Ocean model formulation influences transient climate response

Tido Semmler[1], Johann H Jungclaus[2], Christopher Danek[3], Helge Goessling[1], Nikolay V. Koldunov[4], Thomas Rackow[5], and Dmitry Sidorenko[4]

[1]Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research
[2]Max Planck Institute for Meteorology
[3]Alfred-Wegener-Institute
[4]Alfred Wegener Institute for Polar and Marine Research
[5]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research

November 28, 2022

Abstract

The transient climate response (TCR) is 20% higher in the Alfred Wegener Institute Climate Model (AWI-CM) compared to the Max Planck Institute Earth System Model (MPI-ESM) whereas the equilibrium climate sensitivity (ECS) is only by less than 10% higher in AWI-CM. These results are largely independent of the two considered model resolutions for each model. The two coupled CMIP6 models share the same atmosphere-land component ECHAM6.3 developed at the Max Planck Institute for Meteorology (MPI-M). However, ECHAM6.3 is coupled to two different ocean models, namely the MPIOM sea ice-ocean model developed at MPI-M and the FESOM sea ice-ocean model developed at the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI). A reason for the different TCR is related to ocean heat uptake in response to greenhouse gas forcing. Specifically, AWI-CM simulations show stronger surface heating than MPI-ESM simulations while the latter accumulate more heat in the deeper ocean. The vertically integrated ocean heat content is increasing slower in AWI-CM model configurations compared to MPI-ESM model configurations in the high latitudes. Weaker vertical mixing in AWI-CM model configurations compared to MPI-ESM model configurations seems to be key for these differences. The strongest difference in vertical ocean mixing occurs inside the Weddell Gyre and the northern North Atlantic. Over the North Atlantic, these differences materialize in a lack of a warming hole in AWI-CM model configurations and the presence of a warming hole in MPI-ESM model configurations. All these differences occur largely independent of the considered model resolutions.

Ocean model formulation influences transient climate response

Tido Semmler[1], Johann Jungclaus[2], Christopher Danek[1], Helge F Goessling[1], Nikolay Koldunov[1], Thomas Rackow[1], and Dmitry Sidorenko[1]

[1]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

[2]Max Planck Institute for Meteorology, Hamburg, Germany

Corresponding author: Tido Semmler, e-mail tido.semmler@awi.de

Key points

The transient climate response in two coupled models with the same atmosphere but different ocean components differs by 20%

The upper (deeper) ocean heats faster (slower) in AWI-CM compared to MPI-ESM, independent of model resolution

Vertical mixing in the northern North Atlantic and the Weddell Gyre appears to be key for these differences

Abstract

The transient climate response (TCR) is 20% higher in the Alfred Wegener Institute Climate Model (AWI-CM) compared to the Max Planck Institute Earth System Model (MPI-ESM) whereas the equilibrium climate sensitivity (ECS) is only by less than 10% higher in AWI-CM. These results are largely independent of the two considered model resolutions for each model. The two coupled CMIP6 models share the same atmosphere-land component ECHAM6.3 developed at the Max Planck Institute for Meteorology (MPI-M). However, ECHAM6.3 is coupled to two different ocean models, namely the MPIOM sea ice-ocean model developed at MPI-M and the FESOM sea ice-ocean model developed at the Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research (AWI). A reason for the different TCR is related to ocean heat uptake in response to greenhouse gas forcing. Specifically, AWI-CM simulations show stronger surface heating than MPI-ESM simulations while the latter accumulate more heat in the deeper ocean. The vertically integrated ocean heat content is increasing slower in AWI-CM model configurations compared to MPI-ESM model configurations in the high latitudes. Weaker vertical mixing in AWI-CM model configurations compared to MPI-ESM model configurations seems to be key for these differences. The strongest difference in vertical ocean mixing occurs inside the Weddell Gyre and the northern North Atlantic. Over the North Atlantic, these differences materialize in a lack of a warming hole in AWI-CM model configurations and the presence of a warming hole in MPI-ESM model configurations. All these differences occur largely independent of the considered model resolutions.

Key words

Transient climate response, CMIP6 simulations, vertical mixing

Plain language summary

The transient climate response describes how strongly near-surface temperatures warm in response to gradually increasing greenhouse-gas levels. Here we investigate the role of the ocean which takes up heat and thereby delays the surface warming. Two models of the Coupled Model Intercomparison Project Phase 6 (CMIP6), the Alfred Wegener Institute Climate Model (AWI-CM) and the Max Planck Institute Earth System Model (MPI-ESM), which use the same atmosphere model but different ocean models are selected for this study. In AWI-CM the upper ocean layers heat faster than in MPI-ESM, while the opposite is true for the deep ocean. As a consequence, the transient climate response is 20% stronger in AWI-CM compared to MPI-ESM. We find that weaker vertical ocean mixing in AWI-CM compared to MPI-ESM, especially over the northern North Atlantic and the Weddell Sea, is key for these differences. Our findings corroborate the importance of realistic ocean mixing in climate models when it comes to getting the strength and timing of climate change right.

Introduction

The equilibrium climate sensitivity (ECS) and the transient climate response (TCR) are important metrics of how a climate model reacts to greenhouse gas forcing (Meehl et al., 2020). The ECS has been already computed from climate models as early as in the 1970s (Charney et al., 1979) and the TCR in the 1980s (Stouffer et al., 1989). Both metrics have been used over the generations of model intercomparison projects (MIPs) that serve as a basis for the assessment reports of the Intergovernmental Panel on Climate Change (IPCC). The ECS is commonly defined as the 2 m temperature response to a doubling of CO_2 after equilibration of the climate system. The TCR is the transient 2 m temperature response from a 1% per year CO_2 increase at the time of a doubling of CO_2. The ECS is known to be mainly dependent on the atmosphere model while the TCR also critically depends on the pattern of ocean surface warming (Meehl et al., 2020). The ocean with its high heat capacity plays an important role in delaying the response to a forcing resulting in a clearly lower TCR compared to the ECS (Meehl et al., 2020). Cloud feedbacks and cloud-aerosol interactions in models with prognostic aerosols have a large influence on ECS (Mauritsen & Roeckner, 2020; Meehl et al., 2020; Gettelman et al., 2019) although ocean heat uptake and inhomogeneous Pacific surface warming that may not be properly simulated by CMIP5 models have been recently acknowledged as a source of uncertainty (Sherwood et al., 2020).

Typically, a coupled atmosphere-ocean model needs to be spun up for at least 500 years to account for adjustments on comparably short time scales (Rackow et al., 2016) while it may take 5000 years to achieve an equilibrium of the deep ocean (Li et al., 2012). The e-folding time (time when the difference between simulated and equilibrium temperature of the deep ocean has decayed to 1/e times the difference between initialized and equilibrium temperature) of a previous version of the Alfred Wegener Institute Climate Model (AWI-CM) has been found to be about 800 years (Rackow et al., 2016). For a radiation imbalance, changes of effectiveness of vertical mixing and diffusion at high latitudes as well as the thermohaline circulation have been found to be important for the heat redistribution in the ocean (Banks and Gregory, 2006; Gregory et al., 2000). Depending on the intensity of the circulation and vertical mixing the ocean can be a heat sink of different intensity as a response to greenhouse gas forcing which may also influence the climate sensitivity, especially the transient climate response (TCR). It has been found that the ocean takes up 93% of the current (recent past) radiation imbalance (Rhein et al., 2013).

The role of the ocean for the climate sensitivity, especially TCR, has been investigated in various previous studies: Winton et al. (2014) examine the role of horizontal resolution (eddy parameterizing versus eddy resolving) on the TCR and attribute changes in the TCR to differences in initial Atlantic meridional overturning circulation (AMOC) as well as AMOC decline and Southern Ocean surface warming under increasing greenhouse gas concentrations. Models with weaker AMOC decline tend to show higher TCR to ECS ratios. He et al. (2017) point out that starting from different ocean climates can have an influence on ocean circulation changes. In their study, a simulation with initially weaker overturning shows a smaller overturning decrease than the one with initially stronger overturning. The weaker overturning decrease leads to stronger high latitude surface heating - since the poleward energy transport is less decreased - and thus stronger TCR compared to their simulation with stronger overturning decrease. Furthermore, convection regions in the high latitudes such as Labrador, Weddell, and Ross Sea play an important role for the redistribution of the heat. In addition, TCR may be strongly influenced by the calculation method, and long control simulations are important to get a robust estimate of TCR (Liang et al., 2013). This is in line with investigations of variability in a pre-industrial control state which is stronger than the variability in a transient state with increasing greenhouse gas concentrations (Brierley et al., 2009). Newsom et al. (2020) argue that differences in surface warming and its influences on ventilation especially in the Southern Ocean and subtropical areas are important for the heat uptake of the ocean.

Two models in the CMIP6 archive, MPI-ESM and AWI-CM, offer the opportunity to examine the influence of a different ocean sea-ice model formulation on the climate while leaving the atmosphere and land surface component untouched. In contrast to previous studies (e.g., Winton et al., 2014; Yokohata et al., 2007), we consider not only different horizontal ocean resolutions within one model, but also consider completely different ocean models. Both models utilize the atmosphere and land surface model ECHAM 6.3 / JSBACH 3.2.0. For the ocean and sea ice, MPI-ESM uses MPIOM and AWI-CM FESOM. The models are documented in Mauritsen et al. (2019) and Semmler et al. (2020). For both models, two different configurations are considered in this study: the low resolution version with dynamic vegetation in JSBACH and the high / medium resolution without dynamic vegetation in JSBACH.

The MPI-ESM versions have been tuned to an ECS of 3°C to match the observed historical 2 m temperature development over the last 150 years (Mauritsen and Roeckner 2020). Nevertheless, since in the AWI-CM setups the same ECHAM 6.3 / JSBACH 3.2.0 version as in MPI-ESM has been used without any further tuning, different ECS and TCR in the AWI-CM setups can be interpreted as the influence of the different ocean model formulation and the slightly different land sea masks resulting from the coupling to the different ocean model formulation. A third candidate for this comparison could be the FOCI model (ECHAM6.3 / JSBACH 3 coupled to the ocean model NEMO: Matthes et al., 2020). However, in this case the ECHAM 6.3 model has been differently tuned and therefore the reasons for different developments of the ocean state cannot be purely ascribed to the different ocean model formulation.

In section 2 a brief model description of both coupled systems is given; section 3 compares and explains the results of the different models. In section 4 the results are discussed in the light of previous studies and in section 5 conclusions are given.

Model components

In the following, a brief summary of the model components used in this study is given. We briefly characterize the model components rather than the coupled models since both coupled models share the same atmospheric and land components. Both coupled models employ the coupler OASIS3-MCT_3.0 (Craig et al., 2017).

ECHAM 6.3 / JSBACH 3.2.0

The general circulation model for the atmosphere is ECHAM 6.3 as it is implemented in the CMIP6 version of MPI-ESM (Mauritsen et al., 2019; and Mueller et al., 2018; Mauritsen and Roeckner, 2020). ECHAM consists of a dry spectral-transfer dynamical core, a transport model and a suite of physical parameterizations. The vertical discretization employs a hybrid sigma-pressure coordinate system. ECHAM's design principles and features are described in detail in Stevens et al. (2013) and Mauritsen et al. (2019) account for changes from the CMIP5 to the CMIP6 generation of ECHAM. The model is applied here in two configurations that differ in horizontal and vertical resolution. The low-resolution (LR) version applies a triangular truncation of the spherical harmonics to 63 wave numbers (T63). In physical space this transfers to a grid spacing of roughly 200km. In the vertical, the LR version employs 47 levels. The higher-resolution (HR) set-up has a T127 truncation (roughly 100km) and 95 vertical levels. Both versions resolve the atmosphere up to 0.01 hPa or about 80km.

The land component is represented in the JSBACH model (Reick et al., 2013; Mauritsen et al., 2019; Reick et al., 2021). JSBACH 3.2.0 provides the lower boundary conditions for the atmosphere and includes the land geochemistry, soil hydrology, the terrestrial carbon cycle, and a river-routing scheme. In the low resolution version of JSBACH 3.2.0, dynamic vegetation is considered while in the high resolution version the vegetation is prescribed. While for the AWI model configurations this difference is reflected in the model name of the coupled system (AWI-ESM with dynamic vegetation and AWI-CM without dynamic vegetation) this is not the case for the MPI model configurations (both are called MPI-ESM).

FESOM 1.4

A detailed description of FESOM 1.4 is given by Wang et al. (2014). This model is the first global sea ice ocean model to use unstructured meshes with variable resolution for climate research (Wang et al., 2014). The mesh flexibility allows to increase resolution in dynamically active regions, while keeping a relatively coarse resolution elsewhere. FESOM allows global multi-resolution simulations without traditional nesting. The dynamical core of FESOM 1.4 employs the finite element method to solve the primitive equations. The mesh is composed of horizontal triangles that constitute the faces of 3-dimensional prisms which are cut into tetrahedral elements. The ice module FESIM uses 0-layer thermodynamics and elastic-viscous-plastic rheology. Small-scale mixing along isopycnals as well as tracer stirring through eddies are parameterized (Redi 1982, Gent & McWilliams 1990) and scaled by the local horizontal resolution. Vertical mixing is implemented via KPP (Large et al. 1994). FESOM1.4 is used in this study in two different horizontal resolutions while maintaining the same vertical 46 levels. The low resolution varies between about 25 km in the Arctic and the tropics and around 100 km in the subtropical regions (Fig. 1a). The high resolution features resolutions as high as 8 km over key ocean regions such as the Gulf Stream / North Atlantic Current area, parts of the Southern Ocean as well as coastal regions (Fig. 1b). Over the subtropical areas the horizontal resolution is around 80 km.

MPIOM

MPIOM is the ocean component of MPI-ESM (Jungclaus et al., 2013; Mauritsen et al., 2019; Mueller et al., 2018). MPIOM applies the Boussinesq and hydrostatic approximations and is discretized on a Arakawa-C grid in the horizontal and on a z-level grid in the vertical direction (Marsland et al., 2003). Subgrid-scale parameterizations such as those for lateral mixing on isopycnals and tracer transports by unresolved eddies are described in Jungclaus et al. (2013). Vertical mixing employs a Richardson-number dependent formulation (Pacanowski and Philander, 1981) and directly wind-induced mixing in the mixed layer (Marsland et al., 2003). The configurations used in this study differ in their horizontal grid design (Figure 1). The low-resolution (LR) version uses a bi-polar grid, where one grid pole is located over Antarctica, the other over Greenland. The resulting grid features enhanced resolution in the northern deep water formation regions and the Greenland-Scotland Ridge (Fig. 1c). The nominal 1.5 degree resolution is therefore transformed to less than 20 km near Greenland and almost 200 km in the tropical Pacific. The three-pole "HR" set-up (Jungclaus et al., 2013) has a more uniform resolution of 0.4 degree (Fig. 1d), which can be classified as "eddy-permitting". The vertical dimension is represented in both cases by 40 levels with the first 20 levels covering the upper 700m. The dynamical sea-ice model in MPIOM uses a viscoplastic rheology following Hibler (1979) and the thermodynamic representation of sea ice is based on a simple zero-layer mono-category formulation (Semtner, 1976). The sea-ice model is basically unchanged from the version described in Notz et al. (2013).

Low resolution High resolution

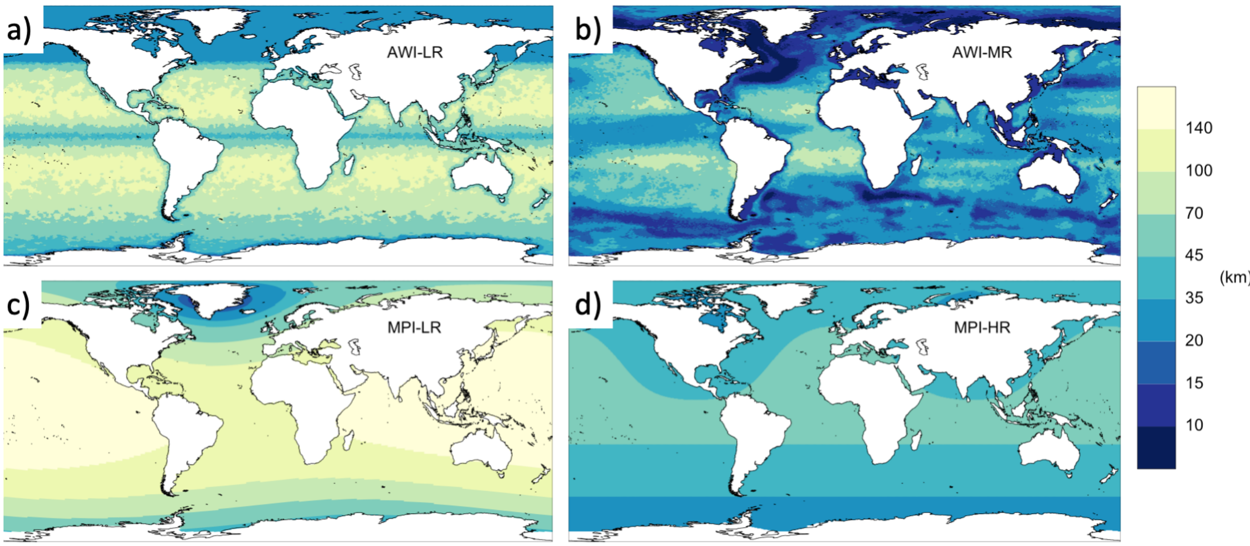


Fig. 1: Resolution of the four ocean grids, computed as the square-root of grid-cell area. (a) AWI-LR, (b) AWI-MR, (c) MPI-LR, (d) MPI-HR.

Methods

Simulations

In this study we use piControl, 1pctCO2, and abrupt-4xCO2 simulations from the Coupled Model Intercomparison Project 6 (CMIP6). A detailed description about the experimental protocol of these "Diagnostics, Evaluation and Characterization of Klima" (DECK) simulations is given in Eyring et al. (2016a). In short, piControl is a long control simulation of at least 500 years with constant 1850 greenhouse gas, aerosol, ozone, and solar forcing. Both the 1pctCO2 and the abrupt-4xCO2 simulations are branched off from the piControl simulation. For 1pctCO2, the CO_2 concentration is increased by 1% / year for 140 years reaching a quadrupling at the end. For abrupt-4xCO2, the CO_2 concentration is quadrupled instantaneously at the beginning of the simulation and the simulation has been carried out for 150 years. The simulations of four model configurations are used in this study: AWI-CM-1-1-MR, AWI-ESM-1-1-LR, MPI-ESM1-2-HR, and MPI-ESM1-2-LR, for simplicity in the following referred to as AWI-MR, AWI-LR, MPI-HR, and MPI-LR. Overviews over the results of these simulations have been published in

Semmler et al. (2020) for AWI-MR, and for MPI-LR and HR in Mauritsen et al. (2019), Müller et al. (2018), and Mauritsen and Roeckner (2020). Data of all four configurations are published at the Earth System Grid Federation (ESGF) (Semmler et al., 2018; Danek et al., 2020; Jungclaus et al., 2019; Wieners et al., 2019).

In the following, we classify AWI-MR and MPI-HR as high resolution configurations and AWI-LR and MPI-LR as low resolution configurations. AWI-MR consists of the high resolution version of FESOM1.4 coupled to the high resolution version of ECHAM6.3. It is called AWI-MR and not AWI-HR because in CMIP6 this set-up is the MR version of AWI-CM. MPI-HR consists of the high resolution version of MPIOM coupled to the high resolution version of ECHAM6.3. Similarly, AWI-LR consists of low resolution FESOM1.4 coupled to low resolution ECHAM6.3 and MPI-LR of low resolution MPIOM coupled to low resolution ECHAM6.3.

Unless otherwise stated, we use years 60-80 of the 1pctCO2 simulation and compare these to the corresponding years 60-80 (after branch-off of the corresponding 1pctCO2 simulation) of the piControl simulation. Since in the 1pctCO2 simulations we are in a strongly transient climate, we opt for a relatively short averaging time period as we otherwise may mix signals from substantially different states of the climate system.

ECS and TCR

For the calculation of the ECS and the TCR, we use the same methodologies as defined in the ESMValTools (Eyring et al., 2016b). The ECS is computed according to the method of Gregory et al. (2004). For each year, the near- surface (2 m) air temperature change and the change in net downward radiative flux between the abrupt- 4xCO_2 and piControl simulations are computed. For this, a linear fit of the piControl simulation to the 150 years corresponding to the years 1-150 of the abrupt-4xCO_2 simulation is calculated. These annual fitted values are subtracted from the abrupt-4xCO_2 simulation annual mean values. To compute the equilibrium temperature difference, a regression is built from all insofar detrended data points and extrapolated to the equilibrium (net shortwave radiation change = 0). The ECS is then obtained by dividing the equilibrium temperature difference by 2. The TCR is computed as the globally averaged 2 m temperature change of the 1pctCO2 simulation versus piControl (CO_2-doubling is reached after ~70 years) averaged over the years 61-80. Here, the linear fit detrending was based on the first 140 years of the piControl simulation counted from the branch point of the 1pctCO2 simulation.

Two-layer energy balance model (EBM)

Simplified climate models relating the global mean surface temperature to a prescribed effective radiative forcing (Held et al., 2010, Geoffroy et al., 2013; Mauritsen et al., 2019) are able to emulate the thermal properties and time-dependent responses seen in coupled atmosphere ocean models (AOGCM). Using analytical solutions, Geoffroy et al. (2013) provided a method to calibrate the parameters of a two-layer energy balance model (EBM) so that it can mimic a specific AOGCM. The first layer corresponds to the atmosphere, the land surface and the upper ocean, the second layer represents the deeper ocean below the mixed layer. The equations can be written as follows:

$$C\frac{dT}{dt} = \mathcal{F} - \lambda T - \gamma(T - T_0) \quad \text{and} \qquad (1)$$

$$C_0\frac{dT_0}{dt} = \gamma(T - T_0). \qquad (2)$$

Here the prognostic variables T and T_0 are the temperature perturbation at the surface and a characteristic temperature perturbation of the deeper ocean, respectively. F is the "effective" radiative forcing. The other symbols are free parameters: C and C_0 are effective heat capacities of the upper and deeper layer, λ the climate feedback parameter and γ is the heat uptake coefficient of the deeper ocean. The heat capacities C and C_0 correspond to equivalent ocean layer depths D and D_0 using equation 22 of Geoffroy et al. (2013).

In a first step the feedback parameter is estimated by a linear regression of the radiative imbalance at the top-of-the-atmosphere as a function of the surface temperature perturbation (Gregory et al., 2004). The second step involves the analytical solution of (1) and (2). For step forcing, the time evolution of surface and deeper ocean temperatures take the form

$$T(t) = T_{eq} - a_f T_{eq} e^{-t/\tau_f} - a_s T_{eq} e^{-t/\tau_s} \quad \text{and}$$

$$T_0(t) = T_{eq} - \phi_f a_f T_{eq} e^{-t/\tau_f} - \phi_s a_s T_{eq} e^{-t/\tau_s}.$$

The parameters can be determined by fitting the global mean surface air temperature response and regression analyses (see Geoffroy et al. (2013) for details). The solution includes the characteristic time scales for fast (τ_f) and slow (τ_s) response. This method is applied to the abrupt 4xCO_2 experiments where the effective radiative forcing is constant in time and is determined following Gregory et al. (2004).

Results

ECS and TCR from the four model configurations are shown in Table 1. The four values of TCR are around the mean value of 2.0°C from the CMIP6 simulations available at ESGF in March 2020 (Meehl et al., 2020) while the four values of ECS are all below the mean value of 3.7°C from these CMIP6 simulations (Meehl et al., 2020). Compared to the range of existing TCR values (1.3 to 3.0 °C) and existing ECS values (1.8 to 5.6 °C) according to Meehl et al. (2020) differences between the four model configurations are rather small (up to about 20% for TCR and up to about 10% for ECS). However, there are some important differences between ocean characteristics of the four model configurations that lead to the differences especially in the TCR which are described in the following.

Table 1: Equilibrium Climate Sensitivity (ECS) and Transient Climate Response (TCR) from the four model configurations, computed according to the definition in the ESMValTools (Eyring et al., 2016b).

	AWI-LR	MPI-LR	AWI-MR	MPI-HR
ECS (°C)	3.29±0.02	3.00±0.02*	3.16±0.03	2.98±0.03*

TCR (°C)	2.11	1.84	2.06	1.66

*tuned to be close to 3.0 (to match the observed historical development of 2 m temperature)

The AWI configurations heat stronger at the surface in most areas compared to the MPI configurations in the 1pctCO2 experiment (Fig. 2), consistent with their higher TCR. In the low resolution, the surface heating is up to around 1.5 °C stronger in the AWI configuration compared to the MPI configuration over the Southern Ocean as well as over parts of the North Atlantic and North Pacific Ocean. There are some limited areas around the gateways to the Arctic as well as in the Agulhas region and south of Australia in which the AWI low resolution configuration shows up to around 1°C less surface heating compared to the MPI low resolution configuration. In the high resolution, strongest positive differences in the surface heating of up to around 2°C in AWI compared to MPI occur over parts of the North Atlantic and North Pacific including the gateways to the Arctic. Small negative differences of up to 0.5°C are mainly restricted to subtropical areas in the Southern Hemisphere.

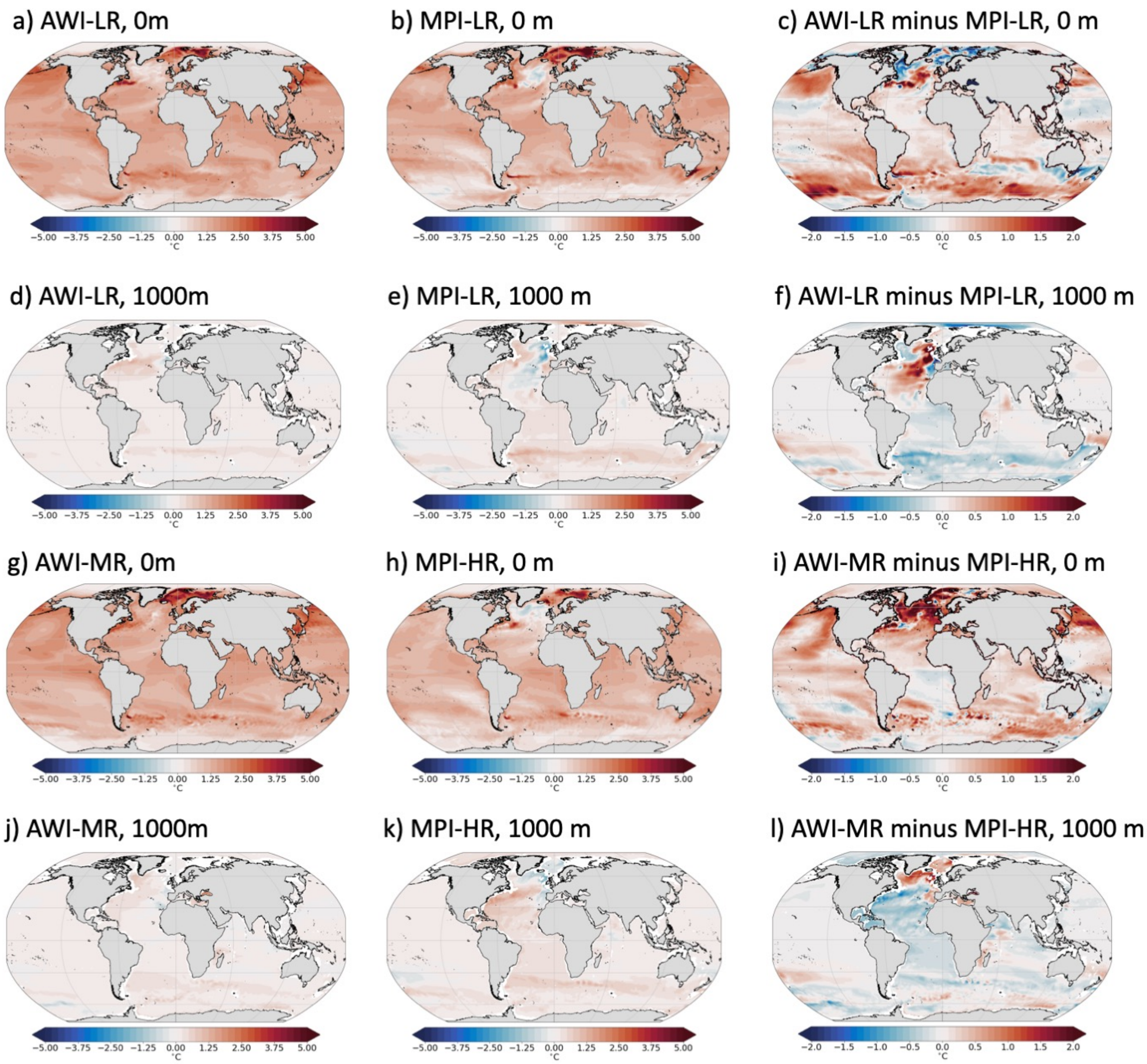


Fig. 2: Ocean potential temperature response at 0 m and 1000 m for years 60-80 (21 years centered around doubling of CO_2) from 1pctCO2 simulation compared to piControl. From left to right: AWI, MPI, difference AWI minus MPI. (a), (b), (c): low resolution at 0 m; (d), (e), (f): low resolution at 1000 m; (g), (h), (i): high resolution at 0 m; (j), (k), (l): high resolution at 1000 m.

While globally averaged there is more surface heating in the AWI configurations compared to the MPI configurations, the opposite is true for the mid-depth ocean (1000 m) (Fig. 2). In the Southern Ocean as well as in the Arctic ocean AWI-LR heats by up to around 1 °C less than

MPI-LR. In the high resolution configuration, the strongest negative differences between AWI and MPI occur over the Atlantic Ocean excluding the very north. There are some areas with positive differences in mid-depth heating between AWI and MPI configurations: Stronger mid-depth heating in AWI compared to MPI configurations occurs over parts of the North Atlantic Ocean. In the low resolution this is the case in mid-latitude and subtropical North Atlantic areas and in the high resolution in the very north of the North Atlantic Ocean.

In conclusion, AWI configurations tend to heat stronger at the surface and weaker in mid-depth compared to MPI configurations with strongest differences occurring in the Southern Ocean, North Atlantic, and North Pacific. The faster surface temperature response in AWI configurations compared to MPI configurations goes along with a faster decline of surface albedo in the 1pctCO2 experiment (Fig. 3b). The stronger albedo decrease in AWI configurations is such that the outgoing longwave radiation (OLR) is reduced compared to the piControl only for a few decades (seen from difference between dotted and solid lines in Fig. 3a); thereafter the OLR is virtually the same as in piControl again, and the imbalance is completely linked to a corresponding change in shortwave radiation: the solid and dotted blue curves almost coincide in case of AWI-MR starting from year 100 onwards. In AWI-LR the solid and dotted curves even cross each other around year 100, meaning that OLR becomes even larger than in piControl afterwards, consistent with the even stronger albedo decrease in AWI-LR. That is not the case in the MPI configurations, where in both cases a reduced OLR still contributes around 30% to the total imbalance at the end of the 1pctCO2 runs.

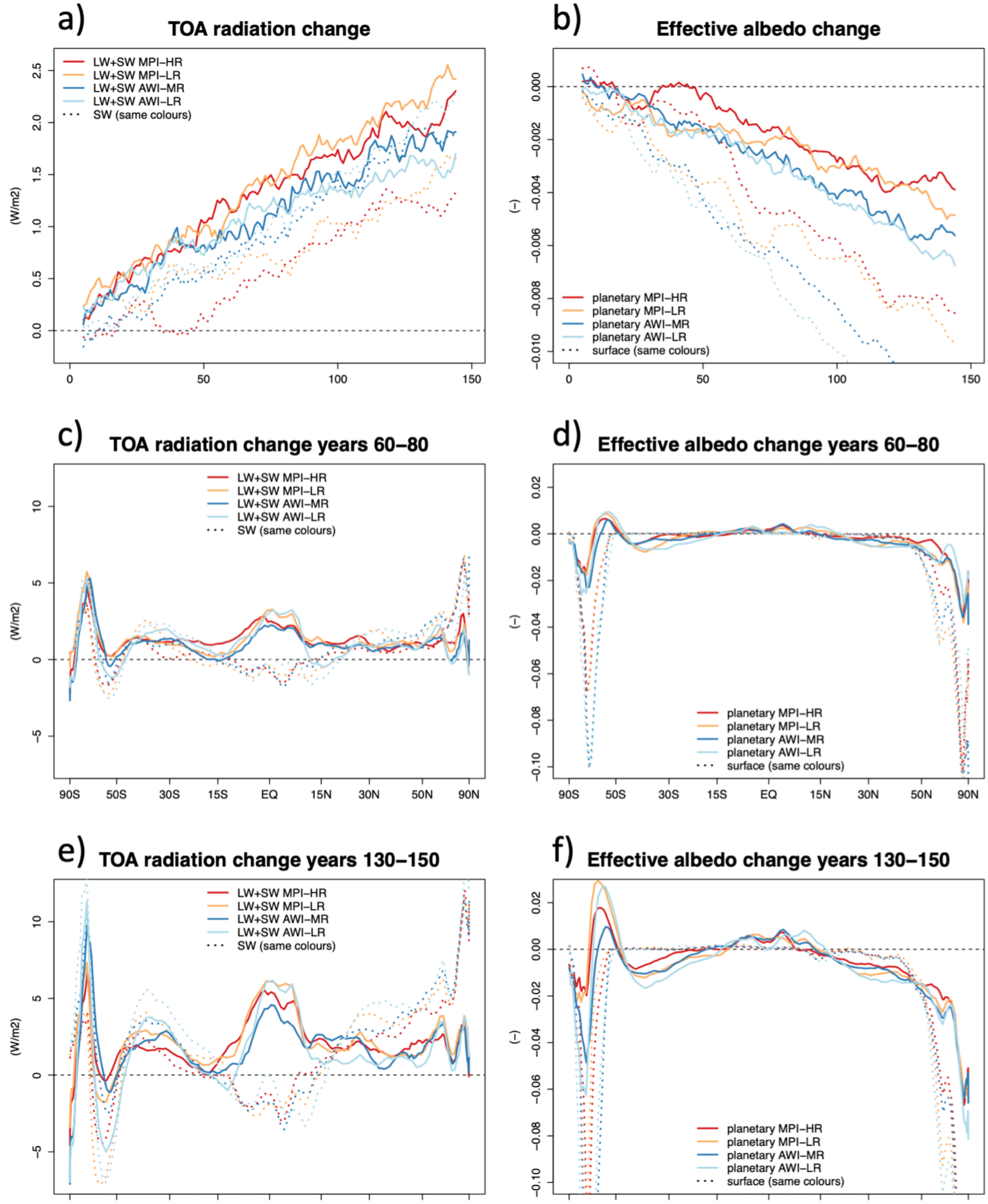


Fig. 3: Annual-mean changes in the energy budget in 1pctCO2 relative to piControl. (a) Top of Atmosphere (TOA) net total (longwave (LW) + shortwave (SW); solid) and net SW (dashed) radiation time series after application of an 11-year running-mean. Positive is downward. (b)

Changes in the effective albedo diagnosed as upwelling divided by downwelling SW radiation at the TOA (planetary; solid) and at the surface (dashed). (c) and (d) same as (a) and (b) but as zonal means averaged over the years 60-80, plotted against sin(latitude) to reflect equal-area global contributions. (e) and (f) same as (c) and (d) but averaged over the years 130-150. Extreme negative values of effective surface albedo changes, reaching about -0.2 in both polar regions, are truncated in (f) to increase the visibility of changes in lower latitudes.

The faster decline of surface albedo in AWI configurations compared to MPI configurations happens in particular in the southern Southern Ocean due to sea ice decline (Fig. A1), which is visible on the surface albedo and to some extent also the planetary albedo (Figs. 3d,f). Slightly further north, the low resolution versions exhibit a stronger planetary albedo increase compared to the high resolution versions, especially in the last 21 years of the 1pctCO2 simulation (Figs. 3d,f). Differences in albedo changes are less pronounced in other latitudes, except for some interesting differences between low and high resolution configurations around 60-70°N. The low resolution configurations have been run with dynamic vegetation, the high resolution configurations not. Therefore, the stronger albedo decrease in those latitudes could be due to a northward extension of vegetation cover in the runs with dynamic vegetation. In fact, in addition to the four configurations discussed in this manuscript, we have run AWI-LR simulations without dynamic vegetation which show a similar albedo response in those latitudes compared to the two high resolution configurations without dynamic vegetation (not shown).

AWI-MR shows a weaker anomalous inner tropical heat uptake through radiation than all other configurations, especially the low resolution configurations (Figs. 3c,e). This is apparently not related to albedo (Figs. 3d,f). At the same time, the low resolution configurations feature the weakest TOA imbalance change around 15°N/15°S.

It is possible to mimic the surface temperature evolution of AOGCMs with a two-layer energy balance model (EBM) containing the atmosphere and upper ocean as one layer and deeper ocean that is still active as an energy reservoir at the considered time scales of 150 years as another layer (Geoffroy et al., 2013). The very deep ocean that is at the considered time scales quasi inactive as an energy sink is excluded in this simple model. AWI configurations show a low effective heat capacity C_0 in the deeper ocean compared to MPI configurations (Table 2). The difference is especially pronounced between the LR configurations and amounts to more than a factor of 2. Similar differences exist for the related parameters: deeper ocean equivalent depth D_0 and slow relaxation time scale τ_s. In contrast, upper ocean parameters C, D, and τ_f are similar between the four different configurations. By applying the coefficients given in Table 2, we can mimic the AOGCM-simulated surface temperature evolution by the EBM (Fig. 4). Roughly consistent with the deeper ocean parameters, AWI-LR shows the strongest surface heating while the two MPI configurations show the weakest surface heating.

Combining these insights with the results from Figure 2, the EBM analysis shows that AWI configurations, especially the LR configuration, simulate a stronger near-surface heating and less vertical heat exchange with the deep ocean (below the two layers of the EBM) compared to MPI configurations.

Table 2: Parameter estimates from the global mean two-layer energy balance model according to Geoffroy et al. (2013) for AWI and MPI configurations. λ is the radiative feedback parameter, γ the heat exchange coefficient, C and C_0 the upper and deeper ocean effective heat capacities, D and D_0 the upper and deeper ocean layer equivalent depths, and τ_f and τ_s the fast and slow relaxation time scales.

Param	AWI-LR	MPI-LR	AWI-MR	MPI-HR
λ [$Wm^{-2}K^{-1}$]	-1.18	-1.42	-1.15	-1.21
γ [$Wm^{-2}K^{-1}$]	0.53	0.71	0.55	0.80
C [$Wyrm^{-2}K^{-1}$]	7.58	8.06	7.1	6.84
C_0 [$Wyrm^{-2}K^{-1}$]	36.05	93.14	50.68	76.42

D [m]	79.38	84.41	74.35	71.2
D_0 [m]	377.5	975.3	530.7	800.3
τ_f [years]	4.34	3.74	3.99	3.35
τ_s [years]	101.4	198.0	131.4	160.6

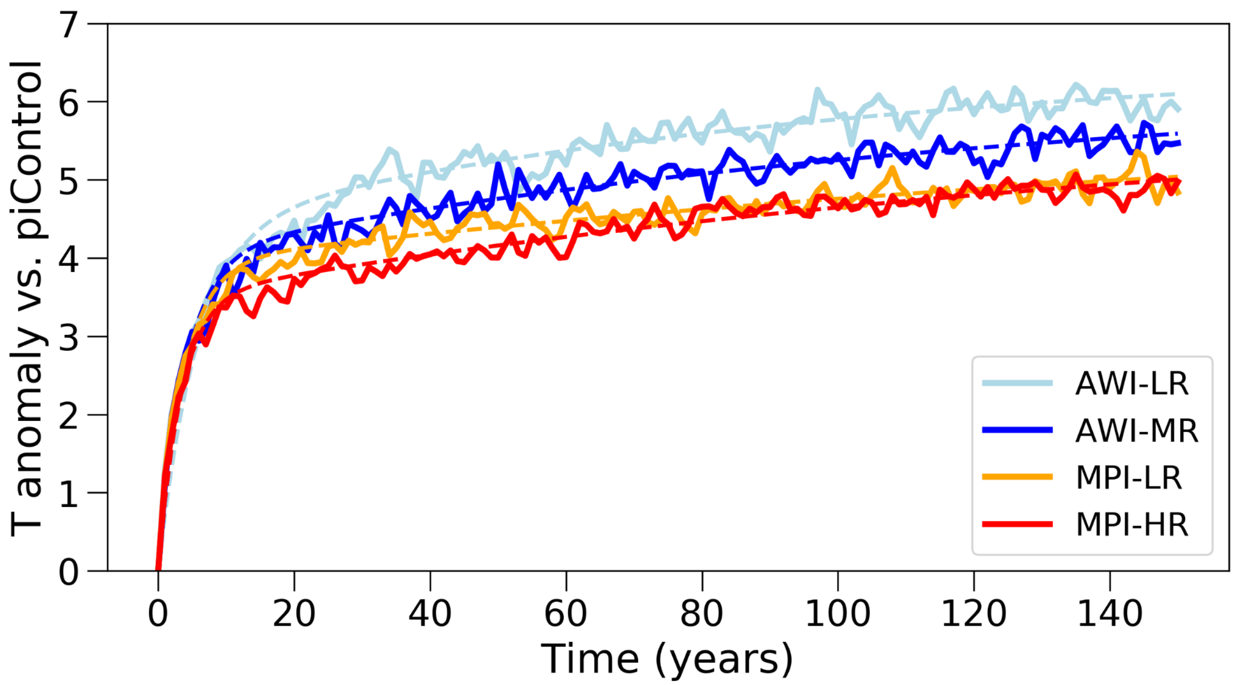


Fig. 4: Surface temperature evolution for AWI-LR (light-blue), MPI-LR (orange), AWI-MR (blue), and MPI-HR (red), from the abrupt-4xCO2 simulations (solid lines) and the upper layer fit to a 2-layer EBM (dashed lines).

However, investigating the total ocean column including the very deep ocean not covered by the 2-layer EBM, it turns out that the vertically integrated ocean heat content increases less in AWI configurations compared to MPI configurations according to the 1pctCO2 simulations in most latitudes, especially in the high latitudes (Fig. 5). Only in limited mid-latitude bands there is a by a factor of up to 2 stronger increase of vertically integrated ocean heat content in AWI-MR compared to MPI-HR configurations (4 Jm-2x1e-9 compared to 2 Jm-2x1e-9 at the end of the simulation). For low resolution configurations mid-latitude heat content increase differences between AWI and MPI models are less pronounced than for high resolution configurations. Over parts of the Southern Ocean the heat content increase is about 3 times less in AWI compared to MPI configurations (3 Jm-2x1e-9 as opposed to 10 Jm-2x1e-9 at the end of the simulation) and over the Arctic Ocean depending on the resolution 1.5 to 3 times less in AWI compared to MPI. This feature is more pronounced in low resolution compared to high resolution configurations. In the tropics, AWI configurations accumulate around 20% less additional energy than MPI configurations. Globally averaged, AWI-LR vertically integrated ocean heat content increases least, followed by AWI-MR, followed by the two MPI configurations (not shown).

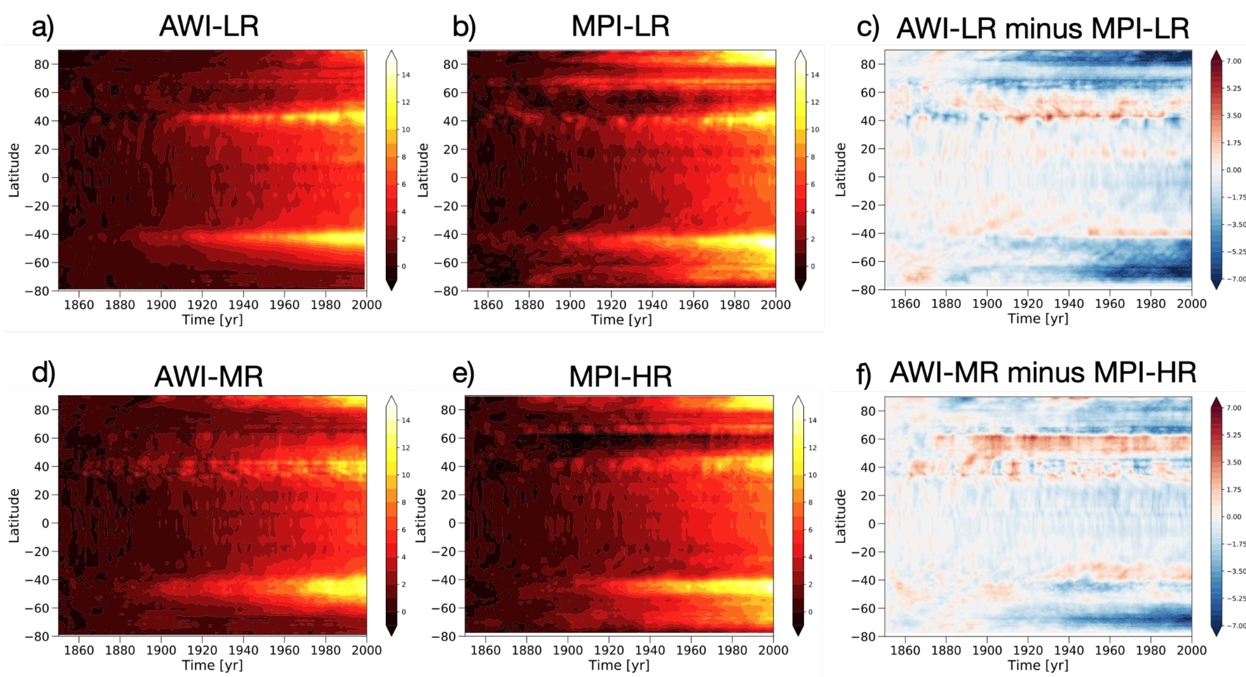


Fig. 5: Temporal evolution of zonal mean heat content density of 1pctCO2 simulation [Jm-2x1e-9] as anomaly with respect to piControl simulation (left and middle columns) for a) AWI-LR, b) MPI-LR , d) AWI-MR, and e) MPI-HR. Right column: differences AWI minus MPI between c) low resolution configurations and f) high resolution configurations .

Differences in mixed-layer depth (Fig. 6) suggest that weaker mixing in AWI compared to MPI configurations may be key for these differences. The strongest difference occurs in the key region Weddell Gyre, but there are also important differences in another key region, the northern North Atlantic. This is consistent with the fact that the strongest differences in ocean heat content increase between AWI and MPI configurations occur in the high latitudes (Fig. 5). Comparing the monthly maximum mixed layer depth with ARGO float measurement data (Holte et al., 2017), it turns out that AWI configurations underestimate mixing in the northern high latitudes while MPI configurations overestimate Southern Ocean mixing, especially in the Weddell Sea. The ARGO float measurement data can neither be compared directly to the piControl nor to the 1pctCO2 simulations at the time of doubling of CO_2 but would be representative for a CO_2 concentration between piControl and doubling of CO_2.

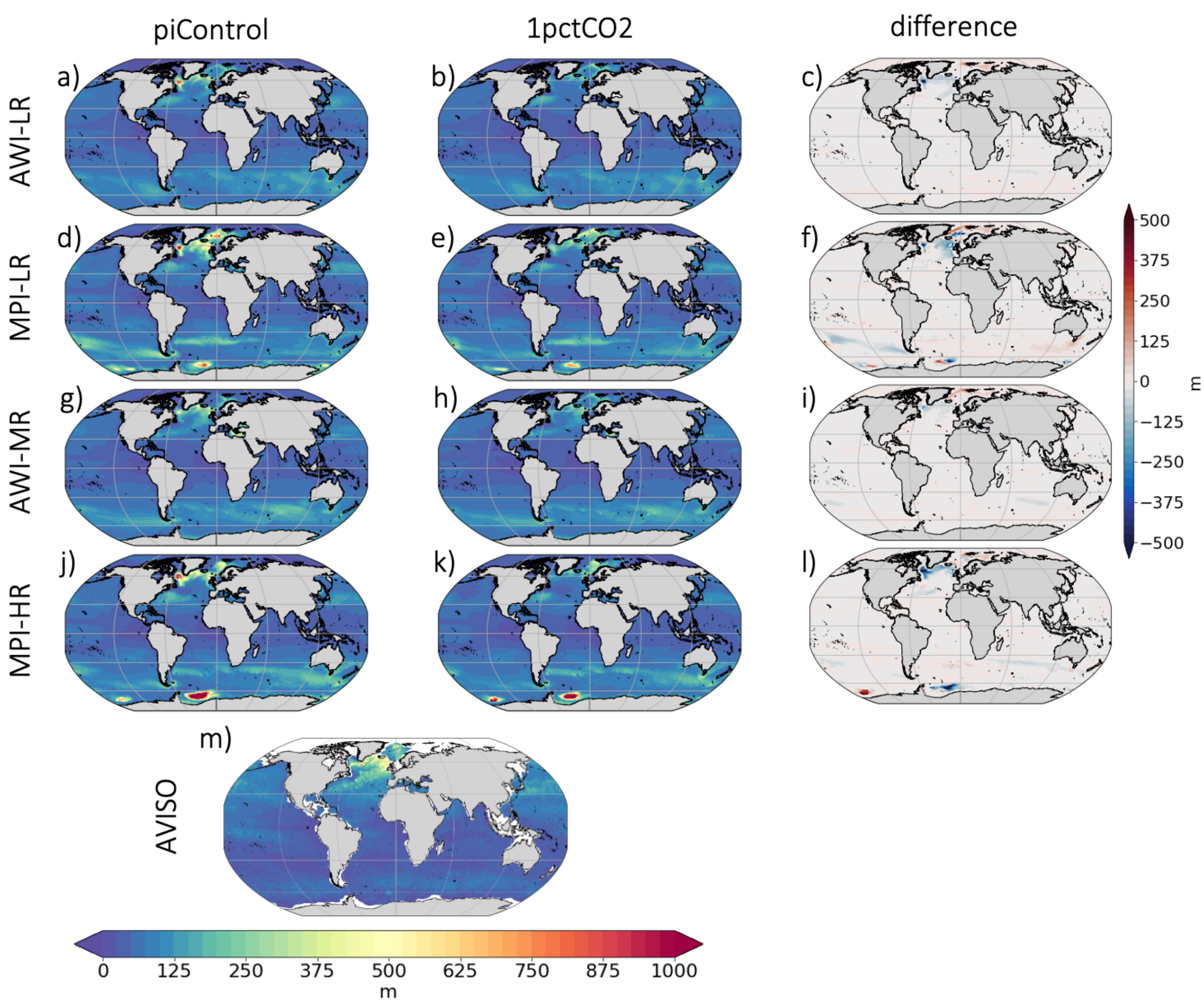
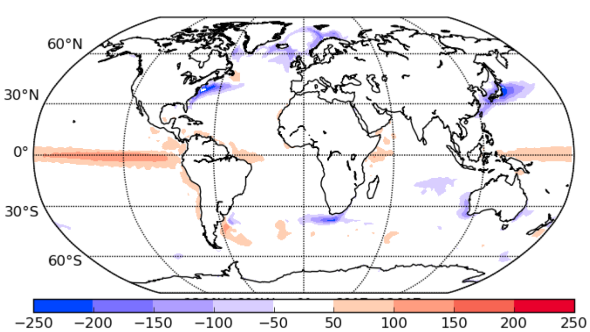


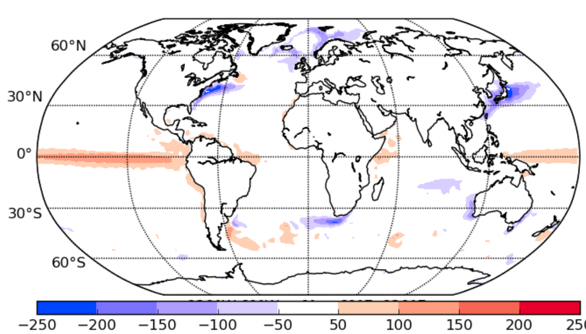
Fig. 6: 21-year mean (years 60-80 of 1pctCO2 simulation and corresponding years of piControl) of monthly maximum mixed layer depth (m; CMIP6 variable omldamax) for (from top to bottom) AWI-LR, AWI-MR, MPI-LR, MPI-HR. piControl (left), 1pctCO2 (middle) and their difference (right). Panel m) shows measurements from ARGO floats (Holte et al., 2017).

The high latitudes, especially key regions such as the Gulf Stream separation, the Kuroshio, the Agulhas, and the Nordic Seas are characterized by heat release at the air-sea interface in all simulations (examples shown in Figs. 7 a,b). In contrast, the tropical Pacific takes up heat at the surface. With increasing greenhouse gas concentrations, none of the four model configurations shows a substantial change in the tropics (Figs. 7 c-f). Strongest affected regions of heat release decrease are the Gulf stream separation and the North Atlantic subpolar gyre. Furthermore, an extended area of heat release decrease exists over the Southern Ocean. This is true for all four model configurations. Differences between AWI and MPI configurations (Fig. 7 g,h) are not as clear-cut as for other investigated parameters; there are small areas of stronger / weaker heat release decrease in AWI compared to MPI configurations close to each other.

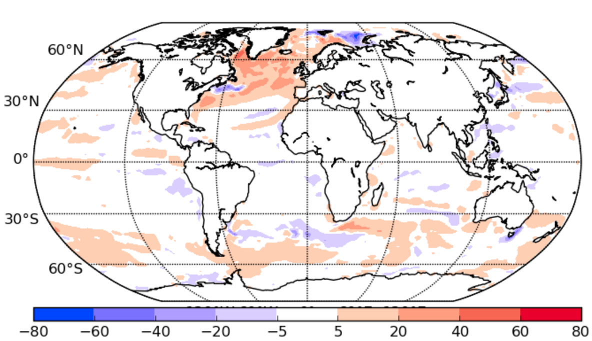
a) AWI-LR piControl



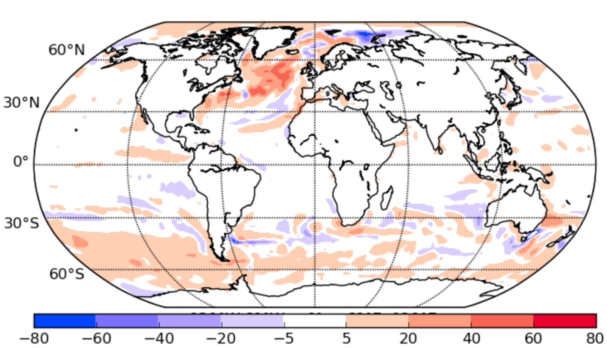
b) AWI-LR 1pctCO2



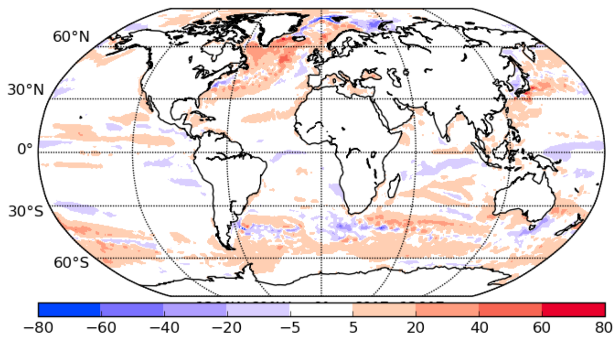
c) AWI-LR response



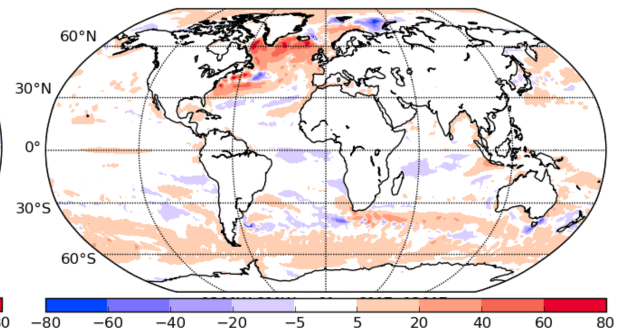
d) MPI-LR response



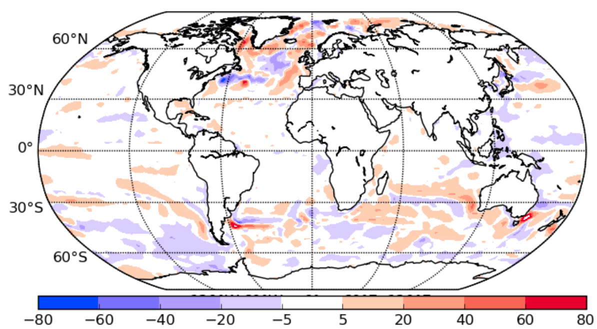
e) AWI-MR response



f) MPI-HR response



g) AWI-LR minus MPI-LR response



h) AWI-MR minus MPI-HR response

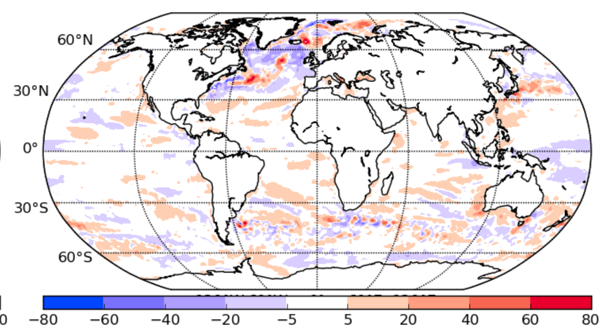


Fig. 7: Ocean heat uptake (W/m^2) computed as the sum of radiative and turbulent heat fluxes at the ocean-atmosphere interface averaged over years 60-80 of the 1pctCO2 simulation and the corresponding 21 years of the piControl simulation. (a) AWI-LR piControl, (b) AWI-LR 1pctCO2, (c) AWI-LR response (1pctCO2 minus piControl), (d) MPI-LR response, (e) AWI-MR response, (f) MPI-HR response, (g) AWI-LR response minus MPI-LR response, (h) AWI-MR response minus MPI-HR response.

The substantially weakening heat release with increasing greenhouse gas concentrations in all four investigated model configurations over the North Atlantic subpolar gyre and surroundings (Figs. 7 c-f) is related to a relatively small ocean surface temperature increase (Figs. 2a, b, g, h) compared to adjacent regions due to a weakening of the Atlantic Meridional Overturning Circulation (AMOC). The AMOC starts at different levels for the different configurations, AWI-LR showing the weakest AMOC and MPI-LR the strongest (Fig. 8). However, here we explain the changes with the AMOC anomalies of the 1pctCO2 experiment with respect to piControl (Fig. 9). The ocean surface temperature increase is particularly small in the MPI configurations (compare Figs. 2b and h for MPI configurations with Figs. 2a and g for AWI configurations); in these configurations there even is a small area over the North Atlantic subpolar gyre in which the ocean surface temperature actually decreases with increasing greenhouse gas concentrations. This phenomenon is referred to as the North Atlantic warming hole (e.g., Chemke et al., 2020; Menary and Wood, 2018). It is likely related to the weakening of the AMOC which is more pronounced in MPI configurations compared to AWI configurations (Figs. 9 and 10). While according to AWI configurations the AMOC at 26°N decreases by around 5 Sv and the northward heat transport by 0.2 PW towards the end of the 140-year period of the 1pctCO2 simulation, stronger decreases of around 8 Sv and 0.3 PW occur, respectively, in the MPI configurations (Figs. 10 and 11).

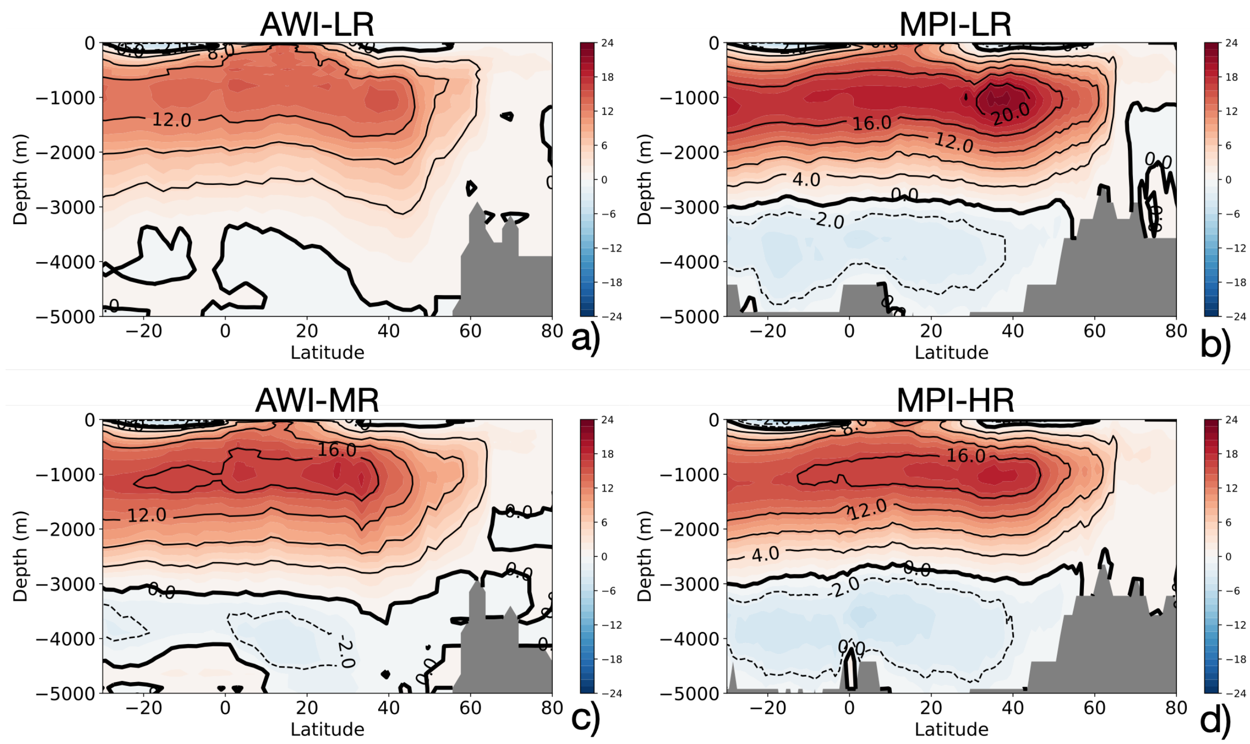


Fig. 8: Latitude-depth plot of the Atlantic meridional overturning circulation (Sv) for the piControl simulations (averaged over the 150 years corresponding to the years of 1pctCO2 simulations) for (a) AWI-LR, (b) MPI-LR, (c) AWI-MR, and (d) MPI-HR.

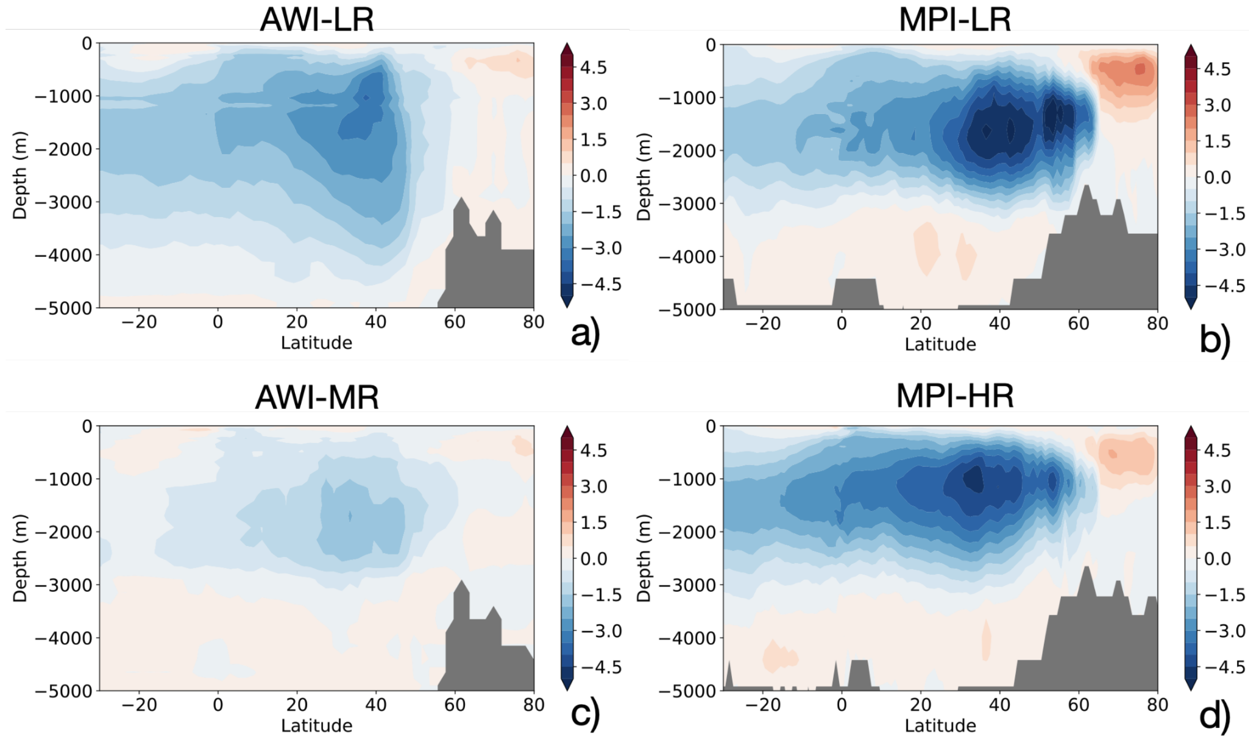


Fig. 9: Latitude-depth difference plot of the Atlantic meridional overturning circulation (Sv) between the 1pctCO2 simulation (averaged over the years 60-80) and the piControl (over the 150 years corresponding to the years of the 1pctCO2 simulation) for (a) AWI-LR, (b) MPI-LR, (c) AWI-MR, and (d) MPI-HR.

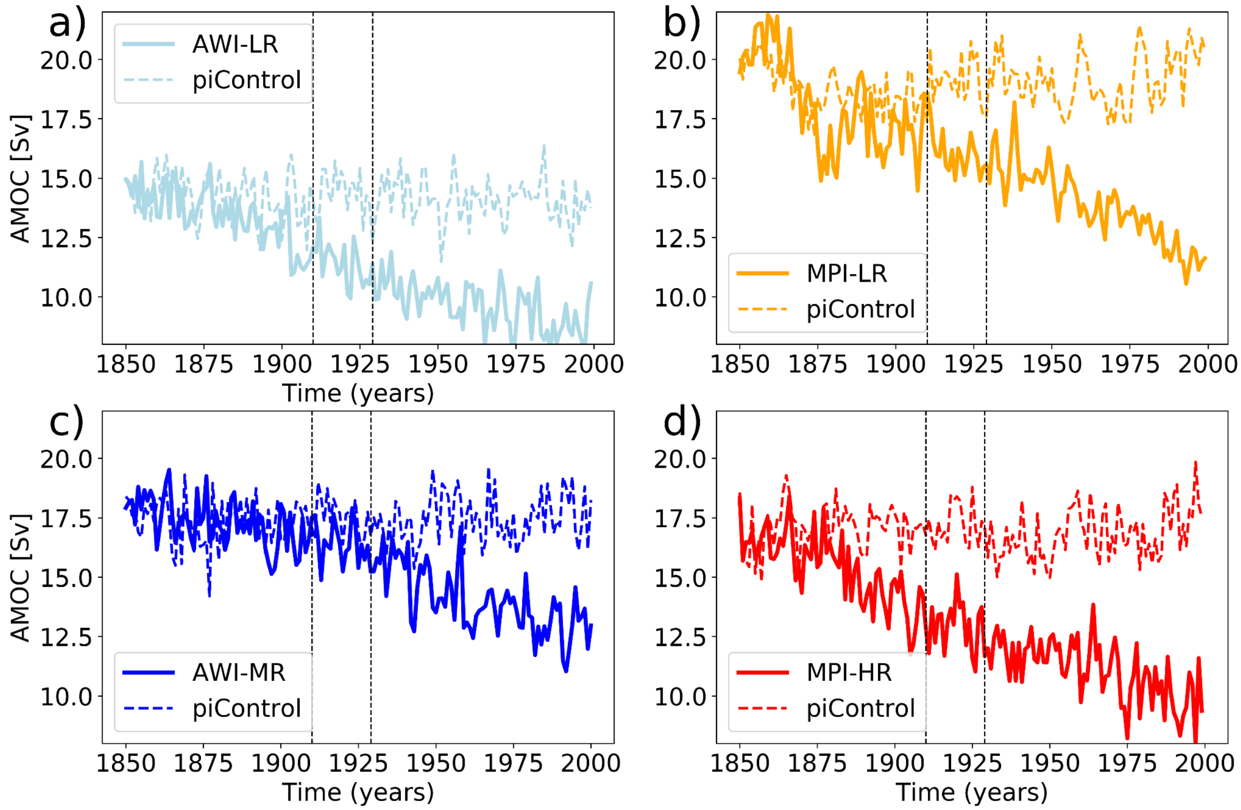


Fig. 10: Time series of the AMOC strength in Sverdrup (1Sv=106m3s-1) at 26N and appr. 1000m depth for the 1pctCO2 experiment (solid lines) and the piControl simulation (dashed lines) for (a) AWI-LR, (b) MPI-LR, (c) AWI-MR, and (d) MPI-HR. Thin black lines denote the period for averaging in Fig. 8.

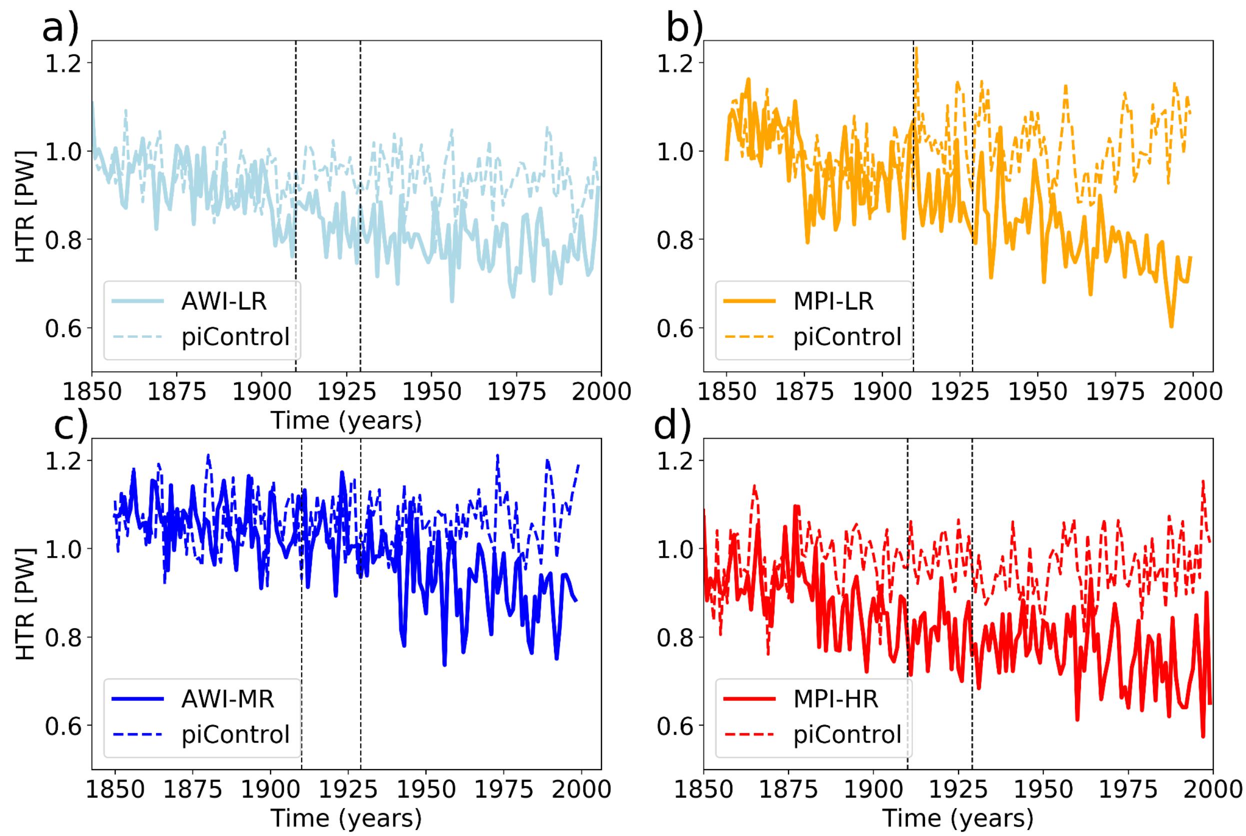


Fig. 11: Time series of the meridional ocean heat transport in Peta Watt (PW) at 26N for the 1pctCO2 experiment (solid lines) and the piControl simulation (dashed lines) for (a) AWI-LR, (b) MPI-LR, (c) AWI-MR, and (d) MPI-HR. Thin black lines denote the period for averaging in Fig. 9.

Discussion and conclusions

This study highlights the importance of the ocean and in particular the intensity of high-latitude ocean mixing for the TCR. In contrast, the ECS is only weakly affected by ocean formulation. The small impact of the ocean formulation on the estimated ECS in our study may even be spurious since the linear Gregory et al. (2004) model applied to only 150 years of data results in inaccuracies of the ECS, or more specifically an underestimation of the ECS (Knutti and Rugenstein, 2015; Rugenstein et al., 2020). In fact, if there is a systematic offset between the estimated ECS and the true ECS depending on how fast the deep ocean takes up heat, the higher estimated ECS for the AWI configurations might be due to differences in mixing (and thus the TCR) rather than true ECS differences. However, true ECS differences cannot be ruled out, one possible reason being differences in cloud feedback induced by the different underlying ocean (Sherwood et al., 2020; Andrews et al., 2012). However, changes in total cloud cover are very similar between the different simulations (not shown). In addition, different warming patterns are known to result in different ECS (Rugenstein et al., 2020 and references therein).

For climate change impact studies relevant for society the TCR is more appropriate as it serves better to predict the development of the climate in the next decades (Knutti et al., 2017). An equilibrium state of the coupled atmosphere-ocean system is reached only after several thousand years (Li et al., 2012) and is therefore not as relevant for political decision-making as the TCR.

In the case of the two considered CMIP6 models in this study we found that the AWI model upper ocean layers heat faster as a response to greenhouse gas forcing compared to the MPI model upper ocean layers while the opposite is true in the deep ocean starting from around 1000 m downward. The faster upper layer ocean warming can be mimicked with a two-layer energy balance model (EBM). AWI configurations, especially AWI-LR, are pumping the heat less vigorously into the deep ocean (1000 m and below) compared to MPI configurations and therefore warm relatively fast in the upper ~1000 m. From the sea surface, more energy is transferred to the atmosphere and then into space as longwave radiation in AWI configurations compared to MPI configurations. As a result, the anthropogenically induced energy imbalance leads to slower heat accumulation in the global ocean column than in MPI configurations. Differences come from the high latitudes while the tropics are hardly affected. We link this to weaker high-latitude ocean mixing, particularly in the key regions Weddell Gyre and Nordic Seas in AWI configurations compared to MPI configurations.

The convection regions of the ocean which are linked to the changes in overturning strength are key for the differences in the heat redistribution of the ocean that we see in our study. Kostov et al. (2014) link the different vertical distribution of ocean heat uptake to differences in the Atlantic meridional overturning circulation (AMOC). Weaker AMOC and weaker decrease of it as response to greenhouse gas forcing have been linked to stronger high latitude surface heating and stronger TCR increase (He et al., 2017; Winton et al., 2014). In our study this relationship is true for the *changes* in, but not necessarily the *initial intensity* of the overturning strength which amounts to 17 to 18 Sv for both the AWI and MPI high-resolution configurations. However, for the low resolution versions both relationships discussed in He et al. (2017) and Winton et al. (2014) do hold since AWI-LR starts with an AMOC of about 15 Sv and MPI-LR with an AMOC of about 20 Sv in pre-industrial conditions. According to AWI configurations the AMOC at 26°N decreases by around 5 Sv and the northward heat transport by 0.2 PW towards the end of the 140-year period of the 1pctCO2 simulation; according to MPI configurations stronger decreases of around 8 Sv and 0.3 PW occur. This leads to a lack of a North Atlantic subpolar gyre warming hole in AWI model configurations and therefore stronger surface warming compared to MPI model configurations. In contrast, MPI model configurations show a clear warming hole over the North Atlantic subpolar gyre with local cooling in this region.

A substantial increase in TCR along with a comparably weak increase in ECS has been found previously by Yokohata et al. (2007) when increasing horizontal ocean resolution; even the values (20% TCR increase along with 10% ECS increase) are very similar to the values reported in our study - with the difference that in our case this does not happen through increasing the horizontal resolution but through replacing the ocean component MPIOM with FESOM. In both cases different high-latitude ocean mixing leads to the differences in TCR and ECS.

In our case the ocean mixed layer depth seems to suffer from different biases in the different models: AWI model configurations tend to show too weak mixing in the northern high latitudes while MPI model configurations are more realistic compared to ARGO floats there (Fig. 6). On the other hand, over the Southern Ocean the MPI model configurations show strong ocean mixing in the Weddell Gyre which is not observed from ARGO floats - here AWI model configurations seem to be more realistic.

There is an interesting aspect of tuning a model towards the observed near-surface temperature increase during the historical simulation through tuning the ECS to a certain value as done for MPI model configurations (Mauritsen and Roeckner, 2020): A higher (lower) ECS model with a stronger (weaker) deep-ocean mixing might yield a similar historical surface temperature evolution, that is, stronger deep-ocean mixing might compensate for a higher ECS when it comes to the TCR. In case one would have wanted to tune the AWI model configurations to match the historical 2 m temperature development, the tuning goal would have been towards a lower ECS compared to MPI model configurations.

Regarding the vertical ocean heat uptake distribution, Gleckler et al. (2016) analysed the ocean heat uptake from pre-industrial times to 2010 in three different layers: 0-700 m, 700-2000 m, below 2000 m. They conclude that the uppermost 700 m have taken around twice as much energy compared to 700-2000 m and around four times as much energy as below 2000 m. For a more recent time period, 1972-2018, von Schuckman et al. (2020) found a similar vertical distribution: the uppermost 700 m have taken almost twice as much energy as the layer between 700-2000 m and almost six times as much as below 2000 m.

A quantitative comparison with our results cannot be made due to the different strength of transient forcing (1% per year in our idealized model simulations; less than half of that in the past few decades). Nevertheless the order of magnitude is similar between both models and the observations. In AWI configurations the uppermost 700 m takes around three times as much energy as the layer between 700-2000 m. For AWI-MR, the uppermost 700 m takes around nine times as much energy as the lowest layer below 2000 m while in AWI-LR the deep ocean heat uptake is negligible and therefore the factor between uppermost 700 m heat uptake and 2000 m and below heat uptake is as large as around 50. Consistent with the stronger ocean mixing in MPI configurations compared to AWI configurations, especially the factor between the two upper layers (uppermost 700 m versus 700-2000 m) is limited to 2 in MPI-HR and even 1.5 in MPI-LR. In MPI-HR the uppermost 700 m takes around 10 times more heat than the layer below 2000 m while in MPI-LR the factor is only 4. Generally, the observations show values between the AWI and the MPI configurations. Having said this, due to the stronger transient forcing in our idealized model simulations compared to the observed greenhouse gas concentration increase, the factors are expected to be higher than in the observations.

Brierley et al. (2009) state that cold ocean states in a pre-industrial climate tend to warm stronger in response to greenhouse gas forcing compared to warm ocean states. While they investigate this for different states of the ocean in one long control run we see this phenomenon in two different versions of the AWI model: AWI-LR is about 1 K colder than AWI-MR in pre-industrial climate and warms stronger as response to greenhouse gas forcing.

The faster albedo response in AWI configurations compared to MPI configurations may be due to the more rapid surface temperature response (due to weaker connection with the deep ocean as seen from EBM results and weaker mixing in AWI configurations compared to MPI configurations) rather than the other way around. This implies that the equilibrium response should be more similar between the models again. Indeed it is the TCR that is substantially different between AWI and MPI model configurations rather than the ECS.

Donohoe et al. (2014) make an attempt to constrain shortwave and longwave radiation changes with satellite observations from the last decades. They come to the conclusion that OLR reduction takes place only for a few decades after greenhouse gas forcing is switched on and that afterwards shortwave radiation feedbacks kick in and lead to enhanced shortwave radiation absorption at the surface. This seems to be more consistent with the AWI simulations (Fig. 3a); however a quantitative comparison is not possible because the considered simulations in our study are idealized 1pctCO2 simulations. This is a faster increasing forcing compared to the observed greenhouse gas concentration increase during the last century.

Even though a quantitative comparison of the AWI and MPI simulations with observations cannot be made due to the pre-industrial and idealized 1pctCO2 rather than historical forcing, it becomes clear from this study that a realistic representation of high-latitude ocean mixing is crucial for constraining the TCR. For future studies historical simulations should be considered in a comparison.

Acknowledgements

The German Climate Computing Centre (DKRZ) granted the computing time and technical support for carrying out the CMIP6 simulations. BMBF provided funding for preparing CMIP6 simulations and supporting the development of postprocessing tools for the publication of the data. We are grateful to Claudia Hinrichs for providing valuable comments on the paper. We thank Oliver Gutjahr for providing ARGO float data. D. Sidorenko was supported by the Helmholtz Climate Initiative REKLIM (Regional Climate Change). H. Goessling was supported by the Federal Ministry of Education and Research of Germany in the framework of SSIP (grant 01LN1701A). The climate model simulation data used in this study are publicly available at the Earth System Grid Federation (ESGF): https://esgf-data.dkrz.de/projects/cmip6-dkrz/. The work described in this paper has received funding from the Helmholtz Association through the project "Advanced Earth System Model Capacity" in the frame of the initiative "Zukunftsthemen." The content of the paper is the sole responsibility of the authors, and it does not represent the opinion of the Helmholtz Association, and the Helmholtz Association is not responsible for any use that might be made of information contained.

References

Andrews, T., Gregory, J. M., Webb, M. J., & Taylor, K. E. (2012). Forcing, feed- backs and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. Geophys. Res. Lett., 39, L09712. https://doi.org/10.1029/2012GL051607

Banks, H. T., & Gregory, J. M. (2006). Mechanisms of ocean heat uptake in a coupled climate model and the implications for tracer based predictions of ocean heat uptake. Geophysical Research Letters, 33, L07608. https://doi.org/10.1029/2005gl025352

Brierley, C. M., Thorpe, A. J., & Collins, M. (2009). An example of the dependence of the transient climate response on the temperature of the modelled climate state. Atmos. Sci. Lett., 10, 23–28. https://doi.org/10.1002/asl.205

Charney, J. G., Arakawa, A., Baker, D. J., Bolin, B., Dickinson, R. E., Goody, R. M., ... & Wunsch, C. I. (1979). Carbon dioxide and climate: a scientific assessment (p. 22). National Academy of Sciences, Washington, DC.

Chemke, R., Zanna, L., & Polvani, L. M. (2020). Identifying a human signal in the North Atlantic warming hole. Nature Communications, 11, 1540. https://doi.org/10.1038/s41467-020-15285-x

Craig, A., Valcke, S., & Coquart, L. (2017). Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3.0, Geoscientific Model Development, 10, 3297-3308, https://doi.org/10.5194/gmd-10-3297-2017

Danek, C., Shi, X., Stepanek, C., Yang, H., Barbi, D., Hegewald, J., & Lohmann, G. (2020). AWI AWI-ESM1.1LR model output prepared for CMIP6 CMIP. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.9301

Donohoe, A., Armour, K. C., Pendergrass, A. G., & Battisti, D. S. (2014). Shortwave and longwave radiative contributions to global warming under increasing CO2. Proceedings of the National Academy of Sciences, 111(47), 16700-16705. https://doi.org/10.1073/pnas.1412190111

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016a). 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.

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., ... & Williams, K. D. (2016b). ESMValTool (v1. 0)–a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP. Geoscientific Model Development, 9(5), 1747-1802. https://doi.org/10.5194/gmd-9-1747-2016.

Gent, P. R., & McWilliams, J. C. (1990). Isopycnal Mixing in Ocean Circulation Models. J. Phys. Oceanogr., 20(1), 150-155. https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2

Geoffroy, O., Saint-Martin, D., Olivié, D. J., Voldoire, A., Bellon, G., & Tytéca, S. (2013). Transient climate response in a two-layer energy-balance model. Part I: Analytical solution and parameter calibration using CMIP5 AOGCM experiments. Journal of Climate, 26(6), 1841-1857. https://doi.org/10.1175/JCLI-D-12-00195.1.

Gettelman, A., Hannay, C., Bacmesiter, J. T., Neale, R. B., Pendergrass, A. G., Danabasoglu, G., Lamarque, J.- F., Fasullo, J. T., Bailey, D. A., & Lawrence, D. M. (2019). High climate sensitivity in the Community Earth System Model Version 2 (CESM2). Geophysical Research Letters, 46, 8329– 8337. https://doi.org/10.1029/2019GL083978.

Gleckler, P. J., Durack, P. J., Stouffer, R. J., Johnson, G. C., & Forest, C. E. (2016). Industrial-era global ocean heat uptake doubles in recent decades. Nature Climate Change, 6(4), 394-398. https://doi.org/10.1038/nclimate2915

Gregory, J. M. (2000). Vertical heat transports in the ocean and their effect on time-dependent climate change. Climate Dynamics, 16(7), 501-515.

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., … & Williams, K. D. (2004). A new method for diagnosing radiative forcing and climate sensitivity. Geophysical Research Letters, 31, L03205. https://doi.org/10.1029/2003GL018747

He, J., Winton, M., Vecchi, G., Jia, L., & Rugenstein, M. (2017). Transient Climate Sensitivity Depends on Base Climate Ocean Circulation, Journal of Climate, 30(4), 1493-1504. https://doi.org/10.1175/JCLI-D-16-0581.1

Hibler III, W. D. (1979). A dynamic thermodynamic sea ice model. Journal of physical oceanography, 9(4), 815-846.

Holte, J., Talley, L. D., Gilson, J., & Roemmich, D. (2017). An Argo mixed layer climatology and database. Geophysical Research Letters, 44(11), 5618-5626. https://doi.org/10.1002/2017GL073426

Johnson, G. C., Lyman, J. M., & Loeb, N. G. (2016). Improving estimates of Earth's energy imbalance, Nat Clim Change, 6, 639–640. https://doi.org/10.1038/nclimate3043

Jungclaus, J., Bittner, M., Wieners, K.-H., Wachsmann, F., Schupfner, M., Legutke, S., … & Roeckner, E. (2019). MPI-M MPI-ESM1.2-HR model output prepared for CMIP6 CMIP piControl. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6674

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

Knutti, R., & Rugenstein, M. A. A. (2015). Feedbacks, climate sensitivity and the limits of linear models. Phil.Trans.R.Soc. A373, 20150146. https://doi.org/10.1098/rsta.2015.0146

Knutti, R., Rugenstein, M. & Hegerl, G. Beyond equilibrium climate sensitivity. Nature Geosci 10, 727–736 (2017). https://doi.org/10.1038/ngeo3017

Kostov, Y., Armour, K. C., & Marshall, J. (2014). Impact of the Atlantic meridional overturning circulation on ocean heat storage and transient climate change. Geophysical Research Letters, 41(6), 2108-2116. https://doi.org/10.1002/2013GL058998

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

Levitus, S., Antonov, J., & Boyer, T. (2005). Warming of the world ocean, 1955–2003. Geophys Res Lett 32. https://doi.org/10.1029/2004GL021592

Li, C., von Storch, JS. & Marotzke, J. (2012). Deep-ocean heat uptake and equilibrium climate response. Clim Dyn 40, 1071–1086. https://doi.org/10.1007/s00382-012-1350-z

Liang, M., Lin, L., Tung, K., Yung, Y. L., & Sun, S. (2013). Transient Climate Response in Coupled Atmospheric–Ocean General Circulation Models. Journal of the Atmospheric Sciences, 70(4), 1291-1296. https://doi.org/10.1175/JAS-D-12-0338.1

Marsland, S. J., Haak, H., Jungclaus, J. H., Latif, M., & Röske, F. (2003). The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates. Ocean Modelling, 5(2), 91-127. https://doi.org/10.1016/S1463-5003(02)00015-X

Matthes, K., Biastoch, A., Wahl, S., Harlaß, J., Martin, T., Brücher, T., ... & Park, W. (2020). The Flexible Ocean and Climate Infrastructure version 1 (FOCI1): mean state and variability. Geoscientific Model Development, 13(6), 2533-2568. https://doi.org/10.5194/gmd-13-2533-2020

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., … & Roeckner, E. (2019). Developments in the MPI- M Earth System Model version 1.2 (MPI- ESM1.2) and its response to increasing CO2. Journal of Advances in Modeling Earth Systems, 11, 998–1038. https://doi.org/10.1029/2018MS001400

Mauritsen, T., & E. Roeckner (2020): Tuning the MPI- ESM1.2 Global Climate Model to Improve the Match With Instrumental Record Warming by Lowering Its Climate Sensitivity. JAMES, https://doi.org/10.1029/2019MS002037

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J-F., Stouffer, R. J., Taylor, K. E., & Schlund, M. (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. Science Advances, 6 (26). https://doi.org/10.1126/sciadv.aba1981

Menary, M. B., & Wood, R. A. (2018). An anatomy of the projected North Atlantic warming hole in CMIP5 models. Climate Dynamics, 50(7- 8), 3063–3080. https://doi.org/10.1007/s00382-017-3793-8

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., & Marotzke, J. (2018). A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). Journal of Advances in Modeling Earth Systems, 10, 1383–1413. https://doi.org/10.1029/2017MS001217

Newsom, E., Zanna, L., Khatiwala, S., & Gregory, J. M. (2020). The influence of warming patterns on passive ocean heat uptake. Geophysical Research Letters, 47 (18), e2020GL088429. https://doi.org/10.1029/2020GL088429

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(2), 173-194. https://doi.org/10.1002/jame.20016

Pacanowski, R. C., & Philander, S. G. H. (1981). Parameterization of vertical mixing in numerical models of tropical oceans. Journal of Physical Oceanography, 11(11), 1443-1451.

Rackow, T., Goessling, H. F., Jung, T., Sidorenko, D., Semmler, T., Barbi, D., & Handorf, D. (2018). Towards multi-resolution global climate modeling with ECHAM6-FESOM. Part II: climate variability. Climate Dynamics, 50(7), 2369-2394. https://doi.org/10.1007/s00382-016-3192-6

Reick, C. H., Raddatz, T., Brovkin, V., & Gayler, V. (2013). Representation of natural and anthropogenic land cover change in MPI- ESM. Journal of Advances in Modeling Earth Systems, 5(3), 459-482. https://doi.org/10.1002/jame.20022

Reick, C. H., Gayler, V., Goll, D., Hagemann, S., Heidkamp, M., Nabel, J. E. M. S., Wilkenskjeld, S. (2021). JSBACH 3 - The land component of the MPI Earth System Model: documentation of version 3.2. Reports on Earth System Science, 240. http://doi.org/10.17617/2.3279802

Rhein, M., Rintoul, S. R., Aoki, S., Campos, E., Chambers, D., Feely, R. A., ... & Talley, L. D. (2013). Observations: ocean. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 255–316, https://doi.org/10.1017/CBO9781107415324.010

Rugenstein, M., Bloch- Johnson, J., Gregory, J., Andrews, T., Mauritsen, T., Li, C., ... & Knutti, R. (2020). Equilibrium climate sensitivity estimated by equilibrating climate models. Geophysical Research Letters, 47(4), e2019GL083898. https://doi.org/10.1029/2019GL083898

Semmler, T., Danilov, S., Rackow, T., Sidorenko, D., Barbi, D., Hegewald, J., Sein, D., Wang, Q., & Jung, T. (2018): AWI AWI-CM 1.1 MR model output prepared for CMIP6 CMIP. Earth System Grid Federation. https://doi.org/10.22033/esgf/cmip6.359

Semmler, T., Danilov, S., Gierz, P., Goessling, H. F., Hegewald, J., Hinrichs, C., ... & Jung, T. (2020). Simulations for CMIP6 with the AWI climate model AWI- CM- 1- 1. Journal of Advances in Modeling Earth Systems, 12(9), e2019MS002009. https://doi.org/10.1029/2019MS002009

Semtner Jr, A. J. (1976). A model for the thermodynamic growth of sea ice in numerical investigations of climate. Journal of Physical Oceanography, 6(3), 379-389.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., ... & Zelinka, M. D. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. Reviews of Geophysics, 58(4), e2019RG000678. https://doi.org/10.1029/2019RG000678

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., ... & Roeckner, E. (2013). Atmospheric component of the MPI- M Earth system model: ECHAM6. Journal of Advances in Modeling Earth Systems, 5(2), 146-172. https://doi.org/10.1002/jame.20015

Stouffer, R.J., S. Manabe and K . Bryan, 1989: Interhemispheric asymmetry in climate response to a gradual increase of atmospheric carbon dioxide. Nature, 342, 660-66. https://doi.org/10.1038/342660a0

Von Schuckmann, K., Cheng, L., Palmer, M. D., Hansen, J., Tassone, C., Aich, V., ... & Wijffels, S. E. (2020). Heat stored in the Earth system: where does the energy go?. Earth System Science Data, 12(3), 2013-2041. https://doi.org/10.5194/essd-12-2013-2020

Wang, Q., Danilov, S., Sidorenko, D., Timmermann, R., Wekerle, C., Wang, X., Jung, T., & Schröter, J. (2014): The Finite Element Sea Ice-Ocean Model (FESOM) v.1.4: formulation of an ocean general circulation model, Geosci. Model Dev., 7, 663–693, https://doi.org/10.5194/gmd-7-663-2014

Wieners, K.-H., Giorgetta, M., Jungclaus, J., Reick, C., Esch, M., Bittner, M., … & Roeckner, E. (2019). MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP piControl. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.6675

Winton, M., Anderson, W. G., Delworth, T. L., Griffies, S. M., Hurlin, W. J., & Rosati, A. (2014). Has coarse ocean resolution biased simulations of transient climate sensitivity? Geophysical Research Letters, 41, 8522– 8529. https://doi.org/10.1002/2014GL061523

Yokohata, T., Emori, S., Nozawa, T., Ogura, T., Okada, N., Suzuki, T., ... & Kimoto, M. (2007). Different transient climate responses of two versions of an atmosphere- ocean coupled general circulation model. Geophysical research letters, 34(2). https://doi.org/10.1029/2006GL027966

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po- Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., & Taylor, K. E. (2020). Causes of higher climate sensitivity in CMIP6 models. Geophysical Research Letters, 47, e2019GL085782. https://doi.org/10.1029/2019GL085782

Figure and Table captions

Fig. 1: Resolution of the four ocean grids, computed as the square-root of grid-cell area. (a) AWI-LR, (b) AWI-MR, (c) MPI-LR, (d) MPI-HR.

Fig. 2: Ocean potential temperature response at 0 m and 1000 m for years 60-80 (21 years centered around doubling of CO_2) from 1pctCO2 simulation compared to piControl. From left to right: AWI, MPI, difference AWI minus MPI. (a), (b), (c): low resolution at 0 m; (d), (e), (f): low resolution at 1000 m; (g), (h), (i): high resolution at 0 m; (j), (k), (l): high resolution at 1000 m.

Fig. 3: Annual-mean changes in the energy budget in 1pctCO2 relative to piControl. (a) Top of Atmosphere (TOA) net total (longwave (LW) + shortwave (SW); solid) and net SW (dashed) radiation time series after application of an 11-year running-mean. Positive is downward. (b) Changes in the effective albedo diagnosed as upwelling divided by downwelling SW radiation at the TOA (planetary; solid) and at the surface (dashed). (c) and (d) same as (a) and (b) but as zonal means averaged over the years 60-80, plotted against sin(latitude) to reflect equal-area global contributions. (e) and (f) same as (c) and (d) but averaged over the years 130-150. Extreme negative values of effective surface albedo changes, reaching about -0.2 in both polar regions, are truncated in (f) to increase the visibility of changes in lower latitudes.

Fig. 4: Surface temperature evolution for AWI-LR (light-blue), MPI-LR (orange), AWI-MR (blue), and MPI-HR (red), from the abrupt-4xCO2 simulations (solid lines) and the upper layer fit to a 2-layer EBM (dashed lines)

Fig. 5: Temporal evolution of zonal mean heat content density of 1pctCO2 simulation [Jm-2x1e-9] as anomaly with respect to piControl simulation (left and middle columns) for a) AWI-LR, b) MPI-LR , d) AWI-MR, and e) MPI-HR. Right column: differences AWI minus MPI between c) low resolution configurations and f) high resolution configurations

Fig. 6: 21-year mean (years 60-80 of 1pctCO2 simulation and corresponding years of piControl) of monthly maximum mixed layer depth (m; CMIP6 variable omldamax) for (from top to bottom) AWI-LR, AWI-MR, MPI-LR, MPI-HR. piControl (left), 1pctCO2 (middle) and their difference (right). Panel m) shows measurements from ARGO floats (Holte et al., 2017).

Fig. 7: Ocean heat uptake (W/m^2) computed as the sum of radiative and turbulent heat fluxes at the ocean-atmosphere interface averaged over years 60-80 of the 1pctCO2 simulation and the corresponding 21 years of the piControl simulation. (a) AWI-LR piControl, (b) AWI-LR 1pctCO2, (c) AWI-LR response (1pctCO2 minus piControl), (d) MPI-LR response, (e) AWI-MR response, (f) MPI-HR response, (g) AWI-LR response minus MPI-LR response, (h) AWI-MR response minus MPI-HR response.

Fig. 8: Latitude-depth plot of the Atlantic meridional overturning circulation (Sv) for the piControl simulations (averaged over the 150 years corresponding to the years of 1pctCO2 simulations) for (a) AWI-LR, (b) MPI-LR, (c) AWI-MR, and (d) MPI-HR.

Fig. 9: Latitude-depth difference plot of the Atlantic meridional overturning circulation (Sv) between the 1pctCO2 simulation (averaged over the years 60-80) and the piControl (over the 150 years corresponding to the years of the 1pctCO2 simulation) for (a) AWI-LR, (b) MPI-LR, (c) AWI-MR, and (d) MPI-HR.

Fig. 10: Time series of the AMOC strength in Sverdrup (1Sv=106m3s-1) at 26N and appr. 1000m depth for the 1pctCO2 experiment (solid lines) and the piControl simulation (dashed lines) for (a) AWI-LR, (b) MPI-LR, (c) AWI-MR, and (d) MPI-HR. Thin black lines denote the period for averaging in Fig. 8.

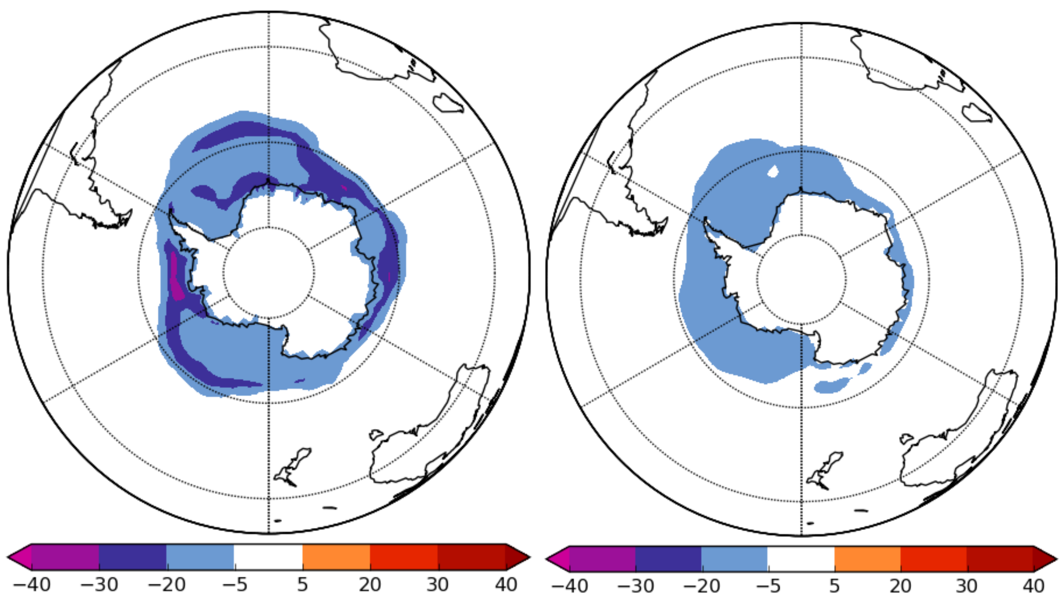
Fig. 11: Time series of the meridional ocean heat transport in Peta Watt (PW) at 26N for the 1pctCO2 experiment (solid lines) and the piControl simulation (dashed lines) for (a) AWI-LR, (b) MPI-LR, (c) AWI-MR, and (d) MPI-HR. Thin black lines denote the period for averaging in Fig. 9.

Table 1: Equilibrium Climate Sensitivity (ECS) and Transient Climate Response (TCR) from the four model configurations, computed according to the definition in the ESMValTools (Eyring et al., 2016b).

Table 2: Parameter estimates from the global mean two-layer energy balance model according to Geoffroy et al. (2013) for AWI and MPI configurations. λ is the radiative feedback parameter, γ the heat exchange coefficient, C and C_0 the upper and deeper ocean effective heat capacities, D and D_0 the upper and deeper ocean layer equivalent depths, and τ_f and τ_s the fast and slow relaxation time scales.

APPENDIX

(a) (b)



(c) (d)

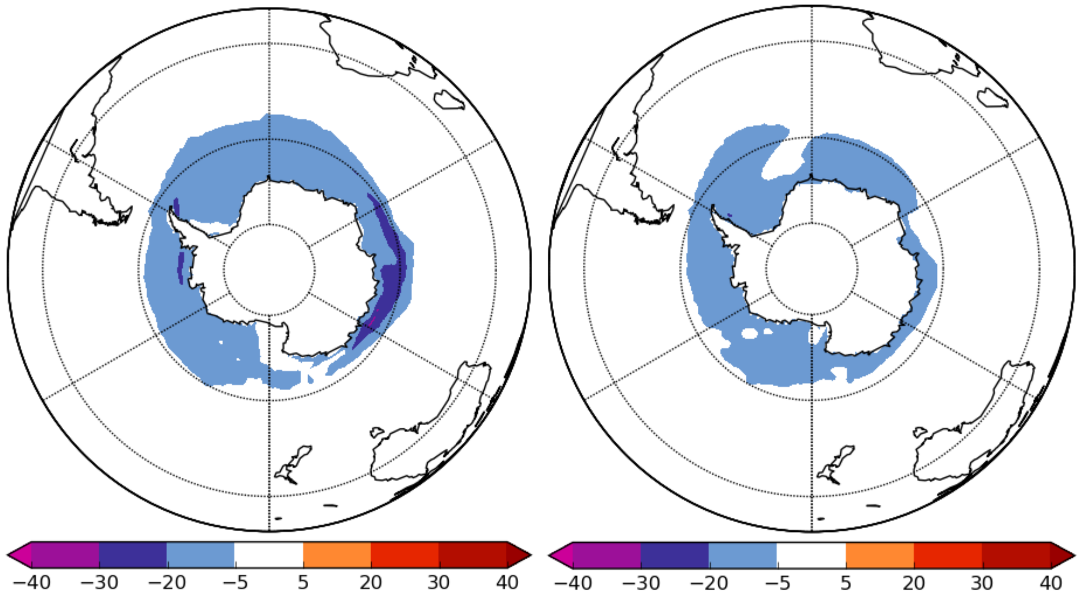


Fig. A1: Antarctic sea ice reduction (%) for years 60-80 of 1pctCO2 simulation compared to corresponding years of piControl simulation for (a) AWI-LR, (b) MPI-LR, (c) AWI-MR, and (d) MPI-HR