Simulations for CMIP6 with the AWI climate model AWI-CM-1-1

Tido Semmler (OrcID: 0000-0002-2254-4901)
Sergey Danilov (OrcID: 0000-0001-8098-182X)
Paul Gierz (OrcID: 0000-0002-4512-087X)
Helge F. Goessling (OrcID: 0000-0001-9018-1383)
Jan Hegewald (OrcID: 0000-0002-3675-0968)
Claudia Hinrichs (OrcID: 0000-0001-5668-9167)
Nikolay Koldunov (OrcID: 0000-0002-3365-8146)
Narges Khosravi (OrcID: 0000-0001-7886-0236)
Longjiang Mu (OrcID: 0000-0001-5668-8025)
Thomas Rackow (OrcID: 0000-0002-5468-575X)
Dmitry V. Sein (OrcID: 0000-0002-1190-3622)
Dmitry Sidorenko (OrcID: 0000-0001-8579-6068)
Qiang Wang (OrcID: 0000-0002-2704-5394)
Thomas Jung (OrcID: 0000-0002-2651-1293)

Key Points

The paper describes contributions of AWI-CM, which employs a sea-ice ocean component formulated on unstructured meshes, to CMIP6

Equilibrium climate sensitivity is similar to average of CMIP5 projections; transient climate response is slightly above average

Response patterns are similar to CMIP5 with more pronounced Arctic sea ice loss and a more stable AMOC compared to other systems

Abstract

The Alfred Wegener Institute Climate Model (AWI-CM) participates for the first time in the Coupled Model Intercomparison Project (CMIP), CMIP6. The sea ice-ocean component, FESOM, runs on an unstructured mesh with horizontal resolutions ranging from 8 to 80 km. FESOM is coupled to the Max-Planck-Institute atmospheric model ECHAM 6.3 at a horizontal resolution of about 100 km. Using objective performance indices, it is shown that AWI-CM performs better than the average of CMIP5 models. AWI-CM shows an equilibrium climate sensitivity of 3.2°C, which is similar to the CMIP5 average, and a transient climate response of 2.1°C which is slightly higher than the CMIP5 average. The negative trend of Arctic sea ice extent in September over the past 30 years is 20-30% weaker in our simulations compared to observations. With the strongest emission scenario, the AMOC decreases by 25% until the end of the century which is less than the CMIP5 average of 40%. Patterns and even magnitude of simulated temperature and precipitation changes at the end of this century compared to present-day climate under the strong emission scenario SSP585 are similar to the multi-model CMIP5 mean. The simulations show a 11°C warming north of the Barents Sea and around 2 to 3°C over most parts of the ocean as well as a wetting of the Arctic, subpolar, tropical and Southern Ocean. Furthermore, in the northern mid-latitudes in boreal summer and autumn as well as in the southern mid-latitudes a more zonal atmospheric flow is projected throughout the year.

Plain Language Summary

The Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (AWI) participates for the first time with a global climate model in the Coupled Model Intercomparison Project 6 (CMIP6). The results of CMIP6 and previous model comparison projects feed into the next assessment report of the Intergovernmental Panel on Climate Change (IPCC). The IPCC assessment reports include information on past and expected climate change in the future and is written for policy- and decision-makers as well as for the general public. The main characteristics of the AWI climate model are described and compared to models from previous intercomparison projects. The projected global warming in AWI-CM is similar to the average warming predicted by climate models in the previous intercomparison project. However, the Arctic sea ice extent declines faster than typical previous estimates. Areas that are wet in present-day climate become wetter, and areas that are dry in present-day climate become drier in the future - consistent with previous climate model simulations. The ocean currents remain rather stable in the AWI climate projections, which leads to a continued warm Gulf stream and therefore an only slightly reduced warming of the North Atlantic and parts of Europe compared to other mid-latitude regions.

Keywords

Global climate model, AWI climate model, Coupled Model Intercomparison Project, climate change, unstructured mesh

1 Introduction

Around 50 institutions worldwide are participating in the current sixth phase of the Coupled Model Intercomparison Project 6 (CMIP6, Eyring et al., 2016). The Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research in Germany, contributes for the first time to CMIP with the novel Finite Element Sea Ice-Ocean Model (FESOM) coupled to the atmosphere model ECHAM6 developed at Max-Planck-Institute (MPI) for Meteorology in Hamburg. The novelty of FESOM lies in the use of global unstructured meshes that only few institutions worldwide are employing at this stage (e.g. Petersen et al., 2019; Korn, 2017). The unstructured-mesh approach allows putting a particular focus on dynamically active regions such as the North Atlantic Current, the Southern Ocean and the tropics while using relatively coarse resolution elsewhere. For the set of “Evaluation and Characterization of Klima” (DECK) and ScenarioMIP experiments a mesh with local refinement of up to 8 km in the North Atlantic Current and the Southern Ocean is used. Coupling the unstructured ocean model FESOM to ECHAM6, which is also used for the MPI-ESM contribution to CMIP6, offers the unique opportunity to investigate the influence of an alternative ocean model formulation on the results which will be exploited in further research.

Many models that participated in CMIP3 and CMIP5 have common descent and share ideas and code with each other (Masson and Knutti 2011, Knutti et al, 2013). This leads to a clustering of results based on model "genealogy" and challenges the assumption of model independence. The ocean part of the AWI-CM is a new unstructured mesh model. It is thus based on a different dynamical core compared to most of the models contributing to CMIP6. Although many parameterisations in FESOM are similar to conventional structured-grid ocean models, and although the ECHAM model has already participated since CMIP3 in the CMIP efforts (Stevens et al., 2013), it can be argued that the use of an unstructured-mesh sea ice-ocean model is an important contribution to the diversity of the CMIP6 ensemble. Large-scale characteristics dominated by the formulation of the atmosphere, such as the equilibrium climate sensitivity, are not expected to be influenced too much by the ocean formulation. In contrast, the ocean has the potential to modulate the transient evolution and regional patterns of the response considerably. This can lead to differences in projected changes of coupled phenomena such as the El Niño-Southern Oscillation (ENSO) as well as sea ice in polar regions.

The aim of this paper is to present the main characteristics of the AWI-CM in the context of the CMIP6 project based on an evaluation of selected atmosphere, ocean, and sea ice parameters for present-day climate as well as for future climate. The evaluation of the unstructured mesh ocean component compared to the traditional mesh ocean component of Max Planck Institute for Meteorology (MPIM) is beyond the scope of this study and will be the topic of a collaborative publication with the MPIM.

In section 2, a brief model description is given along with a summary of the performed DECK and ScenarioMIP simulations, following the CMIP protocol. In section 3 remaining model drift and imbalances are analyzed. Section 4 describes biases in our present-day simulations for some important atmosphere, sea-ice, and ocean variables. The climate change signal is analyzed in detail in section 5. Finally, a discussion of the results and conclusions are presented in sections 6 and 7.

2 Model and simulation description

2.1 Model description

The sea ice-ocean component of AWI-CM is the Finite Element Sea Ice–Ocean Model (FESOM, see Danilov et al. (2015) for the sea ice component and Wang et al. (2014) for the ocean component). It uses unstructured meshes, that allow simulations of ocean and sea ice dynamics with variable grid resolution. This also enables refinement in resolution for areas where small-scale dynamics are prevalent, (e.g. narrow straits and strongly eddying regions, Sein et al., 2016, 2017). Tools have been developed to enable users of FESOM data to perform analysis efficiently (see Appendix A1). Furthermore, selected variables are also available on regular latitude-longitude meshes.

The atmospheric component of AWI-CM is the spectral atmospheric model ECHAM6.3.04p1 from MPIM (Stevens et al., 2013) which is used here without any additional modifications or tuning. This version of ECHAM is also used in the MPIM contribution to CMIP6. Having these two setups, thus will allow future intercomparisons of the coupled systems that share the same atmosphere model but use different sea ice-ocean models.

A more detailed description of the AWI-CM components as well as an evaluation of its mean state and climate variability are provided in Sidorenko et al. (2015) and Rackow et al. (2018), respectively. AWI-CM realistically simulates many aspects of the modern climate, showing an overall performance that is generally better than the most realistic climate models participating in CMIP5.
The CMIP6 version of the code encompasses several changes compared to that described in Sidorenko et al. (2015) and Rackow et al. (2018). The major technical improvement involves the removal of the regular exchange mesh, which in earlier versions was used as an interface between FESOM and the OASIS3-MCT coupler. In the CMIP6 version, the interpolation between unstructured FESOM and structured ECHAM meshes is done by the coupler. Furthermore the coupling between ocean and atmosphere has been sped up remarkably through the use of the parallel support built in OASIS3-MCT.

Updates of physical parameterizations in the ocean sea ice component comprise the inclusion of (1) a salt plume parameterisation (Sidorenko et al. 2018) which improves the simulated sea surface salinity in the Arctic Ocean, (2) modified background diffusivities, as suggested by

Wang et al. (2014) and (3) a K-Profile Parameterization (KPP) for vertical mixing (Large et al. 1994) in the ocean model which has solved shortcomings related to the North Atlantic circulation, pointed out in Sidorenko et al. (2015) and Rackow et al. (2018). Those previous publications were based on simulations on different meshes and with constant rather than transient forcing. This and the fact that these simulations were performed within the CMIP6 framework according to a common protocol calls for documentation of the CMIP6 version of the model and its results presented in this paper.

2.2 CMIP6 simulations

In this paper, the focus is on the DECK and ScenarioMIP simulations, which were defined in the CMIP6 overview paper (Eyring et al., 2016) and are summarized in Table 1. Before starting the 500-year coupled piControl-spinup simulation with constant pre-industrial forcing, a 10-year long ocean-only simulation initialized from the EN4 ocean reanalysis (Good et al., 2013) averaged over 1950–1954 has been performed. In these 10 years of ocean-only simulation the initial adjustment of the ocean state takes place. This pre-spinup helps to ensure a numerically stable adjustment phase of the coupled system. The piControl simulation is a continuation of the piControl-spinup simulation. From the piControl simulation the idealized greenhouse gas forcing simulations 1pctCO2 and abrupt-4xCO2 simulations as well as the historical forcing simulations are branched off at specific years (branch-off point(s), see Table 1). At the end of the historical forcing simulations, that is at the end of the year 2014, the scenario simulations are continued with forcing prescribed from the anthropogenic forcing scenarios. These scenarios are derived from Shared Socioeconomic Pathways (SSP) (Meinshausen et al. 2019).

The idealized and historical forcing simulations have been branched off sufficiently long before the end of the piControl simulation to ensure that every year of the sensitivity simulations (idealized, historical, and scenario simulations) has a corresponding year in the piControl simulation. The climate change signal is always computed following the delta approach (e.g. Lenderink et al., 2007), that is, as the difference between the sensitivity simulation and the corresponding year(s) of the piControl simulation, to account for possible model drift.

Table 1: DECK and ScenarioMIP simulations performed with AWI-CM. The forcing of the ScenarioMIP simulations is described in more detail in Meinshausen et al. (2019).

Experiment	Experiment group	Parent experiment	Years	Branch-off point(s)	Ensemble members
Ocean-only spinup	None	None	10 years	None	1
piControl-spinup	DECK	Ocean-only spinup	500 years	After 10 years	1
piControl	DECK	piControl-spinup	500 years	After 500 years	1
1pctCO2	DECK	piControl	150 years	After 250 years	1
abrupt-4xCO2	DECK	piControl	150 years	After 250 years	1
historical	DECK	piControl	1850-2014	After 150, 175, 200, 225, 250 years	5
ssp126	ScenarioMIP	historical	2015-2100	End of 2014	1
ssp245	ScenarioMIP	historical	2015-2100	End of 2014	1
ssp370	ScenarioMIP	historical	2015-2100	End of 2014	5
ssp585	ScenarioMIP	historical	2015-2100	End of 2014	1

The ECHAM model is run at a spectral resolution of T127L95, where T127 denotes a spectral truncation at total wavenumber 127, which corresponds to about 100 km horizontal resolution in the tropics and higher horizontal (zonal) resolution towards the poles - for example about 25 km in 75° latitude. L95 stands for 95 unevenly spaced model levels with high vertical resolution close to the surface (60 to 300 m in the atmospheric boundary layer) and reaching up to 0.01 hPa corresponding to 80 km (i.e. high-top model version).

In our AWI-CM-1-1-MR CMIP6 contribution (Semmler et al. 2018), the FESOM model is run on a medium-resolution "MR" mesh that follows the mesh design strategy proposed by Sein et al. (2016, 2017) (Fig. 1): The main approach is to locally increase the resolution over areas of high sea surface height (SSH) variability as obtained from satellite data. The horizontal resolution of the mesh varies from 8 km over energetically active areas such as the North Atlantic Current region to 80 km over areas with low SSH variability. The number of surface grid points of the

MR mesh is close to the number of grid points in conventional regular model grids of ¼° resolution. The performance of the “MR-type” meshes in a climate configuration with AWI-CM in comparison to several other FESOM meshes is evaluated in de la Vara et al. (2020), Rackow et al. (2019), Sein et al. (2018).

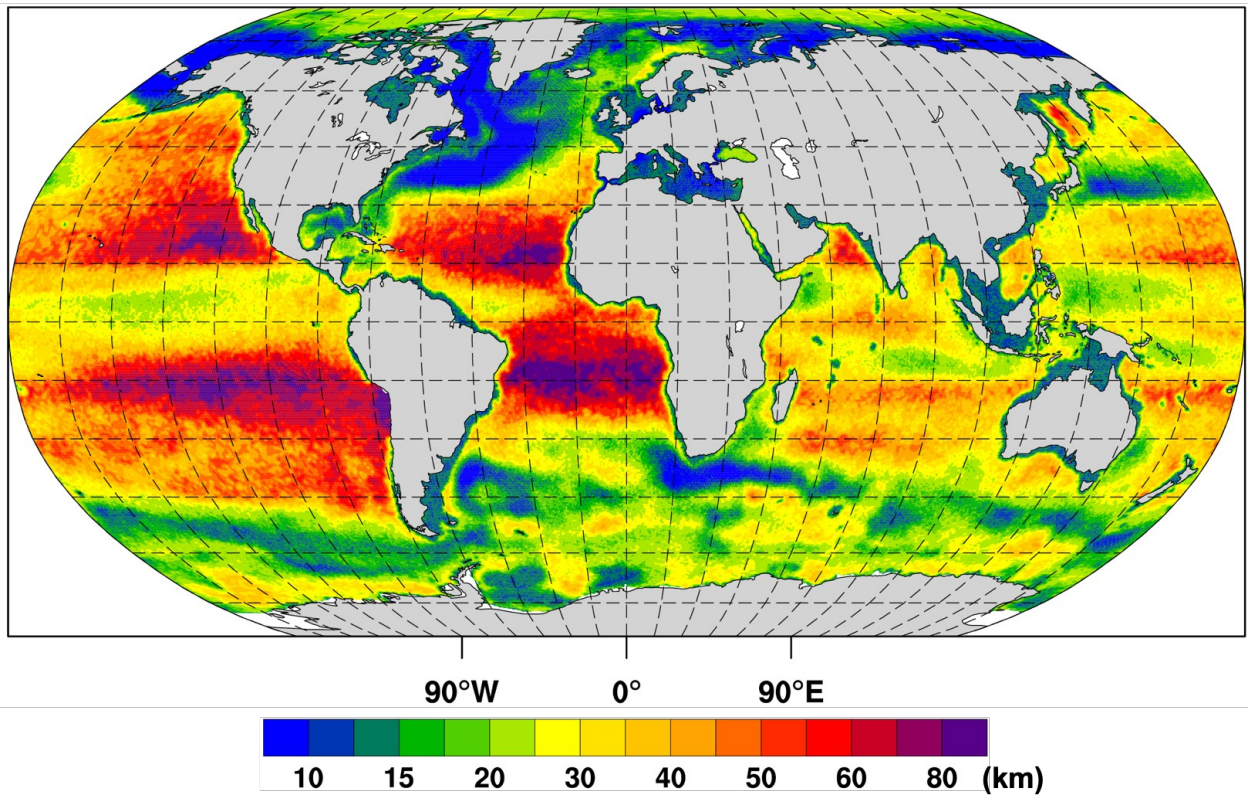


Fig. 1: Spatial resolution (in km) of the FESOM MR grid used in AWI-CM-1-1-MR for the CMIP6 DECK and ScenarioMIP simulations. Resolution is locally increased up to 8 km in regions of high sea surface height (SSH) variability as observed by satellites.

2.3 Cmorization and data publication

CMIP6 is a community project and sharing our experiment result data is an important aspect of the project. To be able to utilize data from other groups, a large set of output data has been defined where the attributes and detailed description for each dataset are put in place as a reference. These are called the CMIP6 CMOR data request (DR) tables (cmip6-cmor-tables 2019). The tables have evolved to a great extent over the past three years. All CMIP6 data are being published through the Earth System Grid Federation (ESGF) (Juckes et al., 2020) - including the AWI-CM CMIP6 data (Semmler et al., 2018).

From a technical point of view, we first chose which variables to generate during our model runs, as re-running the simulations is usually not feasible due to time and resource constraints. We currently produce around 150 variables matching the recent CMIP6 CMOR DR tables. The

model has been optimized to be able to output the data in a very resource efficient manner; this enables us to use less computing resources and complete the simulations more quickly. Due to the many changes of the requirements regarding the output contents and metadata information the CMIP6 CMOR DR tables have undergone, we had to develop a flexible strategy to transform the simulation output into the required publishable format.

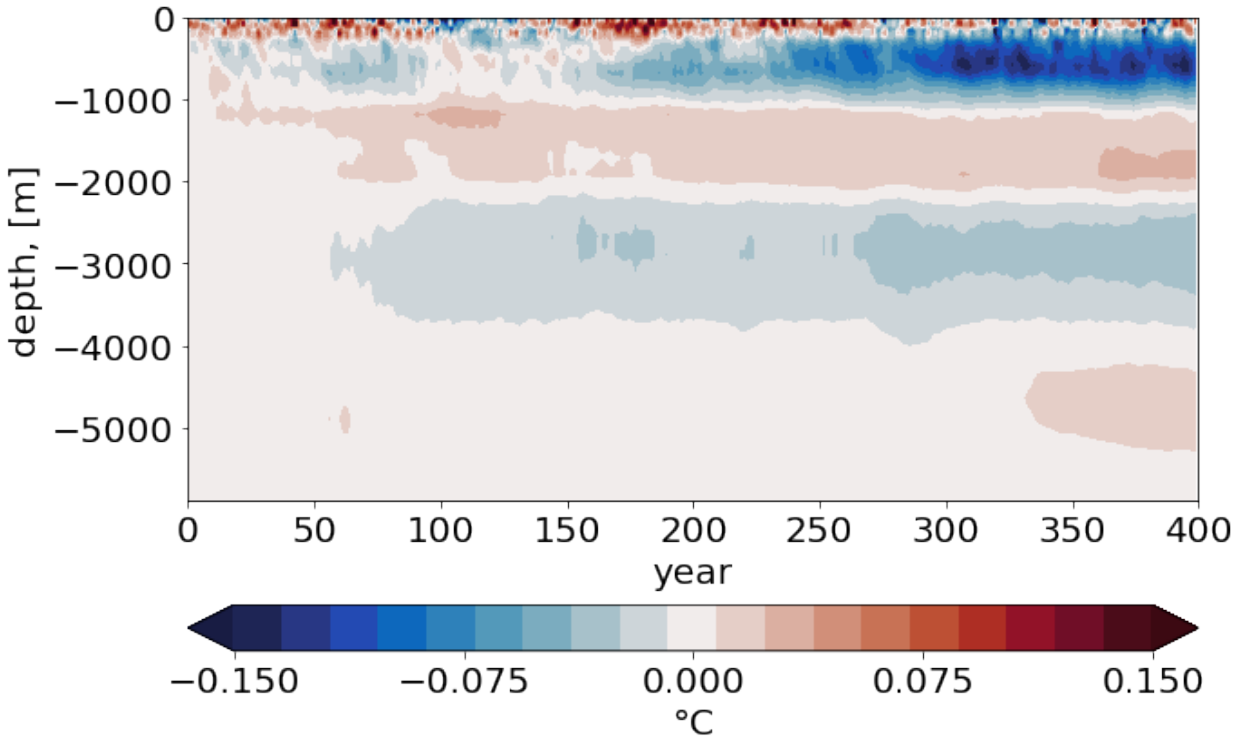
As a result, we now have a post processing software in place, which can directly be fed with the aforementioned DR tables to produce the output accordingly (Hegewald 2019). More details on the procedure and an explanation of how to use unstructured mesh data from the ESGF can be found in the Appendix section A1.

3 Remaining drift and imbalances in the pre-industrial control simulation

In the pre-industrial control simulation AWI-CM is in quasi-equilibrium: The 2 m temperature drift from year 150 to year 400 of the piControl simulation (the time period to which most of the historical, scenario, and idealized CO2 increase experiments need to be compared) amounts to 0.00022 °C/year. Furthermore, sea ice trends are ranging from -0.00069 to -0.00027 million km²/year for the Arctic and from -0.00044 to -0.00026 million km²/year for the Antarctic computed for the years 150 to 400 during March and September, respectively. This suggests that any residual drift of 2 m temperature and sea ice extent in the coupled system is much smaller than the changes anticipated in a warming word.

Fig. 2 shows the Hovmöller diagrams for the global average profiles of oceanic potential temperature and salinity for the last 400 years of the control simulation. The amplitude of the drift is less than 0.15 °C for temperature and 0.05 psu for salinity, respectively, indicating that the system is close to its quasi equilibrium state. The drift in temperature is concentrated at depths of 500, 1,500, 3,000 and 4,500 m, while the drift in salinity happens mainly at depths of 500 and 2,000 m. From inspecting the spatial distribution of the drift (not shown) we conclude that the upper drift zone at 500 m stems primarily from the overall cooling and freshening of the ocean. The drift between 1,500 m and 2,000 m is partly linked to the Mediterranean outflow which spreads into the southern North Atlantic. The simulated outflow is too warm and too salty. At 3,000 m, we observe that the Atlantic and Pacific Oceans become cooler while Indian and Southern Oceans show positive trends in temperature. Simultaneously, salinity in the North Atlantic shows a negative trend at this depth, partly compensating the warming signal there in terms of density. Everywhere else at this depth there is a positive drift in salinity, most pronounced in the Indian Ocean. Finally, the deepest zone of temperature increase at ~4,500 m stems from a warming trend in the Southern Ocean. Although the spatial pattern of non-zero temperature changes implies a small remaining redistribution of heat and salinity, we overall conclude that the system is close to a quasi equilibrium state. Simulated changes in response to greenhouse gas increases are clearly stronger than this residual drift as shown in section 5.4.

(a)



(b)

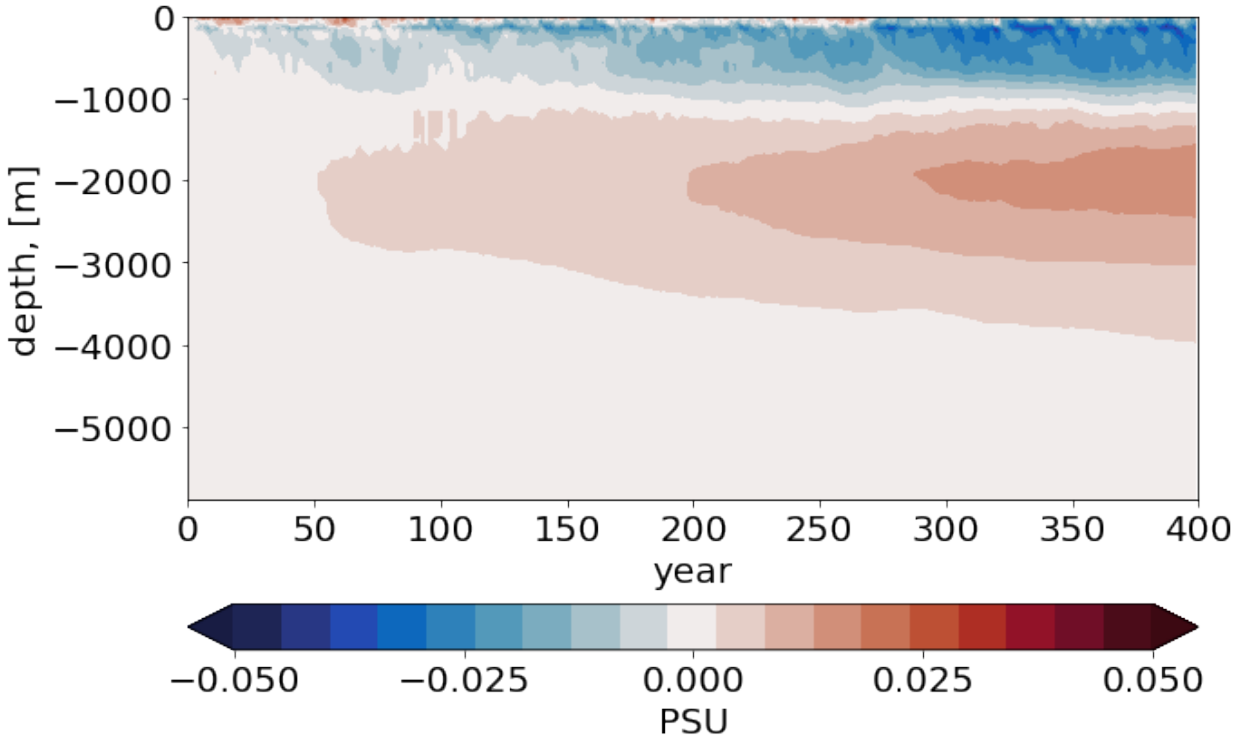


Fig. 2: Vertical profiles of globally averaged ocean temperature and salinity in the last 400 years of the 500 years of piControl simulation (relative to the beginning of this 400 year time period). The last 400 years cover all simulations branched off since the first branch-off point is in year

150 of the piControl simulation, i.e. 350 years before the end of the piControl simulation (see Table 1).

In the last 100 years of the 500 year piControl simulation, which followed the 500 year spinup simulation, there are still imbalances in the top-of-atmosphere (TOA) and net surface radiation. Averaged over these 100 years the TOA radiation imbalance amounts to 0.34 W/m^2 whereas the net surface energy flux consisting of radiation and turbulent heat fluxes amounts to 0.84 W/m^2. Given that changes in the atmospheric energy content on this time scale are much smaller, the discrepancy implies an unphysical atmospheric energy non-conservation of about 0.5 W/m^2. By using the delta approach in the evaluation of the climate change signal as briefly introduced in section 2.2, this non-conservation is cancelled out although one needs to keep in mind the non-linearity of the system.

The gradual energy loss of the ocean over the same time period, diagnosed from changes in the 3D ocean temperature (and sea-ice mass changes), corresponds to a global surface energy flux of -0.01 W/m^2. The deviation from the atmospheric surface flux imbalance by 0.85 W/m^2 cannot be explained by changes in the continental heat content, but points to further deviations from energy conservation that can be related to mismatching grids and coastlines between the model components, inconsistent treatment of temperature, precipitation and runoff (Mauritsen et al. 2012), or other inconsistencies. The atmosphere-related and the surface-related non-conserving energy terms partly compensate each other, resulting in an overall unphysical energy sink of -0.35 W/m^2, and both of them are relatively constant over all simulations (when averaged over decades and longer; not shown).

4 Present-day climate from historical simulations

4.1 Performance indices

In order to objectively characterize the performance of the historical simulations compared to observations, we use modified performance indices by Reichler and Kim (2008) as described in Sidorenko et al. (2015) for the atmosphere and in Rackow et al. (2019) for the ocean. The referenced reanalysis and observation data the model is compared to and a description of the computation of the index are given in the Appendix section A2.

The index measures model error compared to observations relative to the average model error of CMIP5 models. A performance index of 0.5 would indicate an excellent performance as the mean absolute error is halved compared to the CMIP5 models while a performance index of 2 would indicate a doubling of the mean absolute error compared to the CMIP5 models.

Table 2 shows the atmosphere performance indices of the first ensemble member of the historical simulations. For the other four ensemble members of the historical simulations, the results are very similar (not shown). While the performance indices are first computed for each season individually, here, for brevity, we show the annual average. Globally, AWI-CM shows a

good performance in all considered variables and is better than the CMIP5 multi-model mean. Especially Antarctic large-scale circulation and sea ice concentration are very well represented compared to the average of the CMIP5 models. However, there are a few variables such as precipitation, 500 hPa geopotential, and Arctic sea ice which are not in all regions represented better than by the CMIP5 models (only global mean, Arctic and Antarctic shown for brevity). As pointed out in section 5.2.2, the Arctic sea ice *extent* is very well represented both in terms of the mean value and in terms of the trend over the past 3 decades. The sea ice *concentration* is underestimated in boreal summer and autumn in the interior Arctic - see section 4.5 - but the sea ice *extent* is not affected by this since values are generally between 50 and 90% and therefore still well above the threshold of 15%.

Table 2: Atmosphere and sea ice performance indices for years 1985-2014 of the first ensemble member of AWI-CM historical simulations averaged over the four seasons. The set of CMIP5 models consists of CCSM4, MPI-ESM-LR, GFDL-CM3, HadGEM2-ES, and MIROC-ESM.

	t2m	u10m	v10m	toa	pr	tcc	z500	u300	sic	avg
global	0.77	0.81	0.81	0.74	0.99	0.75	0.92	0.72	-	0.81
Arctic	0.92	0.81	0.84	0.69	1.16	0.71	1.24	0.89	1.23	0.94
Antarctic	0.71	0.65	0.84	0.74	1.05	0.70	0.53	0.64	0.50	0.71

From the performance indices for the ocean (Table 3), we can conclude that potential temperature is better represented than in CMIP5. However, this is not the case for salinity. Salinity in the Pacific Ocean as well as in the North Atlantic Ocean deviates more from observations compared to the average of CMIP5 models.

While the performance indices give a quick and objective overview of how a model performs compared to other CMIP5 models, it is necessary to carry out more detailed analysis to investigate if typical errors of climate models such as the Southern Ocean warm bias or the cold bias in the North Atlantic subpolar gyre persist. Regarding the errors in potential temperature and salinity, more analysis is provided in section 4.4.

Table 3: Ocean performance indices for years 1985–2014 of the first ensemble member of AWI-CM historical simulations averaged over the two seasons DJF and JJA. The set of CMIP5 models consists of ACCESS 1.3, BCC-CSM 1.1, BNU-ESM, CanESM2, CCSM4, CMCC-CM, CMCC-CMS, CNRM-CM 5.2, CSIRO-Mk 3.6.0, EC-Earth, MPI-ESM-LR, GFDL-CM3, GISS-E2-H, GISS-E2-R, HadGEM2-ES, IPSL-CM5B, MIROC-ESM, MRI-CGCM3, MRI-ESM1, and NorESM1-ME.

	Potential temperature	Salinity	Average
Global ocean	0.79	1.15	0.97
Southern Ocean	0.96	0.70	0.83
Indian Ocean	0.69	0.94	0.82
North Pacific Ocean	0.92	1.28	1.10
South Pacific Ocean	0.81	1.14	0.98
North Atlantic Ocean	0.70	1.70	1.20
South Atlantic Ocean	0.75	0.79	0.77
Arctic Ocean	0.72	0.90	0.81

4.2 Atmospheric circulation

AWI-CM shows a too strong westerly flow above the Southern Ocean especially in austral summer, indicated by too low mean sea level pressure (MSLP) over the southern high latitudes and too high MSLP over the southern mid-latitudes (Fig. 3). In the Euro-Atlantic sector there is evidence for a southward shift of the jet stream, resulting in a too strong westerly flow over Southern Europe and a too weak westerly flow over Northern Europe in boreal winter and spring. This bias has been found in numerous CMIP5 models (Zappa et al., 2013), and it can be associated with an underestimation of Euro-Atlantic blocking (Jung et al. 2012). Especially in boreal winter, the Aleutian low is too weak. This feature was observed in previous ECHAM6 simulations as well (Stevens et al., 2013). The MSLP biases are not negligible and amount to up to 7 hPa. In the regions they occur, these biases are comparable to the climate change signal indicating that the confidence in projections of circulation changes is low.

The MSLP bias is dependent on the season as shown in Fig. 3 (a) to (d). However, in the following we will also consider the annual mean sea level pressure biases (Fig. 3 (e)) to make our results more comparable with previous studies of CMIP6 models such as Müller et al. (2018) for MPI-ESM (their Fig. 7d), using ECHAM6 as the atmosphere component like AWI-CM. In the annual mean, biases are smaller than 1 hPa over large areas of the tropics, subtropics,

and southern mid-latitudes. . This is consistent with results from Müller et al. (2018). However, differences to Müller et al. (2018) exist over the South Atlantic gyre where AWI-CM shows stronger high pressure biases (2-3 hPa) compared to MPI-ESM (1-2 hPa), and in the south-east Pacific north of West Antarctica where AWI-CM shows negative biases of 1-2 hPa and MPI-ESM positive biases of 2-5 hPa. A thorough comparison between MPI-ESM and AWI-CM, which goes beyond the scope of this study, is planned in collaboration with MPIM.

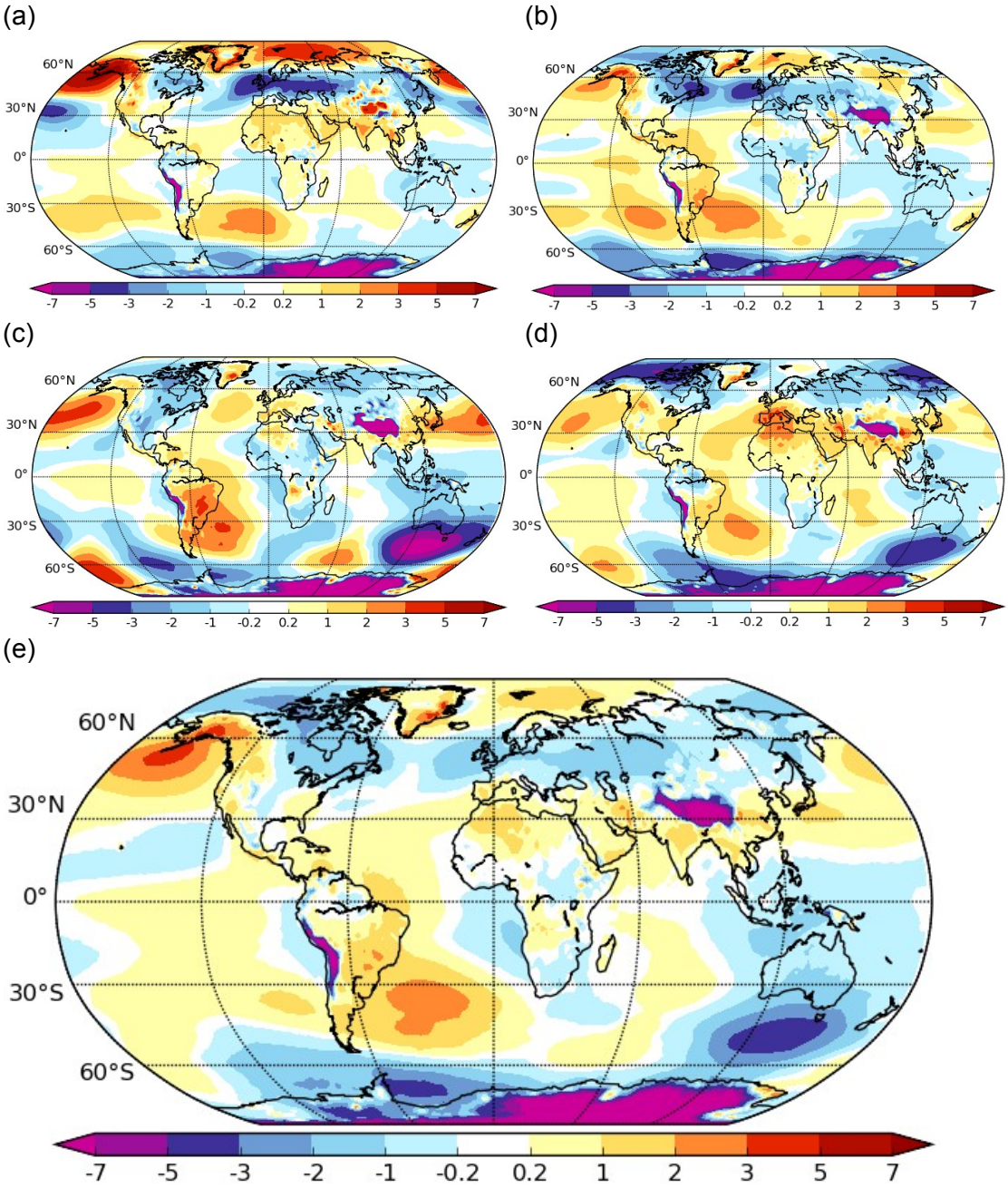


Fig. 3: Mean sea level pressure (MSLP) bias (hPa) as an ensemble mean over the 5 historical realizations for 1985–2014 compared to ERA5 climatology (Copernicus Climate Change Service

(C3S), 2017; Hersbach et al., 2020) from 1985–2014. (a) DJF, (b) MAM, (c) JJA, (d) SON, (e) annual mean

Fig. 4 shows zonal means of temperature and zonal wind biases averaged over the period 1985-2014. In large areas of the troposphere, temperature biases are smaller than 1 °C. Exceptions are the high latitudes with larger positive biases in the north and larger negative biases in the south. Furthermore, in the mid- and high-latitudes there are negative temperature biases of up to around 3 °C in the lower stratosphere around 200 hPa. Not surprisingly, the bias pattern looks very similar to the one from MPIM shown in Müller et al. (2018), their Fig. 9e. The zonal mean zonal wind is generally well represented compared to the ERA5 reanalysis data. Biases are mostly smaller than 2 m/s. Exceptions are the tropical stratosphere, the tropical upper troposphere and the subtropical / mid-latitude stratosphere around 100 hPa and 40-50°N and S. Compared to Müller et al. (2018), their Fig. 9d, biases are generally similar although the subtropical / mid-latitude stratosphere areas of strong biases of more than 2 m/s are smaller in AWI-CM. Furthermore, the negative bias around 60°S extending from 700 to 200 hPa in Müller et al. (2018) does not exist in AWI-CM.

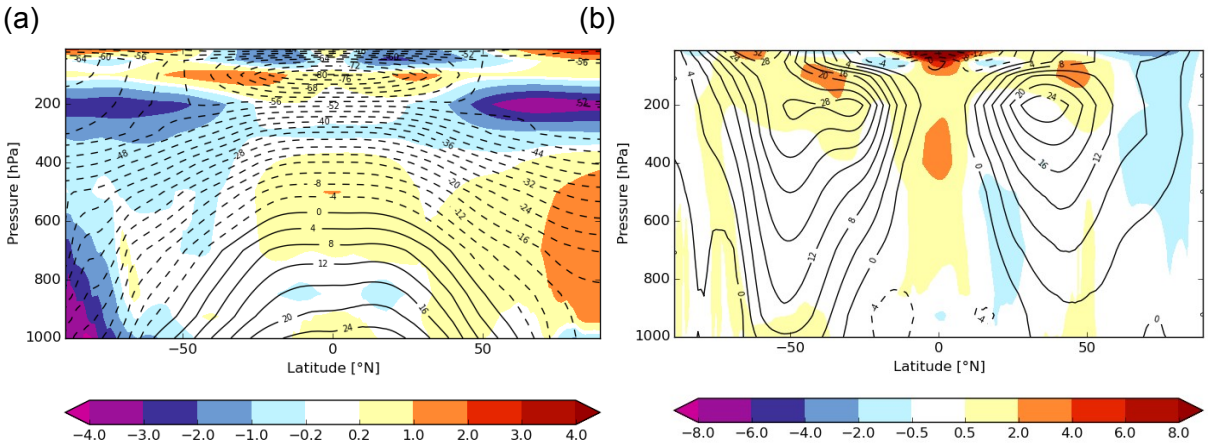


Fig. 4: (a) Annual mean zonal mean temperature (°C) and (b) annual mean zonal mean zonal wind (m/s) as an ensemble mean over the 5 historical realizations for 1985–2014 compared to ERA5 climatology (Copernicus Climate Change Service (C3S), 2017; Hersbach et al., 2020) from 1985–2014. Solid lines represent temperatures at or above 0°C and westerly zonal wind speeds from the ERA5 climatology, dashed lines temperatures below 0°C and easterly wind speeds, and contours represent biases.

4.3 ENSO statistics and phase locking

Sea surface temperature (SST) anomalies in the tropical Pacific associated with the El Niño-Southern Oscillation (ENSO) are of global concern. Since ENSO is the largest signal of interannual variability on Earth (e.g. Timmermann et al., 2018), the realistic simulation of these

SST anomalies, both with respect to their absolute magnitude and temporal behaviour, is crucial for any global climate model.

When comparing area-weighted SST anomalies in the Niño 3.4 box (170°W–120°W, 5°S–5°N) to observations, we find that the five historical ensemble members with AWI-CM show a realistic distribution (Fig. 5). The clear asymmetry between El Niño and La Niña events seen in observations (positive skewness of Niño 3.4 SST anomalies) is also evident in the model. The skewness is 0.15 ± 0.16 (one standard deviation) in the five ensemble members while the observed skewness is 0.36 for 1870-2014. Note that all data have been linearly detrended and the seasonal cycle has been removed before computing the standard deviation. Moreover, the Niño 3.4 index has a significant broad spectral peak, both in the model and in observations for 1870-2014, at a typical period of about 4-7 years when compared to corresponding red-noise processes (Fig. 6). While the distribution of the variance over the frequencies is well reproduced in the model, the total variance is overestimated in all AWI-CM-MR ensemble members (0.75-1.01 K^2 compared to the observed 0.57 K^2).

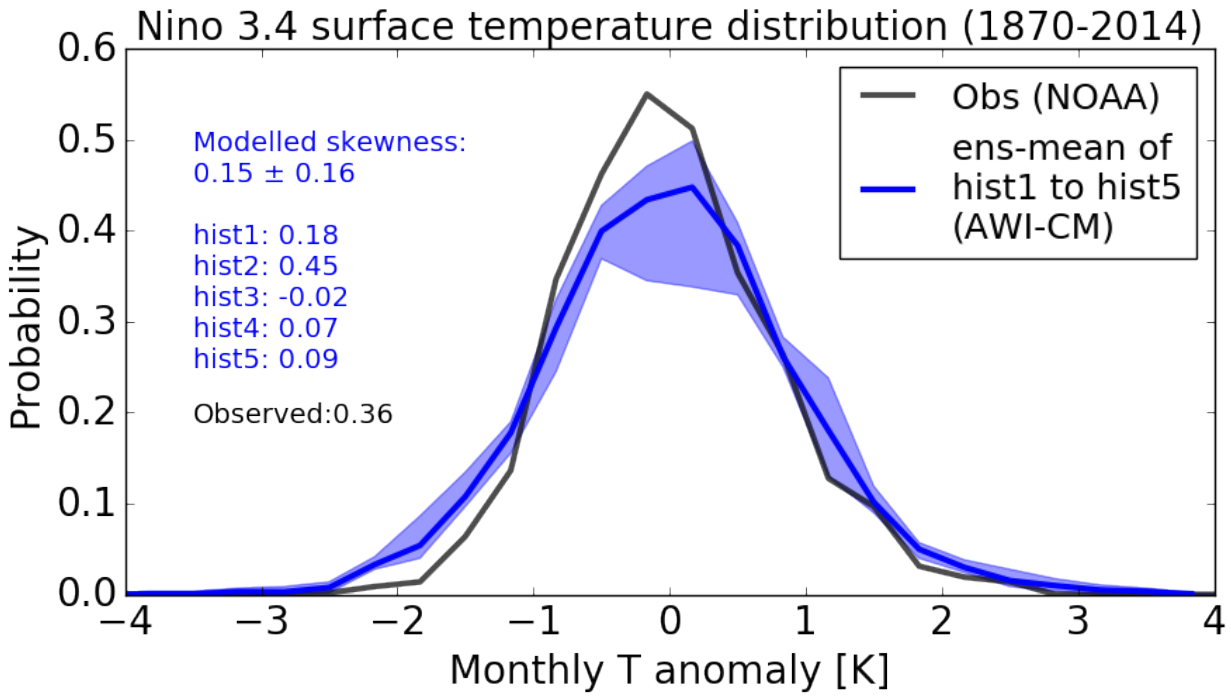


Fig. 5: Probability distribution function (PDF) of sea surface temperature anomalies in the Niño 3.4 region for the historical period 1870–2014. The black line gives the observed Niño3.4 PDF for the period 1870–2014 (Rayner et al., 2003), available for download from the National Oceanic and Atmospheric Administration (NOAA, https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino34/).The ensemble-mean of the five historical members is given in blue; their range (min/max) is shaded in light blue.

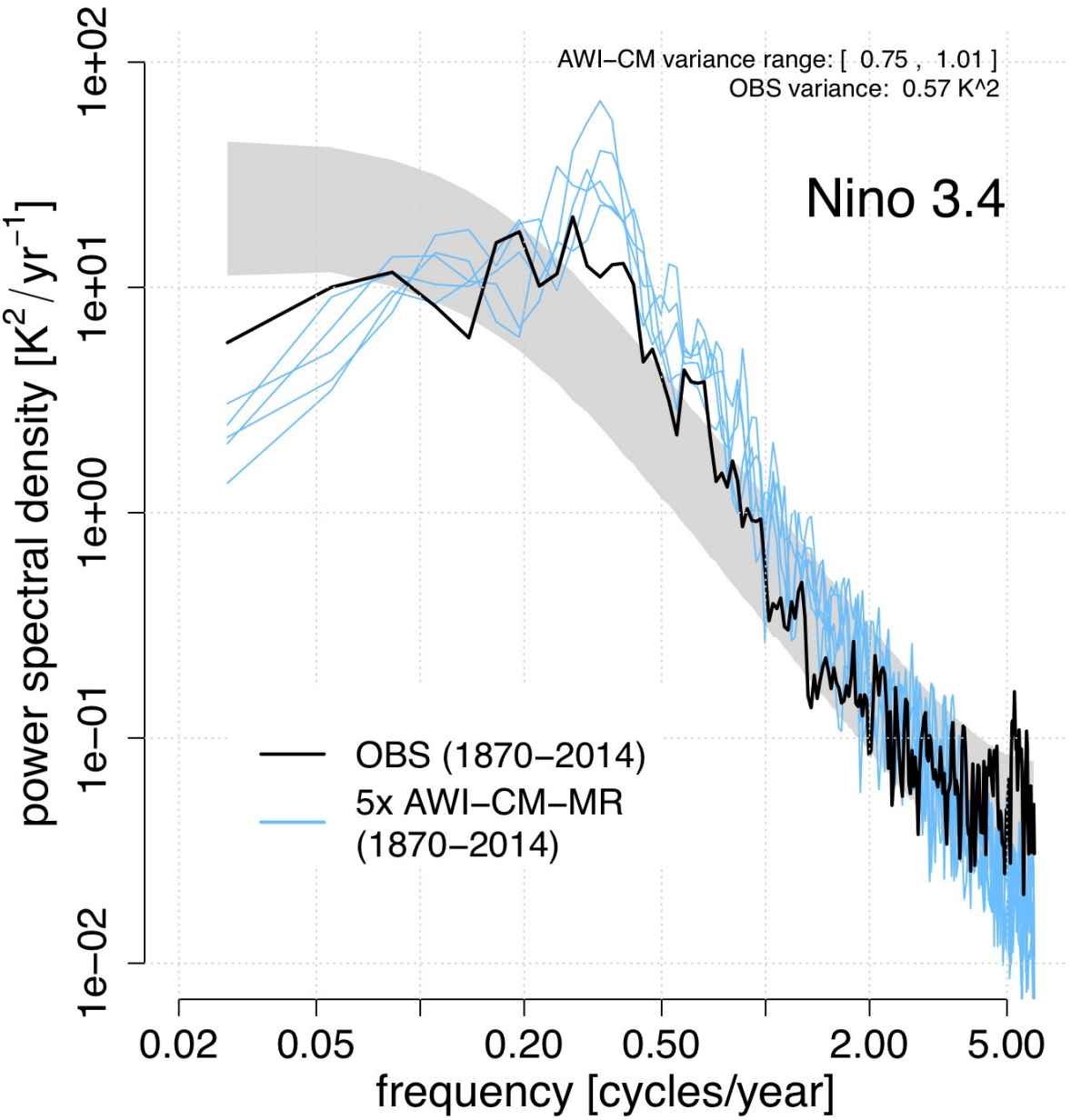


Fig. 6: Power spectral densities (PSDs) of sea surface temperature anomalies in the Niño 3.4 region for the period 1870–2014. The black line gives the observed (OBS) spectrum after Rayner et al. (2003). The five historical ensemble members with AWI-CM are given in blue. Grey shading denotes the 5–95 % confidence interval of an AR(1)-process fitted to OBS, based on a Monte Carlo approach with 10,000 realizations, as detailed by Rackow et al. (2018). The total (integrated) observed Niño 3.4 variance [K^2] is 0.57, for AWI-CM-MR the range is (0.75–1.01).

To assess the temporal behaviour further, we apply a diagnostic that quantifies the seasonal phase locking of Niño 3.4 SST anomalies to the seasonal cycle (Fig. 7). Observed SST variability associated with ENSO, as diagnosed from monthly standard deviation, tends to peak in boreal winter, with a minimum in spring. Especially in boreal winter, the 5 ensemble members

capture the corresponding U-shape and its magnitude relatively well; however, there is a positive bias in spring. A bias of similar magnitude had already been identified in a previous configuration of AWI-CM, using a globally relatively low resolution mesh but with tropical ocean grid refinement at 0.25° (Rackow and Juricke, 2020). The bias appears to be rather sensitive to the applied tropical ocean resolution since the secondary peak in spring is much stronger at a coarser resolution of 1°, using the same atmospheric resolution (see Fig. 6 in Rackow et al., 2014).

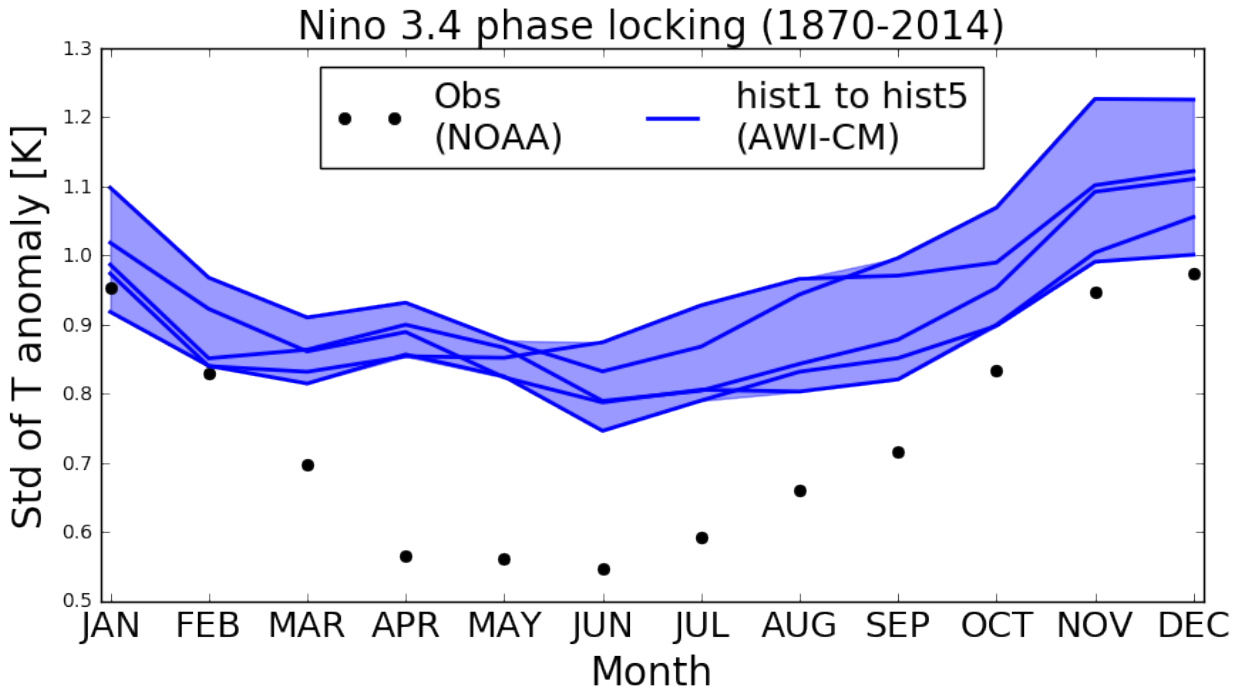


Fig. 7: Seasonal phase locking of sea surface temperature anomalies in the Niño3.4 region for 1870–2014. Black dots give the monthly standard deviation of the observed Niño3.4 index for 1870–2014 (Rayner et al., 2003); blue lines give the standard deviations of the simulated Niño3.4 indices for each of the five historical ensemble members (hist1 to hist5). The range (min/max) spanned by the model results is shaded in light blue. All data have been linearly detrended and the seasonal cycle removed before computing the standard deviation.

4.4 Ocean

Spatial distributions of temperature and salinity biases at the surface and in the interior of the ocean for historical simulations are shown in Fig. 8. Most areas show a small cold bias of 1°C or less in sea surface temperature (SST). There is a pronounced cold bias in the North Atlantic, which is related to the too zonal pathway of the North Atlantic Current; this is a problem that is present in many CMIP climate models (e.g. Wang et al., 2014). If refining the horizontal resolution further to half of the local Rossby radius which for the long time periods of CMIP6 simulations is computationally prohibitive, this bias is largely reduced (Sein et al., 2017). Warm SST biases of up to 1.5°C can be found over the Kuroshio extension, west of Africa as well as very localized close to the equator west of South America, in the Irminger current, over the Labrador Sea, and in the Southern mid-latitudes in the Indian and Atlantic sector. Some of these

biases are typical for climate models such as the cold bias over the North Atlantic subpolar gyre or the warm bias west of Africa. However, over the Southern Ocean no pronounced warm bias is found. This is in stark contrast to MPI-ESM-1.2, the climate model with the same atmospheric component but different ocean model (Müller et al., 2018, their Fig. 2b), and the E3SM model (Golaz et al., 2019, their Fig. 10c), while there are other CMIP models that represent Southern Ocean temperature well.

At the surface, most of the ocean exhibits a fresh bias. In many subtropical and tropical areas this bias amounts to 0.5 to 1 psu; it tends to be weaker in mid-latitude areas. Pronounced but localized salt biases of around 2 psu can be seen close to the coasts of the Eurasian Arctic, in and around the Gulf of Mexico, and in the Bay of Bengal. Smaller salinity biases of up to 0.3 psu can be found over the Southern Ocean and the Pacific warm pool. The general feature of a surface fresh bias in many regions is present also in other climate models such as the E3SM (Golaz et al., 2019), although the regional distribution is not necessarily the same. Features such as the Gulf of Mexico and Bay of Bengal salinity biases are in common with E3SM.

Many CMIP5 models that have coarse ocean resolution suffer from a warm bias at around 1000 m, which is especially strong in the Atlantic Ocean. Increase in the horizontal resolution leads to reduction of this bias, as pointed out by Rackow et al. (2019). Therefore, the performance of AWI-CM in Atlantic temperature is improved compared to other CMIP models. In the AWI-CM simulations discussed in this paper, the magnitude of the warm bias in the South Atlantic is similar to the one over most of the Pacific Ocean (Fig. 8b). The cold and fresh bias in the North Atlantic is related to the outflow and spreading of Mediterranean waters from the Strait of Gibraltar. The reasons for this bias and possible ways to reduce it are discussed in Rackow et al. (2019). The positive temperature and salinity biases in the Indian Ocean are most probably related to excessive supply of warm and salty water from the Red sea. Generally, the biases in temperature and salinity compensate each other in terms of density.

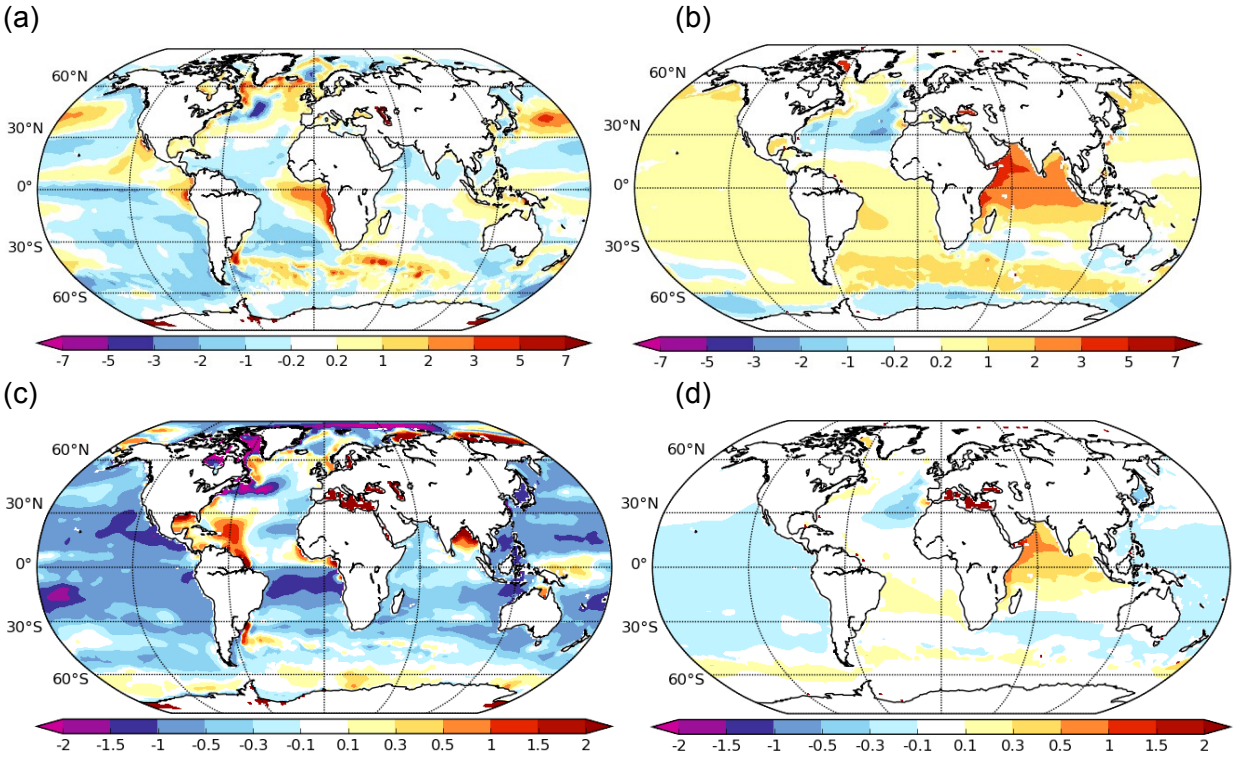


Fig. 8: Bias of the annual mean potential temperature (°C) (a) at the surface, (b) at 1000 m depth averaged over 1985–2014 of the first ensemble member of historical simulations compared to the Polar Science Center Hydrographic Climatology (PHC, updated from Steele et al., 2001). (c) and (d) as (a) and (b) but for salinity (psu).

It turns out that below a depth of about 500 m in the ocean, the mean absolute error of the potential temperature is smaller in AWI-CM than in most of the CMIP5 models (Fig. 9a), while for salinity AWI-CM is comparable to CMIP5 models (Fig. 9c). Compared to the CMIP5 version of MPI-ESM, which shares a slightly older version (6.0 instead of 6.3) of the same atmosphere component and which is run at T63 corresponding to around 200 km horizontal resolution instead of T127 corresponding to around 100 km horizontal resolution, the potential temperature error is smaller in AWI-CM but the salinity error larger. When focusing on the North Atlantic Ocean, potential temperature (Fig. 9b) for which various models show a pronounced warm bias in 1000 to 2000 m (Rackow et al., 2019), AWI-CM performs well. However, for salinity, in the North Atlantic (Fig. 9d) and also in the Pacific (not shown) the mean absolute error is large compared to most of the CMIP5 models including MPI-ESM. Note that Fig. 9 shows results for DJF; for JJA results are very similar below around 300 m.

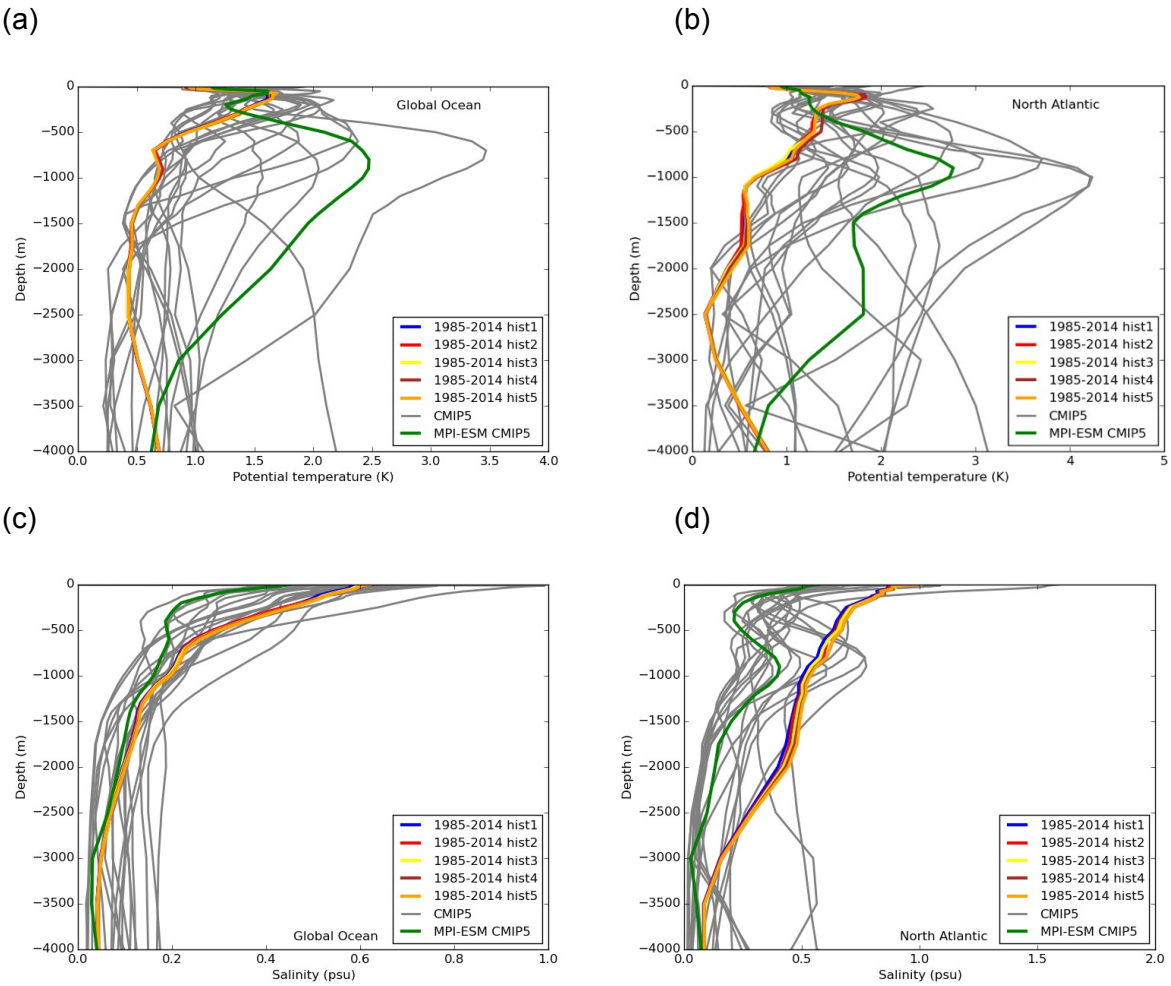


Fig. 9: Profiles of mean absolute error calculated from each grid point for the (a) globe and (b) North Atlantic ocean of potential temperature (°C) for DJF 1985–2014 of the five ensemble members of the historical AWI-CM simulation (in colors) and for DJF 1976–2005 of CMIP5 simulations (in black, each line representing one CMIP5 model, green representing the MPI-ESM CMIP5 model). (c) and (d) as (a) and (b) but for salinity (psu). The reference climatology is the Polar Science Center Hydrographic Climatology (PHC, updated from Steele et al., 2001).

4.5 Sea ice

The general patterns of observed Arctic and Antarctic sea ice concentration are well represented in AWI-CM over the last 30 years of historical simulations (Figs. 10 and 11). Both Arctic and Antarctic sea ice concentration are overestimated in late winter in the marginal ice zones and underestimated in late summer in most areas. This hints to a too pronounced annual cycle of sea ice cover which can also be seen in the sea ice extent as shown in Fig. 15. Nevertheless, Arctic sea ice extent and thickness are remarkably well represented especially over the last few years (Figs. 15 and 16). While late winter Arctic sea ice concentration biases are very similar to MPI-ESM, late winter Antarctic sea ice concentration in MPI-ESM has a

substantial negative bias especially northeast of the Weddell Sea and a slight negative bias in East Antarctic marginal seas (Müller et al., 2018, their Fig. 4) rather than a slight positive bias. This difference is consistent with the reduced Southern Ocean warm bias in AWI-CM compared to MPI-ESM.

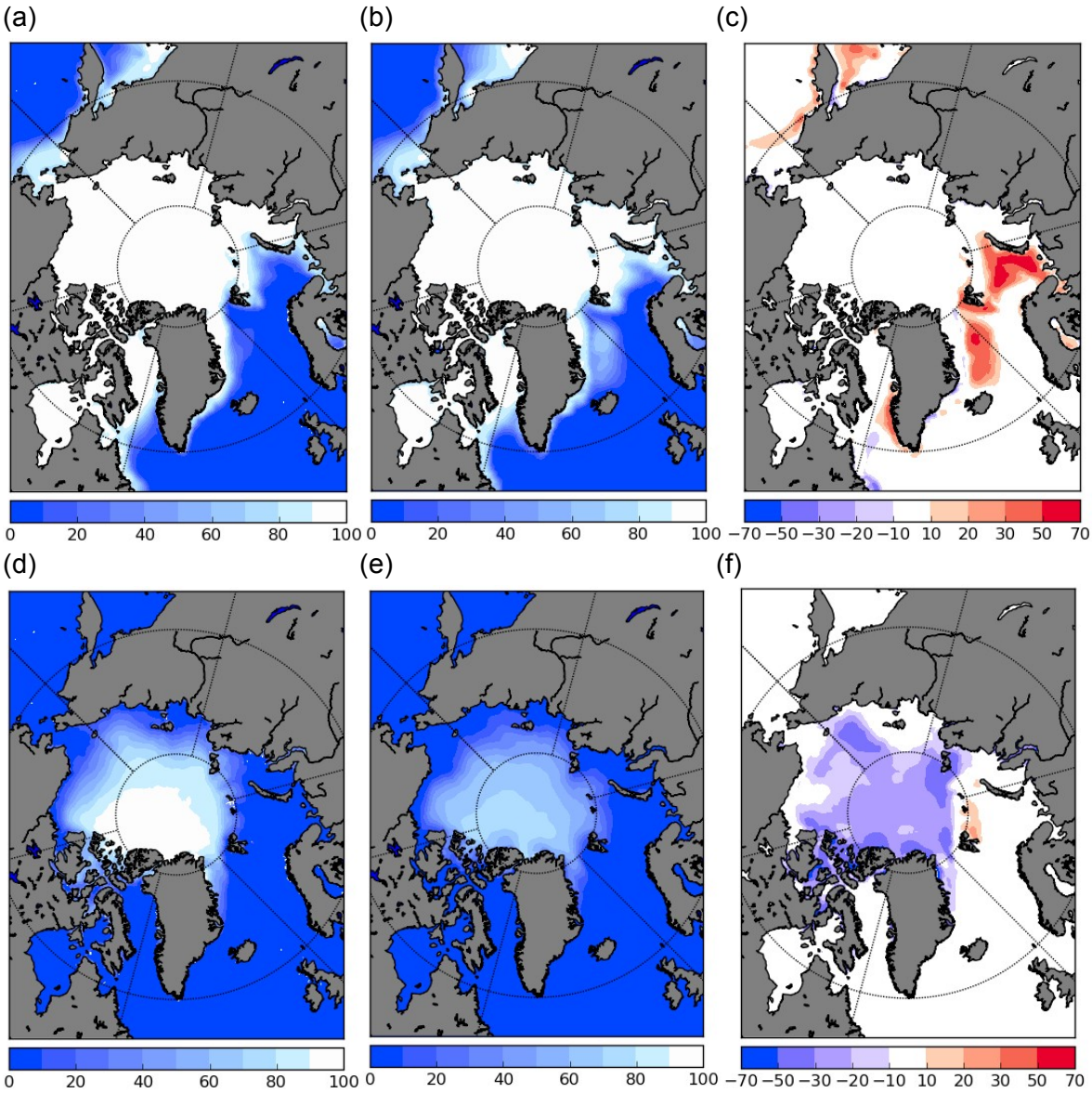


Fig. 10: Arctic sea ice concentration (%) averaged over March 1985 to 2014 from (a) observations from the sea ice portal meereisportal.de (Grosfeld et al., 2016), (b) ensemble mean AWI-CM historical simulations, (c) ensemble mean AWI-CM historical simulation bias. (d) to (f) same as (a) to (c) but for September 1985 to 2014.

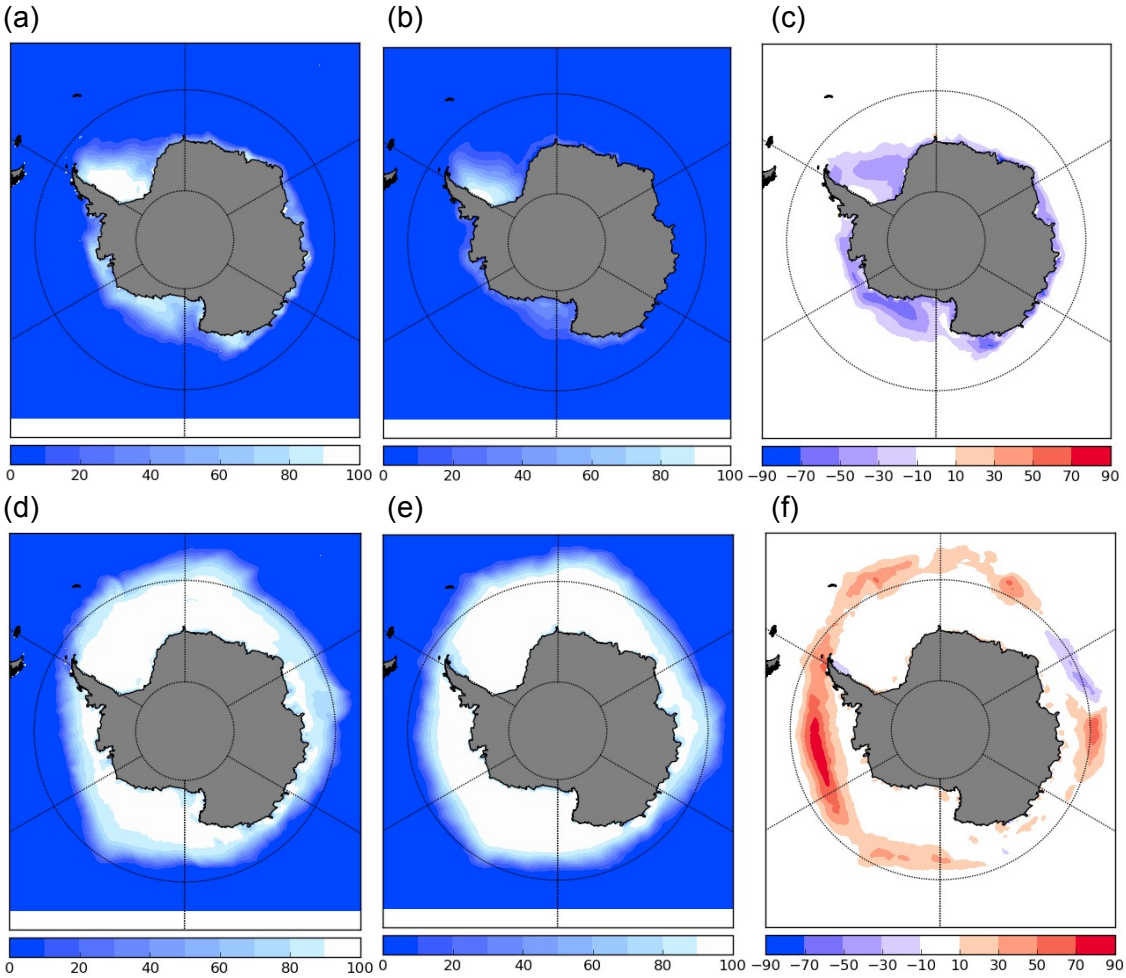


Fig. 11: Same as Fig. 10 but for Antarctica.

5 Climate response

5.1 Climate sensitivity

MPI-ESM has been explicitly tuned to have an equilibrium climate sensitivity (ECS) of 3.0°C (Müller et al., 2018). The ECS is inferred from linear regression of the Top-of-Atmosphere (TOA) imbalance against the temperature response in the 4xCO_2 simulation. AWI-CM uses the same atmospheric component without any extra tuning so that differences both in the ECS and in the transient climate response (TCR), computed as average response over the 30 years around year 70 from the 1pctCO2 simulation, are only due to the different ocean component.

For AWI-CM, the ECS amounts to 3.2°C (Fig. 12, half of the 4xCO_2 equilibrium temperature difference). This is similar to the average over the CMIP5 models (IPCC, 2014) and slightly larger than for the CMIP6 version of MPI-ESM (3.0°C, Müller et al., 2018, Mauritsen et al., 2019, Tokarska et al., 2020). The TCR amounts to 2.1°C, which is slightly stronger than the average over the CMIP5 models (1.8°C, IPCC, 2014) and the CMIP6 version of MPI-ESM (1.7°C,

Tokarska et al., 2020). Note that by considering *changes* in the TOA flux and the global-mean near-surface temperature (delta approach) our estimates for the ECS and the TCR are not affected by the imbalances reported in section 3 (apart from possible non-linear effects).

It seems that AWI-CM absorbs energy in the deep ocean more slowly compared to MPI-ESM. However, this hypothesis needs to be confirmed through a thorough analysis in a joint effort with the Max Planck Institute for Meteorology. Ideally, the ECS should not be affected. However, since the Gregory method to compute ECS is only an approximation, small differences can still occur.

Changes in the energy budget and the role of shortwave feedback in the historical and scenario simulations are detailed in Sect. 5.5.

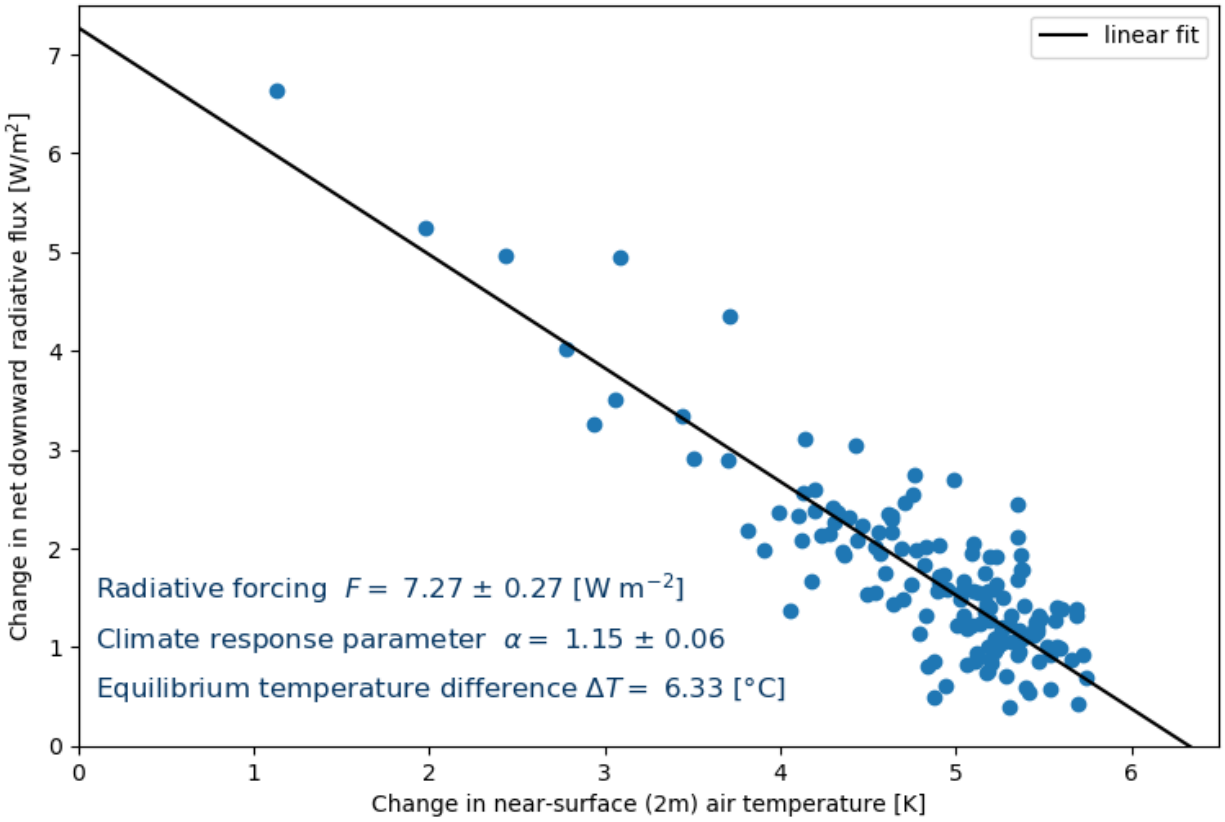


Fig. 12: Gregory plot (Gregory et al., 2004) from the abrupt-4xCO_2 compared to the piControl simulation. For each year, the near-surface (2m) air temperature change between abrupt-4xCO_2 and piControl simulation is plotted against the change in net downward radiative flux between the two simulations. The more the abrupt-4xCO_2 simulation approaches the equilibrium, the smaller the difference in net downward radiative flux compared to the reference simulation becomes. To compute the initial radiative forcing, a regression is built from all data points and extrapolated to a change in near-surface air temperature of 0°C. α is the climate response parameter, indicating the strength of the climate system's net feedback (radiative feedback divided by temperature response). To compute the equilibrium temperature

difference, the regression is extrapolated to the equilibrium (difference in net shortwave radiation = 0).

5.2 Surface response

5.2.1 2 m temperature and precipitation

The evolution of the global and hemispheric mean temperature at 2 m above the surface in the piControl, historical, and scenario simulations is shown in Fig. 13. The piControl simulation shows no discernible trend in temperature, as expected. When considering the anthropogenic forcing, the historical simulations show a warming of 1.1 ± 0.1°C in 2005–2014 compared to 1891–1900 while for the observations the warming amounts to 0.9°C over the same period. Both in the observations and in the historical simulations the Northern (Southern) Hemisphere warming is 0.2°C higher (lower) than the global average. The more pronounced warming over the Northern Hemisphere compared to the Southern Hemisphere is partly due to the higher land partition in the Northern Hemisphere compared to the Southern Hemisphere.

Until the end of the 21st century, the global mean temperature rises by approximately 4°C from today under the strongest emission scenario SSP585. Over the Northern Hemisphere this warming is more pronounced and amounts to approximately 5°C; over the Southern Hemisphere the warming is limited to approximately 3°C. For the weakest emission scenario, SSP126, the global mean warming remains just below 2°C compared to pre-industrial conditions. The SSP126 scenario has been designed to keep global warming below 2°C – a condition that seems to be fulfilled in our simulations. Overall, the temperature increase in the AWI-CM simulations for both the strongest and the weakest emission scenario agrees with the CMIP5 multi-model ensemble mean (IPCC, 2014, their Fig. SPM.6a) and appears to be slightly stronger compared to the CMIP6 version of MPI-ESM - which is expected due to the slightly higher transient climate response in AWI-CM compared to MPI-ESM.

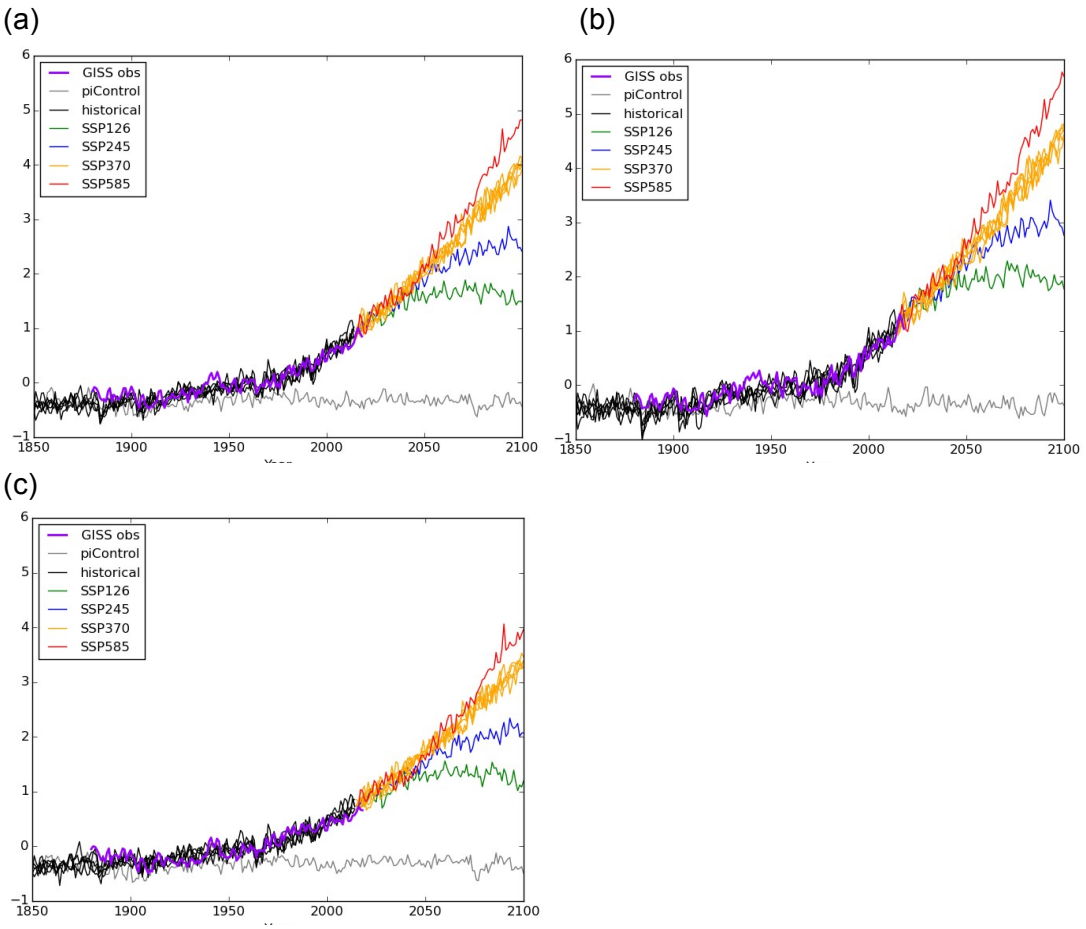


Fig. 13: Mean 2 m temperature anomaly (°C) for (a) the whole globe, (b) Northern Hemisphere, and (c) Southern Hemisphere from piControl, historical, and scenario simulations. Anomalies were computed relative to the period 1951–1980. The purple line indicates the observed 2 m temperature anomaly from Goddard Institute for Space Studies (GISS) Surface Temperature Analysis (GISTEMP v4) (GISS, 2019; Hansen et al., 2010; Lenssen et al., 2019).

Fig. 14 shows the spatial distribution of simulated temperature and precipitation changes until the end of the 21st century according to the strongest emission scenario SSP585. Temperature changes are very robust and exceed the 2 standard deviations of interannual variability of the control simulation over the whole globe (Fig. 14a). Generally, precipitation changes are less robust (Fig. 14b) with the Arctic and the Southern Ocean as well as the African tropics being prominent exceptions. Simulated precipitation changes can be regarded as less robust than temperature changes not only because of large internal variability of the precipitation but also because of large biases in present-day climate which amount to more than 7 mm/day in some tropical areas. Bias patterns for both 2 m temperature and precipitation as well as the magnitude of the biases for present-day climate are not surprisingly very similar to the ones in MPI-ESM (Müller et al., 2018, their Fig. 7f).

The well-known feature of Arctic Amplification, and to a lesser extent also Antarctic Amplification, can clearly be seen from Fig. 14a. According to the SSP585 scenario, the

temperature increases as much as 11°C over the Northern Barents Sea and around Spitsbergen. In the northernmost parts of the European and American continents the warming exceeds 7°C at the end of the century compared to the historical reference period. Large continental areas are affected by temperature increases of more than 5°C. Also over the Weddell Sea and over parts of Antarctica temperature increases of more than 5°C are simulated. Over the ocean, the warming generally amounts to 2 to 3°C.

Over large areas of central Africa and over the tropical Pacific, precipitation increases of more than 50% are simulated. Other areas, with comparable precipitation increases, include the ocean northwest of South Africa as well as northeastern parts of Greenland. Over the whole Arctic a substantial precipitation increase of more than 40% is simulated; over the Southern Ocean adjacent to the Antarctic continent extended areas are affected by precipitation increases of 20 to 30%. These precipitation changes are very robust since they exceed twice the interannual standard deviation of the control simulation. Except for parts of the Amazonas region, simulated precipitation decreases are less robust and are mainly concentrated in subtropical areas. They do not exceed 50% of present-day precipitation.

Compared to the multi-model CMIP5 ensemble (IPCC, 2014, Summary for Policymakers, their Fig. SPM.8), the temperature response in AWI-CM looks very similar, both regarding magnitude (11°C over Northern Barents Sea, more than 5°C over large continental areas as well as Weddell Sea and parts of Antarctica, 2 to 3°C over large parts of the ocean) and pattern of response. However, the warming hole, i.e. a lack of warming over the North Atlantic subpolar gyre, that is present in the CMIP5 ensemble (e.g. Menary and Wood, 2018; Chemke et al., 2020), hardly exists in AWI-CM. Furthermore, the precipitation increase in AWI-CM over the Arctic is less pronounced and the precipitation increase over Africa clearly more pronounced compared to the multi-model CMIP5 ensemble (IPCC, 2014, Summary for Policymakers, their Fig. SPM8). Otherwise, the precipitation response pattern is quite consistent.

It can be concluded that especially the temperature response pattern with strong Arctic and continental as well as weak ocean warming agrees very well with the multi-model ensemble mean of CMIP5 simulations, even in terms of magnitude. Also the feature of wetting polar, subpolar and tropical regions as well as drying subtropical regions agrees with patterns from the multi-model ensemble of CMIP5 simulations although the magnitude of the response is not as consistent as the magnitude of the temperature response.

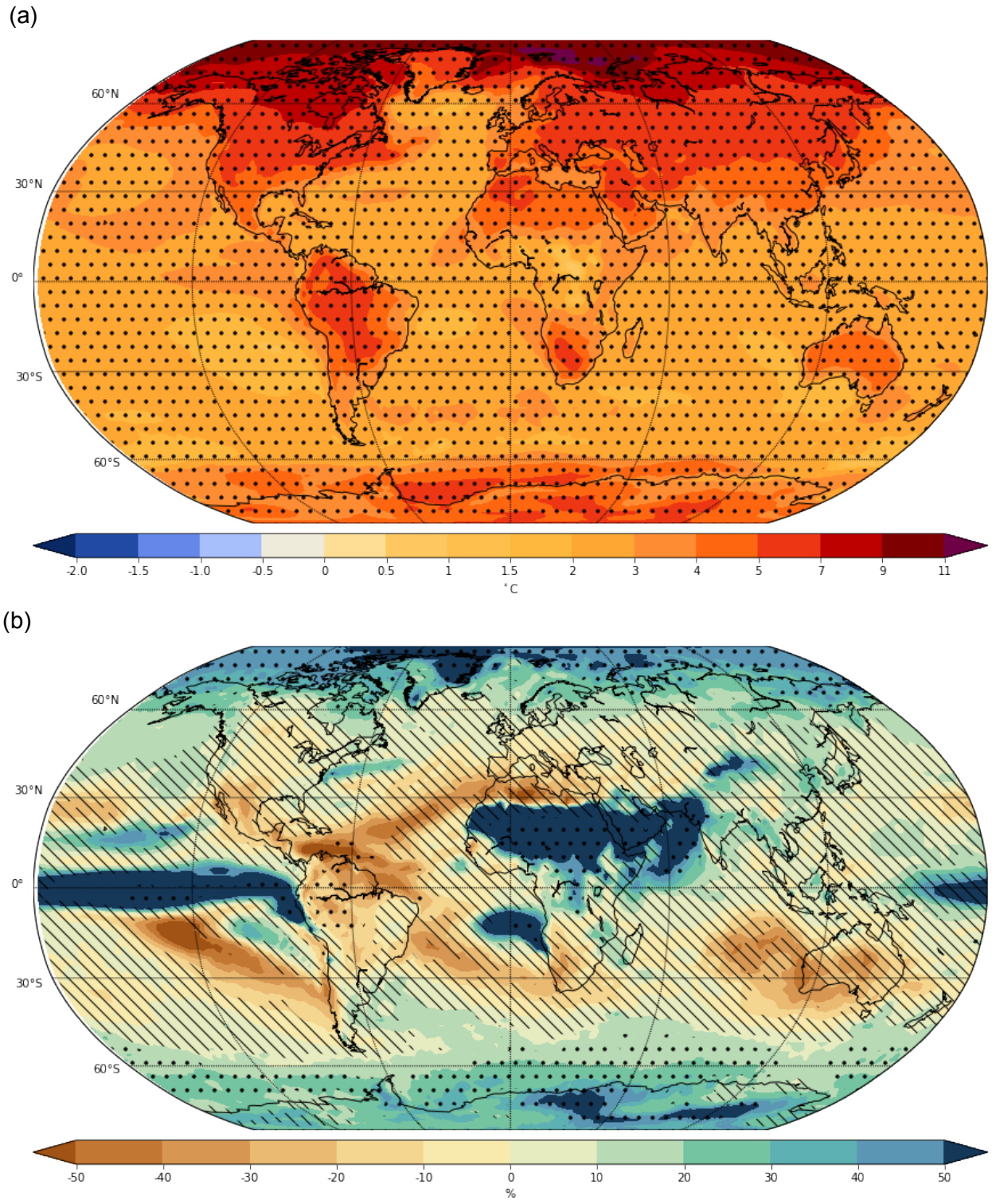


Fig. 14: (a) Annual mean 2 m temperature and (b) precipitation response according to the SSP585 scenario 2071-2100 compared to the historical period 1985-2014. Dotted (hatched) areas represent areas where simulated changes are larger than (smaller than) 2 standard

deviations (1 standard deviation) of the internal variability based on yearly means of the 500-year control simulation.

5.2.2 Sea ice extent

The simulated changes in sea-ice extent are shown in Fig. 15 for the Arctic (a, b) and the Antarctic (c, d) during March and September according to piControl, historical and tier1 scenario experiments (i.e. ssp126, ssp245, ssp370 and ssp585), along with observations of the last decades.

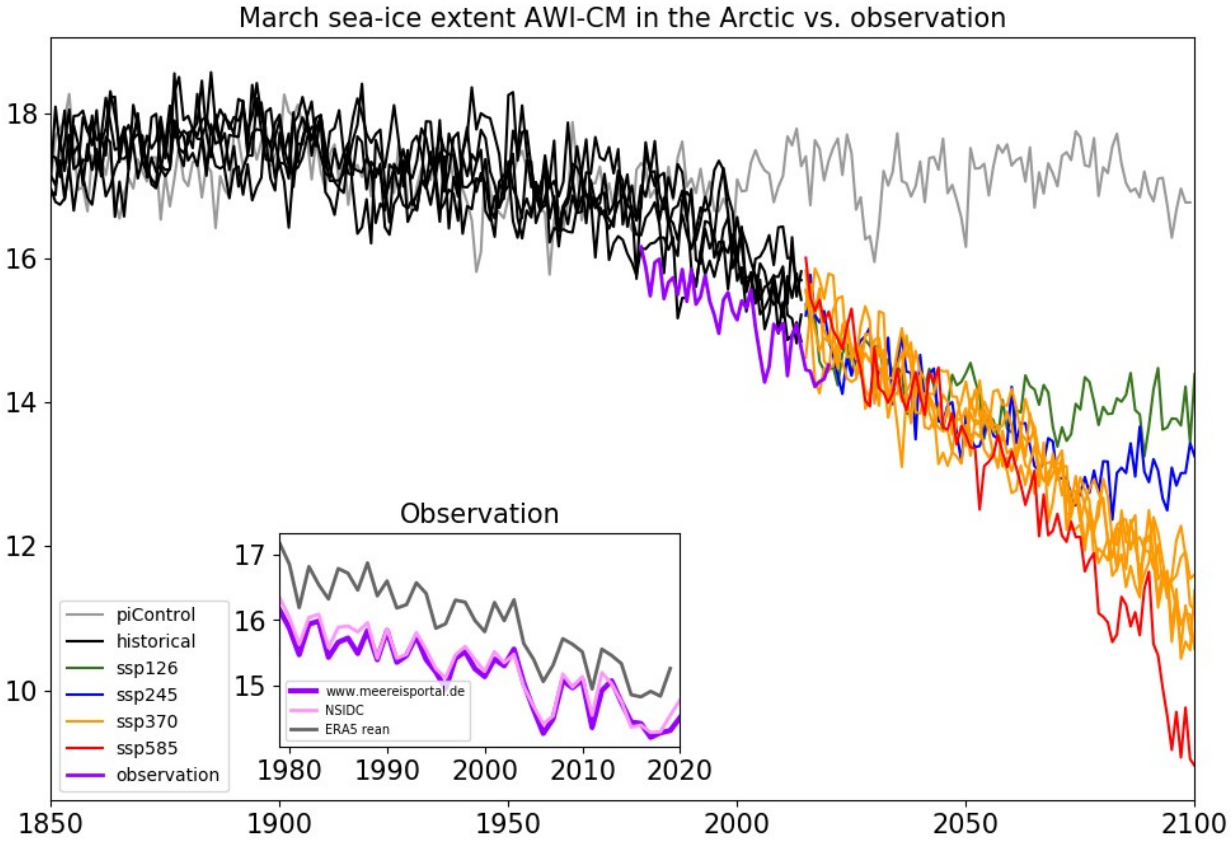
The strongest decline trend in sea-ice extent can be seen in the Arctic, during September (Fig. 15b). Starting between 2025 and 2030 there are isolated years with virtually sea ice free Arctic summers (1 Million km^2 sea ice extent or less) independent of climate change mitigation efforts (see also Notz et al., 2020). Starting from around 2050, except for SSP126, there are subsequent summers of a virtually ice-free Arctic ocean. The observed September sea-ice extent according to AWI's Sea Ice Portal (Grosfeld et al., 2016; derived from the University Bremen AMSR-ASI product, see Spreen et al. 2008) for 1979 to 2019 is shown (in purple) on top of AWI-CM outputs, confirming that AWI-CM sea ice extent agrees well with observations both in terms of the average and in terms of the rate of sea ice decline. However, the September Arctic sea ice *concentration* is underestimated in AWI-CM simulations of the last 30 years as shown in section 4.5. This needs to be taken in consideration when interpreting the projections of the future Arctic sea ice cover. According to the multi-model CMIP5 ensembles of September sea-ice extent, Arctic sea ice was melting even faster than predictions, even though observations remained within the first standard deviation of the models due to high internal variability of the participating models (Stroeve et al., 2015). In comparison to CMIP5, AWI-CM shows stronger sensitivity to the forcings. Unlike multi-model CMIP5 ensembles, ice-free Septembers will be expected not only for SSP585 (corresponding to RCP 8.5 in CMIP5), but also for SSP245 (corresponding to RCP 4.5 in CMIP5) and SSP370 (new pathway). IPCC AR5 reported September sea-ice extent reduction in 2081-2100 with respect to the average of the last 20 years of historical experiments (1986-2005) to be 43% for RCP2.6 and 94% for RCP8.5 (IPCC, 2013, p. 92). According to our simulations, the September Arctic sea-ice extent declines by the end of this century (2081–-2100) with respect to the last 20 years of historical experiments (1995–-2014) according to AWI-CM SSP126 and SSP585 are 64% and 99.99%, respectively. The inter-ensemble variability for both historical and scenario (ssp370) experiments is small. This means that the results are robust against internal variability.

Likewise, Arctic sea-ice extent during March shows a continuously negative trend for historical and scenario experiments (Fig. 15a). This negative trend seems to be independent of the scenario until the mid-21st century which implies that the impact of mitigation efforts might not be seen before that in terms of Arctic winter sea ice. However afterwards, sea-ice extent stabilizes at around 14 million km^2 for SSP126 and SSP245. As detailed in Fig. 15a, scenarios incorporating higher radiative forcings (SSP370 and SSP585) predict accelerating decline of sea-ice extent. According to the high-end scenario of SSP585, by 2100, Arctic March sea-ice extent will be half of its value at the beginning of the century.

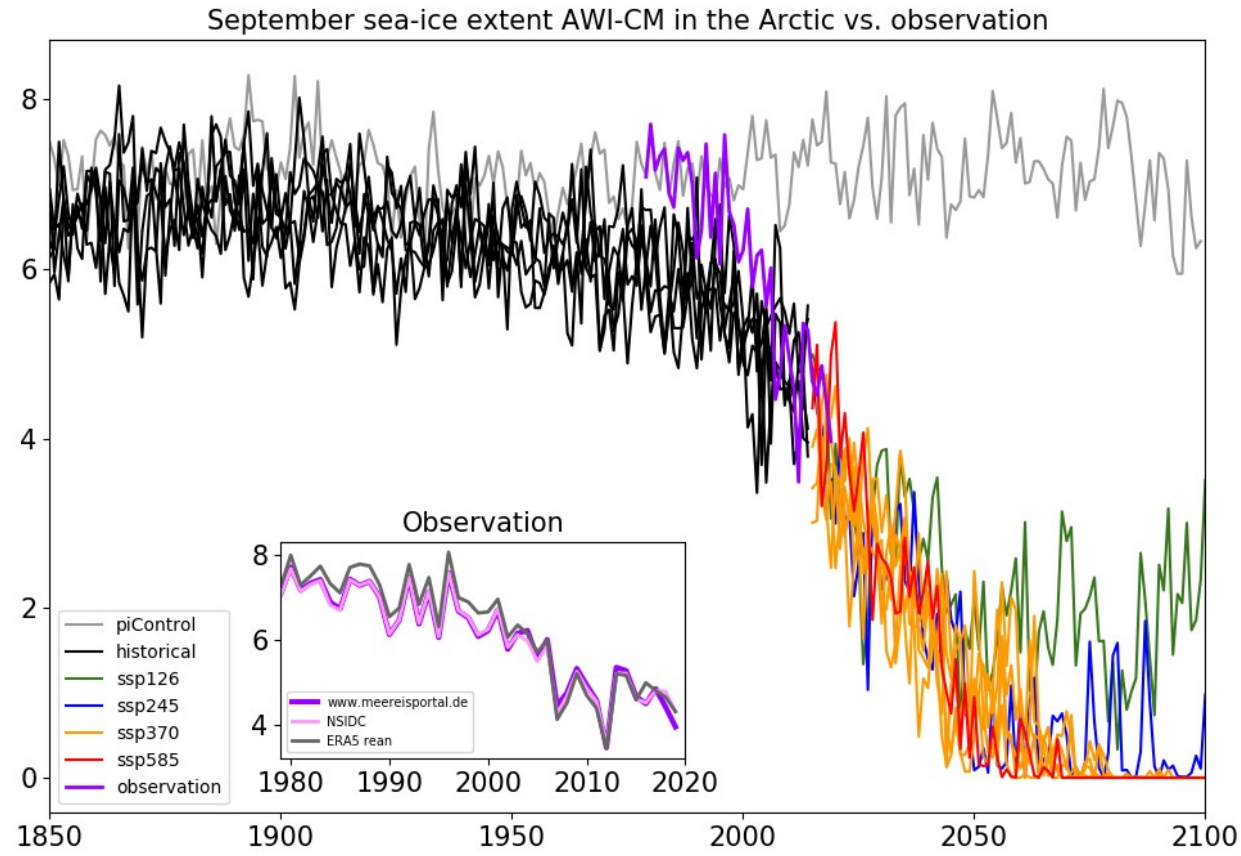
IPCC AR5 (IPCC, 2014: Climate Change 2014: Synthesis Report p.48) reported low confidence in near-term projections of Antarctic sea ice extent. This was due to the mismatch between CMIP5 models (strong simulated decline) and observations (no decline) along with very limited understanding of the origin of this mismatch. According to IPCC AR5, it is suspected that this phenomenon is likely due to regional variability within the Antarctic (IPCC, 2013, p.303). A study of individual CMIP5 models also suggested that although these models cannot replicate the observed Antarctic sea-ice extent trend, the observation still remains within the natural variability of better performing models (Turner et al., 2015). Furthermore, Bintanja (2013) showed that this sea-ice expansion could indeed be due to Antarctic sea ice shelf melting, which is not represented in CMIP5 models.

Similar to CMIP5 models, AWI-CM predicts declining Antarctic sea-ice extent for both September and March over recent decades (Figs. 15c and d) - which is in contrast to observations - and furthermore till the end of the century. In addition, the simulated difference between late winter and late summer Antarctic sea ice extent is more pronounced than in observations. Overall, interannual variability for Antarctic sea-ice extent is larger than for the Arctic, which agrees with the findings regarding CMIP5 models by Turner et al. (2015). Similarly to the Arctic sea ice, mitigation efforts only start to have a noticeable impact from around 2050.

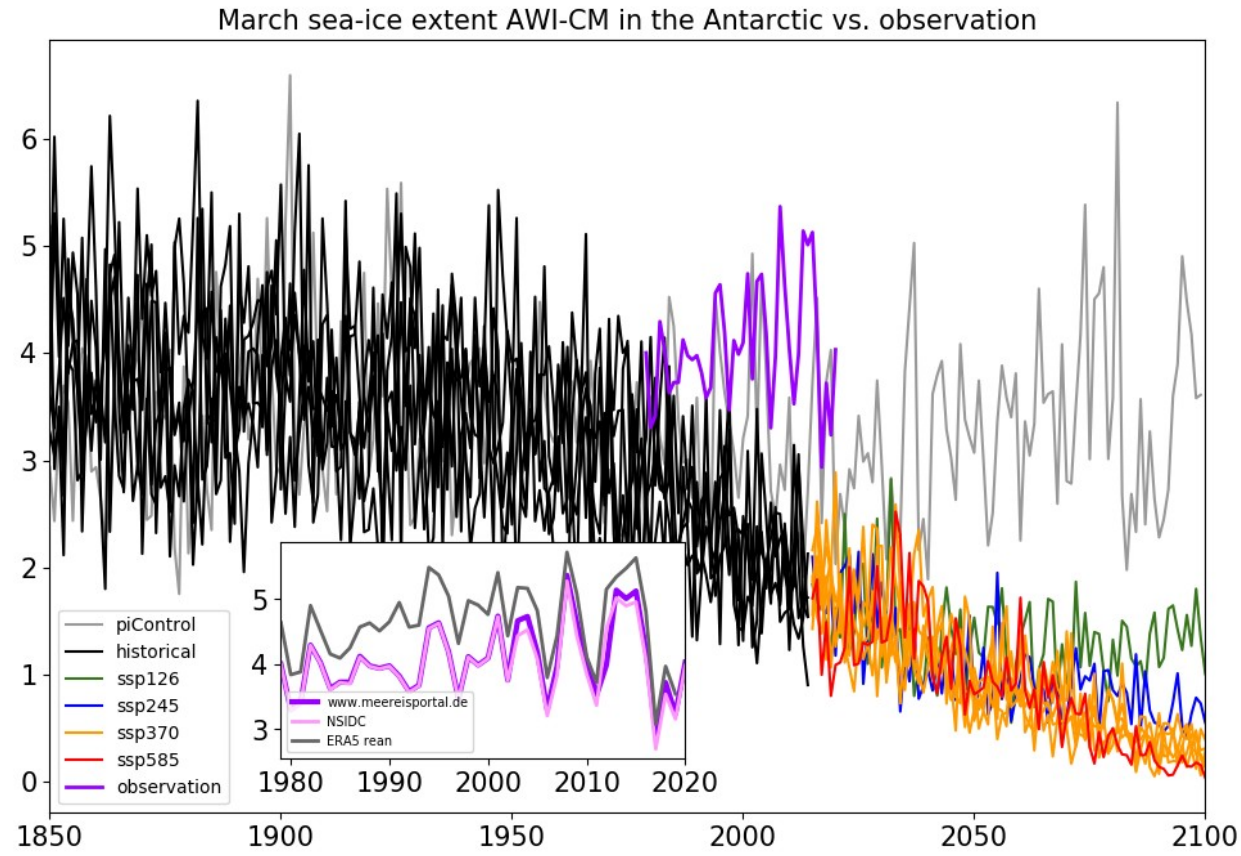
(a)



(b)


September sea-ice extent AWI-CM in the Arctic vs. observation
piControl
historical
ssp126
ssp245
ssp370
ssp585
observation
Observation
www.meereisportal.de
NSIDC
ERA5 rean
0
2
4
6
8
1850
1900
1950
2000
2050
2100
1980
1990
2000
2010
2020

(c)


March sea-ice extent AWI-CM in the Antarctic vs. observation
piControl
historical
ssp126
ssp245
ssp370
ssp585
observation
www.meereisportal.de
NSIDC
ERA5 rean
0
1
2
3
4
5
6
1850
1900
1950
2000
2050
2100
1980
1990
2000
2010
2020

(d)

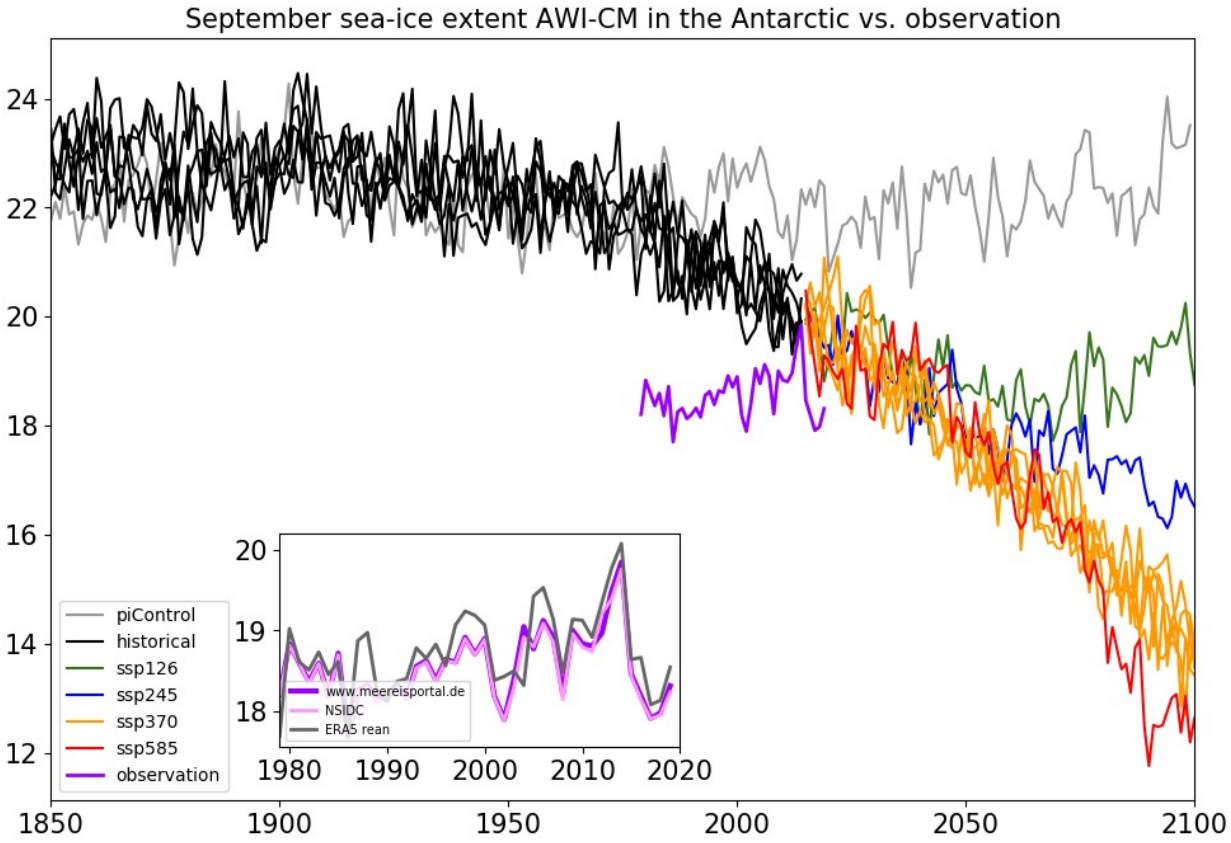


Fig. 15: (a) March and (b) September sea ice extent in the Arctic (Million km^2). (c) and (d) same as (a) and (b) but for the Antarctic region. The purple line indicates the observed sea ice extent from the sea ice portal meereisportal.de (Grosfeld et.al 2016). For comparison, observational National Snow and Ice Data Center (NSIDC) (Fetterer et al. 2017) and ERA5 reanalysis (Copernicus Climate Change Service (C3S), 2017; Hersbach et al., 2020) sea-ice extent are shown in inlays along with the ones from sea ice portal. The observation uncertainty is small and does not affect the conclusions.

Like for the sea-ice extent, the decline of Arctic sea-ice thickness is also evident from the historical simulation during the freezing season, most pronounced from around mid 20th century till recent years (Fig. 16). Simulated sea ice thickness in the Antarctic shows a weaker decline than that in the Arctic. We compare the simulated ensemble mean thickness in the Arctic with recent satellite thickness data from CS2SMOS (Ricker et al., 2017), which is constructed by merging CryoSat-2 and SMOS thickness together using the optimal interpolation method. Sea ice thickness in the historical simulation falls well into the observed range from 2010 to 2013. Basin-scale observations for sea ice thickness in the Antarctic are rather limited. A more

detailed evaluation against observations for Antarctic ice thickness is therefore currently not possible.

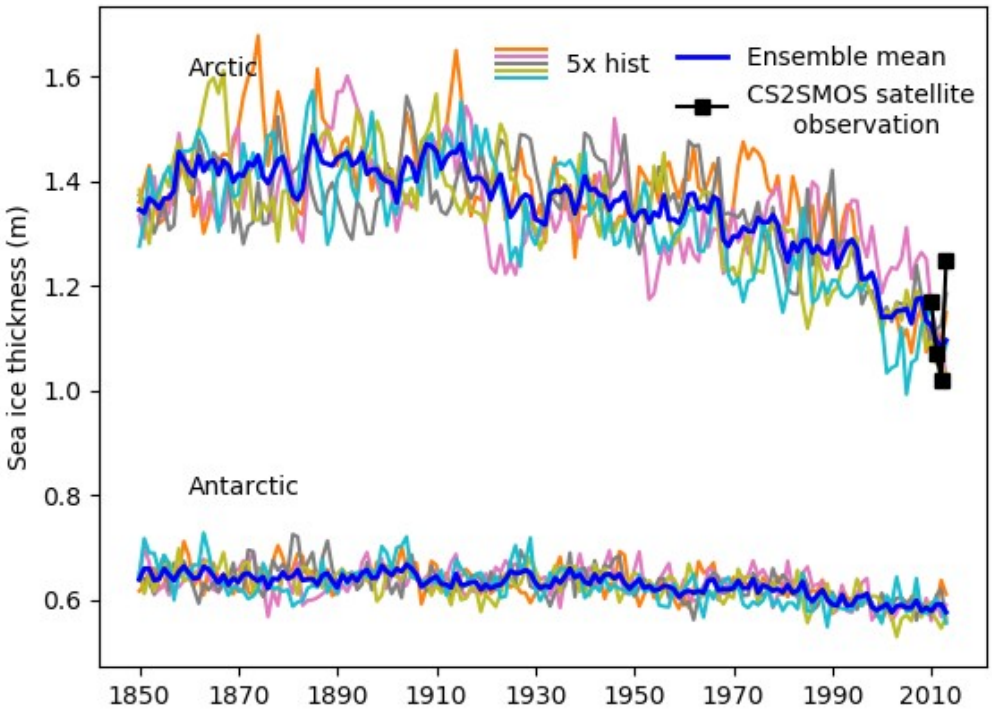


Fig. 16: Sea ice thickness in the historical simulation during the frozen season in the Arctic (averaged over December to March for each year) and the Antarctic (averaged over May to September for each year). The blue lines show the ensemble mean of the 5 members indicated by different colors. Satellite estimates for the Arctic from merged CryoSat-2 and SMOS data (CS2SMOS product) are shown by black squares.

5.3 Large scale circulation response

Similar to other climate models and as stated before, large scale circulation exhibits biases of the same order of magnitude as the simulated response to anthropogenic forcing affecting the reliability of the projections. Nevertheless, a few features are worth mentioning:

The mean sea level pressure (MSLP) response to increasing greenhouse gas concentrations (Fig. 17) is generally characterized by low anomalies over the polar regions and high anomalies in the southern mid-latitudes. Considering the geostrophic balance, this leads to an increase of the westerly flow in the northern and southern mid-latitudes mostly around 60° latitude. Over the Northern Hemisphere this increase is most pronounced in boreal autumn (SON) and winter (DJF). In the North Atlantic region the increase of the westerly flow is located further to the north compared to the CMIP5 ensemble mean as can be seen from Zappa and Shepherd (2017, their Fig. 1), while in the North Pacific region the location of the increase of the westerly flow is comparable. An intensified Aleutian low in boreal winter leads to a shift of the increased

westerly flow over the North Pacific sector towards lower latitudes with a maximum around 45°N. Over the Southern Hemisphere the increased westerly flow is equally present in all seasons with a shift in the African sector towards lower latitudes in austral winter (JJA) and spring (SON).

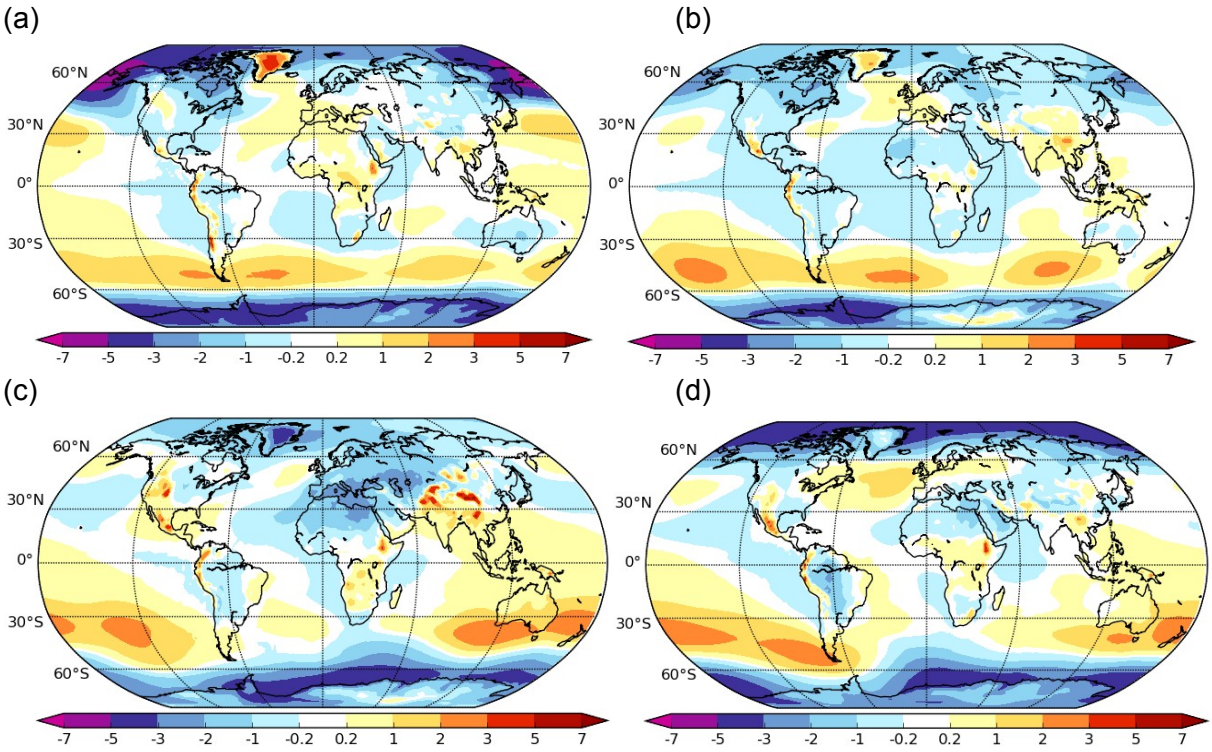


Fig. 17: Mean sea level pressure (MSLP) response in SSP370 scenario simulations (2071–2100) compared to historical simulations (1985–2014). For both the scenario and the historical simulations the 5 member ensemble means have been computed. (a) DJF, (b) MAM, (c) JJA, (d) SON.

Fig. 18 shows the zonal mean temperature and zonal mean zonal wind response to scenario forcing. The typical global warming signature with pronounced upper tropospheric tropical warming and near-surface polar warming occurs in AWI-CM as expected. The strongest warming in excess of 6 °C occurs in the Arctic boundary layer north of 70°N - known as Arctic Amplification. The upper tropospheric tropical warming amounts to 4 to 6 °C while the Antarctic warming is limited to around 4 °C. Strongest zonal mean zonal wind changes occur in the stratosphere around 100 hPa with increases in the westerly wind speed by around 5 m/s in 30 to 40°N and in 30 to 50°S. Around 60°S there are increases in the westerly wind speed by around 1 to 2 m/s throughout the troposphere. This pattern is very similar to the multi-model mean of CMIP5 (IPCC, 2013, chapter 12, their Fig. 12.19).

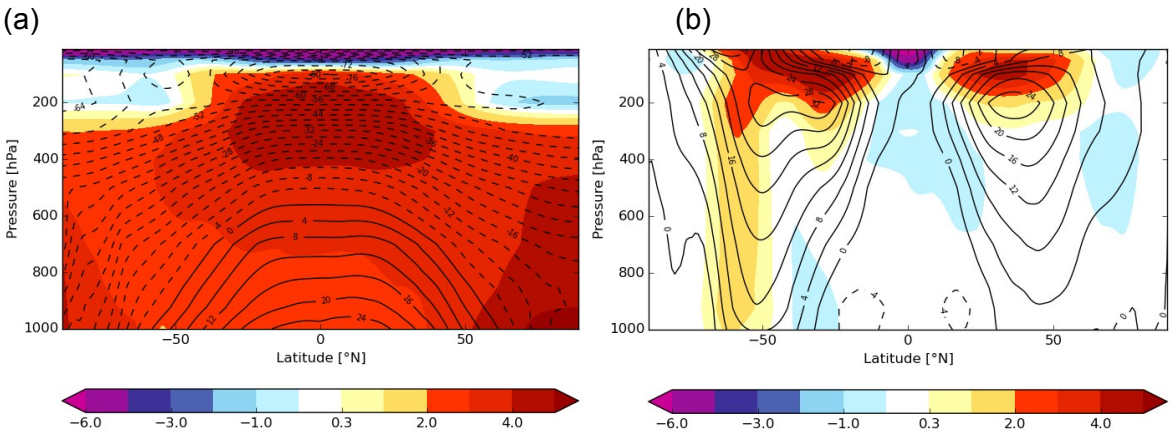


Fig. 18: (a) Zonal mean temperature response (°C), (b) zonal mean zonal wind (m/s) response SSP370 scenario simulations (2071-2100, annual means) compared to historical simulations (1985-2014) (shaded contours). For both the scenario and the historical simulations the 5 member ensemble means have been computed. The solid black lines represent positive values from the historical simulations and the dashed black lines negative values.

There is an ongoing discussion on how the waviness of the atmospheric flow in mid-latitudes will change in the future as a result of changes in the Arctic, through Arctic Amplification, and in the tropics, through upper tropospheric warming. The contrasting driving from the Arctic vs the tropics has been termed a tug of war in the mid-latitudes (e.g. Barnes and Polvani, 2015; Blackport and Kushner, 2017; Chen et al., 2020) Will there be a more zonal flow with a decrease in the intensity of atmospheric waves implying less extreme warm and cold events or will the meridionality of the flow get stronger implying more extreme warm and cold events in the mid-latitudes or will there be no change? To answer this question, various different objective indices have been defined. Cattiaux et al. (2016) defined the sinuosity index (SI) as the length of an isohypse of a specific value divided by the length of the 50°N latitude circle. If due to features such as cut-off lows there are separated isohypses of the specific value, the sum of the lengths of these isohypses is taken. The value of the isohypse is chosen as the area average of z500 over 30 to 70°N to accomodate for seasonal differences and climate change signals. If the SI equals to 1 the flow is zonal since the chosen isohypse is a straight line. The higher the SI, the stronger the meridional component of the atmospheric flow.

Fig. 19 shows the SIs computed for the piControl, historical, scenario simulations, and the ERA5 reanalysis. Overall the differences between the different simulations are smaller than differences between the model and reanalysis data. In all simulations, the waviness of the flow is more pronounced in boreal winter and spring compared to summer and autumn. The annual cycle is shifted compared to the ERA5 reanalysis. While the simulations show the maximum of waviness around February, according to the ERA5 reanalysis it is around May. The minimum of

waviness occurs around August in the simulations and around October according to the ERA5 reanalysis. While the amount of the maximum waviness is well captured in the model compared to the reanalysis, the minimum is too pronounced in the simulations indicating a too zonal flow in late summer.

Generally, a pronounced interannual variability can be seen both in the simulations and in ERA5. With increasing greenhouse gas concentrations, there is a tendency towards a more zonal flow in boreal summer and autumn while in winter and spring there is no robust change. This is consistent with the proposed tug-of-war (e.g. Barnes and Polvani, 2015; Blackport and Kushner, 2017; Chen et al. 2020): the upper tropospheric warming in the tropics leads to an increased meridional temperature gradient, stronger mean westerly flow and decreased waviness. In contrast, in boreal winter the effect of Arctic Amplification leads to a reduced meridional temperature gradient, weaker mean westerly flow and increased waviness offsetting the impact of upper tropospheric warming in the tropics. However, the impact on the waviness is very much under debate and shows very little robustness. Due to the lack of Arctic Amplification in boreal summer, the upper tropospheric warming in the tropics (Fig. 18a) may lead to a stronger zonal and less wavy flow. However, even in boreal summer differences are small compared to the strong interannual variability. Averaged over the year, the zonal mean zonal wind mainly increases in the stratosphere and only to some extent in the upper troposphere in the northern mid-latitudes while in the southern mid-latitudes zonal mean zonal wind increases are present throughout the troposphere, possibly due to the relative lack of Antarctic

Amplification (Fig. 18a).

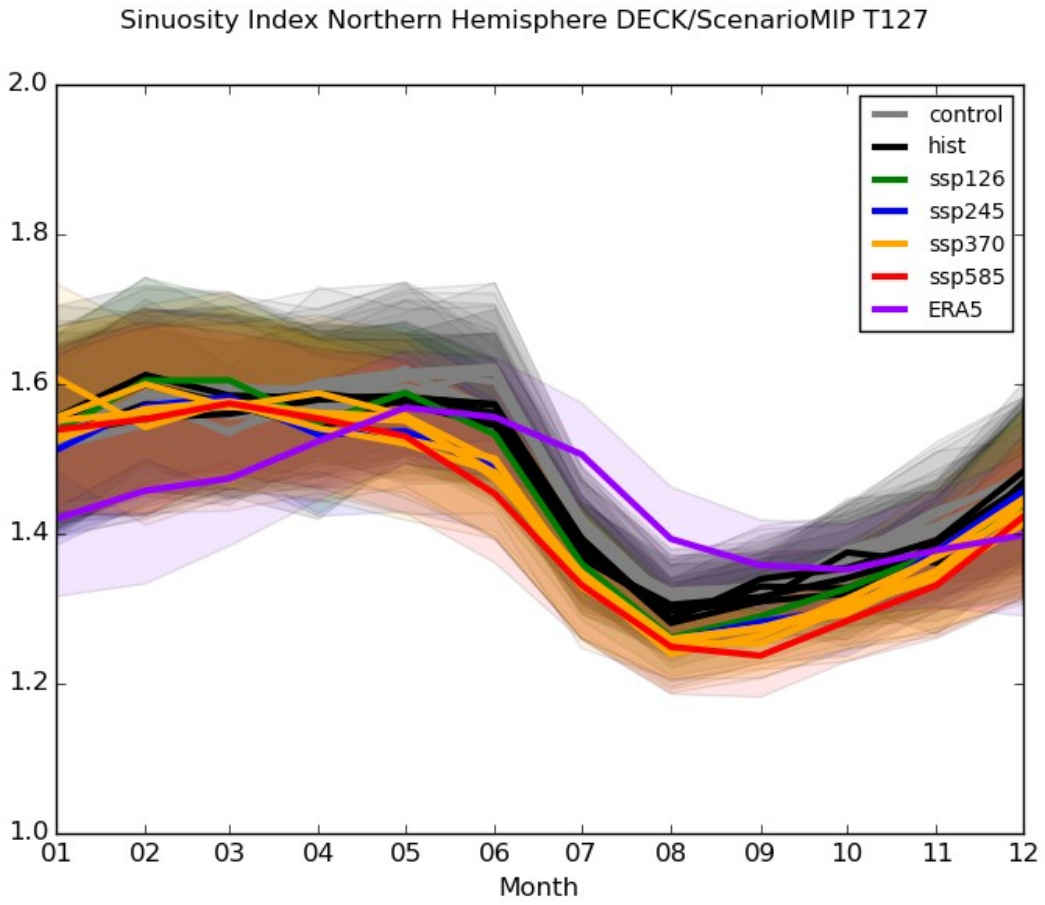


Fig. 19: Monthly sinuosity index (unity) in the Northern Hemisphere according to Cattiaux et al. (2016) from control, historical, scenario simulations, and from ERA5 reanalysis data (Copernicus Climate Change Service (C3S), 2017; Hersbach et al., 2020). Historical and scenario simulations are taken for the 30-year periods 1985–2014 and 2071–2100, respectively. From the control simulation all 30-year periods corresponding to the different ensemble members of historical and scenario simulations are considered resulting in multiple curves. The shaded areas represent the standard deviations of the 30 monthly sinuosity values for each simulation.

5.4 Ocean response

The Atlantic Meridional Overturning Circulation (AMOC) is an important element of the global ocean circulation. Transporting heat from the tropics to the northern North Atlantic, it has profound implications not only for the climate of north-western Europe but for the whole

Northern Hemisphere. It is also associated with ocean heat transport from the South Atlantic to the tropics (Weijer et al., 2019). Figs. 20 and 21 show the maximum AMOC strength at 26°N for piControl, historical, scenario simulations, and RAPID observations (Smeed et al., 2019) as well as for piControl, 1pctCO2 and abrupt-4xCO2 simulations, respectively. For the 15-year record of the RAPID observations, our model agrees well both in terms of the mean value and in terms of the range of interannual variability with the observations. The historical simulation is indistinguishable from the control simulation, i.e. agrees within a standard deviation with the control simulation, even though other parameters such as the Arctic sea ice and near-surface temperature show substantial changes towards the end of the historical period. Furthermore, the development of the AMOC strength according to the weakest scenario SSP126 is indistinguishable from the control simulation until the end of this century. For the three other emission scenarios the signal starts to emerge from the noise later than 2050, i.e. values are continuously lower than the piControl value minus one standard deviation.

In the case of a transient increase of the greenhouse gas forcing (historical, scenario, and 1pctCO2 simulations), the AMOC strength at 26°N gradually decreases by around 20% until the end of the 21st century with the high emission scenario SSP585 and by around 25% within 150 years in the idealized 1pctCO2 simulation. In the abrupt-4xCO2 simulation, the maximum AMOC strength decreases markedly by around 30% over the first 20 to 30 years. Over the remaining 120–130 years of the simulation, it slightly increases again by about 5% and thus reaches values of about 13Sv, which amounts to about 75% of the original AMOC strength.

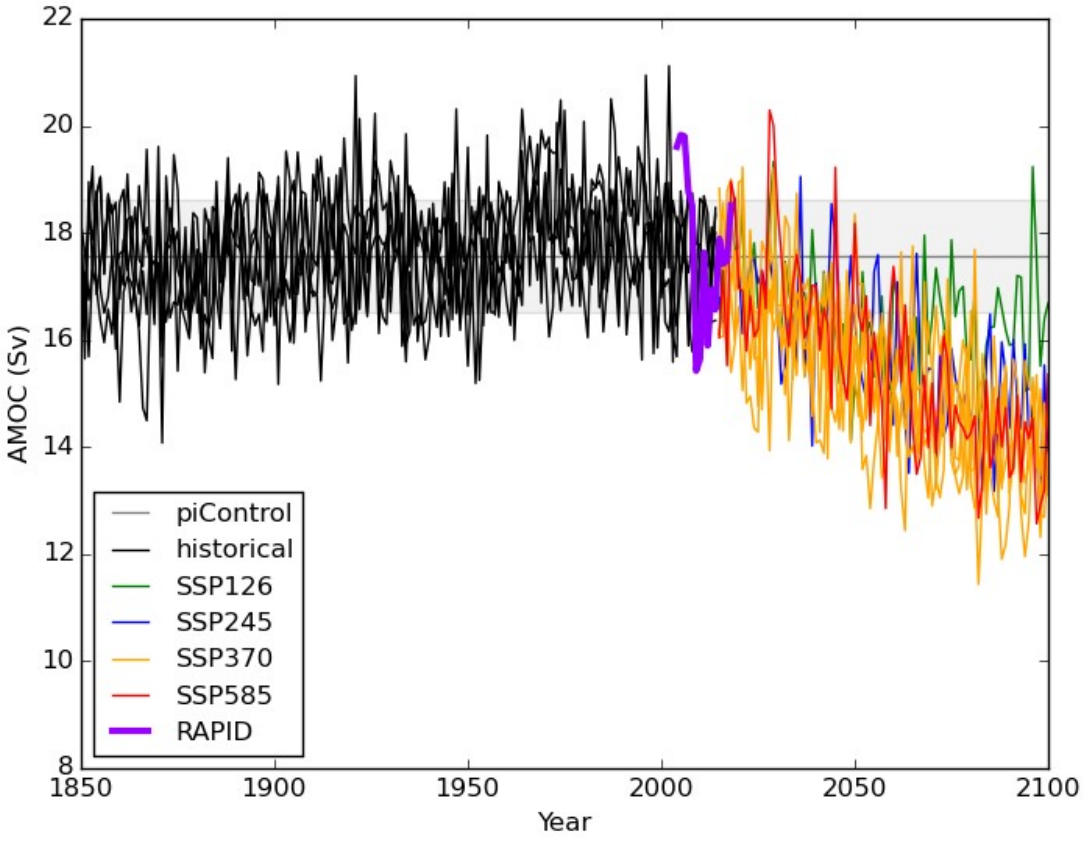


Fig. 20: Maximum Atlantic Meridional Overturning Circulation (AMOC) (Sv) at 26°N for piControl, historical, scenario simulations, and from RAPID observations (Smeed et al., 2019). For piControl, the mean value is indicated by the straight grey line and the standard deviation over the 500 year period is shown by the grey shaded area. Values of individual years are not shown for piControl.

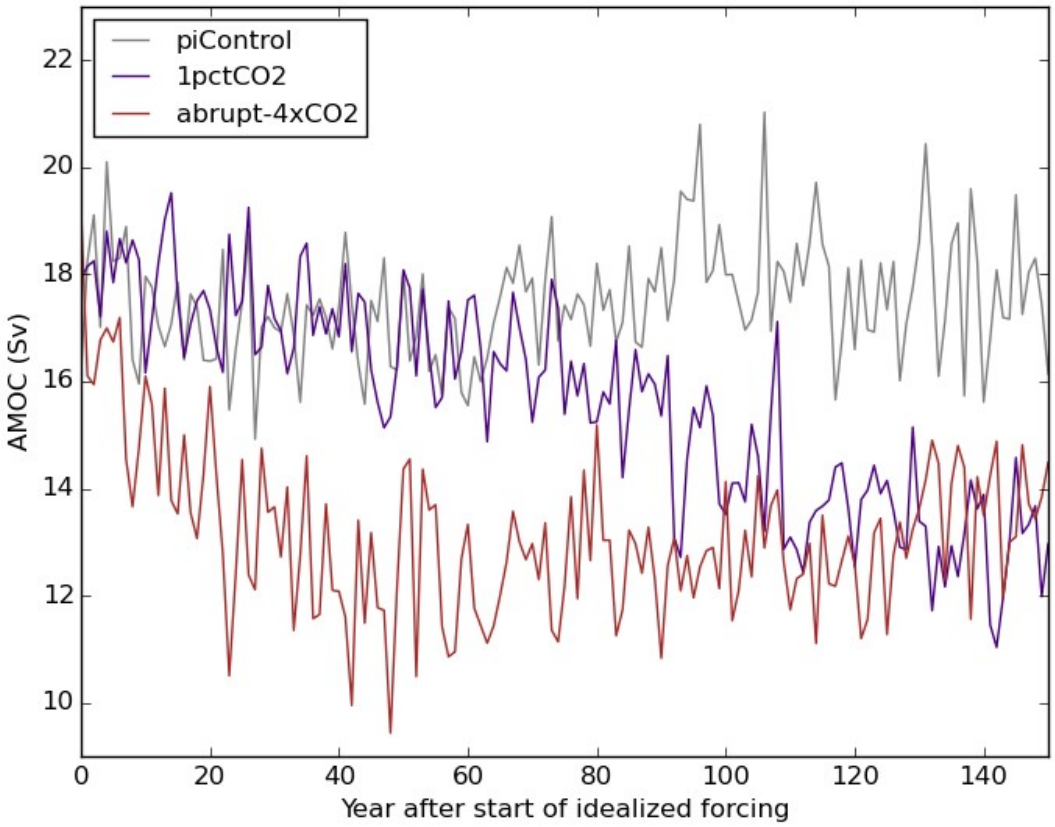


Fig. 21: Maximum Atlantic Meridional Overturning Circulation (AMOC) at 26°N (Sv) for piControl and the idealized simulations 1pctCO2 and abrupt-4xCO2. Please note that the branch-off point for both idealized simulations is year 250 of piControl simulation. The year on the x-axis represents the year after this branch-off point and therefore the year after the start of the idealized forcing.

Table 4 shows the ocean volume transports through some key ocean straits, averaged over all 5 ensemble members and the time period 1985–2014 for the historical runs and over the years 2071–2100 for the SPP370 runs. The historical runs show volume transports that are comparable to observed estimates for most of the ocean straits. For some straits, however, the volume transports are underestimated, including the export from the Arctic Ocean to the North Atlantic measured at the David Strait, the Indonesian Throughflow, and the transport in the Mozambique Channel. The main reason for this underestimation is due to the fact that the model resolution is not fine enough to resolve those narrow straits. In particular, the three main straits in the Canadian Arctic Archipelago (CAA) are only 10, 30 and 50 km wide at their narrowest locations, respectively, which cannot be well resolved with the mesh we used in the CMIP6 simulations. Improved representation of the CAA, and thus of the ocean transport through the Davis Strait, is expected for future coupled model configurations with higher ocean resolution, following promising results with high-resolution stand-alone configurations using FESOM (Wang et al., 2018; Wekerle et al., 2013).

For some ocean straits, the ocean volume transport shows a large response to the climate change in the SPP370 scenario. The Florida Current, for example, decreases by about 15% at the end of the 21st century in the SPP370 scenario, which is consistent with the weakening trend of the AMOC described above. The ocean volume transport in the Indonesian Throughflow and the Mozambique Channel also decreases in a warming climate (by about 20%). This implies that the exchange between the Pacific, Indian and Atlantic Oceans will become weaker. The oceanic linkage between the North Atlantic and the Arctic Ocean, however, is strengthened significantly in a warmer world, as shown by the increase in the volume transport through the Barents Sea Opening (increase by about 40%). Together with the temperature increase in the Atlantic Water, this implies that oceanic heat supply from the North Atlantic to the Arctic Ocean, and hence Atlantification of the Arctic Ocean, will increase in the future. As a consequence of ocean volume conservation, the excess ocean volume inflow through the Barents Sea Opening is balanced by an increased outflow from the Arctic through the Fram Strait.

In a warming climate the strength of the North Atlantic subpolar gyre (SPG) decreases, as shown by the increase in the sea surface height (SSH) in the SPG region (Fig. 22). The weakened SPG brings less Atlantic Water into the gyre circulation from the northeastern North Atlantic, which allows more Atlantic Water to continue to the north into the Nordic Seas. The enhanced northward flow is manifested by the increase in the SSH along the European coast. This can explain the stronger ocean volume transport through the Barents Sea Opening at the end of the 21st century in the SPP370 scenario (Table 4). The SSH on the northwestern side of the Gulf Stream increases in the warming scenario, which indicates a weakening of the Atlantic Current and is consistent with the weakening of the AMOC and the warming off the East Coast of the USA (Fig. 14a).

Table 4: Ensemble mean of ocean volume transport (Sv) through different straits for the historical runs and spp370 runs. (Positive values mean north or eastward flows)

Straits	Historical run (1985-2014)	SSP370 (2071-2100)	Observations	References of observations
Fram Strait	-2.4	-3.1	-2.0 ± 2.7	Schauer et al. (2008)
Davis Strait	-0.6	-0.5	-1.6 ± 0.5	Curry et al. (2014)
Bering Strait	1.3	1.2	0.8 1	Roach et al. (1995) Woodgate (2008)
Barents Sea Opening	1.7	2.4	2	Smedsrud et al. (2010)
Drake Passage	169.2	170.1	136.7 ± 6.9 173.3 ± 10.7	Meredith et al. (2011) Donohue et al. (2016)
Mozambique Channel	-11.7	-9.4	-16.7 ± 8.9	Ridderinkhof et al. (2010)
Indonesian Throughflow	-12.4	-10.0	-15	Gordon et al. (2010)
Florida Bahamas Strait (28°N)	37.9	32.1	31.6	McDonagh et al. (2015)

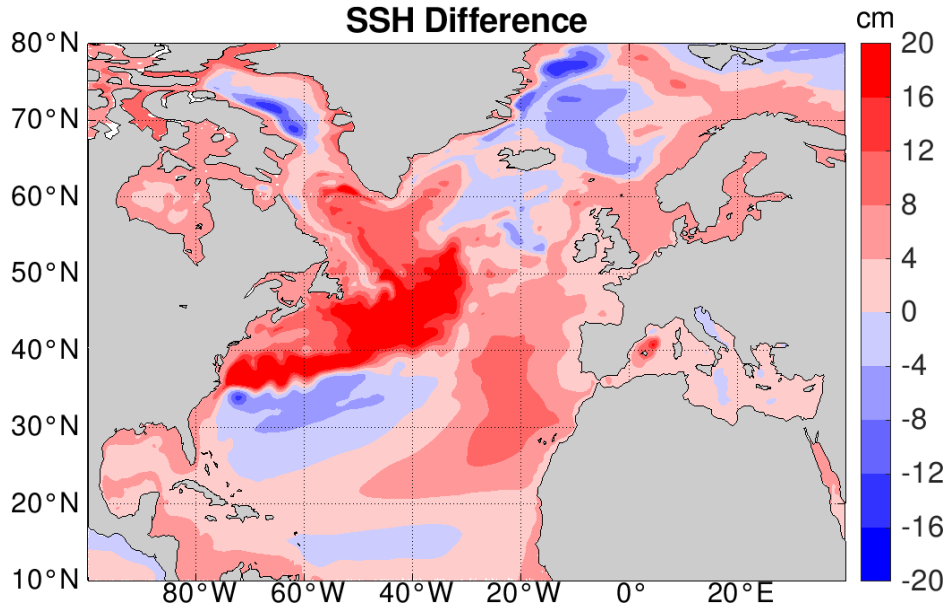


Fig. 22: Difference of the sea surface height (SSH) between the mean over 2071–2100 (SPP370) and the mean over 1985–2014 (historical run) averaged over five ensemble members. Only the dynamical sea level part is shown (i.e. contributions from thermal expansion and water mass changes are not considered).

5.5 Changes in the energy budget

The global-mean net total TOA radiative imbalance remains, on decadal timescales, close to zero in the historical simulation until around 1970, after which it increases to ~0.7 W/m^2 for present-day conditions (Fig. 23a, black solid curve), reflecting the uptake of heat by the climate system. This is less than the observational estimate of 0.9 W/m^2 for the period 2005–2014 by Trenberth et al. (2016), but within the uncertainty bounds (+/- 0.3 W/m^2. It matches the observational estimate by Johnson et al. (2016) who report 0.71 ± 0.1 W/m^2. Compared to CMIP5 and to other CMIP6 models, our simulated 0.7 W/m^2 are below the average (see Wild, 2020, their Figure 6). After the historical period, the net total TOA radiative imbalance decreases gradually in our SSP126 scenario simulation, stabilizes at ~0.9 W/m^2 in the SSP245 scenario simulation, and continues to increase to up to 2.0 W/m^2 in the SSP370 and SSP585 scenario simulations toward the end of the 21st century (Fig. 23a, coloured solid curves). In contrast to the net total TOA radiation, its shortwave component exhibits a negative imbalance varying between 0.0 W/m^2 and -1.0 W/m^2 over the course of the historical simulation (Fig. 23a, black dashed curve), which implies an increased planetary albedo (Fig. 23b, black solid curve). The increased planetary albedo, particularly pronounced during the second half of the 20th century (+0.2%; the absolute simulated planetary albedo is ~28.9%), is not due to changes in surface albedo (Fig. 23b, black dashed curve), but is likely for the largest part due to anthropogenic aerosols that have compensated for a similarly strong positive longwave-radiative imbalance due to increased greenhouse-gas concentrations.

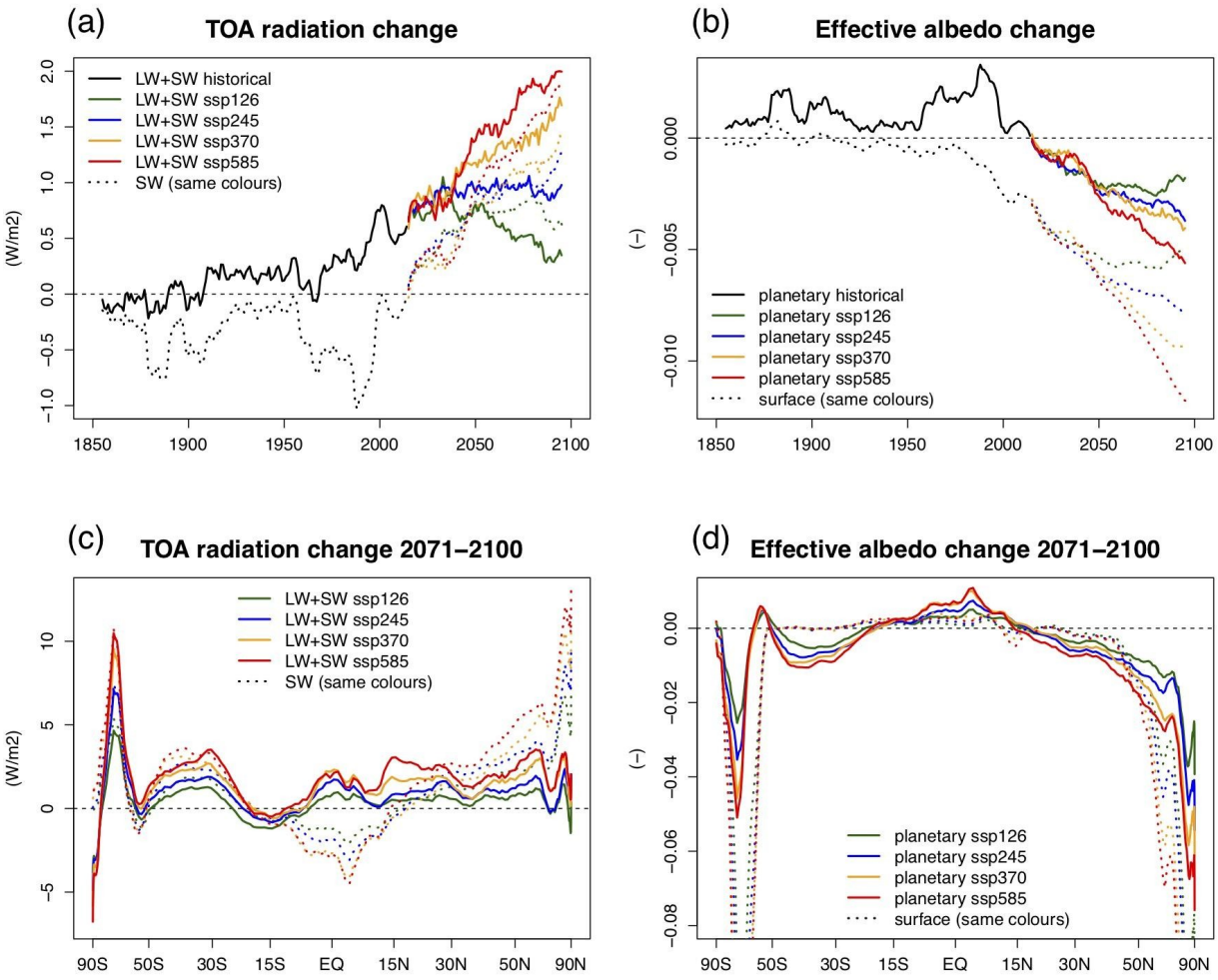


Fig. 23: Annual-mean changes in the energy budget, relative to piControl (delta approach). (a) Top of Atmosphere (TOA) net total (longwave (LW) + shortwave (SW); solid) and net shortwave (dashed) radiation time series after application of an 11-year running-mean, for the historical (black) and scenario (colours) simulations. Positive is downward. (b) Changes in the effective albedo diagnosed as upwelling divided by downwelling shortwave 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 period 2071-2100 (scenario simulations only), plotted against sin(latitude) to reflect equal-area global contributions. Extreme negative values of effective surface albedo changes, reaching about -20% in both polar regions, are truncated in (d) to increase the visibility of changes in lower latitudes. Only the respective first ensemble member of the historical and SSP370 scenario simulations is shown.

While according to our simulations an increased planetary albedo has prevented a stronger warming of the climate system until present-day, the planetary albedo is projected to decrease and thus amplify the future warming in all scenarios (Fig. 23b, coloured solid curves and Fig. 23b, coloured dashed curves). The global-mean effective surface albedo, which seems to have played no major role until around 1980, is projected to decline by more than 1% (the absolute

simulated effective surface albedo is ~13%) until the end of the century in the SSP585 scenario simulation and is thus a significant part of the projected positive shortwave feedback. Interestingly, the global-mean net shortwave radiation is projected to increase faster than the total radiation (Fig. 23a). This implies that, while reduced outgoing longwave radiation (OLR) has caused the warming until present-day, the OLR is projected to *increase* toward the end of the century: Due to the strong shortwave feedback, the positive influence of increasing temperatures on OLR is projected to outweigh the direct negative influence of increased greenhouse-gas concentrations on OLR. This behaviour has been found for most CMIP3 and CMIP5 models (Donohoe et al. 2014).

The surface albedo is decreasing particularly strongly in the regions with declining sea-ice extent, that is, the Southern Ocean (60°S-70°S) and the Arctic (north of 70°N) (Fig. 23d, dashed curves). These changes are clearly reflected in the planetary albedo (Fig. 23d, solid curves), which however also reveals non-surface-related albedo changes in lower latitudes caused by cloud feedbacks. In particular, toward the end of the century the planetary albedo is projected to increase in the tropics between 15°S and 15°N (negative feedback) and to decrease in the subtropics (positive feedback) (Fig. 23d, solid curves).

While the surface-driven changes in the planetary albedo projected toward the end of the century are substantial in both polar regions, the positive net total TOA radiative imbalance is particularly pronounced over the Southern Ocean (Fig. 23c, solid curves). This is consistent with the relatively weak Antarctic and strong Arctic warming (Fig. 14a), leading to strongly enhanced upwelling longwave radiation in the Arctic but not in the Antarctic. This asymmetry in terms of polar amplification and TOA fluxes is consistent with the fact that the Southern Ocean temperature responds much more slowly to changes in atmospheric thermal forcing because of the spatial structure of the global Meridional Overturning Circulation (MOC), with circumpolar upwelling of unperturbed water masses in the south and downwelling in the north (Armour et al. 2016, Rackow et al. 2018).

5.6 Changes in ENSO

Since ENSO is the dominant mode of interannual variability (Timmermann et al., 2018), with pronounced global impacts through far-reaching teleconnections such as the atmospheric bridge (Alexander et al. 2002), an important question is whether the character of ENSO will change under climate change. To address this question, we resort to the five SSP370 projections with AWI-CM until the end of the 21st century. According to these simulations, when compared to the probability distribution of Niño 3.4 SST anomalies for 1870–2014, strong warm and cold SST anomalies become more likely by the end of this century (Fig. 24). The increase of strong cold SST anomalies dominates, so that the clear positive asymmetry between El Niño and La Niña (Fig. 5), as diagnosed from the skewness of modelled Niño 3.4 SST anomalies (1870-2014: 0.15 ± 0.16, observed: 0.36), is reduced under climate change (2071-2100: 0.04 ± 0.17). A reduced positive asymmetry under global warming has been found for most CMIP5

models (for the overlapping Niño3 region; Ham et al., 2017). However, despite sharing the atmospheric model with AWI-CM, in that study MPI-ESM-LR and MPI-ESM-MR showed a strong increase of DJF Niño3 skewness with the RCP4.5 scenario, which might again hint at the different ocean model formulation in AWI-CM and MPI-ESM and should be evaluated in more detail in the future.

Interestingly, the seasonal cycle that has been subtracted to compute the SST anomalies within the Niño 3.4 box consistently changes under climate change in all ensemble members (Fig. 25): when subtracting the different annual means, a stronger positive (negative) peak in April/May (August–October) is evident, suggesting that seasonality within this region will likely increase until the end of the century. Concerning phase locking of Niño 3.4 SST anomalies to the seasonal cycle, the variability increases throughout the entire year (Fig. 26), shifting the characteristic U-shape upwards by the end of the century.

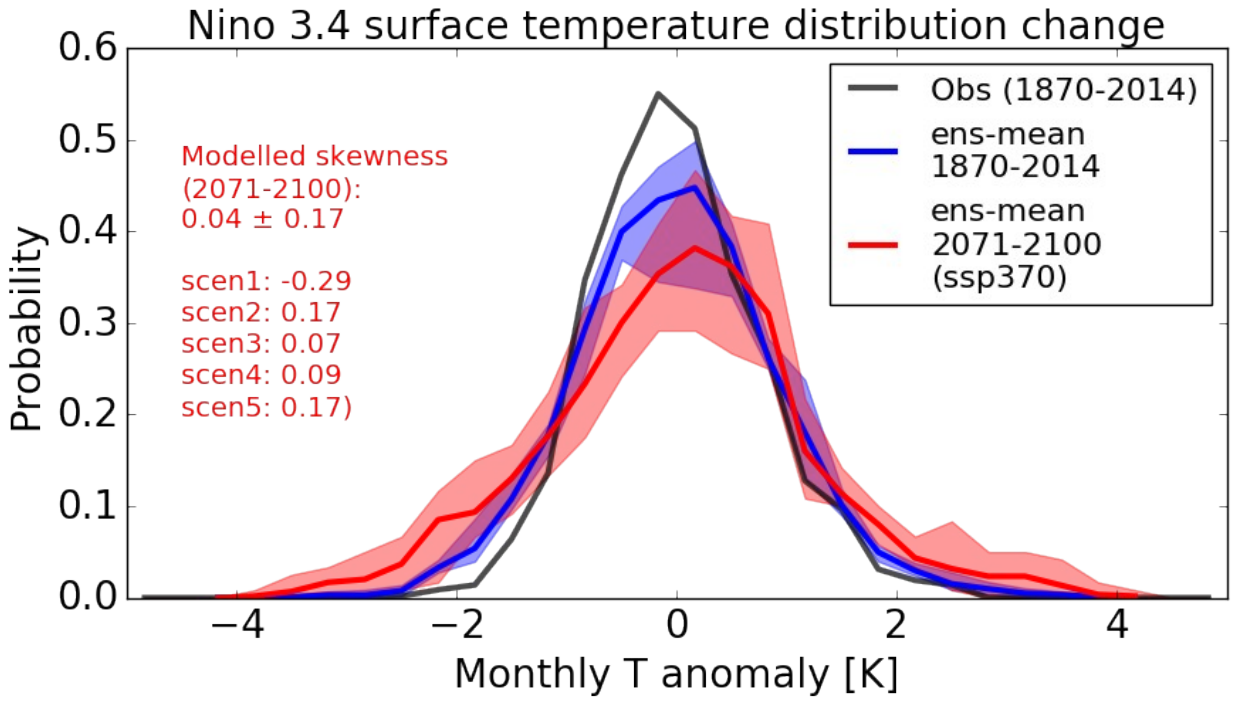


Fig. 24: Change of the probability distribution function of sea surface temperature anomalies in the Niño 3.4 region for 2071–2100 (red line). Compared to 1870–2014 (black and blue lines), extreme anomalies become more likely while the probability of low to medium anomalies decreases. The range of the model results is shaded. ENSO asymmetry (positive skewness) decreases compared to 1870-2014 (see Fig. 5). All data have been linearly detrended and the seasonal cycle has been removed.

*

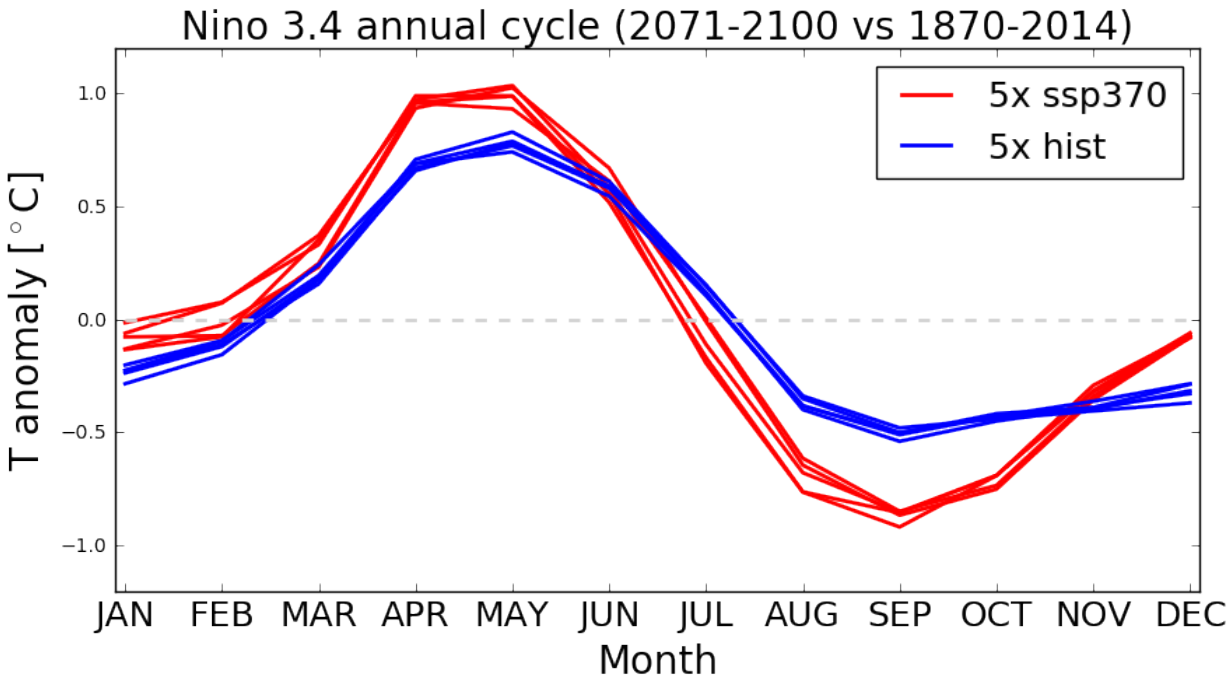


Fig. 25: The annual cycle within the Niño 3.4 box for historical (1870–2014) and future (2071–2100) conditions. 5 ensemble members each are shown for the historical simulations (blue) and the ssp370 scenario (red). The different annual means (historical: 25.4°C, scenario: 28°C) have been removed to highlight the different characteristics of the annual cycles.

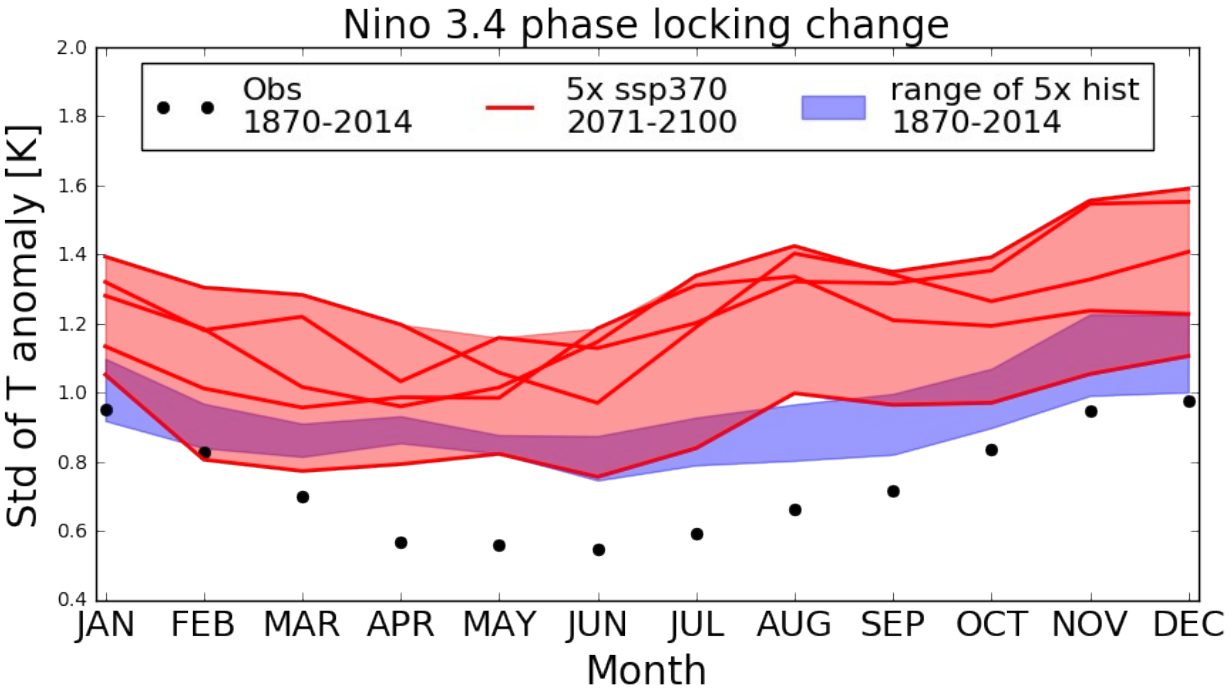


Fig. 26: Seasonal phase locking of temperature anomalies in the Niño3.4 region at the end of the 21st century (2071–2100) and for the period 1870–2014. Black dots give the monthly

standard deviation of the observed Niño3.4 index for 1870–2014 (Rayner et al., 2003). Red lines denote the standard deviations of the simulated Niño3.4 indices for the five ssp370 ensemble members. The range of the model results is shaded in light blue (1870–2014, as in Fig. 7) and in light red (2071–2100).

6 Discussion

The climate model AWI-CM-MR-1-1 presented here has proven to perform well compared to CMIP5 as well as selected CMIP6 models and therefore can be regarded as a solid contribution to the CMIP6 ensemble. Model drift in the control simulation is negligible. While some long standing model biases in AWI-CM such as a too zonal North Atlantic current, a too strong atmospheric westerly flow in the Euro-Atlantic region, a too cold subpolar North Atlantic gyre as well as a warm bias west of Africa are still present, there is a good representation of the North Atlantic ocean temperature profile, i.e. the warm bias in mid-depths is largely alleviated as discussed in Rackow et al. (2018). Furthermore, there is no pronounced Southern Ocean surface warm bias and therefore a very good representation of Antarctic sea ice and circulation compared to CMIP5 models and also compared to the Max Planck Institute for Meteorology CMIP6 model MPI-ESM (Müller et al., 2018): both mean sea level pressure and vertical profiles of zonal mean zonal wind are better represented over the Southern Ocean - except for the area west of the Antarctic peninsula where a positive mean sea level pressure bias is replaced by a negative mean sea level pressure bias.

Since the atmosphere model is the same as in MPI-ESM, it could be hypothesized that the high resolution of the ocean model in the Southern Ocean helps to reduce the long standing biases in this area that is very important for the global ocean circulation as well as heat and carbon uptake (e.g. Frölicher et al., 2015). This is subject to further investigation in the future and in collaboration with the Max Planck Institute for Meteorology. Contrary to observations, Antarctic sea ice decline has been simulated for the past decades, similar to CMIP5 and CMIP6 model results (Roach et al., 2020). .

In terms of the response to increasing greenhouse gases, our model shows very similar outcomes compared to the multi-model ensemble of CMIP5 simulations, both in patterns and in magnitude. Features such as a strong Arctic Amplification along with a weak Antarctic Amplification, Arctic wetting (Bintanja and Selten, 2014), subtropical drying, increased frequency of extreme La Niña & El Niño events (Cai et al., 2015), reduced ENSO asymmetry (Ham et al., 2017), and weakening Atlantic Meridional Overturning Circulation are very similar to the previous simulations.

However, there are potentially important differences that need further investigation: The weakening of the Atlantic Meridional Overturning Circulation is less pronounced compared to the CMIP5 ensemble mean. Since weakening of the Atlantic Meridional Overturning Circulation has been linked to the emergence of a warming hole over the North Atlantic subpolar gyre (e.g. Keil et al., 2020), it is consistent that the warming hole is only weak according to our model results.

The equilibrium climate sensitivity of our model (3.2°C) is slightly lower than the CMIP5 and CMIP6 multi-model means (3.4°C and 3.7°C, respectively, according to Meehl et al., 2020) and slightly higher than the CMIP6 version of the Max Planck Institute for Meteorology model MPI-ESM sharing the same atmosphere component (tuned to be 3.0°C). Our transient climate response is with 2.1°C slightly higher compared to the CMIP5 multi-model mean (1.9°C according to Meehl et al., 2020), slightly lower compared to the CMIP6 multi-model mean (2.2°C according to Meehl et al., 2020), and around 23% higher than in MPI-ESM (1.7°C). This might imply that the deep ocean takes up less energy in our model compared to MPI-ESM.

Furthermore, in our model, the decline of Arctic sea ice extent by the end of the 21st century is stronger than the multi-model mean over CMIP5 simulations, suggesting a higher likelihood of an ice-free Arctic in September even before 2050. According to our simulations, mitigation efforts only start to have an impact after 2050 in terms of Arctic winter sea ice. Surface albedo changes, in particular in the polar regions where sea ice declines, are projected to contribute substantially to a strong positive shortwave feedback. Note however that a recent geoengineering study based on AWI-CM indicates a small impact of the Arctic ice-albedo feedback on temperatures outside the Arctic (Zampieri and Goessling 2019).

While the Arctic sea ice extent trend is still slightly smaller in our model simulation compared to observations over the last few decades, the global mean temperature increase is slightly larger compared to observations. This could either hint at an underestimation of Arctic Amplification in our simulations or that multi-decadal internal variability is superimposed on the observed Arctic climate change. The latter hypothesis is supported by Kay et al. (2011), Ding et al. (2017), and England et al. (2019) stating that around half of the strong negative Arctic sea ice trend over the past decades is explained by internal variability and the other half by the climate change signal although there are strong seasonal and regional differences (England et al., 2019).

The AMOC decreases by around 25% until the end of the 21st century according to the AWI-CM SSP585 scenario simulation, which is less than the multi-model average value of around 40% calculated from CMIP5 models and Earth System Models of Intermediate Complexity (EMICs, Weaver et al., 2012; Cheng et al., 2013). CMIP6 models tend to show even stronger AMOC declines than CMIP5 models (Lyu et al., 2020). Previous studies suggest that the representation of western boundary currents and the Agulhas leakage in higher-resolution ocean models can influence the AMOC strength (e.g. Biastoch et al., 2009, 2018, Sein et al., 2018, Weijer et al., 2019, Hirschi et al., 2020). More dedicated studies in this regard will be carried out in our further work.

7 Conclusions

The Alfred Wegener Institute climate model, AWI-CM, described in this study, contributes to the diversity of climate models with the unstructured mesh approach for its sea ice-ocean component. Biases in AWI-CM tend to be less pronounced than in models contributing to the previous Climate Model Intercomparison Project 5, as shown by objective performance indices.

Even though some long standing biases such as a too zonal pathway of the North Atlantic current, the cold bias over the North Atlantic subpolar gyre or the warm bias west of Africa are still present in AWI-CM, especially Southern Ocean sea surface temperature and the atmospheric temperature above, sea ice concentration around Antarctica as well as North Atlantic ocean temperature profiles are well represented. Furthermore, there is an excellent agreement of the Arctic sea ice thickness in the past years for which observations are available. Therefore, AWI-CM results are a solid contribution to the CMIP6 project. Sea ice-ocean models on unstructured meshes have matured (now contributing to CMIP6) and can be used at high resolutions enabled through excellent scalability characteristics. Our results support the notion that some of the climate change features are robust against model formulation. However, there are some important features that deviate from other CMIP simulations:

1) Despite the smaller Arctic sea ice decline trend compared to observations, as early as starting between 2025 and 2030 there are isolated years with virtually sea ice free Arctic summers (1 Million km^2 sea ice extent or less) independent of climate change mitigation efforts. Only after 2050 mitigation efforts start to play a substantial role and Arctic sea ice can recover to some extent in the SSP126 scenario with strong mitigation efforts.

2) The AMOC decreases by around 25% until the end of the 21st century according to the AWI-CM SSP585 scenario simulation, which is less than the multi-model average value of 40% calculated from CMIP5 models and Earth System Models of Intermediate Complexity (EMICs, Weaver et al., 2012).

The AWI-CM model data is available through the Earth System Grid Federation (ESGF) and includes not only the DECK and ScenarioMIP experiments (Eyring et al., 2016; O'Neill et al., 2016) with AWI-CM-1-1-MR (Semmler et al., 2018) described in the present study. At the time of writing, AWI-CM results from the Polar Amplification Model Intercomparison Project (PAMIP, Smith et al., 2019) and results with AWI-CM-1-1-HR from the High Resolution Model Intercomparison Project (HighResMIP, Haarsma et al., 2016) are available as well (Semmler et al., 2017; Semmler et al., 2019). Furthermore, data publications of AWI-CM simulations are planned for OMIP (Griffies et al., 2016) and PMIP (Kageyama et al., 2018).

Appendix

A1 Postprocessing software for cmorization and easy use of unstructured mesh data for analysis

The atmospheric model data was cmorized with an approach developed by the German Climate Computing Centre (DKRZ). All the FESOM ocean data was also cmorized in a CDO-compatible manner. During the cmorization of sea-ice and ocean data, the user is notified of any incompatibilities between the model output and the designated data request tables. In the conversion process, the most time consuming step would be changes to the bulk variable data itself, but metadata can generally be applied quickly and independent of the data amount in one

file. The time and resources required for changes in the variable data on the other hand will theoretically increase linearly with the size of the mesh and the output frequency of the data. Therefore we had to avoid all steps and utilities which can not alter data in-place, e.g. where no auxiliary file or memory allocation is required.

The overall requirements for a complete CMIP6 formatted file set can be somewhat overwhelming due to its complexity (data request (DR), naming conventions, controlled vocabularies, Earth System Grid Federation (ESGF) requirements, model details). Therefore our CMIP6 CMOR setup has been realized via a compact and human-readable setup with a domain-specific language (DSL). This setup can be executed shared-memory-parallel per conversion task on a HPC system right where the data is stored.

The following steps are necessary, and their processing times scale with the grid size and are limited by the I/O speed of the filesystem (i.e. changes of the bulk data):

- data output frequency change (e.g. due to requirements in the DR)
- unit change (e.g. from deg C to K). Technically this is not a problem, but requires lengthy input / output operations for all affected data.
- data concatenation (merge)
- data compression (e.g. different as the ESGF)
- file format change (e.g. as required by ESGF)

A necessary requirement for using the model data for scientific research is the availability of tools that are able to perform basic pre- and post-processing operations on model output. For ocean models on structured grids, a wide selection of tools is available; for the new generation of global ocean models formulated in unstructured meshes, however, no existing off-the-shelf packages exist. Therefore, we have created a set of python-based analysis and visualization tools for FESOM - pyfesom (https://github.com/FESOM/pyfesom). The pyfesom repository contains a python library and command line tools that allow it to perform basic data analysis and visualisation of FESOM data, and provide ways to interpolate FESOM data onto a regular grid. There is also a documentation with description of the command line tools and a set of Jupyter Notebooks with examples of library usage. Moreover, we have developed the R package spheRlab (https://github.com/FESOM/spheRlab) which facilitates the analysis and visualization of unstructured-mesh data, including functions for the generation of grid description files that enable full compatibility of FESOM data with the more widely used Climate Data Operators (CDO; https://code.mpimet.mpg.de/projects/cdo).

Furthermore, the unstructured mesh files can be further processed with simple cdo commands. To give an example, a typical file is the unstructured potential ocean temperature in *thetao_Omon_AWI-CM-1-1-MR_historical_r1i1p1f1_gn_185001-185012.nc*. A horizontal conservative remapping of this file to a 1°x1° regular grid (“r360x180”) from the command line is a one-liner and as simple as

cdo remapycon,r360x180 thetao_Omon_input.nc thetao_Omon_remapped2D.n

where we shortened the input file name to *thetao_Omon_input.nc* for brevity. An analogous command results in a bilinear interpolation by replacing “remapycon” with “remapbil”. If many years need to be remapped, it is advisable to generate the interpolation weights only once *(cdo genycon,r360x180 thetao_Omon_input.nc weights_unstr_2_r360x180.nc*) and to re-use them for all subsequent remapping commands:

cdo remap,r360x180,weights_unstr_2_r360x180.nc thetao_Omon_input.nc thetao_Omon_remapped2D.nc

Furthermore, additional vertical interpolation to a set of different depth levels can be achieved via, e.g.,

cdo intlevel,10,20,30,50,75,100,125,150,200,250,300,400,500,600,700,800,900,1000,1100,1200,1300,1400,1500, 1750,2000,2500,3000,3500,4000,4500,5000,5500 thetao_Omon_remapped2D.nc thetao_Omon_remapped3D.nc

A2 Reference observation data and computation of objective performance indices

As reference data for the computation of the objective performance indices, various observation and reanalysis data are selected: For the following atmospheric variables, the ERA-40 reanalysis data are used: 2 m temperature (t2m), 10 m u wind component (u10m), 10 m v wind component (v10m), 500 hPa geopotential height (z500), and 300 hPa u component (u300). This is augmented by the following data: CERES for top of atmosphere outgoing longwave radiation (TOA, Loeb et al., 2012), GPCP for precipitation (pr, Huffman et al., 2009), MODIS for total cloud cover (tcc, Platnick et al., 2003), and OSISAF for sea ice concentration (sic, Tonboe et al., 2016). For the ocean, Polar Science Center Hydrographic Climatology (PHC, updated from Steele et al., 2001) is used as a reference for both potential temperature and salinity.

The absolute error is computed for each grid cell and averaged over different regions. For the atmosphere the different regions are Arctic (60–90°N), northern mid-latitudes (30–60°N), tropics (30°S–30°N), southern mid-latitudes (30–60°S), Antarctic (60–90°S), and global. For the ocean the domain is split into the major ocean basins: Arctic Ocean, North Atlantic Ocean, North Pacific Ocean, Indian Ocean, South Atlantic Ocean, South Pacific Ocean, Southern Ocean. Like for the atmosphere, the global ocean is also considered globally in addition. The mean absolute error is computed for each season: for the atmosphere for the four seasons DJF, MAM, JJA, SON, and for the ocean for two seasons DJF and JJA. For the ocean, model data are vertically interpolated to the z-levels of the PHC. Errors are computed for each z-level of the climatology and averaged over the levels. Then the error is normalized with the mean absolute error averaged over a set of CMIP5 models. By doing this, the performance of our new CMIP6 model can be compared objectively using the performance of CMIP5 models in terms of agreement with observation data. A performance index of 1 indicates that the model performs as well as the average of the CMIP5 models; a performance index of smaller than 1 (larger than 1) indicates a better (worse) performance.

Acknowledgments and Data Availability Statement

We used the Niño3.4 index (Rayner et al., 2003) as provided by NOAA on their website at https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Nino34/. Data from the RAPID MOC monitoring project are funded by the Natural Environment Research Council and are freely available from www.rapid.ac.uk/rapidmoc. Furthermore we used ERA5 reanalysis data (Hersbach et al., 2020) provided at https://cds.climate.copernicus.eu/cdsapp#!/home. The simulations were performed at the German Climate Computing Center (DKRZ). We thank the Max-Planck-Institute for Meteorology in Hamburg (MPI) for providing us with the ECHAM6 model code and for the fruitful discussions with various colleagues from this institute. All data are available on the Earth System Grid Federation (ESGF) (Semmler et al., 2018). H. F. Goessling and L. Mu acknowledge the financial support of the Federal Ministry of Education and Research of Germany in the framework of the research group Seamless Sea Ice Prediction (SSIP; Grant 01LN1701A). N. Khosravi and C. Hinrichs have received funding from the European Union's Horizon 2020 Research and Innovation programme through grant agreement No. 727862 APPLICATE. This paper is a contribution to the projects S1 (Diagnosis and Metrics in Climate Models) and S2 (Improved parameterizations and numerics in climate models) of the Collaborative Research Centre TRR 181 "Energy Transfer in Atmosphere and Ocean" funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project no. 274762653. D. Sein was supported by EC Horizon 2020 project PRIMAVERA under the grant agreement no. 641727, the state assignment of FASO Russia theme No. R0149-2019-0015. Q. Wang and D. Sidorenko were supported by the Helmholtz Climate Initiative REKLIM (Regional Climate Change). 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

Alexander, M. A., I. Bladé, M. Newman, J. R. Lanzante, N. Lau, and J. D. Scott (2002), The Atmospheric Bridge: The Influence of ENSO Teleconnections on Air–Sea Interaction over the Global Oceans. *J. Climate, 15*, 2205–2231, https://doi.org/10.1175/1520-0442(2002)015<2205:TABTIO>2.0.CO;2

Armour, K., J. Marshall, J. Scott, et al. (2016), Southern Ocean warming delayed by circumpolar upwelling and equatorward transport, *Nature Geosci, 9,* 549–554, https://doi.org/10.1038/ngeo2731

Barnes, E. A., and L. M. Polvani (2015), CMIP5 Projections of Arctic Amplification, of the North American / North Atlantic circulation, and of their relationship, *Journal of Climate, 28,* 5254-5271, https://doi.org/10.1175/JCLI-D-14-00589.1

Biastoch, A., C. W. Böning, F. U. Schwarzkopf, and J. R. E. Lutjeharms (2009), Increase in Agulhas leakage due to poleward shift in the southern hemisphere westerlies, *Nature, 462,* 495-498, https://doi.org/10.1038/nature08519

Biastoch, A., D. V. Sein, J. V. Durgadoo, Q. Wang, and S. Danilov (2018), Simulating the Agulhas system in global ocean models – nesting vs. multi-resolution unstructured meshes, *Ocean Modelling, 121,* 117-131, https://doi.org/10.1016/j.ocemod.2017.12.002

Bintanja, R., et al. (2013), Important role for ocean warming and increased ice-shelf melt in Antarctic sea-ice expansion, *Nature Geoscience, 6,* 376-379, https://doi.org/10.1038/ngeo1767

Bintanja, R., and F. Selten (2014), Future increases in Arctic precipitation linked to local evaporation and sea-ice retreat, *Nature 509***,** 479–482, https://doi.org/10.1038/nature13259.

Blackport, R., and P. J. Kushner (2017), Isolating the Atmospheric Circulation Response to Arctic Sea Ice Loss in the Coupled Climate System, *Journal of Climate, 30,* 2163-2185, https://doi.org/10.1175/JCLI-D-16-0257.1

Cai, W., A. Santoso, G. Wang, G. et al. (2015), ENSO and greenhouse warming, *Nature Clim Change, 5,* 849–859, https://doi.org/10.1038/nclimate2743

Cattiaux, J., Y. Peings, D. Saint-Martin, N. Trou-Kechout, and S. J. Vavrus (2016), Sinuosity of midlatitude atmospheric flow in a warming world. *Geophysical Research Letters, 43,* 8259-8268, https://doi.org/10.1002/2016GL070309

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

Chen, G., P. Zhang, and J. Lu (2020), Sensitivity of the Latitude of the Westerly Jet Stream to Climate Forcing, *Geophysical Research Letters, 47,* https://doi.org/10.1029/2019GL086563

cmip6-cmor-tables https://github.com/PCMDI/cmip6-cmor-tables access date 2019/12/19

Cheng, W., J. C. H. Chiang, and D. Zhang (2013), Atlantic Meridional Overturning Circulation (AMOC) in CMIP5 Models: RCP and Historical Simulations, *J. Climate, 26,* 7187–7197, https://doi.org/10.1175/JCLI-D-12-00496.1

Copernicus Climate Change Service (C3S) (2017), ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus Climate Change Service Climate Data Store (CDS), *29.07.2020*, https://cds.climate.copernicus.eu/cdsapp#!/home

Curry, B., C. M. Lee, B. Petrie, R. E. Moritz, and R. Kwok (2014) Multiyear Volume, Liquid Freshwater, and Sea Ice Transports through Davis Strait, 2004–10, *J. Phys. Oceanogr., 44,* 1244-1266.

Danilov, S., G. Kivman, J. Schröter (2004), A finite-element ocean model: principles and evaluation, *Ocean Model, 6*, 125–150.

Ding, Q., A. Schweiger, M. L'Heureux, et al. (2017), Influence of high-latitude atmospheric circulation changes on summertime Arctic sea ice, *Nature Clim Change, 7,* 289–295, https://doi.org/10.1038/nclimate3241

Donohoe, A., K. C. Armour, A. G. Pendergrass, and D. S. Battisti (2014), Shortwave and longwave radiative contributions to global warming under increasing CO2, *Proceedings of the National Academy of Sciences, 111,* 16700-16705

Donohue, K. A., K. L. Tracey, D. R. Watts, M. P. Chidichimo, and T. K. Chereskin (2016), Mean Antarctic circumpolar current transport measured in Drake passage, *Geophys. Res. Lett., 43,* 11760-11767

England, M., A. Jahn, and L. Polvani (2019), Nonuniform contribution of internal variability to recent Arctic sea ice loss, *Journal of Climate*, *32,* 4039–4053, https://doi.org/10.1175/JCLI-D-18-0864.1

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

Fetterer, F., K. Knowles, W. N. Meier, M. Savoie, and A. K. Windnagel (2017), updated daily. Sea Ice Index, Version 3.0. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center, https://doi.org/10.7265/N5K072F8. [02.04.2020].

Frölicher, T. L., J. L. Sarmiento, D. J. Paynter, J. P. Dunne, J. P. Krasting, and M. Winton (2015), Dominance of the Southern Ocean in Anthropogenic Carbon and Heat Uptake in CMIP5 Models, *J. Climate, 28*, 862–886, https://doi.org/10.1175/JCLI-D-14-00117.1

GISS (2019), GISS surface temperature analysis GISTEMP v4. Available online at https://data.giss.nasa.gov/gistemp/.

Golaz, J.-C., et al. (2019), The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution, *Journal of Advances in Modeling Earth Systems, 11,* 2089-2129, https://doi.org/10.1029/2018MS001603

Good, S. A., M. J. Martin and N. A. Rayner (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

Gordon, A., J. Sprintall, H. van Aken, D. Susanto, S. Wijffels, R. Molcard, A. Ffield, W. Pranowo, and S. Wirasantosa (2010), The Indonesian throughflow during 2004–2006 as observed by the INSTANT program, *Dynam. Atmos. Oceans, 50,* 115-128, https://doi.org/10.1016/j.dynatmoce.2009.12.002

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

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

Grosfeld, K., R. Treffeisen, J. Asseng, A. Bartsch, B. Bräuer, B. Fritzsch, R. Gerdes, S. Hendricks, W. Hiller, G. Heygster, T. Krumpen, P. Lemke, C. Melsheimer, M. Nicolaus, R. Ricker, and M. Weigelt (2016), Online sea-ice knowledge and data platform <www.meereisportal.de>, Polarforschung, Bremerhaven, Alfred Wegener Institute for Polar and Marine Research & German Society of Polar Research, 85 (2), 143-155, https://doi.org/10.2312/polfor.2016.011

Haarsma R. J., M. J. Roberts, P. L. Vidale, C. A. Senior, A. Bellucci, Q. Bao, P. Chang, S. Corti, N. S. Fuckar, V. Guemas, J. von Hardenberg, W. Hazeleger, C. Kodama, T. Koenigk, L. R. Leung, J. Lu, J.-J. Luo, J. Mao, M. S. Mizielinski, R. Mizuta, P. Nobre, M. Satoh, E. Scoccimarro, T. Semmler, J. Small, and J.-S. von Storch (2016), High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, *Geoscientific Model Development, 9,* 4185-4208, https://doi.org/10.5194/gmd-9-4185-2016

Ham, Y.−G. (2017), A reduction in the asymmetry of ENSO amplitude due to global warming: The role of atmospheric feedback, *Geophys. Res. Lett., 44,* 8576–8584, https://doi.org/10.1002/2017GL074842

Hansen, J., R. Ruedy, M. Sato, and K. Lo (2010), Global surface temperature change, *Rev. Geophys., 48,* RG4004, https://doi.org/10.1029/2010RG000345

Hegewald, J. (2019), seamore - cmorize simulation data to a given CMIP6 data request, https://doi.org/10.5281/zenodo.3585711

Hersbach, H., et al. (2020), The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society,* https://doi.org/10.1002/qj.3803

Hirschi, J. J.−M., B. Barnier, C. Böning, A. Biastoch, A. T. Blaker, A. Coward, et al. (2020), The Atlantic meridional overturning circulation in high−resolution models. *Journal of Geophysical Research: Oceans*, *125,* e2019JC015522, https://doi.org/10.1029/2019JC015522

Huffman, G. J., R. F. Adler, D. T. Bolvin, and G. Gu (2009), Improving the global precipitation record: GPCP version 2.1. *Geophysical Research Letters 36, L17808*

IPCC (2013), 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, 1535 pp.

IPCC (2014), Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Juckes, M., K. E. Taylor, P. J. Durack, B. Lawrence, M. S. Mizielinski, A. Pamment, J. Y. Peterschmitt, M. Rixen, and S. Sénési (2020), The CMIP6 data request (DREQ, version 01.00.31), *Geosci. Model Dev., 13,* 201-224, https://doi.org/10.5194/gmd-13-201-2020

Jung, T., M. J. Miller, T. N. Palmer, P. Towers, N. Wedi, D. Achuthavarier, J. M. Adams, E. L. Altshuler, B. A. Cash, J. L. Kinter III, L. Marx, C. Stan, and K. I. Hodges (2012), High-resolution global climate simulations with the ECMWF model in Project Athena: Experimental design, model climate, and seasonal forecast skill, *J. Climate*, *25,* 3155–3172, https://doi.org/10.1175/JCLI-D-11-00265.1

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

Kageyama, M., P. Braconnot, S. P. Harrison, A. M. Haywood, J. H. Jungclaus, B. L. Otto-Bliesner, J.-Y. Peterschmitt, A. Abe-Ouchi, S. Albani, P. J. Bartlein, C. Brierley, M. Crucifix, A. Dolan, L. Fernandez-Donado, H. Fischer, P. O. Hopcroft, R. F. Ivanovic, F. Lambert, D. J. Lunt, N. M. Mahowald, W. R. Peltier, S. J. Phipps, D. M. Roche, G. A. Schmidt, L. Tarasov, P. J. Valdes, Q. Zhang, and T. Zhou (2018), The PMIP4 contribution to CMIP6 - Part 1: Overview and over-arching analysis plan. *Geosci. Model Dev., 11,* 1033–1057, https://doi.org/10.5194/gmd-11-1033-2018

Kay, J. E., M. M. Holland, and A. Jahn (2011), Inter-annual to multi-decadal Arctic sea ice trends in a warming world. *Geophysical Research Letters, 38,* L15708, https://doi.org/10.1029/2011GL048008

Keil, P., Mauritsen, T., Jungclaus, J. *et al.* (2020), Multiple drivers of the North Atlantic warming hole, *Nat. Clim. Chang., 10,* 667–671, https://doi.org/10.1038/s41558-020-0819-8

Knutti, R., D. Masson, and A. Gettelman (2013), Climate model genealogy: Generation CMIP5 and how we got there, *Geophys. Res. Lett., 40,* 1194– 1199, https://doi.org/10.1002/grl.50256

Korn, P. (2017), Formulation of an Unstructured Grid Model for Global Ocean Dynamics, *Journal of Computational Physics, 339,* 525-552, https://doi.org/10.1016/j.jcp.2017.03.009

Large, W. G., J. C. McWilliams, S. C. Doney (1994), Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. *Rev Geophys, 32,* 363–403

Lenderink, G., A. Buishand, and W. van Deursen (2007), Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. *Hydrol. Earth Syst. Sci., 11,* 1145-1159

Lenssen, N., G. Schmidt, J. Hansen, M. Menne,A. Persin,R. Ruedy, and D. Zyss (2019), Improvements in the GISTEMP uncertainty model, *J. Geophys. Res. Atmos., 124,* 6307-6326, https://doi.org/10.1029/2018JD029522

Loeb N. G., S. Kato, W. Su, T. Wong, F. G. Rose, D. R. Doelling, J. N. Norris, and X. Huang (2012), Advances in understanding top-of-atmosphere radiation variability from satellite observations, *Surv Geophys, 33*, 359-385

Lyu, K., X. Zhang, and J. A. Church (2020), Regional Dynamic Sea Level Simulated in the CMIP5 and CMIP6 Models: Mean Biases, Future Projections, and Their Linkages, *J. Climate*, *33,* 6377–6398, https://doi.org/10.1175/JCLI-D-19-1029.1

Masson, D., and R. Knutti, (2011), Climate model genealogy, *Geophys. Res. Lett., 38,* L08703, https://doi.org/10.1029/2011GL046864

McDonagh, E. L., B. A. King, H. L. Bryden, P. Courtois, Z. Szuts, M. Baringer, S. A. Cunningham, C. Atkinson, and G. McCarthy (2015), Continuous estimate of Atlantic Oceanic freshwater flux at 26.5 N, *J. Climate, 28,* 8888-8906, https://doi.org/10.1175/JCLI-D-14-00519.1

Mauritsen, T., et al. (2012), Tuning the climate of a global model, *J. Adv. Model. Earth Syst., 4,* M00A01, https://doi.org/10.1029/2012MS000154

Mauritsen, T., et al. (2019), Developments in the MPI−M Earth System Model version 1.2 (MPI−ESM1.2) and Its Response to Increasing CO2, *J. Adv. Model. Earth Syst.,* https://doi.org/10.1029/2018MS001400

Meehl, G. A., C. A. Senior, V. Eyring, G. Flato, J.F. Lamarque, R. J., Stouffer, K. E. Taylor, and M. Schlund, (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

Meinshausen, M., Z. Nicholls, J. Lewis, M. J. Gidden, E. Vogel, M. Freund, U. Beyerle, C. Gessner, A. Nauels, N. Bauer, J. G. Canadell, J. S. Daniel, A. John, P. Krummel, G. Luderer, N. Meinshausen, S. A. Montzka, P. Rayner, S. Reimann, S. J. Smith, M. van den Berg, G. J. M. Velders, M. Vollmer, and H. J. Wang (2019), The SSP greenhouse gas concentrations and their extensions to 2500, *Geosci. Model Dev. Discuss.,* https://doi.org/10.5194/gmd-2019-222

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

Meredith, M., P. L. Woodworth, T. K. Chereskin, D. P. Marshall, L. C. Allison, G. R. Bigg, K. Donohue, K. J. Heywood, C. W. Hughes, A. Hibbert, A. M. Hogg, H. L. Johnson, L. Jullion, B. A. King, H. Leach, Y.-D. Lenn, M. A. Morales-Maqueda, D. R. Munday, A. C. Naveira-Garabato, C. Provost, J.-B. Sallée, and J. Sprintall (2011), Sustained monitoring of the Southern Ocean at Drake Passage: past achievements and future priorities, *Rev. Geophys., 49,* L05603, https://doi.org/10.1029/2010RG000348

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

O'Neill B.C., C. Tebaldi, D. P. van Vuuren, V. Eyring, P. Friedlingstein, G. Hurtt, R. Knutti, E. Kriegler, J.-F. Lamarque, J. Lowe, G. A. Meehl, R. Moss, K. Riahi, and B. M. Sanderson (2016), The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geosci. Model Dev., 9,* 3461–3482, https://doi.org/10.5194/gmd-9-3461-2016

Notz, D., and SIMIP community (2020), Arctic Sea Ice in CMIP6. *Geophysical Research Letters,* https://doi.org/10.1029/2019GL086749

Petersen, M. R., X. S. Asay−Davis, A. S. Berres, Q. Chen, N. Feige, M. J. Hoffman, et al. (2019), An evaluation of the ocean and sea ice climate of E3SM using MPAS and interannual CORE−II forcing. J*ournal of Advances in Modeling Earth Systems, 11,* 1438– 1458, https://doi.org/10.1029/2018MS001373

Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riedi, and R. A. Frey (2003), The MODIS cloud products: algorithms and examples from terra, *IEEE Trans Geosci Remote Sens, 41,* 459-473

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

Rackow, T., and S. Juricke (2020), Flow−dependent stochastic coupling for climate models with high ocean−to−atmosphere resolution ratio. *Q J R Meteorol Soc., 146,* 284-300, https://doi.org/10.1002/qj.3674

Rackow T., D. V. Sein, T. Semmler, S. Danilov, N. V. Koldunov, D. Sidorenko, Y. Wang, T. Jung (2019), Sensitivity of deep ocean biases to horizontal resolution in prototype CMIP6 simulations with AWI-CM1.0, *Geosci. Model Dev., 12,* 2635–2656, https://doi.org/10.5194/gmd-12-2635-2019

Rackow, T. , D. Sidorenko, H. F. Goessling, A. Timmermann, and T. Jung (2014), Modeling ENSO with ECHAM6-FESOM: Influence of the ocean resolution, *Ocean Sciences Meeting*

2014, Hawaii Convention Center, Honolulu, USA, 23 February 2014 - 28 February 2014, https://doi.org/10.13140/2.1.2537.9209

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *J. Geophys. Res., 108 (D14),* 4407, https://doi.org/10.1029/2002JD002670

Reichler, T., J. Kim (2008), How well do coupled models simulate today's climate? *Bull Am Meteorol Soc 89,* 303–311

Ricker, R., S. Hendricks, L. Kaleschke, X. Tian-Kunze, J. King, and C. Haas (2017), A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data. *The Cryosphere, 11,* 1607–1623, https://doi.org/10.5194/tc-11-1607-2017

Ridderinkhof, H., P. van der Werf, J. Ullgren, H. van Aken, P. van Leeuwen, and W. de Ruijter (2010), Seasonal and interannual variability in the Mozambique Channel from moored current observations, *J. Geophys. Res., 115,* C06010, https://doi.org/10.1029/2009JC005619

Roach, A., K. Aagard, C. Pease, S. Salo, T. Weingartner, V. Pavlov, and M. Kulakov (1995), Direct measurements of transport and water properties through Bering Strait, *J. Geophys. Res., 100,* 18443-18457.

Roach, L. A., J. Dörr, C. R. Holmes, F. Massonnet, E. W. Blockley, D. Notz, T. Rackow, M. N. Raphael, S. P. O'Farrell, D. A. Bailey, and C. M. Bitz (2020), Antarctic Sea Ice Area in CMIP6, *Geophysical Research Letters, 47,* e2019GL086729, https://doi.org/10.1029/2019GL086729

Schauer, U., A. Beszczynska Moeller, W. Walczowski, E. Fahrbach, J. Piechura, and E. Hansen (2008), Variation of measured heat flow through the Fram Strait between 1997 and 2006, in: Arctic-Subarctic Ocean Fluxes: Defining the Role of the Northern Seas in Climate, edited by: Dickson, R., Springer, 65-85.

Sein, D. V., S. Danilov, A. Biastoch, J. V. Durgadoo, D. Sidorenko, S. Harig, and Q. Wang (2016), Designing variable ocean model resolution based on the observed ocean variability, *Journal of Advances in Modeling Earth Systems, 8,* 904-916.

Sein, D. V., N. V. Koldunov, S. Danilov, Q. Wang, D. Sidorenko, I. Fast, T. Rackow, W. Cabos, and T. Jung (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

Sein, D. V. , N. V. Koldunov, S. Danilov, D. Sidorenko, C. Wekerle, W. Cabos, T. Rackow, P. Scholz, T. Semmler, Q. Wang, and T. Jung (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

Semmler, T., S. Danilov, T. Rackow, D. Sidorenko, J. Hegewald, D. Sein, Q. Wang, T. Jung (2017), AWI AWI-CM 1.1 HR model output prepared for CMIP6 HighResMIP, *Earth System Grid Federation,* http://cera-www.dkrz.de/WDCC/meta/CMIP6/CMIP6.HighResMIP.AWI.AWI-CM-1-1-HR

Semmler, T., S. Danilov, T. Rackow, D. Sidorenko, D. Barbi, J. Hegewald, D. Sein, Q. Wang, T. Jung (2018), AWI AWI-CM 1.1 MR model output prepared for CMIP6 CMIP, version 20191219, *Earth System Grid Federation,* https://doi.org/10.22033/ESGF/CMIP6.359

Semmler, T., E. Manzini, D. Matei, H. K. Pradhan, T. Jung (2019), AWI AWI-CM 1.1 MR model output prepared for CMIP6 PAMIP pdSST-pdSIC, version 20191219, *Earth System Grid Federation,* https://doi.org/10.22033/ESGF/CMIP6.12040

Sidorenko, D., T. Rackow, T. Jung, T. Semmler, D. Barbi, S. Danilov, K. Dethloff, W. Dorn, K. Fieg, H. F. Goessling, D. Handorf, S. Harig, W. Hiller, S. Juricke, M. Losch, J. Schröter, D. V. Sein, Q. Wang, (2015), Towards multi-resolution global climate modeling with ECHAM6–FESOM. Part I: model formulation and mean climate, *Climate Dynamics, 44,* 757-780.

Sidorenko D., N. V. Koldunov, Q. Wang, S. Danilov, H. F. Goessling, O. Gurses, P. Scholz, D. V. Sein, E. Volodin, C. Wekerle, T. Jung (2018), Influence of a salt plume parameterization in a coupled climate model, *Journal of Advances in Modeling Earth Systems, 10,* https://doi.org/10.1029/2018MS001291

Smedsrud, L. H., R. Ingvaldsen, J. E. Ø Nilsen, and Ø. Skagseth (2010), Heat in the Barents Sea: transport, storage, and surface fluxes, *Ocean Sci., 6,* 219-234.

Smeed, D., B. Moat, D. Rayner, W. Johns, M. Baringer, D. Volkov, and E. Frajka-Williams (2019), Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS (RAPID-Meridional Overturning Circulation and Heatflux Array-Western Boundary Time Series) array at 26N from 2004 to 2018, https://doi.org/10.5285/8cd7e7bb-9a20-05d8-e053-6c86abc012c2

Smith D. M., J. A. Screen, C. Deser, J. Cohen, J. C. Fyfe, J. Garcia-Serrano, T. Jung, V. Kattsov, D. Matei, R. Msadek, Y. Peings, M. Sigmond, J. Ukita, J.-H. Yoon, X. Zhang (2018), The Polar Amplification Model Intercomparison Project (PAMIP) contribution to CMIP6: investigating the causes and consequences of polar amplification, *Geoscientific Model Development Discussions,* https://doi.org/10.5194/gmd-2018-82

Spreen, G., L. Kaleschke, and G.Heygster (2008), Sea ice remote sensing using AMSR-E 89 GHz channels, *J. Geophys. Res., 113,* C02S03, doi:10.1029/2005JC003384

Steele, M., R. Morley, and W. Ermold (2001), PHC: A Global Ocean hydrography with a high-quality Arctic Ocean, *Journal of Climate, 14,* 2079-2087

Stevens, B., M. Giorgetta, M. Esch, T. Mauritsen, T. Crueger, S. Rast, M. Salzmann, H. Schmidt, J. Bader, K. Block, R. Brokopf, I. Fast, S. Kinne, L. Kornblueh, U. Lohmann, R. Pincus,

T. Reichler, and E. Roeckner (2013), Atmospheric component of the MPI-M earth system model: ECHAM6. *J Adv Model Earth Syst, 5*, 146–172

Stroeve, J., and D. Notz (2015), Insights on past and future sea-ice evolution from combining observations and models, *Global and Planetary Change, 135*, 119-132

Timmermann, A., S. An, J. Kug, et al. (2018), El Niño–Southern Oscillation complexity. *Nature, 559,* 535–545. https://doi.org/10.1038/s41586-018-0252-6

Tokarska, K. B., M. B. Stolpe, S. Sippel, E. M. Fischer, C. J. Smith, F. Lehner, and R. Knutti (2020), Past warming trend constrains future warming in CMIP6 models, *Science Advances, 6,* https://doi.org/10.1126/sciadv.aaz9549

Tonboe, R. T., S. Eastwood, T. Lavergne, A. M. Sørensen, N. Rathmann, G. Dybkjær, L. T. Pedersen, J. L. Høyer, and S. Kern (2016), The EUMETSAT sea ice concentration climate data record, *The Cryosphere, 10,* 2275–2290, https://doi.org/10.5194/tc-10-2275-2016

Trenberth, K. E., J. T. Fasullo, K. von Schuckmann, and L. Cheng (2016), Insights into Earth's energy imbalance from multiple sources. *Journal of Climate, 29,* 7495-7505, https://doi.org/10.1175/JCLI-D-16-0339.1

Turner, J., J. S. Hoskins, T. J. Bracegirdle, G. J. Marshall, and T. Phillips (2015), Recent changes in Antarctic sea ice, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373,* 20140163, https://doi.org/10.1098/rsta.2014.0163

de la Vara, A., W. Cabos, D. V. Sein, D. Sidorenko, N. V. Koldunov, S. Koseki, P. M. M. Soares, and S. Danilov (2020), On the impact of atmospheric vs oceanic resolutions on the representation of the sea surface temperature in the South Eastern Tropical Atlantic, *Clim Dyn*, https://doi.org/10.1007/s00382-020-05256-9

Wang, C., L. Zhang, S. K. Lee, L. Wu, and C. R. Mechoso (2014), A global perspective on CMIP5 climate model biases. *Nature Climate Change, 4,* 201-205, https://doi.org/10.1038/nclimate2118

Wang, Q., S. Danilov, D. Sidorenko, R. Timmermann, C. Wekerle, X. Wang, T. Jung, J. Schröter (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

Wang, Q., C. Wekerle, S. Danilov, X. Wang, and T. Jung (2018), A 4.5 km resolution Arctic Ocean simulation with the global multi-resolution model FESOM 1.4, *Geosci. Model Dev., 11,* 1229-1255.

Weaver, A. J., J. Sedláček, M. Eby, K. Alexander, E. Crespin, T. Fichefet, G. Philippon-Berthier, F. Joos, M. Kawamiya, K. Matsumoto, M. Steinacher, K. Tachiiri, K. Tokos, M. Yoshimori, K. Zickfeld (2012), Stability of the Atlantic meridional overturning circulation: A model intercomparison. *Geophys. Res. Lett., 39,* L20709.

Weijer, W., W. Cheng, S. S. Drijfhout, A. V. Fedorov, A. V. Hu, L. C. Jackson, W. Liu, E. L. McDonagh, J. V. Mecking, and J. Zhang (2019), Stability of the Atlantic Meridional Overturning Circulation: A Review and Synthesis, *Journal of Geophysical Research: Oceans, 124,* 5336-5375, https://doi.org/10.1029/2019JC015083

Wekerle, C., Q. Wang, S. Danilov, T. Jung, and J. Schröter (2013), The Canadian Arctic Archipelago throughflow in a multiresolution global model: Model assessment and the driving mechanism of interannual variability, *J. Geophys. Res. Oceans, 118,* 4525–4541, https://doi.org/10.1002/jgrc.20330

Wild, M. (2020), The global energy balance as represented in CMIP6 climate models, *Clim Dyn, 55,* 553–577, https://doi.org/10.1007/s00382-020-05282-7

Woodgate, R. A. (2018), Increases in the Pacific inflow to the Arctic from 1990 to 2015, and insights into seasonal trends and driving mechanisms from year-round Bering Strait mooring data, *Progress in Oceanography*, *160*, 124-154

Zampieri, L., and H. F. Goessling (2019), Sea ice targeted geoengineering can delay Arctic sea ice decline but not global warming, *Earth's Future, 7,* https://doi.org/10.1029/2019EF001230

Zappa, G., L. C. Shaffrey, and K. I. Hodges (2013), The ability of CMIP5 models to simulate North Atlantic extratropical cyclones, *Journal of Climate, 26,* 5379-5396, https://doi.org/10.1175/JCLID1200501.1

Zappa, G., and T. G. Shepherd (2017), Storylines of Atmospheric Circulation Change for European Regional Climate Impact Assessment, *Journal of Climate, 30,* 6561-6577, https://doi.org/10.1175/JCLI-D-16-0807.1