Deep mixed ocean volume in the Labrador Sea in HighResMIP models

Torben Koenigk[1,2] · Ramon Fuentes-Franco[1] · Virna L. Meccia[3] · Oliver Gutjahr[4,5] · Laura C. Jackson[6] · Adrian L. New[7] · Pablo Ortega[8] · Christopher D. Roberts[9] · Malcolm J. Roberts[6] · Thomas Arsouze[8] · Doroteaciro Iovino[10] · Marie-Pierre Moine[11] · Dmitry V. Sein[12,13]

Abstract

Simulations from seven global coupled climate models performed at high and standard resolution as part of the high resolution model intercomparison project (HighResMIP) are analyzed to study deep ocean mixing in the Labrador Sea and the impact of increased horizontal resolution. The representation of convection varies strongly among models. Compared to observations from ARGO-floats and the EN4 data set, most models substantially overestimate deep convection in the Labrador Sea. In four out of five models, all four using the NEMO-ocean model, increasing the ocean resolution from 1° to 1/4° leads to increased deep mixing in the Labrador Sea. Increasing the atmospheric resolution has a smaller effect than increasing the ocean resolution. Simulated convection in the Labrador Sea is mainly governed by the release of heat from the ocean to the atmosphere and by the vertical stratification of the water masses in the Labrador Sea in late autumn. Models with stronger sub-polar gyre circulation have generally higher surface salinity in the Labrador Sea and a deeper convection. While the high-resolution models show more realistic ocean stratification in the Labrador Sea than the standard resolution models, they generally overestimate the convection. The results indicate that the representation of sub-grid scale mixing processes might be imperfect in the models and contribute to the biases in deep convection. Since in more than half of the models, the Labrador Sea convection is important for the Atlantic Meridional Overturning Circulation (AMOC), this raises questions about the future behavior of the AMOC in the models.

✉ Torben Koenigk
  torben.koenigk@smhi.se

1   Rossby Centre, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

2   Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden

3   National Research Council of Italy, Institute of Atmospheric Sciences and Climate (CNR-ISAC), Bologna, Italy

4   Institut für Meereskunde, Universität Hamburg, Hamburg, Germany

5   Max Planck Institute for Meteorology, Hamburg, Germany

6   Met Office, Exeter EX1 3PB, UK

7   National Oceanography Centre, Southampton, SO14 3ZH, UK

8   Barcelona Supercomputing Center–Centro Nacional de Supercomputación (BSC), Barcelona, Spain

9   European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, UK

10 Ocean Modeling and Data Assimilation Division, Centro Euro-Mediterraneo sui Cambiamenti Climatici, Bologna, Italy

11 CECI, Université de Toulouse, CERFACS/CNRS, Toulouse, France

12 Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany

13 Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia

1 Introduction

Open-ocean deep convection is a rare phenomenon, occurring only at a few locations in the world's ocean. It provides a vertical link between properties of the surface ocean and the deep ocean. In the North Atlantic, the northward flowing warm and salty water masses become denser because of large heat loss to the atmosphere and sink into the deep ocean. Deep convection ventilates the deep ocean with oxygen and plays an important role for the storage of carbon and heat.

The Labrador Sea is one of the main convection sites in the North Atlantic. Deep convection in the Labrador Sea and in the Irminger Seas (de Jong and de Steur 2016; de Jong et al. 2018) produce the deep water masses called Labrador Sea Water (LSW) (Clarke and Gascard 1983), which together with the dense overflow waters coming through Faroe Bank Channel and Denmark Strait (Dickson and Brown 1994) and originating from convection in the Greenland-Iceland-Norwegian Seas (GIN-Sea, Ronski and Budeus 2005) form the North Atlantic Deep Water.

The deep convection in the Labrador Sea is mainly driven by wintertime buoyancy loss to the atmosphere (Latif et al. 2006; Frankignoul et al. 2009), which is strongly governed by the North Atlantic Oscillation (NAO). Freshwater transports through Fram Strait (Holland et al. 2001; Jungclaus et al. 2005; Koenigk et al. 2006) and Denmark Strait (Ortega et al. 2017) can modify local density stratification as well and thus contribute to the variability of deep convection in the Labrador Sea. Labrador Sea convection varies strongly on interannual to decadal time scales (Yashayaev and Loder 2016) and was rather shallow in the 2000s, but recovered in recent years with mixing depths exceeding 2000 m (Yashayaev and Loder 2017).

Variability and change of deep convection affect local and remote climate. Reduced convection in the Labrador Sea is linked to lower surface salinity and temperature and more sea ice in the Labrador Sea and Davis Strait (Deser et al. 2002) but it has in addition remote effects on the atmospheric circulation (Koenigk et al. 2006). Deep convection and deep water formation, particularly in the Labrador Sea, have been seen for a long time as an important mechanism for the Atlantic Meridional Overturning Circulation (AMOC), which is fundamentally important for the climate in the North Atlantic and adjacent regions such as Western Europe and the Arctic (Manabe and Stouffer 1999; Mahajan et al. 2011; Koenigk et al. 2012; Jackson et al. 2015). Recent studies from Smeed et al. (2014, 2018) indicate a weakening of the AMOC at 26.5 °N after year 2008. However, it is unclear if this observed weakening is caused by climate change or internal decadal variability (e.g. Roberts et al. 2014; Jackson et al. 2016; Robson et al. 2016). Caesar et al.

(2018) used spatial patterns of sea surface temperature as AMOC indicator and suggested that the AMOC has been reduced since the mid-twentieth century, and Thornally et al. (2018) showed based on paleooceanographic data that the AMOC has been in a low state for the last 150 years. Many modelling studies have discussed the potential for a weakening AMOC as a response to global warming (Swingedouw et al. 2007; Cheng et al. 2013; Brodeau and Koenigk 2016; Koenigk and Brodeau 2017) and linked the weakening to a reduction of the deep water formation (Latif et al. 2006; Deshayes et al. 2007; Koenigk et al. 2007; Frankignoul et al. 2009; Langehaug et al. 2012).

However, more recently, the importance of deep convection for the AMOC has been questioned (Lozier et al. 2017, 2019; Sayol et al. 2019). Kuhlbrodt et al. (2007) and Medhaug and Furevik (2011) identified wind-driven upwelling, gyre circulation, and wind and tidal vertical mixing as important processes sustaining the long-term strength of the AMOC, so that a collapse of the deep convection would not necessarily lead to a collapse of the AMOC (Gelderloos et al. 2012; Marotzke and Scott 1999). Menary et al. (2020) showed that the generation of density anomalies in the Irminger Sea, which propagate into the Labrador Sea is of importance. However, still short observational periods and potential lags of several years between AMOC and convection (Roberts et al. 2013; Brodeau and Koenigk 2016) make robust conclusions on the linkage between Labrador Sea convection and AMOC from observations difficult.

The deep convective process is temporally intermittent and spatially compact. This makes this process difficult to observe, and state-of-the-art climate models, such as models participating in the Coupled Model Intercomparison Project, Phase 6 (CMIP6), need to use parameterizations to represent convective processes (Fox-Kemper et al. 2019). Given the importance of the deep convection for climate and its future change, a reliable representation in climate models is highly important. However, Heuzé (2017) found that "the majority of CMIP5 models convect too deeply, over too large an area, too often and too far south". Although CMIP6 models seem to improve the representation of bottom waters, more improvements are required (Heuzé 2020).

Going beyond the horizontal resolution of the standard resolution CMIP6 models, we analyse in this study whether high-resolution models from the CMIP6 High-Resolution Model Intercomparison Project (HighResMIP, Haarsma et al. 2016) improve the representation of deep convection in the Labrador Sea. We use simulations from seven models participating in HighResMIP, which have been undertaken in the EU-H2020-project PRIMAVERA. Most of the high-resolution model versions of these seven models have ocean grid resolutions of around 10–25 km in the Labrador Sea. Since the baroclinic Rossby radius of deformation is small in high latitudes (below 10 km in polar regions and 200 km

in the tropics), mesoscale oceanic eddies are not resolved in the sub-polar convection regions in most of the HighResMIP models. We thus denote them as "eddy permitting" but not "eddy-resolving".

Recent studies found that the increased resolution in the HighResMIP models improved many aspects of the ocean including temperature and salinity biases (Gutjahr et al. 2019), the northward ocean heat transport (Grist et al. 2018), sea ice in the Atlantic sector of the Arctic Ocean (Docquier et al. 2019, 2020), the position of the North Atlantic Current (Sein et al. 2018) as well as Arctic freshwater exports (Fuentes Franco and Koenigk 2019), all of them with the potential to improve the representation of Labrador Sea convection.

After this introduction, we proceed with describing the models, the data and the methods in Sect. 2. Sections 3 to 5 then show the results from this study and we conclude in Sect. 6.

2 Models, data and method

2.1 Models and simulations

In this study we analyze seven global coupled climate models (see e.g. Vannière et al. 2019; Roberts et al. 2020), which contributed to HighResMIP (Haarsma et al. 2016) within the H2020-EU-project PRIMAVERA. These models are ECMWF-IFS (Roberts et al. 2018), HadGEM3-GC31 (Roberts et al. 2019), MPI-ESM1.2 (Gutjahr et al. 2019), CMCC-CM2 (Cherchi et al. 2019), CNRM-CM6.1 (Voldoire et al. 2019), AWI-CM-1.0 (Sidorenko et al. 2014; HR and LR setups: Sein et al. 2016) and EC-Earth3P (Haarsma et al. 2020). The set of HighResMIP experiments is divided into three tiers consisting of atmosphere-only and coupled runs, which span the period 1950–2050 (details in Haarsma et al. 2016). Here, we use the Tier 2 historical coupled simulations from 1950–2014 and the 100-year control simulations from these seven models for our analysis. The control runs used fixed 1950s forcing (greenhouse gases, including ozone and aerosol loading for a 1950s (~ 10 year mean) climatology). Both control and historical runs are initialized from the end of 30–50-year spin-up simulations, which were initialized using 1950-ocean conditions of the EN4 data set (Good et al. 2013), using 1950s-forcing. The 30–50-year spin-up simulations are not long enough to bring the models into equilibrium but they are sufficient to substantially reduce drifts in surface variables such as sea surface temperature and sea ice concentration. However, the deeper ocean might still show drifts throughout the entire historical simulation and these drifts might affect deep convection in the North Atlantic. The control runs are used to evaluate potential model drifts in the historical simulations.

All models performed historical and control simulations in at least two different resolutions. Changes in oceanic and atmospheric tuning parameters are kept to a minimum between low- and high-resolution simulations. All models use the same vertical ocean mixing scheme in their low- and high-resolution versions (Turbulent Kinetic Energy (TKE) mixing scheme or K-profile mixing scheme (KPP), see Table 1). Three models with changing ocean resolution use Gent-McWilliams (GM) parameterization of ocean eddies in their low-resolution ocean version but not in the high-resolution versions (HadGEM-3-GC31, ECMWF-IFS, AWI-CM.1–0) while CNRM-CM6.1 and EC-Earth3P use GM-parameterization in both low and high ocean resolution. The parameter values for e.g. horizontal eddy diffusivity and viscosity, iso-neutral diffusion and eddy-induced velocity coefficient vary with the horizontal resolution in the models (see Table 1 in Roberts et al. 2020). As will be shown in the result section, this does not seem to strongly affect the change of deep convection with resolution and the differences between high- and low-resolution model versions can mainly be attributed to the change in resolution.

The resolution varies among the models. A few of the models change both ocean and atmosphere resolution at the same time while others separately change ocean or atmosphere resolution. This allows us to analyze also the effect of increasing the resolution in only one component of the system. Five of the seven models use the NEMO-ocean model as ocean component. Although the NEMO-model configurations differ quite substantially from each other with different sea ice models (LIM2, LIM3, GELATO, CICE) and differences in parameterizations (e.g. mesoscale eddy closure), we acknowledge that the results we present on the effect of increasing resolution are skewed towards NEMO-models with 1° and 1/4° horizontal resolution. AWI-CM-1–0 and MPI-ESM1.2 use the same atmosphere component but different ocean components.

More details on the models and the simulations used in this study are provided in Table 1 and in Roberts et al. (2020).

2.2 Observational data

As a proxy for convection in the Labrador Sea, we use the simulated mixed layer depth (MLD) and compare it with observations from ARGO-profiles (Holte et al. 2017), which are provided on a 1° grid. Two different ARGO data sets of the MLD are existing and are used in this study: first, the climatological mean MLD in each grid point in March, averaged over all observations in the years 2000–2015; second, the maximum (mean over the three largest observed values in the period 2000–2015) MLD in each grid-point. To calculate the MLD in the ARGO-data, Holte et al. (2017) used the

Table 1 Overview on the model configurations and the simulations used in this study. Further details to the parameter values in the ocean parameterizations of the models are listed in Table 1 of Roberts et al. (2020)

Model	Ocean model resolution	Atmosphere model resolution	Ocean mixing scheme	GM-eddy parameterization	hist runs 1950–2014	100-year ctrl-1950
ECMWF-IFS	NEMO3.4/LIM2	IFS cycle 43r1				
ECMWF-IFS-LR	ORCA1—1°	50 km	TKE	Yes	6	1
ECMWF-IFS-MR	ORCA025—1/4°	50 km	TKE	No	1	1
ECMWF-IFS-HR	ORCA025—1/4°	25 km	TKE	No	4	1
HadGEM3-GC31	NEMO3.6/ CICE5.1	UM				
HadGEM3-GC31-LL	ORCA1—1°	130 km	TKE	Yes	1	1
HadGEM3-GC31-MM	ORCA025—1/4°	60 km	TKE	No	1	1
HadGEM3-GC31-HM	ORCA025—1/4°	25 km	TKE	No	1	1
HadGEM3-GC31-HH	ORCA12—1/12°	25 km	TKE	No	1	1
MPI-ESM1-2	MPIOM1.6.3	ECHAM6.3				
MPI-ESM1-2-HR	TP04—0.4°	T127	KPP	Yes	1	1
MPI-ESM1-2-XR	TP04—0.4°	T255	KPP	Yes	1	1
CMCC-CM2	NEMO3.6/CIC4.0	CAM4				
CMCC-CM2-HR4	ORCA025 -1/4°	100 km	TKE	No	1	1
CMCC-CM2-VHR4	ORCA025—1/4°	25 km	TKE	No	1	1
CNRM-CM6-1	NEMO3.6/GELATO	ARPEGE6.3				
CNRM-CM6-1	ORCA1—1°	T127	TKE	Yes	1	1
CNRM-CM6-1-HR	ORCA025—1/4°	T359	TKE	Yes	1	1
AWI-CM-1-0	FESOM	ECHAM6.3				
AWI-CM-1–0-LR	50 km	T63	KPP	Yes	1 (− 2010)	1
AWI-CM-1–0-HR	25 km	T127	KPP	No	1 (− 2010)	1
EC-Earth3P	NEMO3.6/LIM3	IFS cycle 36r4				
EC-Earth3P	ORCA1—1°	T255	TKE	Yes	1	1
EC-Earth3P-HR	ORCA025—1/4°	T511	TKE	Yes	1	1

de Boyer Montégut et al. (2004) variable density threshold (0.03 kg m^{-3}).

There exist some limitations of the ARGO data. First, only a few ARGO-profiles exist in many grid-points and in some no profiles at all. Second, ARGO-floats generally sample to a depth of 2000 m, thus MLD extending below 2000 m are not captured. Finally, the observational time period is relatively short given the long time scales of variability in deep convection. Thus, the ARGO-observations provide only an estimate of the averaged real world MLD for the period 2000–2015.

In addition to ARGO, we use ocean temperature (T) and salinity (S) from the EN4-data set for the period 1950–2014 (Good et al. 2013). EN4 uses optimal interpolation and relaxation to climatology on fixed horizontal and vertical length scales. For periods with sparse observations, this method might lead to substantial uncertainties. Thus, results from EN4 must be interpreted with cautious as well, especially before year 2000, when no ARGO data were available as input for EN4. However, Jackson et al. (2020) showed that temperature and salinity differences between earlier periods of EN4 and the period after 2000 are relatively small

compared to biases in the models. Given the large decadal to multi-decadal variability of ocean properties in the sub-polar North Atlantic, the 65-year time period from EN4 is useful for comparison to the model results in addition to the comparison to the 15-year ARGO-data. Further, EN4 data are more comparable with models both in temporal and spatial scales, as both come with gridded data of much coarser resolution than the ARGO profiles, and are resolved at monthly time-scales. We calculate the MLD from EN4 based on the monthly mean values of ocean temperature and use a MLD threshold criteria of $\Delta T = 0.5$ K following Obata et al. (1996). Here, we use only the temperature to calculate the MLD since the temperature values are more reliable than the salinity values in EN4. The T criterion that we apply, generally corresponds to the often used density criterion $\Delta\rho = 0.125$ kg m^{-3} (e.g. Monterey and Levitus 1997). However, depending on the criterion and the observational data set that is used, the MLD may vary (Toyoda et al. 2017). In addition, the use of monthly averages compared to individual profiles as in ARGO MLD data leads to uncertainties (Toyoda et al. 2017).

To evaluate the surface heat fluxes in our models, we use monthly mean turbulent latent and sensible heat fluxes from

the global ocean-surface heat flux products (1958–2006) on a 1° grid developed by the Objectively Analyzed air-sea Heat Fluxes project at the Woods Hole Oceanographic Institution (WHOI-OAFlux).

2.3 Method

Several different indices have been defined for the deep convection in the ocean (e.g. Koenigk et al. 2007; Schott et al. 2009; Yashayaev and Loder 2009; L'Hévéder et al. 2012; Lavergne et al. 2014). These indices consider either the maximum MLD and/ or the horizontal extent of the MLD. However, none of them excludes convective events that are too shallow to contribute to deep water formation. To overcome this problem, Brodeau and Koenigk (2016) defined the so-called "Deep Mixed Volume" (DMV) index, which only considers the convective mixing below a specific depth (critical depth z_{crit}) and integrates the volume of these deep mixed water masses in different convection regions of the North Atlantic. In our study, we use the DMV index for monitoring the deep convection using a critical depth of 1000 m for the Labrador Sea (Brodeau and Koenigk 2016). In the Labrador Sea, convection must reach a depth of around 1000 m in order to renew the LSW and contribute to the North Atlantic Deep Water (Yashayaev 2007). We define the Labrador Sea region as 70° W–40° W, 45° N–72° N, which covers the main Labrador Sea convection area in all seven models.

We use monthly mean values of the March MLD of the model simulations to calculate the DMV. Short convection episodes that exceed z_{crit} might thus be missed. The ocean mixed layer thickness in the models (variable mlotst following CMIP6-conventions is used) is defined by the depth at which a change from the surface sigma-t of 0.125 kg/m^3 has occurred (sigma-t criterion, Levitus 1982). This definition differs from the MLD criterion that we use for EN4 data and from the one used in the ARGO-data (compare Sect. 2.2). For additional comparison, we therefore use the $\Delta T = 0.5$ K criterion to calculate the March MLD from monthly mean data of the models.

We calculate the DMV from the ARGO data and EN4 data to compare with the model results. For ARGO, we infilled grid-points with missing data by interpolating the nearest neighbours. To calculate the DMV with $z_{crit} = 1000$ m from ARGO, we use the maximum MLD since usage of climatological MLDs leads likely to an underestimation of the DMV compared to using time-varying data. However, since only few profiles per grid point are available in ARGO, the maximum MLD is often not much larger than the climatological MLD. These issues and those already discussed in Sect. 2.2 make the ARGO estimates for the DMV uncertain. Similarly, the DMV calculated from EN4 data is uncertain as well, and thus the DMV from both observational data sources provide only a rough estimate for the real

world DMV. The difference in DMV between ARGO and EN4 might be interpreted as observational uncertainty of the DMV. The DMV in EN4 in the Labrador Sea in March is 2.45 times larger than the DMV in ARGO in the period 2000–2014. While this is a huge uncertainty, the differences between DMV in many of the models and in ARGO and EN4 is much larger as we will show in Sect. 3.2.

As an additional comparison, we calculate the mean DMV based on critical depths of 0 m, thus considering the full mixed layer. For DMV with $z_{crit} = 0$ m, the climatological MLD from ARGO has been used.

To calculate the density, we follow the definition of Millero and Poisson (1981). For the calculation of the mean density, temperature and salinity, averaged over the Labrador Sea, we exclude all grid points with a salinity below 30 psu to avoid the effect of low surface salinities near coastlines and ice edges on the average values.

For correlations, we calculate the Pearson correlation coefficient (r). We call a correlation significantly different from 0, if the p-value is 0.05 or smaller based on a two-sided student-t distribution.

To calculate the liquid freshwater transport through the Arctic straits, we use the model grid lines on the native grids of the models that are closest to the geographical landmarks that define Fram Strait (across 78° N), Northern Baffin Bay (78° N) and Denmark Strait (66° N). The freshwater is defined as the amount of zero-salinity water required to reach the observed salinity of a seawater sample starting from a reference salinity. Specifically, liquid freshwater transport (fwt, in m^3/s) is estimated as

$$fwt = \int_{p1}^{p2} \int_{D}^{\eta} \left(\frac{S - Sref}{Sref}\right) v \, dz \, dx$$

for salinity S (in practical salinity units) and velocity v (in m/s) perpendicular to the section. As reference salinity Sref we use 34.8 psu for all models. The integration along z is performed from the bottom at depth D to the sea surface at height η (in this case $\eta = 0$). p1 and p2 are the landmarks and the integration is done considering dx as the length (dz as depth) between every grid point. The solid freshwater transport is calculated from the sea ice transports across the sections assuming a constant ice salinity of 5 psu and Sref of 34.8 psu.

3 Deep convection

This section shows first the analysis of the MLD in March, the month with the strongest convection in both observations and models, in the North Atlantic in the different models and in ARGO and EN4. Then, we focus on the DMV in the
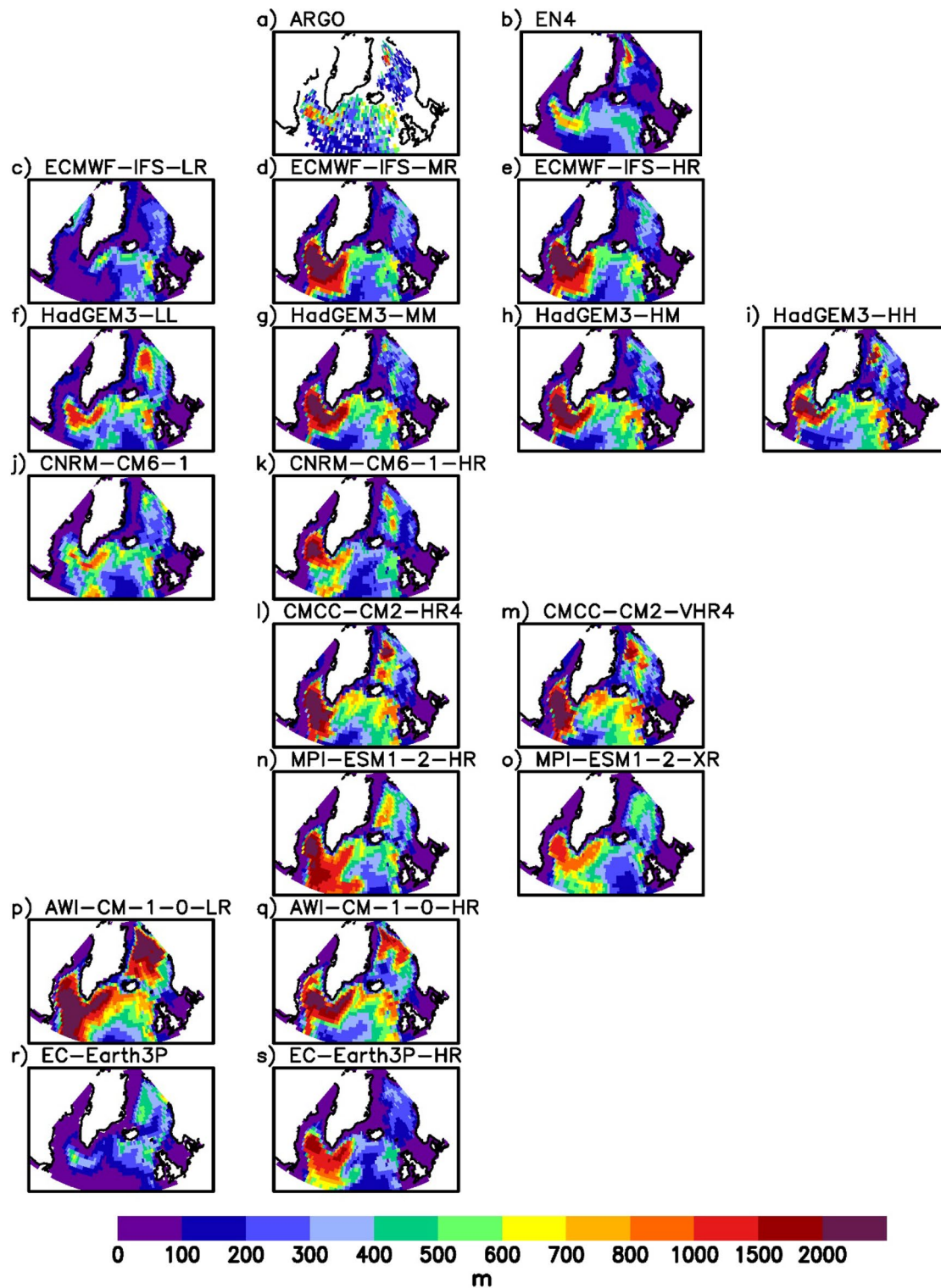
Fig. 1 Mixed layer depth in March in the ARGO-data, averaged over 2000–2015 (**a**), in EN4, averaged over 1950–2014 (**b**) and in the historical low and high-resolution model simulations, averaged over 1950–2014 (**c–s**). Ocean resolution increases from the first column (0.5° and 1°) to the second and third columns (0.25°—0.4°) to the right column (1/12°). Model versions of the same model in the second and third columns have the same ocean resolution but increased atmosphere resolution in column 3

Labrador Sea, potential trends in the historical simulations and differences between models.

3.1 Mixed layer depth in the North Atlantic

Figure 1 shows the average MLD in March from ARGO, EN4 and all historical model simulations. In the period 2000–2015, the ARGO data suggest average MLDs of about 1000 m in the Labrador and GIN Seas (Fig. 1a). The EN4 data show a similar distribution of the MLD in the North Atlantic as ARGO (Fig. 1b). In the models, the MLD differs greatly and shows a strong dependence on the spatial resolution. While ECMWF-IFS-LR and EC-Earth3P show no or very shallow MLD in the main convection areas of the North Atlantic, many of the other simulations strongly overestimate the MLD compared to the observations, especially in the Labrador Sea. The models with particularly deep mixed layers show the largest horizontal extension of convection. Thus, compared to ARGO and EN4, they do not only overestimate the depth of mixed layer but also the area of deep convection. Some models simulate much too strong convection in the Labrador Sea but do not show any deep convection in the GIN Seas while other models overestimate the MLD in both seas. In contrast, in the Irminger Sea, the MLD is more consistent across models and agrees better with ARGO and EN4. Note that we here compare models' MLD averaged over 1950–2014 with ARGO-data from 2000–2015; MLD in EN4 is based on 1950–2014 as in the models. The comparison of MLD across models and between models and observations might depend on the definition of MLD. To analyse this in more detail, we used March mean values of temperature from the model simulations to calculate the mixed layer depth based on the $\Delta T = 0.5$ K criterion (as has been used for MLD in EN4). As expected, the MLD in the models changes based on the definition. Averaged over all models, the MLD based on the $\Delta T = 0.5$ K criterion leads to deeper MLDs, but the change in MLD differs substantially between the models (compare supplementary Figure S1 with Fig. 1). Particularly, in the ECMWF-IFS, MPI-ESM and EC-Earth3P-HR simulations, MLD is deeper while only small changes occur in the HadGEM3-GC31, CMCC-CM2 and AWI-CM1 simulations. Also, the spatial distribution depends partly on the MLD criterion. Particularly, along the southeastern tip of Greenland MLD is increased when using the $\Delta T = 0.5$ K criterion.

As discussed in Sect. 2.2, the realism of the EN4-data depends critically on the available observations and might change over time. Supplementary Figure S2 shows the MLD in EN4 in different 15-year periods between 1950 and 2014. Despite MLD variations between the periods, the biases in the models are substantially larger than the variations in EN4. At least parts of the MLD differences between different

15-year periods in EN4 are likely due to internal climate variability as a comparison to variability in the models between different 15-year periods reveals (supplementary Figures S3 and S4).

The MLD in the Labrador Sea deepens with increasing ocean resolution in all models (ECMWF-IFS, HadGEM3-GC31, CNRM-CM6-1, EC-Earth3P), except for AWI-CM-1–0, independent of the MLD-criterion (Figs. 1 and S1). The models showing deepening MLDs share NEMO as the ocean component, whereas AWI-CM-1–0 has FESOM as its ocean component. On the other hand, even the models with NEMO3.6 as their ocean component (compare HadGEM3-GC31, CNRM-CM6.1, CMCC-CM2 and EC-Earth3P) differ considerably. This suggests that either the different atmospheric or sea ice components or the choice of ocean parameters have a strong influence on the convection in the different NEMO models.

In contrast, the MLD differs little when only the atmosphere resolution is increased (compare ECMWF-IFS-MR with ECMWF-IFS-HR, HadGEM3-GC31-MM with HadGEM3-GC31-HM, and CCCM-CM2-HR4 with CCCM-CM2-VHR4). An exception is MPI-ESM1-2, where an increased atmospheric resolution reduces the MLD. This MLD reduction can be linked to wind forcing, which is too weak in MPI-ESM1-2-XR (Putrasahan et al. 2019).

The place of deepest convection in the Labrador Sea varies somewhat across models, but we do not find any clear linkage to increasing ocean or atmosphere resolution.

To investigate the impact of natural variability on the mean March MLD in the historical period and to quantify the potential contribution of natural variations to the differences in MLD with changing resolution, we use an ensemble of historical simulations with the ECMWF-IFS model. The MLD in the low-resolution version ECMWF-IFS-LR is very shallow in all six ensemble members and there is no deep convection in the historical and control simulations (not shown). Thus, we concentrate in the following on the four members of ECMWF-IFS-HR, which all exhibit pronounced deep mixing in the Labrador Sea. These four ECMWF-IFS-HR members show considerable differences which reflect the impact of natural variability (Fig. 2a–d). The averaged March MLD (1950–2014) deviates in individual ensemble members by up to 200 m from the ensemble mean MLD. Although, this is a considerable amount, given the relatively long averaging period, the MLD differences between NEMO models with 1° and 0.25° resolution are much larger (compare Fig. 2 to differences between 1° and 0.25° simulations in Fig. 1). Even though four members are not sufficient to fully capture the complete range of total natural variability, these results suggest that natural variability cannot explain the differences in MLD, which occur between such simulations which have a change in ocean resolution. Finally, despite the variability across the ensemble members in
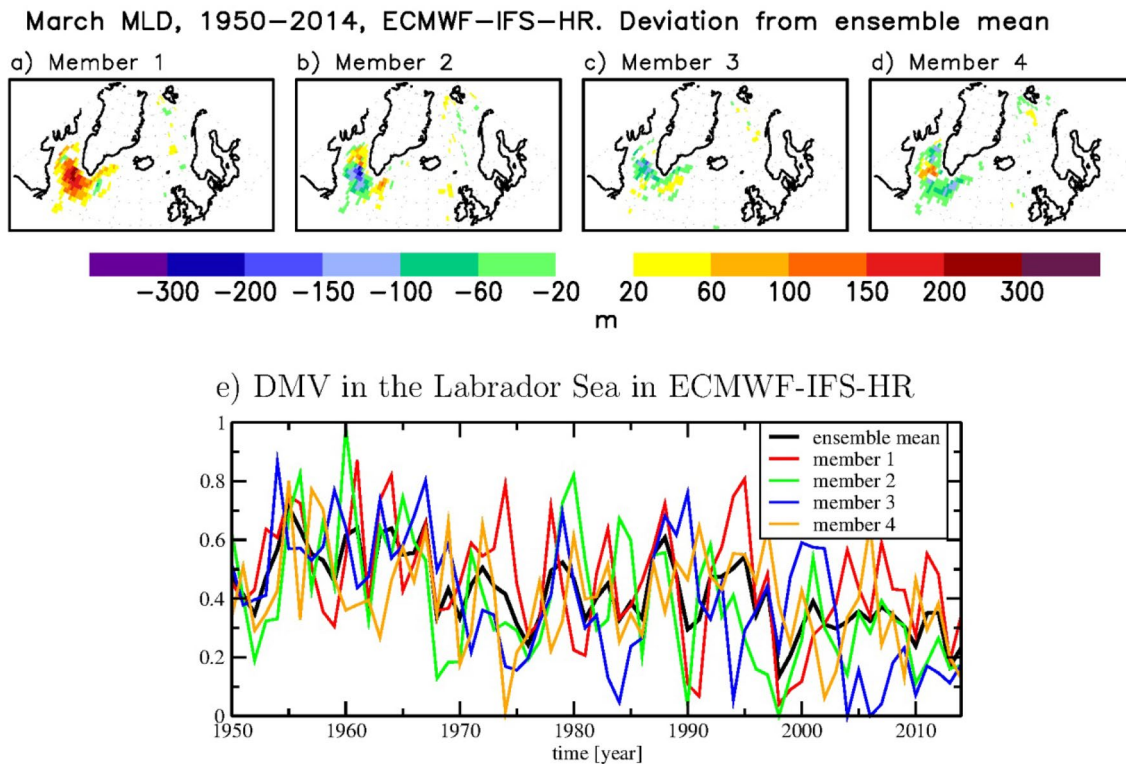
Fig. 2 **a–d** Deviation of mixed layer depth in March in the ensemble members of ECMWF-IFS-HR from the ensemble mean of the four ECMWF-IFS-HR simulations for the time period 1950–2014. **e** DMV in the Labrador Sea (in 10^{15} m^3) in the ensemble mean and single ensemble members of ECMWF-IFS-HR

Fig. 2a–d, the general amplitude, frequency and trends of the DMV are similar in Fig. 2e.

3.2 Deep mixed volume in the Labrador Sea

In the following, we concentrate on the DMV index for a detailed investigation of deep convection in the Labrador Sea. Figure 3 shows the DMV in the Labrador Sea in March in the historical model simulations. In agreement with Fig. 1, increasing the ocean resolution from around 1° to 0.25° leads to a generally higher DMV in all models using NEMO (ECMWF-IFS, HadGEM3-GC31, CNRM-CM6-1, EC-Earth3P, see 2nd and 3rd columns in Table 2), whereas the opposite is true for AWI-CM-1–0. Increasing the ocean resolution further to 1/12° in HadGEM3-GC31-HH does not further increase the DMV. The DMV varies strongly among models: ECMWF-IFS-LR does not show any deep convection events in the entire historical period, CNRM-CM6.1 and EC-Earth3P simulate only a few events with deep convection and AWI-CM-1–0-LR and both CMCC-CM2 versions simulate strong deep convection every winter. As for the MLD, most models overestimate the DMV compared to EN4 (Fig. 3h).

Table 2 compares the average DMV in the historical model simulations with that of ARGO and EN4 for the entire period 1950–2014 and for the period 2000–2014. Generally, the simulated DMV in the Labrador Sea is smaller in 2000–2014 compared to the entire period in the models. On the other hand, natural variability of the DMV is high and thus a 15-year period is short for a robust comparison. The DMV in EN4, averaged over 1950–2014, is around 60% higher than in ARGO in 2000–2014. It is almost 2.5 times larger in EN4 if comparing years 2000–2014 in both EN4 and ARGO. The large difference between ARGO and EN4 is likely related to uncertainties in both data sets (see discussion in Sects. 2.2 and 2.3). EC-Earth3P and ECMWF-IFS-LR show no or rather little deep convection in the Labrador Sea while the other simulations overestimate ARGO with factors of 4 to almost 40 and overestimate EN4 with 2.5 to 25 in the period 1950–2014. If only considering the period 2000–2014 in both models and observations, the overestimation is smaller, however, still 13 out of 17 model versions simulate too large DMV, and all high-resolution models overestimate DMV except for EC-Earth3P-HR and

Fig. 3 Deep Mixed Volume (DMV) in 10^{15} m^3 using a critical depth of 1000 m in the Labrador Sea in March in the models (**a–g**) and in EN4 (**h**) between 1950–2014. For ECMWF-IFS, only member 1 is shown for better visual comparison of the variability across resolutions
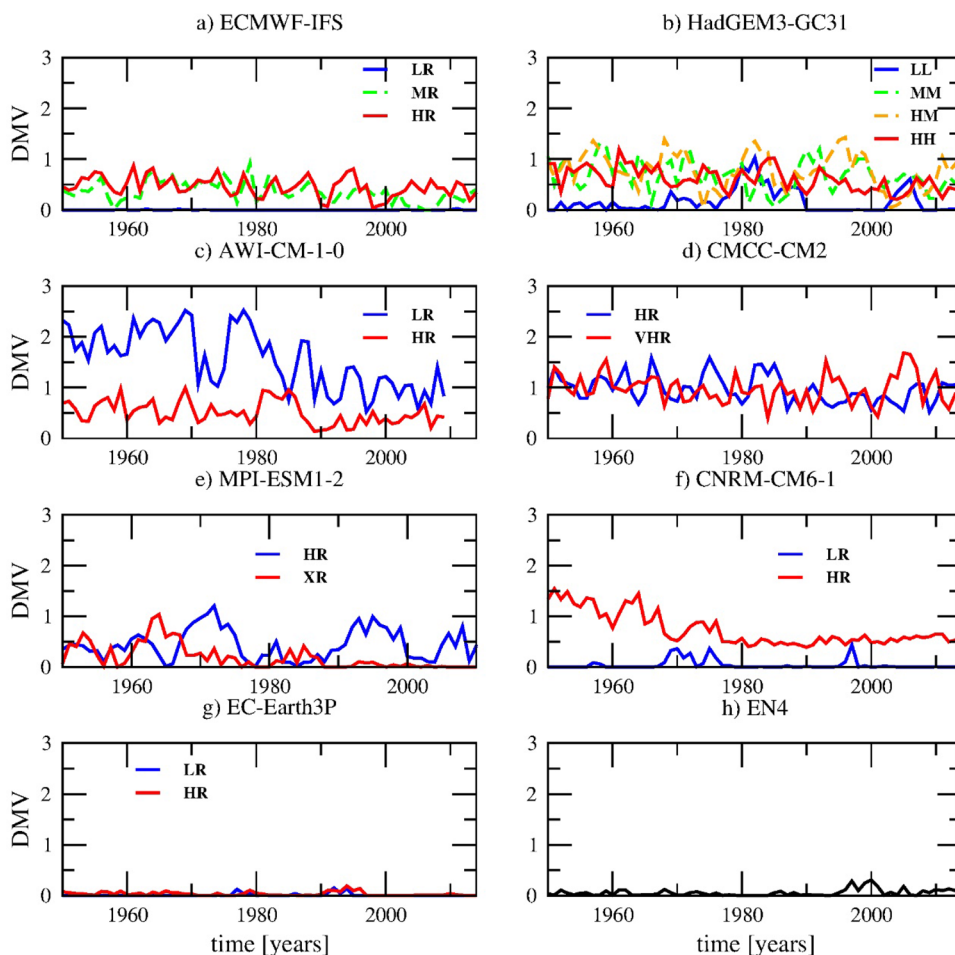
MPI-ESM1-2-XR. Despite the uncertainties in the comparison to the observations, it is clear that the models have problems in realistically simulating the convection in the Labrador Sea. If deep convection occurs in the models, the ocean is often mixed down to the bottom in the models, whereas deep convection rarely exceeds 2000 m in the observations (Yashayaev and Loder 2016, 2017).

If we use a critical depth of $z_{crit} = 0$ m instead of 1000 m in the Labrador Sea and thus consider the total mixed layer depth, the relative deviation of the DMV in the models from the observations is reduced as expected (not shown). However, AWI-CM-1–0-LR and CMCC-CM2 still overestimate the observed DMV (with $z_{crit} = 0$) by a factor of around three and two, respectively. On the other hand, ECMWF-IFS-LR simulates only around 20% of the mixed volume compared to the observations. The comparison between z_{crit0} and $z_{crit1000}$ reveals some non-linearites in the deep convection. While CNRM-CM6.1-h has a nine times higher DMV ($z_{crit1000}$) compared to ARGO, it is only 16% higher for z_{crit0}, whereas the DMV ($z_{crit1000}$) for MPI-ESM1-2-XR is 4.6 times higher compared to ARGO but 14% smaller for z_{crit0}.

3.3 Historical trends of DMV in the Labrador Sea

Ten of 17 simulations indicate a significantly negative trend of the DMV in the historical period (Fig. 3, Table 3). To investigate whether this trend is due to external forcing or due to model drift, we compare the DMV in the historical simulations with that from the 100-year 1950-control simulations (Fig. 4 and Table 3). Most of the control simulations do not show any significant trend, and in 9 out of 17 historical simulations, the DMV trends in the historical simulations are significantly more negative compared to the first 65 years of the control simulations. This indicates external forcing as a major cause for the DMV reduction in the Labrador Sea in these historical simulations. Furthermore, the trends become more negative with higher ocean resolution. In all the models with NEMO as ocean component, the negative trends are stronger in the model versions with higher ocean resolution. In the AWI-CM1-model, both high- and low-resolution model versions show negative trends, however, the control runs show even larger negative trends and thus the negative trends in the historical simulations cannot be attributed to the changes in external forcing.

Table 2 Observed and modeled DMV and SHF in the Labrador Sea, the ratio between model and observed values and correlations between SHF and DMV. Row 2: DMV and SHF in observations, shown are absolute values, ARGO for 2000–2014, EN4 for 1950–2014 (2000–2014 in brackets). Rows 3–9: Ratio of modeled and observed DMV and SHF (Model values divided through observational values; DMV_{model}/DMV_{obs} and SHF_{model}/SHF_{obs}). For ECMWF-IFS, ensemble means are shown. For the DMV (column 2), the first number compares the mean of the entire historical simulation (1950–2014) to ARGO, the number in brackets only the years 2000–2014 of the model simulations. Column 3 compares DMV in the models to EN4 for the entire period 1950–2014 (in brackets 2000–2014). Column 5 (6) shows the correlation between winter SHF (NAO) and March DMV in the Labrador Sea. For the correlations, z_{crit0} has been used to avoid complications with periods without any deep convection

	$DMV_{ratio-ARGO}$	$DMV_{ratio-EN4}$	SHF_{ratio}	Correlation SHF—DMV	Correlation NAO—DMV
ARGO or WHOI EN4	3.95e + 13 m^3	6.38e + 13 m^3 (9.7e + 13 m^3)	129.2 W/m^2	Not enough data 0.19 (SHF from WHOI)	Not enough data 0.20 (NAO from ERA5)
ECMWF-IFS-LR	0.02 (0.03)	0.01 (0.01)	0.53	0.62	0.33
ECMWF-IFS-MR	8.9 (4.9)	5.5 (2.0)	1.15	0.71	0.51
ECMWF-IFS-HR	10.7 (9.9)	6.6 (4.0)	1.21	0.59	0.49
HadGEM3-GC31-LL	4.3 (3.2)	2.7 (1.1)	0.98	0.63	0.39
HadGEM3-GC31-MM	17.1 (13.8)	10.7 (5.6)	1.28	0.70	0.66
HadGEM3-GC31-HM	19.6 (12.5)	12.2 (5.1)	1.39	0.64	0.67
HadGEM3-GC31-HH	17.8 (10.7)	11.1 (4.4)	1.48	0.59	0.62
CMCC-CM2-HR4	24.4 (21.0)	15.2 (8.6)	1.22	0.72	0.67
CMCC-CM2-VHR4	24.8 (25.6)	15.4 (10.4)	1.34	0.59	0.58
CNRM-CM6.1	1.09 (0.13)	0.7 (0.05)	1.15	0.53	0.46
CNRM-CM6.1-HR	9.3 (3.2)	5.8 (1.3)	1.18	0.48	0.49
MPI-ESM1-2-HR	10.6 (7.1)	6.6 (2.9)	1.14	0.61	0.56
MPI-ESM1-2-XR	4.6 (0.30)	2.9 (0.1)	0.98	0.64	0.50
AWI-C-1–0-LR	39.5 (24.9)	24.7 (10.2)	No	No	No
AWI-CM-1–0-HR	12.8 (10.1)	7.9 (4.1)	Data	Data	Data
EC-Earth3P	0.26 (0.05)	0.16 (0.02)	0.63	0.72	0.31
EC-Earth3P-HR	0.95 (0.21)	0.59 (0.09)	1.07	0.50	0.56

A reduction of DMV in the historical period would be in line with some recent studies by Brodeau and Koenigk (2016) and Caesar et al. (2018). In contrast, the trend in the EN4 data for the same period, is significantly positive. However, the trend in EN4 data might be affected by the varying amount and quality of observational data, which are used to produce EN4. For example, from year 2000 onwards ARGO data were used in EN4 but were not available before, potentially making it harder to detect deep convection events in the earlier period.

3.4 Discussion of differences in the mean DMV between models

Figure 5 summarizes the results from Sect. 3.2 on the resolution dependency of the DMV in the Labrador Sea across the models. Each single model shows a clear dependence of the DMV on the oceanic resolution. The differences between models are large, and as discussed before, even models using the same NEMO ocean model exhibit a wide range of solutions. Models with coarse resolution (~100 km) in the ocean produce no or only shallow convection. Models with a resolution of 50 km and higher in the ocean, however, overestimate deep convection compared to ARGO and EN4.

Increasing the atmosphere resolution has a minor effect on the DMV in the Labrador Sea, except for MPI-ESM1.2 and AWI-CM-1–0, where DMV is reduced with increased resolution.

The strength of the deep convection in March is related to the vertical density distribution in the Labrador Sea. We analyse therefore the November density profiles in the Labrador Sea to discuss differences between models (Figs. 6 and 7a, c). We choose November since these profiles represent the stratification prior to the deep winter convection. In November, all models show a near surface layer of low density, mainly due to a combination of low salinity and relatively (compared to late winter) warm water near the surface (Fig. 7). In general, models with lower ocean resolution already show stronger stratification in the upper ocean in November than models with higher resolution (except for AWI-CM-1–0), which is consistent with their comparatively weaker DMV in late winter. This enhanced stratification is caused by colder and fresher surface water masses compared to models with high DMV (Figs. 7d, h). The two model simulations, which do not simulate any deep convection,

Table 3 Trends in the DMV in the Labrador Sea in EN4, the historical simulations and in the first 65 years of the 1950-control simulations and their differences in 10^{12} m^3/year. Trends that are significantly different from 0 at the 95%-confidence level are shown in italic, trends significantly different to the control-runs are bold, and trends significantly different to both 0 and the control-run are italic and bold

Model	DMV- trend hist, 1950–2014	DMV- trend ctrl-1950, year 1–65	Trend— difference: hist—ctrl
EN4	*1.6*		
ECMWF-IFS-LR	0.003	− 0.014	− 0.02
ECMWF-IFS-MR	***− 3.93***	− 0.54	− 3.39
ECMWF-IFS-HR	***− 3.40***	− 0.43	− 2.97
HadGEM3-GC31-LL	0.41	1.58	− 1.17
HadGEM3-GC31-MM	***− 4.43***	− 1.11	− 3.32
HadGEM3-GC31-HM	***− 5.13***	− 3.02	− 2.11
HadGEM3-GC31-HH	***− 6.66***	− 2.28	− 4.38
MPI-ESM1-2-HR	− 1.41	− 0.62	− 0.79
MPI-ESM1-2-XR	***− 7.94***	− 0.06	− 7.88
CMCC-CM2-HR4	***− 5.19***	− 1.08	− 4.11
CMCC-CM2-VHR4	− 1.42	− 3.27	1.85
CNRM-CM6.1	***− 0.54***	0.16	− 0.70
CNRM-CM6.1-HR	***− 6.91***	− 0.31	− 6.60
AWI-CM-1–0-LR	*− 15.0*	*− 16.8*	1.80
AWI-CM-1–0-HR	− 0.96	− 4.98	4.02
EC-Earth3P	0.17	0	0.17
EC-Earth3P-HR	***− 0.39***	*0.922*	*− 1.31*

ECMWF-IFS-LR and EC-Earth3P, show particularly strong density gradients in the upper ocean in November (Fig. 7c). Consequently, a large buoyancy flux would be necessary in winter to initiate deep convection in these two models. Models, which are warmer and saltier in the Labrador Sea in November are generally denser than models which are cooler and fresher. This relationship agrees with findings for CMIP5 models by Menary et al. (2015).

The density profiles of the high-resolution ocean models agree well with ARGO and relatively well with EN4, which show a slightly lower density near the surface than ARGO due to lower salinity (Fig. 7h). The high-resolution models provide a realistic state of both T and S at the surface in the Labrador Sea both in November and March (Figs. 7a, b, d, e). We hypothesize that they transport more warm and salty water masses into the Labrador Sea compared to the lower resolution models (especially than the

NEMO-models using the ORCA1-grid). This hypothesis is supported by recent results from Jackson et al. (2020). Furthermore, we find a clear relationship between the Sub-Polar Gyre (SPG) strength index (computed as the minimum (maximum absolute value) of the barotropic streamfunction between 65°–40° W at 53° N as in Danabasoglu et al. (2014)) and the DMV across the models. A stronger Sub-Polar Gyre in the high-resolution models is related to a stronger northward transport of warm and salty water masses into the Labrador Sea.

Although most of the high-resolution models simulate realistic S and T properties in the upper Labrador Sea, the DMV differs strongly between the models and is substantially higher than in ARGO and EN4. Vertical gradients across the upper 200 m are slightly larger in most high-resolution models than in the observations (Fig. 7c), which would suggest reduced convection compared to ARGO. However, this reduction in convection might be compensated in these models by an excessively shallow surface layer with low density (Fig. 6). This shallower surface layer requires less heat loss to be eroded in these models and could contribute to the overestimation of the deep convection in late winter.

Figure 7 k shows, as expected, a strong dependence of the models' DMV on winter surface heat fluxes (SHF) in the Labrador Sea. The correlation between March DMV and January-March SHF across models reaches 0.76. Figure 7 k shows in addition that all high-resolution models (except for MPI-ESM-1–2-XR) overestimate (meaning too strong fluxes from the ocean to the atmosphere) the SHF compared to WHOI-OAFlux. This overestimation likely contributes to stronger buoyancy loss in the winter and too strong convection in March in the high-resolution models. Models with very large DMV (e.g. both CMCC-CM2 model versions) have rather weak stratification in November as well as strong surface heat fluxes in winter.

While our analysis above explains parts of the differences in the DMV in the different models, it does not explain the reasons for the different biases in the models. Conditions in the Labrador Sea are affected by both large scale ocean (e.g. advection of water masses into the Labrador Sea, position of the North Atlantic Current and its sub-branches, extension of the sub-polar gyre) and atmosphere processes (e.g. exact position, strength and variability of the North Atlantic Oscillation pattern) and descriptions of local processes (e.g. different parameterizations or parameters for mixing and lateral and vertical diffusion processes or sea ice processes). Since the models differ in many aspects in their atmosphere, ocean and sea ice components and partly in their initialization method, it is not possible to identify any specific reason or parameterization

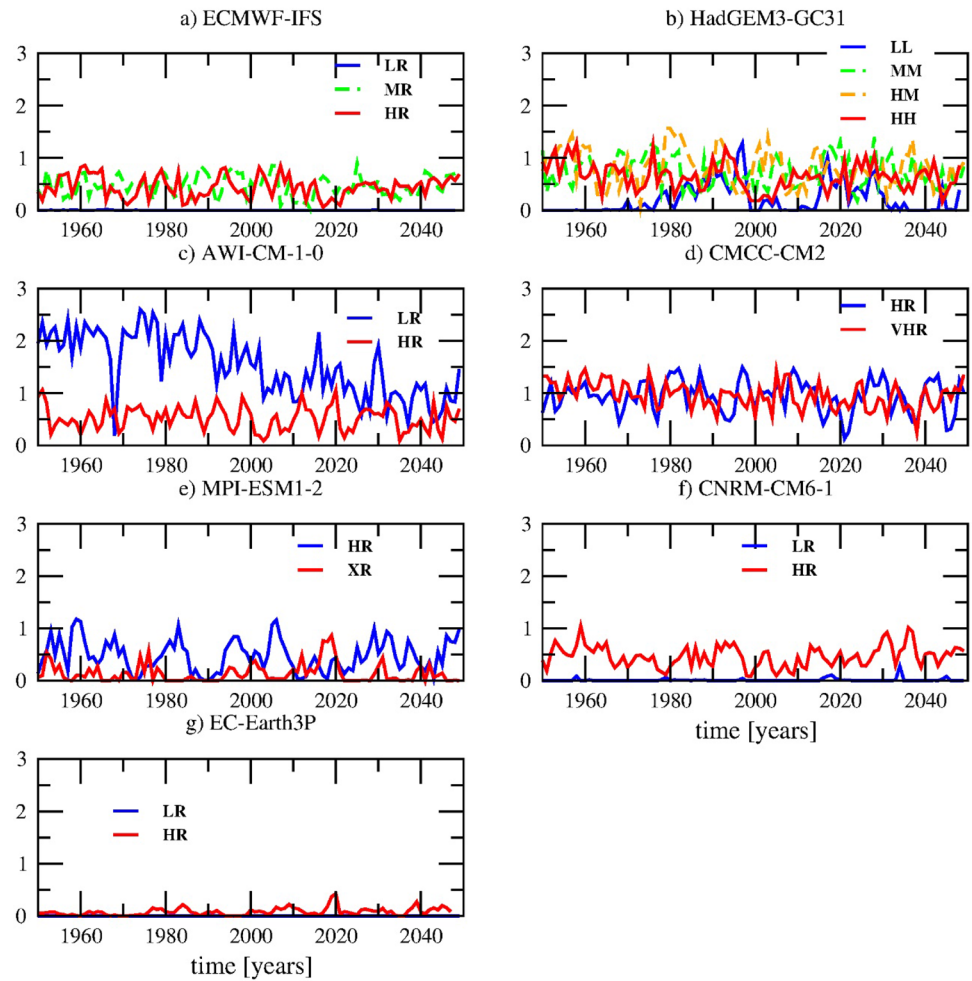**Fig. 4** As Fig. 3 but for the 100-year control simulation

a) ECMWF-IFS

b) HadGEM3-GC31

c) AWI-CM-1-0

d) CMCC-CM2

e) MPI-ESM1-2

f) CNRM-CM6-1

g) EC-Earth3P

time [years]

Fig. 5 DMV (average over 1950–2014 in 10^{12} m^3) in the Labrador Sea in March as a function of the oceanic (left) and atmospheric (right) resolution. Thin lines connect the DMV in simulations with the same model but with different resolution. The black, horizontal lines indicate average DMV values from ARGO (solid) and EN4 (dashed)

DMV Labrador Sea

DMV Labrador Sea
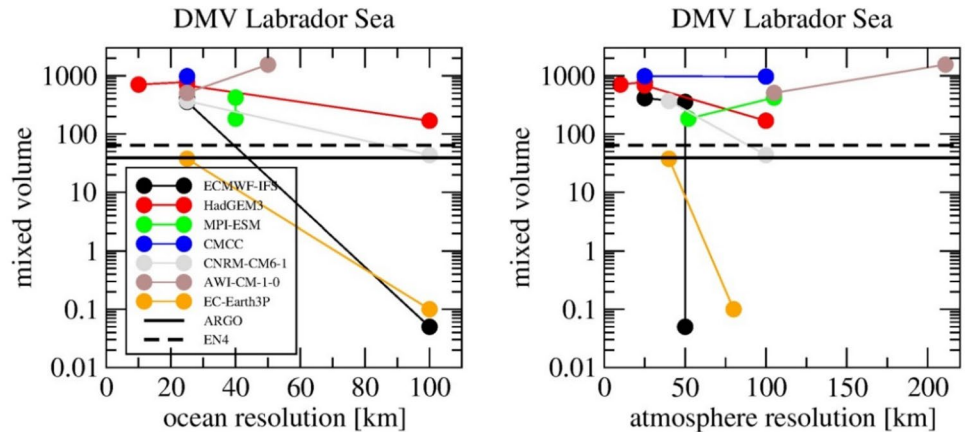
responsible for the differences in S, T and SHF between models as part of this multi-model study. However, representation of mixing processes in ocean models is imperfect and contributes both to the biases in the DMV and to the DMV-differences between models and must be given particular attention in future model developments.

4 Variability of DMV in the Labrador Sea and driving processes

The interannual variability of the DMV is large in all models (Fig. 3). Some of the simulations (EC-Earth3P, HadGEM3-GC31-LL, MPI-ESM1-2, CNRM-CM6.1 and
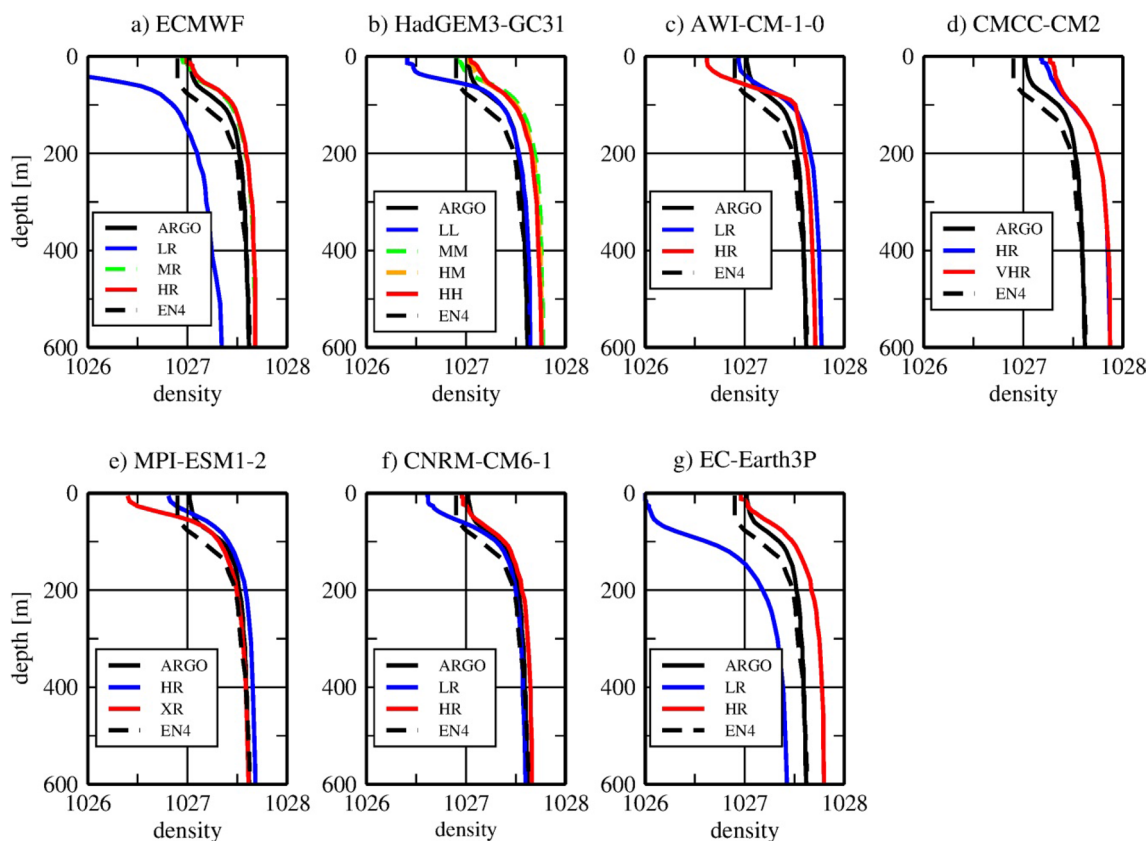
Fig. 6 Density (in kg/m^3) in the upper 600 m averaged over the Labrador Sea in November

ECMWF-IFS-MR) indicate substantial variability at decadal or longer periods, in which phases with and without convection alternate. Such intermittent deep convection was also suggested from observations (Lazier et al. 2002; Yashayaev and Loder 2016) and is visible in the EN4-data as well. However, in most model simulations, deep convection occurs in almost every winter or not at all.

Deep convection depends strongly on the buoyancy of the ocean surface layer in the convection regions, which in turn depends (amongst other things, e.g. advection of heat) on the heat loss to the atmosphere and the influx of fresh water into the convection regions.

4.1 The impact of surface heat fluxes on the variability of deep convection

Brodeau and Koenigk (2016) showed that the turbulent surface heat flux (SHF) is the main driver for interannual variability in the DMV and Fig. 7 k shows that the mean SHF in the Labrador Sea is highly correlated with the mean DMV across models. Thus, we will focus first on the effect of varying SHF on the DMV in the Labrador Sea.

Figure 8 shows the winter (January, February, March) SHF in each of the model simulations. Observations from

WHOI-OAFlux show that the largest SHF of more than 200 W/m^2 is found near the ice edge in the Labrador Sea, with regions of heat loss extending to the southern part of the SPG, south of Iceland and along the southeast coast of Greenland, as well as in the northern Norwegian-Greenland Seas and Barents Sea. The large-scale features of this pattern are reproduced by most of the models. ECMWF-IFS-LR and to a lesser degree EC-Earth3P, however, which both simulate too weak convection, strongly underestimate the SHF in the Labrador Sea. Increased ocean resolution generally improves the representation of the observed SHF pattern. In particular, the extension of high SHF from the Labrador Sea into the southwestern branch of the SPG and the high SHF in the northern Greenland and Norwegian Seas are better simulated. However, a number of models (both CMCC-CM2 versions, HadGEM3-GC31-HH and CNRM-CM6.1) overestimate the SHF in the SPG. In addition, the SHF west and northwest of Scotland is too high in most of the models.

In the Labrador Sea, all high-resolution models with NEMO as the ocean component simulate increased SHF (averaged over the same box as used for the calculation of the DMV) compared to their lower-resolution counterparts (Table 2). In contrast, MPI-ESM1-2 shows reduced SHF with increased atmospheric resolution in line with the
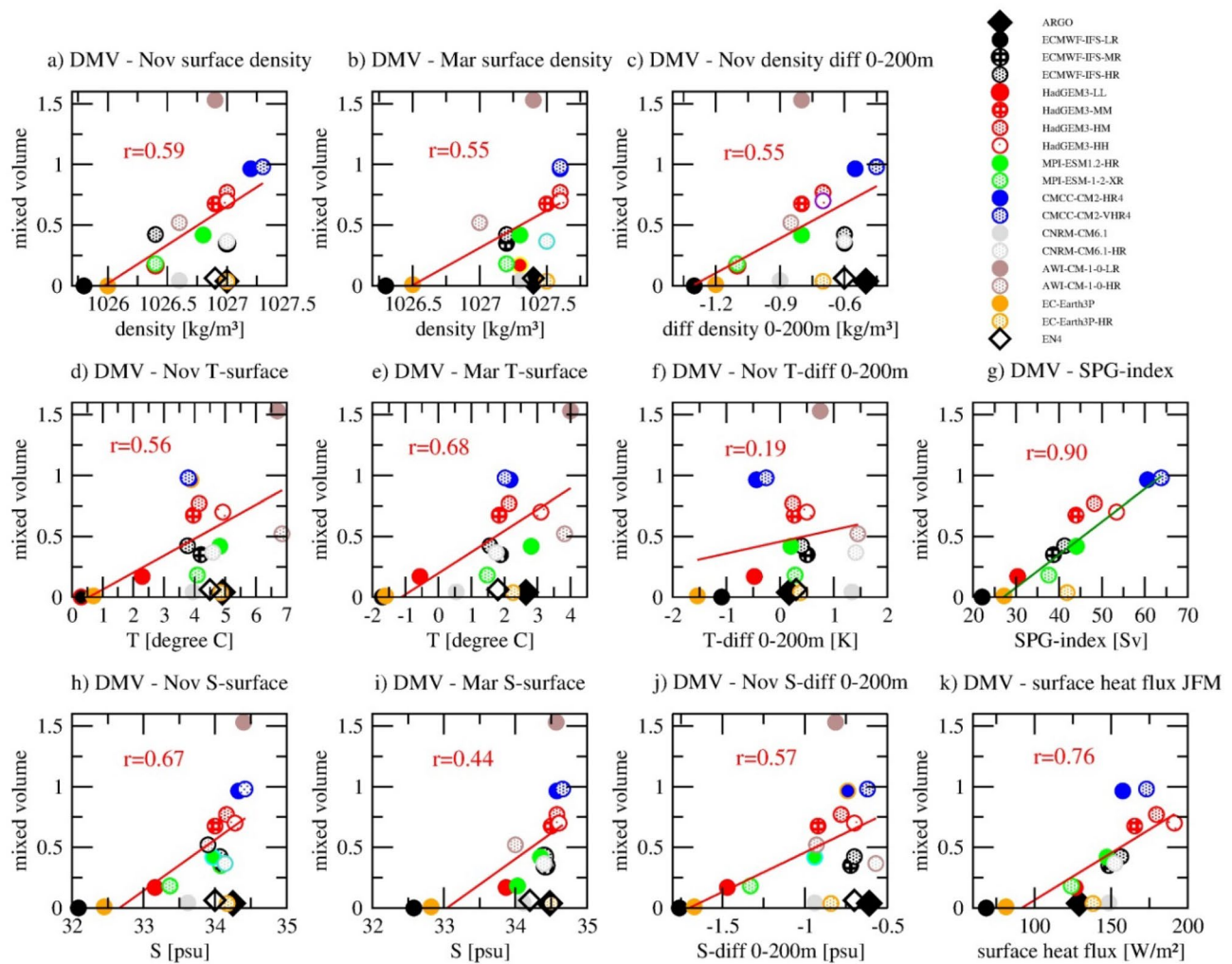
Fig. 7 Relation between DMV in the Labrador Sea in March (in 10^{15} m^3) and surface density (**a**, **b**), surface temperature (**d**, **e**), surface salinity (**h**, **i**) in the Labrador Sea in November and March. **c**, **f**, **j** DMV versus differences between surface and 200 m depth for density, temperature and salinity in November. **g** DMV versus annual mean Sub-Polar Gyre index (SPG-index). AWI-CM-1.0 and CNRM-CM6.1 models are missing due to missing data. **k** DMV versus surface heat flux (positive means flux from the ocean into the atmosphere). AWI-CM-1.0 models are missing due to missing data. Observational data for the SHF are from WHOI-OAFlux Shown are means for the historical simulation 1950–2014. The red line is the linear regression line and the red numbers the correlation coefficient between DMV and the respective variable on the x axis across models

reduced convection. In all models, the interannual variability of winter SHF is significantly positively correlated with the DMV in March. Thus, large ocean heat losses in the winter are linked to strong DMVs in the following March, indicating that large upward surface heat fluxes lead the DMV. The correlation coefficient of SHF and DMV varies from 0.48 in CNRM-CM6.1 to slightly above 0.7 in ECMWF-IFS-MR, EC-Earth3P and CMCC-CM2-HR4. The relation between SHF and DMV is neither resolution nor model dependent.

The winter SHF in the Labrador Sea itself is governed by the atmospheric circulation. In all model simulations north-to-northwesterly winds, which advect cold air towards the Labrador Sea, lead to strong surface heat

fluxes, which can erode the stratification of the ocean and increase convection (Ortega et al. 2011). These north-to-northwesterly winds are linked to the positive phase of the North Atlantic Oscillation (NAO, defined as the leading EOF of geopotential height on the 500 hPa pressure surface over the European/Atlantic sector (80° W–40° E, 20–90° N)). The correlation between NAO and SHF is high in all models except for ECMWF-IFS-LR and EC-Earth3P (Fig. 9) and resemble well the observed correlation pattern between NAO and SHF (Fig. 9 a). In ECMWF-IFS-LR and (to a smaller extent) EC-Earth3P, sea ice extends far to the southeast and covers large part of the Labrador Sea (compare Docquier et al. 2019), and
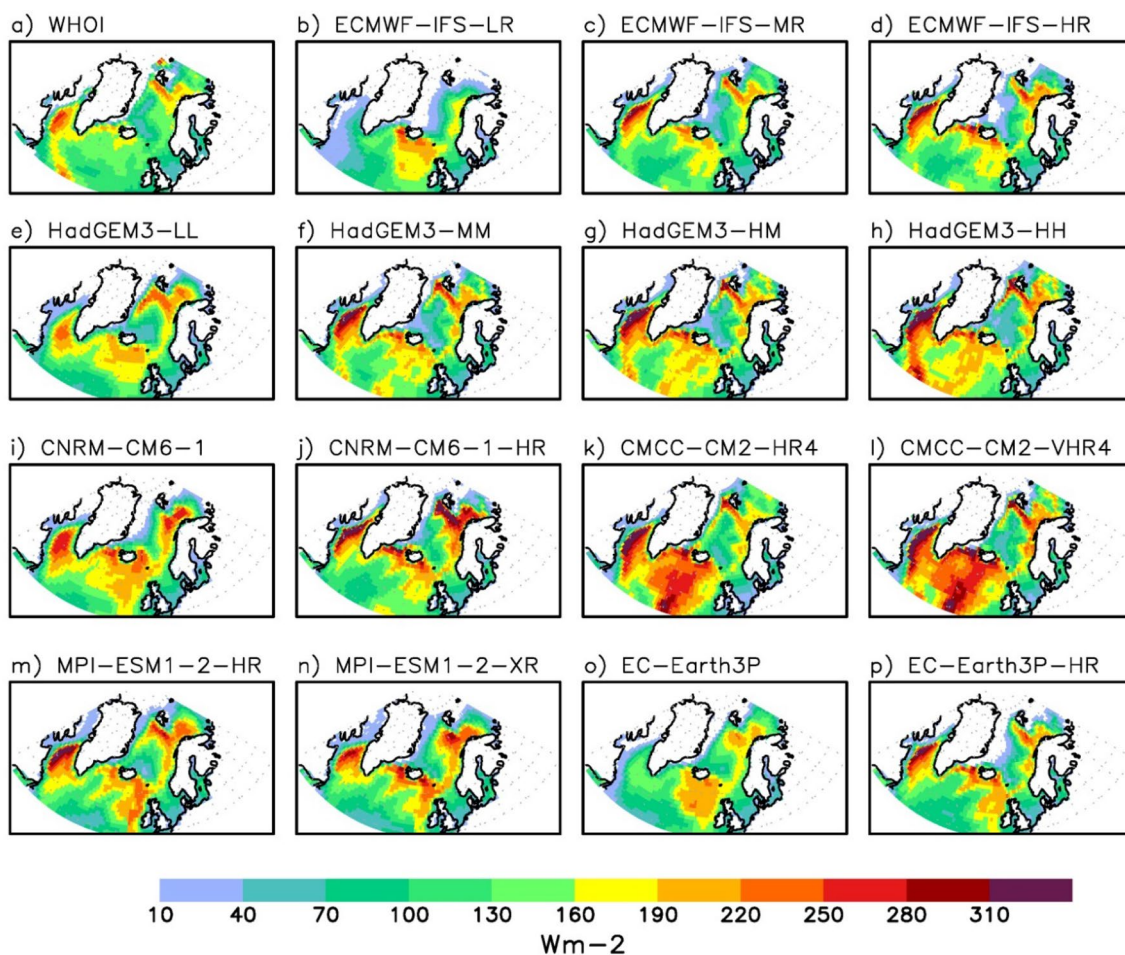
Fig. 8 Turbulent surface heat flux (January, February, March average) in 1950–2014 in the WHOI-OAFlux data and in the model simulations. Positive values mean flux from the ocean to the atmosphere

thus, the correlations between NAO and SHF are smaller in the Labrador Sea. The NAO-index itself, is significantly positively correlated with the DMV in the Labrador Sea in all simulations with values between 0.31 (EC-Earth3P) and 0.67 (HadGEM-GC31-HM and CMCC-CM2-HR4) (Table 2). The models with an ocean resolution around 1° (ECMWF-IFS-LR, EC-Earth3P, HadGEM-GC31-LL, CNRM-CM6-1) show a weaker correlation between NAO and DMV in the Labrador Sea than the models with higher resolution. Interestingly, this is not true for the correlation between SHF and DMV.

The rather high correlation between NAO and both SHF and DMV reveals the importance of the atmospheric circulation for deep convection in the models, however, increased DMV due to strong SHF might feed back on the SHF by continuously bringing warm water to the surface and causing more heat loss.

The spatial imprint of the NAO-index on the 500-hpa geopotential height is shown in Fig. 10. All models reproduce the NAO-pattern of the ERA5-reanalysis data (Hersbach et al. 2020) well. However, the position of the negative pole over Iceland-Greenland and the extension of the positive pole towards Eurasia vary slightly among models. We do not see a systematical difference between the NAO-pattern in the low resolution and the high-resolution models, which could explain why the correlation between NAO and SHF differ between high- and low-resolution models.

4.2 The impact of Arctic freshwater and sea ice exports on the variability of deep convection

A number of studies discussed the effect of Arctic freshwater export as a potential source of variability of the deep water convection in the Labrador Sea (Holland et al. 2001;
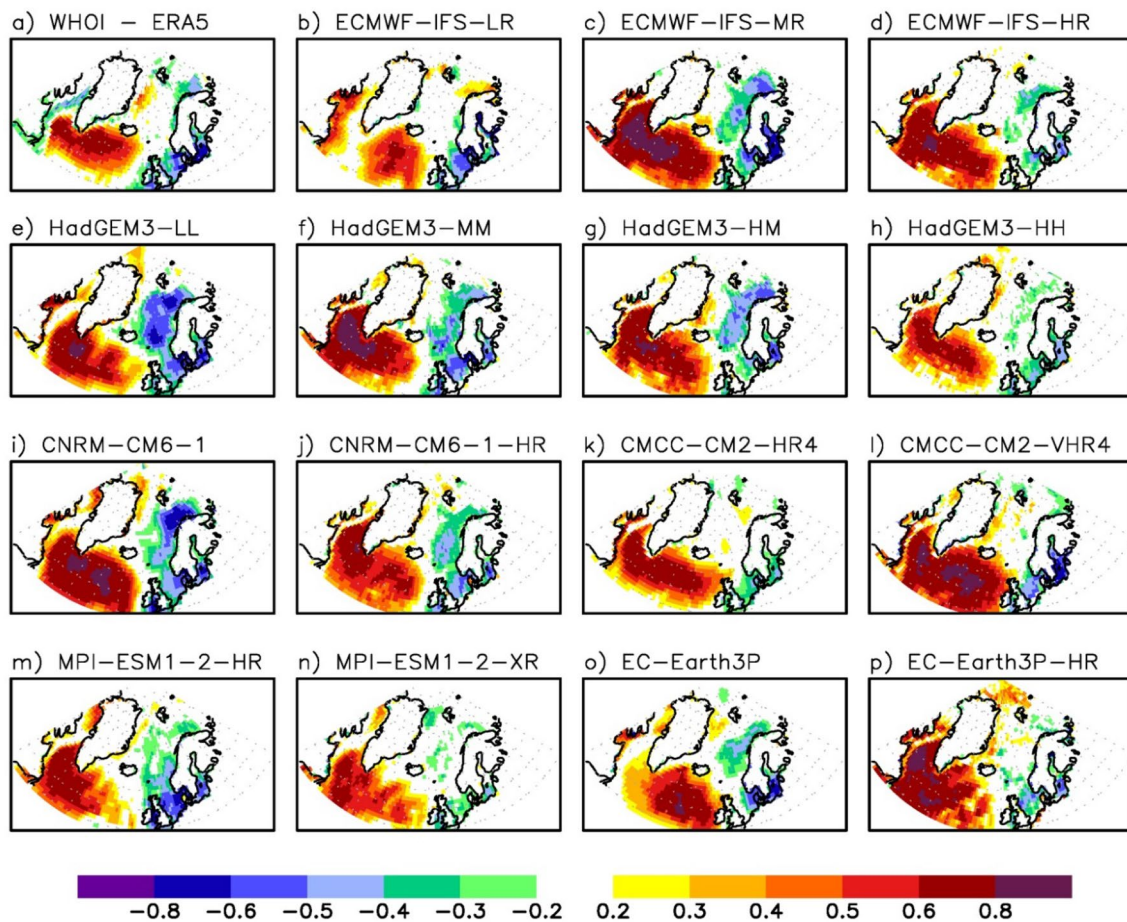
Fig. 9 Correlation between winter NAO-index and turbulent surface heat flux. The periods used were 1979–2013 for correlation between WHOI-OAFlux surface heat flux and NAO from ERA5 data (**a**) and 1950–2014 for the models (**b**–**p**)

Jungclaus et al. 2005; Koenigk et al. 2006). Here, we analyze the correlations between freshwater transports across different sections (Fram Strait, Denmark Strait, northern Baffin Bay) and deep convection in the Labrador Sea in the historical simulations of the models.

Table 4 shows the freshwater exports out of the Arctic Ocean into the North Atlantic through Fram Strait, Baffin Bay and the Denmark Strait. Although differences between models are large, the exports through Fram Strait are generally larger than through Baffin Bay. The total freshwater exports through Fram Strait (liquid + solid export) vary between around 80,000 m³/s in the two CNRM-CM6.1 models and 160,000 m³/s in ECMWF-IFS-LR and HadGEM3-GC31-MM. The ratio between liquid and solid freshwater export through Fram Strait differs strongly in the simulations. While in HadGEM3-GC31-LL and all ECMWF-IFS and EC-Earth3P simulations most of the freshwater leaves the Arctic in the form of sea ice, the liquid and solid fractions are of similar size in the other models. In CNRM-CM6.1-h, the liquid fraction is even larger than the solid fraction. The amount of total freshwater passing

through Denmark Strait is smaller than the export through the Fram Strait in all models except for ECMWF-IFS-LR and MPI-ESM1.2-XR. The liquid part is dominant since large parts of the ice melt in the East Greenland Current on its way from Fram Strait to Denmark Strait leading to large differences in the solid components between these locations.

The low-resolution versions of ECMWF-IFS, HadGEM3-GC31, CNRM-CM6.1 show a larger fraction of solid freshwater exports through Fram Strait and larger liquid transports through Baffin Bay (EC-Earth3P as well for Baffin Bay) compared to their higher resolution counterparts. The sum of freshwater exports through Fram Strait and Baffin Bay differs more between the models than between different versions of single models. Despite the large differences in mean Arctic freshwater exports into the North Atlantic, there is generally no clear linkage to the mean DMV in the Labrador Sea. However, for ECMWF-IFS-LR, we speculate that the very large freshwater fluxes, particularly in the form of sea ice, through Denmark Strait contribute to the low surface density in the Labrador Sea (compare Figs. 6, 7) and consequently to the suppression of deep convection in the Labrador Sea.
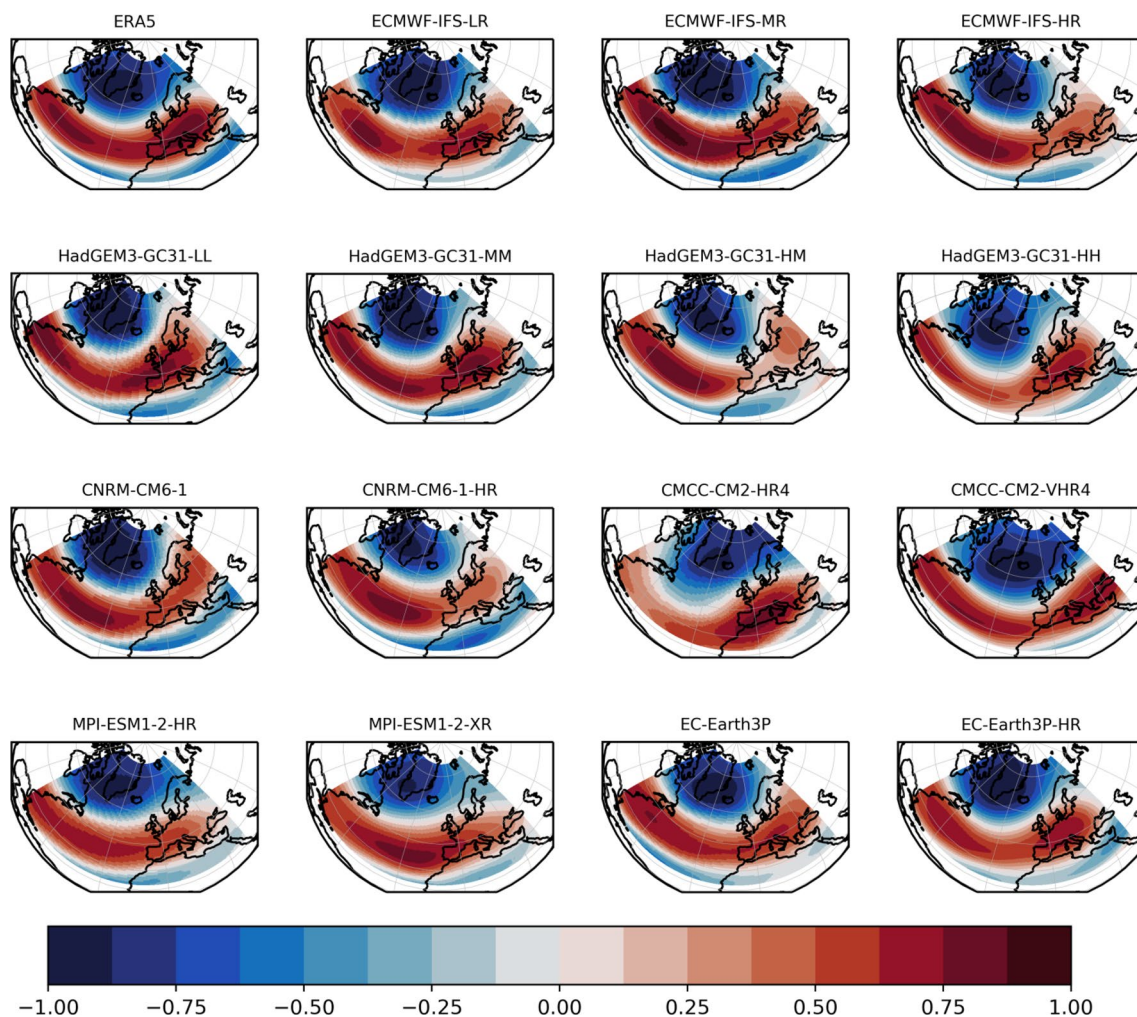
Fig. 10 Correlation between geopotential height at 500 hPa and North Atlantic Oscillation (NAO) index during winter (JFM mean) for ERA5 and the models. The periods used were 1979–2019 for ERA5 and 1950–2014 for the models

To further investigate whether the variability of freshwater exports affects deep convection in the Labrador Sea, we correlate the solid and liquid transports across all sections with the DMV. In all model simulations, the annual mean southward transport of both liquid and solid freshwater across Fram Strait and the liquid transport across Denmark Strait are weakly negatively correlated with the deep convection in the Labrador Sea in March (ranging between −0.1 and −0.4). The highest correlation is reached when the freshwater transport through Fram Strait (and Denmark Strait) leads the convection by one to two years (and zero to one year). Increased southward transport of sea ice and liquid freshwater through Fram Strait along Greenland's east coast and through Denmark Strait leads to more freshwater input into the Labrador Sea, which tends to freshen the surface ocean, increase stratification and reduce the convection. Figure 11 shows for the two model simulations with the highest correlation between freshwater transport through Denmark Strait and DMV in

the Labrador Sea (HadGEM-GC31-LL, EC-Earth3P-HR) that increased freshwater transport reduce the MLD in the Labrador Sea. For most other model simulations, the effect of freshwater transports on the MLD is rather small.

In some models, the southward transport of liquid freshwater through Baffin Bay is positively correlated with the deep convection in the Labrador Sea (up to r = 0.35 in HadGEM3-GC31-LL). This may seem counterintuitive, but while northerly winds in the Baffin Bay cause strong SHF in the Labrador Sea and dominate the convective conditions, they simultaneously lead to increased freshwater transports through Baffin Bay.

We do not find any resolution dependency of the correlation between freshwater exports and convection in the Labrador Sea. This result is in contrast to a recent study from Fuentes Franco and Koenigk (2019) where they analyzed a set of HadGEM3-GC2 simulations at different resolutions and found larger correlations with increased resolution.

Table 4 Liquid and solid freshwater fluxes through Fram Strait, Denmark Strait and northern Baffin Bay in 10^4 m^3/s in the historical simulations averaged over 1950–2014. Positive values mean freshwater exports out of the Arctic. The last column shows the sum of liquid and solid exports through Fram Strait and Baffin Bay. No data were available to calculate transports in the AWI-CM-1.0 model, liquid freshwater transports in HadGEM3-GC3-HH and sea ice transports in the CMCC-CM2 simulations

| Freshwater fluxes in 10^4 m^3/s	Fram Strait liquid	Fram Strait solid	Denmark Strait liquid	Denmark Strait solid	North Baffin Bay liquid	North Baffin Bay solid	Sum Fram Strait + Baffin Bay
ECMWF-IFS-LR	1.27	15.40	6.05	9.21	2.17	0.72	19.56
ECMWF-IFS-MR	5.90	8.51	6.97	1.26	1.76	0.70	16.83
ECMWF-IFS-HR	4.76	10.75	6.89	2.25	1.45	0.74	17.69
HadGEM3-GC31-LL	3.04	8.56	2.53	2.85	5.50	0.70	17.81
HadGEM3-GC31-MM	8.12	8.22	6.95	2.19	1.99	0.66	19.00
HadGEM3-GC31-HM	7.23	7.35	6.18	1.39	1.87	1.03	17.48
HadGEM3-GC31-HH	No data	5.70	No data	1.28	No data	1.71	No data
CMCC-CM2-HR4	7.09	No data	6.05	No data	1.49	No data	No data
CMCC-CM2-VHR4	6.11	No data	1.38	No data	0.62	No data	No data
CNRM-CM6.1	3.97	4.39	3.33	0.87	3.51	0.39	12.26
CNRM-CM6.1-HR	5.23	2.92	4.13	1.54	2.23	0.85	11.24
MPI-ESM1-2-HR	6.68	5.45	9.27	0.85	0.13	0.22	12.49
MPI-ESM1-2-XR	5.48	6.23	11.20	1.24	0.22	0.31	12.26
EC-Earth3P	2.68	9.62	3.04	2.83	2.53	0.58	15.42
EC-Earth3P-HR	2.21	10.59	5.47	2.74	0.18	0.35	13.33
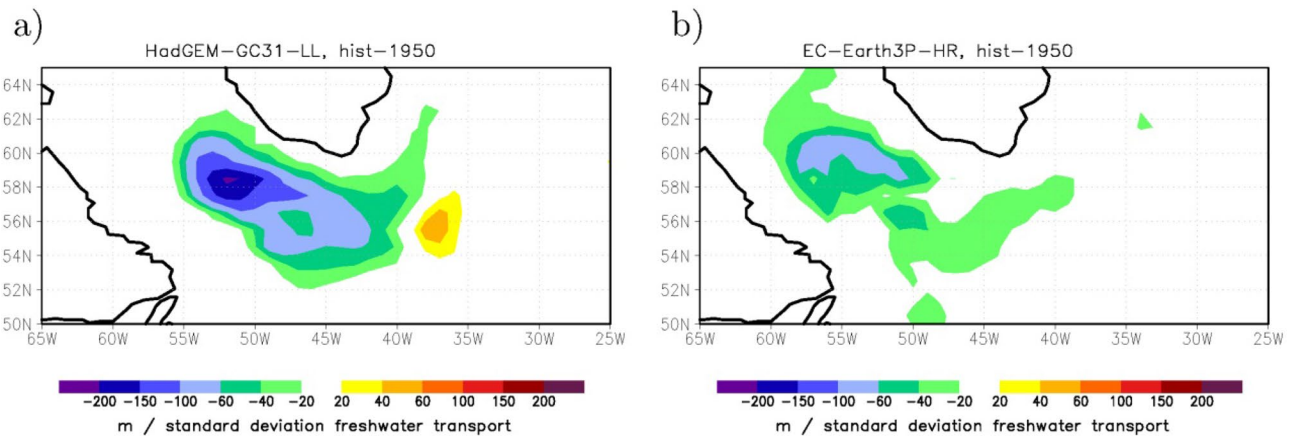
Fig. 11 Regression between annual mean freshwater transport through the Denmark Strait and mixed layer depth in the following March. **a** HadGEM-CG31-LL and **b** EC-Earth3P-HR. Data have been detrended before calculating the regression. These two simulations show the largest correlation between Denmark Strait freshwater transport and DMV in the Labrador Sea (−0.4 and −0.35 for HadGEM-GC31-LL and EC-Earth3P-HR, respectively)

5 The linkage of the DMV to the AMOC

The effect of high resolution on the AMOC in the HighResMIP model simulations has been studied in more detail in Roberts et al. (2020). They found that "the AMOC tends to become stronger as model resolution is enhanced, particularly when the ocean resolution is increased from non-eddying to eddy-present and eddy-rich". Roberts et al. (2020) also analysed the relation between the time-mean values of the DMV and the AMOC in the historical runs. They found that there is a strong relationship between DMV and the AMOC strength across models; models with more deep convection in the Labrador Sea have a stronger AMOC. As shown in our Sect. 3.2, only few models simulate a DMV that is consistent with observed estimates.

However, these models underestimate the AMOC (except for CNRM-CM6-1) compared to the RAPID-MOCHA array observations (Cunningham et al. 2007; Smeed et al. 2004) whereas some of the models (HadGEM3-GC31-MM and -HM, MPI-ESM1.2-h, AWI-CM-1–0-LR) markedly overestimate the DMV in the Labrador Sea but simulate a realistic AMOC. Thus, the linkage between the mean values of the AMOC and the DMV in the models is not consistent with the observations.

To investigate the impact of variability in the deep convection on the variability of the AMOC, we performed cross-correlation analyses between the DMV in Labrador Sea and the AMOC (at 26°N) for lags between −/+ 10 years. In agreement with results by Brodeau and Koenigk (2016), annual values are only rather weakly correlated with each other. We thus focus here on correlations of linearly detrended and 10-year low pass filtered values of DMV and AMOC (Fig. 12) in the 100-year 1950-control simulations. Positive lags mean that the AMOC leads the DMV, negative lags mean that the DMV leads AMOC. Maximum values of the correlations vary between around 0.3 (AWI-CM-1–0-h, ECMWF-IFS-MR, CMCC-CM2-HR, EC-Earth3P-HR) and 0.8 (HadGEM3-GC31-LL, CNRM-CM6-1-h). Most model simulations with significant correlations reach their maximum correlation when the DMV leads the AMOC by 0–5 years. Neither the amplitude of the correlations or the lag of the maximum correlation show any robust resolution dependency. For the HadGEM-GC3.1 model family for example, HadGEM-GC3.1-LL shows a high correlation
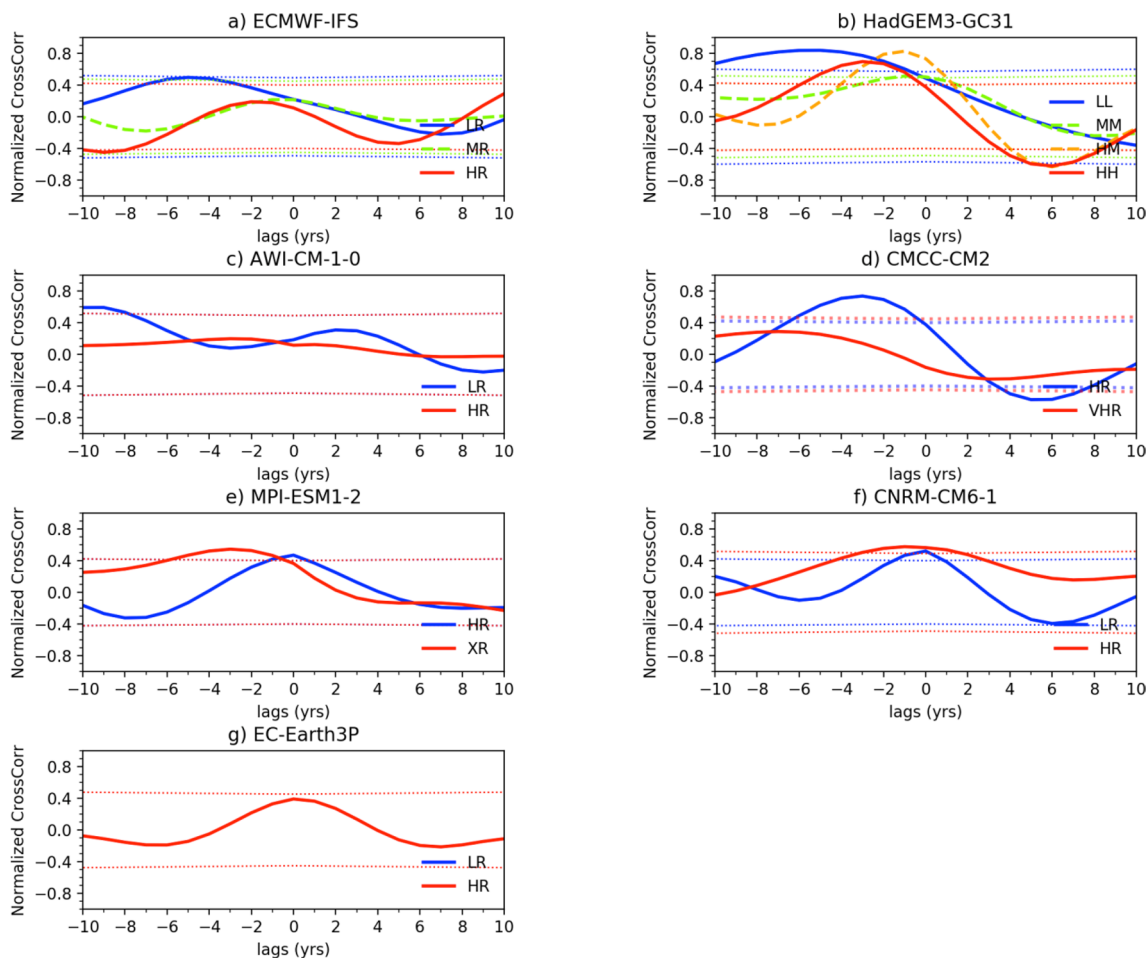

Fig. 12 Correlation between the DMV using a critical depth of 1000 m in the Labrador Sea in March and the AMOC index for the 100-year control simulation. Both timeseries were detrended and filtered with a 10-year low-pass filter. Area enclosed by dotted lines represents the 95% confidence calculated as 2/sqrt(N), where N is the number of independent data based on the time that takes autocorrelation to fall below 1/e. Positive lags mean AMOC leads DMV, negative lags mean DMV leads AMOC. The low-resolution version of EC-Earth3P does not produce any deep convection events in the control simulation

of around 0.8 when the DMV leads by 5 years, HadGEM-GC3.1-MM a correlation of 0.5 at lag 0, HadGEM-GC3.1-HM 0.8 when DMV leads by 1 year and HadGEM-GC3.1-HH shows a correlation of 0.7 when the DMV leads the AMOC by 2–3 years.

6 Conclusions

We have here analyzed historical and 1950-control simulations from seven global climate models, which followed the HighResMIP protocol, and investigated the effect of the increasing resolutions in their ocean and atmosphere components on deep convection in the Labrador Sea.

The ocean resolution clearly affects the deep ocean mixing in the Labrador Sea. Convection activity is enhanced with increasing ocean resolution in four out of five models in this study. However, all these models use NEMO (although in different configurations) as their ocean component and all of them increase the horizontal ocean resolution from 1° to 1/4°. It remains therefore unclear whether global models with other ocean components respond differently to an increased resolution. In the model, in which convection is reduced at higher ocean resolution (AWI-CM-1–0), this reduction results very likely from the simultaneously increased resolution in the atmosphere. The further increase of horizontal resolution in HadGEM-GC31-HH from 1/4° to 1/12° leads to slightly reduced convection. Thus, the conclusions on the effect of increasing ocean resolution on the convection in the Labrador Sea are only valid for models including NEMO as ocean model and for increasing resolution from non-eddy permitting to eddy permitting. Future studies need to explore the effect of increasing the resolution from eddy-permitting to eddy-rich in more detail.

Increasing the ocean resolution from 1° to 1/4° in the models with NEMO as the ocean component has a larger impact on the convection than increasing the atmosphere resolution in these models. In contrast, MPI-ESM1-2, in which only the atmosphere resolution has been increased, and AWI-CM-1–0 (increased resolution in both atmosphere and ocean) show substantially reduced convection in the Labrador Sea at high resolution. Both models (AWI-CM-1–0, MPI-ESM1-2) use the same atmospheric component (ECHAM6.3) and the reduction of DMV with increased atmospheric resolution can probably be linked to weaker winds in the high-resolution version of ECHAM6.3 (Gutjahr et al. 2019; Putrasahan et al. 2019).

Nine out of the 17 different model versions show a significantly negative trend in the Labrador Sea convection in the historical simulation compared to their respective control simulation, thus supporting the view that the trend is externally forced. In five out of the seven models, the

negative trends are stronger in the higher resolution versions of the models. This agrees well with recent finding by Roberts et al. (2020) who showed that the AMOC tends to decline more rapidly in higher-resolution models, since the AMOC is expected to weaken in response to a reduction in Labrador Sea convection in the models.

The models differ strongly in the intensity of the deep convection in the Labrador Sea. While two of the low-resolution models do not simulate any deep convection in the Labrador Sea, other lower resolution and most of the high-resolution models overestimate the deep convection compared to ARGO and EN4 data. We need to keep in mind that uncertainty regarding deep convection exist in both ARGO and EN4. However, our results are qualitatively robust across different time periods and both data sets.

The differences in convection between models can be linked to different states of the ocean vertical stratification in November and to differences in the winter atmospheric forcing (with winter surface heat losses ranging substantially across the models). Generally, we find that larger upper ocean density (mainly due to higher salinity) in November and stronger winter surface oceanic heat loss in the Labrador Sea leads to stronger convection in March. Increasing the ocean resolution improves the vertical stratification of the upper Labrador Sea in late autumn with rather realistic upper ocean temperatures and salinities compared to ARGO and EN4. The models with warmer and more saline upper ocean surface masses in the Labrador Sea show a stronger sub-polar gyre circulation, which could transport these warmer and more saline water masses from lower latitudes into the Labrador Sea. However, despite having reasonably realistic water masses in the Labrador Sea in late autumn, deep convection in March is overestimated in the high-resolution models. As potential causes we identified (1) a too shallow surface layer with low density in late autumn, which can be more easily eroded in the winter, and (2) too high surface heat losses during winter. However, the too shallow upper layer is partly compensated by too strong density gradients in the upper ocean, and the overestimated SHF can partly be a response to too strong convection. Thus, it is likely that other uncertainties in the model descriptions such as parameterizations of sub-gridscale processes in the ocean are contributing to the biases in the convection.

The variability of the DMV is mainly governed by varying surface heat fluxes and their driving large scale atmospheric modes. The NAO is governing the surface heat fluxes in the Labrador Sea and is significantly correlated to the DMV in the models. The density stratification of the Labrador Sea in late autumn plays a role for the variability as well since a strong stratification requires stronger atmospheric forcing to erode the stratification. The upper ocean density is also affected by freshwater transport, however, the variability of Arctic freshwater exports plays in most of the models

a smaller role for the variability of the DMV compared to the surface heat flux variations.

The DMV in the Labrador Sea is highly positively correlated (r = 0.6–0.8) with the AMOC at 26°N in around half of the model simulations at the decadal scale. In these simulations, the DMV leads the AMOC by a few years. In the other simulations, the correlations are positive as well but lower (0.3–0.4) and time lags of the highest correlations are not robust across these simulations. The correlations between the DMV and the AMOC are not dependent on the resolution.

The large bias in the simulation of deep convection in the Labrador Sea in the models is a concern since a realistic simulation of deep convection is important for the large-scale ocean circulation, in particular the AMOC and the North Atlantic subpolar gyre, and for the northward heat transport in the ocean and its related impacts on atmosphere and sea ice. Thus, a realistic representation of convection in the Labrador Sea is an important prerequisite for skillful decadal predictions (Menary and Hermanson 2018). The lack of such a realistic representation also raises serious questions about the future behaviour of the AMOC in climate models and its consequences for local and global climate. Thus, future studies need to analyze potential deficiencies in the ocean parameterizations in more detail to improve the representation of the deep ocean convection in the models.

esgf-node.llnl.gov. Scripts for analyzing the data will be available from the corresponding authors upon reasonable request.

References

Brodeau L, Koenigk T (2016) Extinction of the northern oceanic deep convection in an ensemble of climate model simulations of the 20th and 21st centuries. Clim Dyn 46:2863–2882. https://doi.org/10.1007/s00382-015-2736-5

Caesar L, Rahmstorf S, Robinson A, Feulner G, Saba V (2018) Observed fingerprint of a weakening Atlantic Ocean overturning circulation. Nature 556:191–196. https://doi.org/10.1038/s41586-018-0006-5

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

Cherchi A, Fogli PG, Lovato T, Peano D, Iovino D, Gualdi S, Masina S, Scoccimarro E, Materia S, Bellucci A, Navarra A (2019) Global mean climate and main patterns of variability in the CMCC-CM2 coupled model. J Adv Model Earth Syst 11:185–209. https://doi.org/10.1029/2018MS001369

Clarke R, Gascard J (1983) The formation of Labrador Sea Water. Part 1: large-scale processes. J Phys Ocean 13:1764–1778

Cunningham SA, Kanzow T, Rayner D, Baringer MO, Johns WE, Marotzke J, et al. (2007) Temporal variability of the Atlantic meridional overturning circulation at 26.5 °N. Science 317(5840):935–938. https://doi.org/10.1126/science.1141304

Danabasoglu G, Yeager S, Bailey D, Behrens E, Bentsen M, Bi D et al (2014) North Atlantic simulations in coordinated ocean-ice reference experiments phase II (CORE-II). Part I Mean States Ocean Model 73:76–107. https://doi.org/10.1016/j.ocemod.2013.10.005

de Jong MF, de Steur L (2016) Strong winter cooling over the Irminger Sea in winter 2014–2015, exceptional deep convection, and the emergence of anomalously low SST. Geophys Res Lett 43:7106–7113. https://doi.org/10.1002/2016GL069596

de Boyer Montégut C, Madec G, Fischer AS, Lazar A, Iudicone D (2004) Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology. J Geophys Res. https://doi.org/10.1029/2004JC002378

de Jong MF, Oltmanns M, Karstensen J, de Steur L (2018) Deep convection in the Irminger Sea observed with a dense mooring array. Oceanography 31(1):50–59. https://doi.org/10.5670/oceanog.2018.109

Deser C, Holland M, Reverdin G, Timlin M (2002) Decadal variations in Labrador Sea ice cover and North Atlantic sea surface temperatures. J Geophys Res. https://doi.org/10.1029/2000JC000683

Deshayes J, Frankignoul C, Drange H (2007) Formation and export of deep water in the Labrador and Irminger seas in a GCM. Deep Sea Res 54:510–532

Dickson RR, Brown J (1994) The production of North Atlantic Deep Water: sources, rates, and pathways. J Geophys Res 99(C6):12319–12341. https://doi.org/10.1029/94jc00530

Docquier D, Grist JP, Roberts MJ, Roberts CD, Semmler T, Ponsoni L, Massonnet F, Sidorenko D, Sein DV, Iovino D, Bellucci A, Fichefet T (2019) Impact of model resolution on Arctic sea ice and North Atlantic Ocean heat transport. Clim Dyn 53:4989–5017. https://doi.org/10.1007/s00382-019-04840-y

Docquier D, Fuentes-Franco R, Koenigk T, Fichefet T (2020) Sea ice-ocean interactions in the Barents Sea modeled at different resolutions. Front Earth Sci 8:172. https://doi.org/10.3389/feart.2020.00172

Fox-Kemper B, Adcroft A, Böning CW, Chassignet EP, Curchitser E, Danabasoglu G, Eden C et al (2019) Challenges and prospects in ocean circulation models. Front Mar Sci. https://doi.org/10.3389/fmars.2019.00065

Frankignoul C, Deshayes J, Curry R (2009) The role of salinity in the decadal variability of the North Atlantic meridional overturning circulation. Clim Dyn 33:777–793. https://doi.org/10.1007/s00382-008-0523-2

Fuentes Franco R, Koenigk T (2019) Sensitivity of the Arctic fresh water budget to model resolution. Clim Dyn 53:1765–1781. https://doi.org/10.1007/s00382-019-04735-y

Gelderloos R, Straneo F, Katsman CA (2012) Mechanisms behind the temporary shutdown of deep convection in the Labrador Sea: lessons from the great salinity anomaly years 1968–71. J Clim 25:6743–6755. https://doi.org/10.1175/JCLI-D-11-00549.1

Good SA, Martin MJ, Rayner NA (2013) EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. J Geophys Res Oceans 118(12):6706–6716. https://doi.org/10.1002/2013JC009067

Grist JP, Josey SA, New AL, Roberts M, Koenigk T, Iovino D (2018) Increasing Atlantic ocean heat transport in the latest generation coupled ocean-atmosphere models: the role of air-sea interaction. J Geophys Res Oceans. https://doi.org/10.1029/2018JC014387

Gutjahr O, Putrasahan D, Lohmann K, Jungclaus JH, von Storch J-S, Brüggemann N, Haak H, Stössel A (2019) Max planck institute earth system model (MPI-ESM1.2) for the high-resolution model intercomparison project (HighResMIP). Geosci Model Dev 12:3241–3281. https://doi.org/10.5194/gmd-12-3241-2019

Haarsma RJ, Roberts MJ, Vidale PL, Senior CA, Bellucci A, Bao Q, Chang P, Corti S, Fučkar NS, Guemas V, von Hardenberg J, Hazeleger W, Kodama C, Koenigk T, Leung LR, Lu J, Luo J-J, Mao J, Mizielinski MS, Mizuta R, Nobre P, Satoh M, Scoccimarro E, Semmler T, Small J, von Storch J-S (2016) High resolution model intercomparison project (HighResMIP v1.0) for CMIP6. Geosci Model Dev 9:4185–4208. https://doi.org/10.5194/gmd-9-4185-2016

Haarsma R, Acosta M, Bakhshi R, Bretonnière P-A, Caron L-P, Castrillo M, Corti S, Davini P, Exarchou E, Fabiano F, Fladrich U, Fuentes Franco R, García-Serrano J, von Hardenberg J, Koenigk T, Levine X, Meccia V, van Noije T, van den Oord G, Palmeiro FM, Rodrigo M, Ruprich-Robert Y, Le Sager P, Tourigny É, Wang S, van Weele M, Wyser K (2020) HighResMIP versions of EC-Earth: EC-Earth3P and EC-Earth3P-HR. Description, model performance, data handling and validation. Geosci Model Dev 13:3507–3527. https://doi.org/10.5194/gmd-2019-350

Hersbach H, Bell B, Berrisford P, Hirahara S, Horanyi A et al (2020) The ERA5 global reanalysis. Quat J R Met Soc 146(73):1999–2049. https://doi.org/10.1002/qj.3803

Heuzé C (2017) North Atlantic deep water formation and AMOC in CMIP5 models. Ocean Sci 13:609–622. https://doi.org/10.5194/os-13-609-2017

Heuzé C (2020) Antarctic bottom water and North Atlantic deep water in CMIP6 models. Ocean Sci Discuss. https://doi.org/10.5194/os-2020-66

Holland MM, Bitz CM, Eby M, Weaver AJ (2001) The role of ice-ocean interactions in the variability of the North Atlantic thermohaline circulation. J Clim 14:656–675. https://doi.org/10.1175/1520-0442(2001)014

Holte J, Talley LD, Gilson J, Roemmich D (2017) An Argo mixed layer climatology and database. Geophys Res Lett 44:5618–5626. https://doi.org/10.1002/2017GL073426

Jackson LC, Kahana R, Graham T, Ringer MA, Woollings T, Mecking JV, Wood RA (2015) Global and European climate impacts of a slowdown of the AMOC in a high resolution GCM. Clim Dyn 45:3299–3316. https://doi.org/10.1007/s00382-015-2540-2

Jackson LC, Peterson K, Roberts C, Wood RA (2016) Recent slowing of Atlantic overturning circulation as a recovery from earlier strengthening. Nat Geosci 9:518–522. https://doi.org/10.1038/ngeo2715

Jackson LC, Roberts MJ, Hewitt HT, Iovino D, Koenigk T, Meccia VL, Roberts CD, Ruprich-Robert Y, Wood RA (2020) Impact of ocean resolution and mean state on the rate of AMOC weakening. Clim Dyn 55:1711–1732. https://doi.org/10.1007/s00382-020-05345-9

Jungclaus J, Haak H, Latif M, Mikolajewicz U (2005) Arctic-North Atlantic interactions and multidecadal variability of the meridional overturning circulation. J Clim 18:4013–4031. https://doi.org/10.1175/JCLI3462.1

Koenigk T, Brodeau L (2017) Arctic climate and its interaction with lower latitudes under different levels of anthropogenic warming in a global coupled climate model. Clim Dyn 49(1–2):471–492. https://doi.org/10.1007/s00382-016-3354-6

Koenigk T, Mikolajewicz U, Haak H, Jungclaus J (2006) Variability of Fram Strait sea ice export: causes, impacts and feedbacks in a coupled climate model. Clim Dyn 26:17–34. https://doi.org/10.1007/s00382-005-0060-1

Koenigk T, Mikolajewicz U, Haak H, Jungclaus J (2007) Arctic freshwater export in the 20th and 21st centuries. J Geophys Res. https://doi.org/10.1029/2006JG000274

Koenigk T, König Beatty C, Caian M, Döscher R, Wyser K (2012) Potential decadal predictability and its sensitivity to sea ice albedo parameterization in a global coupled model. Clim Dyn 38(11–12):2389–2408. https://doi.org/10.1007/s00382-011-1132-z

Kuhlbrodt T, Griesel A, Montoya M, Levermann A, Hofmann M, Rahmstorf S (2007) On the driving processes of the Atlantic meridional overturning circulation. Rev Geophys. https://doi.org/10.1029/2004RG001662007

L'Hévéder B, Li L, Sevault F, Somot S (2012) Interannual variability of deep convection in the Northwestern Mediterranean simulated with a coupled AORCM. Clim Dyn 41(3–4):937–960. https://doi.org/10.1007/s00382-012-1527-5

Langehaug HR, Medhaug I, Eldevik T, Otterå OH (2012) Arctic/Atlantic exchanges via the Subpolar Gyre. J Clim 25(7):2421–2439. https://doi.org/10.1175/jcli-d-11-00085.1

Latif M, Böning C, Willebrand J, Biastoch A, Dengg J, Keenlyside N, Schweckendiek U, Madec G (2006) Is the thermohaline circulation changing? J Clim 19:4631–4637. https://doi.org/10.1175/JCLI3876.1

Lavergne C, Palter JB, Galbraith ED, Bernardello R, Marinov I (2014) Cessation of deep convection in the open southern ocean under anthropogenic climate change. Nat Clim Change 4:278–282. https://doi.org/10.1038/nclimate2132

Lazier J, Hendry R, Clarke A, Yashayaev I, Rhines P (2002) Convection and restratification in the Labrador Sea, 1990–2000. Deep Sea Res Part I 49(10):1819–1835. https://doi.org/10.1016/S0967-0637(02)00064-X

Levitus S (1982) Climatological atlas of the world ocean. U.S. Government Printing Office 13, NOAA, Washington, D.C., 163 pp

Lozier MS, Bacon S, Bower AS, Cunningham SA, de Jong MF, de Steur L, de Young B, Fischer J et al (2017) Overturning in the subpolar North Atlantic Program: a new international ocean observing system. BAMS 98(4):737–752. https://doi.org/10.1175/BAMS-D-16-0057.1

Lozier MS, Li F, Bacon S, Bahr F, Bower AS, Cunningham SA, de Jong MF, de Steur L et al (2019) A sea change in our view of overturning in the subpolar North Atlantic. Science 363:516–521

Mahajan S, Zhang R, Delworth T (2011) Impact of the atlantic meridional overturning circulation (AMOC) on arctic surface air temperature and sea ice variability. J Clim 24:6573–6581. https://doi.org/10.1175/2011JCLI4002.1

Manabe S, Stouffer RJ (1999) The role of thermohaline circulation in climate. Tellus B 51:91–109. https://doi.org/10.1034/j.1600-0889.1999.00008.x

Marotzke J, Scott JR (1999) Convective mixing and the thermohaline circulation. J Phys Oceanogr 29(11):2962–2970. https://doi.org/10.1175/1520-0485(1999)029

Medhaug I, Furevik T (2011) North Atlantic 20th century multidecadal variability in coupled climate models: sea surface temperature and ocean overturning circulation. Ocean Sci 7:389–404. https://doi.org/10.5194/os-7-389-2011

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

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

Menary MB, Jackson LC, Lozier MS (2020) Reconciling the relationship between the AMOC and Labrador Sea in OSNAP observations and climate models. Geophys Res Lett 47: e2020GL089793. https://doi.org/10.1029/2020GL089793

Millero FJ, Poisson A (1981) International one-atmosphere equation of state of seawater. Deep Sea Research Part A. Oceanogr Res Pap 28(6): 625–629. https://doi.org/10.1016/0198-0149(81)90122-9

Monterey G, Levitus S (1997) Seasonal variability of mixed layer depth for the world ocean. NOAA ATLAS, NESDlS 14, Washington, D.C., 96 pp

Obata A, Ishizaka J, Endoh M (1996) Global verification of critical depth theory for phytoplankton bloom with climatological in situ temperature and satellite ocean color data. J Geophys Res 101:20657–20667

Ortega P, Hawkins E, Sutton R (2011) Processes governing the predictability of the Atlantic meridional overturning circulation in a coupled GCM. Clim Dyn 37:1771–1782. https://doi.org/10.1007/s00382-011-1025-1

Ortega P, Robson J, Sutton RT, Andres MB (2017) Mechanisms of decadal variability in the Labrador Sea and the wider North Atlantic in a high-resolution climate model. Clim Dyn 49: 2625–2647. https://link.springer.com/article/10.1007/s00382-016-3467-y

Putrasahan DA, Lohmann K, von Storch J-S, Jungclaus JH, Gutjahr O, Haak H (2019) Surface flux drivers for the slowdown of the atlantic meridional overturning circulation in a high-resolution global coupled climate model. J Adv Model Earth Syst 11:1349–1363. https://doi.org/10.1029/2018MS001447

Roberts CD, Garry FK, Jackson LC (2013) A Multimodel study of sea surface temperature and subsurface density fingerprints of the atlantic meridional overturning circulation. J Clim 26:9154–9174. https://doi.org/10.1175/JCLI-D-12-00762.1

Roberts CD, Jackson L, McNeall D (2014) Is the 2004–2012 reduction of the Atlantic meridional overturning circulation significant? Geophys Res Lett 41:3204–3210. https://doi.org/10.1002/2014GL059473

Roberts CD, Senan R, Molteni F, Boussetta S, Mayer M, Keeley SPE (2018) Climate model configurations of the ECMWF integrated forecasting system (ECMWF-IFS cycle 43r1) for HighResMIP. Geosci Model Dev 11:3681–3712. https://doi.org/10.5194/gmd-11-3681-2018

Roberts MJ, Baker A, Blockley EW, Calvert D, Coward A, Hewitt HT, Jackson LC, Kuhlbrodt T, Mathiot P, Roberts CD, Schiemann R, Seddon J, Vannière B, Vidale PL (2019) Description of the resolution hierarchy of the global coupled HadGEM3-GC3.1 model as used in CMIP6 HighResMIP experiments. Geosci Model Dev 12:4999–5028. https://doi.org/10.5194/gmd-12-4999-2019

Roberts M.J, Jackson LC, Roberts CD, Meccia V, Docquier D, Koenigk T, Ortega P, Moreno-Chamarro E, Bellucci A, Coward A, Drijfhout S, Exarchou E, Gutjahr O, Hewitt H, Iovino D, Lohmann K, Schiemann R, Seddon J, Terray L, Xu X, Zhang Q, Chang P, Yeager SG, Castruccio FS, Zhang S, Wu L (2020) Sensitivity of the atlantic meridional overturning circulation to model resolution in CMIP6 HighResMIP simulations and implications for future changes. J Adv Model Earth Syst. https://doi.org/10.1029/2019MS002014

Robson J, Ortega P, Sutton R (2016) A reversal of climatic trends in the North Atlantic since 2005. Nature Geosci 9:513–517. https://doi.org/10.1038/ngeo2727

Ronski S, Budéus G (2005) Time series of winter convection in the Greenland Sea. J Geophys Res. https://doi.org/10.1029/2004JC002318

Sayol JM, Dijkstra H, Katsman C (2019) Seasonal and regional variations of sinking in the subpolar North Atlantic from a high-resolution ocean model. Ocean Sci 15:1033–1053. https://doi.org/10.5194/os-15-1033-2019

Schott FA, Stramma L, Giese BS, Zantopp R (2009) Labrador Sea convection and subpolar North Atlantic deep water export in the SODA assimilation model. Deep Sea Res 56:926–938. https://doi.org/10.1016/j.dsr.2009.01.001

Sein DV, Danilov S, Biastoch A, Durgadoo JV, Sidorenko D, Harig S, Wang Q (2016) Designing variable ocean model resolution based on the observed ocean variability. J Adv Model Earth Syst 8(2):904–916. https://doi.org/10.1002/2016MS000650

Sein DV, Koldunov NV, Danilov S, Sidorenko D, Wekerle C, Cabos W et al (2018) The relative influence of atmospheric and oceanic model resolution on the circulation of the North Atlantic ocean in a coupled climate model. J Adv Model Earth Syst 10(8):2026–2041. https://doi.org/10.1029/2018MS001327

Sidorenko D, Rackow T, Jung T, Semmler T, Barbi D, Danilov S et al (2014) Towards multi-resolution global climate modeling with ECHAM6–FESOM. Part I: model formulation and mean climate. Clim Dyn 44:757–780. https://doi.org/10.1007/s00382-014-2290-6

Smeed D, McCarthy G., Rayner D, Moat BI, Johns WE, Baringer MO, Meinen CS 2017: Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS (RAPID-meridional overturning circulation and heatflux array-western boundary time series) array at 26°N from 2004 to 2017. https://doi.org/10.5285/5acfd143-1104-7b58-e053-6c86abc0d94b

Smeed DA, McCarthy GD, Cunningham SA, Frajka-Williams E, Rayner D, Johns WE, Meinen CS, Baringer MO, Moat BI, Duchez A, Bryden HL (2014) Observed decline of the Atlantic meridional overturning circulation 2004–2012. Ocean Sci 10:29–38. https://doi.org/10.5194/os-10-29-2014

Smeed DA, Josey SA, Beaulieu C, Johns WE, Moat BI, Frajka-Williams E, Rayner D, Meinen CS, Baringer MO, Bryden HL, McCarthy GD (2018) The North Atlantic Ocean is in a state of reduced overturning. Geophys Res Lett 45:1527–1533. https://doi.org/10.1002/2017GL07635

Swingedouw D, Braconnot P, Delecluse P, Guilyardi E, Marti O (2007) Quantifying the AMOC feedbacks during a 2xCO$_2$ stabilization experiment with land-ice melting. Clim Dyn 29:521–534. https://doi.org/10.1007/s00382-007-0250-0

Thornalley D, 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 556:227–231. https://doi.org/10.1038/s41586-018-0007-4

Toyoda T, Fujii Y, Kuragano T et al (2017) Intercomparison and validation of the mixed layer depth fields of global ocean syntheses. Clim Dyn 49:753–773. https://doi.org/10.1007/s00382-015-2637-7

Vannière B, Demory ME, Vidale PL, Schiemann R, Roberts MJ, Roberts CD, Matsueda M, Terray L, Koenigk T, Senan R (2019) Multi-model evaluation of the sensitivity of the global energy budget and hydrological cycle to resolution. Clim Dyn 52(11):6817–6846. https://doi.org/10.1007/s00382-018-4547-y

Voldoire A, Saint-Martin D, Sénési S, Decharme B, Alias A, Chevallier M et al (2019) Evaluation of CMIP6 DECK experiments with CNRM-CM6-1. J Adv Model Earth Syst 11:2177–2213. https://doi.org/10.1029/2019MS001683

Yashayaev I (2007) Hydrographic changes in the Labrador Sea, 1960–2005. Prog Oceanogr 73(3–4):242–276. https://doi.org/10.1016/j.pocean.2007.04.015

Yashayaev I, Loder JW (2009) Enhanced production of Labrador Sea Water in 2008. Geophys Res Lett. https://doi.org/10.1029/2008gl036162

Yashayaev I, Loder JW (2016) Recurrent replenishment of Labrador Sea Water and associated decadal-scale variability. J Geophys Res Oceans 121:8095–8114. https://doi.org/10.1002/2016JC012046

Yashayaev I, Loder JW (2017) Further intensification of deep convection in the Labrador Sea in 2016. Geophys Res Lett 44:1429–1438. https://doi.org/10.1002/2016GL071668

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