The North Atlantic mean state in eddy-resolving coupled models: a multimodel study

Amanda Frigola[1], Eneko Martin-Martinez[1], Eduardo Moreno-Chamarro[1,2], Margarida Samsó[1], Saskia Loosvelt-Tomas[1], Pierre-Antoine Bretonnière[1], Daria Kuznetsova[1], Xia Lin[3], Pablo Ortega[1]

[1]Barcelona Supercomputing Center, Barcelona, Spain
[2]now at Max Planck Institute for Meteorology, Hamburg, Germany
[3]Nanjing University of Information Science and Technology, Nanjing, China

Correspondence to: Amanda Frigola (amanda.frigola@bsc.es)

Abstract. Ocean mesoscale structures, which are parameterized in standard resolution models, play an important role in featuring global ocean dynamics. Here we study the effects of increasing model ocean resolution to eddy-resolving scales on the representation of the North Atlantic mean state, by comparing an ensemble of four HighResMIP coupled historical simulations with nominal ocean resolutions of at least 1/10º – corresponding to the models CESM1-CAM5-SE-HR, EC-Earth3P-VHR, HadGEM3-GC31-HH, and MPI-ESM1-2-ER – to a baseline of 39 Coupled Model Intercomparison Project phase 6 (CMIP6) simulations at coarser resolution. Despite a warm and salty bias at the surface, we find generally improved vertical stratification in the Labrador (LS) and western Irminger Seas (IS) in the high resolution ensemble, leading to stronger deep water convection in that region, in closer agreement with observations. Both the overturning and barotropic circulations of the North Atlantic present significant improvements in the eddy-resolving models: the Atlantic Meridional Overturning Circulation (AMOC) is weaker than for lower resolution models and closer to RAPID observations; the paths, strength, and structure of the Gulf Stream (GS), North Atlantic Current (NAC), and subpolar gyre (SPG) are also improved, resulting into reduced surface temperature and salinity biases north of Cape Hatteras (NCH) and in the Central North Atlantic (CNA).

1 Introduction

The North Atlantic is a key region with multiple impacts on the global climate system. One of its main roles is the redistribution of heat from low to high latitudes through the Atlantic Meridional Overturning Circulation (AMOC). At 26.5º N the Atlantic Ocean transports ~1.2 PW of heat, which represents ~60–65 % of the combined contributions from the Atlantic and the Pacific at those latitudes (Ganachaud and Wunsch, 2000; Johns et al., 2023; Lumpkin and Speer, 2007; Trenberth and Fasullo, 2017). Indeed, heat transport by the AMOC explains the milder temperatures in the Northern Hemisphere compared to the Southern Hemisphere (Buckley and Marshall, 2016). The North Atlantic is also an important anthropogenic carbon sink, contributing to reducing atmospheric global warming (Brown et al., 2021). This region exhibits

the highest global uptake rate of anthropogenic carbon per area, which is related to enhanced vertical penetration via the AMOC upper cell (Gruber et al., 2019).

Changes in the AMOC in the past have been associated with abrupt changes in climate (Ng et al., 2018), and climate projections indicate consistent AMOC weakening at increased CO_2 levels (Jackson et al., 2020), with important effects upon
35   climate, such as Northern Hemisphere drying and cooling, and a southward shift in the intertropical convergence zone (ITCZ; Bellomo and Mehling, 2024; Liu et al., 2020). Thus, considering its fundamental role within the climate system, the dynamics of the North Atlantic need to be appropriately represented in climate models, in order to trustfully evaluate the future impacts of climate change.

The North Atlantic circulation is influenced by a series of elements and processes that are strongly interconnected: the
40   strength, paths, and water properties of the AMOC northward limb, which determine the heat and salinity content of the waters reaching the subpolar North Atlantic (SPNA; Marzocchi et al., 2015); deep water convection at high latitudes, which is linked to the formation of dense waters (Koenigk et al., 2021); surface water mass transformation in the SPNA, associated with cooling through atmospheric heat fluxes (Jackson and Petit, 2023); densification of waters at the subpolar gyre (SPG) boundary (Katsman et al., 2018); or mixing with Arctic waters entering the SPG through the Greenland–Scotland Ridge
45   (Dey et al., 2024).

The horizontal resolution of ocean models is crucial for accurately representing these processes. A realistic bathymetry is key in characterizing ocean throughflows and their properties, in particular those of the Greenland–Scotland Ridge (Katsman et al., 2018; Dey et al., 2024). Additionally, ocean horizontal resolution determines the representation of mesoscale features, the so-called ocean eddies. These structures impact the ocean dynamics through their fundamental role in
50   the transport of heat and salt (Sun et al., 2019; Treguier et al., 2012). Ocean eddies' average horizontal scale is smaller at high latitudes, continental shelves, and areas of weak stratification (Hallberg, 2013). Models with ocean resolutions of at least 1/10º are known as eddy-resolving models and are capable of resolving eddies in extensive areas of the North Atlantic – with limitations in regions of weak stratification or shallow bathymetry. Eddy-permitting models, instead, have resolutions in the order of 1/4º and are only able to resolve eddies in the tropics. In models with ocean resolutions of 1º or coarser, the
55   contribution of eddies is parameterized.

Associated with a better characterization of the ocean mesoscale, increasing the ocean resolution to eddy-permitting scales has been shown to improve the representation of boundary and frontal currents, such as the Gulf Stream (GS) and North Atlantic Current (NAC), both in terms of location and structure, with significant further improvement at eddy-resolving scales (Hewitt et al., 2017; Marzocchi et al., 2015). This better characterization of the GS and NAC leads to
60   reduced sea surface temperature (SST) and salinity (SSS) biases north of Cape Hatteras (NCH) and in the Central North Atlantic (CNA) at eddy-resolving scales (Marzocchi et al., 2015). These reduced surface temperature biases are reflected in the atmosphere mean state as well: the winter stormtrack bias generally present at high latitudes in eddy-parameterized models is reduced at eddy-resolving scales, associated to a weaker meridional temperature gradient in the North Atlantic

(Moreno-Chamarro et al., 2024); and also the local negative bias in precipitation associated to the cold CNA bias is reduced
65   at high resolution (Moreno-Chamarro et al., 2022).

Increasing ocean resolution to (at least) eddy-permitting scales has also been shown to improve air-sea interactions in the
North Atlantic. More specifically, eddy-permitting models, when run together with an atmospheric component of equivalent
resolution, exhibit more realistic near surface wind stress divergence and curl fields over the GS and NAC compared to
eddy-parameterized models (Tsartsali et al., 2022). Similarly, the representation of the covariance between SST and heat
70   fluxes is improved in that region at eddy-permitting scales (Bellucci et al., 2021).

This study has its focus on the North Atlantic, and aims at assessing the impact of explicitly resolving mesoscale ocean
eddies in the representation of its mean state in historical simulations, by comparing an ensemble of four Coupled Model
Intercomparison Project phase 6 (CMIP6) HighResMIP (Sect. 2.1; Haarsma et al., 2016) coupled eddy-resolving models –
namely CESM1-CAM5-SE-HR (Chang et al., 2020), EC-Earth3P-VHR (Moreno-Chamarro et al., 2024), HadGEM3-GC31-
75   HH (Roberts et al., 2019), and MPI-ESM1-2-ER (Gutjahr et al., 2019) – with a second ensemble of 39 CMIP6 coupled non-
eddy-resolving models. This work focuses on describing the dynamics of the North Atlantic, as well as the properties that
impact them, such as the biases in temperature and salinity, the stratification, and the deep ocean convection.

To our knowledge, only a few multimodel comparisons of coupled historical experiments with a focus on the North
Atlantic ocean exist, that include eddy-resolving simulations (e.g. Roberts et al., 2020; Koenigk et al., 2021), although none
80   of them specifically addresses the impact of resolving mesoscale ocean eddies. In that context, our study stands out for its
particular focus on the added value of these eddies, featuring the largest ensemble of coupled eddy-resolving simulations
considered so far. This ensemble allows us to evaluate more consistently which aspects of the mean climate are improved at
that resolution.

This manuscript is structured as follows: the data and methodological approach employed are described in Sect. 2. The
85   main results of the study are presented in Sect. 3, including a characterization of SST and SSS biases in the North Atlantic
for the high and low resolution ensembles (Sect. 3.1), the stratification (Sect. 3.2) and mixing in the regions of deep water
formation (Sect 3.3), the AMOC (Sect 3.4), and the gyre circulations, including the NAC and the SPG (Sect. 3.5). Finally, in
Sect. 4 we delve deep into the discussion of the main results, relate them to the current literature, and present our
conclusions.

2 Data and methods

2.1 Model data and methodological approach

In order to assess the impact of increased horizontal resolution on the representation of the North Atlantic mean state, we
analyze the outputs from four CMIP6-endorsed HighResMIP (Haarsma et al., 2016) coupled eddy-resolving historical
simulations (hist-1950; HR-HIST hereinafter) – corresponding to the models CESM1-CAM5-SE-HR (Chang et al., 2020),

95 EC-Earth3P-VHR (Moreno-Chamarro et al., 2024), HadGEM3-GC31-HH (Roberts et al., 2019), and MPI-ESM1-2-ER
(Gutjahr et al., 2019) – and compare them to a baseline ensemble of 39 CMIP6 coupled historical runs (Eyring et al., 2016)
performed at coarser resolution (LR-HIST hereinafter). We reckon this standard resolution ensemble as a more rigorous
benchmark than the low-resolution HighResMIP counterparts of the four eddy-resolving models, given its much larger size.
More details on the models considered are provided in Tables 1 and B1.

100

HR-HIST models	ocean component	ocean grid	atm. component	atm. grid	reference
CESM1-CAM5-SE-HR	POP2	1/10°; tripolar; 3600x2400 lon/lat; 62 levels;	CAM5.2	25 km; 30 levels;	Chang et al. (2020)
EC-Earth3P-VHR	NEMO3.6	1/12°; ORCA12 tripolar; 4322 x 3059 lon/lat; 75 levels;	IFS cy36r4	16 km; 91 levels;	Moreno-Chamarro et al. (2024)
HadGEM3-GC31-HH	NEMO-HadGEM3-GO6.0	1/12°; eORCA12 tripolar; 4320 x 3604 lon/lat; 75 levels;	MetUM-HadGEM3-GA7.1	50 km; 85 levels;	Roberts et al. (2019)
MPI-ESM1-2-ER	MPIOM	1/10°; TP6M tripolar; 3602 x 2394 lon/lat; 40 levels;	ECHAM6.3	103 km; 95 levels;	Gutjahr et al. (2019)
LR-HIST models	–	25–250 km (mainly 100 km)	–	100–500 km (mainly 100–250 km)	–

Table 1: Overview of models used in the current study. Ocean grid details include: nominal resolution; grid type; size of horizontal grid; and number of vertical levels. Details about the individual LR-HIST models can be found in Table B1 in Appendix B.

105 All models in the HR-HIST ensemble have a nominal ocean resolution of at least 1/10º, allowing them to represent the
ocean mesoscale in extensive areas of the North Atlantic. By contrast, in LR-HIST, the ocean mesoscale is, at best, only
resolved in the tropics and more generally parameterized, since ocean resolution in that ensemble ranges from 25 to 250 km,
being 100 km the most common resolution across models. Atmospheric resolution is also generally higher in the HR-HIST
ensemble compared to LR-HIST, ranging from 15 to 100 km for HR-HIST, and from 100 to 500 km for LR-HIST (Tables 1
110 and B1). We note the heterogeneity in model components employed across the ensembles, which might avoid a dominant
contribution of specific individual model biases within the ensembles.

Model selection criteria is based on the availability of three-dimensional temperature and salinity, and the overturning
mass streamfunction (either msftmz or msftyz) for the Atlantic Ocean as output variables. Although in most ocean grids the
y-grid direction might differ from the meridional direction at high northern latitudes, in general the overturning mass
115 streamfunction calculated along lines of constant y (msftyz) provides a good approximation of that calculated along lines of

constant latitude (msftmz) (Griffies et al., 2016). In our study, we select msftmz over msftyz when available. The models for which only msftyz is available as output variable (marked with an * in Fig. 9) present significant grid rotation only from ≥ 40º N northwards (except for MPI-ESM1-2-ER, which presents rotation already at 30º N). For caution, we restrict our analysis to latitudes below 40º N, which nevertheless allows us to extract valuable information.

120   Forcing fields in the HighResMIP coupled historical simulations (HR-HIST ensemble) are almost identical to those in the CMIP6 historical simulations (LR-HIST ensemble). The only significant difference concerns land use, which is fixed in time in HighResMIP and representative of the present-day period (around year 2000) (Haarsma et al., 2016), and time-varying in CMIP6 historical simulations (Eyring et al., 2016). We do not expect this to cause important differences in the ocean variables considered in our analysis. It is also worth noting that HighResMIP historical simulations are run without

125   interactive aerosols, but this is also the case for several CMIP6 historical simulations (e.g. CMCC-CM2-HR4, EC-Earth3, FGOALS-f3-L, GISS-E2-2-G, IPSL-CM6A-LR, MPI-ESM1-2-HR).

More significant differences concern model initialization and spin-up. Due to the high computational costs of high-resolution modelling, in HighResMIP, initial conditions for historical runs are taken from a short spin-up (~30–50 years) with fixed 1950's radiative forcings and ocean initial conditions (Haarsma et al., 2016), instead of from a long pre-industrial

130   control representative of 1850 conditions. Thus, in the case of the HighResMIP experiments, the substantially shorter spin-up and historical period covered (1950–2014) can leave some lingering drifts. Nevertheless, in some eddy-resolving HighResMIP simulations, the ocean seems to equilibrate faster (Moreno-Chamarro et al., 2024; Roberts et al., 2019) compared to their lower resolution counterparts.

Data analyses are carried out using the Earth System Model Evaluation Tool (ESMValTool v2.10.0), a Python package

135   designed for model intercomparison purposes (Andela et al., 2023a, b; Righi et al., 2020). We note that the CMIP6 model ICON-ESM-LR was excluded from some of the analyses due to an incompatibility of the data with ESMValTool. Climatologies are computed for the last 35 years of the historical runs (1980–2014) to reduce the potential effect of model drifts, which are expected to be larger in the earlier years of the HighResMIP simulations.

2.2 Specific diagnostics

140   Potential density anomalies with respect to a reference pressure of 0 dbar (σ_0) are calculated from temperature and salinity monthly means, using the polynomial approximation of the TEOS-10 equation of state for Boussinesq models (Roquet et al., 2015). Mixed layer depth (MLD) is defined and calculated as the shallowest depth level at which monthly potential density σ_0 exceeds by a threshold of 0.03 kg m^{-3} its value at a reference depth of 10 m, as described in de Boyer Montégut et al. (2004). This method is preferred over employing direct MLD model outputs that use instantaneous values

145   and a range of different definitions, to ensure a consistent comparison across models and observations.

2.3 Observational references

Observational and reanalysis data are employed to evaluate model performance. For temperature and salinity, the Met Office Hadley Centre EN.4.2.2 dataset (Good et al., 2013) with the Gouretski and Reseghetti (2010) expendable bathythermograph and Gouretski and Cheng (2020) mechanical bathythermograph corrections is used, which has a resolution of 1º. We opt for this three-dimensional dataset to jointly assess biases at the surface and depth. We also use it to derive an observational reference for the MLD that is physically consistent with EN4 salinity and temperature fields.

For the Atlantic overturning mass streamfunction and barotropic streamfunction, ORAS5m reanalysis data (Tietsche et al., 2020) are used as our reference (1/4° resolution; period 1980–2014). ORAS5m is an improved version of the 5th ECMWF ocean reanalysis system ORAS5 (Zuo et al., 2019), with reduced SST nudging and increased weight to coastal observations. This version improves the representation of the AMOC and leads to reduced biases in winter reforecasts of the North Atlantic.

As a complementary reference of direct observational data, the climatological vertical profile of the RAPID array AMOC streamfunction is employed to validate the simulated AMOC at 26.5° N. RAPID is a monitoring programme providing time series of AMOC based on temperature, salinity and pressure profiles from a mooring array crossing the Atlantic from west to east at 26.5º N (Johns et al., 2023; Moat et al., 2023). The climatology employed corresponds to the period April 2004 – February 2022.

Monthly averaged absolute dynamic topography data (sea surface height above geoid) are also employed from AVISO observations at 1/4° resolution for the period February 1993 – December 2014.

3 Results

3.1 Sea surface biases

Temperature and salinity biases, through their impact on the zonal and vertical density gradients, are important for the realism of the ocean circulation and deep water formation in the North Atlantic. The mean SST biases of the individual LR-HIST and HR-HIST models are shown in Fig. 1, and their respective multi-model means in Fig. 2. In general, the HR-HIST ensemble mean displays warmer surface waters in the SPNA, compared to the LR-HIST one. The LR-HIST ensemble shows two main SST biases of opposite sign and similar magnitude. The first is a warm bias located along the North American coast, at NCH, with temperatures 2–5° C warmer than observations (Fig. 2). This bias has previously been associated with a misrepresentation of the position of the GS separation from the coast (Marzocchi et al., 2015). The other is a cold bias in the CNA (2–5º C), which earlier studies have linked to an overly weak NAC and an underestimation of the horizontal heat transport into the CNA domain (Lin et al., in review). The NCH bias has been shown to have an important impact on the global atmospheric circulation, through a Rossby wave response to local changes in vertical motion in the troposphere (Lee

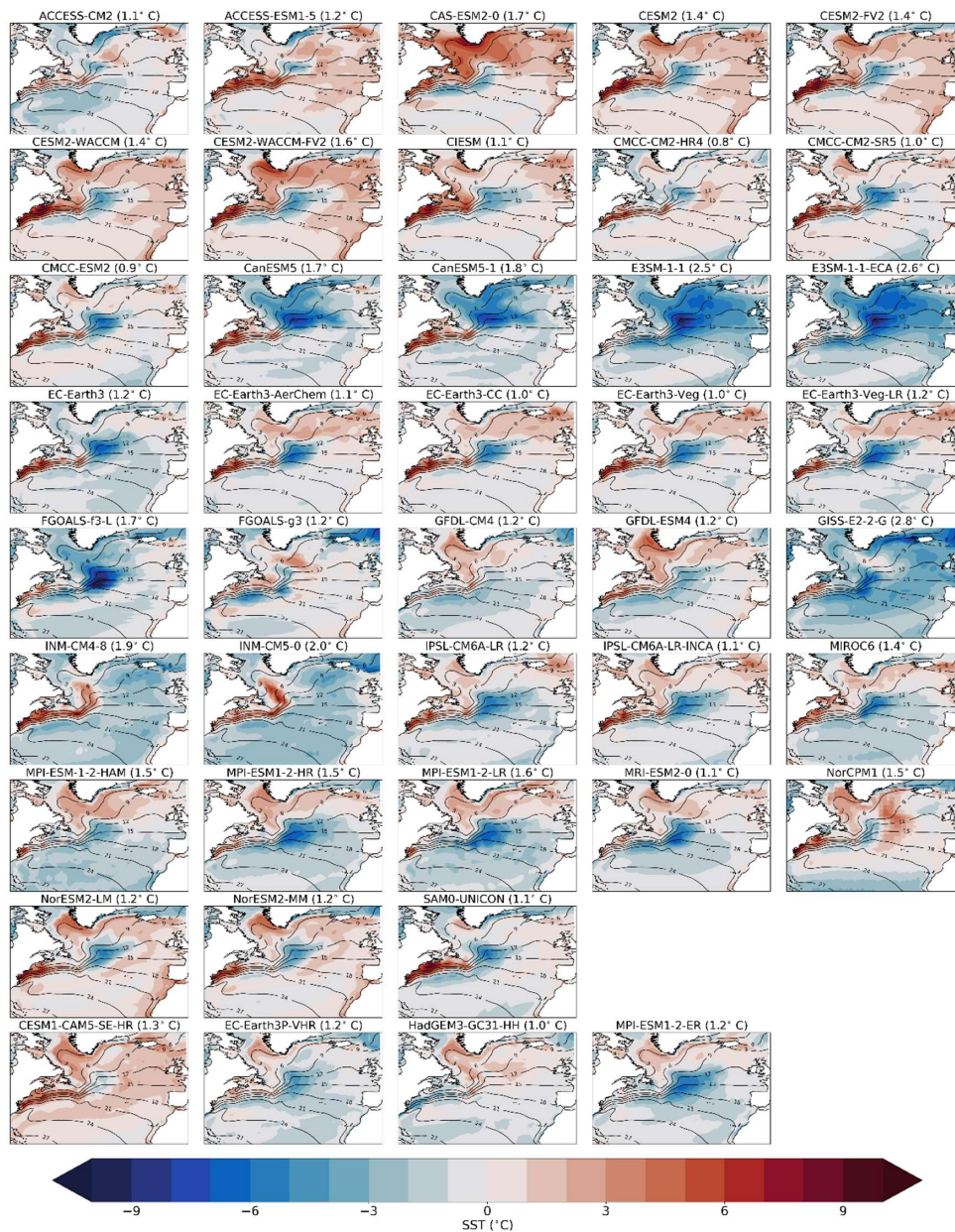
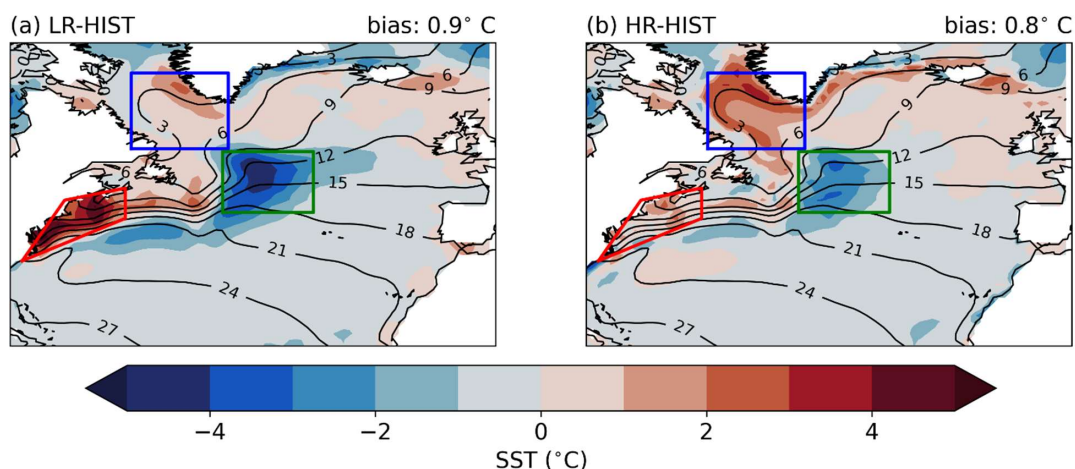



Figure 1. Sea surface temperature (SST) bias (shading; in ºC) for the individual LR-HIST (rows 1–8) and HR-HIST (last row) models with respect to EN4 for the period 1980–2014. EN4 climatology shown in contour lines (in ºC). Values in parenthesis in each subfigure header show the spatially averaged absolute mean bias for each individual model.

180

et al., 2018); the CNA bias has an effect on local precipitation (Moreno-Chamarro et al., 2022). In the HR-HIST mean, the
NCH and CNA biases are significantly reduced compared to LR-HIST. By contrast, the HR-HIST mean shows a significant
positive bias of 1–3º C in the Labrador Sea (LS), which is weaker in the LR-HIST mean (Fig. 2). We note, however, that the
range of biases is larger in the LR-HIST ensemble compared to the HR-HIST one, with some LR-HIST models showing

185 larger positive biases than the HR-HIST ones (Fig. 1; e.g. CAS-ESM2-0 and CESM2-WACCM-FV2), while other LR-HIST
models present substantial biases of opposite sign (e.g. CanESM5-1 and E3SM-1-1).

190 **Figure 2.** SST bias for the multi-model mean of the (a) LR-HIST and (b) HR-HIST ensembles. Plotting details as in Fig. 1. Coloured
polygons delineate the main bias regions addressed in the paper: north of Cape Hatteras (NCH) in red [edges: (78º W, 34º N), (61º W, 41º
N), (61º W, 46º N), (71º W, 44º N)], Central North Atlantic (CNA) in green (30º–45º W, 42º–52º N), and Labrador Sea (LS) in blue (44º–
60º W, 52.5º–65º N).

195 Analogously to the SSTs, SSS biases from the individual models and the corresponding ensemble means are described in
Figs. 3 and 4. The multi-model mean SSS biases show a similar pattern to the temperature ones (Fig. 4). LR-HIST presents a
positive salinity bias of 1–3 at NCH, and a negative bias of 0.5–1.5 in the CNA. Note that salinities are presented on the
practical salinity scale throughout the manuscript, with no associated units. In contrast to the SST biases, the SSS CNA
negative bias is not a common feature in all LR models, although it is indeed dominant across them (Fig. 3). For HR-HIST,

200 the NCH and CNA biases are significantly reduced with respect to LR-HIST, although a positive bias of 0.5–1 appears in the
LS that is not present in the LR-HIST ensemble mean, probably because biases of different models compensate with each
other. We note also that in the LS, models tend to show SST and SSS biases of the same sign, with CIESM, GISS-E2-2-G,
INM-CM4-8, and INM-CM5-0 as exceptions. This might lead to a compensating contribution to the surface density biases.
Despite the apparent LS degradation for HR-HIST, the spatially-averaged absolute SSS biases in the North Atlantic are

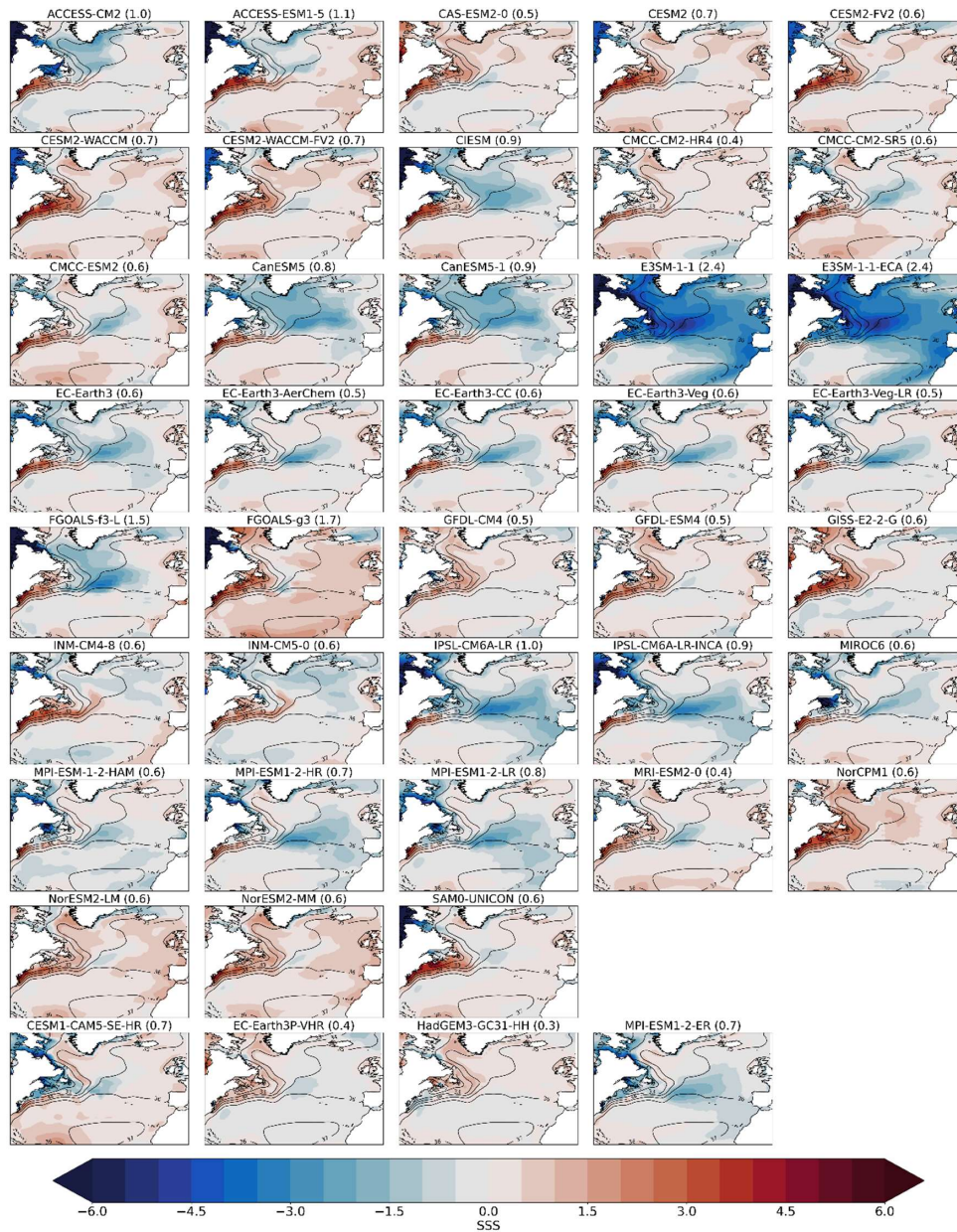



Figure 3. As in Fig. 1 but for the sea surface salinity (SSS) biases.

205

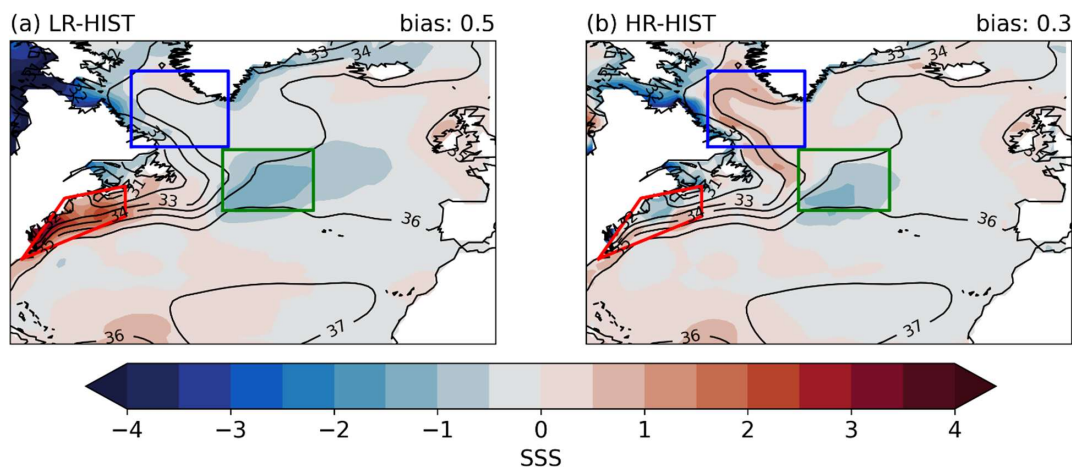



Figure 4. As in Fig. 2 but for the SSS biases.

210

substantially lower in HR-HIST simulations than in LR-HIST (i.e. 0.3 in HR-HIST vs 0.5 in LR-HIST), supporting the overall beneficial effect of the enhanced resolution.

3.2 Stratification in the Labrador/Irminger Sea

LS and Irminger Sea (IS) vertical water properties are important for deep water formation and connected to the AMOC
215 strength (Ortega et al., 2021). In Fig. 5, vertical profiles of temperature, salinity, and density for the Labrador/Irminger Sea (LIS) box shown in Figs. 7 and 8, are plotted and compared to observations, to characterize their related biases and assess the differences across ensembles. The LIS domain contains the LS and the western part of the IS, and is defined to cover the area of weakest stratification in the North Atlantic (Ortega et al., 2021). Although the HR-HIST multi-model mean temperature profile (Fig. 5a, thick red curve) displays a larger surface bias in the LIS box compared to the LR-HIST one, when the whole
220 vertical column is considered, HR-HIST temperature profiles are closer to EN4 than LR-HIST. This better agreement with observations is further supported by the respective root mean square errors (RMSE) and correlation values in the vertical dimension against EN4 (Fig. 6a). Interestingly, the EC-Earth3P-VHR and MPI-ESM1-2-ER models exhibit the most realistic temperature profiles of all models, with CESM1-CAM5-SE-HR showing a very good vertical structure but a relatively high RMSE.

225 In terms of salinity, the HR-HIST ensemble displays a positive bias in the whole water column in the LIS box, while the mean for LR-HIST exhibits a negative bias in the upper 150 m, where it remains closer to the EN4 reference (Fig. 5b). However, below 150 m, the LR-HIST ensemble mean salinity bias turns positive and is stronger than for HR-HIST. Overall, the vertical salinity profile is more realistic in HR-HIST, as supported by the higher correlation coefficient and smaller

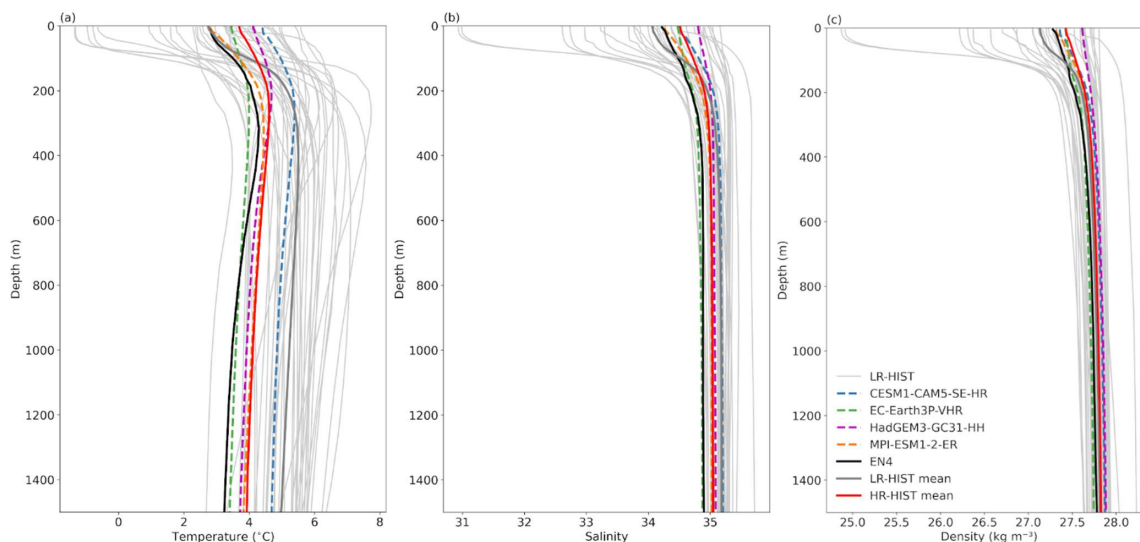



Figure 5. Vertical profiles of (a) temperature (in ºC), (b) salinity, and (c) density (in kg/m³), averaged over (35º–60º W, 50º–65º N), the Labrador/Irminger Sea (LIS) box shown in Figs. 7 and 8.

RMSE against EN4 indicated in Fig. 6b. The previous temperature and salinity profiles determine the realism of the climatological density profile in the LIS box in both ensembles. Figure 5c shows density biases of opposite sign and similar magnitude in the top ~100 m of HR-HIST and LR-HIST, and positive biases of comparable magnitude below. More interestingly, HR-HIST exhibits a more realistic density stratification, which is largely overestimated in LR-HIST. Indeed, the density profile for HR-HIST is closer in shape to the EN4-derived one, as supported by Pearson correlation coefficients in Fig. 6c, which are very close to one in all HR-HIST models. Although some individual LR-HIST models present similar (or improved) density profiles compared to HR-HIST, the LR-HIST ensemble shows significant spread, with many models being far from observations both in terms of correlation and RMSE (Fig. 6c).

The comparatively lower density stratification in HR-HIST is explained by a relative reduction in salinity stratification (which is partly counterbalanced by the relative increase in temperature stratification; Fig. 5). The improved shape of the HR-HIST mean density profile is expected to reflect on the vertical mixing, which is addressed in the next section.

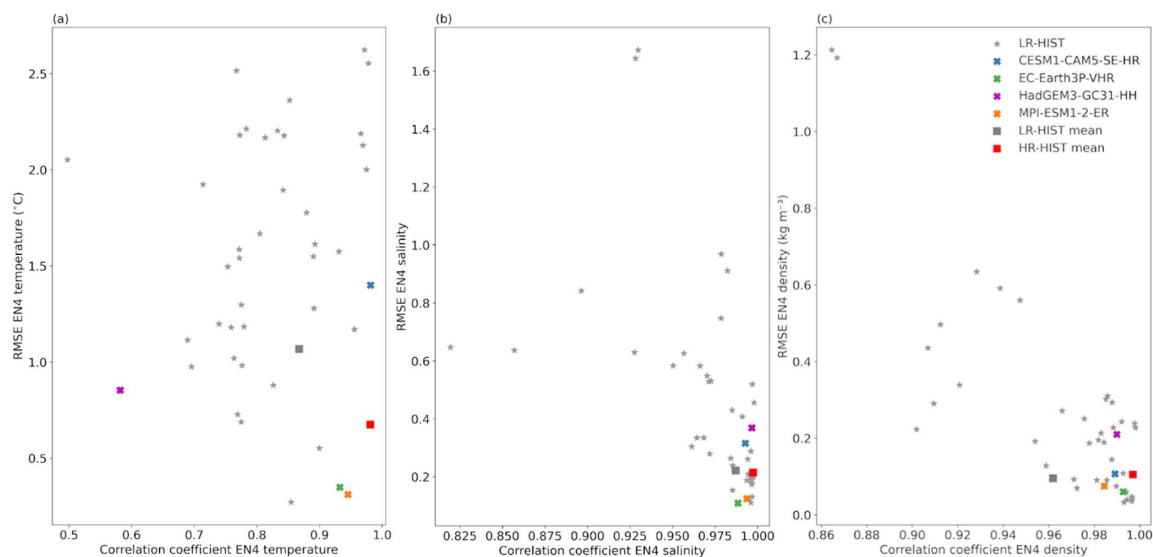



Figure 6. Pearson correlation coefficient (horizontal axis; units as in Fig. 5) and Root Mean Square Error (RMSE) (vertical axis) of the (a) temperature, (b) salinity and (c) density profiles in Fig. 5 against EN4, in the vertical dimension.

3.3 Mixed layer depth

MLD is generally used as a proxy for deep water convection, it is correlated to AMOC strength (Martin-Martinez et al., in review), and in the North Atlantic achieves its maximum in March. Its March climatology for the two multi-model ensembles and EN4-derived values is shown in Figs. 7 and 8. In the LS, the multi-model mean of HR-HIST shows a deeper mixed layer than the LR-HIST mean (Fig. 8), as anticipated due to its relatively weaker density stratification (Sect. 3.2). HR-HIST is also in better agreement with EN4-derived estimates. The LS March mixed layer reaches 1000–1200 m deep in LR-HIST, 1800–2000 m in HR-HIST, and 2000–2200 m in EN4. If we check the individual models (Fig. 7) we note that all the HR-HIST models show deep mixed layers in the LS, meanwhile ~25 % of the LR-HIST models show little or no convection in that area. Notice too, that the convection area along the East Greenland current, in the western IS, is also deeper and better connected with the mixed layers of the LS in the HR-HIST ensemble with respect to LR-HIST, better resembling the EN4 pattern.

Additional deepening of the mixed layer at higher resolution is found in the Nordic Seas, where the multi-model means reach depths of 1400–1600 m for HR-HIST versus only 1000–1200 m for LR-HIST. EN4-derived values are again the

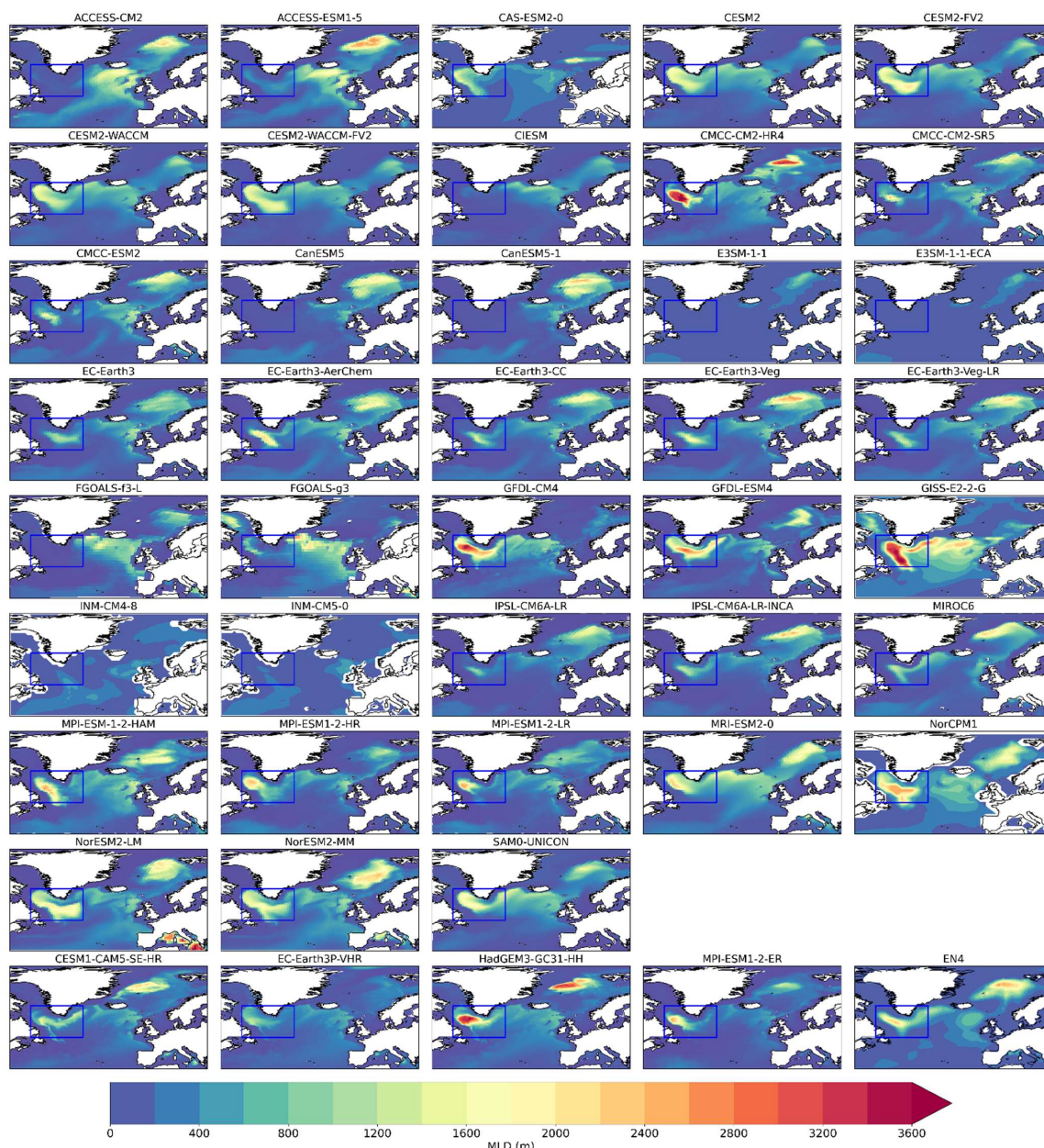
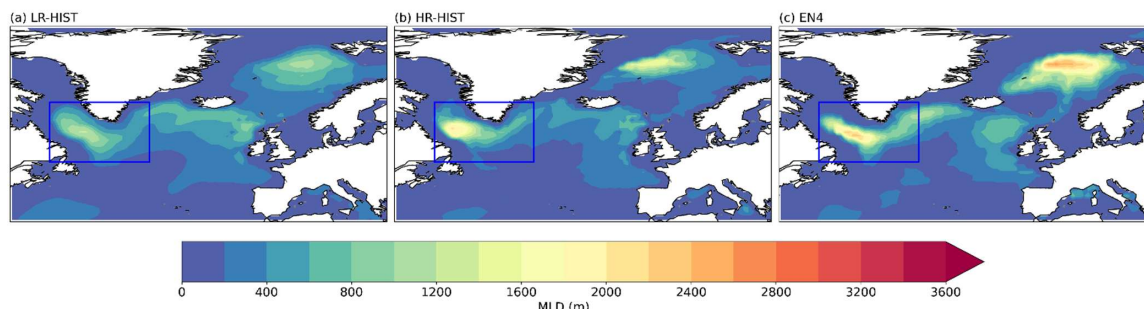



Figure 7. March mixed layer depth (MLD; in m) for LR-HIST (rows 1–8) and HR-HIST models and EN4 (row 9). MLD is calculated
using the potential density threshold method described in de Boyer Montégut et al. (2004), with a threshold value of 0.03 kg m^{-3} and a
reference depth of 10 m. The blue box (35º–60º W, 50º–65º N) indicates the region used in the vertical profiles calculations in Sect. 3.2.

270

largest, reaching down to 2200–2400 m. We note that the core of the deep mixing area in the Nordic Seas in HR-HIST is displaced westward compared with both EN4 and LR-HIST.

275   In the eastern IS and in the Iceland basin, the MLD ensemble mean shows values down to 600–800 m for LR-HIST and 400–600 m for HR-HIST. EN4 values are again larger, reaching down to 1000–1200 m. Next section will address if these overall improvements in deep mixing for HR-HIST are accompanied by an enhanced representation of the AMOC.

280   **Figure 8.** March MLD (in m) by groups, for (a) LR-HIST, (b) HR-HIST, and (c) EN4. Plotting details as in Fig. 7.

3.4 Atlantic meridional overturning circulation

        The AMOC streamfunction is a measure of the northward ocean volume transport, integrated zonally over the Atlantic basin and cumulatively from the top of the ocean as a function of depth. The AMOC in the depth space for the individual
285   models and reanalysis is shown in Fig. 9, and their LR-HIST and HR-HIST ensemble means are compared with ORAS5m reanalysis in Fig. 10. The AMOC is weaker in the multi-model HR-HIST mean compared to the LR-HIST one, and closer in strength and structure to the ORAS5m reanalysis data. This improved agreement with ORAS5m can partly be attributed to the use of an eddy-permitting ocean model in ORAS5m, allowing for the representation of some eddies. In both LR-HIST and HR-HIST, the upper cell of AMOC is shallower compared to ORAS5m, with the return branch reaching depths of 3000
290   m in models vs. 3700 m in reanalysis data. In the multi-model plot (Fig. 9), we observe that members are much more homogeneous in terms of AMOC structure and intensity within HR-HIST than within LR-HIST, with maximum climatological values close to 35º N in all four HR-HIST models.

        In order to compare model results with direct observational evidence, we examine the AMOC at 26.5º N, where RAPID volume transports are available for the 2004–2022 period (Fig. 11a). The vertical profile at 26.5º N indicates a weaker
295   AMOC for the multi-model mean of HR-HIST (max. 17 Sv) compared to LR-HIST (max. 18.8 Sv), which is also closer to observations (max. 16.6 Sv) and the reanalysis (max. 15.3 Sv), in line with our results from Fig. 10. The HR-HIST mean

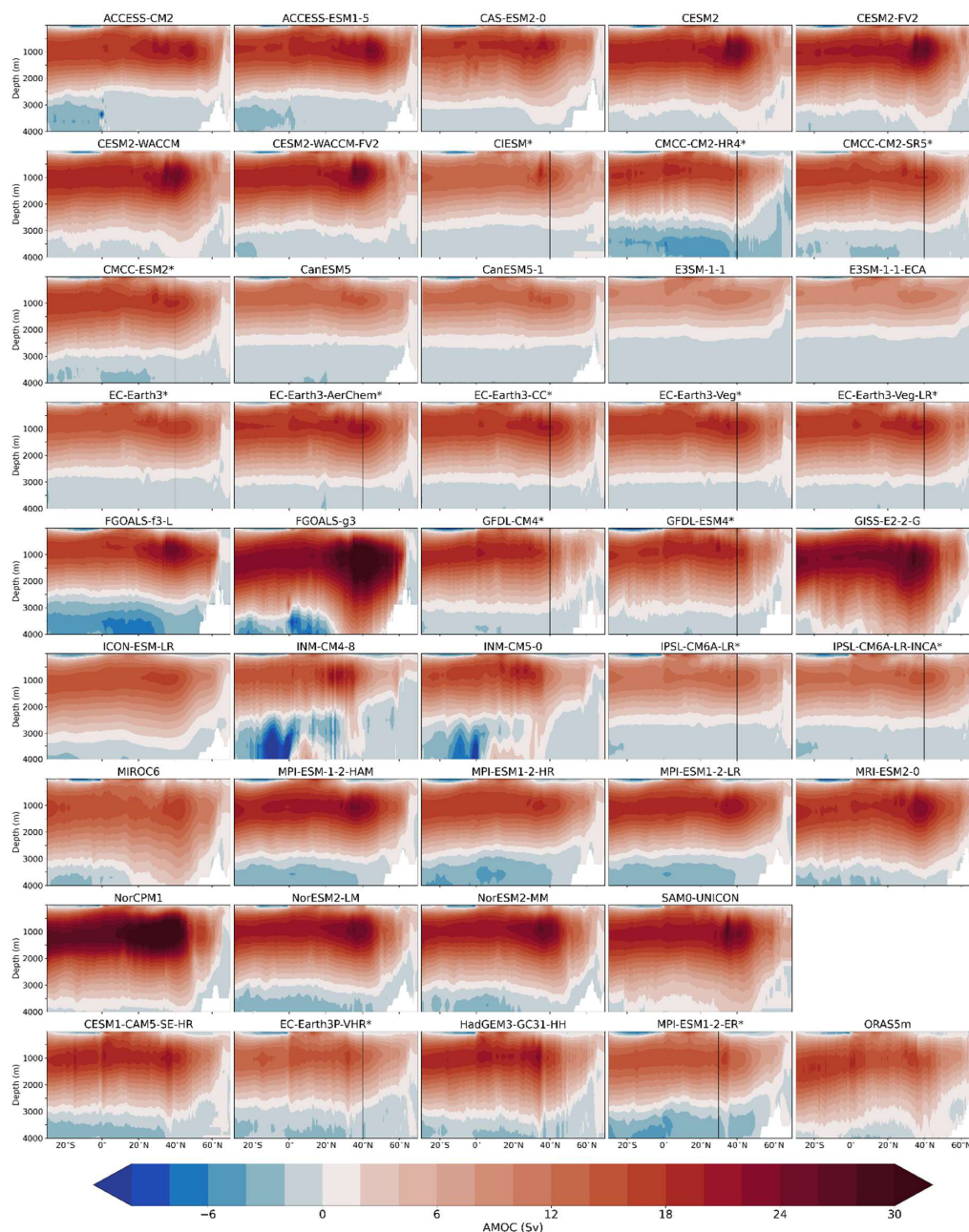



Figure 9. AMOC (in Sv) for LR-HIST (rows 1–8), and HR-HIST models and reanalysis (last row). The asterisks indicate when msftyz is displayed. Values at latitudes north of the black vertical lines might be affected by grid rotation (Sect. 2.1).

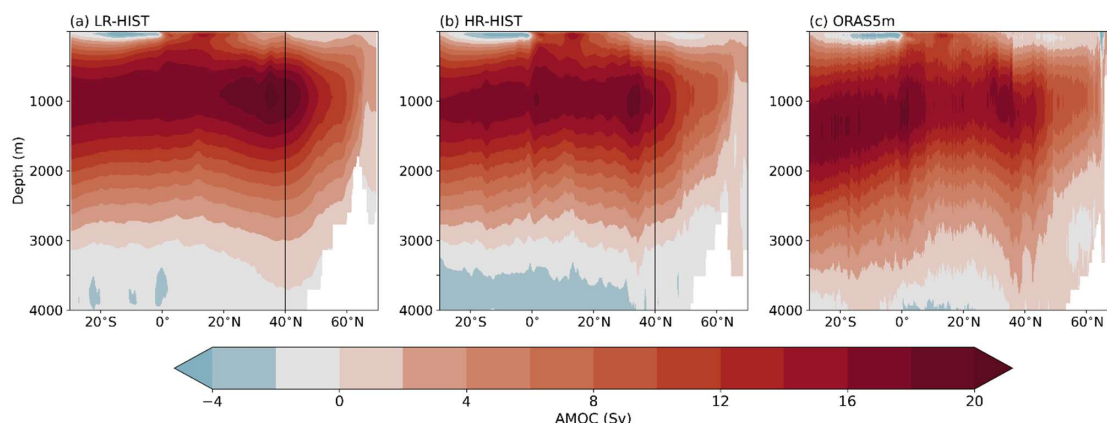



Figure 10. AMOC (in Sv) by groups, for (a) LR-HIST, (b) HR-HIST and (c) ORAS5m reanalysis. The black vertical line indicates the latitude from which grid rotation could have an effect (Sect 2.1).

profile shows a particularly good fit with the RAPID array one above ~1000 m, although, in general, the AMOC is too

305 shallow both for LR-HIST and HR-HIST. This is in part due to differences in the methodological approach (Danabasoglu et al., 2021). HR-HIST models remain relatively close to the RAPID data, and the main outliers both in terms of under- and overestimation of the AMOC are LR-HIST models (Fig. 11a).

To perform a quantitative comparison between the two ensembles, we extend the analysis of AMOC profiles by computing two metrics that measure the degree of agreement of the different models with observations, as diagnosed by the

310 Pearson correlation (x-axis in Fig. 11b) and the RMSE (y-axis in Fig. 11b) across the vertical dimension. Figure 11b confirms that, although none of the HR-HIST models is systematically better than all the LR-HIST ones, HR-HIST models lie within the range of best performing models, both in terms of vertical correlation and RMSE against RAPID, with HR-HIST models concentrated close to the bottom-right corner of the figure. To complement this analysis of the impact of resolution on the overturning circulation, next section looks at the impact on the gyre circulations in the North Atlantic.

315 **3.5 Gyre circulations**

In this section, the main gyre circulations of the North Atlantic are examined, as described by the barotropic streamfunction (BSF), a measure of the vertically integrated volume transport. The gyre circulations play a key role in climate in terms of northward ocean heat and freshwater transport and deep water formation. In order to validate the position of the GS and NAC in models, we also plot the zero contour line of absolute dynamic topography from AVISO observations,

320 which delimits the intergyre boundary (dashed lines in Figs. 12 and 13). Ideally, this zero line in observations would overlap the zero line of the model BSF, as for ORAS5m in Fig. 13c.

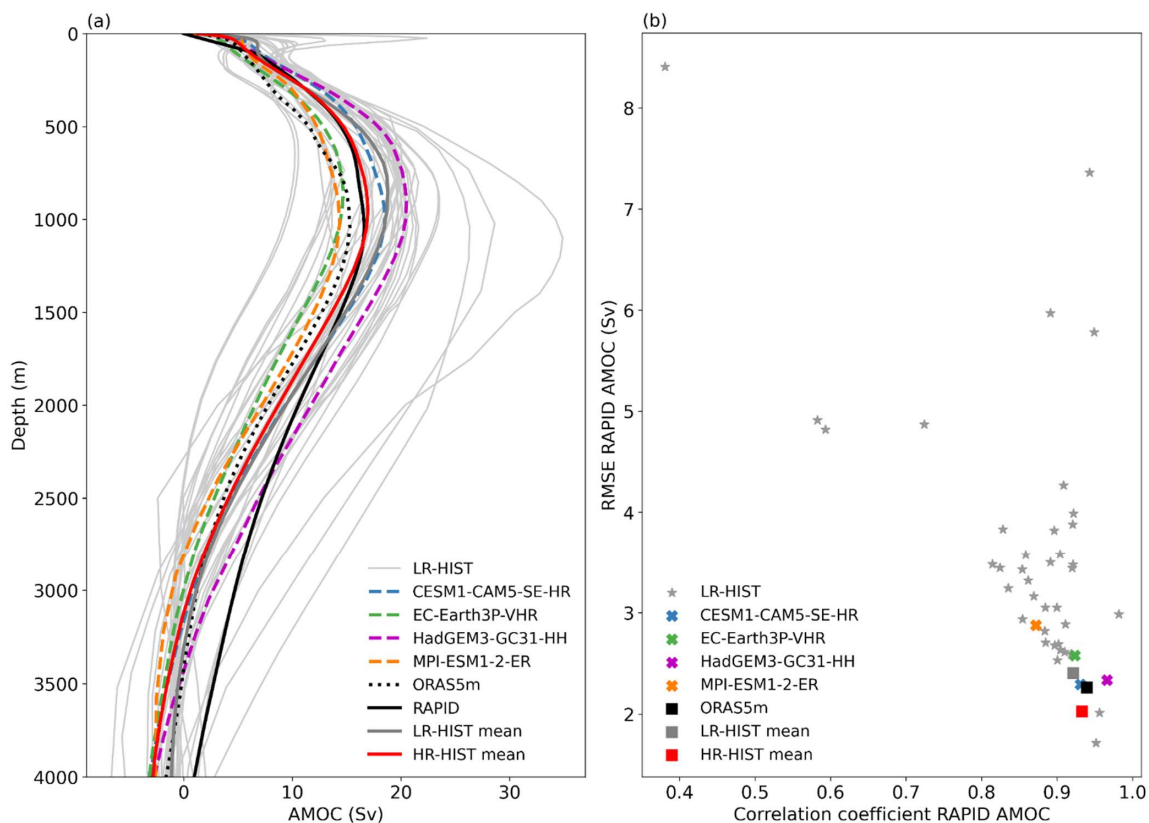



Figure 11. (a) Climatological AMOC profile at 26.5º N (in Sv). Model and reanalysis (ORAS5m) data correspond to the interval 1980–2014 (monthly data). RAPID observations are averaged over the period April 2004 – February 2022. (b) Pearson correlation coefficient (horizontal axis) and Root Mean Square Error (RMSE; in Sv) (vertical axis) of AMOC profiles at 26.5º N against RAPID, both estimated across the vertical dimension.

In the multi-model mean of LR-HIST, the GS separates too far north from the American coast compared to AVISO, which implies that its NCH bias region (Fig. 13a, red polygon) is only influenced by warmer, more saline waters of southern origin, in contrast with ORAS5m, where slight entrainment of colder, fresher waters from the north occurs. This can therefore explain the positive temperature and salinity biases described in Sect. 3.1.

In the multi-model mean of HR-HIST, the GS separates from the coast further south compared to LR-HIST, but also slightly compared to AVISO, and the NCH region is partially influenced by waters of northern origin, as in ORAS5m (Fig. 13b, c). These results can explain why the HR-HIST models show comparably reduced SST and SSS biases with respect to LR-HIST (Figs. 2 and 4). Furthermore, HR-HIST displays an improved GS structure (e.g. in the Florida Current) with a

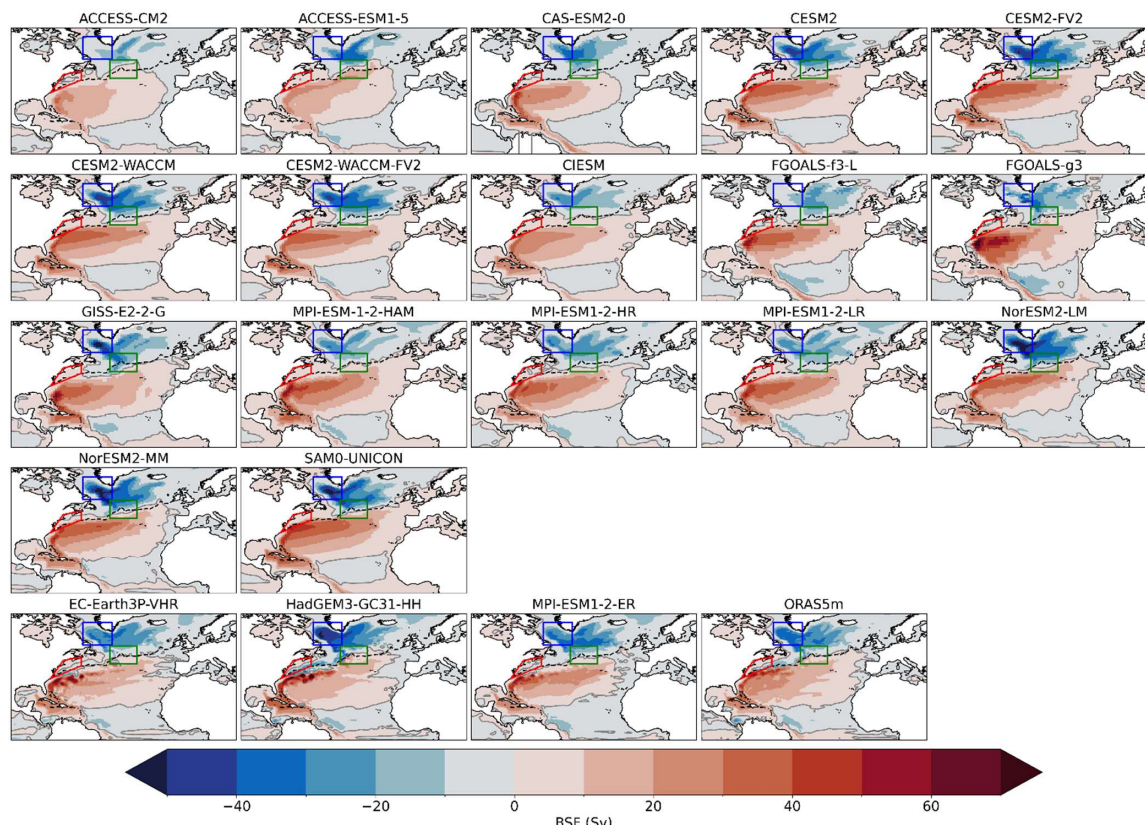



Figure 12. Barotropic streamfunction (BSF; in Sv) for LR-HIST (rows 1–4) and HR-HIST models and reanalysis (ORAS5m) (last row). Zero contour of absolute dynamic topography from AVISO observations (dashed black line) corresponding to the period 1993–2014 is also shown. Boxes as in Figs. 2 and 4: NCH in red, CNA in green, LS in blue. Note: some models did not present a BSF output (e.g. CESM1-CAM5-SE-HR) or this presented unrealistic values, hence the reason fewer models are shown in this figure.

narrower and locally stronger current than LR-HIST (Fig. 13b), in closer agreement with ORAS5m, although this could be partly explained by the fact that ORAS5m was produced with an eddy-permitting ocean.

The NAC is too zonal in most LR-HIST models (Fig. 12). In the LR-HIST ensemble mean, the CNA region is only touched at its southern edge by the warmer/saltier NAC waters, remaining predominantly exposed to the influence of the SPG (Fig. 13a). By contrast for ORAS5m and HR-HIST, the SPG has a more restrained influence on that region in favor of a less zonal NAC (Fig. 13b, c), which could explain the reduced biases at high resolution in the CNA for HR-HIST (Sect. 3.1.). This is a common feature in all available individual HR-HIST models for which the BSF is available: EC-Earth3P-VHR, HadGEM3-GC31-HH, and MPI-ESM1-2-ER.

The SPG is generally stronger in HR-HIST compared to LR-HIST, and closer in strength and structure to ORAS5m (Figs. 12 and 13), although some individual LR-HIST models like NorESM2-LM and SAM0-UNICON have gyres of

comparable intensity to HadGEM3-GC31-HH, which has the strongest SPG across HR-HIST models. We note that a stronger SPG is consistent with the stronger convection identified in Sect. 3.3, as this latter enhances the presence of dense waters in the deep ocean, subsequently strengthening the baroclinic pressure gradient that drives the gyre (Yashayaev and Loder, 2009).

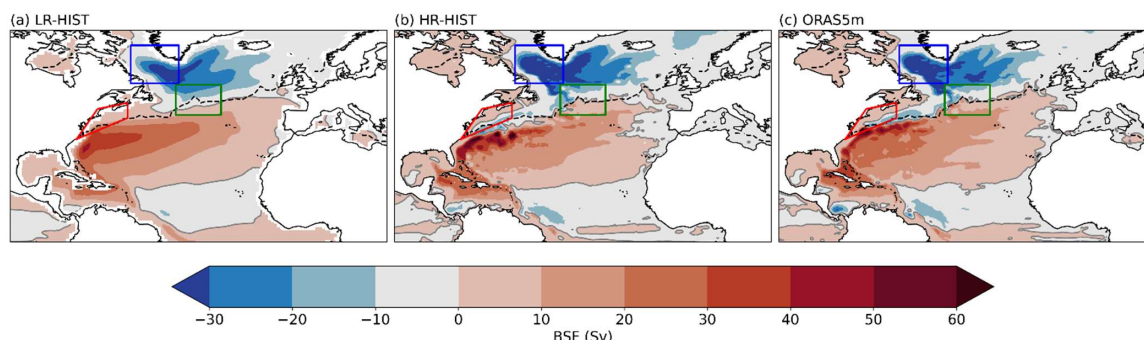



Figure 13. BSF (in Sv) by groups for (a) LR-HIST, (b) HR-HIST and (c) reanalysis (ORAS5m). Plotting details as in Fig. 12.

4 Main conclusions and discussion

In this study, we analyse the impact of horizontal resolution on the representation of the North Atlantic mean state, by comparing two ensembles of coupled historical simulations: four HighResMIP experiments at eddy-resolving scales (HR-HIST ensemble; at least 1/10º nominal resolution) and 39 CMIP6 experiments with eddy-parameterized or eddy-permitting ocean resolutions (LR-HIST ensemble).

The main biases of key thermodynamic and dynamical variables for the North Atlantic are analysed for the two ensembles. In particular we examine i) the main surface temperature and salinity biases; ii) stratification and iii) deep water convection in the North Atlantic; iv) the representation of the AMOC; and v) the gyre circulations of the North Atlantic, including the GS, NAC, and the SPG. In the following, the main findings of the paper are described and their implications discussed in light of the previous literature.

Three main SST and SSS bias regions are found in the simulations, located at NCH, the CNA, and the LS, which show significant differences across the two ensembles. In the NCH region, we find reduced positive temperature and salinity surface biases for the multi-model HR-HIST mean with respect to LR-HIST, associated with a more southward position of the GS separation, in agreement with previous individual model studies (Roberts et al., 2019; Gutjahr et al., 2019; Marzocchi et al., 2015).

Then, the CNA cold and fresh biases in the multi-model LR-HIST mean are also reduced in HR-HIST – as in Gutjahr et al. (2019) and Marzocchi et al. (2015) – which describes a less zonal NAC and a more restricted influence of SPG waters in

that region, in closer agreement with observations and reanalysis. Sein et al. (2018) found similar results when comparing HighResMIP coupled simulations obtained with the AWI-CM model, showing that ultimately it was an increase in ocean resolution that shifted the NAC path towards the north, with no significant influence of atmospheric resolution. Surface
380 biases in the NCH and CNA regions might be connected with each other due to their ultimate link with the GS/NAC dynamics. As expected, the correlation between the surface biases of these two regions increases when the HR-HIST models are excluded from the computation (Fig. A1), which reflects that the NCH region is outside the GS domain in HR-HIST models.

The LS region stands out for a warm and salty bias in the HR-HIST multi-model mean. For temperature, it is also present
385 in the LR-HIST ensemble mean, although much weaker, whereas for salinity, it does not show in the LR-HIST ensemble mean, even if it is present in some of the individual models. Indeed, some individual LR-HIST models show LS biases of comparable strength to the HR-HIST ones, but their signal is compensated in the multi-model mean by models with biases of opposite sign. The origin and dynamical impacts of the LS biases are a current matter of debate (Menary et al., 2015; Roberts et al., 2020). These surface biases might have an effect on LS deep water convection through their decisive influence on
390 vertical stratification and, therefore, correcting them might help obtaining more realistic present-day AMOC estimates and reliable future projections. Further efforts should thus aim at identifying the source/origin of these biases. Our study hints that CNA and LS biases might be actually related. Scatterplots of SSS biases between both regions (Fig. A2) indicate a strong correlation between them (r = 0.86, p < 0.001), which suggests a potential link between LS salinity biases and the NAC, through the effect of the NAC on the northward salinity transport. We note, however, that the LS and CNA are also
395 connected through the SPG circulation, which could also partly explain why their SSS biases are related. The correlation between the SST biases of the LS and CNA regions is also significant, although weaker compared to the SSS biases (r = 0.54, p < 0.001; Fig. A2), which might indicate additional differences between the mechanisms exerting control over the SSTs in both regions, like the local atmospheric forcing. Studies such as Chang et al. (2020) and Roberts et al. (2019) report increased heat transport by the AMOC in eddy-resolving models, further supporting the idea of increased northward
400 transport as a potential origin for the LS biases.

Interestingly, we find improved vertical profiles of temperature and salinity in the LIS box for HR-HIST: despite the larger biases found at the surface with respect to LR-HIST, the subsurface is colder and fresher in HR-HIST. This might be related to increased vertical (upwards) heat and salt transports by ocean mesoscale eddies (Hewitt et al., 2017), providing a potential explanation for the warm and salty surface biases in the LS. The mean density profile, which is the key property
405 controlling the vertical mixing, is also improved for the LIS in HR-HIST with respect to LR-HIST, showing comparatively reduced stratification. This explains why HR-HIST presents deeper mixed layers in the LS and along the east Greenland coast than LR-HIST, in closer agreement with EN4-derived values, an improvement that is also found in the Nordic Seas. Since EN4 might present some uncertainties in the ocean subsurface, in particular before the Argo period (1999 to present), a comparison with other more recent MLD estimates is warranted to assess the consistency of our results. A MLD climatology
410 for the 2000–2016 period based on Argo density profiles by Holte et al. (2017) shows mixed layers down to a maximum of

1800 m in the LS. Time-varying estimates of winter maximum MLDs in the LS obtained from Argo floats, the AR7W line, and moored measurements, suggest values around 1200–1700 m in the 2002–2015 interval (Yashayaev and Loder, 2016), showing an intensification in recent years, with a record value of 2100 m in 2016 (Yashayaev and Loder, 2017). Overall, the maximum values in the HR-HIST ensemble mean for the LS (1800–2000 m) are slightly larger than the direct observational

415   estimates, excluding the record values in Yashayaev and Loder (2017). However, it is important to note that some differences are expected as our HR-HIST values are computed from 1980 to 2014, and use generally smoother profiles associated with the coarser temporal resolution of the model data compared to the individual profiles from observational studies.

The AMOC is weaker in the multi-model HR-HIST mean compared to LR-HIST and its strength and structure exhibit a

420   better fit with RAPID observations and the ocean reanalysis ORAS5m. Additionally, the HR-HIST AMOC presents sharper features, as shown e.g. in Sein et al. (2018), better resembling reanalysis data. Nevertheless, as pointed out in previous studies (Roberts et al., 2020; Hirschi et al., 2020), the AMOC remains too shallow compared to RAPID both in the LR-HIST and HR-HIST ensemble means.

The role of resolution in AMOC strength is a current matter of debate in the literature, with different individual model

425   studies pointing at different results. Winton et al. (2014) report AMOC strengthening with increased ocean resolution for the GFDL CM2.6 and CM2.5FLOR models at 0.1º and 1º resolution, respectively. They also find AMOC strength is sensitive to horizontal friction and mesoscale eddy parameterizations. Hewitt et al. (2016) show strengthening in the mean AMOC at a concomitant increase in ocean (from 1/4º to 1/12º) and atmospheric resolution (from 60 to 25 km) in the GC2.1 model. Similar results are found by Moreno-Chamarro et al. (2024) for the EC-Earth3P model when increasing ocean and

430   atmospheric resolution from 0.25º to 0.08º and from ~54 km to ~12 km, respectively. On the other hand, a study assessing the separate effects of enhanced atmospheric and ocean resolution on AMOC behaviour with the AWI-CM model, describes a weakening at increased atmospheric resolution (from 1.9º to 0.9º) associated with reduced winds, but both a weakening at ~45º N and a strengthening at ~20º N related to ocean grid refinement (from 1º to 1/4° nominal resolutions, and the later with grid refinements in eddy-rich areas; Sein et al., 2018). Furthermore, Gutjahr et al. (2019) show little difference in AMOC

435   strength between MPI-ESM1-2-HR and MPI-ESM1-2-ER, which use the same atmosphere and vertical mixing parameterization yet different ocean resolution (0.4º vs 0.1º, respectively). That study also shows AMOC can be very sensitive to the vertical mixing scheme.

Multimodel studies on this topic have also been conducted (Roberts et al., 2020; Hirschi et al., 2020). Hirschi et al. (2020) analyze 28 models (22 ocean-only and six coupled models) with ocean resolutions ranging from 2º to 0.05º and find

440   increased AMOC strength at eddy-resolving scales (their Fig. 2). Roberts et al. (2020) compare the AMOC in HighResMIP simulations with seven different coupled models, not finding a consistent effect of enhancing ocean and/or atmospheric resolution on the AMOC strength in the depth space. This also applies for the two simulations at eddy-resolving scales in that study, performed with HadGEM3-GC31 and CESM1.3 (our CESM1-CAM5-SE), the first showing a stronger AMOC in

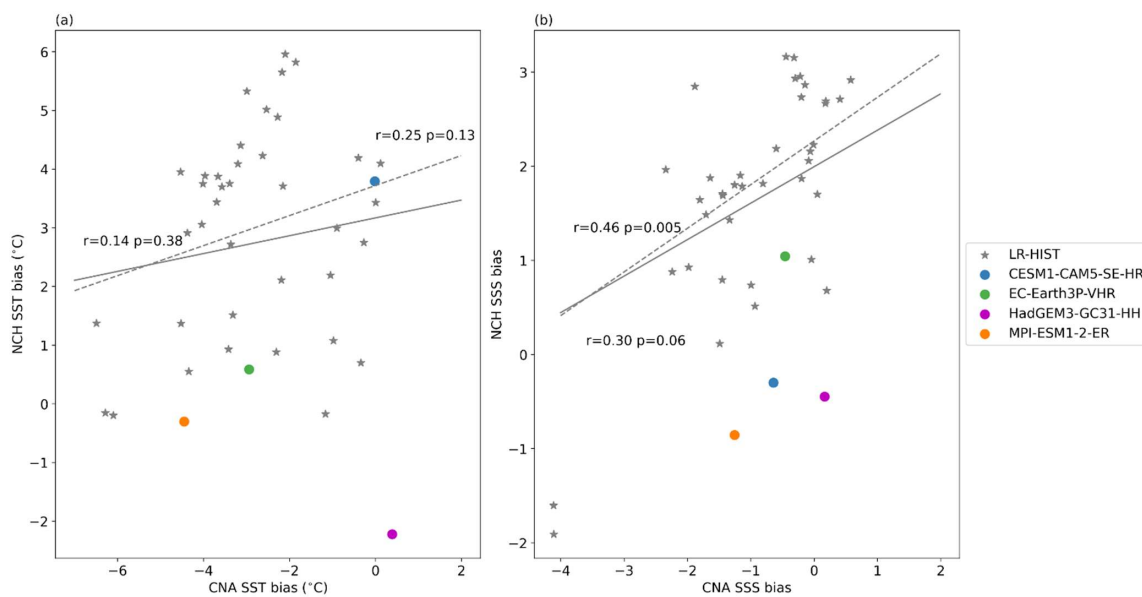
depth-space than its low resolution counterpart, and the second showing a weaker AMOC instead. Interestingly, in that study
445 results converge towards a stronger AMOC in the eddy-resolving simulations when density coordinates are used instead.

Meanwhile, our results show a weaker AMOC at eddy-resolving scales. Nevertheless, we note that our AMOC profiles of eddy-resolving models at 26.5º N display similar values to those in Roberts et al. (2020) and Hirschi et al. (2020). The differences in the results lie rather in the characteristics of the low resolution model ensembles, which in those studies show considerable lower AMOC values compared to ours, probably in relation with the high sensitivity of AMOC to model
450 schemes and parameterizations (see e.g. Winton et al., 2014; Gutjahr et al., 2019 above). Nevertheless, all three multimodel studies – Hirschi et al. (2020), Roberts et al. (2020), and our study – point at an improved AMOC mean-state representation at enhanced resolution.

Although a link exists between AMOC strength and SPNA densities/mixed layers (Ortega et al., 2021; Menary et al., 2020; Martin-Martinez et al., in review), in our study, the deeper mixed layers in the multi-model HR-HIST mean with
455 respect to LR-HIST despite a weaker AMOC, reflect a different representation of deep water sinking mechanisms in high resolution models, as described in Katsman et al. (2018). That study shows that deep water sinking in eddy-permitting models occurs only at the continental slopes – at the boundary current of the SPG – and not also in the open ocean where MLDs reach their maximum depths, as in 1º ocean models. The sinking mechanism described for eddy-permitting models can be explained by buoyancy loss along the boundary current path, triggering a cross-shore baroclinic flow and subsequent
460 sinking forced by mass conservation (Katsman et al., 2018; Straneo, 2006; Spall and Pickart, 2001). The more realistic SPG structure and MLDs near the continental boundaries in HR-HIST, e.g along the Greenland Current, might thus be key factors for the improvement in modeled AMOC strength in HR-HIST. Results by Martin-Martinez et al. (in review) show a faster southward propagation of the MLD signal into the AMOC in eddy-resolving models, indicating that ocean resolution affects also the timescales of the dynamics of the North Atlantic. Finally, a stronger SPG is a robust feature in HR-HIST with
465 respect to LR-HIST, as already pointed out by Hirschi et al. (2020) when comparing a range of ocean resolutions from 1º to 0.08º, and has been shown to be highly correlated with a deeper mixed layer (Koenigk et al., 2021).

To summarize, compared to a large low-resolution CMIP6 ensemble, vertical stratification in the LS and western IS is improved in eddy-resolving models, leading to stronger deep water convection closer to observations. Surface temperature and salinity biases are reduced in the NCH and CNA regions associated with improved GS and NAC paths. The AMOC is
470 weaker and closer to RAPID and the SPG is stronger and presents an improved structure.

475

Appendix A: Scatterplots of surface biases

480   **Figure A1.** Scatterplots of (a) SST (in ºC) and (b) SSS biases between the Central North Atlantic (CNA) and North Cape Hatteras (NCH) regions defined in Figs. 1 and 2. Biases are calculated as spatially averaged temporal means in the model minus the corresponding EN4 values in each of the selected boxes. The corresponding correlation coefficients and their p-values are shown next to the fit lines. Dashed lines are regression lines obtained after removal of HR-HIST models. Note: all regression lines and values in (b) are calculated excluding the two outliers at the bottom left of the figure.

485

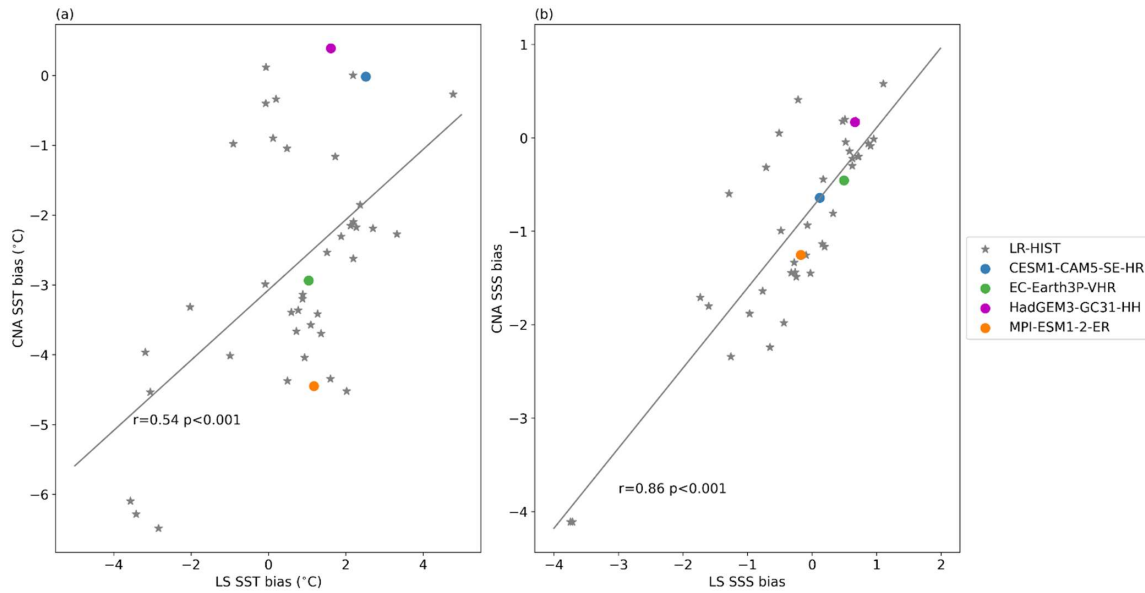



Figure A2. As in Fig. A1 but between the Central North Atlantic (CNA) and the Labrador Sea (LS).

490

495

500

505

Appendix B: Additional model details

	ocean component	ocean grid	atm. component	atm. grid
HR-HIST				
CESM1-CAM5-SE-HR	POP2	1/10°; tripolar; 3600x2400 lon/lat; 62 levels;	CAM5.2	25 km; 30 levels;
EC-Earth3P-VHR	NEMO3.6	1/12°; ORCA12 tripolar; 4322 x 3059 lon/lat; 75 levels;	IFS cy36r4	16 km; 91 levels;
HadGEM3-GC31-HH	NEMO-HadGEM3-GO6.0	1/12°; eORCA12 tripolar; 4320 x 3604 lon/lat; 75 levels;	MetUM-HadGEM3-GA7.1	50 km; 85 levels;
MPI-ESM1-2-ER	MPIOM	1/10°; TP6M tripolar; 3602 x 2394 lon/lat; 40 levels;	ECHAM6.3	103 km; 95 levels;
LR-HIST				
ACCESS-CM2	ACCESS-OM2	100 km; GFDL-MOM5 tripolar; 360 x 300 lon/lat; 50 levels;	MetUM-HadGEM3-GA7.1	250 km; 85 levels;
ACCESS-ESM1-5	ACCESS-OM2	100 km; MOM5 tripolar; 360 x 300 lon/lat; 50 levels;	HadGAM2	250 km; 38 levels;
CAS-ESM2-0	LICOM2.0	100 km; 362 x 196 lon/lat; 30 levels;	IAP AGCM 5.0	100 km; 35 levels;
CESM2	POP2	100 km; gx1v7 displaced pole; 320x384 lon/lat; 60 levels;	CAM6	100 km; 32 levels;
CESM2-FV2	POP2	100 km; gx1v7, displaced pole; 320 x 384 lon/lat; 60 levels;	CAM6	250 km; 32 levels;
CESM2-WACCM	POP2	100 km; gx1v7 displaced pole; 320 x 384 lon/lat; 60 levels;	WACCM6	100 km; 70 levels;
CESM2-WACCM-FV2	POP2	100 km; gx1v7 displaced pole; 320 x 384 lon/lat; 60 levels;	WACCM6	250 km; 70 levels;
CIESM	CIESM-OM	100 km; mod. POP2 displ. pole; 320 x 384 lon/lat; 60 levels;	CIESM-AM (modified CAM5)	100 km; 30 levels;
CMCC-CM2-HR4	NEMO3.6	25 km; ORCA0.25; 1442 x 1051 lon/lat; 50 levels;	CAM4	100 km; 26 levels;
CMCC-CM2-SR5	NEMO3.6	100 km; ORCA1 tripolar; 362 x 292 lon/lat; 50 levels;	CAM5.3	100 km; 30 levels;
CMCC-ESM2	NEMO3.6	100 km; ORCA1 tripolar; 362 x 292 lon/lat; 50 levels;	CAM5.3	100 km; 30 levels;
CanESM5	NEMO3.4.1	100 km; ORCA1 tripolar; 361 x 290 lon/lat; 45 levels;	CanAM5	500 km; 49 levels;
CanESM5-1	NEMO3.4.1	100 km; ORCA1 tripolar; 361 x 290 lon/lat; 45 levels;	CanAM5.1	500 km; 49 levels;
E3SM-1-1	MPAS-Ocean (v6.0)	30-60 km; oEC60to30 unstructured; 60 levels;	EAM (v1.1)	100 km; 72 levels;

Table B1. Overview of individual models used in the current study. Ocean grid details include: nominal resolution; grid type; size of horizontal grid; and number of vertical levels. Expanded version of Table 1 (part1).

	ocean component	ocean grid	atm. component	atm. grid
E3SM-1-1-ECA	MPAS-Ocean (v6.0)	30-60 km; oEC60to30 unstructured; 60 levels;	EAM (v1.1)	100 km; 72 levels;
EC-Earth3	NEMO3.6	100 km; ORCA1 tripolar; 362 x 292 lon/lat; 75 levels;	IFS cy36r4	100 km; 91 levels;
EC-Earth3-AerChem	NEMO3.6	100 km; ORCA1 tripolar; 362 x 292 lon/lat; 75 levels;	IFS cy36r4	100 km; 91 levels;
EC-Earth3-CC	NEMO3.6	100 km; ORCA1 tripolar; 362 x 292 lon/lat; 75 levels;	IFS cy36r4	100 km; 91 levels;
EC-Earth3-Veg	NEMO3.6	100 km; ORCA1 tripolar; 362 x 292 lon/lat; 75 levels;	IFS cy36r4	100 km; 91 levels;
EC-Earth3-Veg-LR	NEMO3.6	100 km; ORCA1 tripolar; 362 x 292 lon/lat; 75 levels;	IFS cy36r4	250 km; 62 levels;
FGOALS-f3-L	LICOM3.0	100 km; tripolar; 360 x 218 lon/lat; 30 levels	FAMIL2.2	100 km; 32 levels;
FGOALS-g3	LICOM3.0	100 km; tripolar; 360 x 218 lon/lat; 30 levels;	GAMIL3	250 km; 26 levels;
GFDL-CM4	GFDL-OM4p25	25 km; GFDL-MOM6 tripolar; 1440 x 1080 lon/lat; 75 levels;	GFDL-AM4.0.1	100 km; 33 levels;
GFDL-ESM4	GFDL-OM4p5	50 km; GFDL-MOM6 tripolar; 720 x 576 lon/lat; 75 levels;	GFDL-AM4.1	100 km; 49 levels;
GISS-E2-2-G	GISS Ocean	100 km; GO1; 360 x 180 lon/lat; 40 levels;	GISS-E2.2	250 km; 102 levels;
ICON-ESM-LR	ICON-O	50 km; icosahedral/triangles; 40 levels;	ICON-A	250 km; 47 levels;
INM-CM4-8	INM-OM5	100 km; shifted North Pole; 360 x 318 lon/lat; 40 levels;	INM-AM4-8	100 km; 21 levels;
INM-CM5-0	INM-OM5	50 km; shifted North Pole; 720 x 720 lon/lat; 40 levels.	INM-AM5-0	100 km; 73 levels;
IPSL-CM6A-LR	NEMO-OPA	100 km; eORCA1.3 tripolar; 362 x 332 lon/lat; 75 levels;	LMDZ	250 km; 79 levels;
IPSL-CM6A-LR-INCA	NEMO-OPA	100 km; eORCA1.3 tripolar; 362 x 332 lon/lat; 75 levels;	LMDZ	250 km; 79 levels;
MIROC6	COCO4.9	100 km; tripolar; 360 x 256 lon/lat; 63 levels;	CCSR AGCM	250 km; 81 levels;
MPI-ESM-1-2-HAM	MPIOM1.63	250 km; bipolar GR1.5; 256 x 220 lon/lat; 40 levels;	ECHAM6.3	250 km; 47 levels;
MPI-ESM1-2-HR	MPIOM1.63	50 km; tripolar TP04; 802 x 404 lon/lat; 40 levels;	ECHAM6.3	100 km; 95 levels;

LR-HIST

Table B1. Continuation (part2).

510

		ocean component	ocean grid	atm. component	atm. grid
LR–HIST	MPI-ESM1-2-LR	MPIOM1.63	250 km; bipolar GR1.5; 256 x 220 lon/lat; 40 levels;	ECHAM6.3	250 km; 47 levels;
	MRI-ESM2-0	MRI.COM4.4	100 km; tripolar; 360 x 364 lon/lat; 61 levels;	MRI-AGCM3.5	100 km; 80 levels;
	NorCPM1	MICOM1.1	100 km; displaced pole; 320 x 384 lon/lat; 53 levels;	CAM-OSLO4.1	250 km; 26 levels;
	NorESM2-LM	MICOM	100 km; tripolar; 360 x 384 lon/lat; 70 levels;	CAM-OSLO	250 km; 32 levels;
	NorESM2-MM	MICOM	100 km; tripolar; 360 x 384 lon/lat; 70 levels;	CAM-OSLO	100 km; 32 levels;
	SAM0-UNICON	POP2	100 km; displaced pole; 320 x 384 lon/lat; 60 levels;	CAM5.3 with UNICON	100 km; 30 levels;

Table B1. Continuation (part3).

Code and data availability. The ESMValTool code will be made available in the revised version. Model data used in this study can be found on the Earth System Grid Federation (ESGF) site (https://esgf-ui.ceda.ac.uk/cog/search/cmip6-ceda/),

515   except for: 1) HadGEM3-GC31-HH AMOC and BSF data, which are available upon request via the CEDA-JASMIN platform (https://www.ceda.ac.uk/services/jasmin/), 2) CESM1-CAM5-SE-HR AMOC data, which are available from NCAR's Climate and Global Dynamics lab (https://www.cgd.ucar.edu/) upon request, 3) MPI-ESM1-2-ER data, which are archived by the Max Planck Institute for Meteorology and can be obtained by contacting publications@mpimet.mpg.de, and 4) EC-Earth3P-VHR data, which will be published on ESGF soon, while at the moment are available upon request from the

520   Barcelona Supercomputing Center (BSC). ORAS5m data are available upon request from the ECMWF file storage system. AVISO absolute dynamic topography data (MADT-H) can be directly downloaded from this link https://www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/global/gridded-sea-level-anomalies-mean-and-climatology.html. RAPID AMOC data can be downloaded from https://rapid.ac.uk/data/. EN4 data are available from https://www.metoffice.gov.uk/hadobs/en4/index.html.

525   *Author contributions.* AF carried out the analysis with ESMValTool and wrote the manuscript. SLT and EMM provided technical support and guidance in ESMValTool. SLT made improvements for memory usage in ESMValTool. EMM wrote the ESMValTool code to calculate density and mixed layer from temperature and salinity. EMC performed the ECEarth-3P HighResMIP runs, provided scientific guidance, as well as input to the draft. PO provided scientific guidance, as well as input to the draft. Xia Lin provided input to the draft. MS, AF, PAB, and DK performed the data assemblage and formatting.

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

Acknowledgements. The authors thank Nikolay Koldunov for his ESMValTool arctic ocean diagnostics (Khosravi et al., 2022) shared on GitHub (https://github.com/ESMValGroup/ESMValTool/tree/main//esmvaltool/diag_scripts/arctic_ocean) under an Apache 2.0 license. We would also like to thank Steffen Tietsche and Hao Zuo for kindly providing the ORAS5m reanalysis data; Dian Putrasahan and Katja Lohmann for sharing all the MPI-ESM1-2-ER data; Malcolm Roberts for the

535   HadGEM3-GC31-HH AMOC and BSF data; as well as Gokhan Danabasoglu, James Hurrell, Frederic Castruccio, and Gary Strand for the CESM1-CAM5-SE-HR AMOC data. Further acknowledgements go to Albert Vila and David Vicente for the technical and HPC support at BSC, and to the BSC itself for providing the computing resources. AF thanks Malcolm Roberts and Michael Lai for the science discussions and seminars at Met Office.

Financial support. This research has been funded by the Spanish STREAM project (grant no. PID2020-114746GB-I00) and

540   the Horizon Europe EERIE project (grant no. 101081383).

References

Andela, B., Broetz, B., de Mora, L., Drost, N., Eyring, V., Koldunov, N., Lauer, A., Predoi, V., Righi, M., Schlund, M., Vegas-Regidor, J., Zimmermann, K., Bock, L., Diblen, F., Dreyer, L., Earnshaw, P., Hassler, B., Little, B., Loosveldt-Tomas, S., Smeets, S., Camphuijsen, J., Gier, B. K., Weigel, K., Hauser, M., Kalverla, P., Galytska, E., Cos-Espuña, P.,

545   Pelupessy, I., Koirala, S., Stacke, T., Alidoost, S., Jury, M., Sénési, S., Crocker, T., Vreede, B., Soares Siqueira, A., Kazeroni, R., Bauer, J., Beucher, R., and Benke, J.: ESMValCore, Zenodo [code], https://doi.org/10.5281/zenodo.10406626, 2023a.

Andela, B., Broetz, B., de Mora, L., Drost, N., Eyring, V., Koldunov, N., Lauer, A., Mueller, B., Predoi, V., Righi, M., Schlund, M., Vegas-Regidor, J., Zimmermann, K., Adeniyi, K., Arnone, E., Bellprat, O., Berg, P., Bock, L., Bodas-Salcedo,

550   A., Caron, L.-P., Carvalhais, N., Cionni, I., Cortesi, N., Corti, S., Crezee, B., Davin, E. L., Davini, P., Deser, C., Diblen, F., Docquier, D., Dreyer, L., Ehbrecht, C., Earnshaw, P., Gier, B., Gonzalez-Reviriego, N., Goodman, P., Hagemann, S., von Hardenberg, J., Hassler, B., Heuer, H., Hunter, A., Kadow, C., Kindermann, S., Koirala, S., Kuehbacher, B., Lledó, L., Lejeune, Q., Lembo, V., Little, B., Loosveldt-Tomas, S., Lorenz, R., Lovato, T., Lucarini, V., Massonnet, F., Mohr, C. W., Amarjiit, P., Pérez-Zanón, N., Phillips, A., Russell, J., Sandstad, M., Sellar, A., Senftleben, D., Serva, F., Sillmann, J.,

555   Stacke, T., Swaminathan, R., Torralba, V., Weigel, K., Sarauer, E., Roberts, C., Kalverla, P., Alidoost, S., Verhoeven, S., Vreede, B., Smeets, S., Soares Siqueira, A., Kazeroni, R., Potter, J., Winterstein, F., Beucher, R., Kraft, J., Ruhe, L., and Bonnet, P.: ESMValTool, Zenodo [code], https://doi.org/10.5281/zenodo.10408909, 2023b.

Bellomo, K. and Mehling, O.: Impacts and State-Dependence of AMOC Weakening in a Warming Climate, Geophys. Res. Lett., 51, e2023GL107624, https://doi.org/10.1029/2023GL107624, 2024.

560   Bellucci, A., Athanasiadis, P. J., Scoccimarro, E., Ruggieri, P., Gualdi, S., Fedele, G., Haarsma, R. J., Garcia-Serrano, J., Castrillo, M., Putrahasan, D., Sanchez-Gomez, E., Moine, M.-P., Roberts, C. D., Roberts, M. J., Seddon, J., and Vidale, P. L.: Air-Sea interaction over the Gulf Stream in an ensemble of HighResMIP present climate simulations, Clim. Dyn., 56, 2093–2111, https://doi.org/10.1007/s00382-020-05573-z, 2021.

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D.: Mixed layer depth over the global ocean:

565   An examination of profile data and a profile-based climatology, J. Geophys. Res. Oceans, 109, C12003, https://doi.org/10.1029/2004JC002378, 2004.

Brown, P. J., McDonagh, E. L., Sanders, R., Watson, A. J., Wanninkhof, R., King, B. A., Smeed, D. A., Baringer, M. O., Meinen, C. S., Schuster, U., Yool, A., and Messias, M.-J.: Circulation-driven variability of Atlantic anthropogenic carbon transports and uptake, Nat. Geosci., 14, 571–577, https://doi.org/10.1038/s41561-021-00774-5, 2021.

570   Buckley, M. W. and Marshall, J.: Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review, Rev. Geophys., 54, 5–63, https://doi.org/10.1002/2015RG000493, 2016.

Chang, P., Zhang, S., Danabasoglu, G., Yeager, S. G., Fu, H., Wang, H., Castruccio, F. S., Chen, Y., Edwards, J., Fu, D., Jia, Y., Laurindo, L. C., Liu, X., Rosenbloom, N., Small, R. J., Xu, G., Zeng, Y., Zhang, Q., Bacmeister, J., Bailey, D. A.,

Duan, X., DuVivier, A. K., Li, D., Li, Y., Neale, R., Stössel, A., Wang, L., Zhuang, Y., Baker, A., Bates, S., Dennis, J.,
575 Diao, X., Gan, B., Gopal, A., Jia, D., Jing, Z., Ma, X., Saravanan, R., Strand, W. G., Tao, J., Yang, H., Wang, X., Wei, Z.,
and Wu, L.: An Unprecedented Set of High-Resolution Earth System Simulations for Understanding Multiscale Interactions
in Climate Variability and Change, J. Adv. Model. Earth Syst., 12, e2020MS002298,
https://doi.org/10.1029/2020MS002298, 2020.

Danabasoglu, G., Castruccio, F. S., Small, R. J., Tomas, R., Frajka-Williams, E., and Lankhorst, M.: Revisiting AMOC
580 Transport Estimates From Observations and Models, Geophys. Res. Lett., 48, e2021GL093045,
https://doi.org/10.1029/2021GL093045, 2021.

Dey, D., Marsh, R., Drijfhout, S., Josey, S. A., Sinha, B., Grist, J., and Döös, K.: Formation of the Atlantic Meridional
Overturning Circulation lower limb is critically dependent on Atlantic-Arctic mixing, Nat. Commun., 15, 7341,
https://doi.org/10.1038/s41467-024-51777-w, 2024.

585 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the
Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9,
1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Ganachaud, A. and Wunsch, C.: Improved estimates of global ocean circulation, heat transport and mixing from
hydrographic data, Nature, 408, 453–457, https://doi.org/10.1038/35044048, 2000.

590 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, J. Geophys. Res. Oceans, 118, 6704–6716,
https://doi.org/10.1002/2013JC009067, 2013.

Gouretski, V. and Cheng, L.: Correction for Systematic Errors in the Global Dataset of Temperature Profiles from
Mechanical Bathythermographs, J. Atmos. Ocean. Technol., 37, 841–855, https://doi.org/10.1175/JTECH-D-19-0205.1,
595 2020.

Gouretski, V. and Reseghetti, F.: On depth and temperature biases in bathythermograph data: Development of a new
correction scheme based on analysis of a global ocean database, Deep-Sea Research Part I: Oceanogr. Res. Pap., 57, 812–
833, https://doi.org/10.1016/j.dsr.2010.03.011, 2010.

Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Böning, C. W., Chassignet, E. P., Curchitser,
600 E., Deshayes, J., Drange, H., Fox-Kemper, B., Gleckler, P. J., Gregory, J. M., Haak, H., Hallberg, R. W., Heimbach, P.,
Hewitt, H. T., Holland, D. M., Ilyina, T., Jungclaus, J. H., Komuro, Y., Krasting, J. P., Large, W. G., Marsland, S. J.,
Masina, S., McDougall, T. J., Nurser, A. J. G., Orr, J. C., Pirani, A., Qiao, F., Stouffer, R. J., Taylor, K. E., Treguier, A. M.,
Tsujino, H., Uotila, P., Valdivieso, M., Wang, Q., Winton, M., and Yeager, S. G.: OMIP contribution to CMIP6:
experimental and diagnostic protocol for the physical component of the Ocean Model Intercomparison Project, Geosci.
605 Model Dev., 9, 3231–3296, https://doi.org/10.5194/gmd-9-3231-2016, 2016.

Gruber, N., Clement, D., Carter, B. R., Feely, R. A., van Heuven, S., Hoppema, M., Ishii, M., Key, R. M., Kozyr, A.,
Lauvset, S. K., Lo Monaco, C., Mathis, J. T., Murata, A., Olsen, A., Perez, F. F., Sabine, C. L., Tanhua, T., and Wanninkhof,

R.: The oceanic sink for anthropogenic $CO2$ from 1994 to 2007, Science, 363, 1193–1199, https://doi.org/10.1126/science.aau5153, 2019.

610   Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., von Storch, J.-S., Brüggemann, N., Haak, H., and Stössel, A.: 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, 2019.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J.,
615   Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and von Storch, J.-S.: 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, 2016.

Hallberg, R.: Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects, Ocean Model., 72, 92–103, https://doi.org/10.1016/j.ocemod.2013.08.007, 2013.

620   Hewitt, H. T., Roberts, M. J., Hyder, P., Graham, T., Rae, J., Belcher, S. E., Bourdallé-Badie, R., Copsey, D., Coward, A., Guiavarch, C., Harris, C., Hill, R., Hirschi, J. J.-M., Madec, G., Mizielinski, M. S., Neininger, E., New, A. L., Rioual, J.-C., Sinha, B., Storkey, D., Shelly, A., Thorpe, L., and Wood, R. A.: The impact of resolving the Rossby radius at mid-latitudes in the ocean: results from a high-resolution version of the Met Office GC2 coupled model, Geosci. Model Dev., 9, 3655–3670, https://doi.org/10.5194/gmd-9-3655-2016, 2016.

625   Hewitt, H. T., Bell, M. J., Chassignet, E. P., Czaja, A., Ferreira, D., Griffies, S. M., Hyder, P., McClean, J. L., New, A. L., and Roberts, M. J.: Will high-resolution global ocean models benefit coupled predictions on short-range to climate timescales?, Ocean Model., 120, 120–136, https://doi.org/10.1016/j.ocemod.2017.11.002, 2017.

Hirschi, J. J.-M., Barnier, B., Böning, C., Biastoch, A., Blaker, A. T., Coward, A., Danilov, S., Drijfhout, S., Getzlaff, K., Griffies, S. M., Hasumi, H., Hewitt, H., Iovino, D., Kawasaki, T., Kiss, A. E., Koldunov, N., Marzocchi, A., Mecking, J. V.,
630   Moat, B., Molines, J.-M., Myers, P. G., Penduff, T., Roberts, M., Treguier, A.-M., Sein, D. V., Sidorenko, D., Small, J., Spence, P., Thompson, L., Weijer, W., and Xu, X.: The Atlantic Meridional Overturning Circulation in High-Resolution Models, J. Geophys. Res. Oceans, 125, e2019JC015522, https://doi.org/10.1029/2019JC015522, 2020.

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

635   Jackson, L. C. and Petit, T.: North Atlantic overturning and water mass transformation in CMIP6 models, Clim. Dyn., 60, 2871–2891, https://doi.org/10.1007/s00382-022-06448-1, 2023.

Jackson, L. C., Roberts, M. J., Hewitt, H. T., Iovino, D., Koenigk, T., Meccia, V. L., Roberts, C. D., Ruprich-Robert, Y., and Wood, R. A.: 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, 2020.

640   Johns, W. E., Elipot, S., Smeed, D. A., Moat, B., King, B., Volkov, D. L., and Smith, R. H.: Towards two decades of Atlantic Ocean mass and heat transports at 26.5° N, Philos. Trans. R. Soc. A, 381, 20220188, https://doi.org/10.1098/rsta.2022.0188, 2023.

Katsman, C. A., Drijfhout, S. S., Dijkstra, H. A., and Spall, M. A.: Sinking of Dense North Atlantic Waters in a Global Ocean Model: Location and Controls, J. Geophys. Res. Oceans, 123, 3563–3576, https://doi.org/10.1029/2017JC013329, 645   2018.

Khosravi, N., Wang, Q., Koldunov, N., Hinrichs, C., Semmler, T., Danilov, S., and Jung, T.: The Arctic Ocean in CMIP6 Models: Biases and Projected Changes in Temperature and Salinity, Earth's Future, 10, e2021EF002282, https://doi.org/10.1029/2021EF002282, 2022.

Koenigk, T., Fuentes-Franco, R., Meccia, V. L., Gutjahr, O., Jackson, L. C., New, A. L., Ortega, P., Roberts, C. D., 650   Roberts, M. J., Arsouze, T., Iovino, D., Moine, M.-P., and Sein, D. V.: Deep mixed ocean volume in the Labrador Sea in HighResMIP models, Clim. Dyn., 57, 1895–1918, https://doi.org/10.1007/s00382-021-05785-x, 2021.

Lee, R. W., Woollings, T. J., Hoskins, B. J., Williams, K. D., O'Reilly, C. H., and Masato, G.: Impact of Gulf Stream SST biases on the global atmospheric circulation, Clim. Dyn., 51, 3369–3387, https://doi.org/10.1007/s00382-018-4083-9, 2018.

655   Liu, W., Fedorov, A. V., Xie, S.-P., and Hu, S.: Climate impacts of a weakened Atlantic Meridional Overturning Circulation in a warming climate, Sci. Adv., 6, eaaz4876, https://doi.org/10.1126/sciadv.aaz4876, 2020.

Lumpkin, R. and Speer, K.: Global Ocean Meridional Overturning, J. Phys. Oceanogr., 37, 2550–2562, https://doi.org/10.1175/JPO3130.1, 2007.

Marzocchi, A., Hirschi, J. J.-M., Holliday, N. P., Cunningham, S. A., Blaker, A. T., and Coward, A. C.: The North 660   Atlantic subpolar circulation in an eddy-resolving global ocean model, J. Mar. Syst., 142, 126–143, https://doi.org/10.1016/j.jmarsys.2014.10.007, 2015.

Menary, M. B., Hodson, D. L. R., Robson, J. I., Sutton, R. T., Wood, R. A., and Hunt, J. A.: 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.

665   Menary, M. B., Jackson, L. C., and Lozier, M. S.: 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, 2020.

Moat, B. I., Smeed, D., Rayner, D., Johns, W. E., Smith, R. H., Volkov, D. L., Baringer, M. O., Collins, J.: Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS (RAPID-Meridional Overturning Circulation 670   and Heatflux Array-Western Boundary Time Series) array at 26N from 2004 to 2022 (v2022.1), NERC EDS British Oceanographic Data Centre NOC [data set], https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/04c79ece-3186-349a-e063-6c86abc0158c, 2023.

Moreno-Chamarro, E., Caron, L.-P., Loosveldt Tomas, S., Vegas-Regidor, J., Gutjahr, O., Moine, M.-P., Putrasahan, D., Roberts, C. D., Roberts, M. J., Senan, R., Terray, L., Tourigny, E., and Vidale, P. L.: Impact of increased resolution on long-standing biases in HighResMIP-PRIMAVERA climate models, Geosci. Model Dev., 15, 269–289, https://doi.org/10.5194/gmd-15-269-2022, 2022.

Moreno-Chamarro, E., Arsouze, T., Acosta, M., Bretonnière, P.-A., Castrillo, M., Ferrer, E., Frigola, A., Kuznetsova, D., Martin-Martinez, E., Ortega, P., and Palomas, S.: The very-high resolution configuration of the EC-Earth global model for HighResMIP, Geosci. Model Dev. Discuss. [preprint], https://doi.org/10.5194/gmd-2024-119, 2024.

Ng, H. C., Robinson, L. F., McManus, J. F., Mohamed, K. J., Jacobel, A. W., Ivanovic, R. F., Gregoire, L. J., and Chen, T.: Coherent deglacial changes in western Atlantic Ocean circulation, Nat. Commun., 9, 2947, https://doi.org/10.1038/s41467-018-05312-3, 2018.

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 Syst. Dyn., 12, 419–438, https://doi.org/10.5194/esd-12-419-2021, 2021.

Righi, M., Andela, B., Eyring, V., Lauer, A., Predoi, V., Schlund, M., Vegas-Regidor, J., Bock, L., Brötz, B., de Mora, L., Diblen, F., Dreyer, L., Drost, N., Earnshaw, P., Hassler, B., Koldunov, N., Little, B., Loosveldt Tomas, S., and Zimmermann, K.: Earth System Model Evaluation Tool (ESMValTool) v2.0 – technical overview, Geosci. Model Dev., 13, 1179–1199, https://doi.org/10.5194/gmd-13-1179-2020, 2020.

Roberts, M. J., Baker, A., Blockley, E. W., Calvert, D., Coward, A., Hewitt, H. T., Jackson, L. C., Kuhlbrodt, T., Mathiot, P., Roberts, C. D., Schiemann, R., Seddon, J., Vannière, B., and Vidale, P. L.: 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, 2019.

Roberts, M. J., Jackson, L. C., Roberts, C. D., 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., Putrasahan, D., Schiemann, R., Seddon, J., Terray, L., Xu, X., Zhang, Q., Chang, P., Yeager, S. G., Castruccio, F. S., Zhang, S., and Wu, L.: Sensitivity of the Atlantic Meridional Overturning Circulation to Model Resolution in CMIP6 HighResMIP Simulations and Implications for Future Changes, J. Adv. Model. Earth Syst., 12, e2019MS002014, https://doi.org/10.1029/2019MS002014, 2020.

Roquet, F., Madec, G., McDougall, T. J., and Barker, P. M.: Accurate polynomial expressions for the density and specific volume of seawater using the TEOS-10 standard, Ocean Model., 90, 29–43, https://doi.org/10.1016/j.ocemod.2015.04.002, 2015.

Sein, D. V., Koldunov, N. V., Danilov, S., Sidorenko, D., Wekerle, C., Cabos, W., Rackow, T., Scholz, P., Semmler, T., Wang, Q., and Jung, T.: 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, 2026–2041, https://doi.org/10.1029/2018MS001327, 2018.

Spall, M. A. and Pickart, R. S.: Where Does Dense Water Sink? A Subpolar Gyre Example, J. Phys. Oceanogr., 31, 810–826, https://doi.org/10.1175/1520-0485(2001)031<0810:WDDWSA>2.0.CO;2, 2001.

Straneo, F.: On the Connection between Dense Water Formation, Overturning, and Poleward Heat Transport in a
710   Convective Basin, J. Phys. Oceanogr., 36, 1822–1840, https://doi.org/10.1175/JPO2932.1, 2006.

Sun, B., Liu, C., and Wang, F.: Global meridional eddy heat transport inferred from Argo and altimetry observations, Sci. Rep., 9, 1345, https://doi.org/10.1038/s41598-018-38069-2, 2019.

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

715   Treguier, A. M., Deshayes, J., Lique, C., Dussin, R., and Molines, J. M.: Eddy contributions to the meridional transport of salt in the North Atlantic, J. Geophys. Res. Oceans, 117, C05010, https://doi.org/10.1029/2012JC007927, 2012.

Trenberth, K. E. and Fasullo, J. T.: Atlantic meridional heat transports computed from balancing Earth's energy locally, Geophys. Res. Lett., 44, 1919–1927, https://doi.org/10.1002/2016GL072475, 2017.

Tsartsali, E. E., Haarsma, R. J., Athanasiadis, P. J., Bellucci, A., de Vries, H., Drijfhout, S., de Vries, I. E., Putrasahan,
720   D., Roberts, M. J., Sanchez–Gomez, E., and Roberts, C. D.: Impact of resolution on the atmosphere–ocean coupling along the Gulf Stream in global high resolution models, Clim. Dyn., 58, 3317–3333, https://doi.org/10.1007/s00382-021-06098-9, 2022.

Winton, M., Anderson, W. G., Delworth, T. L., Griffies, S. M., Hurlin, W. J., and Rosati, A.: Has coarse ocean resolution biased   simulations   of   transient   climate   sensitivity?,   Geophys.   Res.   Lett.,   41,   8522–8529,
725   https://doi.org/10.1002/2014GL061523, 2014.

Yashayaev, I. and Loder, J. W.: Enhanced production of Labrador Sea Water in 2008, Geophys. Res. Lett., 36, L01606, https://doi.org/10.1029/2008GL036162, 2009.

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

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

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 Sci., 15, 779–808, https://doi.org/10.5194/os-15-779-2019, 2019.

735