https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Mechanisms Driving Subpolar North Atlantic Upper-Ocean Heat Content Predictability in CMIP6 Decadal Prediction Systems

Dylan C. Oldenburg[1], Stephen Yeager[1], Gokhan Danabasoglu[1], Isla Simpson[1], Who Kim[1], Jon Robson[2], Niamh O'Callaghan[3], and Panos Athanasiadis[4]

[1]National Science Foundation National Center for Atmospheric Research, Boulder, CO USA
[2]National Centre for Atmospheric Science, University of Reading, Reading, UK
[3]Department of Meteorology, University of Reading, Reading, UK
[4]CMCC, Bologna, Italy

Correspondence: Dylan C. Oldenburg (oldend@ucar.edu)

Abstract. Previous work has indicated that the subpolar North Atlantic (SPNA) Ocean exhibits particularly high decadal predictability, influenced by both external forcing and predictable internal variability as a result of large-scale ocean processes. Here we examine and compare a set of thirteen DP systems with a wide spread in skill at predicting SPNA upper ocean heat content (UOHC) at multiyear lead times, seeking to understand mechanisms driving differences in skill between the models. The well-documented SPNA warming event of the mid-1990s to early 2000s is of particular interest as it is the largest source of forecast error spread over the verification period across these systems. The mechanism in CESM2-DP is consistent with that previously documented for CESM1-DPLE, with initialized Labrador Sea Water thickness anomalies propagating eastwards towards the mid-Atlantic ridge over the course of the decadal simulations, generating strong ensemble mean signals in zonal sea surface height (SSH) gradient and associated anomalous northward flow at long leads. We find that less skilful systems tend to exhibit weaker coupling between the lower and upper limbs of mid-latitude, density-space AMOC together with a weaker correspondence between the SSH and abyssal thickness anomalies stemming from the Labrador Sea. The comparison of mechanisms at work in CMIP6 DCPP systems is therefore consistent with the hypothesis that skilful decadal prediction of SPNA UOHC is linked to predictable AMOC and SSH variations associated with LSW initialisation.

1 Introduction

Previous evaluation of decadal prediction systems from the Coupled Model Intercomparison Project phases 5 and 6 (CMIP5 and CMIP6) indicates that the subpolar North Atlantic (SPNA) is characterized by particularly high decadal predictability as a result of both external forcing and predictable internal variability linked to ocean dynamics (Yeager and Robson, 2017; Yeager, 2020; Borchert et al., 2021; Carmo-Costa et al., 2022), with CMIP6 systems showing even greater sea-surface temperature (SST) predictability in the subpolar gyre (SPG) than in CMIP5 (Borchert et al., 2021). Studies of these decadal prediction (DP) hindcasts, which are initialized from observation-based data and reanalyses, consistently highlight the importance of initialization in providing high SST predictive skill in the SPG (e.g., Müller et al. 2012; Kirtman et al. 2013; Doblas-Reyes et al. 2013; Smith et al. 2019; Carmo-Costa et al. 2022). Multidecadal variability in the Atlantic Meridional Overturning

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Circulation (AMOC) is a likely source of decadal ocean memory underpinning high skill in the SPNA (e.g., Griffies and Bryan 1997; Collins et al. 2006; Robson et al. 2012b; Yeager et al. 2012; Pohlmann et al. 2013; Msadek et al. 2014).

The mechanisms underpinning long-lasting predictability in the SPNA upper-ocean heat content (UOHC, 0-300m) in the Community Earth System Model Version 1 Decadal Prediction Large Ensemble (CESM1-DPLE; Yeager et al. 2018) have been studied extensively (Yeager et al., 2015; Yeager, 2020). In that system, high and long-lasting skill (i.e., up to 10 years after initialisation) in the SPNA UOHC has been linked to initialized Labrador Sea Water (LSW) thickness anomalies which propagate predictably through the interior ocean towards the mid-Atlantic ridge (MAR), where they persist (Yeager, 2020). Positive LSW thickness anomalies are associated with LSW production events in the Labrador Sea and Irminger-Iceland Basins, primarily via surface-forced water-mass transformation. Large positive LSW thickness anomalies in the western SPG lead to depressed sea-surface height (SSH) there, inducing buoyancy-driven SSH gradients and spinup of the barotropic gyre circulation (Yeager, 2015). An enhanced, mid-latitude zonal SSH gradient is associated with anomalously strong northward near-surface geostrophic flow and increased transport of warm, salty water into the subpolar gyre (Yeager, 2020). In CESM1-DPLE, AMOC in density space exhibits high potential predictability for both its upper and lower limbs even out to decadal lead times (Yeager, 2020). This predictability mechanism, whereby SPNA UOHC change is the end result of a decadal causal chain that links low frequency North Atlantic Oscillation (NAO) forcing and LSW production to subpolar SSH and time-delayed variations of the AMOC lower and upper limbs, has also been identified as the dominant mechanism of internal decadal AMOC variability in a high-resolution configuration of CESM1 that exhibits relatively low bias in the subpolar North Atlantic (Yeager et al., 2021). Given its prominence and detailed investigation in simulations using CESM1, we will refer to this causal sequence as the "CESM1" decadal variability/predictability mechanism.

However, the generality of the CESM1 predictability mechanism remains unclear. Decadal prediction systems run as part of the CMIP6 DCPP intercomparison (Boer et al. 2016) exhibit a wide range of skill for SPNA upper 700m ocean heat content (Carmo-Costa et al., 2025). Multi-model analyses have indicated that model biases in the Labrador Sea upper-ocean stratification and temperature and salinity co-variability can lead to decreased decadal North Atlantic variability and predictability (Menary et al., 2015; Menary and Hermanson, 2018; Ortega et al., 2021; Kim et al., 2023). Large model drifts and initialization shocks can also impact SPNA skill in hindcast simulations (Kröger et al., 2018; Polkova et al., 2019; Bilbao et al., 2021; Polkova et al., 2023), though it is still not clear whether initialization methods designed to minimize drift necessarily produce greater skill compared to full-field initialization (Smith et al., 2013; Hazeleger et al., 2013; Volpi et al., 2017; Kröger et al., 2018; Polkova et al., 2023).

The NAO, which has been shown to play a key role in driving variability in AMOC as well as North Atlantic SSTs (Eden and Jung, 2001; Danabasoglu, 2008; Danabasoglu et al., 2012; Mecking et al., 2015; Delworth et al., 2016; Delworth and Zeng, 2016; Kim et al., 2018, 2020, 2024; Patrizio et al., 2025), appears to exhibit decadal predictability in large ensemble DP systems (Athanasiadis et al., 2020; Smith et al., 2020), albeit with much too weak amplitude. The NAO skill is associated with the positive feedback between SPNA SSTs and NAO (Patrizio et al., 2025). Kim et al. (2024) investigated the behaviour of three different coupled models subject to observed NAO forcing, and the response was found to be similar to the CESM1 predictability mechanism discussed above, with persistent, positive NAO inducing anomalous dense water production in the

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Labrador Sea and Irminger-Iceland Basins along with a densification of the AMOC lower limb at subpolar latitudes. In each of the models examined, deep density anomalies propagate southward to intergyre latitudes with a several year time delay, whereupon they induce a zonal surface pressure gradient, enhanced North Atlantic Current flow, and increased ocean heat transport and UOHC. Even if DP systems cannot predict NAO signals of realistic amplitude, which would help to extend the predictability horizon of AMOC and UOHC, we would expect them to show some baseline predictability associated with the strong imprint of NAO on the historical ocean state used for initialization.

We hypothesize that skilful decadal prediction systems will, in general, exhibit a qualitatively similar mechanism to that first identified in CESM1 simulations, and that its presence or absence may help to explain the diversity of prediction system performance in the SPNA. After first quantifying the SPNA UOHC skill and its lead time dependence in 13 DCPP systems, we classify the models based on their skill over lead years 6-10, because we are interested in skill that persists towards the end of the simulations. We create two subsets - one with the three highest skill models, and one with the two lowest skill models. These subsets are the primary focus for our more in-depth analysis. The well-documented 1990s-2000s SPNA warming event (e.g., Holliday et al. 2008; Marsh et al. 2008; Sarafanov et al. 2008; Grist et al. 2010; Reverdin 2010; Robson et al. 2012b, a) is a key case study that reveals stark differences between forecast evolution in high versus low skill systems. We analyse fields that reflect the thermohaline gyre mechanism of CESM1, including: UOHC, AMOC in density space, the barotropic streamfunction (BSF), LSW thickness, and SSH.

2 Methods

2.1 CESM and CMIP6 Decadal Prediction Systems

We use data from the CESM1 and CESM2 Decadal Prediction systems (CESM1-DPLE and CESM2-DP), as well from all available CMIP6 DCPP 10-year hindcast simulations (Boer et al., 2016). The DCPP systems include: CanESM5, HadGEM3-GC31-MM, MIROC6, NorCPM1, CMCC-CM2-SR5, EC-Earth3, IPSL-CM6A-LR, BCC-CSM2-MR, FGOALS-f3-L and MPI-ESM1-2-HR, in addition to IPSL-CM6-ESM, whose data was obtained through personal communication with the modelling team. All hindcasts were initialized each year from different reanalysis and observation-based products (see Table 1 for further details).

2.2 Observation-based benchmarks

For our observational benchmarks, we primarily use upper-ocean potential temperature and absolute salinity from the Hadley Centre's EN.4.2.2 (Good et al., 2013). For our analysis of AMOC, LSW thickness, the BSF, upper ocean meridional velocity, and SSH, observational fields are either non-existent or too sparse (in space, time, or both) to permit robust skill evaluation. In order to provide a plausible historical benchmark for the decadal variability in these fields, we use data from low (1°) and high-resolution (0.1°) CESM2-based forced ocean–sea-ice (FOSI) simulations, forced by atmospheric reanalysis, similar to those used to initialized CESM DP systems. Such FOSI simulations are constrained by observation-based surface fields and have been shown to yield good comparisons to available ocean measurements in many different ocean models (Danabasoglu et al.,

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Model (Reference)	Initialisation procedure	Atmos., ocean resolution & model	Init. years	Num. ens. mem.
CESM1-DP (Yeager et al., 2018)	OCN/ICE: FOSI. ATM: model restart	∼1° (CAM5) ∼1° (POP2)	1954–2017	40
CESM2-DP (Yeager et al., 2022)	OCN/ICE: CESM2 FOSI; ATM: JRA55	∼1° (CAM6) ∼1° (POP2)	1958–2023	20
CanESM5 (Sospedra-Alfonso et al., 2021)	ATM: ERA5. OCN: ORAS5. SST: OISST and ERSSTv3	∼2.8° (CanAM5) ∼1° (CanNEMO)	1960–2016	10
HadGEM3-GC31-MM (Dunstone et al., 2016; Choi et al., 2026)	Full-field	∼0.25° (NIMOv3.6)	1960–2018	10
MIROC6 (Kataoka et al., 2020)	OCN: objective analysis. ICE/ATM: reanalysis	∼1.4° ∼1° (COCO4.9)	1960–2021	10
NorCPM1 (Bethke et al., 2021)	Anomaly assimilation of SST, T/S	∼2° (CAM4) ∼1° (isopycnic, MICOM)	1960–2018	10
CMCC-CM2-SR5 (Nicolì et al., 2023)	Full-field, surface nudging & bias correction	∼1° (CAM5) ∼1° (NEMOv3.6)	1960–2019	40
IPSL-CM6-ESM (No reference)	SST/SSS anomaly nudging	1.3°×2.5° (LMDZ) ∼1° (NEMO, ORCA1)	1980–2025	10
EC-Earth3 (Bilbao et al., 2021)	Full-field T/S, surface nudging	∼80 km ∼1° (NEMO3.6, ORCA1)	1960–2018	10
IPSL-CM6A-LR (Estella-Perez et al., 2020)	SST/SSS anomaly nudging	1.3°×2.5° (LMDZ) ∼1° (NEMO, ORCA1)	1960–2016	10
BCC-CSM2-MR (Wu et al., 2019; Xin et al., 2019)	SODA nudging experiments	∼1.1° ∼1° (1/3° in tropics; MOM4)	1961–2014	8
FGOALS-f3-L (Hu et al., 2023)	SST, T/S anomaly assimilation	∼1° (FAMIL2) ∼1° (LICOM3)	1960–2016	6
MPI-ESM1-2-HR (Müller et al., 2018; Mauritsen et al., 2019)	Institute's protocol	T127 (∼100 km) 0.4° (TP04, MPIOM)	1960–2018	5

Table 1. Description of the decadal prediction systems used in this study. Column 1 lists the model name alongside the corresponding references. Column 2 indicates the initialisation procedure. Column 3 shows the atmospheric and ocean resolutions alongside their model names. Column 4 indicates the initialisation years. The final column shows the number of ensemble members for each model. More details about the initialisation procedures can be found in the Appendix.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

2016). The FOSI simulations were run in a configuration designed to isolate the impact of interannually-varying buoyancy forcing (using wind fields that vary only subannually) in order to highlight decadal SPNA variability, as shown in previous studies (Biastoch et al., 2008; Yeager and Danabasoglu, 2014). Hereafter, these simulations are referred to as FOSI-LR-B and FOSI-HR-B. Figures that compare hindcasts to FOSI-LR and FOSI-HR, which include interannual wind forcing, can be found in the supplementary information (Figs. S6-S8). Here we focus our comparison on the FOSI-B simulations, because we want to evaluate the lower-frequency (buoyancy-driven) variability in an ensemble of simulations. We further consider data from the last cycle of CMCC-CM2-SR5's offline ocean model simulation contributed to the Ocean Model Intercomparison Project 2 (OMIP2, Tsujino et al. 2020), in order to include a non-CESM model which is also the model with the third highest DP skill for SPNA UOHC for lead years 6-10. Although we could theoretically use the initialization ('assim') datasets for each DP system as our observation-based benchmark, these datasets were not available for all of the systems examined.

2.3 Methods for calculations

Before analysing model results, we first remove the lead-dependent climatology (i.e., drift; Boer et al. 2016) as well as the historical linear trend from observed and hindcast time series. The SPNA UOHC is defined here as the average temperature over the upper 300m of the water column. We define the predictive skill as the anomaly correlation coefficient (ACC) between each model's ensemble mean hindcast time series and the corresponding observation-based (EN4 in this case) time series.

AMOC in density coordinates is calculated for all models using the following equation:

$$\psi(y,\sigma,t) = -\int_{\sigma_b}^{\sigma}\int_{x_w}^{x_e} v(x,y,\sigma,t)\,\frac{\partial z}{\partial \sigma}\,\mathrm{d}x\,\mathrm{d}\sigma \tag{1}$$

where v is the monthly-mean meridional velocity, positive northward; t is time; z is the vertical coordinate, positive upward; x is the zonal coordinate, positive eastward; y is the meridional coordinate, positive northward; and σ is the potential density, computed using monthly-mean potential temperature and absolute salinity. The subscripts e and w denote the eastern and western limits of the zonal integral, respectively; the subscript b denotes the bottom (or the highest density), reflecting that the integration is performed from the bottom up; and $\partial z/\partial \sigma$ is the isopycnal thickness. For this calculation we take into account the different staggered grids used in each ocean model.

The BSF is either taken from model output (if available) or estimated by integrating the y-ward velocity v eastward from the western boundary of the Atlantic basin. Though the y-ward velocity is not entirely meridional in the North Atlantic, this method is still able to largely reproduce monthly-mean model BSF output where available (not shown). SSH is obtained from monthly-mean output. Before computing the large-scale SSH gradients, we apply a 90-km smoothing to all models to remove small scale gradients. 90 km was chosen here it approximates the spacing at 45°N in the low-resolution models.

The LSW thickness is calculated using the same method from Yeager 2020, i.e., by calculating the thickness of the layer corresponding to the density class based on the initialisation dataset's water-mass formation (WMF) anomaly from 1991-1995 relative to climatology. Plots depicting the WMF for the models can be found in Figs. S1-S5. For CanESM5, because the

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

freshwater flux was unavailable for the initialisation dataset, our WMF calculation is based on the heat flux component. An analysis of the two components in the hindcast simulations determined that the freshwater component does not contribute substantially to the WMF (not shown). WMF plots for each model can be seen in Figs. S1-S5. The pathways for the LSW layers and where they sit in the various initialisation datasets used for these models can be found in Fig. S6.

3 Predictive skill of the subpolar North Atlantic upper ocean heat content

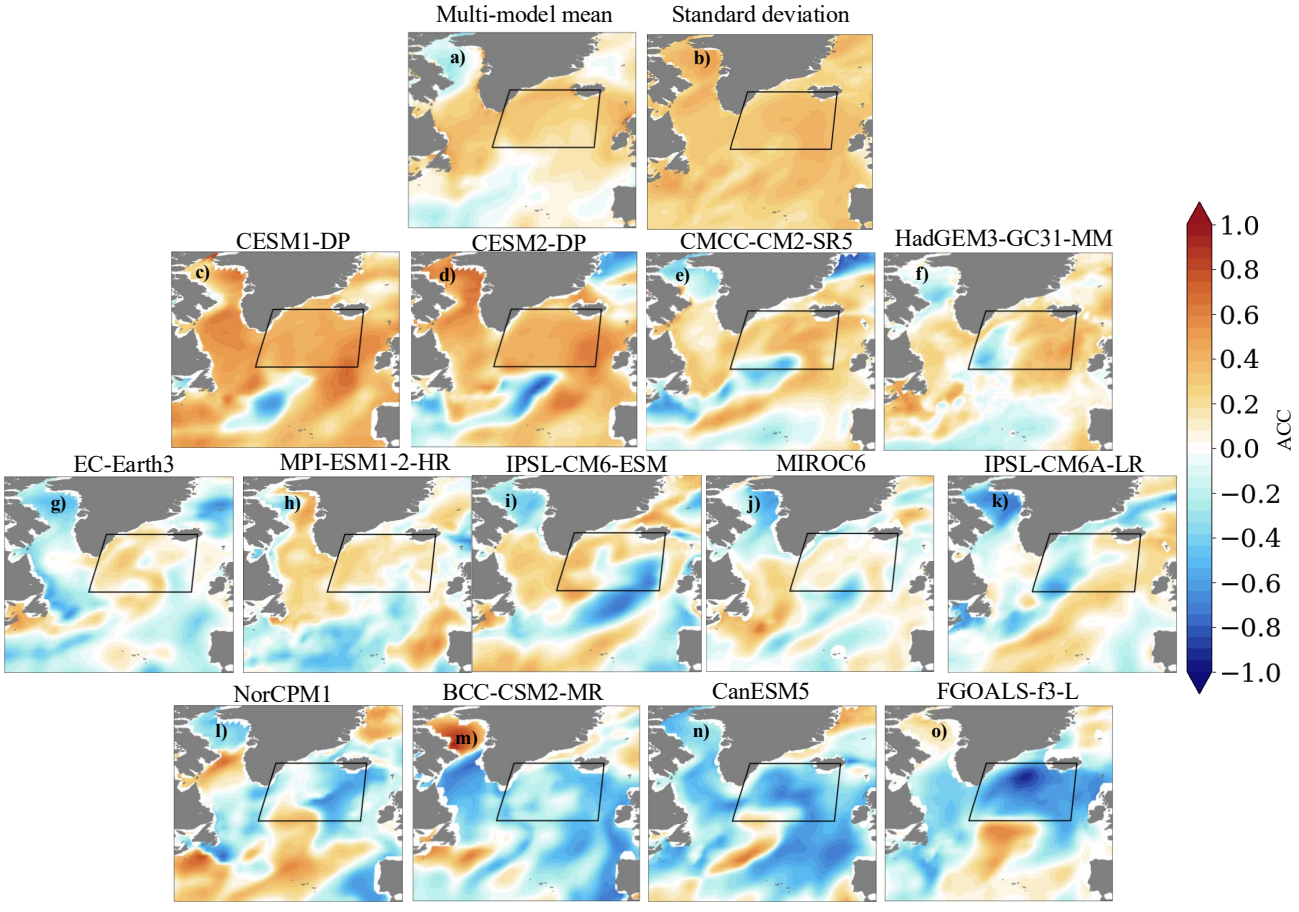


Figure 1. Ensemble-mean subpolar North Atlantic upper ocean heat content (0-300m) predictive skill over forecast years 6-10 for CMIP6 decadal prediction systems. **a)** Multi-model mean of the anomaly correlation coefficient (ACC) computed with EN4. **b)** Standard deviation of the ACC across models. **c-o)** Deviation of ACC from multi-model mean. The boxes here represent what we define as the SPNA region, i.e. 43°W-15°W, 52°N-65°N.

Maps of the SPNA UOHC ACC skill for each individual model's DCPP relative to EN4 data for forecast years 6-10 are presented in Fig. 1. CMIP6 DP systems exhibit a wide range of skill in the North Atlantic, with the northern SPG showing

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

particularly high prediction skill across models (Fig. 1a). CESM1-DP and CESM2-DP show the highest skill (Fig. 1c-d), and NorCPM1, BCC-CSM2-MR, CanESM5, and FGOALS-f3-L exhibit the lowest skill (Fig. 1m-o). Both CMCC-CM2-SR5 and HadGEM3-GC31-MM stand out for good SPNA performance, as models with high skill in the far eastern SPG, but much lower skill farther west (Fig. 1f). The black outlined boxes shown in Fig. 1 represent the area that we define as the SPNA region. We chose this particular box because it is an area where the ACC exhibits particularly high inter-model variability, and is also a region where advection is expected to play an important role in driving variability (Moat et al., 2019; Yeager, 2020; Ortega et al., 2021). The SPNA ACC in the models does change somewhat if the box is shifted slightly, and hence a quantitative analysis of averaged skill may depend on the choice of box. However, the order of the models in terms of their skill remains unchanged (not shown).

Plotting the skill averaged across the SPNA box as a function of forecast year highlights the persistently high skill of CESM1-DP and CESM2-DP, as well as a steep decline over time in skill across most models between year 1 and years 2-6 (Fig. 2a). All models with the exception of IPSL-CM6A-LR and BCC-CSM2-MR show decent skill (ACC$\geq$0.5) at the first month, suggesting relatively accurate initialisation of UOHC. NorCPM1, FGOALS-f3-L and CanESM5 in particular stand out in that they initially capture the SPNA UOHC variability very well before rapidly losing their skill at later forecast years (Fig. 2a). Over forecast years 6-10, which we focus on in our analysis, CESM1-DP, CESM2-DP, and CMCC-CM2-SR5 show the highest skill, with ACC values equal to 0.54, 0.5 and 0.37, respectively. NorCPM1, CanESM5, FGOALS-f3-L, and BCC-CSM2-MR on the other hand show the lowest skill, with ACCs equal to 0.07, -0.08, -0.19 and -0.03, respectively. BCC-CSM2-MR, however, shows relatively poor skill (compared to the multi-model mean) even at early forecast years; hence, the poor skill is likely due to initialization issues. Because of this, we do not focus on it in the rest of our analysis. IPSL-CM6-ESM and IPSL-CM6A-LR also show fairly poor skill (i.e., significantly below the multi-model mean), with IPSL showing weak ACC at all forecast times (Fig. 2a).

It is interesting to also consider the upper-ocean salinity skill in the SPNA, given that variability in upper ocean salinity and temperature are typically linked in this region via ocean advection which carries warm, salty water into the SPG. The skill spread for SPNA upper ocean salinity content (UOSC) at decadal lead times largely reflects that for UOHC, such that models that show high L6-10 skill for UOHC also tend to exhibit high UOSC skill (and vice versa; Fig. 2b). A notable exception is CMCC-CM2-SR5 which has a considerably lower skill for UOSC than for UOHC. The distribution of scores at month 1 for UOSC is also lower than for UOHC, likely related to the challenge of accurately reconstructing historical ocean salinity from sparse salinity observations. The high correspondence between UOSC skill and UOHC skill across systems suggests that the skill level for both is related to the representation of ocean advective processes.

In the rest of this analysis, we refer to the top three highest skill models (CESM1-DP, CESM2-DP, and CMCC-CM2-SR5) as "high skill" models, and the bottom two models as "low skill" models. Although CMCC-CM2-SR5 is less skilful than CESM-based DP systems, it is the third highest skill model, and it is included to make sure conclusions drawn from the high-skill group do not just depend on one modelling system. For our bottom two models, we use NorCPM1 and CanESM5, because these are both low-skill models with the data available to calculate the LSW thickness. The 95% confidence intervals for the

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

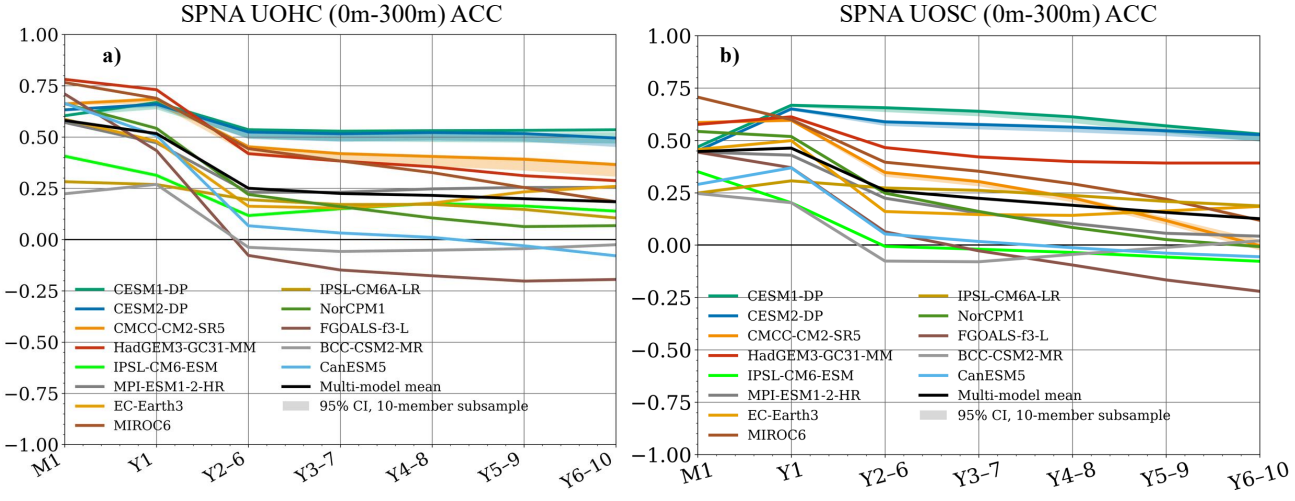


Figure 2. Ensemble-mean anomaly correlation coefficient predictive skill evaluated against EN4 averaged over the SPNA region at the first month after initialisation (M1) and across forecast years (Y) for **a)** upper ocean heat content (UOHC) and **b)** upper ocean salinity content (UOSC). Shading here for CESM1-DP, CESM2-DP and CMCC-CM2-SR5 represents the 95% confidence interval for the 10-member subsampled skill. All variables are detrended before the calculation of skill.

10-member subsampled skills of CESM1-DP, CESM2-DP and CMCC-CM2-SR5 remain well above the other models, and hence the higher skill in these models is not simply due to a larger ensemble size than other models.

We next look at the full field representation of the SPNA UOHC without removing the lead-dependent climatology in the first month after initialization, when we expect most models to align well with EN4 observation-based data, regardless of their skill at later forecast years (Fig. 3). Since the focus of the rest of our analysis is on the high and low skill models, we include those here. We can see from the raw time series that models vary significantly in their representation of SPNA UOHC climatology

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

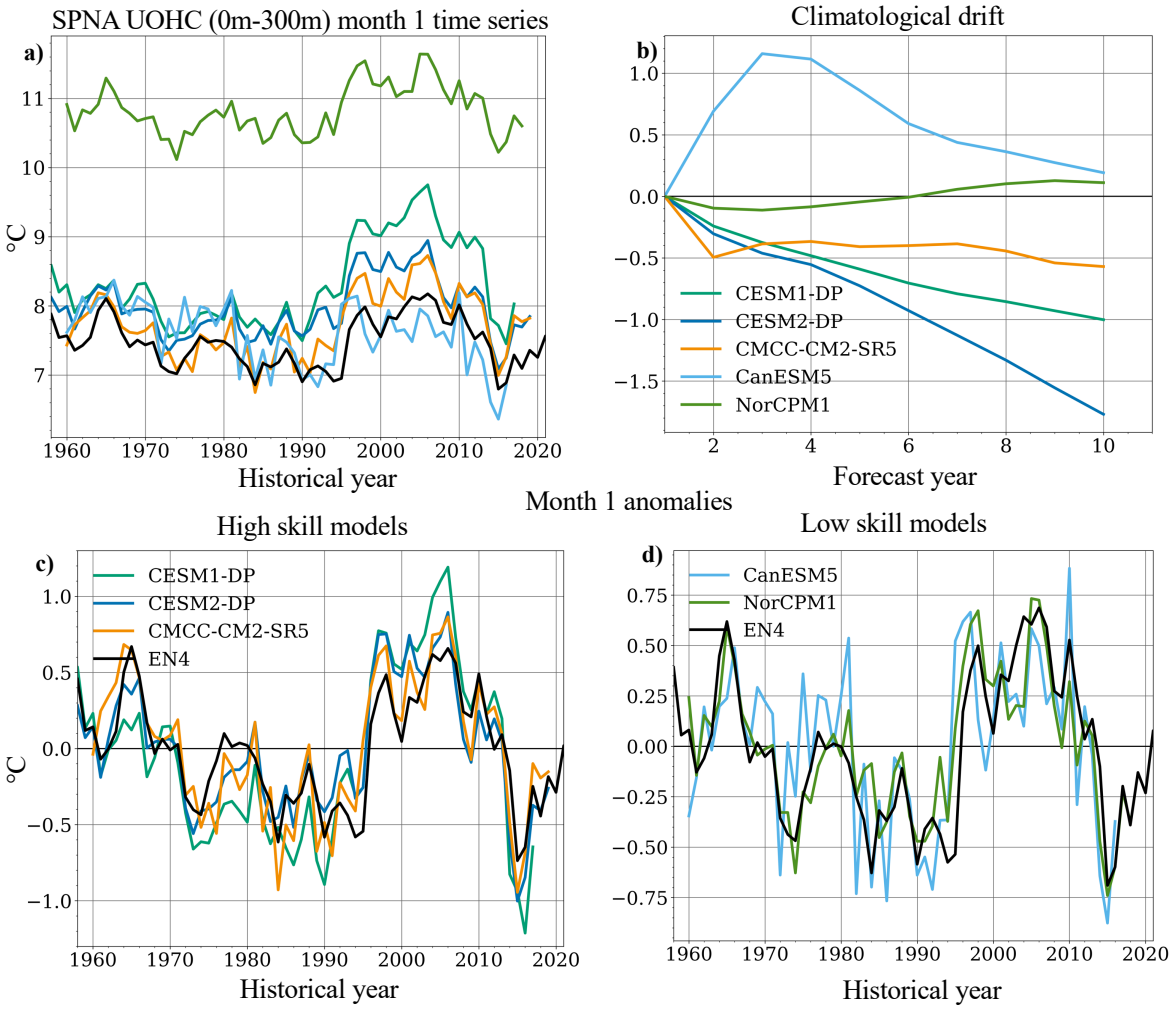


Figure 3. Top row: **a)** Raw ensemble-mean SPNA UOHC (0m-300m) time series for the highest and lowest skill models for the first month after initialisation (defined as forecast month 1) plotted alongside EN4 January data. **b)** Ensemble-mean climatological drift for the different models over forecast years. Bottom row: Ensemble-mean detrended, de-drifted time series of the initial SPNA UOHC for the **c)** highest skill models and **d)** lowest skill models.

early in the forecasts (Fig. 3a), despite the fact that all of the models capture the variability quite well (Fig. 3c, d). We also see that the highest skill models included here, i.e., CESM1-DP and CESM2-DP, substantially overestimate UOHC compared to EN4 (Fig. 3a), indicating that realistic variability can be captured even if a model's climatology is poor. This is further explored by examining the change in climatology with forecast year (i.e., drift). Several models show a strong drift in UOHC climatology over decadal forecast leads, but there is no indication that higher skill models exhibit reduced drift compared to

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

low skill models (Fig. 3b). Indeed, CESM2-DP exhibits the largest drift of all the systems examined, despite being one of the highest skill models (Fig. 3b).

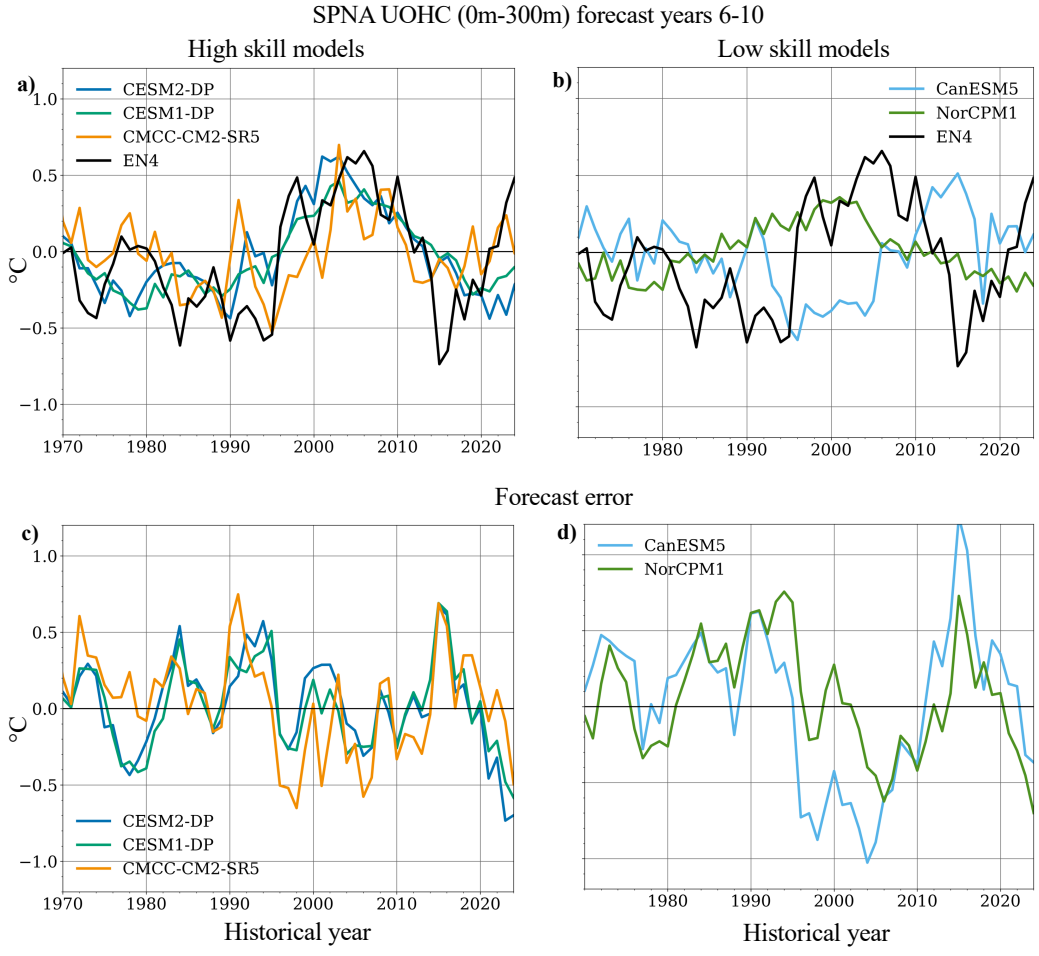


Figure 4. Top row: Historical time series of the ensemble-mean SPNA UOHC averaged over forecast years 6-10 plotted alongside EN4 annual-mean values for **a)** High skill and **b)** low skill models. Bottom row: Forecast errors for **c)** high skill and **d)** low skill models.

Figure 4 shows the time series of the SPNA UOHC at forecast years 6-10, to reveal when the UOHC prediction breaks down in low skill models. CanESM5 diverges greatly from observation-based data (and higher skill models) starting in the mid-1990s, it fails to capture the rapid warming event in the SPNA, and instead experience warming later in the mid-2000s (Fig. 4b, d). NorCPM1, on the other hand, is too warm in the early 1990s and cools too early in the 2000s compared to EN4 (Fig. 4b, d). The high skill models, on the other hand, capture the warming and subsequent cooling quite well (Fig. 4a, c). These

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

results motivate a focus on the 1990s warming event as an important case study for understanding the diversity in UOHC skill across DCPP systems.

4 Time evolution of AMOC, LSW thickness, SSH, and meridional velocity

The time evolution of mid-latitude AMOC in the high skill models as well as in the three benchmark FOSI simulations can be seen in Fig. 5. Here we focus on time evolution at 45°N; full contour map plots for the 1990s event can be seen in the supplementary information (Figs. S7-S11). All of the benchmark simulations show a broad strengthening of AMOC in the 1990s, with high density (lower limb, σ_2>36.5 kg m^3) anomalies leading low density (upper limb, σ_2<35.7 kg m^3) anomalies by several years (Fig. 5a-c). In the case of FOSI-LR-B and CMCC-CM2-SR5, this broadly strong AMOC extends into the mid-2000s (Fig. 5a, c), while in FOSI-HR-B it rapidly dissipates after 2001 with the exception of very high densities ($\sigma_2 \geq 36.9$; Fig. 5b).

In both CESM1-DP and CESM2-DP at forecast year 1, there is also a broad strengthening in the 1990s, starting with an initial strengthening of the lower limb which then expands to the upper limb as well (Fig. 5d-e). The upper limb anomalies in CESM1-DP and CESM2-DP show good consistency across forecast lead times from year 1 to years 6-10, all showing a transition to positive anomalies in the mid 1990s (Fig. 5d, e, g, h, j, k). This contrasts with the notable lead dependence of lower limb anomalies which transition to positive in the mid 1980s at short leads (forecast year 1) but after 1990 at later leads (forecast years 6-10). This is consistent with the poor long-lead skill at predicting the anomalous LSW formation that directly impacts lower limb variability, but good skill at predicting the delayed surface expression of initialized LSW anomalies, as discussed in Yeager (2020). The CMCC-CM2-SR5 system also shows consistent predictions of decadal variability in the upper limb over lead times, but not the lower limb, which suggests the presence of a similar LSW-related mechanism controlling near-surface flow variability (Fig. 5f, i, l). In CMCC-CM2-SR5, however, the AMOC in year 1 shows much stronger interannual variability than in the CESM systems which show stronger decadal variability (Fig. 5f). This suggests that the mid-latitude AMOC in CESM DP systems is more buoyancy than wind dominated (Biastoch et al., 2008; Yeager and Danabasoglu, 2014), thus more strongly controlled by the LSW variability, than in CMCC-CM2-SR5. Still, all three high skill systems exhibit a similar decadal modulation of AMOC upper limb variability at multiyear leads that appears consistent with their years 6-10 predictions of UOHC change (Fig. 4a).

The propagation of abyssal LSW thickness and SSH anomalies towards the MAR (near 30°W south of 50°N, see Fig. S1a) can be seen by looking at their time evolutions at 45°N (Fig. 6). In both FOSI-LR-B and CMCC-CM2-SR5 OMIP2, there are strong LSW production events in the late 1980s and 1990s, and there is also a clear connection between depressed SSH and positive LSW thickness anomalies which appear in the late 1990s west of the MAR ($\sim$ 29°W, Fig. 6a, c), though this is less clear in FOSI-HR-B, where SSH is much noisier (Fig. 6b).

In CESM1-DP and CESM2-DP (Fig. 6d, e), we also see strong LSW production events associated with depressed sea level west of the MAR at L1, though in CESM2-DP LSW thickness anomalies reach much farther east of the MAR than in CESM1-DP. At L3-7 and L6-10, the predicted LSW thickness anomalies are much weaker along the shelf of Newfoundland

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

(far western longitudes), which is consistent with a lack of prediction skill for the production events which occur upstream in the Labrador and Irminger Seas (Yeager, 2020). However, strong LSW thickness (and opposite signed SSH) signals are still evident in the interior in these systems, which implies considerable multiyear memory of the state of LSW at initialization (Fig. 6g, h, j, k), consistent with the CESM1 predictability mechanism. In CMCC-CM2-SR5, the SSH field is much noisier, and a connection between SSH and abyssal thickness anomalies is much less clear, even at L1 (Fig. 6f). Again, this suggests that wind forcing exerts a stronger control on SPNA SSH than buoyancy forcing in that system. In CMCC-CM2-SR5, the correspondence between SSH and LSW thickness is not as strong as in the CESM systems (Fig. 6i, l).

Positive zonal SSH gradient anomalies imply northward surface flow anomalies through geostrophy. The SSH zonal gradient, dSSH/dx, is shown alongside the upper 300m-average meridional velocity at 45°N in Fig. 7. As expected, we see that positive SSH gradients are associated with anomalous northward flow near the MAR in the benchmark simulations, although there is much greater variability in eddy-resolving FOSI-HR-B than in the other two simulations, and in CMCC the strong SSH gradient is farther west than in the FOSI simulations (Fig. 7a-c). In both CESM1-DP and CESM2-DP, a similar SSH gradient appears at a similar longitude, with this signal persisting even at forecast years 6-10 (Fig. 7d, e, g, h, j, k). In CMCC-CM2-SR5, however, the strongest SSH gradient and northward flow anomalies are located farther west at year 1 than in the CESM systems, and cover a longer time period (Fig. 7f). At later lead times and with multiyear averaging, the positive surface flow anomaly in the CMCC system shifts east, yielding a closer match to the CESM systems, but it is delayed, emerging only around 2000 (Fig. 7i, l). This delay relative to CESM systems and the benchmarks is consistent with the delayed spinup of the upper limb (Fig. 5) and the delayed increase in SPNA UOHC (Fig. 4) seen in CMCC-CM2-SR5.

It is also useful to look at the correlations between the smoothed LSW thickness and SSH in order to further ascertain where there is a close link between the two. These correlations can be seen in Fig. 8. Here we can see that in CESM1-DP at both early and late lead times, LSW thickness anomalies are anticorrelated with SSH west of the MAR, with a sign change on the crest of the ridge (Fig. 8a, d, g). Hence, a thicker LSW layer west of the ridge correlates with a depressed SSH there, while the reverse is true east of the crest. In other models, the same dipole exists, but is shifted to different longitudes. In the case of CESM2-DP, it sits east of the crest, between 42°N and 50°N (Fig. 8b, e, h). In CMCC-CM2-SR5, the sign shift occurs near the ridge crest, but the dipole pattern is not as pronounced as in CESM1 and CESM2 (Fig. 8f, i). In its initialisation dataset, there is no indication of any dipole pattern (Fig. 8c).

We now focus on the low skill models which show high skill at early forecast years but whose performance rapidly degrades with lead time. The AMOC changes at 45°N can be seen in Fig. 9. In CanESM5, at forecast year 1, there are long periods of time throughout much of the 1980s and 1990s of very strong AMOC with little to no lag between the strengthening of the lower limb (representing a densification of the deep water) and a broader AMOC strengthening (Fig. 9e). The lack of a lagged relationship between lower and upper limb AMOC anomalies suggests that the CESM1 mechanism is not operating in this model. At later forecast years, there is no broad AMOC strengthening at all in the 1990s, AMOC actually weakens during that period except in the lower limb (Fig. 9g, i), which is likely responsible for the cooling in the SPNA that occurs during that period (Fig. 9). In NorCPM1 at L1, there is a notable strengthening of the lower limb of AMOC in the mid-1980s, followed by a broad strengthening throughout the 1990s, then a weakening in the early 2000s until the end of the simulations (Fig. 9d). At

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

L3-7 and L6-10, the broad strengthening still occurs throughout the 1990s but AMOC weakens sooner compared to L1 (Fig. 9f, h).

Unlike in the high-skill models, CanESM5 shows no eastward propagation of abyssal thickness/density anomalies at 45°N during the 1990s (Fig. 10e, g, i). In that model, there is a steric-induced northward flow during the early 2000s portion of the SPNA warming event at forecast year 1 (Fig. 12e), but at later forecast years there is actually anomalous southward flow between 40°W and 30°W in the late 1990s until about 2003 (Fig. 11g, i). In NorCPM1, the SSH field shows some evidence of a decadal, steric merchanism at L1 (Fig. 10d), though the largest SSH gradients are located west of the MAR (Fig. 11d). At L3-7 and L6-10, there is no indication of a connection between SSH and LSW thickness, and SSH gradients near the MAR drive southward flow, with northward flow occurring west of there (Fig. 10f, h; Fig. 11f, h).

LSW thickness-SSH correlation plots reveal that in NorCPM1, the sign shift between positive and negative correlations occurs significantly west of the MAR (Fig. 12a, c, e), implying that an LSW thickening event in the North Atlantic (as occurs in 1990s, Fig. 10d, f, h; Fig. S11) would drive depressed SSH west of about 35°W and enhanced SSH east of there, leading to enhanced northward geostrophic flow towards the western part of the SPNA box, and towards the LS. By L6-10, the geostrophic flow driven by such an event would carry warm water too far west to reach much of the SPNA (Fig. 12c). In the case of CanESM5, there is no indication of any dipole pattern; in fact, the Irminger-Iceland Basins and regions near the MAR do not host any LSW, even in the initialisation dataset (Fig. 12b, d, f). At 45°N, unlike in other models, the LSW layer in the assim run remains trapped below 3000m, and does not reach the Irminger-Iceland Basins at all (Fig. S6). Hence LSW layers in the different regions are much thinner than in other models, and show no significant change in the west MAR region during the 1990s event, unlike in other models (Fig. S6j).

We further create plots similar to that in Fig. 10 of (Yeager, 2020) for these models, to determine whether LSW thickness and the MAR-averaged upper ocean geostrophic meridional velocities (v_{MAR}) are well predicted relative to the initialisation runs and whether there are clear lead-lag relationships between the LSW thickness and v_{MAR}. Here, in the case of CESM2-DP, CMCC-CM2-SR5 and NorCPM1, we shift the MAR box to where the dipole in Figs. 8 and 12 is located in each model. In the case of CanESM5, since there is no dipole pattern either west nor east of the MAR, this cannot be done.

We see here that in both CESM1-DP and CESM2-DP, there is a clear lead-lag relationship between the LSW thickness at L1-5 and v_{MAR} at L5-9 (Fig. 13a-f). Both models' ΔLSW and v_{MAR} show close alignment with their respective FOSI runs as well (Fig. 13 b, c, e, f). However, in CMCC-CM2-SR5, ΔLSW is not nearly as well aligned with its initialisation dataset, showing much more noise than the other models (Fig. 13h). In the case of v_{MAR}, CMCC's hindcast does not show alignment with its initialisation dataset at all, and the lead-lag relationship with the preceding ΔLSW is not at all clear. v_{MAR} does not become positive until the early 2000s, several years after the SPNA starts to warm in the mid-1990s. This discrepancy in CMCC suggests that the SPNA UOHC changes at these lead times are not primarily driven by v_{MAR} variability, and may instead reflect other processes, such as advection of temperature anomalies by the mean circulation or the response to external forcing.

The same figures for the low-skill models can be seen in Fig. 14. For NorCPM1, because the dipole is shifted significantly west of the MAR (where geostrophic velocity would drive heat transport only to the westernmost part of the SPNA and to the

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

LS), we include figures for both the CESM1 MAR box as well as the dipole box. In NorCPM1, if the original MAR box is used, LSW anomalies are very well predicted at L1-5 relative to assim (Fig. 14b), but v_{MAR} variability is weak and shows little connection to the preceding LSW thickness changes (Fig. 14c). If the dipole box is used, LSW anomalies are only well predicted after about 1985, and v_{MAR} values are positive throughout the late 20th century all the way until about 2004 (Fig. 14e, f), after which they decline until the end of the simulations. There is similarly no clear connection between LSW anomalies and v_{MAR}. In CanESM5, LSW anomalies and v_{MAR} are completely out of step with assim and satellite observations (Fig. 14h-i).

5 Conclusions

In this study, we have computed the anomaly correlation coefficient predictive skill of the subpolar North Atlantic upper ocean heat content in 13 different decadal prediction systems, which allowed us to sort them by their skill. Here we have labelled the three models with the highest skill in the SPNA as 'high-skill' and the two with the lowest skill as 'low-skill'. For the rest of the analysis, we have focussed on determining whether the predictability mechanism found in CESM1-DP is also present in other models with high SPNA UOHC skill by analysing the historical time evolution of AMOC, SSH, Labrador Sea Water thickness, UOHC and the barotropic streamfunction.

Our results indicate that across the best performing initialized decadal prediction systems available from DCPP of CMIP6 for SPNA UOHC, predictable variability at later forecast years is consistently associated with a "buoyancy-driven gyre" mechanism (Yeager and Robson, 2017) whereby large AMOC and BSF anomalies at initialization result in persistent advective heat transport signals out to decadal lead times. This CESM-identified mechanism links abyssal thickness anomalies, transported southeastwards from the Labrador Sea towards the MAR, to steric-induced anomalous northward surface flow into the SPNA. When viewed in AMOC in density space, it manifests as predictable upper limb (low density) flow anomalies that lag same-signed lower limb (high density) flow anomalies at intergyre latitudes. CESM2-DP shows close alignment with this CESM1 predictability mechanism (Yeager, 2020; Yeager et al., 2021), with strong, persistent initialized LSW thickness anomalies underpinning high skill in sea-surface height gradients and anomalous northward surface geostrophic flow at multiyear lead times. This alignment is not unexpected given the similarity between the CESM1-DPLE and CESM2-DP systems, which differ in terms of component model physics but share a common initialization methodology (see Table 1).

In CMCC-CM2-SR5, the third highest skill model, we also find evidence of a prominent role for initialized AMOC and gyre anomalies in generating persistent long lead signals in near surface northward flow that coincide with the SPNA warming of the late 1990s. We found that steric-induced sea level gradient change at midlatitudes contributes to the skilful prediction of anomalously warm SPNA conditions in the 2000s in this system, although the onset of the warming in the mid 1990s is not well captured. In CMCC-CM2-SR5, the AMOC and LSW thickness anomalies are considerably weaker than in CESM systems, and they show less correspondence with the BSF and SSH fields. While AMOC in density space shows some long lead persistence of upper limb variability that contributes to the advective warming of the SPNA, the connection to preceding

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

lower limb variations is also less clear. Overall, we conclude that CMCC-CM2-SR5 exhibits a weak version of the CESM1 mechanism, which may account for its lower UOHC skill (and considerably lower UOSC skill).

Two decadal prediction systems which experience the large degradations in skill over forecast lead years, i.e., CanESM5 and NorCPM1 exhibit quite distinct behaviours from one another, and from the high skill models. In CanESM5, the SPNA UOHC is initially well represented relative to observation-based data, but its skill degrades rapidly after forecast year 1. The rapid degradation of skill has previously been attributed to its use of Ocean Reanalysis System 5 (ORAS5, Zuo et al. 2019) for ocean initialization, which exhibits non-stationary anomalies in the SPNA (Sospedra-Alfonso et al., 2021; Carmo-Costa et al., 2025). The initially high skill in the upper ocean in CanESM5 is likely related to its additional usage of SST relaxation towards Optimum Interpolation Sea Surface Temperature (OISST, Huang et al. 2021) and Extended Reconstructed Sea Surface Temperature (ERSSTv5, Huang et al. 2017) values, which provide accurate surface properties in the North Atlantic (Huang et al., 2017, 2021). The AMOC initially shows qualitatively similar multidecadal variability as seen in OMIP simulations, but with much stronger amplitude fluctuations. However, there is little AMOC memory in the system, such that the sign of predicted AMOC anomalies in later forecast years is incorrect in the critical decade leading up to the late 1990s SPNA warming event. Abyssal thickness anomalies track very well with the SSH field in this model at forecast year 1, but only lead to anomalous northward flow near the MAR starting around 2000. At later forecast years, there is still evidence of steric SSH gradients at midlatitudes, but they are anomalously southward near the MAR in the 1990s. Thus, while elements of the CESM1 mechanism are present in CanESM5, for unknown reasons they act to degrade rather than enhance skill at long leads. This model is also unique in that its initialised decadal hindcasts show lower skill in the SPNA than uninitialised historical simulations (Carmo-Costa et al., 2025).

In NorCPM1, a clear LSW production event and subsequent eastward propagation of LSW thickness anomalies towards the MAR does not appear to lead to a corresponding large SSH gradient near the MAR, with anomalous northward flow occurring farther west. In this model, SSH anomalies particularly at later forecast years appear to be much more closely tied to the upper ocean properties, rather than abyssal thickness anomalies (not shown).

In IPSL-CM6A-LR and BCC-CSM2-MR, skill is already poor at the time of initialization, which suggests that the initialization method is problematic. In IPSL, only surface properties are assimilated, and deep ocean conditions may not be adequately constrained (Swingedouw et al., 2026). In BCC-CSM2-MR, where skill is already poor at initialization and degrades further with lead time, it is less clear why the initial state is not aligned with observation-based benchmarks, given the use of subsurface assimilation in that model. In the case of HadGEM3-GC31-MM, the SPNA variability is generally weaker, and prediction skill in the SPNA is significantly greater in the eastern part of the SPNA than in the western part. The AMOC exhibits realistic multidecadal variability at short leads, but again with relative weak amplitude (Fig. S9). The AMOC predictions exhibit very low amplitude signals, and (as in CESM systems) tend to be phase shifted so that the decline begins only in 2010. The SSH field is very noisy with much greater interannual variability than most other models (Fig. S10). All this suggests that the deep LSW signals controlling AMOC, SSH, and UOHC might simply be underrepresented in this system. EC-Earth3 shows a reasonable AMOC close to initialization, but it also weakens considerably during the course of the forecasts (Fig. S9). It shows only a weak steric-induced SSH gradient (Fig. S10).

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

There are several caveats to this analysis. Though both CESM1 and CESM2 emerge as the highest skill models in the SPNA, their agreement in terms of the predictability mechanism may be due to common biases in the Labrador Sea, i.e., the fact that Labrador Sea convection is too strong in both models (Danabasoglu et al., 2020; Yeager et al., 2021). Furthermore, our definition of LSW thickness is based on the Labrador Sea density stratification during the first forecast year of each hindcast simulation and hence may not entirely capture the actual LSW layer in these models. However, recent analyses of the SCOTIA array, which indicate that the strengthened overturning during 2016-2020 was driven by enhanced southward transport of denser LSW, demonstrate this predictability mechanism occurring in reality (Fox et al., 2026).

This intercomparison of SPNA UOHC prediction skill in DCPP decadal predictions reveals a clear contrast in the lead-dependent mechanisms at work in high skill compared to low skill systems. Sustained predictability in this region appears to be linked to the ability to predict strong positive anomalies in AMOC upper limb transport at 45°N in the 1990s. In systems that yield negative ACC scores for forecast years 6-10, there is a lack of strong northward transport over the MAR, and instead strong southward transport is predicted. This results in poor skill at predicting the abrupt, large increase in SPNA UOHC that was observed in the late 1990s. In contrast, all of the systems that exhibit good SPNA skill at long leads predict some version of steric-induced northward transport anomalies over the MAR at multiyear leads. The location timing, and magnitude of that AMOC upper limb signal varies from system to system and these variations may contribute to the spread in system performance.

Here, we have found that in models that have poor skill in the subpolar North Atlantic, the CESM1 predictability mechanism is largely absent, while in the model that has the next highest skill to the CESM systems displays some evidence of the CESM1 predictability mechanism, but with reduced amplitude. The consensus found amongst high skill models, as well as the recent results from SCOTIA observations demonstrating this mechanism occurring in reality, indicates that this mechanism may hold true generally.

Data availability. Data for initialisation datasets and DCPP simulations can be accessed from the ESGF CMIP6 archive.

Appendix A: Descriptions of models and initialisation procedures for DCPP simulations

CESM1-DP (Yeager et al., 2018): Ocean–sea-ice are initialised from a forced ocean-sea ice (FOSI) simulation; atmosphere from a CESM1 Large Ensemble member restart (not directly observationally constrained). Hindcasts initialised on 1 November.

CESM2-DP (Yeager et al., 2022): Ocean–sea-ice initialised from an updated CESM2-based FOSI; atmosphere from JRA55 reanalyses. Hindcasts initialised on 1 November.

CanESM5 (Sospedra-Alfonso et al., 2021): Full-field nudging of assimilation run to ECMWF 6-hourly ERA-Interim atmospheric reanalyses and to ORAS5 ocean T/S; SST also nudged to daily interpolations of OISST weekly values from 1981-2016, and to monthly values from ERSSTv3 from 1958-1981. Hindcasts initialised on 1 January.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

HadGEM3-GC31-MM (Dunstone et al., 2016; Choi et al., 2026): Full-field initialization (DePreSys4): atmosphere nudged to reanalysis; ocean/sea ice to full-field ocean analysis from Hermanson et al. 2023. Hindcasts initialised on 1 November.

MIROC6 (Kataoka et al., 2020): Ocean T/S assimilated to monthly global objective analysis from Ishii et al. 2006 and Ishii and Kimoto 2009; full-field sea ice concentration; atmospheric initial states from reanalysis. Hindcasts initialised on 1 November.

NorCPM1 (Bethke et al., 2021): Anomaly assimilation of SST observations (HadISST2.1.0.0 and OISST when HadISST is not available) and T/S hydrographic profiles (EN4.2.1) via Ensemble Kalman Filter (EnKF) into NorESM1. Data within mixed layer of the first ensemble member not assimilated. Hindcasts initialised on 1 October.

CMCC-CM2-SR5 (Nicolì et al., 2023): Full-field (direct) initialisation from CHOR (1960-2010) and CGLORSv7 ocean reanalyses (2011-2020) using a 3D data assimilation system, surface nudging, and a bias correction scheme.

IPSL-CM6-ESM: SST/SSS anomaly nudging (as for IPSL-CM6A-LR). Hindcasts initialised on 1 January.

EC-Earth3 (Bilbao et al., 2021): Full-field initialization: dedicated coupled assimilation of atmosphere (ERA) and NEMO-LIM reconstructed ocean and sea ice states forced at the surface with DRAKKAR fluxes. A standard surface nudging approach is applied to ocean temperature and salinity fields, which are restored to ORAS4 data. Subsurface assimilation is also applied to create these ocean-sea ice reconstructions. Hindcasts initialised on 1 November.

IPSL-CM6A-LR (Estella-Perez et al., 2020): SST/SSS anomaly nudging. Hindcasts initialised on 1 January.

BCC-CSM2-MR (Wu et al., 2019; Xin et al., 2019): The initial states are derived from two nudging experiments with the oceanic temperature and salinity relaxing to the monthly observed anomaly data from the Simple Ocean Data Assimilation (SODA). The two nudging experiments were integrated from 1958 to 2014 and initialised from the historical experiment of BCC-CSM-MR (January 1 of 1958). Hindcasts initialised on 1 January. FGOALS-f3-L (Hu et al., 2023): Weakly coupled EnOI-IAU anomaly assimilation of SST (HadISST1.1) and 0-1000 m T/S profiles (EN.4.2.2). Hindcasts initialised on 1 November. MPI-ESM1-2-HR (Müller et al., 2018; Mauritsen et al., 2019): Follows the institute's operational decadal prediction protocol; a dedicated methods paper was not identified. Hindcasts initialised on 1 November.

Author contributions. SY and GD conceived the study and designed the analysis. DO downloaded and processed the DCPP hindcast data and performed the predictability analysis. SY, GD, IS, WK and JR contributed to the interpretation of the results. NO calculated and provided the overturning circulation data for HadGEM3-GC31-MM's decadal prediction system. PA provided CMCC-CM2-SR5's initialisation dataset. DO wrote the manuscript with input from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The National Science Foundation (NSF) National Center for Atmospheric Research (NCAR) is a major facility sponsored by the NSF under Cooperative Agreement No. 1852977. The authors are grateful for support from the National Science Foundation

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.



(NSF) through Grant 2406511. We thank the modelling team at IPSL for providing us with the IPSL-CM6-ESM hindcast simulation data which we used in our analysis.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Athanasiadis, P. J., Yeager, S., Kwon, Y.-O., Bellucci, A., Smith, D. W., and Tibaldi, S.: Decadal predictability of North Atlantic blocking and the NAO, npj Climate and Atmospheric Science, 3, 20, https://doi.org/10.1038/s41612-020-0120-6, 2020.

Bethke, I., Wang, Y., Counillon, F., Keenlyside, N., Kimmritz, M., Fransner, F., Samuelsen, A., Langehaug, H., Svendsen, L., Chiu, P.-G., Passos, L., Bentsen, M., Guo, C., Gupta, A., Tjiputra, J., Kirkevåg, A., Olivié, D., Seland, Ø., Solsvik Vågane, J., Fan, Y., and Eldevik, T.: NorCPM1 and its contribution to CMIP6 DCPP, Geoscientific Model Development, 14, 7073–7116, https://doi.org/10.5194/gmd-14-7073-2021, 2021.

Biastoch, A., Böning, C. W., Getzlaff, J., Molines, J.-M., and Madec, G.: Causes of Interannual–Decadal Variability in the Meridional Overturning Circulation of the Midlatitude North Atlantic Ocean, Journal of Climate, 21, 6599–6615, https://doi.org/10.1175/2008JCLI2404.1, 2008.

Bilbao, R., Wild, S., Ortega, P., Acosta-Navarro, J., Arsouze, T., Bretonnière, P.-A., Caron, L.-P., Castrillo, M., Cruz-García, R., Cvijanovic, I., Doblas-Reyes, F. J., Donat, M., Dutra, E., Echevarría, P., Ho, A.-C., Loosveldt-Tomas, S., Moreno-Chamarro, E., Pérez-Zanon, N., Ramos, A., Ruprich-Robert, Y., Sicardi, V., Tourigny, E., and Vegas-Regidor, J.: Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth, Earth System Dynamics, 12, 173–196, https://doi.org/10.5194/esd-12-173-2021, 2021.

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., Kushnir, Y., Kimoto, M., Meehl, G. A., Msadek, R., Mueller, W. A., Taylor, K. E., Zwiers, F., Rixen, M., Ruprich-Robert, Y., and Eade, R.: The Decadal Climate Prediction Project (DCPP) contribution to CMIP6, Geoscientific Model Development, 9, 3751–3777, https://doi.org/10.5194/gmd-9-3751-2016, 2016.

Borchert, L. F., Menary, M. B., Swingedouw, D., Sgubin, G., Hermanson, L., and Mignot, J.: Improved Decadal Predictions of North Atlantic Subpolar Gyre SST in CMIP6, Geophysical Research Letters, 48, e2020GL091 307, https://doi.org/10.1029/2020GL091307, e2020GL091307 2020GL091307, 2021.

Carmo-Costa, T., Bilbao, R., Ortega, P., Teles-Machado, A., and Dutra, E.: Trends, variability and predictive skill of the ocean heat content in North Atlantic: an analysis with the EC-Earth3 model, Clim Dyn, 58, 1311–1328, https://doi.org/10.1007/s00382-021-05962-y, 2022.

Carmo-Costa, T., Bilbao, R., Robson, J., Teles-Machado, A., and Ortega, P.: A multi-model analysis of the decadal prediction skill for the North Atlantic ocean heat content, Earth System Dynamics, 16, 1001–1028, https://doi.org/10.5194/esd-16-1001-2025, 2025.

Choi, J., Jun, S.-Y., Son, S.-W., Hyun, Y.-K., Lee, J.-R., Lee, J., Boo, K.-O., and Park, B.-J.: Near-term Climate Prediction of Agricultural Thermal Conditions in East Asia, Advances in Atmospheric Sciences, 43, 631–644, https://doi.org/10.1007/s00376-025-4471-0, 2026.

Collins, M., Botzet, M., Carril, A. F., Drange, H., Jouzeau, A., Latif, M., Masina, S., Otteraa, O. H., Pohlmann, H., Sorteberg, A., Sutton, R., and Terray, L.: Interannual to Decadal Climate Predictability in the North Atlantic: A Multimodel-Ensemble Study, Journal of Climate, 19, 1195 – 1203, https://doi.org/10.1175/JCLI3654.1, 2006.

Danabasoglu, G.: On Multidecadal Variability of the Atlantic Meridional Overturning Circulation in the Community Climate System Model Version 3, Journal of Climate, 21, 5524 – 5544, https://doi.org/10.1175/2008JCLI2019.1, 2008.

Danabasoglu, G., Yeager, S. G., Kwon, Y.-O., Tribbia, J. J., Phillips, A. S., and Hurrell, J. W.: Variability of the Atlantic Meridional Overturning Circulation in CCSM4, Journal of Climate, 25, 5153 – 5172, https://doi.org/10.1175/JCLI-D-11-00463.1, 2012.

Danabasoglu, G., Yeager, S. G., Kim, W. M., Behrens, E., Bentsen, M., Bi, D., Biastoch, A., Bleck, R., Böning, C., Bozec, A., Canuto, V. M., Cassou, C., Chassignet, E., Coward, A. C., Danilov, S., Diansky, N., Drange, H., Farneti, R., Fernandez, E., Fogli, P. G., Forget, G., Fujii, Y., Griffies, S. M., Gusev, A., Heimbach, P., Howard, A., Ilicak, M., Jung, T., Karspeck, A. R., Kelley, M., Large, W. G., Leboissetier, A.,

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Lu, J., Madec, G., Marsland, S. J., Masina, S., Navarra, A., Nurser, A. G., Pirani, A., Romanou, A., Salas y Mélia, D., Samuels, B. L., Scheinert, M., Sidorenko, D., Sun, S., Treguier, A.-M., Tsujino, H., Uotila, P., Valcke, S., Voldoire, A., Wang, Q., and Yashayaev, I.: North Atlantic simulations in Coordinated Ocean-ice Reference Experiments phase II (CORE-II). Part II: Inter-annual to decadal variability, Ocean Modelling, 97, 65–90, https://doi.org/10.1016/j.ocemod.2015.11.007, 2016.

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., van Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, e2019MS001 916, https://doi.org/https://doi.org/10.1029/2019MS001916, e2019MS001916 2019MS001916, 2020.

Delworth, T. L. and Zeng, F.: The Impact of the North Atlantic Oscillation on Climate through Its Influence on the Atlantic Meridional Overturning Circulation, Journal of Climate, 29, 941–962, https://doi.org/10.1175/JCLI-D-15-0396.1, 2016.

Delworth, T. L., Zeng, F., Vecchi, G. A., Yang, X., Zhang, L., and Zhang, R.: The North Atlantic Oscillation as a driver of rapid climate change in the Northern Hemisphere, Nature Geosci, 9, 509–512, http://dx.doi.org/10.1038/ngeo2738, 2016.

Doblas-Reyes, F., Andreu-Burillo, I., and Chikamoto, Y. e. a.: Initialized near-term regional climate change prediction, Nat Commun, 4, 1715, https://doi.org/10.1038/ncomms2704, 2013.

Dunstone, N., Smith, D., Scaife, A., Hermanson, L., Eade, R., Robinson, N., Andrews, M., and Knight, J.: Skilful predictions of the winter North Atlantic Oscillation one year ahead, Nature Geosci, 9, 809–814, https://doi.org/10.1038/ngeo2824, 2016.

Eden, C. and Jung, T.: North Atlantic Interdecadal Variability: Oceanic Response to the North Atlantic Oscillation (1865–1997), Journal of Climate, 14, 676–691, https://doi.org/10.1175/1520-0442(2001)014<0676:NAIVOR>2.0.CO;2, 2001.

Estella-Perez, V., Mignot, J., Guilyardi, E., Swingedouw, D., and Reverdin, G.: Advances in reconstructing the AMOC using sea surface observations of salinity, Climate Dynamics, 55, 975–992, https://doi.org/10.1007/s00382-020-05304-4, 2020.

Fox, A. D., Fraser, N. J., Burmeister, K., Jones, S. C., Cunningham, S. A., Drysdale, L. A., Dilmahamod, A. F., and Karstensen, J.: The Scotland-Canada overturning array (SCOTIA): twenty years of meridional overturning in the subpolar North Atlantic, Ocean Science, 22, 1439–1456, https://doi.org/10.5194/os-22-1439-2026, 2026.

Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, Journal of Geophysical Research: Oceans, 118, 6704–6716, https://doi.org/10.1002/2013JC009067, 2013.

Griffies, S. M. and Bryan, K.: Predictability of North Atlantic Multidecadal Climate Variability, Science, 275, 181–184, https://doi.org/10.1126/science.275.5297.181, 1997.

Grist, J., Josey, S., Marsh, R., Good, S. A., Coward, A. C., de Cuevas, B. A., Alderson, S. G., New, A. L., and Madec, G.: The Roles of surface heat flux and ocean heat transport convergence in determining Atlantic Ocean temperature variability, Climate Dyn., 60, 771–790, https://doi.org/10.1007/s10236-010-0292-4, 2010.

Hazeleger, W., Guemas, V., Wouters, B., Corti, S., Andreu–Burillo, I., Doblas–Reyes, F. J., Wyser, K., and Caian, M.: Multiyear climate predictions using two initialization strategies, Geophysical Research Letters, 40, 1794–1798, https://doi.org/10.1002/grl.50355, 2013.

Hermanson, L., Dunstone, N., Eade, R., and Smith, D.: An ensemble reconstruction of ocean temperature, salinity, and the Atlantic meridional overturning circulation 1960–2021, Q. J. R. Meteorol. Soc., 150, 98–111, https://doi.org/10.1002/qj.4587, 2023.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Holliday, N. P., Hughes, S. L., Bacon, S., Beszczynska-Möller, A., Hansen, B., Lavín, A., Loeng, H., Mork, K. A., Østerhus, S., Sherwin, T., and Walczowski, W.: Reversal of the 1960s to 1990s freshening trend in the northeast North Atlantic and Nordic Seas, Geophys. Res. Lett., 35, L03 614, https://doi.org/10.1029/2007GL032675, 2008.

Hu, S., Wu, B., Wang, Y., Zhou, T., Yu, Y., He, B., Lin, P., Bao, Q., Liu, H., Chen, K., and Zhao, S.: CAS FGOALS-f3-L Model Datasets for CMIP6 DCPP Experiment, Adv. Atmos. Sci., 40, 1911–1922, https://doi.org/10.1007/s00376-023-2122-x, 2023.

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S., and Zhang, H.-M.: Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons, Journal of Climate, 30, 8179 – 8205, https://doi.org/10.1175/JCLI-D-16-0836.1, 2017.

Huang, B., Liu, C., Banzon, V., Freeman, E., Graham, G., Hankins, B., Smith, T., and Zhang, H.-M.: Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1, Journal of Climate, 34, 2923 – 2939, https://doi.org/10.1175/JCLI-D-20-0166.1, 2021.

Ishii, M. and Kimoto, M.: Reevaluation of historical ocean heat content variations with time-varying XBT and MBT depth bias corrections, Journal of Oceanography, 65, 287–299, https://doi.org/10.1007/s10872-009-0027-7, 2009.

Ishii, M., Kimoto, M., Sakamoto, K., and Iwasaki, S.-I.: Steric sea level changes estimated from historical ocean subsurface temperature and salinity analyses, Journal of Oceanography, 62, 155–170, https://doi.org/10.1007/s10872-006-0041-y, 2006.

Kataoka, T., Tatebe, H., Koyama, H., Mochizuki, T., Ogochi, K., Naoe, H., Imada, Y., Shiogama, H., Kimoto, M., and Watanabe, M.: Seasonal to Decadal Predictions With MIROC6: Description and Basic Evaluation, Journal of Advances in Modeling Earth Systems, 12, e2019MS002 035, https://doi.org/10.1029/2019MS002035, e2019MS002035 2019MS002035, 2020.

Kim, H. J., An, S. I., and Park, J. e. a.: North Atlantic Oscillation impact on the Atlantic Meridional Overturning Circulation shaped by the mean state, npj Clim Atmos Sci, 6, https://doi.org/10.1038/s41612-023-00354-x, 2023.

Kim, W. M., Yeager, S., Chang, P., and Danabasoglu, G.: Low-Frequency North Atlantic Climate Variability in the Community Earth System Model Large Ensemble, Journal of Climate, 31, 787–813, https://doi.org/10.1175/JCLI-D-17-0193.1, 2018.

Kim, W. M., Yeager, S., and Danabasoglu, G.: Atlantic Multidecadal Variability and Associated Climate Impacts Initiated by Ocean Thermohaline Dynamics, Journal of Climate, 33, 1317–1334, https://doi.org/10.1175/JCLI-D-19-0530.1, 2020.

Kim, W. M., Ruprich-Robert, Y., Zhao, A., Yeager, S., and Robson, J.: North Atlantic Response to Observed North Atlantic Oscillation Surface Heat Flux in Three Climate Models, Journal of Climate, 37, 1777 – 1796, https://doi.org/10.1175/JCLI-D-23-0301.1, 2024.

Kirtman, B., Power, S., Adedoyin, J., Boer, G., Bojariu, R., Camilloni, I., Doblas-Reyes, F., Fiore, A., Kimoto, M., Meehl, G., Prather, M., Sarr, A., Schär, C., Sutton, R., van Oldenborgh, G., Vecchi, G., and Wang, H.: Near-term climate change: projections and predictability, Cambridge University Press, Cambridge. The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, chap. 11, https://doi.org/10.1017/CBO9781107415324.023, 2013.

Kröger, J., Pohlmann, H., Sienz, F., Marotzke, J., Baehr, J., Köhl, A., Modali, K., Polkova, I., Stammer, D., Vamborg, F. S., and Mu¨ller, W. A.: Full-field initialized decadal predictions with the MPI earth system model: an initial shock in the North Atlantic, Clim. Dynam., 51, 2593–2608, https://doi.org/10.1007/s00382-017-4030-1, 2018.

Marsh, R., Josey, S. A., de Cuevas, B. A., Redbourn, L. J., and Quartly, G. D.: Mechanisms for recent warming of the North Atlantic: Insights gained with an eddy-permitting model, J. Geophys. Res., 113, C04 031, https://doi.org/10.1029/2007JC004096, 2008.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L.,

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J.-S., Tian, F., Voigt, A., Vrese, P., Wieners, K.-H., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2, Journal of Advances in Modeling Earth Systems, 11, 998–1038, https://doi.org/10.1029/2018MS001400, 2019.

Mecking, J. V., Keenlyside, N. S., and Greatbatch, R. J.: Multiple timescales of stochastically forced North Atlantic Ocean variability: A model study, Ocean Dynamics, 65, 1367–1381, https://doi.org/10.1007/s10236-015-0868-0, 2015.

Menary, M. and Hermanson, L.: Limits on determining the skill of North Atlantic Ocean decadal predictions, Nat Commun, 9, 1694, https://doi.org/10.1038/s41467-018-04043-9, 2018.

Menary, M. B., R.Hodson, D. L., I.Robson, J., T.Sutton, R., A.Wood, R., and A.Hunt, J.: Exploring the impact of CMIP5 model biases on the simulation of North Atlantic decadal variability, Geophys. Res. Lett., 42, 5926–5934, https://doi.org/10.1002/2015GL064360, 2015.

Moat, B. I., Sinha, B., Josey, S. A., Robson, J., Ortega, P., Sévellec, F., Holliday, N. P., McCarthy, G. D., New, A. L., and Hirschi, J. J.-M.: Insights into Decadal North Atlantic Sea Surface Temperature and Ocean Heat Content Variability from an Eddy-Permitting Coupled Climate Model, Journal of Climate, 32, 6137–6161, https://doi.org/10.1175/JCLI-D-18-0709.1, 2019.

Msadek, R., Delworth, T. L., Rosati, A., Anderson, W., Vecchi, G., Chang, Y.-S., Dixon, K., Gudgel, R. G., Stern, W., Wittenberg, A., Yang, X., Zeng, F., Zhang, R., and Zhang, S.: Predicting a Decadal Shift in North Atlantic Climate Variability Using the GFDL Forecast System, Journal of Climate, 27, 6472 – 6496, https://doi.org/10.1175/JCLI-D-13-00476.1, 2014.

Müller, W. A., Baehr, J., Haak, H., Jungclaus, J. H., Kröger, J., Matei, D., Notz, D., Pohlmann, H., von Storch, J. S., and Marotzke, J.: Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology, Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL053326, 2012.

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F., and Marotzke, J.: A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR), Journal of Advances in Modeling Earth Systems, 10, 1383–1413, https://doi.org/https://doi.org/10.1029/2017MS001217, 2018.

Nicolì, D., Bellucci, A., Ruggieri, P., Athanasiadis, P. J., Materia, S., Peano, D., Fedele, G., Hénin, R., and Gualdi, S.: The Euro-Mediterranean Center on Climate Change (CMCC) decadal prediction system, Geoscientific Model Development, 16, 179–197, https://doi.org/10.5194/gmd-16-179-2023, 2023.

Ortega, P., Robson, J. I., Menary, M., Sutton, R. T., Blaker, A., Germe, A., Hirschi, J. J.-M., Sinha, B., Hermanson, L., and Yeager, S.: Labrador Sea subsurface density as a precursor of multidecadal variability in the North Atlantic: a multi-model study, Earth System Dynamics, 12, 419–438, https://doi.org/10.5194/esd-12-419-2021, 2021.

Patrizio, C. R., Athanasiadis, P. J., Smith, D. M., and Nicolì, D.: Ocean-atmosphere feedbacks key to NAO decadal predictability, npj Climate and Atmospheric Science, 8, 146, https://doi.org/10.1038/s41612-025-01027-7, 2025.

Pohlmann, H., Smith, D., Balmaseda, M., Keenlyside, N. S., Masina, S., Matei, D., Müller, W. A., and Rogel, P.: Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system, Climate Dynamics, 41, 775–785, https://doi.org/10.1007/s00382-013-1663-6, 2013.

Polkova, I., Brune, S., Kadow, C., Romanova, V., Gollan, G., Baehr, J., Glowienka-Hense, R., Greatbatch, R. J., Hense, A., Illing, S., Köhl, A., Kröger, J., Müller, W. A., Pankatz, K., and Stammer, D.: Initialization and Ensemble Generation for Decadal Climate Predictions: A

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Comparison of Different Methods, Journal of Advances in Modeling Earth Systems, 11, 149–172, https://doi.org/10.1029/2018MS001439, 2019.

Polkova, I., Swingedouw, D., Hermanson, L., Köhl, A., Stammer, D., Smith, D., Kröger, J., Bethke, I., Yang, X., Zhang, L., Nicolì, D., Athanasiadis, P. J., Karami, M. P., Pankatz, K., Pohlmann, H., Wu, B., Bilbao, R., Ortega, P., Yang, S., Sospedra-Alfonso, R., Merryfield, W., Kataoka, T., Tatebe, H., Imada, Y., Ishii, M., and Matear, R. J.: Initialization shock in the ocean circulation reduces skill in decadal predictions of the North Atlantic subpolar gyre, Frontiers in Climate, Volume 5 - 2023, https://doi.org/10.3389/fclim.2023.1273770, 2023.

Reverdin, G.: North Atlantic subpolar gyre surface variability (1895–2009), J. Climate, 23, 4571–4584, https://doi.org/10.1175/2010JCLI3493.1, 2010.

Robson, J., Sutton, R., Lohmann, K., Smith, D., and Palmer, M. D.: Causes of the Rapid Warming of the North Atlantic Ocean in the Mid-1990s, Journal of Climate, 25, 4116 – 4134, https://doi.org/10.1175/JCLI-D-11-00443.1, 2012a.

Robson, J., Sutton, R., and Smith, D.: Initialized decadal predictions of the rapid warming of the North Atlantic Ocean in the mid 1990s, Geophysical Research Letters, https://doi.org/10.1029/2012GL053370, 2012b.

Sarafanov, A., Falina, A., Sokov, A., and Demidov, A.: Intense warming and salinification of intermediate waters of southern origin in the eastern subpolar North Atlantic in the 1990s to mid-2000s, J. Geophys. Res., 113, C12 022, https://doi.org/10.1029/2008JC004975, 2008.

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

Smith, D., Eade, R., and Scaife, A.: Robust skill of decadal climate predictions, npj Clim Atmos Sci, 2, 13, https://doi.org/10.1038/s41612-019-0071-y, 2019.

Smith, D. M., Scaife, A. A., Eade, R., Athanasiadis, P., Bellucci, A., Bethke, I., Bilbao, R., Borchert, L. F., Caron, L. P., Counillon, F., Danabasoglu, G., Delworth, T., Doblas-Reyes, F. J., Dunstone, N. J., Estella-Perez, V., Flavoni, S., Hermanson, L., Keenlyside, N., Kharin, V., Kimoto, M., Merryfield, W. J., Mignot, J., Mochizuki, T., Modali, K., Monerie, P. A., Müller, W. A., 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., and Zhang, L.: North Atlantic climate far more predictable than models imply, Nature, 583, 796–800, https://doi.org/10.1038/s41586-020-2525-0, 2020.

Sospedra-Alfonso, R., Merryfield, W. J., Boer, G. J., Kharin, V. V., Lee, W.-S., Seiler, C., and Christian, J. R.: Decadal climate predictions with the Canadian Earth System Model version 5 (CanESM5), Geoscientific Model Development, 14, 6863–6891, https://doi.org/10.5194/gmd-14-6863-2021, 2021.

Swingedouw, D., Reiter, R., Mignot, J., Gastineau, G., Torres, O., and Alkama, R.: Evaluation of two versions of the IPSL-EPOC decadal prediction systems, Frontiers in Climate, Volume 8 - 2026, https://doi.org/10.3389/fclim.2026.1803256, 2026.

Tsujino, H., Urakawa, L. S., Griffies, S. M., Danabasoglu, G., Adcroft, A. J., Amaral, A. E., Arsouze, T., Bentsen, M., Bernardello, R., Böning, C. W., Bozec, A., Chassignet, E. P., Danilov, S., Dussin, R., Exarchou, E., Fogli, P. G., Fox-Kemper, B., Guo, C., Ilicak, M., Iovino, D., Kim, W. M., Koldunov, N., Lapin, V., Li, Y., Lin, P., Lindsay, K., Liu, H., Long, M. C., Komuro, Y., Marsland, S. J., Masina, S., Nummelin, A., Rieck, J. K., Ruprich-Robert, Y., Scheinert, M., Sicardi, V., Sidorenko, D., Suzuki, T., Tatebe, H., Wang, Q., Yeager, S. G., and Yu, Z.: Evaluation of global ocean-sea-ice model simulations based on the experimental protocols of the Ocean Model Intercomparison Project phase 2 (OMIP-2), Geoscientific Model Development, 13, 3643–3708, https://doi.org/10.5194/gmd-13-3643-2020, 2020.

Volpi, D., Guemas, V., and Doblas-Reyes, F.: Comparison of full field and anomaly initialisation for decadal climate prediction: towards an optimal consistency between the ocean and sea-ice anomaly initialisation state, Clim Dyn, 49, 1181–1195, https://doi.org/10.1007/s00382-016-3373-3, 2017.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., Jie, W., Zhang, J., Liu, Y., Zhang, L., Zhang, F., Zhang, Y., Wu, F., Li, J., Chu, M., Wang, Z., Shi, X., Liu, X., Wei, M., Huang, A., Zhang, Y., and Liu, X.: The Beijing Climate Center Climate System Model (BCC-CSM): the main progress from CMIP5 to CMIP6, Geoscientific Model Development, 12, 1573–1600, https://doi.org/10.5194/gmd-12-1573-2019, 2019.

Xin, X., Wei, M., Li, Q., Zhou, W., Luo, Y., and Zhao, Z.: Decadal Prediction Skill of BCC-CSM1.1 with Different Initialization Strategies, Journal of the Meteorological Society of Japan, 97, 733–744, https://doi.org/10.2151/jmsj.2019-043, 2019.

Yeager, S.: Topographic Coupling of the Atlantic Overturning and Gyre Circulations, Journal of Physical Oceanography, 45, 1258–1284, https://doi.org/10.1175/JPO-D-14-0100.1, 2015.

Yeager, S.: The abyssal origins of North Atlantic decadal predictability, Climate Dynamics, https://doi.org/10.1007/s00382-020-05382-4, 2020.

Yeager, S. and Danabasoglu, G.: The Origins of Late-Twentieth-Century Variations in the Large-Scale North Atlantic Circulation, Journal of Climate, 27, 3222–3247, https://doi.org/10.1175/JCLI-D-13-00125.1, 2014.

Yeager, S., Karspeck, A., Danabasoglu, G., Tribbia, J., and Teng, H.: A Decadal Prediction Case Study: Late Twentieth-Century North Atlantic Ocean Heat Content, Journal of Climate, 25, 5173 – 5189, https://doi.org/10.1175/JCLI-D-11-00595.1, 2012.

Yeager, S., Castruccio, F., Chang, P., Danabasoglu, G., Maroon, E., Small, J., Wang, H., Wu, L., and Zhang, S.: An outsized role for the Labrador Sea in the multidecadal variability of the Atlantic overturning circulation, Science Advances, 7, eabh3592, https://doi.org/10.1126/sciadv.abh3592, 2021.

Yeager, S. G. and Robson, J. I.: Recent Progress in Understanding and Predicting Atlantic Decadal Climate Variability, Current Climate Change Reports, 3, 112–127, https://doi.org/10.1007/s40641-017-0064-z, 2017.

Yeager, S. G., Karspeck, A. R., and Danabasoglu, G.: Predicted slowdown in the rate of Atlantic sea ice loss, Geophysical Research Letters, 42, 10,704–10,713, https://doi.org/10.1002/2015GL065364, 2015.

Yeager, S. G., Danabasoglu, G., Rosenbloom, N. A., Strand, W., Bates, S. C., Meehl, G. A., Karspeck, A. R., Lindsay, K., Long, M. C., Teng, H., and Lovenduski, N. S.: Predicting Near-Term Changes in the Earth System: A Large Ensemble of Initialized Decadal Prediction Simulations Using the Community Earth System Model, Bulletin of the American Meteorological Society, 99, 1867 – 1886, https://doi.org/10.1175/BAMS-D-17-0098.1, 2018.

Yeager, S. G., Rosenbloom, N., Glanville, A. A., Wu, X., Simpson, I., Li, H., Molina, M. J., Krumhardt, K., Mogen, S., Lindsay, K., Lombardozzi, D., Wieder, W., Kim, W. M., Richter, J. H., Long, M., Danabasoglu, G., Bailey, D., Holland, M., Lovenduski, N., Strand, W. G., and King, T.: The Seasonal-to-Multiyear Large Ensemble (SMYLE) prediction system using the Community Earth System Model version 2, Geoscientific Model Development, 15, 6451–6493, https://doi.org/10.5194/gmd-15-6451-2022, 2022.

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

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

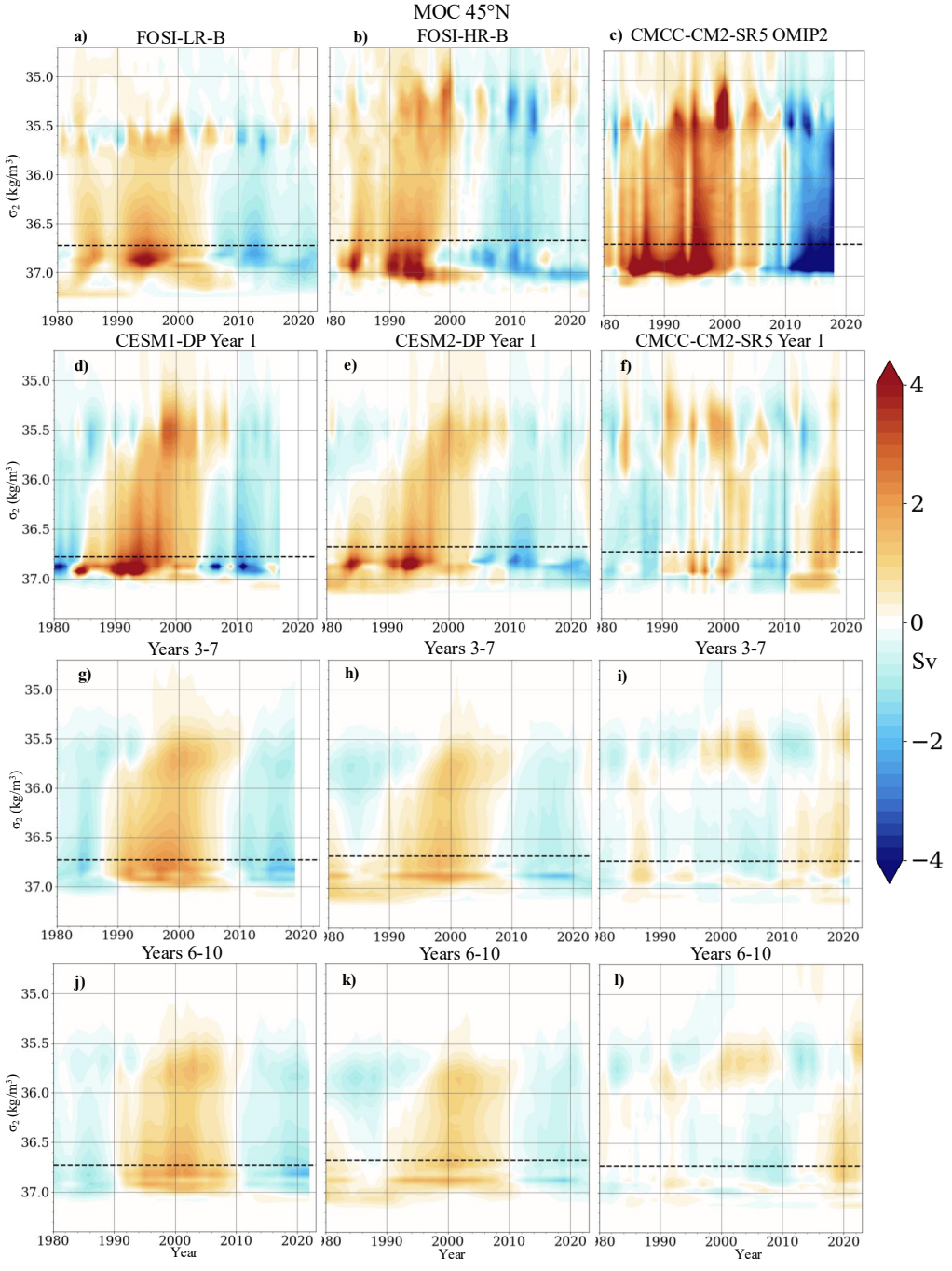


Figure 5. Time series of AMOC in density coordinates at 45°N for **a)** FOSI-LR-B, **b)** FOSI-HR-B, **c)** CMCC-CM2-SR5 OMIP2, as well as historical time series for different forecast lead year periods for **d, g, j)** CESM1-DP, **e, h, k)** CESM2-DP and **f, i, l)** CMCC-CM2-SR5. Dashed lines indicate the density where the lead-dependent maximum climatological AMOC at 45°N occurs in each model.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

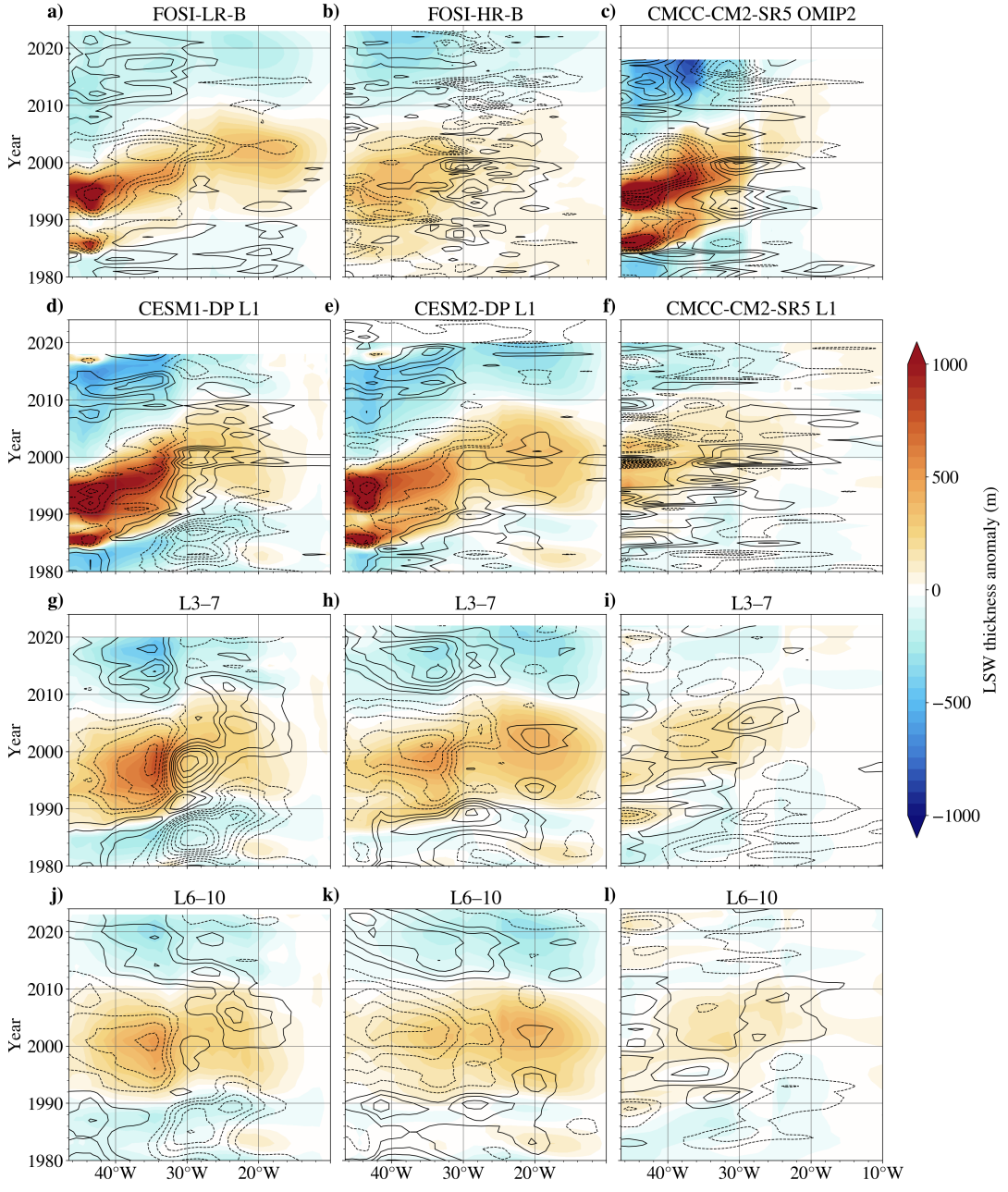


Figure 6. Time series of Labrador Sea Water thickness (colours) and sea-surface height (contours) at 45°N as a function of longitude for **a)** FOSI-LR-B, **b)** FOSI-HR-B, **c)** CMCC-CM2-SR5 OMIP2, as well as historical time series for different forecast lead year (L) periods for **d, g, j)** CESM1-DP, **e, h, k)** CESM2-DP and **f, i, l)** CMCC-CM2-SR5. Contour intervals used for SSH are 0.02 m.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

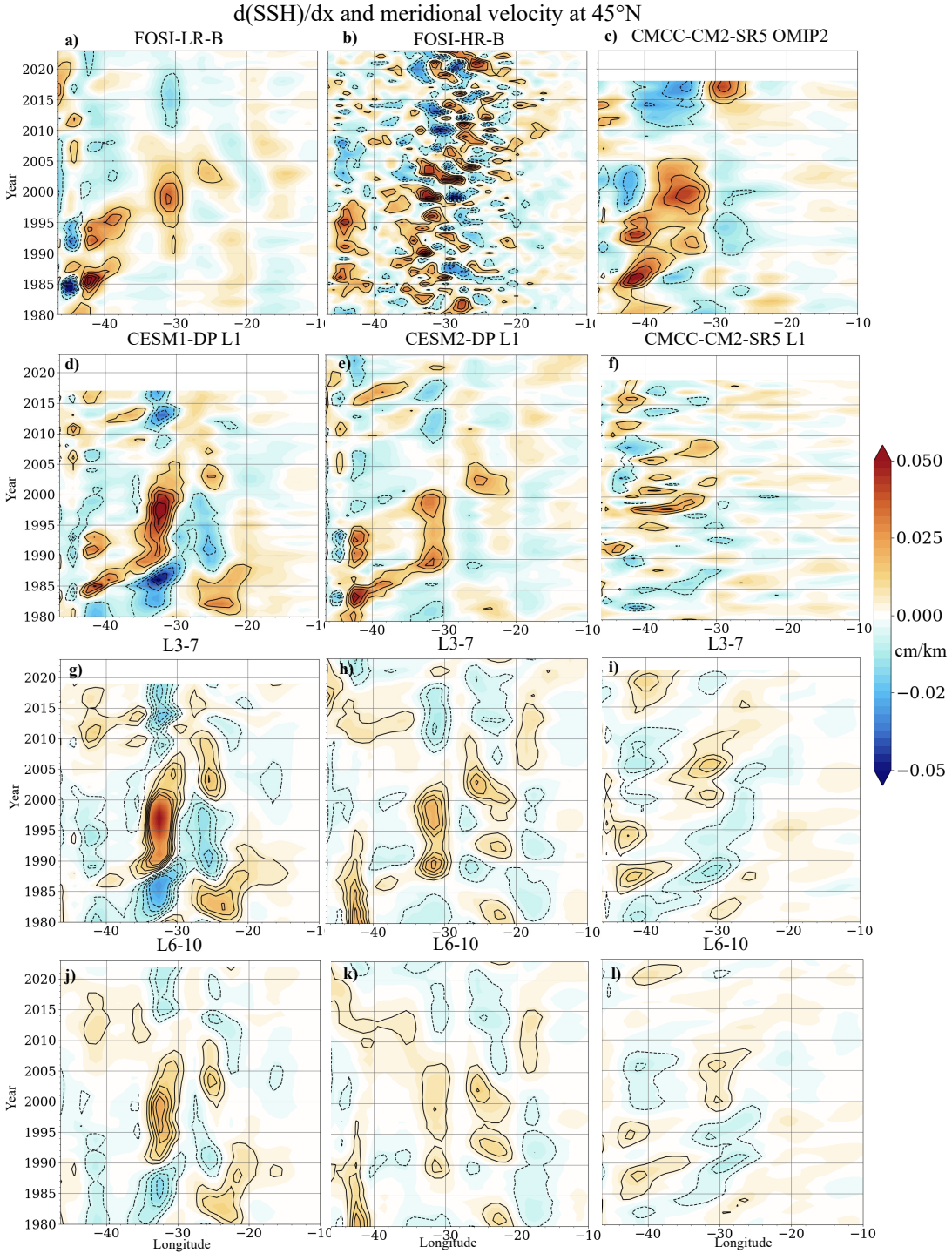


Figure 7. Hovmöller time series of sea-surface height zonal gradient (colours) and meridional velocity averaged over 0-300m (m/s, contours) at 45°N as a function of longitude for **a)** FOSI-LR-B, **b)** FOSI-HR-B, **c)** CMCC-CM2-SR5 OMIP2, as well as historical time series for different forecast lead year periods for **d, g, j)** CESM1-DP, **e, h, k)** CESM2-DP and **f, i, l)** CMCC-CM2-SR5.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

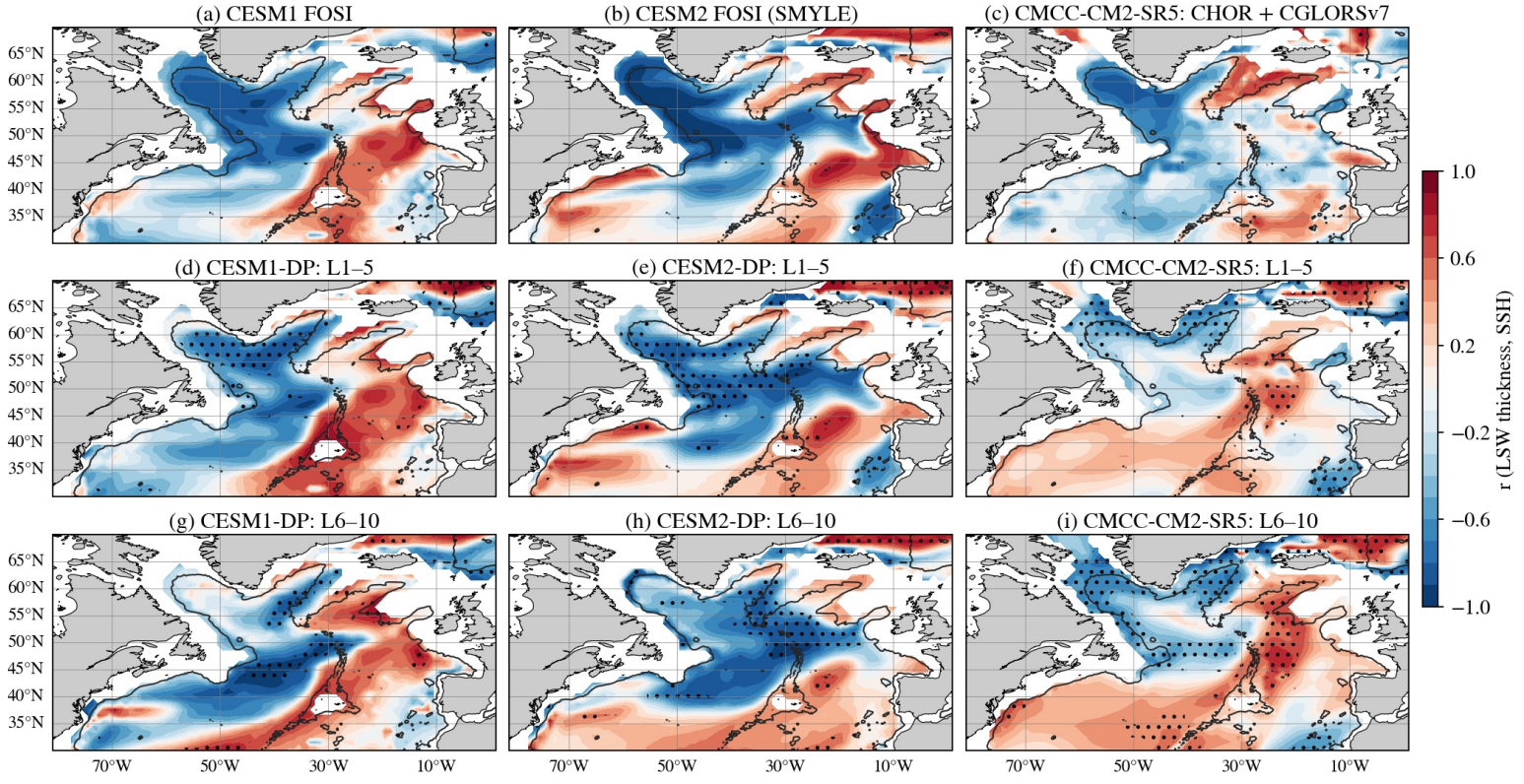


Figure 8. Correlations between 5-year smoothed LSW thickness and SSH for **a)** CESM1 FOSI, **b)** CESM2 FOSI (SMYLE), **c)** CMCC-CM2-SR5 initialisation dataset, as well as for L1-5 of the hindcast runs (middle row) and L6-10 (bottom row) for the three models. The thin black contours indicate the 2500m isobath. Dots indicate areas of statistical significance.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

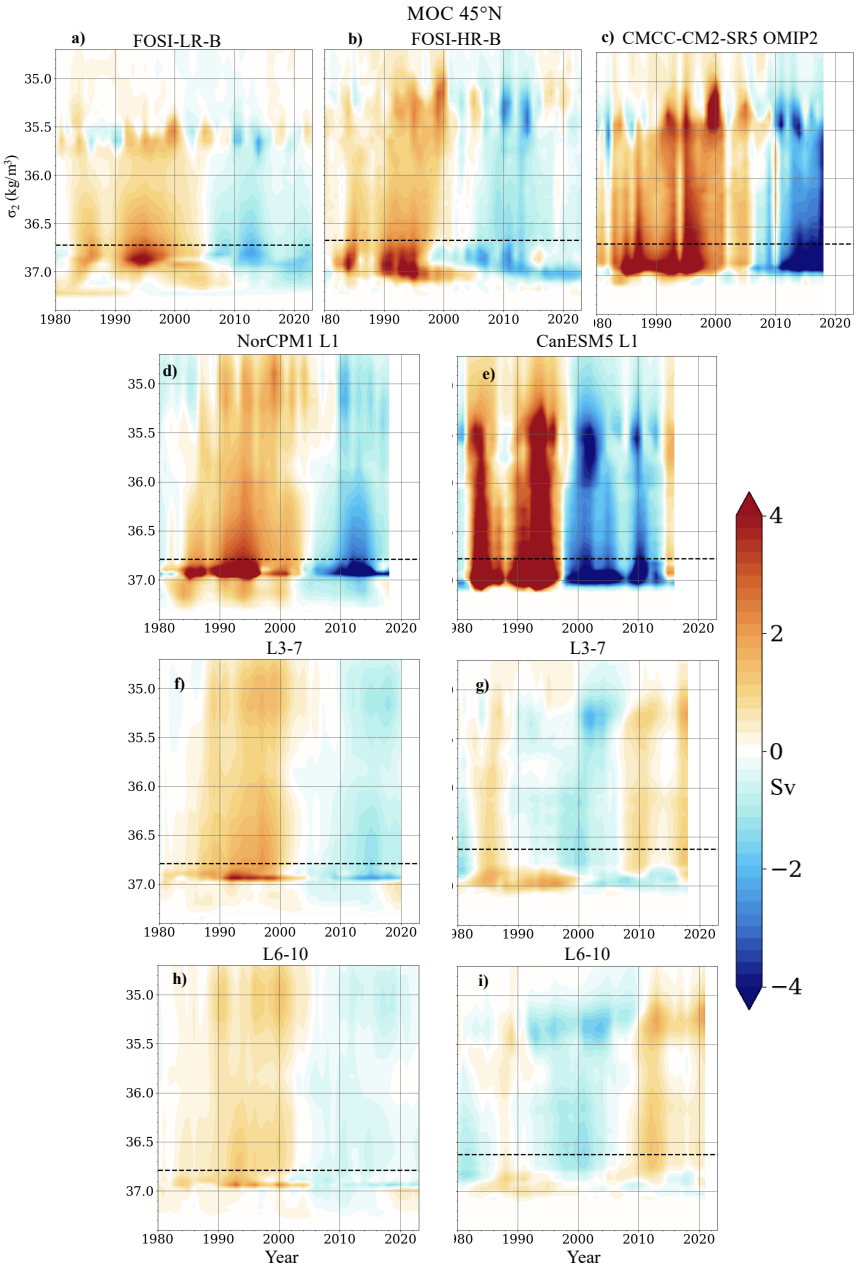


Figure 9. Time series of AMOC in density coordinates at 45°N for **a)** FOSI-LR-B, **b)** FOSI-HR-B, **c)** CMCC-CM2-SR5 OMIP2, as well as historical time series for different forecast lead year periods for **d, f, h)** NorCPM1 and **e, g, i)** CanESM5. Dashed lines indicate the density where the lead-dependent maximum climatological AMOC at 45°N occurs in each model.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

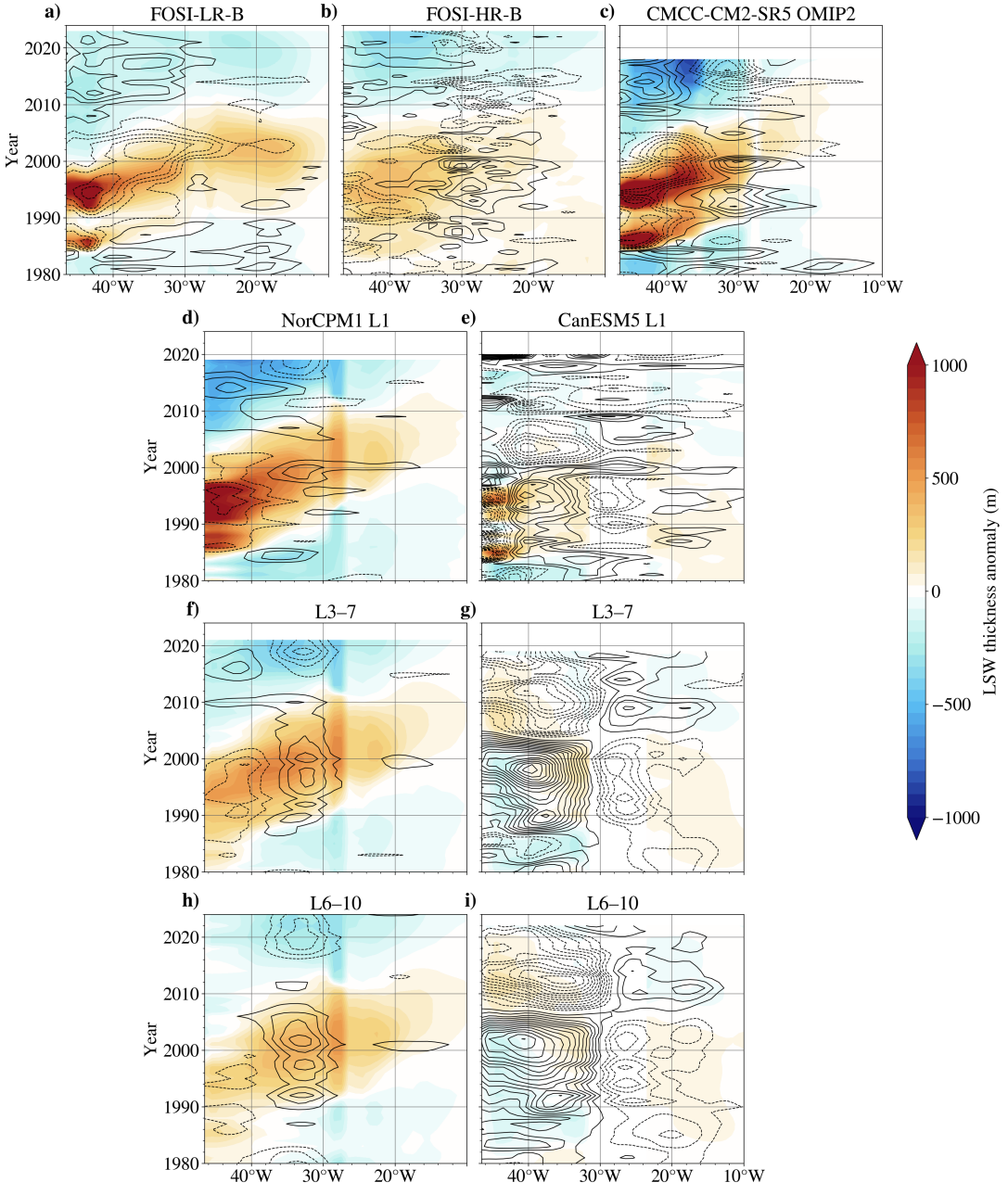


Figure 10. Time series of Labrador Sea Water thickness (colours) and sea-surface height (contours) at 45°N as a function of longitude for **a)** FOSI-LR-B, **b)** FOSI-HR-B, **c)** CMCC-CM2-SR5 OMIP2, as well as historical time series for different forecast lead year (L) periods for **d, f, h)** NorCPM1 and **e, g, i)** CanESM5. Contour intervals used for SSH are 0.02 m.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

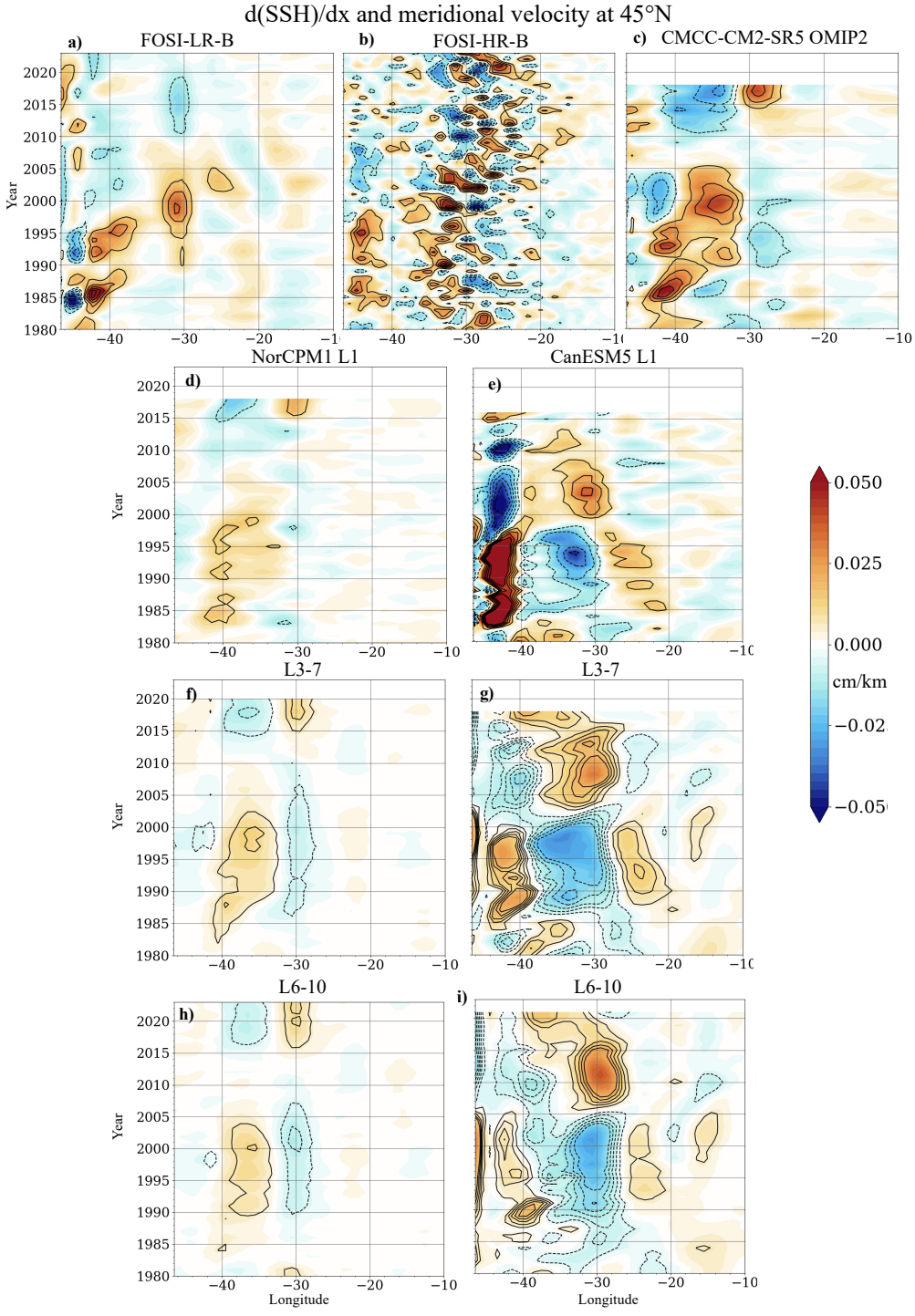


Figure 11. Time series of sea-surface height gradient (colours) and meridional velocity averaged over 0-300m (m/s, contours) at 45°N for **a)** FOSI-LR-B, **b)** FOSI-HR-B, **c)** CMCC-CM2-SR5 OMIP2, as well as historical time series for different forecast lead year (L) periods for **d, f, h)** NorCPM1 and **e, g, i)** CanESM5. Contour intervals used for SSH are 0.02 m.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

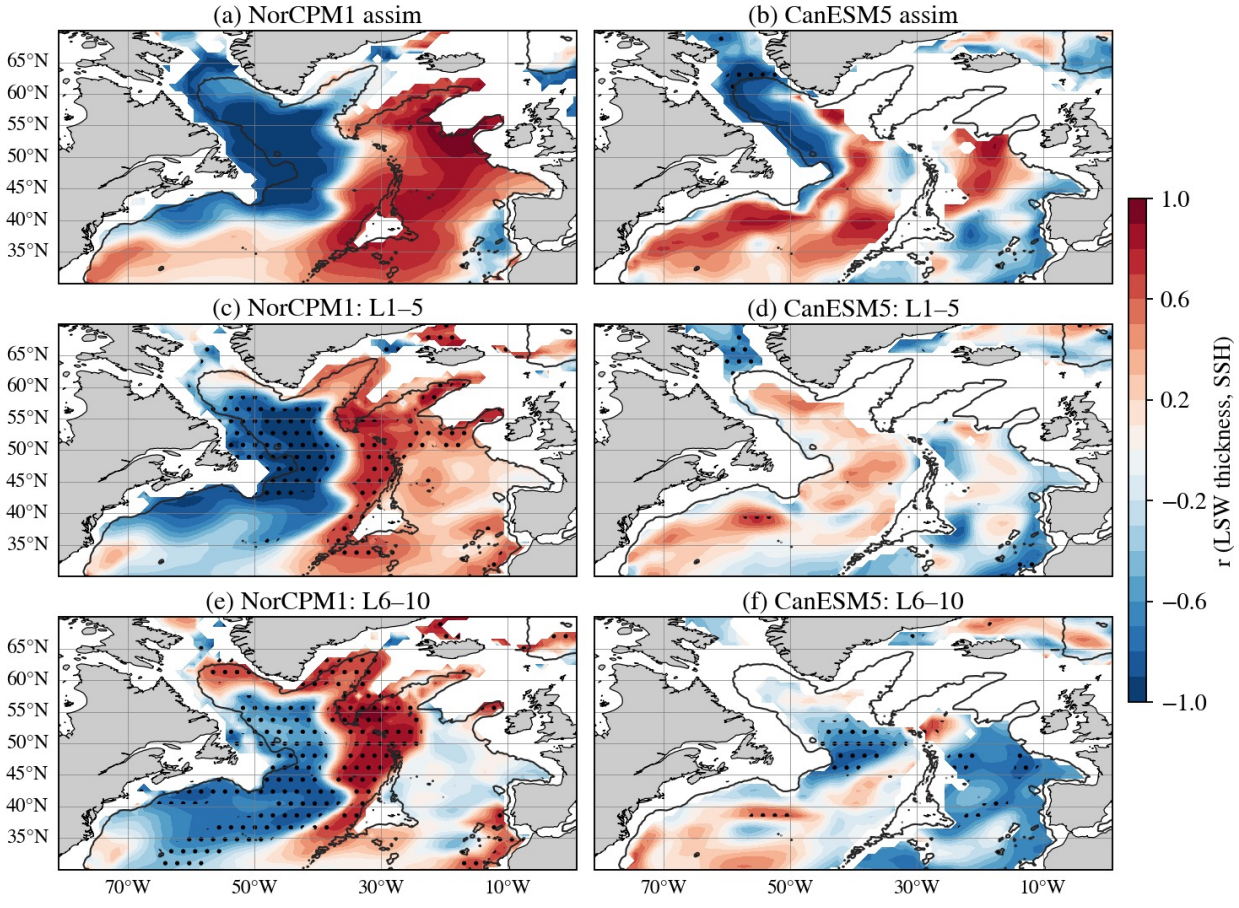


Figure 12. Correlations between 5-year smoothed LSW thickness and SSH for **a)** NorCPM1 assim, **b)** CanESM5 assim, as well as for L1-5 of the hindcast runs (middle row) and L6-10 (bottom row) for the two models. The thin black contours indicate the 2500m isobath. Dots indicate areas of statistical significance.

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

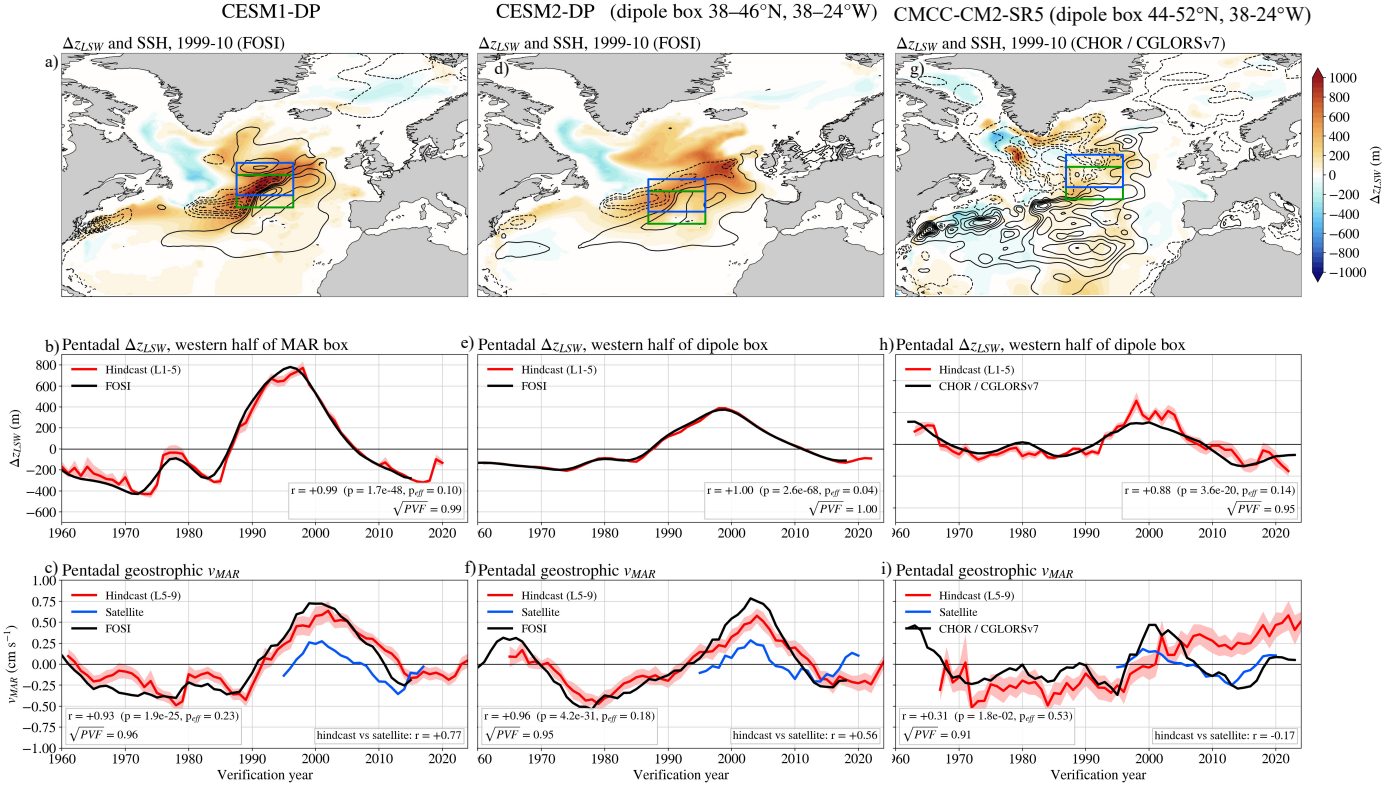


Figure 13. a, d, g) LSW thickness anomalies (colours) from October 1999 alongside 5-year low-pass filtered SSH anomalies (contours) from each model's initialisation dataset. **b, e, h)** Pentadal predicted LSW thickness anomalies for L1-5 plotted alongside values from initialisation datasets for each model. **c, f, i)** Predicted pentadal geostrophic v_{MAR} over L5-9 plotted alongside values from the initialisation datasets as well as those computed from Copernicus Marine Environment Monitoring Service (CMEMS; http://marine.copernicus.edu) satellite observations. The MAR box used for observations in c (blue box) is shifted slightly north of that used for model fields (green box). The shading around CESM-DPLE curves represents ensemble uncertainty ($\pm 1\sigma$). Correlations (and corresponding p-values) are given for each decadal prediction system verified against its respective initialisation dataset, which can be compared to the square root of the predictable variance fraction (PVF)

https://doi.org/10.5194/egusphere-2026-6427
Preprint. Discussion started: 19 November 2026
© Author(s) 2026. CC BY 4.0 License.

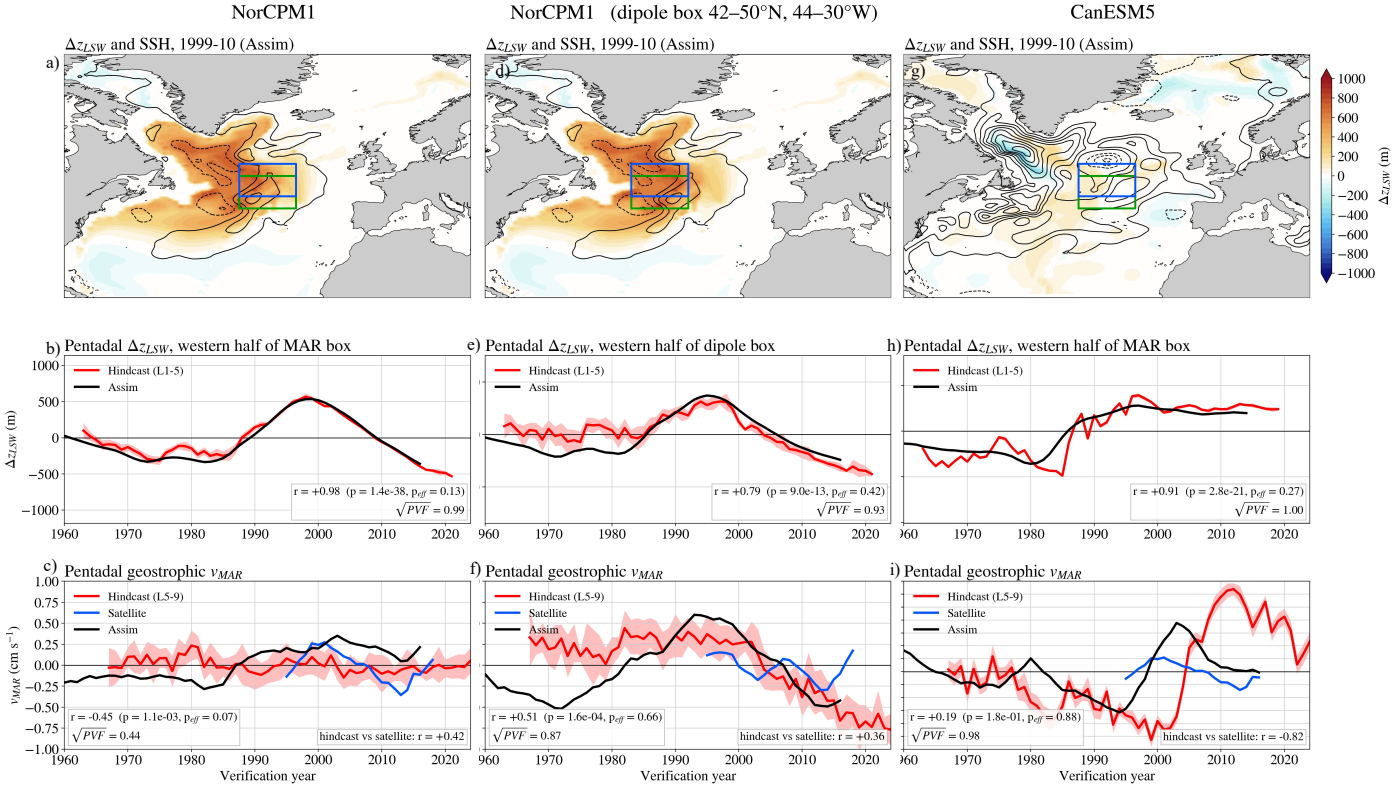


Figure 14. a, d, g) LSW thickness anomalies (colours) from October 1999 alongside 5-year low-pass filtered SSH anomalies (contours) from each model's initialisation dataset. **b, e, h)** Pentadal predicted LSW thickness anomalies for L1-5 plotted alongside values from initialisation datasets for each model. **c, f, i)** Predicted pentadal geostrophic v_{MAR} over L5-9 plotted alongside values from the initialisation datasets as well as those computed from Copernicus Marine Environment Monitoring Service (CMEMS; http://marine.copernicus.edu) satellite observations.