JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2019MS001696

Evaluation of FESOM2.0 Coupled to ECHAM6.3: Preindustrial and HighResMIP Simulations

D. Sidorenko[1], H.F. Goessling[1], N.V. Koldunov[1,2], P. Scholz[1], S. Danilov[1,3,4], D. Barbi[1], W. Cabos[7], O. Gurses[1], S. Harig[1], C. Hinrichs[1], S. Juricke[3], G. Lohmann[1,2,6], M. Losch[1,2], L. Mu[1], T. Rackow[1], N. Rakowsky[1], D. Sein[1,5], T. Semmler[1], X. Shi[1], C. Stepanek[1], J. Streffing[1], Q. Wang[1], C. Wekerle[1], H. Yang[1], and T. Jung[1,6]

[1]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany, [2]MARUM—Center for Marine Environmental Sciences, University of Bremen, Bremen, Germany, [3]Department of Mathematics and Logistics, Jacobs University, Bremen, Germany, [4]A. M. Obukhov Institute of Atmospheric Physics, Russian Academy of Science, Moscow, Russia, [5]Shirshov Institute of Oceanology, Russian Academy of Science, Moscow, Russia, [6]Institute of Environmental Physics, University of Bremen, Bremen, Germany, [7]Department of Physics and Mathematics, University of Alcala, Alcala, Spain

Key Points:
- A new climate model setup using an unstructured-mesh ocean-sea ice component has been developed
- Mean state and long-term drift under preindustrial climate conditions are evaluated
- Modeled climates using coarse and eddy-resolving ocean configurations are compared

Correspondence to:
D. Sidorenko,
Dmitry.Sidorenko@awi.de

Citation:
Sidorenko, D., Goessling, H. F., Koldunov, N. V., Scholz, P., Danilov, S., Barbi, D., et al (2019). Evaluation of FESOM2.0 coupled to ECHAM6.3: Preindustrial and HighResMIP simulations. *Journal of Advances in Modeling Earth Systems*, *11*, 3794–3815. https://doi.org/10.1029/2019MS001696

Received 22 MAR 2019
Accepted 17 OCT 2019
Accepted article online 24 OCT 2019
Published online 25 NOV 2019

Abstract A new global climate model setup using FESOM2.0 for the sea ice-ocean component and ECHAM6.3 for the atmosphere and land surface has been developed. Replacing FESOM1.4 by FESOM2.0 promises a higher efficiency of the new climate setup compared to its predecessor. The new setup allows for long-term climate integrations using a locally eddy-resolving ocean. Here it is evaluated in terms of (1) the mean state and long-term drift under preindustrial climate conditions, (2) the fidelity in simulating the historical warming, and (3) differences between coarse and eddy-resolving ocean configurations. The results show that the realism of the new climate setup is overall within the range of existing models. In terms of oceanic temperatures, the historical warming signal is of smaller amplitude than the model drift in case of a relatively short spin-up. However, it is argued that the strategy of "de-drifting" climate runs after the short spin-up, proposed by the HighResMIP protocol, allows one to isolate the warming signal. Moreover, the eddy-permitting/resolving ocean setup shows notable improvements regarding the simulation of oceanic surface temperatures, in particular in the Southern Ocean.

1. Introduction

The fifth and earlier phases of the Coupled Model Intercomparison Project (CMIP) have led to a great number of scientific works, including the assessment reports of the Intergovernmental Panel on Climate Change, which have shed light on the drivers of the observed climate change. However, the spread of uncertainties between individual models indicates substantial shortcomings in simulating a realistic ocean state (see, e.g., Pithan et al., 2014; Sgubin et al., 2017). Common model biases in ocean models are attributed to insufficient spatial resolution (Delworth et al., 2012; Scaife et al., 2011; Sein et al., 2017; Sterl et al., 2012; Wekerle et al., 2017) and shortcomings of parameterizations that aim to represent unresolved processes. Increasing the resolution in traditional ocean models requires a large amount of computational resources that are not yet available. This explains the recent interest in climate modeling with components that operate on unstructured meshes and allow for variable resolution in key areas of the global ocean (Ringler et al., 2013; Skamarock et al., 2012). One of such models is the Finite Element Sea Ice-Ocean Model (FESOM1.4; Danilov et al., 2015; Wang et al., 2014). FESOM1.4 constitutes the ocean component of the climate system model AWI-CM (Rackow et al., 2016; Sidorenko et al., 2015) where it is coupled to the atmosphere model ECHAM6.3 (Stevens et al., 2013). AWI-CM is contributing to several CMIP6-endorsed model intercomparison projects (Eyring et al., 2016).

While FESOM1.4 has been used in various applications, faster numerical solutions operating on unstructured meshes have been sought for, leading to a new dynamical core for FESOM (Danilov et al., 2017; Scholz et al., 2019). FESOM version 2.0 provides up to 3 times speedup compared to its predecessor version 1.4, ensuring a throughput similar to that of regular-mesh models, while promising larger mesh flexibility and good scalability characteristics (Koldunov et al., 2019). The dynamical core of FESOM2.0 is based on

©2019. The Authors.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

finite volume discretization compared to the finite element discretization used for FESOM1.4; the abbreviation FESOM now reads as the Finite-volumE Sea ice–Ocean Model.

FESOM2.0 in standalone sea ice-ocean mode has been evaluated in a series of configurations with the conclusion that its physical performance is comparable to that of FESOM1.4. Given the numerical efficiency of FESOM2.0, a natural step has been to update the climate configuration (AWI-CM) accordingly. The goal of this paper is to evaluate the performance of FESOM2.0 in the climate setting when coupled to ECHAM6.3. Our intention is to analyze the new setup regarding (1) the climate mean state and long-term drift under pre-industrial (PI) climate conditions, (2) the fidelity in simulating the historical warming over the period 1950–2014, and (3) the influence of the oceanic resolution on the simulated climate. Since the atmospheric component is left unchanged, in the analysis below we focus on the performance in simulating sea ice and ocean characteristics.

The paper is organized as follows: Section 2 describes the climate model, section 3 describes the simulation setups, section 4 analyzes the long-term climate drift, section 5 analyzes the model fidelity in simulating the historical warming using the HighResMIP protocol, section 6 presents a discussion, and section 7 provides conclusions.

2. Model Description

The design of the coupled model follows the formulation described in Sidorenko et al. (2015) and Rackow et al. (2016). The atmosphere is simulated by ECHAM version 6.3.04p1 (Stevens et al., 2013). FESOM version 2.0 has been described in Scholz et al., 2019. The coupling between FESOM and ECHAM is achieved via the parallel OASIS3-MCT coupler (Valcke, 2013). The computation of momentum, heat, and freshwater fluxes is done by the atmosphere model, which computes 12 fluxes in total and passes them to the ocean. For all fluxes except momentum, we use distance-weighted interpolation between the atmosphere and ocean grids (DISTWGT option in OASIS3-MCT). In view of the fact that FESOM and ECHAM have substantially different representations of coastlines, geometry, and the flux formalisms, we did not apply the conservative remapping option available in OASIS3-MCT but enforce the conservation of heat and freshwater fluxes between model components. This is achieved via scaling of individual fluxes interpolated onto the ocean grid such that their net values are identical to those computed on the atmospheric grid (see, e.g., Sidorenko et al., 2015).

For momentum fluxes a bicubic interpolation is used (BICUBIC option in OASIS3-MCT), which is a more expensive option of the coupler (Rackow et al., 2019). However, using this option is necessary in order to obtain a smooth representation of the wind stress curl when it is computed on the ocean mesh (Valcke, 2013). The ocean is driven by the atmospheric fluxes and communicates its surface state to the atmosphere. In total, four ocean fields are sent to ECHAM: sea surface temperature (SST), sea ice thickness (SIT), sea ice concentration, and snow on sea ice.

In all experiments ECHAM was used with T63 spectral resolution (1.88° at the equator) and 47 vertical levels. The ECHAM time step is 450 s; depending on the ocean mesh the FESOM time step is 30 or 10 min (see sections 3.1 and 3.2), and the ocean-atmosphere coupling frequency is 3 hourly, which is sufficient to account for the diurnal cycle.

In the previous ECHAM-FESOM1.4 setup by Sidorenko et al. (2015) the Pacanowski-Philander (PP; Pacanowski & Philander, 1981) vertical mixing scheme was used and the same scheme was also adopted in MPI-ESM (see, e.g., Jungclaus et al., 2013). In the new FESOM2.0 as well as in the most recent version of FESOM 1.4, which contributes to CMIP6, the K-Profile Parameterization (KPP; Large et al., 1994) scheme is employed.

3. Experimental Setups

All experiments discussed in this paper are summarized in Table. 1.

3.1. PI

The first experiment is a PI control simulation following the Diagnostic, Evaluation and Characterization of Klima protocol, which is the fundamental part of the CMIP experimental design (Eyring et al., 2016). We

Table 1
Overview of Experiments Conducted With Low (LR) and High (HR) Meshes

Name	Oce ini. state	Ext. forcing	Years
PI control	spinup	1850	1,000
control-1950-LR	EN4	1950	170
hist-LR-1	control-1950-LR (year 50)	hist. (1950–2014)	64
hist-LR-2	control-1950-LR (year 100)	hist. (1950–2014)	64
control-1950-LR-1	control-1950-LR (year 50)	1950	64
control-1950-LR-2	control-1950-LR (year 100)	1950	64
control-1950-HR	EN4	1950	114
hist-HR	control-1950-HR (year 50)	hist. (1950–2014)	64

Note. The control experiment for preindustrial climate using LR (PI control) has been run for 1,000 years. HighResMIP control simulations control-1950-LR and control-1950-HR have been run for 170 and 114 years, respectively. From those the historical scenario and sensitivity runs have been started.

integrate FESOM2.0-ECHAM6.3 for 1,000 years with external forcing from the year 1850. In this experiment, FESOM is configured on a mesh with resolution varying from nominal 1° in the interior of the ocean to (1/3)° in the equatorial belt and 24 km north of 50°N (see Figure 1, left panel). The ocean surface is discretized with about 127,000 grid points, and 46 vertical levels are used. The ocean time step is set to 30 min. FESOM is initialized with the final state of a 60-year spin-up run under CORE-II atmospheric forcing with surface salinity restoring (Large & Yeager, 2009; Wang et al., 2014). We estimate the model bias and long-term drift by comparing several climate diagnostics to those from other models participating in CMIP.

3.2. HighResMIP

In the second set of experiments FESOM is used with two meshes differing in horizontal resolution (Figure 1). The first mesh (LR hereafter) is the one used for the PI simulation. The second mesh (HR hereafter) resolves the regions of high eddy activity with 10 km, which is better or close to two grid points per internal Rossby radius in most parts of the global ocean. The regions of high eddy activity were diagnosed from the variance of sea surface height

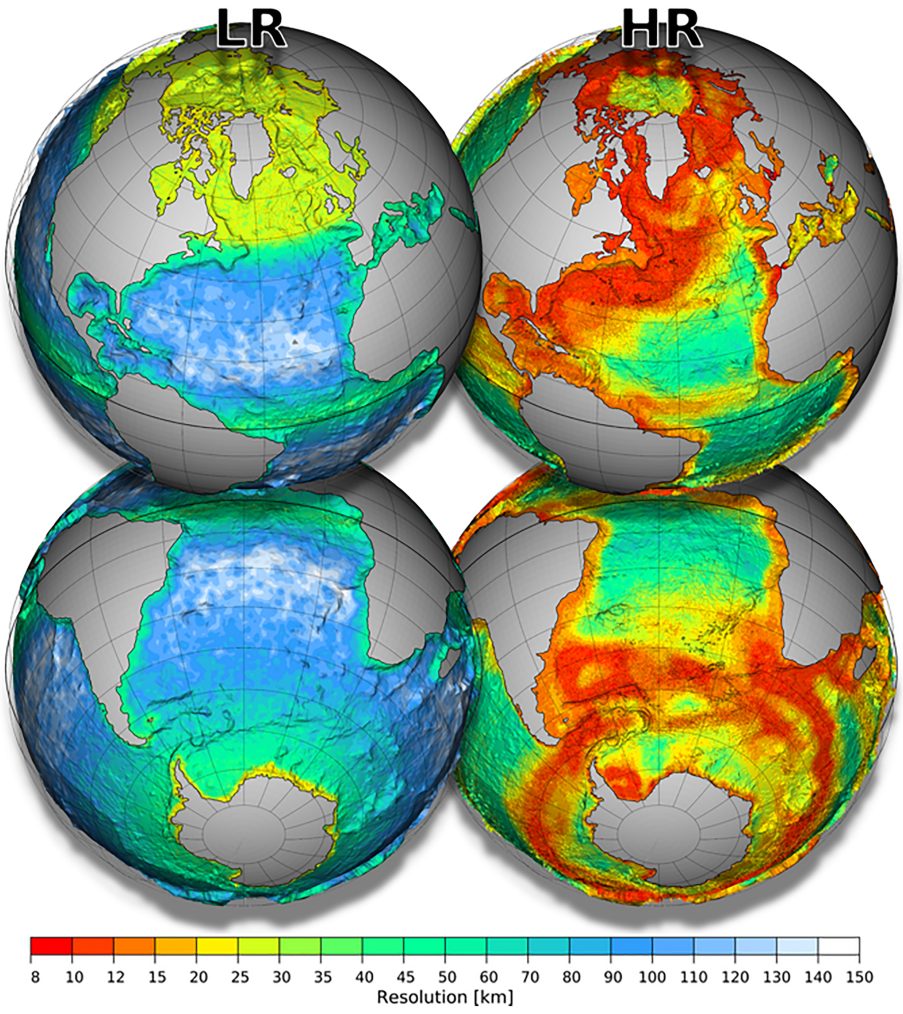

Figure 1. Resolution in the LR and HR ocean meshes. The number of surface vertices is 126,858 in LR and 1,306,775 in HR.

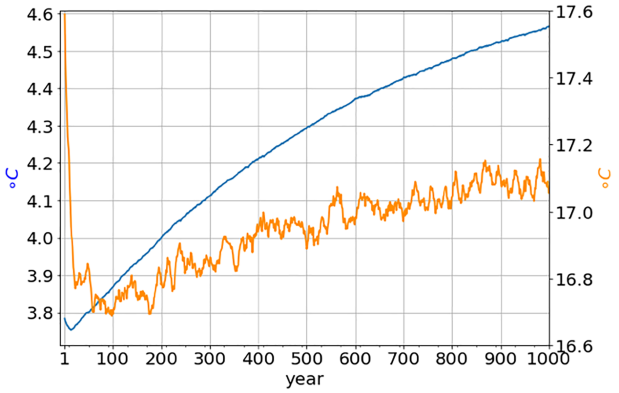

Figure 2. Time series of global potential temperature (blue line) and sea surface temperature (SST) (orange line, a 10-year moving average filter has been used) in PI control. The global volume-averaged potential temperature increases by ~0.8 °C over the 1,000 years. Consistently, the simulated SST drops by ~0.9 °C over the first 200 years and then increases gradually by ~0.35 °C until the end of the simulation.

as derived from satellite altimetry; additional data for sea ice extent and mixed layer depth were used to determine the mesh resolution. The detailed mesh design is described in Sein et al. (2016, 2017). The ocean surface in HR is discretized with about 1,300,000 grid points, and the same 46 vertical levels as in LR are used. To put this in context, a typical 0.25° global regular mesh includes about 900,000 wet grid points. The ocean time step is reduced to 10 min to maintain numerical stability for the HR mesh.

Following the HighResMIP protocol FESOM is initialized with the EN4 reanalysis data from the UK Met Office (Good et al., 2013). ECHAM, including the land surface scheme JSBACH, was initialized with the (quasi arbitrary) default initial state.

Using LR and HR meshes, we run several model simulations following the protocol of the High Resolution Model Intercomparison Project (HighResMIP; Haarsma et al., 2016). We first compute the control-1950 constant-forcing simulations using external forcing from 1950 (control-1950-LR and control-1950-HR). Then, starting from the year 50 of these simulations, we compute historical runs for the period of 1950–2014 (hist-LR-1 and hist-HR). Furthermore, using LR we conducted a second hististorical simulation (hist-LR-2) starting at year 100 of control-1950-LR.

For the hist-1950 simulations only FESOM was restarted from control-1950, while cold starts were used for ECHAM for technical reasons. This discontinuity was considered acceptable because the inertia of the ocean is much greater than that of the atmosphere. However, in order to quantify to which extent the incomplete restart of the climate model contaminates the climate signal, two additional control-1950 runs were conducted using the LR mesh. In these runs (with indices 1 and 2, hereafter) we restarted two additional control-1950 simulations after years 50 and 100 with cold starts for ECHAM (control-1950-LR-1 and -control-1950-LR-2).

With this set of experiments we explore the model fidelity in simulating the historical 1950–2014 climate and address the model sensitivity to initial conditions in the atmosphere and land surface components.

4. PI Control Simulation

In this section we evaluate the FESOM2.0 state and its drift in the 1,000 years of PI control.

4.1. Hydrography

The drift in ocean temperature is shown in Figure 2. The global volume-averaged potential temperature warms by ~0.8 °C over the 1,000 years. Consistently, the simulated SST drops by ~0.9 °C over the first 200 years and then increases gradually by ~0.35 °C until the end of the simulation. Although a slowdown of the increase can be seen toward the end of the simulation period, 1,000 years are clearly insufficient for reaching an equilibrium state.

Indeed, when initialized from an observed climatology of the ocean state, most climate models are characterized by an ocean warming drift even under PI climate conditions (e.g., Griffies et al., 2011; Lucarini & Ragone, 2011). A similar ocean warming was also observed in Sidorenko et al. (2015) with an older version of FESOM1.4 and using external forcing from the year 1990. In Sidorenko et al. (2015) (see their Figure 10) the excessive heat was stored primarily at a depth of ~1,000 m. Figure 3 (upper panel) shows the Hovmöller diagram for the global profile of oceanic potential temperature anomaly with respect to initial conditions. It illustrates that in the present simulations the extra heat also reaches the deeper ocean that exhibits an additional maximum at a depth of ~5,000 m.

The behavior of the new setup is more consistent with what has been observed in the GFDL model (see, e.g., Delworth et al., 2006, 2012). On the other hand the behavior of the previous FESOM1.4-ECHAM setup is similar to that of the MPI-ESM (see, e.g., Jungclaus et al., 2013). Interestingly the appearance of the deep maximum in heat storage correlates with the use of the turbulent closure for the vertical mixing scheme.

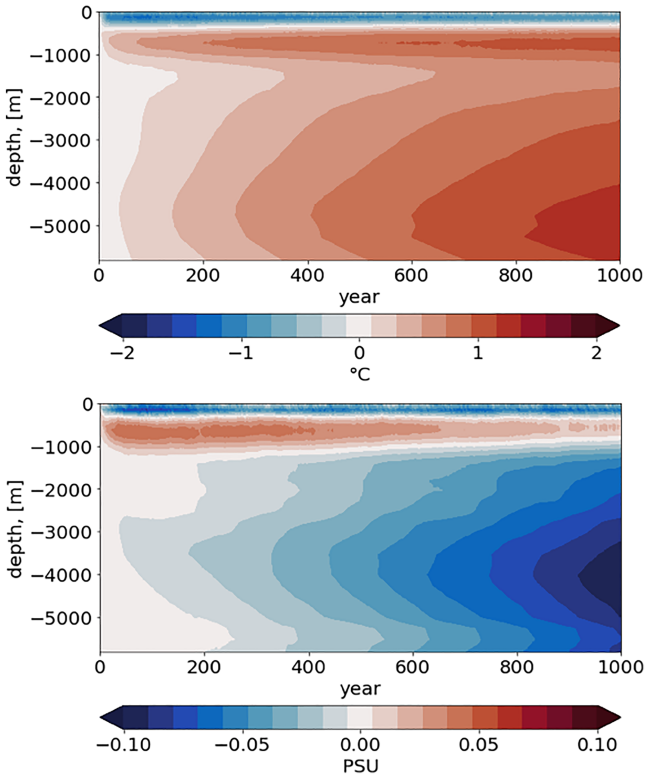

Figure 3. Hovmöller diagrams for global profiles of annual mean potential temperature (upper panel) and salinity (lower panel) anomalies with respect to initial conditions in PI control. In terms of buoyancy the temperature and salinity drifts partially compensate each other within the upper 1,000 m. The deeper ocean becomes more buoyant.

As already mentioned in section 2, in the older version ECHAM-FESOM1.4 the PP vertical mixing scheme was used and the same scheme was also adopted in MPI-ESM (see, e.g., Jungclaus et al., 2013). In the present version, the KPP scheme is employed, as in Delworth et al. (2012). We thus speculate that more intense vertical mixing through KPP brings more excessive heat to the deeper ocean. Note that using PP led to sporadic multidecadal collapses of the mixed layer depth in the Labrador Sea as has been described in Sidorenko et al. 2015 and also noticed in other models.

From the Hovmöller diagram for salinity anomalies (lower panel of Figure 3) we observe that in terms of buoyancy the decrease in salinity tends to compensate signals of cooling in the upper 200 m. At depths between 200 and 1,000 m the model shows a salinity increase during the first years of integration and a gradual decrease thereafter. At the end of the simulation a positive anomaly still remains and partially counteracts the buoyancy effect of increased temperature over the same depth range. Interestingly, the effects of both temperature and salinity changes lead to less dense water in deep layers where the ocean becomes fresher and warmer.

The spatial distribution of the model drift is depicted in Figure 4 as the departure of the modeled hydrography averaged over years 901–1,000 from the years 51–150. Based on the profiles of the subsurface drift in hydrography we focus on three depth levels: surface, 1,000 and 2,000 m.

SSTs increase on average by 0.36 °C. The warming is most pronounced over the Southern Ocean. Over 1,000 years this pattern of warming continues through the entire ocean depth. The largest positive heat anomalies are found in the deep North Atlantic at about 4,000-m depth (not shown) and reach ~2 °C. In contrast, some cooling is found at the subsurface of the Weddell Sea and Mediterranean Sea and in the middepths over the northern part of the Indian Ocean.

The complementary decrease in sea surface salinity (SSS) over most of the Atlantic Ocean, Western Indian Ocean, and in the Southern Ocean around the Ross Sea contributes to an overall increase of the surface buoyancy there. However, an overall decrease in salinity is observed in the deeper ocean, being the largest in the Weddell Sea. Positive salinity anomalies at depth are found over the Southern and Tropical Atlantic Ocean.

4.2. Sea Ice

The EVP solver for sea ice dynamics in FESOM2.0 has been updated according to Danilov et al. (2015) allowing for a smaller number of substeps. In standalone simulations it has been shown that this approach allows the simulation of linear kinematic features in sea ice (Wang et al., 2016a, 2016b). Using the same ocean mesh, these features were also present in FESOM2.0 during the ocean spinup before coupling. In the coupled PI simulation, however, these features were not formed anymore, likely due to the increased ice thickness. The disappearance of the linear kinematic features also suggests that the resolution of the mesh used here is still too coarse.

In Figure 5 we show the ice thickness patterns in September and March in both hemispheres. Given the rapid recent Arctic climate change, a direct comparison with observations can be misleading because satellite-derived ice thickness fields are limited to the recent past and to certain times of the year and come with relatively large uncertainties. However, a general picture of the observed Arctic ice thickness distribution can be found, for example, in Lindsay and Schweiger (2015).

The simulated ice thickness in the Arctic Ocean in March is above 3 m for most of the Central Arctic. The ice is more than 1 m thicker than that reported in Sidorenko et al. (2015) for 1990 conditions. Given the increased strength associated with the thicker sea ice, the momentum stress from the atmosphere is not sufficient to maintain a strong transpolar drift toward the north of Greenland. As a result the Arctic sea ice in

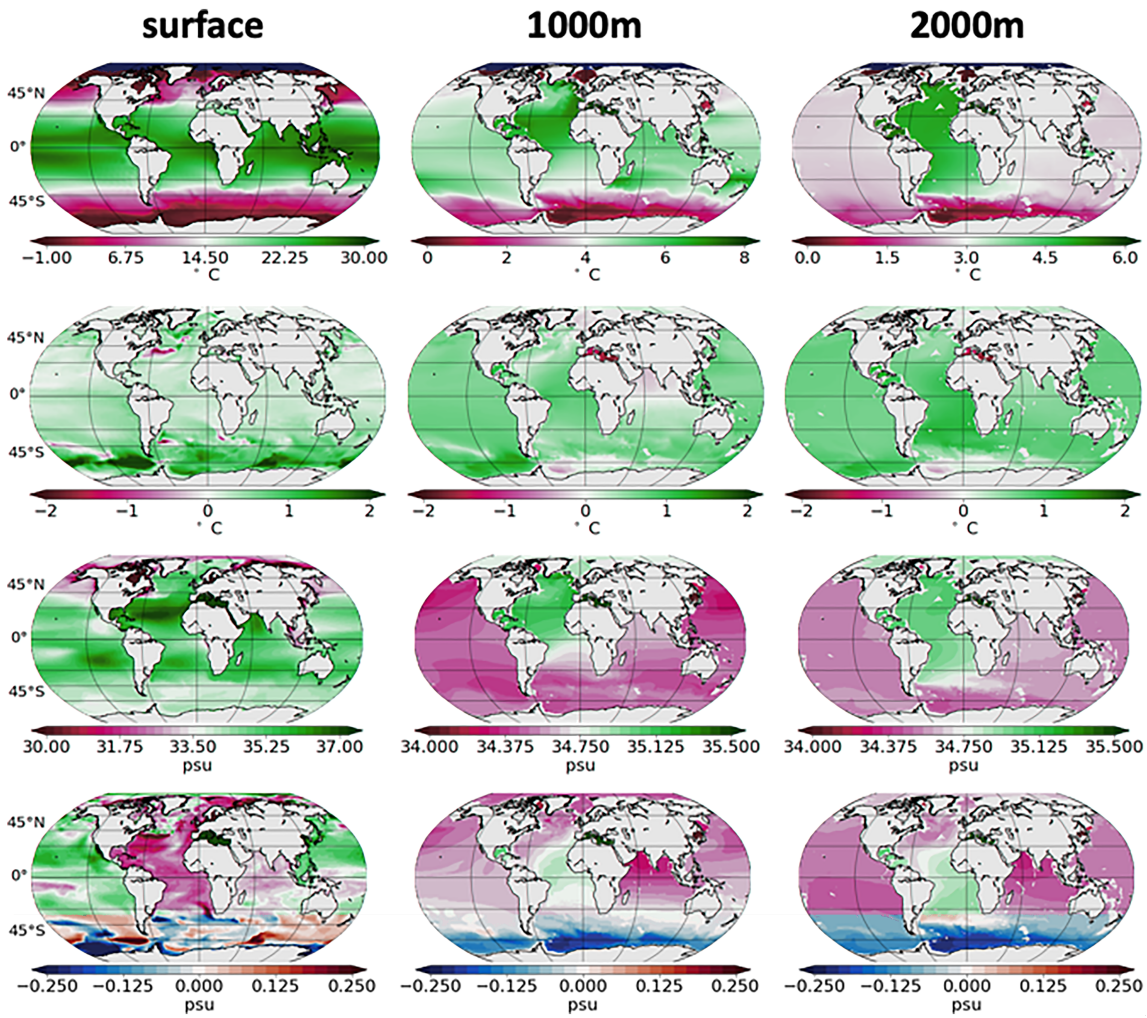

Figure 4. Potential temperature and salinity fields and drift in PI control. The first row shows the potential temperature averaged over the last 100 years at different depths. The second row shows the corresponding drift in potential temperature computed as the difference between years 901–1,000 and 51–150 (the initial adjustment within the first 50 years is excluded). The third and fourth rows show the same as the upper two but for salinity.

March tends to be more homogeneously distributed. Similar sea ice patterns in March were reported in Koldunov et al. (2010) and Notz et al. (2013) for the historical period 1979–2007 when using ECHAM5/MPIOM. We assume that the improper representation of the anticyclonic atmospheric circulation over the Beaufort Gyre (not shown) is primarily responsible for the unrealistic spatial distribution of ice thickness. Notz et al. (2013) reported on more realistic simulated sea ice after updating their atmospheric component to ECHAM6.

In September the Arctic sea ice is on average ~2 m thick in the Central Arctic. Being thinner than in March, it is pushed against Greenland, resulting in an increase of the SIT to about 3 m there.

In the Southern Hemisphere the sea ice volume remains low and is similar to previous versions of FESOM-ECHAM. Indeed, given the average SST warming of ~1 °C over the Southern Ocean (see Figure 4), the sea ice hardly reaches 0.25-m thickness in the central Weddell Sea in September. Concurrently, the sea ice is on the lower side in March.

The time series of integrated sea ice area is shown in Figure 6 for the whole simulation period. In the Northern Hemisphere the ice area equilibrates within the first 50 years, showing no significant drift during the rest of the simulation. In the Southern Hemisphere, however, a strong reduction of the September sea ice area is observed within the first 500 years, which, although at a lower rate, continues toward the end of the

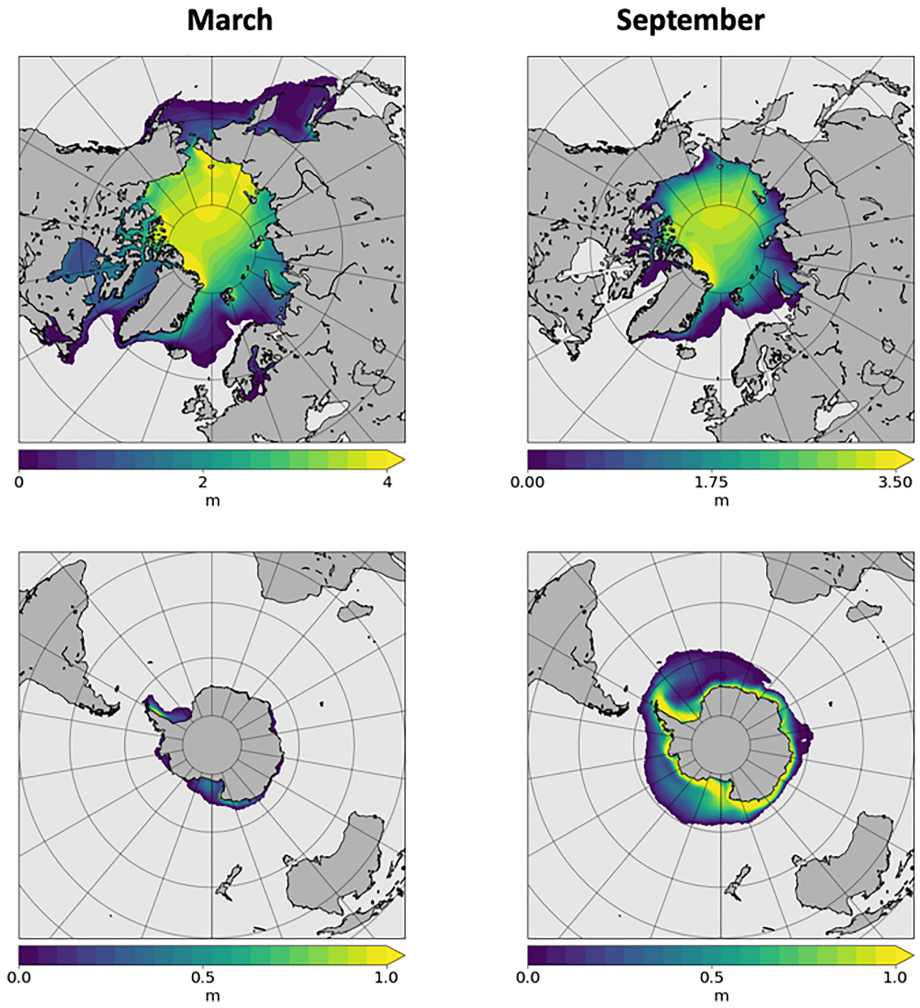

Figure 5. Sea ice thickness patterns in PI control (mean over the last 100 years).

simulation. A similar asymmetry between the simulated Antarctic and Arctic sea ice had been seen in a previous ECHAM6-FESOM1.4 (albeit present-day) control simulation (Rackow et al., 2019), which was attributed to the fact that the Southern Ocean forms the upwelling branch of the MOC (Marshall & Speer, 2012). Antarctic sea ice is thus stronger influenced by the long-timescale warming of the deep ocean than Arctic sea ice. This decay is also accompanied by the continuous surface warming of the Southern Ocean, which is an artifact inherent in many climate models contributing to CMIP.

4.3. Meridional Overturning Circulation

As the last diagnostic we inspect the meridional overturning circulation (MOC), which provides the most general characteristic of the water mass transformation and production. Figure 7 depicts the simulated global and Atlantic (AMOC). The former is expressed by the basinwide middepth cell of ~20 Sv at ~40°N. The bottom cell, induced by the circulation of the Antarctic Bottom water, is also well reproduced with a maximum of ~10 Sv. In the Northern Hemisphere the MOC is primarily formed by its contribution from the Atlantic Ocean as it can be stated from inspecting the pattern of AMOC. Even though large differences in temperature and salinity from the observed climatology are found, the simulated MOC and AMOC show the canonical pictures as known from other works. This indicates that although biases in the representation of water mass properties and ventilation mechanisms are present, they still result in a reasonable density distribution which maintains realistic transports.

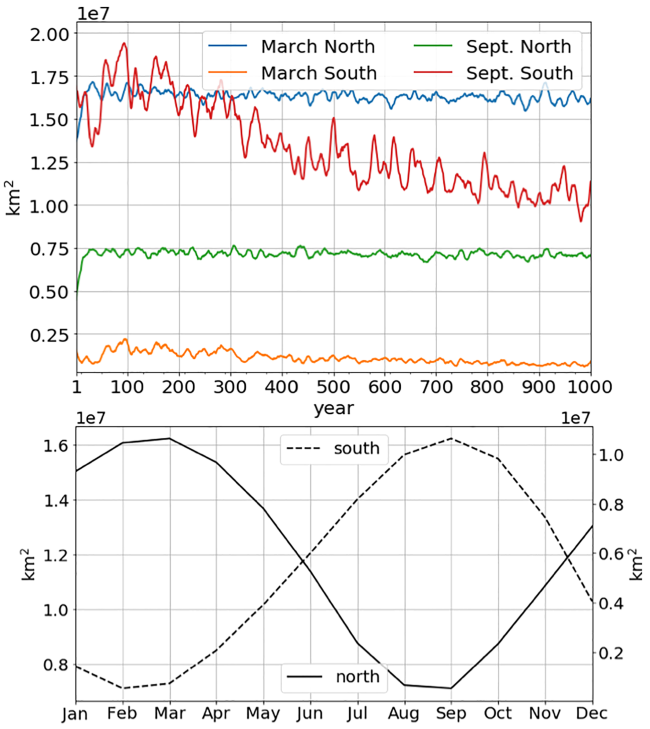

Figure 6. Time series of total sea ice area in PI control (top) and its mean seasonal cycle over the last 100 years seasonal cycle of sea ice area (bottom). The time series depicts the reduction with time of the southern sea ice in response to the pronounced warming signal there. The northern sea ice area shows nearly no drift.

The time series of Antarctic Circumpolar Current (ACC) as well as the AMOC at 26.5°N and 45°N are also provided in Figure 7. After 100 years of adjustment there is not much trends seen in above transports. Nonetheless, they depict strong interdecadal variability being in amplitude of ~3 Sv for AMOC and ~10 Sv for ACC.

We conclude that the simulated PI control climate is well within the spread of other climate models. This finding is not surprising considering that FESOM versions 1.4 and 2.0 show similar performance in standalone configurations (Scholz et al., 2019). From the observed drift over the whole period of the PI control we also conclude that the integration time of 1,000 years is not sufficient for reaching an equilibrated PI climate over the whole water column.

5. HighResMIP Simulations Using Different Resolutions in the Ocean

A locally eddy-resolving (permitting) FESOM2.0 configuration has been used along with the reference mesh to conduct runs according to the HighResMIP protocol (Haarsma et al., 2016). Between these setups we compare the simulated climate state, drift, and the historical warming.

5.1. Ocean Mean State and Drift

The temperature drift in the control runs is shown in Figure 8. Interestingly, the increase in the global volume-averaged ocean temperature is larger in the high-resolution setup with ~0.25 °C after 114 years compared to ~0.2 °C in the low-resolution setup. The total net heating over common integration period in LR and HR is 0.79 and 0.95 W/m^2, respectively. This agrees with the overall lower SST in HR versus LR which causes positive flux anomalies into the ocean. However, the averaged SST in HR gradually approaches that in LR during the last 100 years of simulation.

From the Hovmöller diagram for temperature and salinity drift (see Figure 9) it can be seen that the anomaly with excessive heat is less localized in-depth at ~800 m in HR compared to LR. In salinity, the formation of a freshwater anomaly is found in the deep ocean in LR but not in HR. In contrast, the SSS drift is larger in the HR setup. Despite the differences in terms of initialization and spin-up, the patterns of the drift in the HighResMIP simulations are similar to those in the PI control simulation reported in section 4.1.

The spatial distribution of the modeled temperature bias in LR is depicted in Figure 10 (upper panel) as the departure of the modeled temperature from the Polar Science Center Hydrographic Climatology (Steele et al., 2001). For consistent comparison we use the modeled hydrography from the historical simulations averaged over years 1994–2013 and show three depths of interest: surface, 1,000 m, and 2,000 m.

In LR, the SST is on average by ~1° cooler, with the largest differences found in the North Atlantic around Grand Banks and in the North Pacific Subtropical Gyre. Accompanying warming is found in the Southern Ocean, along the East Greenland Current and in the coastal upwelling zones. The latter is most pronounced in the South Atlantic. Interestingly, the observed SST difference from climatology resembles a canonical pattern shared by most of the climate models contributing to CMIP.

A strong warm bias at ~1,000 m is found in the Labrador Sea and in the Tropical Atlantic. From there it further propagates into the Pacific Ocean as we derived from observing the anomaly animation. A cold bias is found in the central North Atlantic, likely caused by too cold Mediterranean outflow, and in the Indian Ocean. At depths below ~2,000 m the ocean is persistently warmer than the Polar Science Center Hydrographic Climatology. The initial warming appears in the North Atlantic and then is communicated to other oceans (Rackow et al., 2019).

The salinity bias in LR is shown in Figure 11 (upper panel). It indicates that everywhere except for the North Atlantic it contributes to a reduction of buoyancy. In the North Atlantic, however, halosteric and

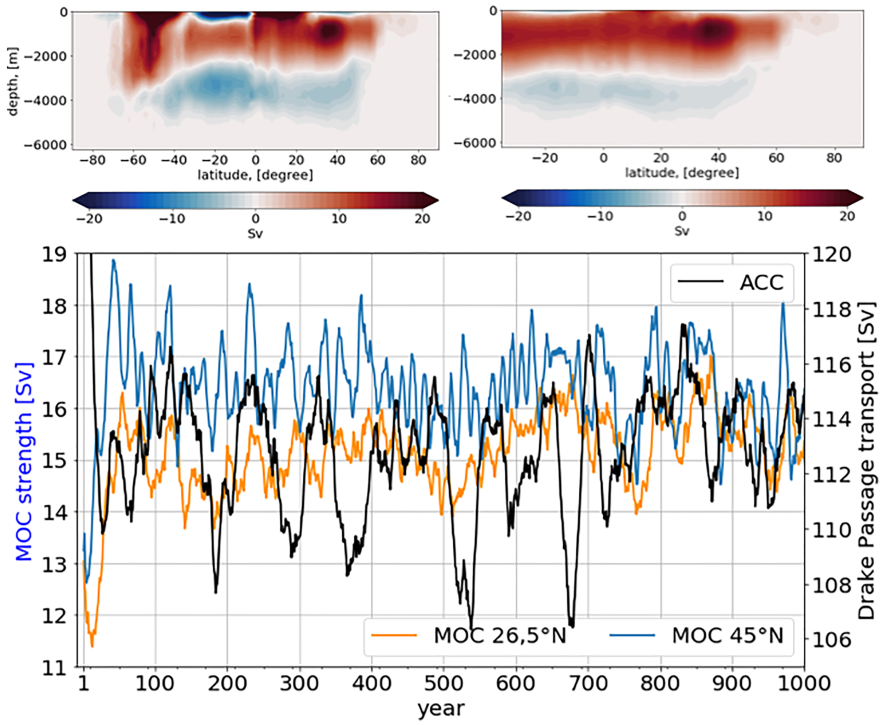

Figure 7. Top panel from left to right: Global meridional overturning circulation (MOC) and the Atlantic meridional overturning ocean (AMOC) streamfunctions in PI control including the eddy-induced transports (mean over the last 100 years). The streamfunctions depicts a canonical pattern as known from the literature with a maximum of ~20 Sv at ~45°N. Bottom panel shows the time series for the maximum MOC at 26,6°N, 45°N (orange and blue lines, respectively) and for the volume transport across Drake Passage (black curve). A 10-year moving average filter has been used.

thermosteric effects tend to compensate each other. SSS shows large positive differences in the vicinity of river mouths and in the Arctic Ocean. The latter bias was alleviated by a salt plume parameterization introduced in FESOM1.4 (Sidorenko et al., 2018). The salt plume parameterization scheme has not yet been tested in the new setup but will become available in future releases.

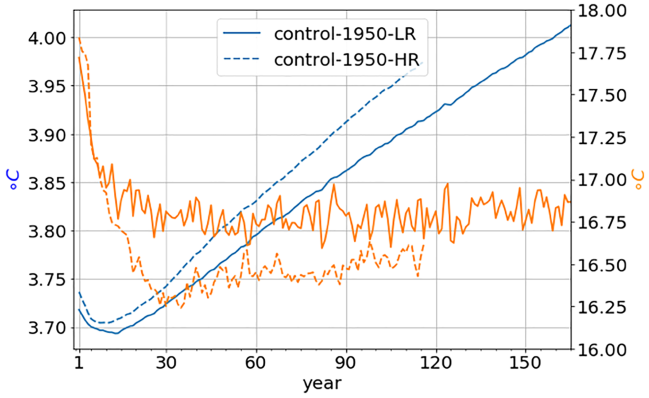

Figure 8. Time series of global volume-averaged potential temperature (blue) and sea surface temperature (yellow) in the control-1950-LR (solid lines) and HR (dashed lines) runs. The increase in the global mean temperature is larger in HR and is ~0.25 °C after 114 years compared to ~0.2 °C in LR. The total net heating over the first 114 years in LR and HR is 0.79 and 0.95 W/m^2, respectively. This agrees with the overall lower SST in the HR versus LR that causes stronger positive heat flux anomalies into the ocean.

Even though the RMS error (RMSE) of the SST bias is larger in HR (RMSE = 1.59) than in LR (RMSE = 1.43), which is primarily caused by the cooler temperatures in the Southern Pacific, there are several improvements in the simulated surface temperature in HR compared to LR. First, the cold bias around Newfoundland as well as over the area of the Gulf Stream is reduced by ~1.5 °C. This agrees with other studies (e.g., Scaife et al., 2011) showing that an eddy-permitting resolution in the North Atlantic is required for a realistic representation of the path of the North Atlantic Current (NAC). Second, the biases associated with coastal upwelling are smaller in HR compared to LR. This is most notably seen in the South Atlantic where the positive SST bias associated with a shift of the Malvinas Current has been remarkably reduced. Finally, the warm SST bias in the Southern Ocean associated with decreased sea ice area in the Southern Hemisphere is largely absent in the HR setup. The latter is consistent with the finding by Rackow et al. (2019). As shown below, these differences are also reflected in the patterns of the atmospheric circulation.

At depth, the temperature biases also differ between the setups, with HR being warmer than LR along the northern edge of the ACC and in the central Indian and Pacific Oceans. Based on the drift along the southern edge

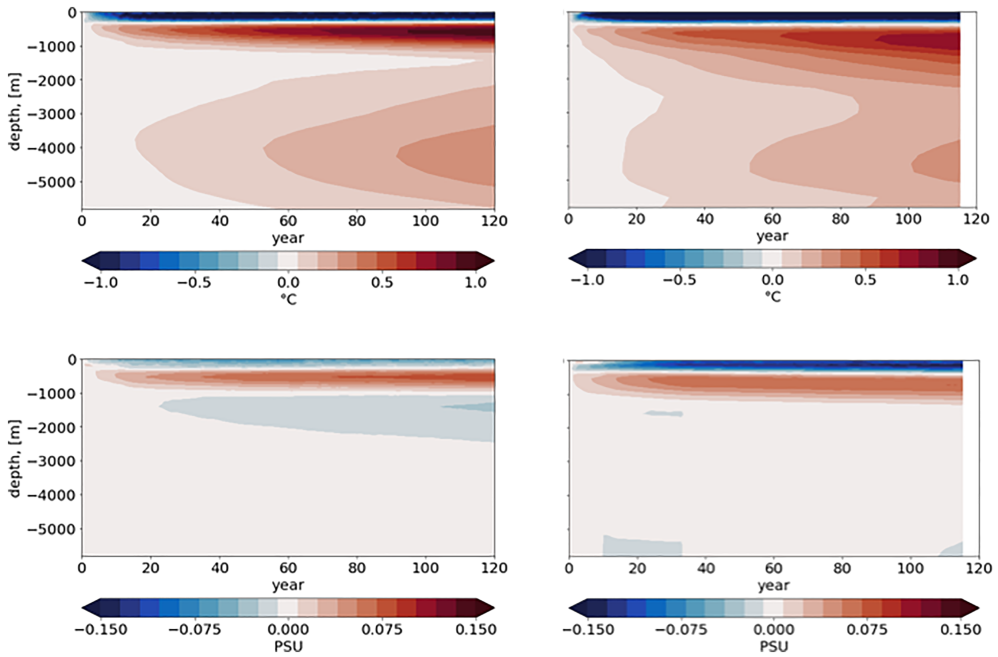

Figure 9. Hovmöller diagrams for global profiles of potential temperature (top) and salinity (bottom) anomalies with respect to initial conditions in control-1950-LR (left) and control-1950-HR (right). The positive heat anomaly is less localized in depth at ~800 m in HR compared to LR.

of the ACC we speculate that the resolution in HR is still insufficient to simulate realistic eddy dynamics in the SO. Indeed, the Gent-McWilliams (Gent et al., 1995; Gent & McWilliams, 1990) parameterization, although scaled with grid size (see Scholz et al., 2019, for the implementation of Gent-McWilliams and

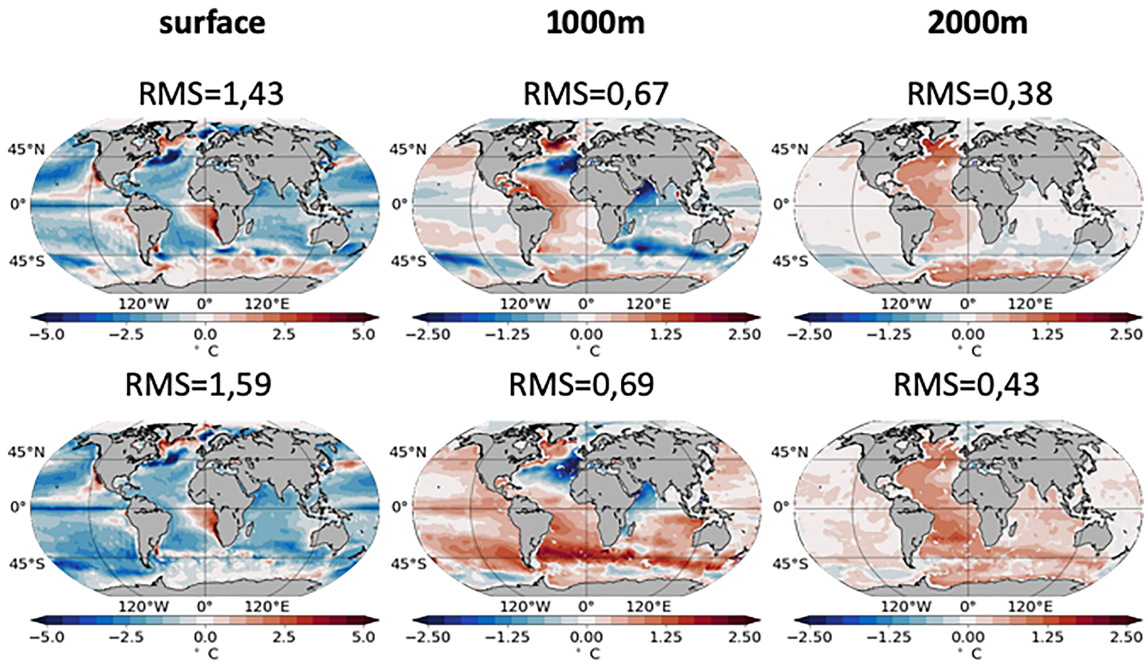

Figure 10. Temperature differences from WOA Climatology 2013 (mean over years 1994–2013) in hist-LR-1 (top) and hist-HR-1 (bottom) shown from left to right for depths at 0, 1,000, and 2,000 m. The SST pattern is improved in the Southern Ocean in HR. The largest differences at depth are also found there along the ACC front.

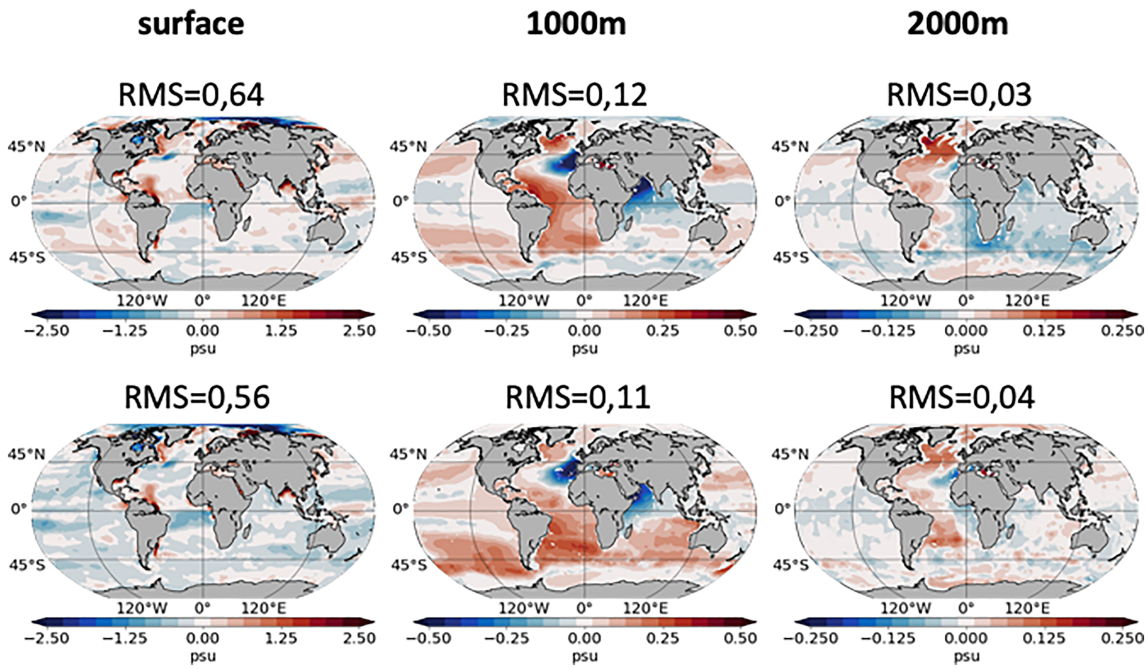

Figure 11. Same as Figure 10 but for salinity.

Redi diffusivity), was still active in HR and could have damped the simulated eddy variability. Due to the relatively high cost of HR we did not perform additional sensitivity studies but kept the model configuration files identical for both setups. Generally, in both setups the positive difference in temperature is accompanied by an increase in the salinity over the same area reducing the effect on density.

We note that the bias in modeled hydrography in both LR and HR is comparable or even smaller than that in the long-term climate drift over 1,000 years discussed in section 4.1. Furthermore, this bias is similar in amplitude to the standalone ocean configuration discussed in Scholz et al. (2019). This means that the largest part of the ocean differences from climatology presented here are accumulated during the ocean spinup and during the initial phase of the coupled simulation.

5.2. Sea Ice

The resulting sea ice patterns are shown in Figures 12 and 13 for the Northern and Southern Hemispheres, respectively. In the Northern Hemisphere the SIT is systematically lower in HR in March as well as in September. It is further confirmed by Figure 14 showing the seasonal cycle of the integrated sea ice area and is consistent with an intensification of the NAC and the AMOC (not shown). In September there is no accumulation of sea ice north of Greenland in both HR and LR. Furthermore, for both months HR agrees more closely with the older version of ECHAM5/MPIOM where significant positive biases in the Northern Hemisphere surface pressure patterns were thought to be the main source of the sea ice deficiencies (Koldunov et al., 2010; Notz et al., 2013). In the Southern Hemisphere, the HR setup simulates thicker and more cover sea ice, which is in line with the reduced warming of the SST in the Southern Ocean.

5.3. Atmosphere

From the last 70 years of the control runs we computed the difference in zonal 10-m winds and in Sea Level Pressure which are shown in Figure 15 The comparison between the LR and HR runs indicates that the increase of ocean resolution causes a poleward shift of the jet stream in the South. This is in line with the improvement in the simulated ocean surface hydrography as an SST bias is reduced in HR compared to LR. The interplay between the atmosphere and the ocean results in overall cooler surface temperatures in HR and more simulated sea ice.

In contrast, in the North the HR setup exhibits weaker SLP gradients. Reduced SLP in the Arctic matches the increase of the oceanic SST and the reduction of the sea ice there. Positive (negative) SLP differences

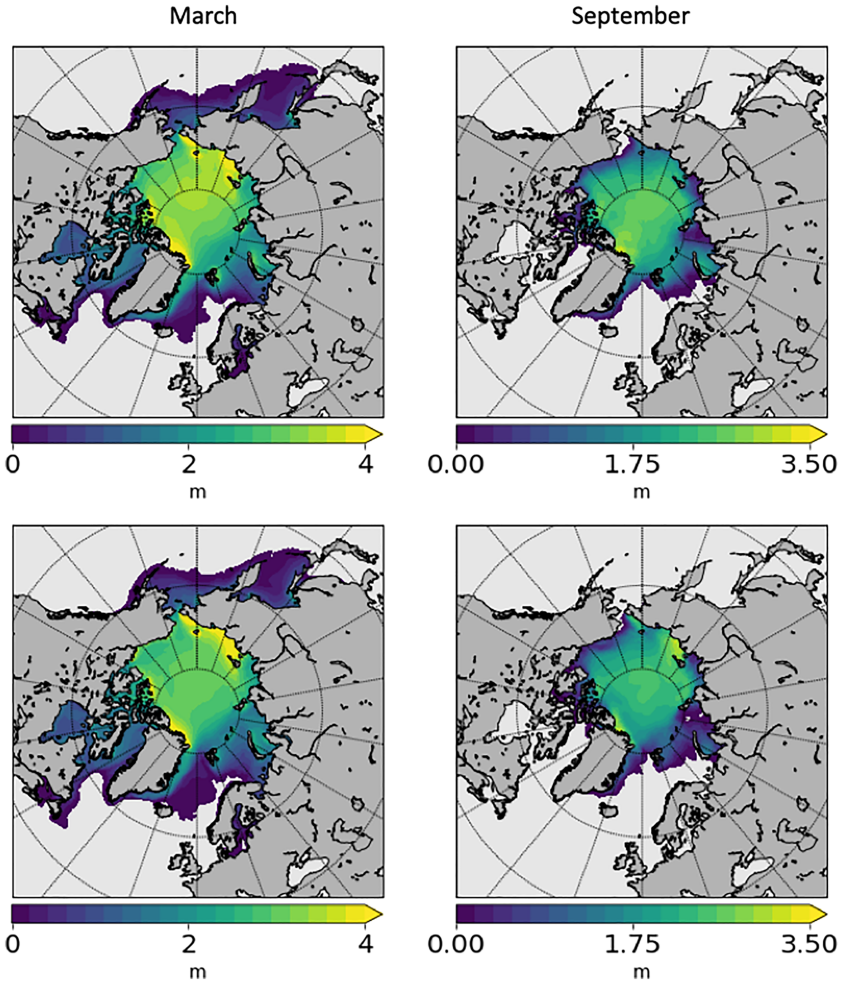

Figure 12. Northern sea ice thickness in March (left) and September (right) in control-1950-LR-1 (top) and control-1950-HR (bottom). Years 95–114 were used for averaging the control simulations. HR simulates thinner sea ice compared to LR, which is in line with the increased MOC.

coincide with positive (negative) differences in 500-hPa geopotential height (not shown) everywhere except in the Arctic. There the increase of oceanic SST and the reduction of the sea ice leads to negative SLP anomalies (Figure 14) and positive 500-hPa geopotential height anomalies (not shown), which is a typical signature of an anomalous heat low.

We quantify the atmospheric performance by computing performance indices as introduced by Reichler and Kim (2008) and further modified by Sidorenko et al. (2015) for the mean climate state and by Rackow et al. (2016, 2019) for the interannual variability. The performance indices grade climate model simulations of various atmospheric and oceanic parameters. Index of 1 indicates same performance as CMIP5 ensemble, while an index of less than 1 indicates better performance. The performance indices are shown in Table 2 for the five regions which are defined by the following latitudinal belts: the Arctic (60–90°N), the Northern Hemisphere midlatitudes (30–60°N), the Tropics (30°N–30°S), the Southern Hemisphere midlatitudes (30–60°S), and the Antarctic (60–90°S). Improvements in HR versus LR linked to the reduction of the surface temperature bias around Newfoundland are found in the North Atlantic and in the Arctic Ocean. Despite the fact that the simulated mean SST is more realistic in HR, the performance over the tropics and the Southern Hemisphere is slightly worse regarding the mean but slightly better regarding the interannual variability. We conclude that the atmospheric state has not been significantly altered by the new ocean component and is within the spread of other climate models.

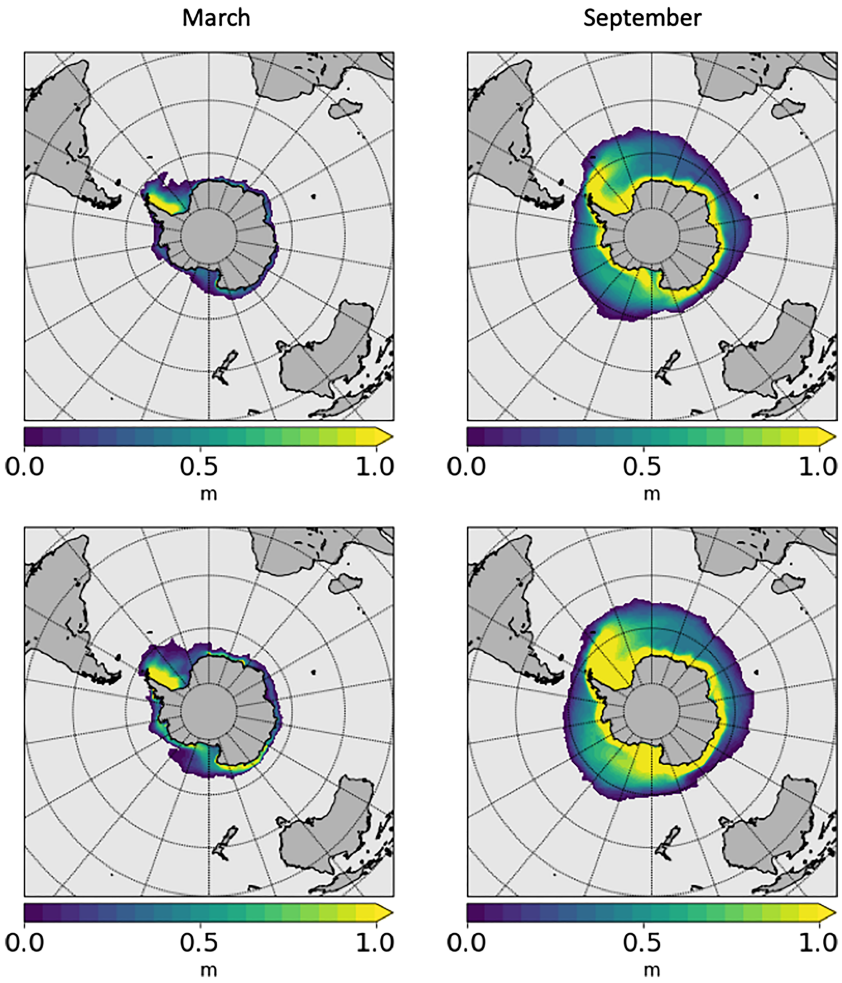

Figure 13. The same as Figure 12 but for the Southern Hemisphere. HR simulates thicker sea ice, which is in line with the reduced warming of the SSTs in the Southern Ocean.

5.4. Historical Warming

We compute the warming signal as the difference between the historical simulation and the control run, both averaged over the last 20 years 1994–2014. Considering that control-1950 is too short for reaching the equilibrated ocean state one may speculate that similar drift is also present in the historical simulation. It is expected that subtracting one run from another would result in the "de-drifted" warming signal.

The respective differences for atmospheric 2-m temperature and SST are shown in Figure 16. In general all runs show a similar warming pattern. The globally averaged atmospheric 2-m temperature increase varies from 0.71 to 0.83 °C between the simulations and compares to the HadCRUT4 (Morice et al., 2012) data of 0.61 °C. The accompanying increase in ocean SST varies from only 0.33 to 0.39 °C because the amplified warming of land areas and of the Arctic is not reflected in SST. It compares to the HadSST3 (Kennedy et al., 2011a, 2011b) data of 0.40 °C.

The largest differences between HR and LR are seen in SST and are found in the South Atlantic upwelling system, around Gulf Stream, along the path of the NAC and in the Pacific Ocean. It is striking that the temperature change patterns are more coherent between the second realization of LR (LR-2) and HR than between the two realizations of LR in many places. This indicates that the largest part of the differences is due to the natural variability of the climate system and that more realizations would be required to identify the robust impacts of ocean resolution on the warming pattern. Note that the eddy dynamics in HR SST is still visible in the difference pattern despite the 20 years average.

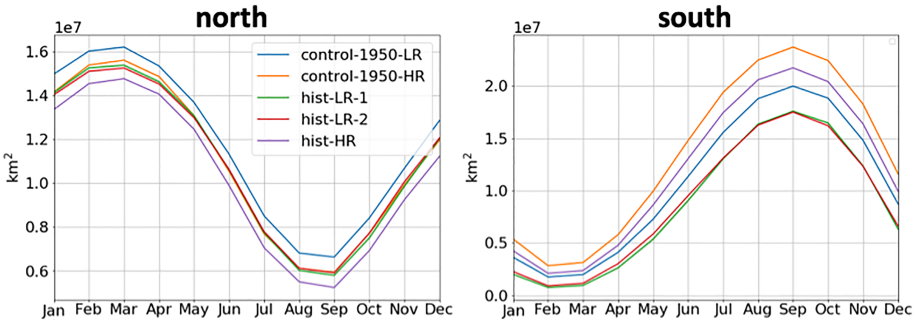

Figure 14. Averaged over years 1994–2014 mean seasonal cycle of the total sea ice area in historical and control runs. In the north all historical runs show a reduction of the sea ice area by ~7%, which depicts a uniform seasonal distribution. In the south the sea ice area is reduced by ~30% showing the largest decrease in September.

The reduction of SIT is shown in Figure 14 for the seasonal cycle and in Figures 17 and 18 for the patterns of spatial distribution. Independently on ocean resolution all historical runs show a reduction of the sea ice area by ~7%, which depicts a uniform seasonal distribution. In the south the sea ice area is reduced by ~30% showing the largest decrease in September. In the Northern Hemisphere the patterns of decrease look rather different between the LR and HR simulations. In LR the decrease varies from ~20 cm in the interior of the Arctic Ocean to about 1 m to the north and along the eastern part of Greenland. In contrast, in HR the maximum of the sea ice decrease is rather confined to the Siberian shelf. Such a discrepancy is probably linked to the difference in representing the mean sea ice distribution described in section 4.2. On the other hand, one may expect that more realizations are required for obtaining a consistent picture of the sea ice changes.

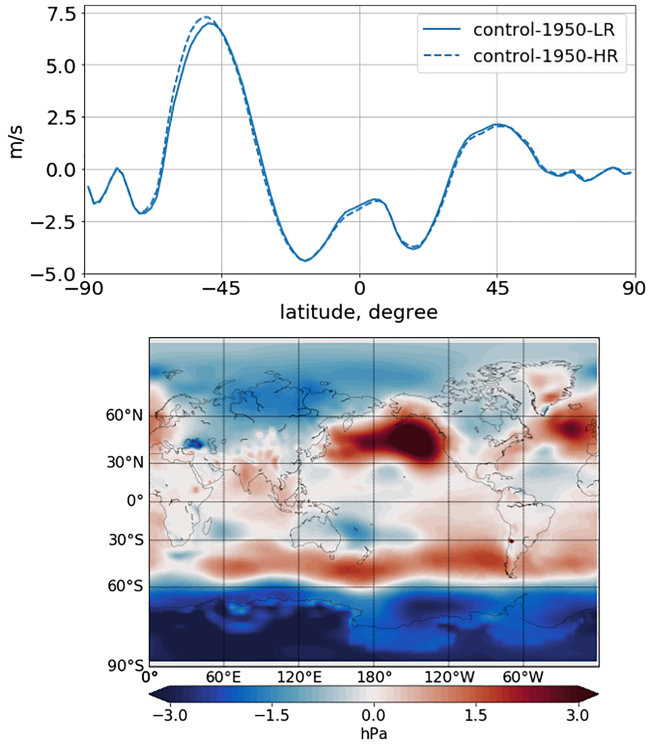

Figure 15. Zonal-mean zonal 10-m wind in control-1950-HR and control-1950-LR (top) and the difference in SLP (bottom). Years 51–114 have been used for averaging. The increase of the ocean resolution causes a poleward shift of the jet stream in the south. The interplay with the ocean makes the surface temperatures colder there, and more sea ice is simulated. In the north there is a weakening of the SLP gradients in HR along with the increase of the oceanic SSTs and the reduction of the sea ice.

Table 2
Mean [Interannual Standard Deviation] Modified Performance Index for Five Regions in the Historical Climate Run With FESOM2.0–ECHAM6.3

Model run	60–90°S	30–60°S	30°N–30°S	30–60°N	60–90°N
hist-LR-1	0.78 [1.08]	0.90 [1.21]	0.90 [1.30]	0.99 [1.08]	1.06 [1.03]
hist-LR-2	0.79 [1.03]	0.90 [1.16]	0.93 [1.16]	1.04 [1.10]	1.07 [1.03]
hist-HR	0.85 [0.99]	0.99 [1.09]	0.97 [1.11]	0.96 [1.01]	1.02 [0.98]

Note. Indices below (above) 1 indicate that a model performs better (worse) than the average of the CMIP5 models MPI-ESM-LR, HadGEM2, CCSM4, GFDL-CM3, and MIROC-ESM.

5.5. Inertia From Atmosphere and Land Hydrology

It is well known that for long timescales the inertia of the ocean is larger than that of the atmosphere and land. Therefore, when initializing a historical simulation, it may be acceptable to restart only the ocean component and perform a cold start (a start from arbitrary initial conditions) for the atmosphere and land. In the following we investigate the impact of such a partial cold start, which we performed for technical reasons.

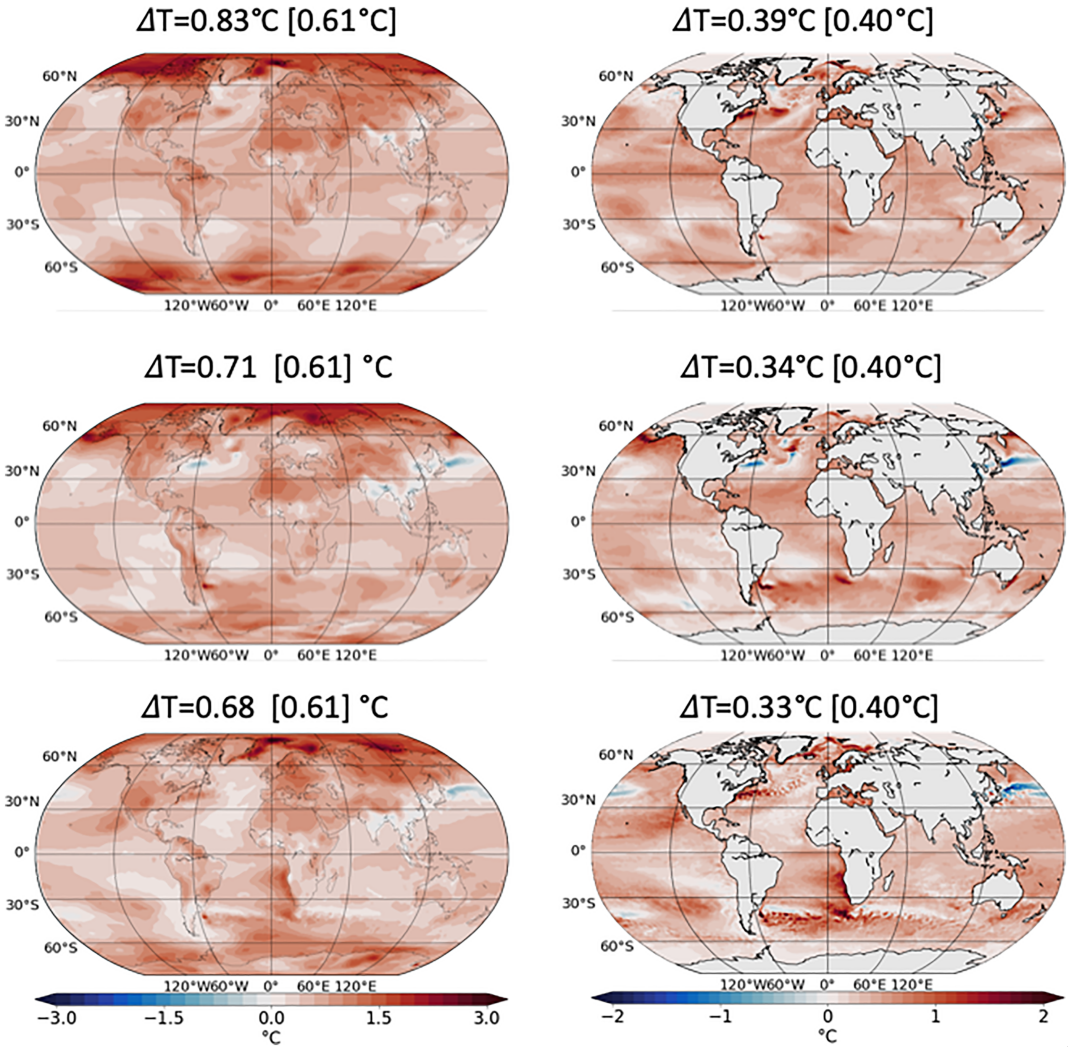

Figure 16. Warming signal in atmospheric 2-m temperature (left) and ocean SST (right) during years 1994–2014 computed as a difference between hist-LR-1 minus control-1950-LR-1 (top row), hist-LR-2 minus control-1950-LR-2 (middle row), and hist-HR minus control-1950-HR (bottom row). The values indicate the average warming signal that is compared to the data (shown in brackets) as derived from the HadCRUT4 (Morice et al., 2012) and HadSST3 (Kennedy et al., 2011a, 2011b) data sets. The globally averaged atmospheric 2-m temperature varies from 0.71 and 0.83 °C (0.61 °C, HadCRUT4) between the simulations, while the averaged ocean SST varies from 0.33 to 0.39 °C (0.4 °C, HadSST3).

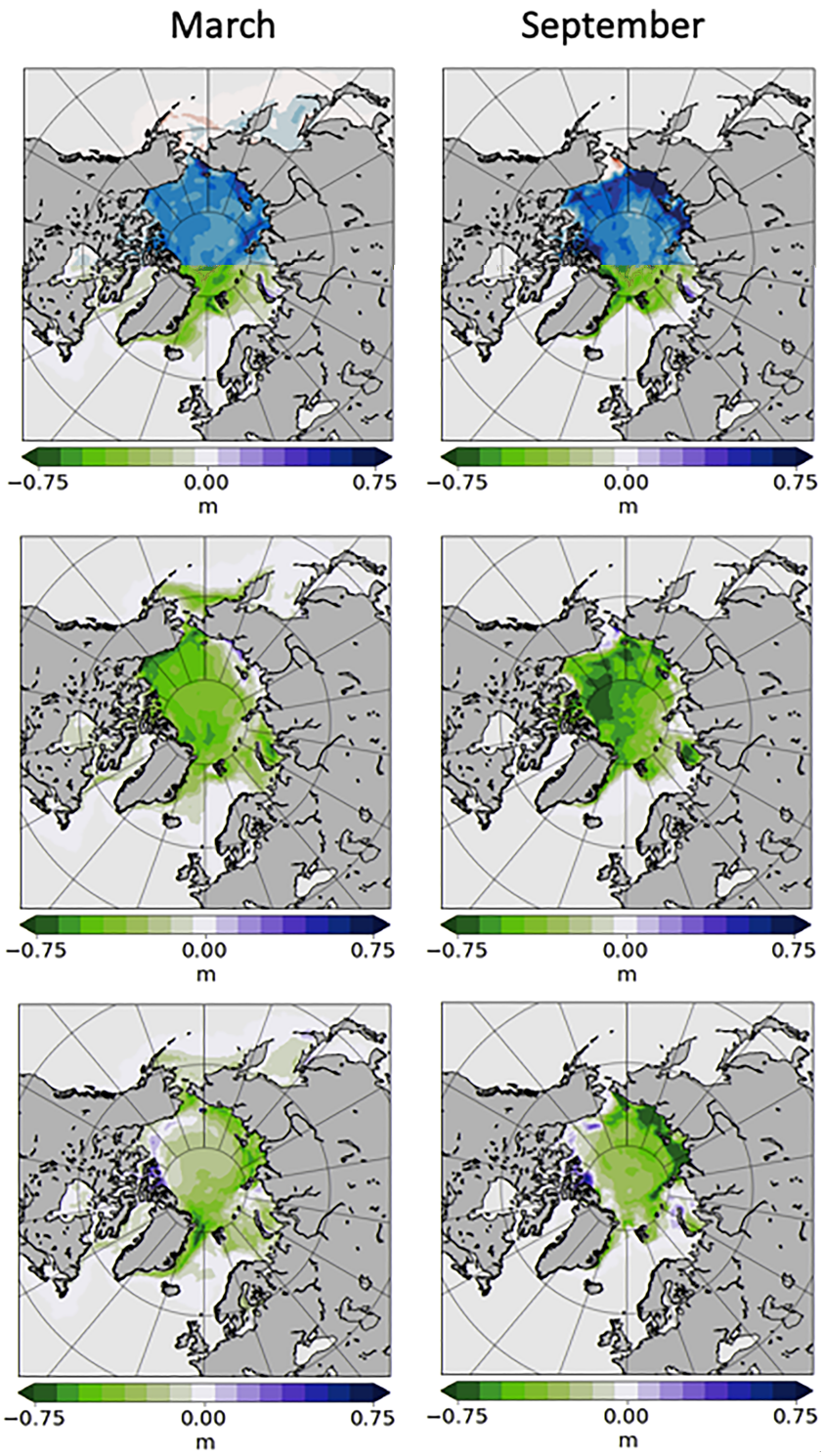

Figure 17. Difference in northern sea ice between hist-LR-1 and control-1950-LR-1 (top), hist-LR-2 and control-1950-LR-2 (middle), and hist-HR and control-1950-HR (bottom). Years 1994–2014 were used for averaging the historical simulations. The patterns of decrease look rather different between the LR and HR simulations. More realizations are required for obtaining a robust picture of the signal.

The time-depth departure of the simulated global temperature profile from that in a continuous control run is shown for all three realizations in Figure 19 (left column). The amplitude of the forced climate signal, although persistent, is smaller than the model drift which is depicted in Figures 3 and 9. Within the first

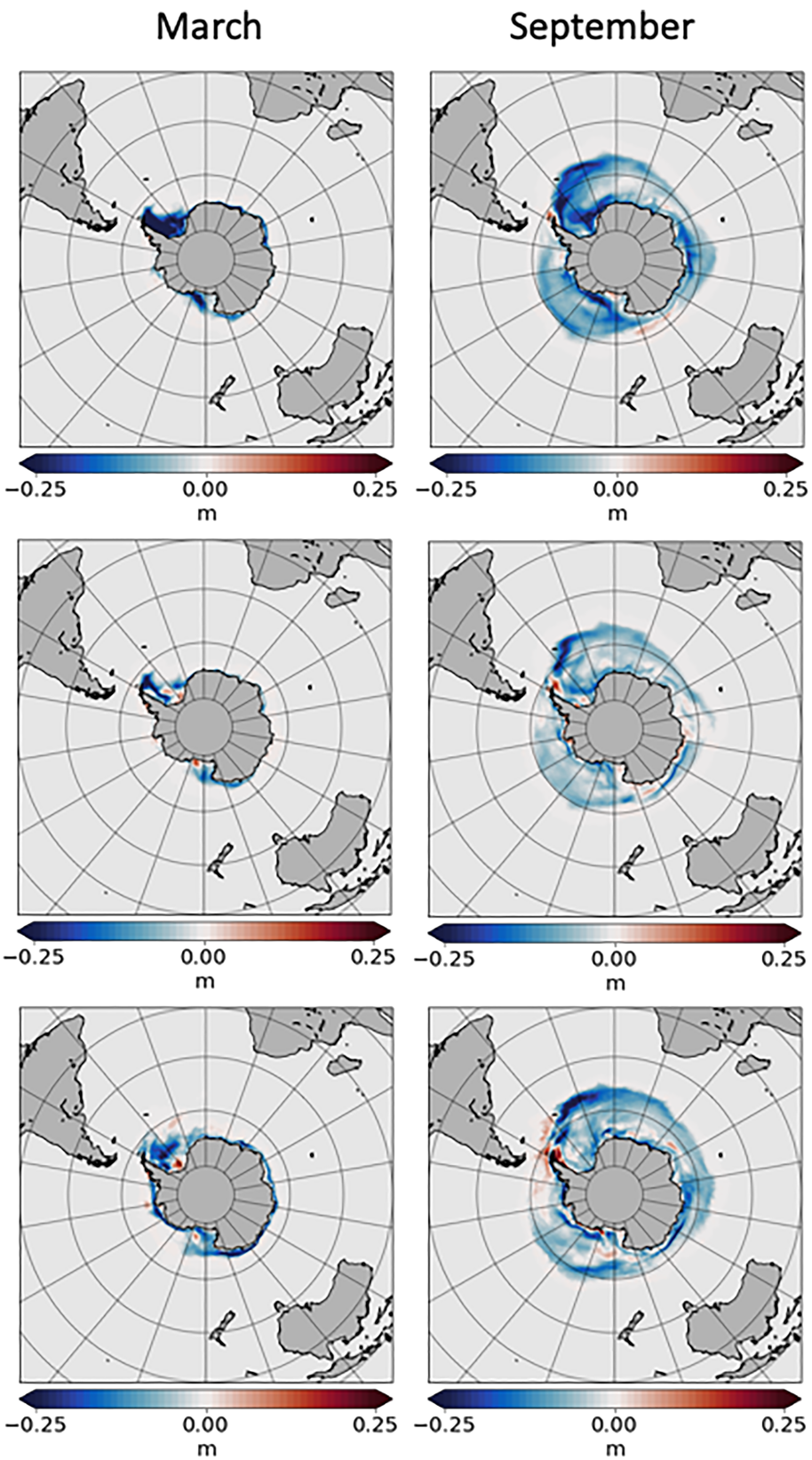

Figure 18. The same as figure 16 but for the Southern Hemisphere.

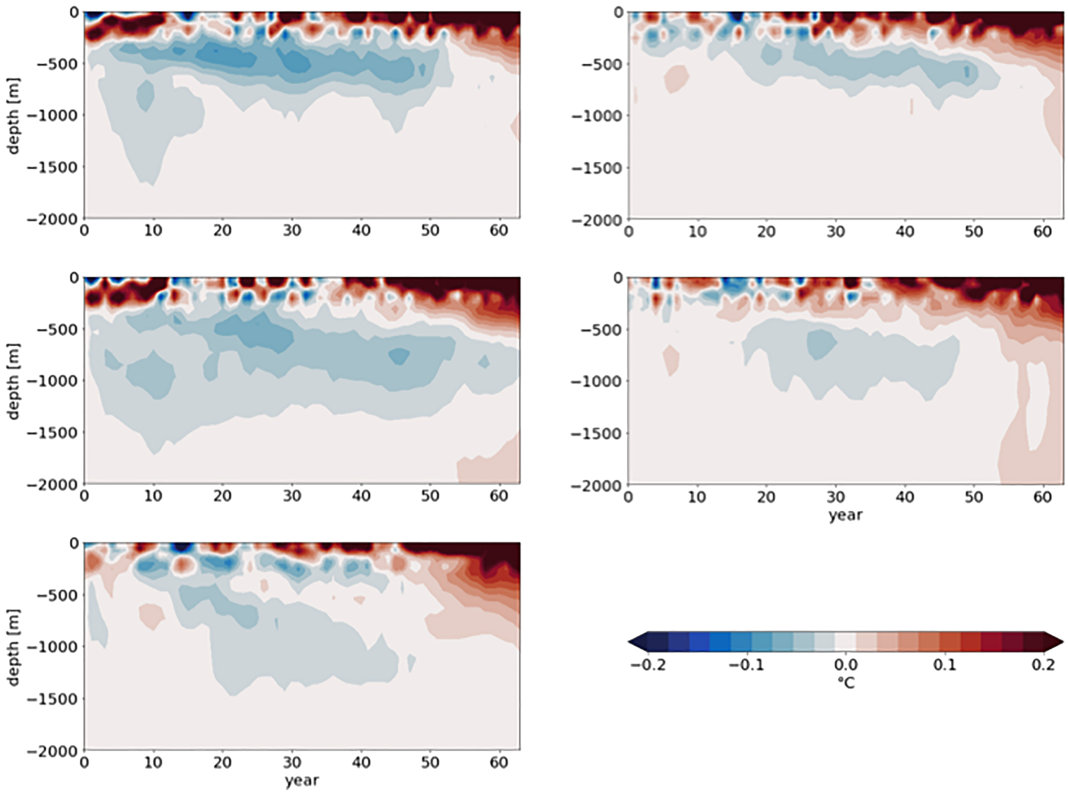

Figure 19. Hovmöller diagrams for differences in ocean global temperature profiles. Left column from top to bottom: hist-LR-1 minus control-1950-LR, hist-LR-2 minus control-1950-LR, and hist-HR minus control-1950-HR. In LR the first 15 years are contaminated by the cold start for the atmosphere and land. Right column from top to bottom: hist-LR-1 minus control-1950-LR-1 and hist-LR-2 minus control-1950-LR-2. The effect from the cold start for the atmosphere and land is removed if the control is perturbed accordingly.

15 years both LR runs simulate a cooler surface ocean that is accompanied by a subsurface warming. Such a behavior, but to a smaller extent, is also observed in HR. In all historical simulations the actual warming signal starts to appear after ~30 years. The repeatedly changing signal of warming and cooling at the surface and subsurface of the ocean superimposing the long-term evolution stems primarily from the El Niño–Southern Oscillation mode of variability (compare Deser et al., 2010; Rackow et al., 2016), which is, as expected, not coherent between the runs.

The question to which extent the inconsistency in the initialization contributes to these anomalies needs to be answered. To do this we performed two additional simulations with LR where we continued control-1950-LR but perturbed it with a cold start in ECHAM (control-1950-LR-1 and control-1950-LR-2). The years where these simulations are started coincide with those chosen for the start of the historical simulations. The signal from improper initialization nearly vanishes if we subtract the historical simulation from the perturbed control run as shown in Figure 19 (right panels) for the two LR realizations. The same behavior has been observed for other climate quantities that are not shown in this paper. At the same time, the warming signal over the last 20 years of the historical runs remains largely unchanged. We thus conclude that the cold start in the atmosphere-land component leads to a perturbation signal with the inertia timescale of ~15 years. After this time, for the climate quantities we have analyzed, effects from the improper initialization are below the noise level due to internal variability.

6. Discussion

In this paper, we performed a first evaluation of the new FESOM2.0-ECHAM6.3 coupled model. The dynamical core of FESOM2.0 is based on the finite volume discretization and provides speedups up to a factor of three compared to its preceding version 1.4, enabling a throughput similar to that of regular-mesh models (Scholz et al., 2019). The atmosphere, hydrology, and vegetation have been simulated with ECHAM6.3

Table 3
FESOM2.0 Performance Using LR and HR Meshes in Uncoupled and Coupled Configurations

Mesh (grid points)	Cores	Time step (min)	Uncoupled (SYPD)	Coupled (SYPD)
LR (126,858)	288	45	~144	~48
HR (1,306,775)	1,152	10	~12	~9

Note. ECHAM has been configured with 576 CPU cores in all runs. The performance has been computed on Ollie (AWI Bremerhaven), equipped with CPUs: Intel Xeon E5-2697 v4 (Broadwell) 2.3 GHz; Network: OmniPath, 100 Gbit/s. Intel 19.0.1 IntelMPI 2018.4 has been used.

that is also used for AWI-CM (Rackow et al., 2016; Sidorenko et al., 2015) and MPI-ESM (Giorgetta et al., 2013; Gutjahr et al., 2018) contributions to CMIP6. The coupling has been realized via the parallel OASIS3-MCT coupler.

We investigated (1) the mean state and long-term drift by analyzing a 1,000-year simulation under PI external forcing, (2) the fidelity in simulating the historical warming using the HighResMIP protocol, and (3) the influence of coarse versus eddy-resolving (permitting) ocean resolution.

The PI simulation exhibits a volume-averaged warming by ~0.8 °C over 1,000 years. The largest long-term warming occurs in the Southern Ocean (over timescales of ~60 years) and is associated with a sea ice decrease. Although this behavior is consistent with other models, it also raises the question of whether climate scenario simulations are reliable. On the one hand, running a climate model long enough to attain a quasi-equilibrium will result in an unrealistic climate state, in particular in the deeper ocean. On the other hand, imbalances that remain after starting a climate scenario shortly after the initialization might contribute spurious signals associated with model drift. The protocol suggested by HighResMIP, employed here, follows the second route.

Our analysis for the HiResMIP experiments shows that for some quantities the simulated signal associated with the historical warming is significantly smaller than the model drift, especially during the initialization phase. This becomes apparent from comparing Figures 9 and 19 for the profiles of global potential temperatures. Nevertheless, despite the presence of strong drift, the simulated climate warming shows rather persistent patterns of temperature change in both the ocean and the atmosphere (see Figures 16 and 19). We conclude that the "de-drifting" approach for the climate scenario simulations works reasonably well.

Increased ocean resolution results in more realistic SSTs, especially over the Southern Ocean. In this context, the oceanic eddies play an important role in maintaining slopes of isopycnals and thus the water mass structure. In turn, the simulated sea ice becomes more realistic. In the North Atlantic, the simulated SST bias around Newfoundland is less pronounced in HR compared to LR. Large biases are still present in HR in the deeper ocean and are different from those in LR. One reason for this might be improper ventilation at the surface caused by insufficient resolution in the atmosphere; in fact, Sein et al. (2018) have shown that an increase in oceanic resolution has to be complemented by an increase in atmospheric resolution.

The throughput of the new climate setup is presented in Table 3. The HR mesh is about 10 times larger than LR in terms of surface grid points. Only a rough estimate of the load balancing between FESOM and ECHAM was made for LR before the experiments were conducted. For all experiments here, the configuration of ECHAM is kept unchanged. ECHAM was set up using the same amount of 576 cores in all runs. FESOM was run using 288 cores and 1,152 cores for LR and HR meshes, respectively. It is worth mentioning that FESOM2.0 shows almost linear scalability up to ~300 surface grid points per core (Koldunov et al., 2019), implying that the throughputs of the ocean component presented here could easily be increased.

However, the throughput in LR drops from 144 simulated years per day (SYPD) to 48 SYPD from the uncoupled to the coupled configuration. In HR, the throughput drops from 12 SYPD to 9 SYPD. Increasing the computational resources for both components, however, did not improve the performance of the respective climate configurations. Indeed, the total speed of a climate setup is not only determined by the slowest component but also by the latency of storage access and the coupling. A detailed analysis of individual contributions and solutions for improvements will be discussed elsewhere. The main advantage of the new setup is that the ocean part requires less computational resources and thus allows higher resolutions in the ocean compared to older versions of FESOM.

7. Conclusions

A new climate model, using FESOM2.0 for the ocean and ECHAM6.3 for the atmosphere and land, has been developed. Using the new version of FESOM in the climate model promises higher computational efficiency and thus more efficient use of computational resources compared to its predecessor based on FESOM1.4. The new setup makes it possible to use spatial resolutions in the ocean that were not feasible before. The climate model has been evaluated in terms of its state and long-term drift under PI forcing and in terms of the fidelity in reproducing the historical climate change.

The results show that the performance of the new climate model is well within the range of existing models. We find that in terms of oceanic temperatures the historical warming is of smaller amplitude than the model drift. We also confirm that the approach of "de-drifting" short climate simulations, as suggested by the HighResMIP protocol, allows to isolate the warming signal. Furthermore we demonstrated that the new configuration can be used with a locally eddy-permitting (resolving) ocean. The oceanic mesh with ~1.3 million surface nodes exhibits notable improvements regarding the simulation of oceanic SSTs in the Southern Ocean compared to the low-resolution ocean mesh. Given the central role Southern Ocean temperatures play when it comes to Antarctic sea ice, bottom water formation, ice sheet processes, and thus sea level rise, we expect that these improvements translate into more reliable climate change projections, justifying the additional computational costs of higher ocean resolution.

Data Availability Statement

Data sets related to this article can be found at

https://swiftbrowser.dkrz.de/public/dkrz_035d8f6ff058403bb42f8302e6badfbc/Sidorenko_FESOM2.0_ECHAM6.3_2019/

Acknowledgments
We thank two anonymous reviewers for their comments which helped to considerably improve the manuscript. The work was supported by the Helmholtz Climate Initiative REKLIM (Regional Climate Change) (Q. Wang and D. Sidorenko), the APPLICATE (Advanced Prediction in Polar regions and beyond: modelling, observing system design and LInkages associated with a Changing Arctic climaTE, grant H2020-BG-2016-1) (T. Jung), the FRAM (FRontiers in Arctic marine Monitoring program) (C. Wekerle), contribution to the projects M3 and S2 of the Collaborative Research Centre TRR 181 "Energy Transfer in Atmosphere and Ocean" funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project 274762653 (S. Danilov, N. Koldunov, and S. Juricke), the EC Horizon 2020 project PRIMAVERA under the grant agreement 641727, the state assignment of FASO Russia theme 0149-2019-0015 (D. Sein), and the BMBF project SSIP (grant 01LN1701A; H. Goessling). The simulations were performed at the North-German Supercomputing Alliance (HLRN).

References

Danilov, S., Sidorenko, D., Wang, Q., & Jung, T. (2017). The Finite-volumE Sea ice–Ocean Model (FESOM2). *Geoscientific Model Development*, *10*(2), 765–789. https://doi.org/10.5194/gmd-10-765-2017

Danilov, S., Wang, Q., Timmermann, R., Iakovlev, N., Sidorenko, D., Kimmritz, M., et al. (2015). Finite-element sea ice model (FESIM), version 2. *Geoscientific Model Development*, *8*(6), 1747–1761. https://doi.org/10.5194/gmd-8-1747-2015

Delworth, T. L., Broccoli, A. J., Rosati, A., Stouffer, R. J., Balaji, V., Beesley, J. A., et al. (2006). GFDL's CM2 Global coupled climate models. Part I: formulation and simulation characteristics. *Journal of Climate*, *19*(5), 643–674. https://doi.org/10.1175/JCLI3629.1

Delworth, T. L., Rosati, A., Anderson, W., Adcroft, A. J., Balaji, V., Benson, R., et al. (2012). Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model. *Journal of Climate*, *25*(8), 2755–2781. https://doi.org/10.1175/JCLI-D-11-00316.1

Deser, C., Alexander, M. A., Xie, S. P., & Phillips, A. S. (2010). Sea surface temperature variability: Patterns and mechanisms. *Annual Review of Marine Science*, *2*, 115–143. https://doi.org/10.1146/annurev-marine-120408-151453

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Gent, P. R., & McWilliams, J. C. (1990). Isopycnal mixing in ocean circulation models. *Journal of Physical Oceanography*, *20*(1), 150–155. https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2

Gent, P. R., Willebrand, J., McDougall, T. J., & McWilliams, J. C. (1995). Parameterizing eddy-induced tracer transports in Ocean circulation models. *Journal of Physical Oceanography*, *25*(4), 463–474. https://doi.org/10.1175/1520-0485(1995)025<0463:PEITTI>2.0.CO;2

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., et al. (2013). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems*, *5*, 572–597. https://doi.org/10.1002/jame.20038

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

Griffies, S. M., Winton, M., Donner, L. J., Horowitz, L. W., Downes, S. M., Farneti, R., et al. (2011). The GFDL CM3 Coupled Climate Model: Characteristics of the ocean and sea ice simulations. *Journal of Climate*, *24*(13), 3520–3544. https://doi.org/10.1175/2011JCLI3964.1

Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., von Storch, J.-S., Brüggemann, N., et al., 2018: Max Planck Institute Earth System Model (MPI-ESM 1.2) for High-Resolution Model Intercomparison Project (HighResMIP), *Geoscientific Model Development Discussion*. https://doi.org/10.5194/gmd-2018-286

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., et al. (2016). High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6. *Geoscientific Model Development*, *9*(11), 4185–4208. https://doi.org/10.5194/gmd-9-4185-2016

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., et al. (2013). Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *Journal of Advances in Modeling Earth Systems*, *5*, 422–446. https://doi.org/10.1002/jame.20023

Kennedy, J. J., Rayner, N. A., Smith, R. O., Saunby, M., & Parker, D. E. (2011a). Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 1: measurement and sampling errors. *Journal of Geophysical Research*, *116*, D14103. https://doi.org/10.1029/2010JD015218

Kennedy, J. J., Rayner, N. A., Smith, R. O., Saunby, M., & Parker, D. E. (2011b). Reassessing biases and other uncertainties in sea-surface temperature observations since 1850 part 2: Biases and homogenisation. *Journal of Geophysical Research*, *116*, D14104. https://doi.org/10.1029/2010JD015220

Koldunov, N. V., Aizinger, V., Rakowsky, N., Scholz, P., Sidorenko, D., Danilov, S., & Jung, T. (2019). Scalability and some optimization of the Finite-volumE Sea ice–Ocean Model, Version 2.0 (FESOM2). *Geoscientific Model Development*, *12*(9), 3991–4012. https://doi.org/10.5194/gmd-12-3991-2019

Koldunov, N. V., Stammer, D., & Marotzke, J. (2010). Present-day Arctic sea ice variability in the coupled ECHAM5/MPI-OM model. *Journal of Climate*, *23*(10), 2520–2543. https://doi.org/10.1175/2009JCLI3065.1

Large, W. G., McWilliams, J. C., & Doney, S. C. (1994). Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Reviews of Geophysics*, *32*(4), 363–403. https://doi.org/10.1029/94RG01872

Large, W. G., & Yeager, S. G. (2009). The global climatology of an interannually varying air–sea flux data set. *Climate Dynamics*, *33*(2-3), 341–364. https://doi.org/10.1007/s00382-008-0441-3

Lindsay, R., & Schweiger, A. (2015). Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations. *The Cryosphere*, *9*(1), 269–283. https://doi.org/10.5194/tc-9-269-2015

Lucarini, V., & Ragone, F. (2011). Energetics of climate models: Net energy balance and meridional enthalpy transport. *Reviews of Geophysics*, *49*, RG1001. https://doi.org/10.1029/2009RG000323

Marshall, J., & Speer, K. (2012). Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geoscience*, *5*(3), 171–180. https://doi.org/10.1038/ngeo1391

Morice, C. P., Kennedy, J. J., Rayner, N. A., & Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 dataset. *Journal of Geophysical Research*, *117*, D08101. https://doi.org/10.1029/2011JD017187

Notz, D., Haumann, F. A., Haak, H., Jungclaus, J. H., & Marotzke, J. (2013). Arctic sea-ice evolution as modeled by Max Planck Institute for Meteorology's Earth system model. *Journal of Advances in Modeling Earth Systems*, *5*, 173–194. https://doi.org/10.1002/jame.20016

Pacanowski, R., & Philander, S. (1981). Parameterization of vertical mixing in numerical models of tropical oceans. *Journal of Physical Oceanography*, *11*(11), 1443–1451. https://doi.org/10.1175/1520-0485(1981)011<1443:POVMIN>2.0.CO;2

Pithan, F., Medeiros, B., & Mauritsen, T. (2014). Mixed-phase clouds cause climate model biases in Arctic wintertime temperature inversions. *Climate Dynamics*, *43*(1-2), 289–303. https://doi.org/10.1007/s00382-013-1964-9

Rackow, T., Goessling, H. F., Jung, T., Sidorenko, D., Semmler, T., Barbi, D., & Handorf, D. (2016). Towards multi-resolution global climate modeling with ECHAM6-FESOM. Part II: climate variability. *Climate Dynamics*, *50*(7-8), 2369–2394. https://doi.org/10.1007/s00382-016-3192-6

Rackow, T., Sein, D. V., Semmler, T., Danilov, S., Koldunov, N. V., Sidorenko, D., et al. (2019). Sensitivity of deep ocean biases to horizontal resolution in prototype CMIP6 simulations with AWI-CM1.0. *Geoscientific Model Development*, *12*(7), 2635–2656. https://doi.org/10.5194/gmd-2018-192

Reichler, T., & Kim, J. (2008). How well do coupled models simulate today's climate? *Bulletin of the American Meteorological Society*, *89*(3), 303–312. https://doi.org/10.1175/BAMS-89-3-303

Ringler, T., Petersen, M., Higdon, R., Jacobsen, D., Maltrud, M., & Jones, P. (2013). A multi-resolution approach to global ocean modelling. *Ocean Modelling*, *69*, 211–232. https://doi.org/10.1016/j.ocemod.2013.04.010

Scaife, A. A., Copsey, D., Gordon, C., Harris, C., Hinton, T. J., Keeley, S. P., et al. (2011). Improved Atlantic winter blocking in a climate model. *Geophysical Research Letters*, *38*, L23703. https://doi.org/10.1029/2011GL049573

Scholz P., Sidorenko D., Gurses O., Danilov S., Koldunov N., Wang Q., et al., 2019: Assessment of the Finite VolumE Sea Ice Ocean Model (FESOM2.0), Part I: Description of selected key model elements and comparison to its predecessor version. *Geoscientific Model Development Discussions*. https://doi.org/10.5194/gmd-2018-329, 1–42

Sein, D. V., Danilov, S., Biastoch, A., Durgadoo, J. V., Sidorenko, D., Harig, S., & Wang, Q. (2016). Designing variable ocean model resolution based on the observed ocean variability. *Journal of Advances in Modeling Earth Systems*, *8*, 904–916. https://doi.org/10.1002/2016MS000650

Sein, D. V., Koldunov, N. V., 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. *Journal of Advances in Modeling Earth Systems*, *10*, 2026–2041. https://doi.org/10.1029/2018MS001327

Sein, D. V., Koldunov, N. V., Danilov, S., Wang, Q., Sidorenko, D., Fast, I., et al. (2017). Ocean modeling on a mesh with resolution following the local Rossby radius. *Journal of Advances in Modeling Earth Systems*, *9*, 2601–2614. https://doi.org/10.1002/2017MS001099

Sgubin, G., Swingedouw, D., Drijfhout, S., Mary, Y., & Bennabi, A. (2017). Abrupt cooling over the North Atlantic in modern climate models. *Nature Communications*, *8*, 14375. https://doi.org/10.1038/ncomms14375

Sidorenko, D., Rackow, T., Jung, T., Semmler, T., Barbi, D., Danilov, S., et al. (2015). Towards multi-resolution global climate modeling with ECHAM6–FESOM. Part I: Model formulation and mean climate. *Climate Dynamics*, *44*(3-4), 757–780. https://doi.org/10.1007/s00382-014-2290-6

Sidorenko, D., Koldunov, N., Wang, Q., Danilov, S., Goessling, H. F., Gurses, O., et al. (2018). Influence of a salt plume parameterization in a coupled climate model. *Journal of Advances in Modeling Earth Systems*. https://doi.org/10.1029/2018MS00129

Skamarock, W. C., Klemp, J. B., Duda, M. G., Fowler, L. D., Park, S. H., & Ringler, T. D. (2012). A multiscale nonhydrostatic atmospheric model using centroidal Voronoi tesselations and C-grid staggering. *Monthly Weather Review*, *140*(9), 3090–3105. https://doi.org/10.1175/MWR-D-11-00215.1

Steele, M., Morley, R., & Ermold, W. (2001). PHC: A global ocean hydrography with a high-quality Arctic Ocean. *Journal of Climate*, *14*(9), 2079–2087. https://doi.org/10.1175/1520-0442(2001)014<2079:PAGOHW>2.0.CO;2

Sterl, A., Bintanja, R., Brodeau, L., Gleeson, E., Koenigk, T., Schmith, T., et al. (2012). A look at the ocean in the EC-Earth climate model. *Climate Dynamics*, *39*(11), 2631–2657. https://doi.org/10.1007/s00382-011-1239-2

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., et al. (2013). Atmospheric component of the MPI-M Earth System Model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, *5*, 146–172. https://doi.org/10.1002/jame.20015

Valcke, S. (2013). The OASIS3 coupler: A European climate modelling community software. *Geoscientific Model Development*, *6*(2), 373–388 . https://doi.org/10.5194/gmd-6-373-2013

Wang, Q., Danilov, S., Sidorenko, D., Timmermann, R., Wekerle, C., Wang, X., et al. (2014). The Finite Element Sea Ice-Ocean Model (FESOM) v. 1.4: Formulation of an ocean general circulation model. *Geoscientific Model Development*, *7*(2), 663–693. https://doi.org/10.5194/gmd-7-663-2014

Wang, Q., Ilicak, M., Gerdes, R., Drange, H., Aksenov, Y., Bailey, D. A., et al. (2016a). An assessment of the Arctic Ocean in a suite of interannual CORE-II simulations. Part I: Sea ice and solid freshwater. *Ocean Modelling*, *99*, 110–132. https://doi.org/10.1016/j.ocemod.2015.12.008

Wang, Q., Ilicak, M., Gerdes, R., Drange, H., Aksenov, Y., Bailey, D. A., et al. (2016b). An assessment of the Arctic Ocean in a suite of interannual CORE-II simulations. Part II: Liquid freshwater. *Ocean Modelling*, *99*, 86–109. https://doi.org/10.1016/j.ocemod.2015.12.009

Wekerle, C., Wang, Q., Danilov, S., Schourup-Kristensen, V., von Appen, W. J., & Jung, T. (2017). Atlantic water in the Nordic Seas: Locally eddy-permitting ocean simulation in a global setup. *Journal of Geophysical Research: Oceans*, *122*, 914–940. https://doi.org/10.1002/2016JC012121