node (https://esgf.bsc.es/thredds/catalog/esgcet/catalog.html).

2.2 Reference observation-based datasets

Since ocean temperature observations, especially in the deep ocean, are temporally and spatially sparse, ocean reanalyses provide a physically consistent representation of the past OHC variability by assimilating temperature observations to models (among other variables). However, the sparsity of deep ocean temperature observations, model uncertainties in the representation of sub-scale processes and limitations of data assimilation, make these reanalysis inherently uncertain.

To take into account and quantify this uncertainty when evaluating the forecast skill of the EC-Earth3 simulations, we use monthly averaged ocean temperatures from three different reanalyses:

1. The Ensemble Coupled Data Assimilation experiment v3.1 (ECDA; Chang et al. 2013) from the Geophysical Fluid Dynamics Laboratory (GFDL).
2. The European Centre for Medium-Range Weather Forecasts (ECMWF) Ocean Reanalysis System 4 (ORAS4; Balmaseda et al. 2013).
3. The latest ocean reanalysis product from ECMWF (ORAS5; Zuo et al. 2019) and it's respective backward extension.

These three products are combined to produce a multi-reanalysis ensemble, which will be hereafter our observational-based reference, as it can smooth out some of the uncertainty specific to the individual reanalyses. The differences between the products provide an indirect measure of the observational uncertainty, which obey, among other things, to differences in model resolution, the model components, the forcing employed, the assimilation schemes and the assimilated observations. For further details on these datasets, please refer to Table 1.

One limitation, however, of the multi-reanalysis ensemble is that its mean might not be physically consistent. For this reason, some analyses have been repeated using ORAS4 as a reference, the reanalysis from which the initial conditions have been created and for which higher skill values are therefore expected.

2.3 Data pre-processing and diagnostics

The upper ocean heat content for the reanalyses and EC-Earth3 simulations was computed for each individual grid-cell from the three-dimensional ocean temperature field T (K) following the equation:

$$OHC(t) = \sum_{i=0}^{h} T(t)_i \, c_p \, \rho_0 \, \Delta x \, \Delta y \, \Delta z_i \qquad (1)$$

where c_p (J/kg/K) is the specific heat capacity of seawater, ρ_0 (kg/m^3) is the reference seawater density and Δx, Δy, Δz (m) are the individual cell dimensions. Vertically, the ocean temperature was integrated from the surface to 700 m or to the ocean floor, whichever came first. The upper 700 m ocean heat content will be referred to hereinafter as OHC700.

All datasets (reanalyses and model simulations) were interpolated to a common regular 1° × 1° resolution grid. The bilinear interpolation was performed using the Climate Data Operators package (from the Max-Planck Institute for Meteorology, https://code.zmaw.de/projects/cdo, version 1.7.2).

2.4 Trend analysis

The trends in OHC700 were computed by linear least-squares regression against time. OHC700 trend maps were produced in order to depict the OHC700 change patterns in the North Atlantic basin. To determine whether the trends are significantly different from zero, we use a two-sided t-test at the 95% confidence level.

The trends were computed for the period 1970–2014. This choice responds to two reasons. First, the decision to start in 1970 is motivated by constrains in the analysis of the predictions. Because we want to elucidate whether and how trends contribute to predictive skill at different forecast times, we need to define a fixed common period that is adequate to compute the skill for lead times from 1 to 10 years. Since the first start year in PRED is 1960, the first year in which we can evaluate the skill for its longest lead time is 1970. Second, HIST experiments finish in year 2014, and even if they could be concatenated with their corresponding SSP2-4.5

Table 1 Summary of features of the reanalyses used as benchmark here

Dataset name	ECDA3.1	ORAS4	ORAS5 (incl. backward extension)
Period covered	1961–2016	1959–2017	1958–2020
Horizontal resolution	1°	1°	1/4°
Vertical layers	50	42	75
Ocean model	MOM4	NEMO 3.0	NEMO 3.4
Sea ice model	Sea ice simulator	Prescribed	LIM2
Atmospheric data	Temperature, U, V (NCEP)	–	–
Forcing fluxes	–	ERA40, ERA-Interim	ERA40, ERA-Interim
Ocean/sea ice data	SSTs, TS profiles	SSTs, TS profiles, sea level anomalies	SSTs, TS profiles, sea level anomalies, sea ice concentrations
References	Chang et al. (2013)	Balmaseda et al. (2013)	Zuo et al. (2019)

scenario simulations, that would introduce uncertainty in the externally forced signals, and in particular in those of natural origin, which are not accounted for in the scenarios. Since the effects of natural forcings (such as volcanic aerosols) are expected to manifest at shorter time-scales, their contribution to the linear trends over the longer period of 1970–2014 is expected to be minor.

To illustrate the consistency of the trends across the reanalyses, and also within the ensemble members of each experiment, we use stippling to highlight the grid points where the trends for all members or reanalyses have the same sign.

2.5 Forecast drift correction and skill evaluation

Since the initialisation method used in PRED is full-field, as the forecasts progress they experience a spurious drift from the observed initial state towards the model attractor (e.g. Meehl et al. 2014). To correct for this drift, anomalies are computed using the "mean drift correction" method, which consists of computing the anomalies relative to the forecast-time-dependent climatology (e.g. Goddard et al. 2013). This is also the recommended approach to account for model drifts in DCPP (Boer et al. 2016).

In the case of the HIST experiment the anomalies are computed by subtracting the average of all years in the period of interest 1970–2014, for each ensemble member. The same procedure was done for the reanalyses.

To evaluate the forecast quality of the EC-Earth3 experiments we use the temporal anomaly correlation coefficient (ACC) as our skill metric. The statistical significance of

ACC differences is assessed following the methodology proposed by Siegert et al. (2017), a statistical test developed for cases—such as this study—where competing forecasting systems are strongly correlated with one another. Additionally, to evaluate the agreement of OHC700 patterns we use area-weighted spatial correlations, and the statistical significance is determined by a t-test at the 95% confidence level and taking into account the temporal autocorrelation.

3 Results

3.1 Ocean heat content trends and variability in reanalyses

The geographical pattern of OHC700 trends in the multi-reanalysis mean (Fig. 1a) shows that the North Atlantic has warmed throughout most of the basin in the past decades (1970–2014), with the main exception of the CSPNA, which experienced a long-term cooling. While the large-scale warming is consistent in the three reanalyses, as indicated by the stippling, the cooling trends are only consistent in a few grid points of the CSPNA, with the whole region showing a general lack of agreement in the sign of the trends. This calls for caution when interpreting the CSPNA cooling and its exact location, which is represented differently in the three reanalyses (Supp. Figure 1). While ORAS5 supports a widespread cooling, negative trends are very localised (and non-significant) in ECDA. It is important to note that the CSPNA cooling region in ORAS5 is very close to the area where a non-stationary bias in surface temperature
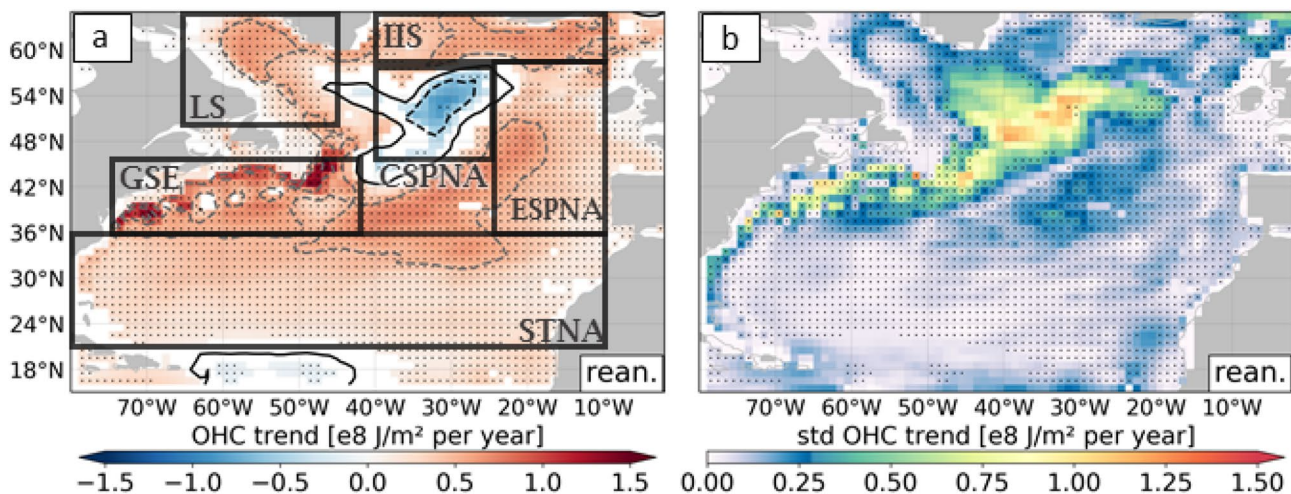

Fig. 1 a Map of the multi-reanalysis OHC700 mean trends over the period 1970–2014. Black boxes delimit the regions of interest specifically addressed in this study (see Table 2). Stippling is used to indicate the grid-points where the sign of the trend is the same in all the individual reanalysis. All trend values for which the trend is not significant are masked out in white. The full contour lines represent the zero trend and dashed lines show the subsequent trends in increments of 0.5×10^8 J/m² per year (grey/black lines indicate-positive/negative trends). **b** Map of the standard deviation in the OHC700 trends across the reanalyses. The same stippling as in panel **a** is included

has been reported in that same reanalysis product (Tietsche et al. 2020), linked to an excessively strong AMOC. For this reason, the realism of ORAS5's widespread strong cooling must be put, at least partially, into question.

Other regions also show noteworthy discrepancies across reanalyses. These differences are highlighted in Fig. 1b by the standard deviation of the OHC700 trends among the reanalyses. Apart from the CSPNA, where the largest values are found, the neighboring areas of the wider SPNA region, as well as the Gulf Stream also show considerable standard deviation values (over 0.25×10^8 J/m^2 per year). In the Gulf Stream, these are mostly associated to inter-reanalysis discrepancies in the magnitude of the warming trend, as the sign appears to be robust in most of the region.

Based on the spatial trends, we have selected several regions of interest to investigate how variability (beyond

trends) is reproduced in the reanalyses (defined in Table 2). The selection of these regions was inspired by the study of Maze et al. (2017), that classified local Argo floats temperature profiles into different clusters with coherent vertical heat distribution (Fig. 11 in Maze et al. 2017). These regions represent areas in which key ocean processes occur (e.g. the deep open ocean mixing in the Labrador and Irminger Seas, or the subtropical and subpolar gyre circulations), important currents are present (i.e. the Gulf Stream) or unique recent trends have developed (i.e. the CSPNA). The zonal and meridional boundaries were adjusted from Maze et al. (2017) to match the main features of the multi-reanalysis mean trends for the time period studied (boxes in Fig. 1a).

Figure 2 depicts the time-series of the OHC700 averaged in the North Atlantic (NA) and in the selected regions (see Table 2 for details on the exact boundaries considered). The whole NA basin displays a clear and significant warming trend that is consistent among the reanalyses (Fig. 2a and Table 3). In particular, OHC700 remains rather flat from the 1970s until the early 1990s, when it develops a strong positive trend that peaks at about year 2005. Since then, it has a slightly negative trend. Year to year OHC700 variations are generally small throughout the whole period.

At the regional level, interannual and decadal variability are generally more prominent than for the NA, as well as the range of the OHC700 variability (note the different y-axes in the panels). The Irminger-Iceland Seas (IIS) and the Eastern Subpolar North Atlantic (ESPNA) show similar variability,

Table 2 Regions of interest of this study and their boundaries

Acronym	Name	Boundary coordinates
NA	North Atlantic	[15–65] °N [80–2] °W
IIS	Irminger-Iceland Seas	[58–65] °N [40–10] °W
ESPNA	Eastern Subpolar North Atlantic	[36–58] °N [25–10] °W
CSPNA	Central Subpolar North Atlantic	[45–57] °N [40–25] °W
LS	Labrador Sea	[50–65] °N [65–45] °W
GSE	Gulf Stream Extension	[36–45] °N [75–42] °W
STNA	Subtropical North Atlantic	[22–36] °N [79–10] °W

Fig. 2 Time-series of the spatially averaged OHC700 across **a** the NA and **b–g** the individual selected sub-regions, computed from the reanalyses ORAS4, ORAS5 and ECDA, expressed as an anomaly per volume unit. The multi-reanalysis ensemble mean is also included (dashed grey line). Sub-regions are described in Fig. 1 and Table 2. Note that y-scale is different for panel **a**. The degree of agreement in OHC700 variability across the reanalyses is illustrated for each region by the temporal mean of the inter-reanalysis spread (defined by the range between the maximum and the minimum OHC700 values), indicated at the bottom corner of each panel
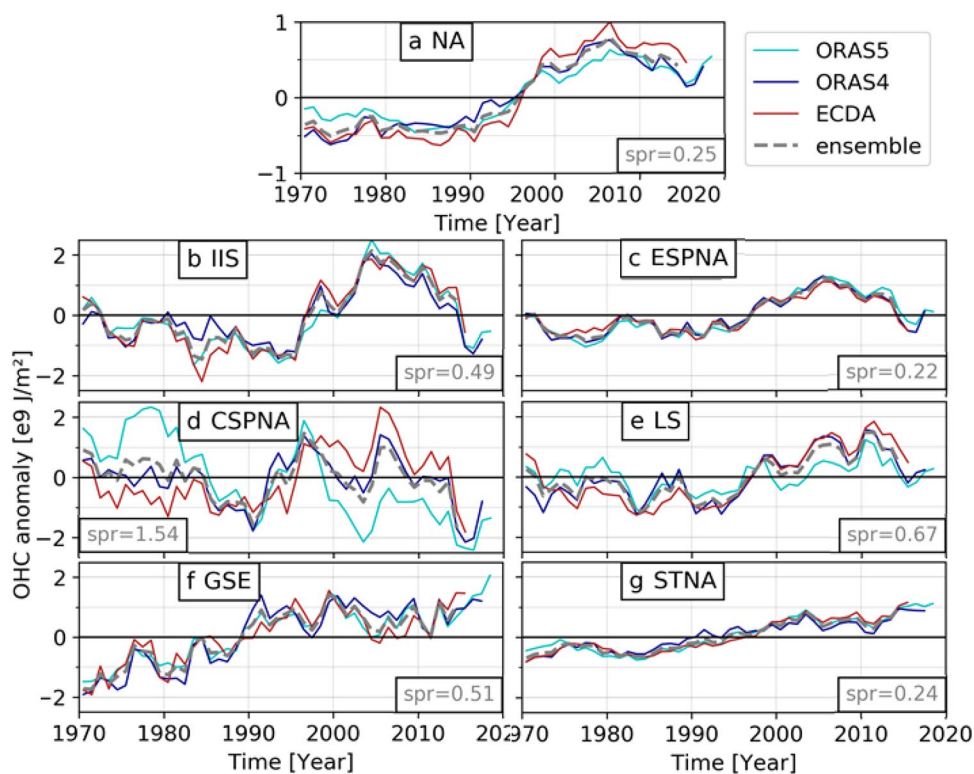
Table 3 Spatially averaged OHC700 trends for several North Atlantic regions in the reanalyses, expressed in 10^7 J/m^2 per year

Region	ECDA	ORAS5	ORAS4	Multi reanalysis mean
NA	3.82 (0.77)	2.27 (0.67)	3.10 (0.85)	3.06 (0.79)
IIS	5.46 (0.39)	4.70 (0.31)	4.08 (0.36)	4.74 (0.37)
ESPNA	3.03 (0.51)	4.24 (0.64)	3.71 (0.57)	3.66 (0.59)
CSPNA	3.82 (0.28)	−7.26 (0.54)	0.19* (< 0.01)	−1.08* (0.04)
LS	4.97 (0.49)	1.30 (0.11)	4.37 (0.52)	3.55 (0.44)
GSE	5.13 (0.59)	5.05 (0.67)	6.56 (0.67)	5.58 (0.68)
STNA	3.66 (0.81)	3.26 (0.72)	3.04 (0.81)	3.32 (0.82)

Trends refer to the period 1970–2014. The respective time-series can be seen in Fig. 2. All trend values are statistically significant at the 0.95 level, unless followed by an asterisk (*). The percentage of the total variance explained by the trends (characterised by the R^2, between the linear trend and the original time-series) is given between brackets

marked by decadal oscillations (i.e. a cooling during the 80s followed by a warming until 2005 and a subsequent cooling until present) superimposed on the long-term warming trend. Even if in both cases the trends are stronger than for NA (Table 3), they explain a smaller percentage of the total OHC700 variance (as indicated by the R^2 values in brackets). No major differences across the reanalyses are found for the OHC700 in these two regions.

The Labrador Sea (LS), and in particular the CSPNA, are the regions where the largest discrepancies between reanalyses are found. These differences are quantified by the temporal mean of the inter-reanalysis spread (indicated in the bottom-right corners of Fig. 2). The major differences come from ORAS5 for which the sign and/or the magnitude of the long-term trends is substantially different than in the other reanalyses. In the CSPNA, ORAS5 is the only reanalysis showing a (remarkably strong) cooling trend, while ECDA exhibits a significant warming trend, and the trend in ORAS4 is not statistically significant. There are indications that this strong trend in ORAS5 might not be realistic, since the CSPNA region is likely to be affected by the non-stationary bias reported in Tietsche et al. (2020), which occurs North East of the Grand Banks. These are relevant differences that lead to a multi-reanalysis mean that is negative and not significant (Table 3). It is important to remark that all three reanalyses show negative trends over the region, although they differ in the exact location, extension and strength, as already seen in Supp. Figure 1. The variance explained by the regional trend also varies substantially across the reanalyses, from above 50% in ORAS5 to virtually zero in ORAS4. In the LS the three reanalyses support a statistically significant warming trend (Table 3), accounting overall for a 44% of the total variance, although with a lower value in ORAS5 (i.e. 11%).

In the two remaining regions, the GSE and the Subtropical North Atlantic (STNA), the trends present a good degree of agreement across reanalyses. These are also the two regions in which the trends explain the largest fractions of variability, associated with reduced decadal-scale oscillations. While GSE exhibits the strongest trends, it also presents some interannual variations, not so evident in STNA, for which the trend explains on average about 80% of the total OHC700 variability, the largest value of all the regions. In both regions trends are rather linear, and could therefore be dominated by the changes in the external radiative forcings.

3.2 Forecast quality and added value of initialisation in North Atlantic

Even though the broad North Atlantic displays clear signs of warming in all reanalyses, the previous section has also revealed important regional differences regarding the magnitude and significance of the trends, and the presence of decadal variability in the upper OHC. This raises the questions of (i) whether and for which regions these upper OHC changes are predictable, (ii) whether trends contribute to predictability, and (iii) to what extent initialising internal variability matters for the predictive skill.

The first column of Fig. 3 shows the ACC for the EC-Earth3 PRED computed against the multi-reanalysis mean for various forecast years (FY; 1, 4, 7 and 10), revealing values that are generally high throughout the basin, especially in the first forecast year. In subsequent forecast years there is a rapid loss of skill in the vicinity of the CSPNA, the region with the largest differences across reanalyses. This rapid loss of skill could reflect that EC-Earth represents differently the dynamical processes that give rise to the local negative trends when compared to the reanalyses. Another region that exhibits a significant loss in skill with forecast time is the western STNA. Other regions, such as LS and IIS, also experience an initial moderate loss in skill, but it is recovered at the longest forecast years, which suggests that the temporary loss in skill could be related to the effect of initial shocks reported in Bilbao et al. (2020). A large part of this skill, in particular in the ESPNA and STNA areas, arises from the externally-forced signals, as shown by Supp. Figure 2, and could be enhanced with larger prediction ensembles.

The second column of Fig. 3 shows the ACC difference between the predictions and the historical simulations, which allows us to determine the added value of initialisation on skill. In the first FY, initialisation is largely beneficial, except for some parts of the eastern STNA and the sea North of Scotland, where it leads to skill degradation. In the subsequent FYs, only the ESPNA, a region that in other models seems to be controlled by advective processes linked to the AMOC (Borchert et al. 2018), shows a consistent added
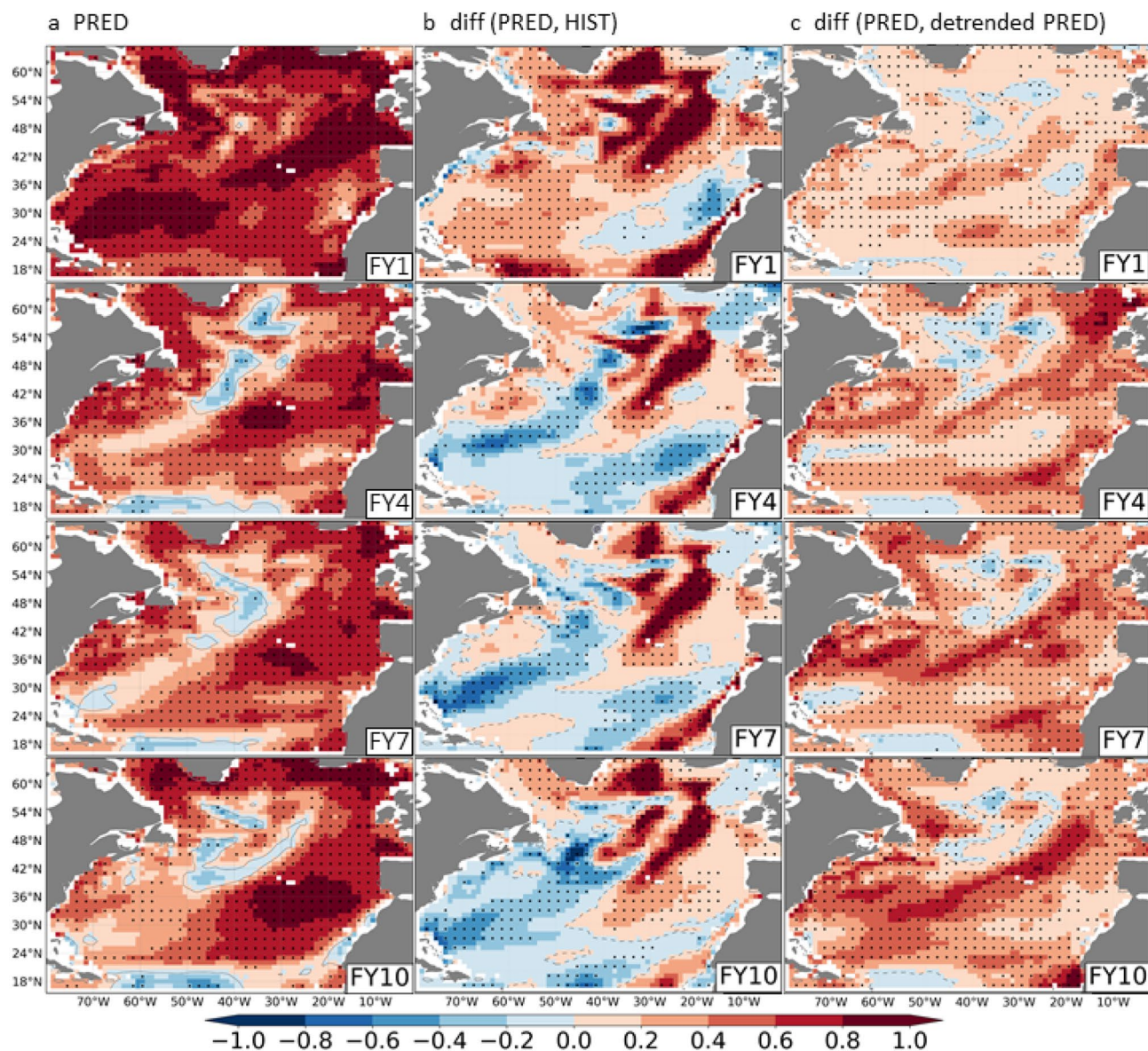
Fig. 3 **a** Anomaly correlation coefficient (ACC) maps of OHC700 in the initialised EC-Earth predictions for the forecast years 1, 4, 7 and 10. Stippling indicates where the correlation is statistically significant at the 95% confidence level. Stippling is only applied in every 4th grid cell for the sake of visibility. **b** Difference in the ACC values for the initialised and uninitialised predictions (PRED and HIST, respec-tively). **c** Difference in the ACC values in PRED for the undetrended and the detrended OHC700 anomalies. In **b**, **c** stippling highlights correlations that are significantly different at the 95% confidence level. All ACC values are evaluated against the multi-reanalysis ensemble mean for the period 1970–2014

value of initialisation on the predictive skill. Interestingly, the latest FYs show again improved skill with initialisation in the IIS and LS regions, which would be consistent with the recovery from the initial shock effects previously men-tioned. This recovery of skill could be linked to the skillfully predicted ESPNA OHC700 anomalies from earlier FYs, which are advected westward by the mean subpolar gyre circulation. By contrast, skill in the southwestern side of the basin is substantially hindered by initialisation, in particular

over the Gulf Stream area, where the initialisation shock might have a longer-lasting effect.

The third column of Fig. 3 shows the contribution of the linear long-term trends to ACC skill for the same forecast years. This is done by calculating the difference in terms of ACC between the raw anomalies of PRED and its detrended anomalies. In the first FY, for which we have shown that the influence of initialisation is notably higher, trends play a significant role in skill mainly south of the subpolar lati-tudes, including over the GSE and the STNA. At subsequent

FYs, as the added value of initialisation on forecast skill gets confined to the ESPNA, stronger and more significant contributions from trends are seen all across the basin, except in the CSPNA. The fact that the relative contribution of the trends to the final predictive skill increases with forecast time suggests that, at least for some regions like the LS, IIS and ESPNA, the trends in the initial model state are incompatible with the multi-reanalysis trends. The realism of these trends along the forecast is further explored in Sect. 3.3.

Very similar results in terms of predictive skill, added value of initialization, and predictive role of the trends are obtained when using ORAS4 as a reference (Fig. 3 vs Supp. Figure 3), the main difference being in the first forecast year, for which, as expected, skill is substantially higher when evaluated against ORAS4.

We now evaluate the skill of the mean OHC700 in the North Atlantic and in the individual regions and how it evolves for forecast years 1–10 in PRED in comparison to HIST (Fig. 4). We include two indicators of statistical significance: (i) all the individual ACC values are represented with cyan dots when they are significantly different from zero at the 95% confidence level; and (ii) ACC values for HIST are additionally highlighted with a red cross when their value is significantly different from the PRED one. The role of the trend is also tested by recomputing the skill after both the reanalyses and PRED anomalies have been detrended (dashed line in Fig. 4). For these ACC values, the

same significance indicators as for the historical ensemble are included.

Figure 4a shows that the whole NA is highly predictable at all forecast years, a result that is largely but not exclusively due to the external forcing influence. HIST exhibits high ACC values, although they are significantly higher in PRED at almost all FYs, supporting a small but beneficial effect of initialisation on skill. We also show that most of the skill comes from the predictability of the linear trend, given the very low ACC values obtained for the detrended series. This linear trend is largely dominated by a rapid warming in the mid-1990s (Fig. 2a), which has been previously associated with a prolonged strengthening of the AMOC since the 1960s, leading to increased ocean heat transport into the region, an instantaneous response to a strong negative NAO phase in 1995–1996 (Robson et al. 2012). Only the effects of the former on the OHC are deemed to be predictable, and thus thought to be reproduced in the EC-Earth simulations.

The individual regions show contrasting results. In the IIS, PRED has also significant skill at all FYs, although with an initial drop and a subsequent recovery during the 2nd to 7th FYs, which is likely related to the initialisation shock (Fig. 4b). ACC values are significantly lower for HIST than PRED in this region, suggesting that internal variability contributes decisively to the local predictive skill. The external forcing also provides significant skill, although this can only be detected when using the

Fig. 4 ACC skill assessment of the spatially averaged OHC700 in **a** the North Atlantic, and **b–g** all the selected individual regions. Skill values are shown for the PRED (blue lines) and HIST ensembles (grey lines) and are evaluated against the multi-reanalyses mean. In PRED, skill is also computed after detrending both the forecast anomalies and the reanalysed anomalies (detrended PRED; dashed blue lines). Cyan dots indicate ACC values that are significantly different from zero at the 95% confidence level. Red crosses indicate that the HIST or the detrended PRED ACC values are significantly different from the PRED ones
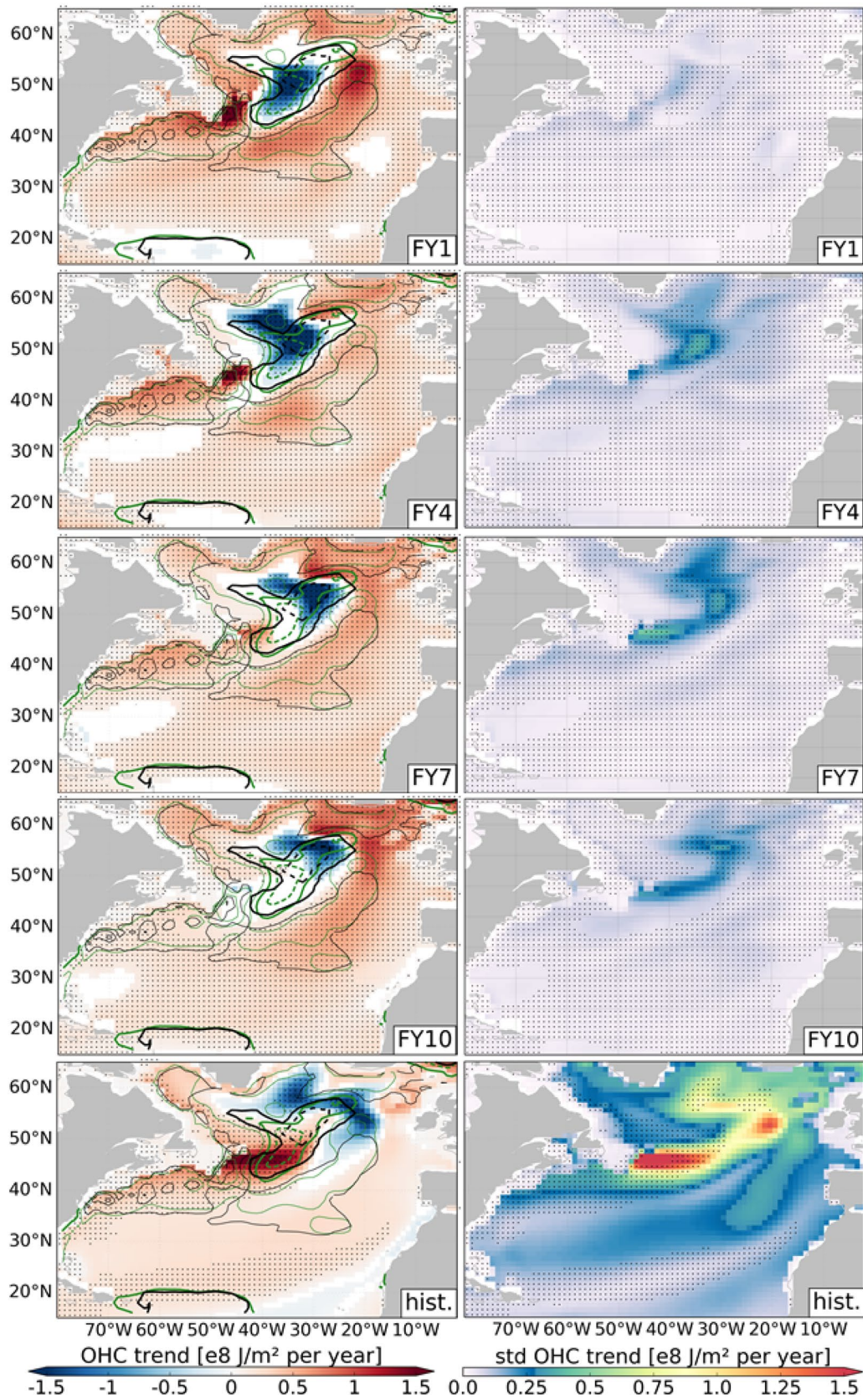
large HIST ensemble (Supp. Table 1). This is a region that clearly benefits from initialisation. The trends explain part but not all of the PRED skill, given the significantly lower values for the detrended timeseries. Similar results are obtained over the ESPNA (Fig. 4c), although with no apparent signal of initialisation shock effects, as it experiences a lower and more gradual decrease in skill, with no recovery at the longest FYs.

The CSPNA exhibits poor skill (only significant in the first FY of PRED) in both the HIST and PRED experiments (Fig. 4d). This is also the region in which the trends seem to play the smallest role in the predictive skill (given that solid and dashed lines are virtually identical), which is not surprising given that its multi-reanalysis mean trend was not significant (Table 3). The lack of skill might derive from limitations in the prediction system to represent realistically the local variability. However, caution is recommended because the large uncertainties across reanalyses seem to critically affect the skill assessment. When re-evaluating the skill in the region against each of the individual reanalyses, the results are drastically different (Supp. Figure 5), and seem to be affected by the sign and magnitude of the reanalysed trends (Table 3). Reducing the uncertainty of the OHC over this region is therefore essential to evaluate trustworthily its predictive skill. By contrast, the lack of skill for HIST does seem to be a robust result, independent of the reduced ensemble size (Supp. Table 1), and also of the reanalysis used as a reference (grey line in Supp. Figure 5), which suggests a minor role of the external forcing in the CSPNA.

The LS (Fig. 4e), the region where the initial shock is triggered, shows clear evidence of its effects, with a longer-lasting decrease in skill than in the IIS (the minimum occurs in FY6, compared to FY4 in IIS). Another important difference with respect to IIS is that LS skill is largely arising from the external forcings (as indicated by the high ACC values in HIST), which might be affecting vertical mixing through buoyancy-induced changes in local stratification. Interestingly, after the skill recovery in FY9–10, PRED skill becomes significantly higher than for HIST, something that only occurred in the first FY. This would be consistent with the aforementioned advection of OHC700 anomalies from the ESPNA, a region in which initialisation matters decisively for skill.

The two remaining regions, the GSE and STNA, show similar ACC features (Fig. 4f, g respectively). In both cases the predictive skill seems to come almost exclusively from the external forcings. Indeed, PRED skill is only significantly higher than in HIST in the first FY. Both are also regions with important long-term trends (as previously seen in Table 3) that are critical to achieve high levels of skill.

In summary, this analysis reveals that the regions in which the external forcings play a dominant role on the prediction skill are the GSE, the STNA and to a lesser extent, the LS. This latter, together with the IIS and ESPNA, show the largest benefits of initialisation, which are still significant at the longest FYs. In all these five regions, trends contribute substantially to the final prediction skill. The same conclusions can be drawn when computing the skill against ORAS4 instead of the multi-reanalysis mean (Supp. Figure 4), which further supports the results. In the following section we evaluate the realism of these trends in the predictions and historical ensemble, assessing to what extent they are externally forced, and whether and how they change with forecast time.

3.3 Representation of upper OHC long-term trends in the EC-Earth3 predictions and historical simulations

We have shown that, at the regional level, trends can contribute substantially to the skill, with different weights from both external radiative forcing changes and internal variability. We now explore how the long-term OHC700 linear trends are represented in the predictions as a function of forecast lead time and also in the HIST ensemble (both in shaded colors), and how they compare with the multi-reanalysis mean trends (left column of Fig. 5). Large areas of consistency between the predicted and reanalysis trends are evident, especially in the earliest forecast years, in which initialisation corrects spurious trends in the mean HIST ensemble (shown in the bottom row of Fig. 5), that will be discussed further ahead. In FY1 of PRED, the region of negative trends is well collocated with the multi-reanalysis mean pattern. This region is flanked at the east by an area of strong positive trends, also present in ORAS4 (Supp. Figure 1), the reanalysis that was used to generate the initial conditions of PRED. These locally large trends are, however, not supported by the other two reanalyses, leading to an initial disagreement between PRED and the multi-reanalysis mean trends. This disagreement is reduced at subsequent FYs in which the positive trends are smoothed.

The trend pattern evolves with forecast time as climate noise grows, eroding the signal from initialisation, which makes different members diverge. Interestingly, the trends in PRED show similar spatial features to the multi-reanalysis throughout the forecast, with some local differences. For example, along the Gulf Stream Extension, warming trends

that are initially strong seem to fade away at the latest FYs. Likewise, the initial cooling region over the CSPNA expands northwestwards towards the Labrador and Irminger Seas. This region of cooling trends is also the one exhibiting the largest intra-ensemble spread (right column of Fig. 5), not only in the forecasts but also, much more importantly, in the historical experiments.

The larger intra-ensemble spread in HIST than in PRED trends can explain the discrepancies between the mean trends in the HIST ensemble and in the multi-reanalysis mean. Even though the HIST ensemble mean also exhibits a cooling, it happens over a wider area, located within the IIS and ESPNA regions. All individual HIST experiments agree (as indicated by the stippling) in representing a cooling trend over a small band North of CSPNA. This suggests that the misplacement of the location with respect to the reanalyses corresponds to a structural model bias. The same happens along the Gulf Stream Extension in which the HIST trend is overly warm compared to the multi-reanalysis mean, and appears to extend too far eastward into the CSPNA, a feature for which all historical members also systematically agree. We note that in HIST there is little to no intra-ensemble agreement regarding the sign of the trends in regions like the Labrador Sea, the IIS and the ESPNA.

By forecast year 10 of PRED, the spread in the representation of the trends remains small compared to the HIST ensemble spread (right column of Fig. 5). This suggests that initialisation has a long-lasting beneficial effect on the representation of the long-term trends, both in time and space. Nevertheless, the trend patterns are similar in HIST, PRED and the multi-reanalysis mean, characterised by a large-scale warming and a regional cooling in the northern NA. These findings suggests that the overall pattern is mostly driven by external forcing, and initialisation is essential to represent the trends in the correct location.

An alternative and more synthesised way to evaluate the impact of initialisation in the representation of the trends is to directly compare the spatial patterns from PRED and HIST ensemble with the reanalyses patterns. Figure 6a shows the area-weighted spatial correlation for the North Atlantic region between the multi-reanalysis mean OHC700 trends and those in individual members and the ensemble-mean of PRED (thin and thick red lines, respectively), as a function of FY. The corresponding correlations for the HIST are also shown (in grey). The spatial correlations for PRED show systematically higher values than for HIST at all forecast times and for all ensemble members. Also note that the inter-member spread in PRED (red envelop in Fig. 6a) remains substantially smaller than for HIST (grey envelop), indicating that the initialisation of internal variability helps constraining the long-term trends for more than a decade. The low spatial correlation values in HIST can be explained by a combination of two key factors already discussed in light of Fig. 5: (i) that internal variability plays an important role in some local trends that cannot be properly represented in the forced uninitialised experiments; and (ii) that systematic model biases are fully developed in some areas, preventing the model to represent the forced response in the right location. An additional analysis computing the spatial correlations against each of the individual reanalyses (Supp. Figure 6) shows that the previous results, and in particular the important role of internal variability in the development of the trends, are supported by both ORAS5 and ORAS4 but not by ECDA, for which the spatial correlations remain low and comparable between PRED and HIST. This is probably due to the fact that in ECDA large positive trends dominate in the CSPNA region, where only a very reduced area shows a long-term (and very weak) cooling. To corroborate if internal variability generally matters for understanding the spatial trends, the spatial correlations were recomputed in Fig. 6b
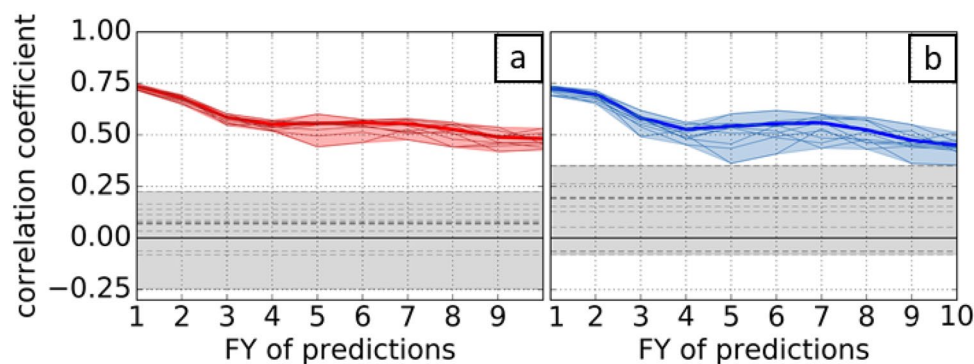

Fig. 6 **a** Area-weighted spatial correlations as a function of forecast time between the OHC700 trend patterns in the EC-Earth experiments (red for PRED, grey for HIST), and the trends in the multi-reanalysis mean. Thin lines represent the correlations for individual members in PRED and HIST, and the thick lines the corresponding values for their ensemble mean. The ensemble spread is indicated respectively by the red and grey shading. **b** The same as in **a**, but with spatial correlations only calculated over the stippled grid points in Fig. 1 (cells in which all reanalyses support a trend of the same sign)

after masking out the regions where the reanalyses do not agree in the sign of the trend (based on stippled areas of the multi-reanalysis ensemble mean), to thus exclude the areas with large observational uncertainties. The figure confirms that internal variability does matter in the areas where the reanalyses tend to agree.

To further assess the contributions of internal and external sources of variability to the representation of the local trends, we compare the regionally averaged trends in the reanalyses with the corresponding trends in PRED (as a function of forecast time) and HIST (Fig. 7). In the broad North Atlantic, the long-term trends are systematically underestimated in both DCPP and HIST experiments, although with important differences between them. The mean HIST trend is half of the multi-reanalysis mean, and its ensemble spread does not overlap with any of the individual reanalysis trends. The trends in PRED start close to the multi-reanalysis mean trend, and are fully included in the multi-reanalysis spread for the first forecast year (Fig. 7a). However, for subsequent forecast years NA trends rapidly weaken, dropping down by half by FY4 to similar values than in HIST. This quick reduction could be due to the initialisation shock, because in later forecast years the NA trend steadily increases, overlapping again with the reanalysis spread. The fact that all reanalyses display a positive trend in the NA indicates that

there is certainly a forced origin in the trend, but internal variability seems to be essential to explain the magnitude of the trend.

At the individual regions, other differences between the trends in PRED and HIST emerge. In both the IIS and the ESPNA the trends in PRED remain rather close to the reanalysed trends, agreeing particularly well for the latter region (Fig. 7b, c). In the ESPNA, the prediction and multi-reanalyses ensemble spreads overlap at all forecast years and are narrow, indicating a reduction in observational and model uncertainty (Fig. 7c), compared to the IIS region. In both IIS and ESPNA HIST shows a mean negative trend that is very close to zero, with individual members supporting both positive and negative trend values. The large spread in HIST trends, which overlaps with the much narrower multi-reanalysis and PRED spreads in IIS (Fig. 7b), and is very close to the corresponding spreads in ESPNA (Fig. 7c), suggests that trends in both regions are largely controlled by internal variability. PRED results indicate that internal variability can be correctly initialised and that it contributes substantially to the local predictive skill, as shown in Fig. 4, and also identified in other decadal prediction systems (Meehl et al. 2014; Yeager et al. 2018).

The large differences between the individual reanalysis (ECDA supporting a significantly positive trend, ORAS5 a

Fig. 7 **a** Spatially averaged North Atlantic OHC700 trends in the multi-reanalysis, PRED and HIST ensembles evaluated over the period 1970–2014, as a function of forecast time. The ensemble mean trends are represented by the dashed thick lines, and the trends for the individual members by the thin lines. Grid cells with non-significant trends are masked out before the regional averages are computed. Note that the y-axis is not the same for all panels, to improve the comparability of the different ensembles. **b**–**g** The same as in panel a but for the selected NA regions: IIS, ESPNA, CSPNA, GSE and STNA
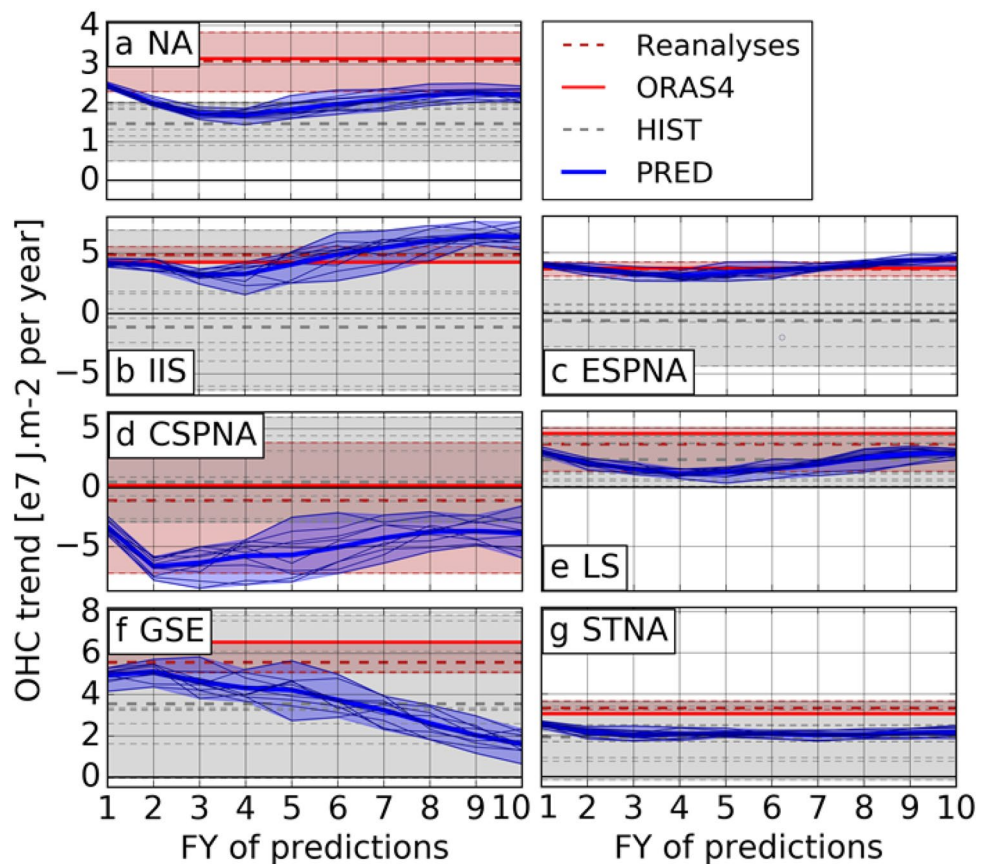
significantly negative trend, and ORAS4 a trend not significantly different from zero; Table 3) hamper the interpretation of the CSPNA results (Fig. 7d), as it is unclear whether HIST and PRED represent them realistically. From the wide spread in HIST trends, together with their mean value close to zero, we can at least conclude that simulated trends in this region are largely influenced by internal variability. Different representations of internal variability processes in response to the assimilated observations might cause the differences across reanalyses. The trends in PRED are small and negative at the beginning of the forecast, similar to the ORAS4 value, and exhibit an abrupt change to more negative values in the second year, presumably resulting from the initialisation shock. At subsequent forecast years the trends recover slowly towards the initial trend values but remain significantly negative, like in ORAS5.

In the LS, GSE, and STNA regions (Fig. 7e–g), the reanalyses, PRED and HIST show consistently positive trend values for all their individual members, thus supporting an important contribution of the external forcing to the trends, and through it, on the predictive skill for the OHC700 (Fig. 4). Yet, some differences among the regions can also be seen. The LS is the second region with the largest discrepancies across the reanalyses (Fig. 2, but unlike for CSPNA, all reanalyses support a positive OHC700 trend. The PRED ensemble mean starts and ends very close to the multi-reanalysis mean, showing considerably lower trends from forecast years 2–6, the same years in which the initialisation shock manifests locally (Bilbao et al. 2020). In the GSE, the trends in PRED also starts close to the reanalysis ones, overlapping with them up to FY6, but after that they decrease quickly and show no final sign of recovery. Another difference with respect to the LS is that in the GSE, PRED trends do not converge towards the HIST mean trend, as they reach considerably lower trends by FY10. This might be related to a longer-term effect of the initialisation shock, which have been shown in Bilbao et al. (2020) to bring the model to a different equilibrium state. Additionally, a large spread of the HIST can be seen in Fig. 7f, indicating that internal variability in the region also contributes substantially to the multi-decadal trends. Finally, the STNA trends start slightly weaker than in the reanalyses and by FY2 they stabilize around the mean HIST value. Unlike in the two other regions, it shows no sign of initial shock effects.

4 Conclusions and final remarks

This study has explored the trends in upper 700 m ocean heat content (OHC700) for the period 1970–2014 in a set of decadal climate predictions and an ensemble of historical experiments performed with the EC-Earth3 model, using reanalyses as benchmarks to evaluate the skill of the model. The main findings of the paper are described as follows:

- A comparison of OHC700 variability and trends in different reanalyses has revealed that the North Atlantic (NA) ocean, for its most part, has been progressively warming since the 1970s, albeit with regional differences. Changes in two major regions stand out: the Central Subpolar North Atlantic (CSPNA), where the multi-reanalysis ensemble reveals a cooling trend, although its extent, location and intensity varies substantially across the reanalyses; and the Gulf Stream Extension (GSE), in which reanalyses exhibit the strongest warming, but with differences in terms of the magnitude.

- A skill assessment of OHC700 in the two EC-Earth3 experiments has shown a high level of predictive skill at all forecast ranges in the North Atlantic, with higher values in the Eastern Subpolar North Atlantic (ESPNA) and the Subtropical North Atlantic (STNA). Other regions like the Labrador Sea (LS), the Irminger-Iceland Seas (IIS) and the GSE also show high initial levels of skill, which degrade after some forecast years due to an initialisation shock affecting the decadal predictions that were considered in this study.

- An important part of the skill comes from initialisation, specially in the ESPNA and in the IIS regions, in line with previous studies documenting a positive impact on skill of ocean initialisation (e.g. Doblas-Reyes et al. 2013). By contrast, predictive skill in regions like the LS, GSE and the STNA are found to be dominated, although not explained exclusively, by the influence of the external forcings, especially in both GSE and STNA regions, where all skill is lost when removing the linear trend.

- Very limited skill has been found to predict the CSPNA OHC700 variability, only skilful for the first forecast year when the multi-reanalysis mean is used as a reference. When evaluated against the individual reanalyses, very different skill levels are obtained for this region, which reflects that large uncertainties across reanalyses (and observations) prevent a proper skill assessment in the region, and are potentially hampering its correct initialisation. Extra care should be taken when evaluating skill in this region with ORAS5 due to this reanalysis' non-stationary bias (Tietsche et al, 2020).

- Initialisation has been found to be key to represent the long-term trends in the right location, with predicted trends largely outperforming the realism of the historical ones, even after ten forecast years. The historical simulations place the regions of the maximum warming and cooling trends in the wrong geographical location, which could be due to the effect of structural model biases. OHC700 trends, which depending on the region are mostly forced or arise from internal decadal

variability, have been found to contribute decisively to the final OHC700 prediction skill in the whole North Atlantic.

Despite the high levels of skill identified in EC-Earth3 and other forecast systems to predict OHC700 changes in the North Atlantic a decade in advance, and the numerous studies linking North Atlantic variability with numerous climate signals over North America, Northern Africa and Europe (e.g. Hodson and Sutton 2005; Folland et al. 1986; Zhang and Delworth 2006; Kushnir et al. 2010), there is very little evidence of multi-year predictive skill over these continents (Yeager and Robson 2017b). This might indicate that the key atmospheric teleconnection mechanisms enabling these impacts are not well represented in models. It could also be related to the fact that state-of-the-art models used for climate prediction tend to substantially underestimate the amplitude of the predictable signal, a problem that is particularly important in the North Atlantic sector (Scaife and Smith 2018). A recent study (Smith et al. 2020) has shown that this problem can be partly circumvented through the exploitation of large ensembles of decadal climate predictions, which led to skilful predictive capacity for the North Atlantic Oscillation and its climate fingerprints over the continents on decadal timescales, and also enhanced predictive skill for the AMV. Such large ensembles of predictions can also be exploited to better disentangle the forced and internal sources of predictability for the North Atlantic OHC, and to ascertain to what extent the results illustrated in this study are model-dependent. This will be the task of a follow-up study, that will also investigate the regional skill differences across models, and if these can be traced back to specific mean state model properties like the strength of the meridional and barotropic circulations, the western boundary currents, or the Labrador Sea stratification. These are critical aspects to consider in the tuning of the future model versions used for climate prediction purposes.

Declarations

References

Abraham JP, Baringer M, Bindoff NL, Boyer T, Cheng LJ, Church JA, Conroy JL, Domingues CM, Fasullo JT, Gilson J, Goni G, Good SA, Gorman JM, Gouretski V, Ishii M, Johnson GC, Kizu S, Lyman JM, Macdonald AM, Minkowycz WJ, Moffitt SE, Palmer MD, Piola AR, Reseghetti F, Schuckmann K, Trenberth KE, Velicogna I, Willis JK (2013) A review of global ocean temperature observations: implications for ocean heat content estimates and climate change. Rev Geophys 51(3):450–483. https://doi.org/10.1002/rog.20022, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/rog.20022

Balmaseda MA, Mogensen K, Weaver AT (2013) Evaluation of the ECMWF ocean reanalysis system ORAS4. Q J R Meteorol Soc 139:1132–1161. https://doi.org/10.1002/qj.2063

Barnston AG (1994) Linear statistical short-term climate predictive skill in the northern hemisphere. J Clim 7(10):1513–1564. https://doi.org/10.1175/1520-0442(1994)007<1513:LSSTCP>2.0.CO;2,

https://journals.ametsoc.org/view/journals/clim/7/10/1520-0442_1994_007_1513_lsstcp_2_0_co_2.xml

Bilbao RAF, Gregory JM, Bouttes N, Palmer MD, Stott P (2019) Attribution of ocean temperature change to anthropogenic and natural forcings using the temporal, vertical and geographical structure. Clim Dyn 53(9):5389–5413. https://doi.org/10.1007/s00382-019-04910-1

Bilbao R, Wild S, Ortega P, Acosta-Navarro J, Arsouze T, Bretonnière PA, Caron LP, Castrillo M, Cruz-García R, Cvijanovic I, Doblas-Reyes FJ, Donat M, Dutra E, Echevarría P, Ho AC, Loosveldt-Tomas S, Moreno-Chamarro E, Pérez-Zanon N, Ramos A, Ruprich-Robert Y, Sicardi V, Tourigny E, Vegas-Regidor J (2020) Assessment of a full-field initialised decadal climate prediction system with the CMIP6 version of EC-Earth. Earth Syst Dyn Discuss 2020:1–30. https://doi.org/10.5194/esd-2020-66, https://esd.copernicus.org/preprints/esd-2020-66/

Boer GJ, Smith DM, Cassou C, Doblas-Reyes F, Danabasoglu G, Kirtman B, Kushnir Y, Kimoto M, Meehl GA, Msadek R, Mueller WA, Taylor KE, Zwiers F, Rixen M, Ruprich-Robert Y, Eade R (2016) The decadal climate prediction project (DCPP) contribution to CMIP6. Geosci Model Dev 9(10):3751–3777. https://doi.org/10.5194/gmd-9-3751-2016

Borchert LF, Müller WA, Baehr J (2018) Atlantic ocean heat transport influences interannual-to-decadal surface temperature predictability in the North Atlantic region. J Clim 31(17):6763–6782

Borchert LF, Menary MB, Swingedouw D, Sgubin G, Hermanson L, Mignot J (2021) Improved decadal predictions of North Atlantic subpolar gyre SST in CMIP6. Geophys Res Lett 48(3):e2020GL091,307

Brodeau L, Barnier B, Treguier AM, Penduff T, Gulev S (2010) An ERA40-based atmospheric forcing for global ocean circulation models. Ocean Model 31(3–4):88–104. https://doi.org/10.1016/j.ocemod.2009.10.005

Buckley MW, Marshall J (2016) Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: a review. Rev Geophys 54:5–63. https://doi.org/10.1002/2015RG000493

Buckley MW, DelSole T, Susan Lozier M, Li L (2019) Predictability of North Atlantic Sea surface temperature and upper-ocean heat content. J Clim 32(10):3005–3023. https://doi.org/10.1175/JCLI-D-18-0509.1

Chang YS, Zhang S, Rosati A, Delworth TL, Stern WF (2013) An assessment of oceanic variability for 1960–2010 from the GFDL ensemble coupled data assimilation. Clim Dyn 40(3–4):775–803. https://doi.org/10.1007/s00382-012-1412-2

Chemke R, Zanna L, Polvani LM (2020) Identifying a human signal in the North Atlantic warming hole. Nat Commun 11(1):1–7. https://doi.org/10.1038/s41467-020-15285-x

Chen X, Tung KK (2014) Varying planetary heat sink led to global-warming slowdown and acceleration. Science 345(6199):897–903. https://doi.org/10.1126/science.1254937

Cheng W, Chiang JC, Zhang D (2013) Atlantic meridional overturning circulation (AMOC) in CMIP5 models: RCP and historical simulations. J Clim 26(18):7187–7197. https://doi.org/10.1175/JCLI-D-12-00496.1

Craig A, Valcke S, Coquart L (2017) Development and performance of a new version of the OASIS coupler, OASIS3-MCT-3.0. Geosci Model Dev 10(9):3297–3308. https://doi.org/10.5194/gmd-10-3297-2017

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars AC, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, Mcnally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F (2011) The

ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656):553–597. https://doi.org/10.1002/qj.828

Doblas-Reyes FJ, Andreu-Burillo I, Chikamoto Y, García-Serrano J, Guemas V, Kimoto M, Mochizuki T, Rodrigues LRL, van Oldenborgh GJ (2013) Initialized near-term regional climate change prediction. Nat Commun 4(1):1–9. https://doi.org/10.1038/ncomms2704

Döscher R, Acosta M, Alessandri A, Anthoni P, Arneth A, Arsouze T, Bergmann T, Bernadello R, Bousetta S, Caron LP, Carver G, Castrillo M, Catalano F, Cvijanovic I, Davini P, Dekker E, Doblas-Reyes FJ, Docquier D, Echevarria P, Fladrich U, Fuentes-Franco R, Gröger M, v. Hardenberg J, Hieronymus J, Karami MP, Keskinen J-P, Koenigk T, Makkonen R, Massonnet F, Ménégoz M, Miller PA, Moreno-Chamarro E, Nieradzik L, van Noije T, Nolan P, O'Donnell D, Ollinaho P, van den Oord G, Ortega P, Prims OT, Ramos A, Reerink T, Rousset C, Ruprich-Robert Y, Le Sager P, Schmith T, Schrödner R, Serva F, Sicardi V, Sloth Madsen M, Smith B, Tian T, Tourigny E, Uotila P, Vancoppenolle M, Wang S, Wårlind D, Willén U, Wyser K, Yang S, Yepes-Arbós X, Zhang Q (2021) The EC-Earth3 Earth system model for the climate model intercomparison project 6. Geosci Model Dev (Accepted). https://doi.org/10.5194/gmd-2020-446

Drijfhout S, van Oldenborgh GJ, Cimatoribus A (2012) Is a decline of AMOC causing the warming hole above the North Atlantic in observed and modeled warming patterns? J Clim 25(24):8373–8379. https://doi.org/10.1175/JCLI-D-12-00490.1

Eyring V, Bony S, Meehl GA, Senior CA, Stevens B, Stouffer RJ, Taylor KE (2016) Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci Model Dev 9(5):1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Folland CK, Palmer TN, Parker DE (1986) Sahel rainfall and worldwide sea temperatures, 1901–85. Nature 320(6063):602–607. https://doi.org/10.1038/320602a0

Frajka-Williams E, Beaulieu C, Duchez A (2017) Emerging negative Atlantic Multidecadal Oscillation index in spite of warm subtropics. Sci Rep 7(11224). https://doi.org/10.1038/s41598-017-11046-x

Frankignoul C, Gastineau G, Kwon YO (2017) Estimation of the SST response to anthropogenic and external forcing and its impact on the Atlantic multidecadal oscillation and the Pacific decadal oscillation. J Clim 30(24):9871–9895. https://doi.org/10.1175/JCLI-D-17-0009.1, https://journals.ametsoc.org/view/journals/clim/30/24/jcli-d-17-0009.1.xml

Gastineau G, Mignot J, Arzel O, Huck T (2018) North Atlantic Ocean internal decadal variability: role of the mean state and ocean-atmosphere coupling. J Geophys Res Oceans 123:1–22. https://doi.org/10.1029/2018JC014074

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 CA, Stephenson DB, Meehl GA, Stockdale T, Burgman R, Greene AM, Kushnir Y, Newman M, Carton J, Fukumori I, Delworth T (2013) A verification framework for interannual-to-decadal predictions experiments. Clim Dyn 40(1–2):245–272. https://doi.org/10.1007/s00382-012-1481-2

Ham YG, Kim JH, Luo JJ (2019) Deep learning for multi-year ENSO forecasts. Nature 573(7775):568–572. https://doi.org/10.1038/s41586-019-1559-7

Hazeleger W, Severijns C, Semmler T, Ştefănescu 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, Willén U (2010) EC-Earth. Bull Am Meteorol Soc 91(10):1357–1364. https://doi.org/10.1175/2010BAMS2877.1, http://journals.ametsoc.org/doi/10.1175/2010BAMS2877.1

Hermanson L, Bilbao R, Dunstone N, Ménégoz M, Ortega P, Pohlmann H, Robson JI, Smith DM, Strand G, Timmreck C et al (2020) Robust multiyear climate impacts of volcanic eruptions in decadal prediction systems. J Geophys Res Atmos 125(9):e2019JD031,739

Hewitt HT, Bell MJ, Chassignet EP, Czaja A, Ferreira D, Griffies SM, Hyder P, McClean JL, New AL, Roberts MJ (2017) Will high-resolution global ocean models benefit coupled predictions on short-range to climate timescales? Ocean Model 120:120–136. https://doi.org/10.1016/j.ocemod.2017.11.002, http://www.sciencedirect.com/science/article/pii/S1463500317301774

Hodson DLR, Sutton RT (2005) Atlantic Ocean forcing of North American and European summer climate. Science 309(5731):115. http://www.sciencemag.org/cgi/content/abstract/sci;309/5731/115

Johnson GC, Lyman JM (2020) Warming trends increasingly dominate global ocean. Nat Clim Change 10(8):757–761. https://doi.org/10.1038/s41558-020-0822-0

Keil P, Mauritsen T, Jungclaus J, Hedemann C, Olonscheck D, Ghosh R (2020) Multiple drivers of the North Atlantic warming hole. Nat Clim Change 10(7):667–671. https://doi.org/10.1038/s41558-020-0819-8

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):1–4. https://doi.org/10.1029/2005GL024233

Kushnir Y, Seager R, Ting M, Naik N, Nakamura J (2010) Mechanisms of tropical Atlantic SST influence on North American precipitation variability. J Clim 23(21):5610–5628. https://doi.org/10.1175/2010JCLI3172.1

Levitus S, Antonov JI, Boyer TP, Baranova OK, Garcia HE, Locarnini RA, Mishonov AV, Reagan JR, Seidov D, Yarosh ES, Zweng MM (2012) World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. Geophys Res Lett 39(10):1–5. https://doi.org/10.1029/2012GL051106

Lyman JM, Good SA, Gouretski VV, Ishii M, Johnson GC, Palmer MD, Smith DM, Willis JK (2010) Robust warming of the global upper ocean. Nature 465(7296):334–337. https://doi.org/10.1038/nature09043

Madec G, the NEMO team, (2016) NEMO ocean engine. Note du Pôle de modélisation de l'Institut Pierre-Simon Laplace. Tech Rep 27:1288–1619

Mann ME, Steinman BA, Brouillette DJ, Miller SK (2021) Multidecadal climate oscillations during the past millennium driven by volcanic forcing. Science 371(6533):1014–1019

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(24):8502–8523. https://doi.org/10.1175/JCLI-D-11-00633.1, https://journals.ametsoc.org/view/journals/clim/25/24/jcli-d-11-00633.1.xml

Maze G, Mercier H, Fablet R, Tandeo P, Radcenco ML, Lenca P, Feucher C, Goff CL (2017) Coherent heat patterns revealed by unsupervised classification of Argo temperature profiles in the North Atlantic Ocean. Prog Oceanogr 151:275–292. https://doi.org/10.1016/j.pocean.2016.12.008

Meehl GA, Goddard L, Murphy J, Stouffer RJ, Boer GJ, 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 skilful? Bull Am Meteorol Soc 90(10):1467–1486. https://doi.org/10.1175/2009BAMS2778.I

Meehl GA, Moss R, Taylor KE, Eyring V, Stouffer RJ, Bony S, Stevens B (2014) Climate model intercomparisons: preparing for the next phase. Eos 95(9):77–78. https://doi.org/10.1002/2014EO090001

Mignot J, García-Serrano J, Swingedouw D, Germe A, Nguyen S, Ortega P, Guilyardi E, Ray S (2016) Decadal prediction skill in the ocean with surface nudging in the IPSL-CM5A-LR climate model. Clim Dyn 47(3):1225–1246. https://doi.org/10.1007/s00382-015-2898-1

Milinski S, Maher N, Olonscheck D (2020) How large does a large ensemble need to be? Earth Syst Dyn 11(4):885–901. https://doi.org/10.5194/esd-11-885-2020, https://esd.copernicus.org/articles/11/885/2020/

Mogensen K, Alonso Balmaseda M, Weaver A (2012) The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4. Tech Memo 668(February):1–59

Ortega P, Mignot J, Swingedouw D, Sévellec F, Guilyardi E (2015) Reconciling two alternative mechanisms behind bi-decadal variability in the North Atlantic. Prog Oceanogr 137:237–249. https://doi.org/10.1016/j.pocean.2015.06.009

Piecuch CG, Ponte RM, Little CM, Buckley MW, Fukumori I (2017) Mechanisms underlying recent decadal changes in subpolar North Atlantic Ocean heat content. J Geophys Res Oceans 122:7181–7197. https://doi.org/10.1002/2017JC012845

Rahmstorf S, Box JE, Feulner G, Mann ME, Robinson A, Rutherford S, Schaffernicht EJ (2015) Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. Nat Clim Change 5(5):475–480. https://doi.org/10.1038/nclimate2554

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(12):4116–4134. https://doi.org/10.1175/JCLI-D-11-00443.1, https://journals.ametsoc.org/view/journals/clim/25/12/jcli-d-11-00443.1.xml

Robson J, Ortega P, Sutton R (2016) A reversal of climatic trends in the North Atlantic since 2005. Nat Geosci. https://doi.org/10.1038/ngeo2727

Robson J, Polo I, Hodson DLR, Stevens DP, Shaffrey LC (2018) Decadal prediction of the North Atlantic subpolar gyre in the HiGEM high-resolution climate model. Clim Dyn 50(3):921–937. https://doi.org/10.1007/s00382-017-3649-2

Rousset C, Vancoppenolle M, Madec G, Fichefet T, Flavoni S, Barthélemy A, Benshila R, Chanut J, Levy C, Masson S, Vivier F (2015) The Louvain-La-Neuve sea ice model LIM3.6: global and regional capabilities. Geosci Model Dev 8(10):2991–3005. https://doi.org/10.5194/gmd-8-2991-2015

Ruiz-Barradas A, Chafik L, Nigam S, Häkkinen S (2018) Recent subsurface North Atlantic cooling trend in context of Atlantic decadal-to-multidecadal variability. Tellus Ser A Dyn Meteorol Oceanogr 70(1):1–19. https://doi.org/10.1080/16000870.2018.1481688

Scaife AA, Smith D (2018) A signal-to-noise paradox in climate science. NPJ Clim Atmos Sci 1(1):28. https://doi.org/10.1038/s41612-018-0038-4

Schlesinger ME, Ramankutty N (1994) An oscillation in the global climate system of period 65–70 years. Nature 367(6465):723–726. https://doi.org/10.1038/367723a0

Schuckmann KV, Cheng L, Palmer MD, Hansen J, Tassone C, Aich V, Adusumilli S, Beltrami H, Boyer T, Cuesta-Valero FJ, Desbruyères D, Domingues C, García-García A, Gentine P, Gilson J, Gorfer M, Haimberger L, Ishii M, Johnson GC, Killick R, King BA, Kirchengast G, Kolodziejczyk N, Lyman J, Marzeion B, Mayer M, Monier M, Monselesan DP, Purkey S, Roemmich D, Schweiger A, Seneviratne SI, Shepherd A, Slater DA, Steiner AK, Straneo F, Timmermans ML, Wijffels SE (2020) Heat stored in the Earth system: where does the energy go? Earth Syst Sci Data 12(3):2013–2041. https://doi.org/10.5194/essd-12-2013-2020

Siegert S, Bellprat O, Ménégoz M, Stephenson DB, Doblas-Reyes FJ (2017) Detecting improvements in forecast correlation skill: statistical testing and power analysis. Mon Weather Rev 145(2):437–450. https://doi.org/10.1175/MWR-D-16-0037.1

Smith B, Wärlind D, Arneth A, Hickler T, Leadley P, Siltberg J, Zaehle S (2014) Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. Biogeosciences 11(7):2027–2054. https://doi.org/10.5194/bg-11-2027-2014

Smith DM, Scaife AA, Eade R, Athanasiadis P, Bellucci A, Bethke I, Bilbao R, Borchert LF, Caron LP, Counillon F, Danabasoglu G, Delworth T, Doblas-Reyes FJ, Dunstone NJ, Estella-Perez V, Flavoni S, Hermanson L, Keenlyside N, Kharin V, Kimoto M, Merryfield WJ, Mignot J, Mochizuki T, Modali K, Monerie PA, Müller WA, Nicolí D, Ortega P, Pankatz K, Pohlmann H, Robson J, Ruggieri P, Sospedra-Alfonso R, Swingedouw D, Wang Y, Wild S, Yeager S, Yang X, Zhang L (2020) North Atlantic climate far more predictable than models imply. Nature 583(7818):796–800. https://doi.org/10.1038/s41586-020-2525-0

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. Nat Commun 6(1):1–12

Thornalley DJ, Oppo DW, Ortega P, Robson JI, Brierley CM, Davis R, Hall IR, Moffa-Sanchez P, Rose NL, Spooner PT, Yashayaev I, Keigwin LD (2018) Anomalously weak Labrador Sea convection and Atlantic overturning during the past 150 years. Nature 556(7700):227–230. https://doi.org/10.1038/s41586-018-0007-4

Tietsche S, Balmaseda M, Zuo H, Roberts C, Mayer M, Ferranti L (2020) The importance of North Atlantic Ocean transports for seasonal forecasts. Clim Dyn 55:1995–2011. https://doi.org/10.1007/s00382-020-05364-6

Uppala SM, Kållberg PW, Simmons AJ, Andrae U, da Costa Bechtold V, 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 AC, van de Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PA, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The ERA-40 re-analysis. Q J R Meteorol Soc 131(612):2961–3012. https://doi.org/10.1256/qj.04.176

Wang G, Cheng L, Abraham J, Li C (2018) Consensuses and discrepancies of basin-scale ocean heat content changes in different ocean analyses. Clim Dyn 50(7–8):2471–2487. https://doi.org/10.1007/s00382-017-3751-5

Yang X, Rosati A, Zhang S, Delworth TL, Gudgel RG, Zhang R, Vecchi G, Anderson W, Chang YS, DelSole T, Dixon K, Msadek R, Stern WF, Wittenberg A, Zeng F (2013) A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system. J Clim 26(2):650–661. https://doi.org/10.1175/JCLI-D-12-00231.1, https://journals.ametsoc.org/view/journals/clim/26/2/jcli-d-12-00231.1.xml

Yeager SG, Robson JI (2017a) Recent progress in understanding and predicting Atlantic decadal climate variability. Curr Clim Change Rep 3(2):112–127. https://doi.org/10.1007/s40641-017-0064-z

Yeager SG, Robson JI (2017b) Recent progress in understanding and predicting Atlantic decadal climate variability. Curr Clim Change Rep 3(2):112–127. https://doi.org/10.1007/s40641-017-0064-z

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(15):5173–5189. https://doi.org/10.1175/JCLI-D-11-00595.1, https://journals.ametsoc.org/view/journals/clim/25/15/jcli-d-11-00595.1.xml

Yeager S, Kim W, Robson J (2016) What caused the Atlantic cold blob of 2015? US CLIVAR Var 14(2):24–31

Yeager SG, Danabasoglu G, Rosenbloom NA, Strand W, Bates SC, Meehl GA, Karspeck AR, Lindsay K, Long MC, Teng H, Lovenduski NS (2018) Predicting near-term changes in the earth system: a large ensemble of initialized decadal prediction simulations using the community earth system model. Bull Am Meteorol Soc 99(9):1867–1886. https://doi.org/10.1175/BAMS-D-17-0098.1

Zanna L, Khatiwala S, Gregory JM, Ison J, Heimbach P (2019) Global reconstruction of historical ocean heat storage and transport. Proc Natl Acad Sci USA 116(4):1126–1131. https://doi.org/10.1073/pnas.1808838115

Zhang R, Delworth TL (2006) Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. Geophys Res Lett 33(17):1–5. https://doi.org/10.1029/2006GL026267

Zuo H, Balmaseda MA, Tietsche S, Mogensen K, Mayer M (2019) The ECMWF operational ensemble reanalysis-analysis system for ocean and sea ice: a description of the system and assessment. Ocean Sci 15(3):779–808. https://doi.org/10.5194/os-15-779-2019

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.