A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR)

W. A. Müller[1,2] (iD), J. H. Jungclaus[1] (iD), T. Mauritsen[1] (iD), J. Baehr[3] (iD), M. Bittner[1] (iD), R. Budich[1], F. Bunzel[1] (iD), M. Esch[1], R. Ghosh[1] (iD), H. Haak[1], T. Ilyina[1] (iD), T. Kleine[3], L. Kornblueh[1], H. Li[1] (iD), K. Modali[1], D. Notz[1] (iD), H. Pohlmann[1] (iD), E. Roeckner[1], I. Stemmler[1], F. Tian[1,4], and J. Marotzke[1] (iD)

[1]Max Planck Institute for Meteorology, Hamburg, Germany, [2]Deutscher Wetterdienst, Hamburg, Germany, [3]Institute of Oceanography CEN, University Hamburg, Germany, [4]National Centre for Atmospheric Science, Department of Meteorology, University of Reading, Reading, UK

Key Points:
- A higher-resolution version of MPI-ESM1.2 is presented, which has a well-balanced radiation budget and stable ocean circulation
- The higher atmospheric resolution improves North Atlantic storm tracks, blocking frequency, and NAO representation
- The higher computational costs remain manageable and enable studies of seasonal to decadal predictions and climate impacts

Abstract The MPI-ESM1.2 is the latest version of the Max Planck Institute Earth System Model and is the baseline for the Coupled Model Intercomparison Project Phase 6 and current seasonal and decadal climate predictions. This paper evaluates a coupled higher-resolution version (MPI-ESM1.2-HR) in comparison with its lower-resolved version (MPI-ESM1.2-LR). We focus on basic oceanic and atmospheric mean states and selected modes of variability, the El Niño/Southern Oscillation and the North Atlantic Oscillation. The increase in atmospheric resolution in MPI-ESM1.2-HR reduces the biases of upper-level zonal wind and atmospheric jet stream position in the northern extratropics. This results in a decrease of the storm track bias over the northern North Atlantic, for both winter and summer season. The blocking frequency over the European region is improved in summer, and North Atlantic Oscillation and related storm track variations improve in winter. Stable Atlantic meridional overturning circulations are found with magnitudes of ~16 Sv for MPI-ESM1.2-HR and ~20 Sv for MPI-ESM1.2-LR at 26°N. A strong sea surface temperature bias of ~5°C along with a too zonal North Atlantic current is present in both versions. The sea surface temperature bias in the eastern tropical Atlantic is reduced by ~1°C due to higher-resolved orography in MPI-ESM-HR, and the region of the cold-tongue bias is reduced in the tropical Pacific. MPI-ESM1.2-HR has a well-balanced radiation budget and its climate sensitivity is explicitly tuned to 3 K. Although the obtained reductions in long-standing biases are modest, the improvements in atmospheric dynamics make this model well suited for prediction and impact studies.

1. Introduction

The increase in resolution of coupled climate models results in improved atmospheric and oceanic dynamics and reduces biases in mean state and variability of many quantities. Atmospheric blocking and storm tracks, for example, are substantially improved by the increase in atmospheric and oceanic resolutions (Scaife et al., 2011; Zappa et al., 2013). On the other hand, increase in resolution of a coupled climate model often comes with unexpected problems such as a very weak Atlantic meridional overturning circulation (AMOC), which is an important quantity in the Atlantic-Arctic climate system (e.g., Polyakov et al., 2010). A stable coupled climate model can therefore only be obtained by consideration of both climate components, which must be well-balanced in the atmospheric radiation budget and the ocean circulation. The slowdown of the AMOC has been a major reason for the provision of the coupled Max Planck Institute Earth System Model (MPI-ESM) in a relatively low resolution for the Coupled Model Intercomparison Project phase 5 (CMIP5; Giorgetta et al., 2013; Jungclaus et al., 2013) and current seasonal and decadal predictions (Baehr et al., 2015; Marotzke et al., 2016). Here we describe the coupled MPI-ESM version 1.2 in a higher-resolution configuration (MPI-ESM1.2-HR), which maintains a stable coupled climate and contributes to the upcoming CMIP6 experiments (Eyring et al., 2016) and to the next generation of seasonal and decadal predictions.

The MPI-ESM1.2-HR employs twice the horizontal resolution in its atmospheric component (T127, ~100 km) compared to its precursors provided for CMIP5 (T63, ~200 km). The impact of increasing the resolution of the atmospheric component of MPI-ESM (European Centre/Hamburg, ECHAM) has been documented in several Atmospheric Model Intercomparison Project (AMIP) types of experiments (Hertwig et al., 2015; Roeckner

et al., 2006). Roeckner et al. (2006) investigate the systematic error convergence in ECHAM5 in different resolutions (T42–T159). In terms of the atmospheric mean state, they show that the mean-squared errors of the zonal mean climate and stationary waves are reduced with increasing horizontal resolution. The increase in horizontal resolution further results in a warming of the troposphere in the middle and high latitudes, confirming earlier works by Stratton (1999), and leads to an intensification and poleward shift of the jets. Hertwig et al. (2015) examine the errors in ECHAM6 at the horizontal resolutions T63, T127 (as used here), and T255 (~50 km). Normalized errors indicate a mean improvement of the mean state for various variables of 14% in T127 and 19% in T255 compared to the T63 resolution. The improvements vary dependent on regions and variables, but the results are consistent with other publications and basically show improved extratropical key variables. The mean state errors have also been reduced by ~30% from ECHAM5 to ECHAM6.3 due to improved physics. This is shown in a forthcoming manuscript describing the developments in the MPI-M ESM version 1.2.

The increase in atmospheric horizontal resolution also affects higher-order moments and variability. The error in variance of several variables is reduced by ~10% in T127 compared to the T63 resolution in AMIP-type experiments within ECHAM6 (Hertwig et al., 2015). Weather regimes over the North Atlantic, such as atmospheric blocking, are also improved with increasing atmospheric resolution (e.g., Berckmans et al., 2013; Dawson & Palmer, 2015). Atmospheric blocking is typically characterized by the reversal of the upper-level geopotential height gradient (e.g., Tyrilis & Hoskins, 2007), persists up to 20 days, and leads to extreme temperatures (Buehler et al., 2011). However, global climate models tend to underestimate the blocking frequency (Anstey et al., 2013; D'Andrea et al., 1998; Masato et al., 2013). Dawson and Palmer (2015) point out that the simulation of quasi-persistent regimes, such as atmospheric blocking, are unrealistic in CMIP5-like resolutions and require atmospheric resolutions typically used for numerical weather predictions. Berckmans et al. (2013) show that European blocking frequency is improved with increasing atmospheric resolution using AMIP-type experiments, owing to improved simulation of variability patterns and the mean state. They argue that the improvements stem from a better representation of the orography, and a positive effect of increasing resolution on transient eddy momentum forcing. Others point out that improved parameterization of orographic wave drag rather than increasing of resolution can alleviate typical circulation biases in low-resolution climate models (Pithan et al., 2016).

The combined increase in atmospheric and oceanic resolution of coupled climate models has been tested by many modeling centers, and basically the biases can be reduced in first and higher moments of different key quantities. Mean sea surface temperature (SST) bias in the North Atlantic is reduced by increasing the resolution in the atmospheric and ocean components of the Hadley Centre Global Environmental Model version 3 (Shaffrey et al., 2009). Here the horizontal resolutions of the atmospheric and oceanic components are increased to ~0.8° and 0.25°. A marked improvement in regional precipitation patterns is found by using higher model resolutions, for example, in the GFDL CM2.5 coupled model (50 km in the atmosphere, 28 km in the ocean, Delworth et al., 2012), and the Community Climate System Model (0.5° in the atmosphere, Gent et al., 2010). Others have investigated the use of high resolution in the ocean to assess the impact of ocean eddies on the mean state (e.g., Kirtman et al., 2012).

This paper serves as documentation of MPI-ESM1.2-HR, which is being prepared for CMIP6, and for further developments of the seasonal and decadal prediction systems. This model version is also compared with coupled model integrations of MPI-ESM1.2 in lower resolution (MPI-ESM1.2-LR), which is also used for CMIP6. The developments of the atmospheric component ECHAM6.3 and MPI-ESM1.2-LR are described in a forthcoming manuscript on MPI-ESM1.2. We start with the description of the coupled model, tuning process, and spin-up experiments (section 2), followed by an evaluation of a selected set of variables describing the mean states in the ocean (section 3). The ocean mean states are examined for temperature, salinity, sea ice, and prominent diagnostics for the large-scale circulation such as ocean gyres and transports through distinct passages, and the ocean biochemistry. A special emphasis is placed on the AMOC. The AMOC transports warm and saline water masses northward into high latitudes and is responsible for approximately a quarter of the global northward heat transport (Cunningham & Marsh, 2010; Ganachaud & Wunsch, 2000). As such a stable AMOC that is as close as possible to observational estimates is a basic requirement for mean state and variability estimates of the North Atlantic coupled climate system and its predictability. The atmospheric mean state is evaluated in section 4. Here the horizontal distribution of surface air temperature, precipitation, sea level pressure (SLP), and the vertical structure of temperature and zonal wind are examined. Special emphasis is placed on storm track climatology and blocking frequency, which is expected to be improved with increasing resolution, as mentioned above.

In section 5, two prominent modes of variability are discussed, the El Niño/Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO). ENSO is the dominant mode of interannual climate variability, and its representation in CMIP5 models has been documented in Bellenger et al. (2014). The models in CMIP5 show an improvement against CMIP3 for some basic characteristics, such as SST amplitude, life cycle, the location of surface temperature anomalies, and phase locking. Others, such as feedback mechanisms, are underestimated, and a key characteristic, the ENSO spectrum, shows large variation of frequency of occurrence among the models (Bellenger et al., 2014). Because ENSO is an important diagnostic for seasonal prediction, it is examined in more detail. The NAO is a further important climate diagnostic, with distinct signatures in the North Atlantic atmosphere-ocean coupling and impact on continental scale key parameters (for review see Hurrell et al., 2003; Pinto & Raible, 2012). The NAO describes about one third of total variance and thus is the most important mode on a seasonal timescale over the North Atlantic-European sector. Predicting the NAO is of special interest, and recently, progress has been made in achieving predictive skill (Butler et al., 2016; Dobrynin et al., 2018; Scaife et al., 2014). Therefore, the NAO is also described in more detail. Finally, a summary and a discussion are given in sections 6 and 7.

2. Model, Tuning and Spin-Up, Experiments and Data

2.1. Model Description

For MPI-ESM1.2-HR (hereafter called MPI-ESM-HR) we use the atmospheric model ECHAM6.3 coupled to the ocean model MPIOM version 1.6.2. The atmospheric component ECHAM6.3 dynamical core is spectrally resolved with a truncation at wavenumber 127 (~100 km). The model's physics are represented on a regular Gaussian grid in the horizontal and 95 vertical levels, with a relatively highly resolved stratosphere and a model lid at 0.01 hPa. This vertical resolution was also used in a mixed resolution for CMIP5 (MPI-ESM-MR, Giorgetta et al., 2013) and allows for a detailed description of tropical waves and an improvement in a wide range of stratospheric processes such as the quasi-biennial oscillation (Krismer et al., 2013). ECHAM6.3 builds upon the previous version ECHAM6.1 used for CMIP5, and a detailed description of the model physics and changes made are described in the forthcoming manuscript of MPI-ESM1.2. In the current version, the ECHAM6.3 radiation scheme has changed to a Monte Carlo independent column approximation (Pincus & Stevens, 2013). Furthermore, a stratocumulus parametrization is activated. Also, several bug fixes are implemented in the cloud cover scheme accounting for energy conservation in the atmosphere. Changes in ECHAM6.3 have led to several improvements in the bias structure, globally and on a regional scale.

For the ocean component, MPIOM is set up on a tripolar (TP) grid at nominal 0.4° horizontal resolution (TP04, Jungclaus et al., 2013). The TP grid merges a quasi-isotropic grid north of the equator, for which the singularity at the North Pole is replaced over Canada and Siberia, and a Mercator grid south of the equator. A similar grid is applied for eddy-resolving studies described by von Storch et al. (2012). TP04 can be classified as "eddy permitting." In the vertical, 40 unevenly spaced levels are used. The first 20 levels are distributed over the top 700 m. The bottom topography is described by a partial-step formulation (Wolff et al., 1997). The model's physics of the current version of MPIOM remain unchanged with respect to the CMIP5 version, and a detailed description is found in Jungclaus et al. (2013). An overview of the model resolutions and references is given in Table 1.

MPI-ESM1.2 includes an ocean and land carbon cycle, comprising the latest ocean biogeochemistry module Hamburg Model of the Ocean Carbon Cyle (HAMOCC; Ilyina et al., 2013; Paulsen et al., 2017) and the land surface scheme JSBACH (Reick et al., 2013). New components in HAMOCC are the sinking velocity of detritus as a function of depth according to Martin et al. (1987) and prognostic nitrogen fixers (Paulsen et al., 2017). The new sinking velocity of detritus increases linearly with depth, and it varies from 3.5 m/day in the upper 100 m to 77.5 m/day at 6,000 m. New components of JSBACH include a five-layer soil scheme compared to a bucket scheme in CMIP5 and an improved hydrology and soil model in the land biogeochemistry module (Goll et al., 2015). The JSBACH five-layer scheme accounts for a root zone, which is differentiated into several layers. Soil-moisture memory is increased and regionally enlarged during dry seasons (Hagemann & Stacke, 2015; Stacke & Hagemann, 2016). The five-layer scheme led to improved seasonal predictions of the European summer climate (Bunzel et al., 2017). The dynamical vegetation is turned off in MPI-ESM-HR but turned on in MPI-ESM1.2-LR (hereafter called MPI-ESM-LR). The river runoff is calculated by a horizontal discharge model (Hagemann & Dümenil-Gates, 2003).

Table 1
Model Setup of MPI-ESM1.2 (CMIP6), MPI-ESM-1.1 (CMIP5), Their Atmospheric and Oceanic Components, and Respective Resolutions and References

	MPI-ESM1.2-HR	MPI-ESM1.2-LR	MPI-ESM.1.0-LR
Atmosphere	ECHAM6.3	ECHAM6.3	ECHAM6.1
Horizontal resolution	T127 (~100 km)	T63 (~200 km)	T63 (~200 km)
Vertical resolution	L95	L47	L47
Reference			Stevens et al. (2013)
Ocean	MPIOM 1.6.3	MPIOM 1.6.3	MPIOM 1.0.0
Horizontal resolution	0.4°	Nominal 1.5°	Nominal 1.5°
Vertical resolution	L40	L40	L40
Reference	Jungclaus et al. (2013)	Jungclaus et al. (2013)	Jungclaus et al. (2013)
Reference (coupled)			Giorgetta et al. (2013)

Note. MPI-ESM = Max Planck Institute Earth System Model; CMIP = Coupled Model Intercomparison Project.

The oceanic and atmospheric components are coupled without flux adjustment through the Ocean-Atmosphere-Sea-Ice coupler version 4. The coupling frequency is increased to 1 hr, in contrast to the CMIP5 version where the coupling frequency was 1 day. A higher coupling frequency improves the diurnal cycle of convection in the tropical Pacific and enhances the higher SST that appears during El Niño events. This has an effect on the skewness of the distribution of ENSO events (Tian et al., 2017). The increase in resolution in MPI-ESM-HR results in a higher demand of the computer resources compared to MPI-ESM-LR. MPI-ESM-HR currently has a throughput of ~15 model years per day and thus makes it possible to perform climate simulations. For this, however, the model is run on 106 nodes (each with 36 cores). MPI-ESM-LR is considerably faster and requires only 16 nodes for a throughput of ~50 model years per day. The data storage increases by a factor of 5 from MPI-ESM-LR to MPI-ESM-HR considering 6-hourly model output. This is mainly due to the doubling of horizontal resolution and the number of vertical levels used form MPI-ESM-LR to MPI-ESM-HR.

It is noteworthy that the presented model versions are tuned with respect to the CMIP5 forcing, because CMIP6 forcing was not available when the tuning experiments were carried out. In fact, when the CMIP6 forcing became available and was adopted, the global mean temperature (GMT) increased and drifted away from the target of 13.8°C. For MPI-ESM-HR, the GMT increased by only 0.2°K. However, the top of the atmosphere shortwave incoming and longwave outgoing radiation were balanced, and no further changes appeared in the ocean circulation (e.g., ~18 Sv of mean AMOC strength at 30–40°N and 1,000-m depth). Therefore, no further tuning was required for MPI-ESM-HR with CMIP6 forcing. For MPI-ESM-LR the CMIP6 forcing led to large increase in GMT, and tuning was required to achieve the target values (this is shown in a forthcoming manuscript describing the MPI-ESM1.2).

2.2. Tuning and Spin-Up

Tuning of MPI-ESM-HR is done following Mauritsen et al. (2012) starting with the AMIP-type experiments of ECHAM6 for the period 1979–2005. This is relatively unproblematic, with only minor adjustments compared to ECHAM6 of MPI-ESM-LR. Tuned parameters in the cloud scheme are the critical relative humidity, the fall velocity of ice particles, and the threshold of cloud ice on the Bergeron-Findeisen process. The convection scheme is adapted through the conversion rate for cloud water to rain and the overshooting fraction. Also, the timescale of removal of convective available potential energy is changed. The actual values are given in Table 2. A correction of the orographic wave drag is not necessary, because Northern Hemispheric winds and surface pressure biases are already adjusted and are similar to the lower resolution.

The equilibrium climate sensitivity (ECS) of ECHAM6.3 is explicitly tuned to 3 K. This target is chosen to better match the observed centennial warming. A series of test experiments has been conducted with ECHAM6 to achieve this target. Within these experiments several parameters were identified to control the ECS, such as the entrainment rate of shallow convection, which made shallow clouds in the tropics more stable (this is shown in a forthcoming manuscript on tuning the climate sensitivity in MPI-ESM1.2). This parameter setting is carried over to the coupled models MPI-ESM-HR and MPI-ESM-LR, for which

Table 2
Specific Parameters and Their Active Values Used for the Tuning Procedure for MPI-ESM1.2-HR and MPI-ESM1.2-LR

Parameter	MPI-ESM1.2-HR	MPI-ESM1.2-LR
Critical relative humidity, surface (crs)	0.994	0.968
Fall velocity parameter (cvtfall)	3.0	2.5
WBF threshold (csecfrl)	1.0×10^{-5}	5.0×10^{-6}
Convection conversion rate for cloud water to rain (cprcon)	1.5×10^{-4}	2.5×10^{-4},
Convective overshooting fraction (cmfctop)	0.23	0.20
Timescale for removal of CAPE (cmftau)	3600	7200

Note. MPI-ESM = Max Planck Institute Earth System Model.

the ECS is about 3 K. Here global mean yearly mean surface temperature anomalies are regressed to surface solar and terrestrial radiation. The anomalies are calculated for an abrupt4xCO2 experiment subtracted from the piControl climatology.

After tuning ECHAM6, in total a 2,000-year spin-up with the coupled model of MPI-ESM-HR is performed with preindustrial conditions. This spin-up procedure is necessary to run the climate system in to radiative equilibrium and to reduce the drift in the ocean. The ocean component of MPI-ESM-HR is started from an already tuned version of MPI-ESM-LR. The Arctic sea ice volume appeared to be too large and, based on experience its magnitude, is tuned to reach an annual mean value of about 2×10^{13} m³. This is performed by changing the geometry of melting and freezing processes as described in Mauritsen et al. (2012). Two further changes are applied to the model during this spin-up. First, the dynamical vegetation module is turned off due to an increasing bias in the vegetation distribution. Second, the daily coupling is replaced by an hourly coupling. However, this change is not relevant for either the tuning of radiation equilibrium or the drift in the ocean. In HAMOCC, switching to hourly air-sea coupling produces a diurnal cycle of light, compared to daily mean light used in MPI-ESM-LR. This leads to a reduced phytoplankton growth. Therefore, the zooplankton grazing rate is adjusted to a lower value. The appropriate weathering rates, which are used to compensate the loss of carbon and nutrient from the water column to the sediment, were calculated according to diagnosis of the spin-up run.

The final steady state global mean surface air temperature exhibits no drift and stays on the target value of 13.8°C. The net of incoming shortwave and outgoing longwave radiation at the top of the atmosphere is 0.1 W/m² for MPI-ESM-HR and 0.4 W/m² for MPI-ESM-LR, for the spin-up and piControl experiments, and indicate a well-balanced radiation budget (Table 3).

2.3. Experiments and Data

The Diagnosis, Evaluation, and Characterization of Klima experiments recommended for CMIP (Eyring et al., 2016) are performed for MPI-ESM-HR. The experiments include an AMIP-type experiment, a preindustrial control run (piControl), 1% per year increase in carbon dioxide concentration (1pctCO$_2$), quadrupling of CO$_2$ (abrupt4xCO$_2$), and historical simulations with nineteenth and twentieth century forcing (historical). The forcing components are chosen from CMIP5, as by the time of writing CMIP6 forcing was not available. The forcing and experiment set ups are similar to the ones described in Giorgetta et al. (2013).

Table 3
Top of the Atmosphere Global Mean Shortwave and Longwave Radiation for MPI-ESM1.2-HR and MPI-ESM1.2-LR

W/m²	Spin-up		piControl		Historical	
	HR	LR	HR	LR	HR	LR
Shortwave radiation	242.0	240.0	242.0	240.0	241.5 ± 0.1	239.5 ± 0.2
Longwave radiation	−241.9	−239.6	−241.9	−239.6	−241.1 ± 0.1	238.7 ± 0.2
net	0.1	0.4	0.1	0.4	0.4	0.7

Note. Shown are values for spin-up and piControl experiments, ensemble mean of historical simulations (1979–2005), and standard deviations derived from respective ensemble members. Units are in W/m².

Table 4
Maximum Values of Time Mean Barotropic Stream Function in Selected Regions for MPI-ESM1.2-HR and MPI-ESM1.2-LR

Region	MPI-ESM1-2-LR	MPI-ESM1.2-HR	Observations	References
Subpolar gyre (North Atlantic)	29	34	26–40	Clarke (1984). Bersch (1995), Bacon (1997), Lherminier (2007), and Holliday et al. (2009)
Subtropical gyre (North Atlantic)	54	48	46–61	Johns et al. (1995)
Subtropical gyre (Pacific)	66	79	42 ± 2.5	Imawaki et al. (2001)

Note. Shown are values for piControl experiments. Also shown are observations with respective references. The range in observations result from different measurement campaigns southeast of Greenland (for Atlantic subpolar gyre) and 73°W and 68°W (for Atlantic subtropical gyre), and measurement error estimates (for Pacific subtropical gyre). Units are in Sverdrup (Sv).

Here we analyze the piControl and historical experiments. The piControl experiment is started directly from the end of the 2,000-year spin-up and run for a period of 500 years for MPI-ESM-HR and MPI-ESM-LR. The forcing is constant in time with orbital parameters set to 1850 values. Greenhouse gas and ozone concentrations are fixed at 1850 levels. No volcanic aerosol forcing is applied. Only tropospheric natural aerosol is used (see Giorgetta et al., 2013). For the historical experiments, an ensemble of five experiments is performed and each member is starting from piControl with a lag of 50 years. Each historical experiment is run from 1850 to 2005 under natural and anthropogenic forcing following CMIP5 (Taylor et al., 2012). The natural forcing consists of solar irradiance, seasonal varying tropospheric aerosols, stratospheric aerosols from volcanic eruptions, ozone, and greenhouse gases (details described in Giorgetta et al., 2013). The anthropogenic forcing includes CO_2 from fossil fuel burning (Giorgetta et al., 2013). For comparison, a similar ensemble (five members) is performed for MPI-ESM-LR. MPI-ESM-HR and MPI-ESM-LR are forced by prescribed atmospheric CO_2 concentrations without interactive carbon cycle, that is, the calculated air-sea and air-land CO_2 fluxes do not interactively alter the atmospheric CO_2 concentrations.

The historical warming in the past decades and century has been an objective in model development and tuning. Overall, the historical experiments of MPI-ESM-HR reproduce the warming trends during the twentieth century reasonably well as compared to HadCRUT data (shown in forthcoming manuscript on the development on MPI-ESM1.2). Notably is the warming in the midtwentieth century and stagnation during the 1950s and 1960s, which is comparable to observations. The trend from the 1970s onward is also well captured, though during the last hiatus period during the 2000s the simulations warm too rapidly. It also deserves mention that the model is consistently cooler during periods with strong volcanism. A feature that has been detected also in the previous version of MPI-ESM (see Giorgetta et al., 2013).

Vertical profiles of ocean salinity and temperature are evaluated against data from the Polar Science Center Hydrographic Climatology (Steele et al., 2001). Surface salinity and SST are compared against Levitus (Levitus et al., 1998). Transports through distinct passages and barotropic stream functions are compared with observational estimates for selected regions (see Tables 4 and 5). For AMOC, the model runs are compared with the RAPID array (Smeed et al., 2016). Note that for RAPID, AMOC is derived by adding the measured components of Ekman, Florida Current, and upper-mid-ocean transport (UMO), whereas in MPI-ESM-LR and MPI-ESM-HR UMO is not a direct model output and derived by AMOC minus Ekman minus Florida Current. Since the Strait of Florida is geographically not resolved in MPI-ESM-LR, the entire northward flowing western boundary current is integrated to calculate the Florida Current transport. For MPI-ESM-HR, the Strait of Florida is geographically resolved, and the UMO is also shown for calculating the Florida Current based on model output.

For the atmosphere, many variables are evaluated with respect to ERA-Interim (Dee et al., 2011), such as surface air temperature, SLP, zonal mean temperature and wind, storm tracks and blocking frequency, and the NAO variability. The precipitation is compared with Global Precipitation Climatology Project data (Adler et al., 2003). Additionally, HadISST1.1 (Rayner et al., 2003) is used to compare the observed and simulated ENSO.

In section 5 we analyze the transient eddy mean flow interactions as the scalar product of $E \cdot D$. E is the E vector from Hoskins et al. (1983) with the vector components $(v'^2 - u'^2)/2$ and $-u'v'$. u' and v' are the 2- to 6-day

Table 5
Time-Mean Transport Through Selected Passages for MPI-ESM1.2-HR and MPI-ESM1.2-LR and Observations With References

Section	MPI-ESM1.2-LR	MPI-ESM1.2-HR	Observations	References
Bering Strait	0.67	0.78	0.8 (range from 0.7 to 1.1)	Woodgate et al. (2006, 2012)
Denmark Strait	5.2	3.3	4.6	Hansen et al. (2008)
			3.4 ± 1.4	Jochumsen et al. (2012)
			3.2 ± 0.5	Jochumsen et al. (2017)
Iceland Scotland	5.3	4.4	4.8	Hansen et al. (2008)
			4.6 ± 0.25	Rossby and Flagg (2012)
Florida Current	—	15	31.7	Kanzow et al. (2010)
			31.6 ± 2.7	McDonagh et al. (2015)
Indonesian Throughflow	15.5	12.7	11.6–15.7	Gordon et al. (2010)
Mozambique Channel	21.8	13.8	5–26	DiMarco et al. (2002)
			16.7 ± 8.9	Ridderinkhof et al. (2010)
Drake Passage	171	165	134 ± 14	Nowlin and Klinck (1986)
			137 ± 8	Cunningham et al. (2003)
			136.7 ± 6.9	Meredith et al. (2011)
			173.3 ± 10.7	Donohue et al. (2016)

Note. Shown are values for piControl experiments. Units are in Sverdrup (Sv).

band-pass-filtered zonal and meridional winds. D is the deformation vector of the time mean flow with the components $(U_x - V_y)$ and $(V_x + U_y)$. U and V are the seasonal mean zonal and meridional winds. $E \cdot D$ describes the exchange of the kinetic energy between the eddy and the mean flow (Mak & Cai, 1989). Negative values of $E \cdot D$ denote that the eddy flow fields is providing kinetic energy to the mean flow and positive values denote the opposite.

For the ocean biogeochemical part, climatological mean observational data are used to evaluate the modeled distribution of respective variables. These data include ocean surface pCO$_2$ compiled by Takahashi et al. (2009), surface phosphate from World Ocean Atlas (Garcia et al., 2014), and total alkalinity (TA) from the Global Data Analysis Project version 2 (Lauvset et al., 2016; Olsen et al., 2016).

The variables of the ocean models are interpolated from the native nonregular grid to a 1° × 1° regular grid prior to the evaluation. Similarly, the observations and reanalyses are gridded with 1° x 1° resolution. For the evaluation of the atmospheric mean states, the reanalyses are interpolated on to the respective model grids.

3. Evaluation of Ocean Mean State

3.1. Ocean Temperature and Salinity and Sea Ice

Because of the long memory of the ocean, millennia-long spin-up runs are required to achieve a steady state of the coupled climate. However, spin-up runs and piControl are typically shorter, and drifts can prevail for many centuries. Multicentennial model drifts of ocean temperature and salinity are now examined, to see whether the ocean component is in an approximate steady state. The vertical profiles of trends of the GMT for the 500-year piControl experiments range between ±0.5 × 10^{-4} °C/year for both MPI-ESM-HR and MPI-ESM-LR (Figure 1a). This is a factor 500–1,000 smaller than the standard deviation of the year-to-year variations and indicates that the model has reached an approximate degree of steady state. For MPI-ESM-HR, the values are also smaller compared to the CMIP5 MPI-ESM-MR version, which has the same ocean grid. The values of MPI-ESM-LR are comparable to the CMIP5 MPI-ESM-LR version. The trends of the global mean salinity in the upper 1,500 m range between −1.0 × 10^{-5} and 1.5 × 10^{-5} g/kg per year, for both model versions (Figure 1b). This is also a factor 1,000 smaller than the standard deviation of the year-to-year variations. For layers below 1,500 m the drift of the salinity ranges within ±0.5 × 10^{-5} g/kg per year.

Vertical profiles of GMT and salinity biases are also shown in Figure 1. The GMT bias shows that MPI-ESM-HR is too cold in the upper layers with a bias of up to −1.7°C. Within the 500–1,500-m layer, temperature shows a warm bias in both model versions. In MPI-ESM-HR values reach 1.8°C and are about 0.5°C less than in MPI-ESM-LR. The bias further decreases in layers below 1,500 m and remains smaller in MPI-ESM-HR compared to MPI-ESM-LR. Different mesoscale eddy heat transport may contribute to a smaller temperature bias in
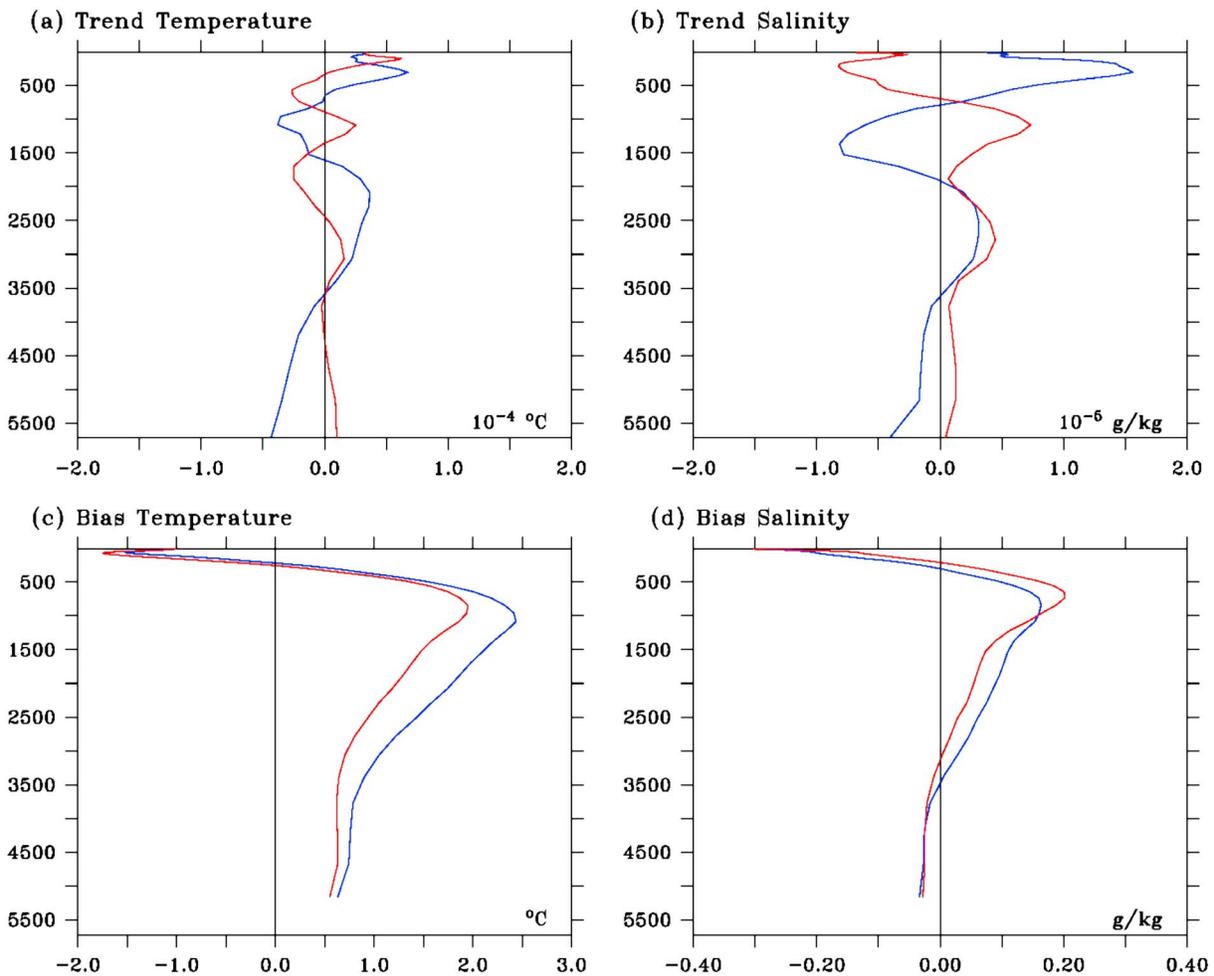
Figure 1. Vertical profiles of (a, b) trends and (c, d) mean bias of global mean (a, c) temperature and (b, d) salinity. Figure shows profiles from 500-year piControl experiments for MPI-ESM1.2-LR (blue) and MPI-ESM1.2-HR (red). The units are (a) 10^{-4} degree Celsius per year (°C/yr), (b) 10^{-5} g/kg per year, (c) degree Celsius (°C), and (d) g/kg. The biases in (c) and (d) are calculated with respect to PHC3.0 data.

the ocean interior (Griffies et al., 2015). Griffies et al. (2015) indicate that mesoscale eddies transport heat upward and thereby compensate downward heat transport by the mean. In particular, they show that a stronger mesoscale transient eddy advection results in a cooling of the ocean interior and thereby reduces the warm bias.

The distribution of the global mean salinity bias is similar to that in temperature. Negative biases are found in the upper layers with magnitudes of −0.3 g/kg. Maximum biases occur in the 500 to 1,000-m layer with values of up to 0.2 g/kg. A decrease of the bias occurs below 1,000 m. Similar bias structures were found in the previous versions of MPI-ESM (Jungclaus et al., 2013). Jungclaus et al. (2013) inferred that this reflects the inability of the model to simulate the correct water masses and noted that this is a prominent feature in many climate models.

Next we consider SST and salinity, which are directly related to atmosphere-ocean fluxes and thus are important quantities describing the coupled climate. Biases are calculated on the basis of the ensemble mean of historical runs and are compared against present-day estimates of Levitus and PCH3.0. Here the period 1979–2005 is considered, for which the historical runs overlap with the observational records. SST biases occur in the prominent areas in the North Atlantic, tropical Pacific cold-tongue and coastal upwelling regions (Figure 2). The SST bias structure does not change when a longer period is considered. Similar to previous versions of the model the largest bias is found in the North Atlantic at ~40°N. Here both model versions

(a) SST, MPI-ESM1.2-LR minus Levitus
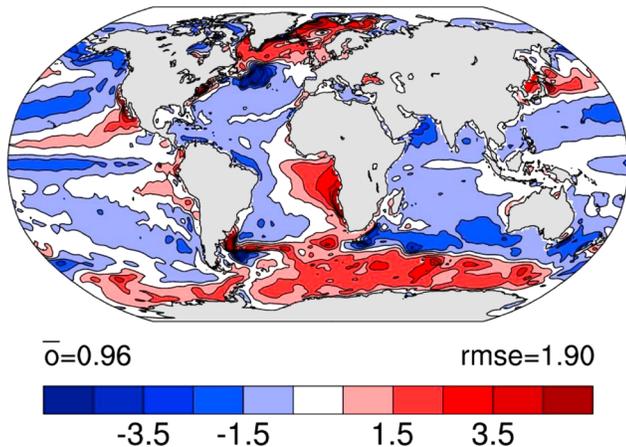
(b) SST, MPI-ESM1.2-HR minus Levitus
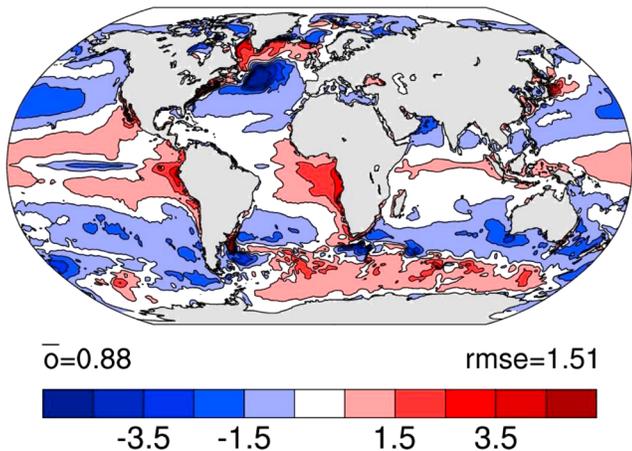
Figure 2. Sea surface temperature (SST) bias defined as (a) ensemble mean historical MPI-ESM1.2-LR minus Levitus and (b) ensemble mean historical MPI-ESM1.2-HR minus Levitus for the period 1979–2005. Units are in degree Celsius (°C). Numbers show mean bias and root-mean-squared error averaged over all grid points.

have deviations from observations with magnitudes up to 6°C. Both model versions further produce a too zonal North Atlantic Current as shown by the time mean sea surface salinity (Figure 3), which indicates that cold and fresh water masses from the Labrador Sea intrude too far southeastward compared to observations. Moreover, warm and saline water masses from the subtropics intrude too far north in the models compared to observations. There is no improvement of this bias in MPI-ESM-HR compared to MPI-ESM-LR. The cold-tongue bias in the equatorial Pacific persists in most climate models, though some improvements have been made during CMIP5 (Flato et al., 2013). The cold-tongue bias occurs in the MPI-ESM-HR version with magnitudes of ~2°C. However, the cold-tongue bias region is smaller in MPI-ESM-HR compared to MPI-ESM-LR, which may be related to a general warming in the Tropics in MPI-ESM-HR. This is consistent for all ensemble members. Also, the bias region does not cross the dateline in MPI-ESM-HR, whereas the cold-tongue bias region extends up to ~150°E in MPI-ESM-LR. This has implications for the skewness of ENSO variability as discussed in section 5.1. Other factors affecting the amplitude of ENSO variability are related to background vertical diffusivity and sharpness of equatorial thermocline (e.g., Meehl et al., 2001). The equatorial mean zonal SST gradient, as used for diagnostic, however, becomes stronger in MPI-ESM-HR compared to MPI-ESM-LR and HadISST and indicates a sharper thermocline in MPI-ESM-HR (not shown). Given the coupled nature of the tropical Pacific atmosphere-ocean climate, the cold-tongue bias also has an effect on precipitation patterns, wind stress, and formation of clouds, a point taken up in section 4.

The SST bias is large in the tropical upwelling regions with magnitudes of ~3°C near the South American and African coasts. However, the SST bias is reduced by about 1°C in the upwelling region near the African coast in MPI-ESM-HR compared to MPI-ESM-LR. This is consistent for all ensemble members. In a recent work, including an atmospheric model with even higher (T255) resolution, Milinski et al. (2016) has shown that an increase in atmospheric resolution and better resolved orography leads to an improvement of wind stress in the upwelling region and a reduction of the SST bias.

A negative SST bias occurs in the Nordic Seas in MPI-ESM-HR, in contrast to MPI-ESM-LR where a positive SST bias prevails. This is closely related to the Arctic sea ice. Maximum sea ice thickness is found in March in the Beaufort Sea and northern Greenland in both model versions (not shown). The magnitudes of the sea ice thickness are ~3 m and are similar to the PIOMAS reanalysis (Schweiger et al., 2011) and the CMIP5 version of the model (Jungclaus et al., 2013; Notz et al., 2013). However, the Atlantic late-winter sea ice extends far too south and reaches the northern Icelandic coast in MPI-ESM-HR compared to OSI SAF satellite data (Figure 4c). This bias is not present in MPI-ESM-LR, which shows a reduction of sea ice extension in the Greenland and Iceland Sea. This can explain the negative SST bias in MPI-ESM-HR and has implications for the Arctic-Atlantic freshwater exchange as shown in the next subsection. For Antarctic winter, MPI-ESM-HR shows a decrease of the bias of the of sea-ice fraction compared to MPI-ESM-LR, in the Ross and Weddell Seas, and from 0° to 180°E (Figures 4b and 4d). This may be affected by the decrease of SST bias in the Southern Ocean in MPI-ESM-HR.

3.2. Ocean Circulation

The large-scale ocean circulation is described by the barotropic stream function, transport through distinct passages, and the AMOC. Both model versions simulate the well-known features of the horizontal circulation (Table 4). The subpolar gyre and subtropical gyre in the North Atlantic show maximum values of 34 and 48 Sv, respectively, for MPI-ESM-HR and are within the range of observational estimates (26–40 Sv for the subpolar
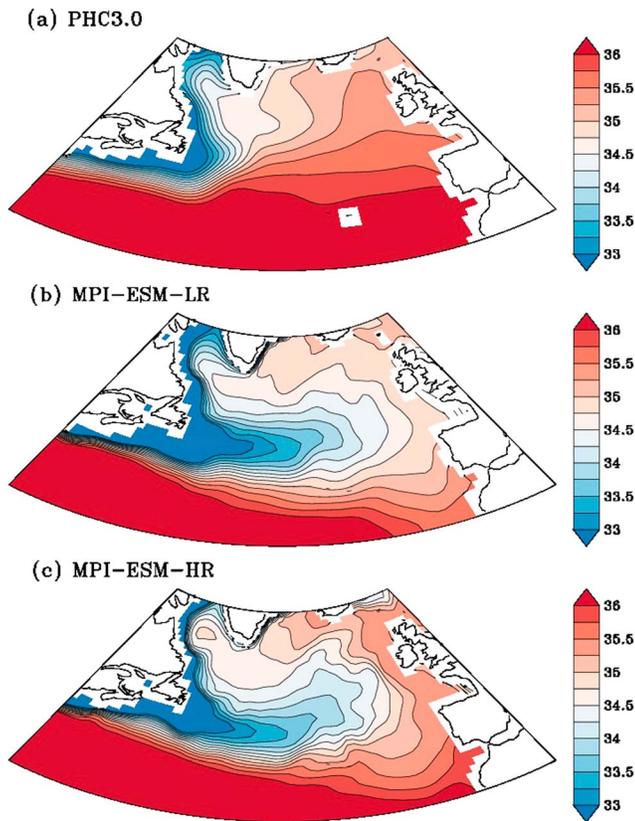
Figure 3. Time mean surface salinity for (a) PHC3.0 climatology and ensemble mean of (b) MPI-ESM1.2-LR and (c) MPI-ESM1.2-HR. For the models, the period 1979–2005 is considered. Units are in g/kg.

gyre strength at the southern tip of Greenland, Clarke, 1984; Bersch, 1995; Bacon, 1997; Lherminier et al., 2007; Holliday et al., 2009, and 46–61 Sv for the subtropical gyre strength, Johns et al., 1995). For MPI-ESM-LR, the maximum barotropic stream functions are 29 Sv for the subpolar gyre and 54 Sv for the subtropical gyre. Hence, both model versions represent the mean horizontal circulations in the North Atlantic reasonably well. For the Florida current, the increase in resolution in MPI-ESM-HR enables to resolve the Strait of Florida, and the Florida Current can be directly calculated. Though the value of the Florida Current is ~15 Sv and about half of the observations, the AMOC strength in the model is similar to observations (see below). This means that the reduced transport through the Strait of Florida is compensated by increased transport along the eastern coast of the Bahamas islands. For the subtropical Pacific, both model versions show larger gyre strength compared to observations. Both MPI-ESM-HR and MPI-ESM-LR overestimate the gyre strength with values of 79 and 66 Sv, respectively, whereas observational estimates reach values of 42 ± 2.5 Sv (Imawaki et al., 2001).

The transports through important passages are shown in Table 5. The transports via Bering Strait indicate values of 0.78 Sv for MPI-ESM-HR and 0.67 Sv for MPI-ESM-LR, which match observations (0.7–1.1 Sv, Woodgate, 2006, 2012). The Atlantic-Arctic exchange is shown by Denmark Strait transports and the transport over the Iceland-Scotland ridge. The Denmark Strait transport is 3.3 Sv in MPI-ESM-HR and comparable to recent observational estimates of 3.4 ± 1.4 Sv in Jochumsen et al. (2012) and 3.2 ± 0.5 Sv in Jochumsen et al. (2017). Hansen et al. (2008) suggest a higher value of 4.6 Sv. The transport over the Iceland-Scotland ridge is 4.4 Sv in MPI-ESM-HR, and also comparable to the observational values (4.8 Sv, Hansen et al., 2008, and 4.6 ± 0.25, Rossby & Flagg, 2012). For MPI-ESM-LR the values are higher compared to MPI-ESM-HR (5.2 Sv for Denmark Strait transport and 5.3 Sv for transport over the Iceland-Scotland ridge). The smaller transports in MPI-ESM-HR are related to the sea ice distribution, which extends too far south and reaches the Icelandic coast (Figure 4), and a reduced maximum mixed-layer depth and convection strength in the region north of Iceland (not shown).

Other transport estimates are shown for the Indonesian throughflow, which is an important characteristic of the thermohaline circulation, the Mozambique Channel, and the Drake Passage. In contrast to the CMIP5 version, transports through the Indonesian throughflow in MPI-ESM-HR and MPI-ESM-LR fit well to observational estimates. The transports show values of 12.7 and 15.5 Sv for MPI-ESM-HR and MPI-ESM-LR and are within the observational uncertainties (11.6–15.7 Sv, Gordon et al., 2010). The transports through the Mozambique Channel are increased from MPI-ESM-HR (13.8 Sv) to MPI-ESM-LR (21.8 Sv), but both fit in with the uncertainties in the observational estimates (5–26 Sv, DiMarco et al., 2002, and 16.7 ± 8.9 Sv, Ridderinkhof et al., 2010). For the Drake Passage, both model versions show relative large transports with values of 165 Sv for MPI-ESM-HR and 171 Sv for MPI-ESM-LR compared to observational estimates (134 ± 14Sv, Nowlin and Klinck, 1986; 137 ± 8 Sv, Cunningham et al., 2003; 136.7 ± 6.9 Sv, Meredith et al., 2011). However, a recent observational estimate reveals Drake Passage transport of about 173.3 ± 10.7 Sv (Donohue et al., 2016).

The mean AMOC is shown in Figure 5 and reveals the well-known single-cell structure indicative of a northward mass transport in the upper to intermediate layers and deepwater southward flow. Peak values are found for both versions at 30–40°N and ~1,000-m depth with magnitudes of ~18 Sv in MPI-ESM-HR and ~23 Sv in MPI-ESM-LR. With respect to the RAPID AMOC estimate (McCarthy et al., 2015; Smeed et al., 2016), MPI-ESM-HR shows both an improved time-mean vertical structure and an improved mean annual cycle compared to MPI-ESM-LR (Figures 5c and 5d). The time-mean of the AMOC at 26°N and 1,000 m is around 16 Sv in MPI-ESM-HR and comparable to 17 Sv of the RAPID AMOC estimate. The time-mean of the

(a) MPI-ESM-LR, March   **(b) MPI-ESM-LR, September**

(c) MPI-ESM-HR, March   **(d) MPI-ESM-HR, September**
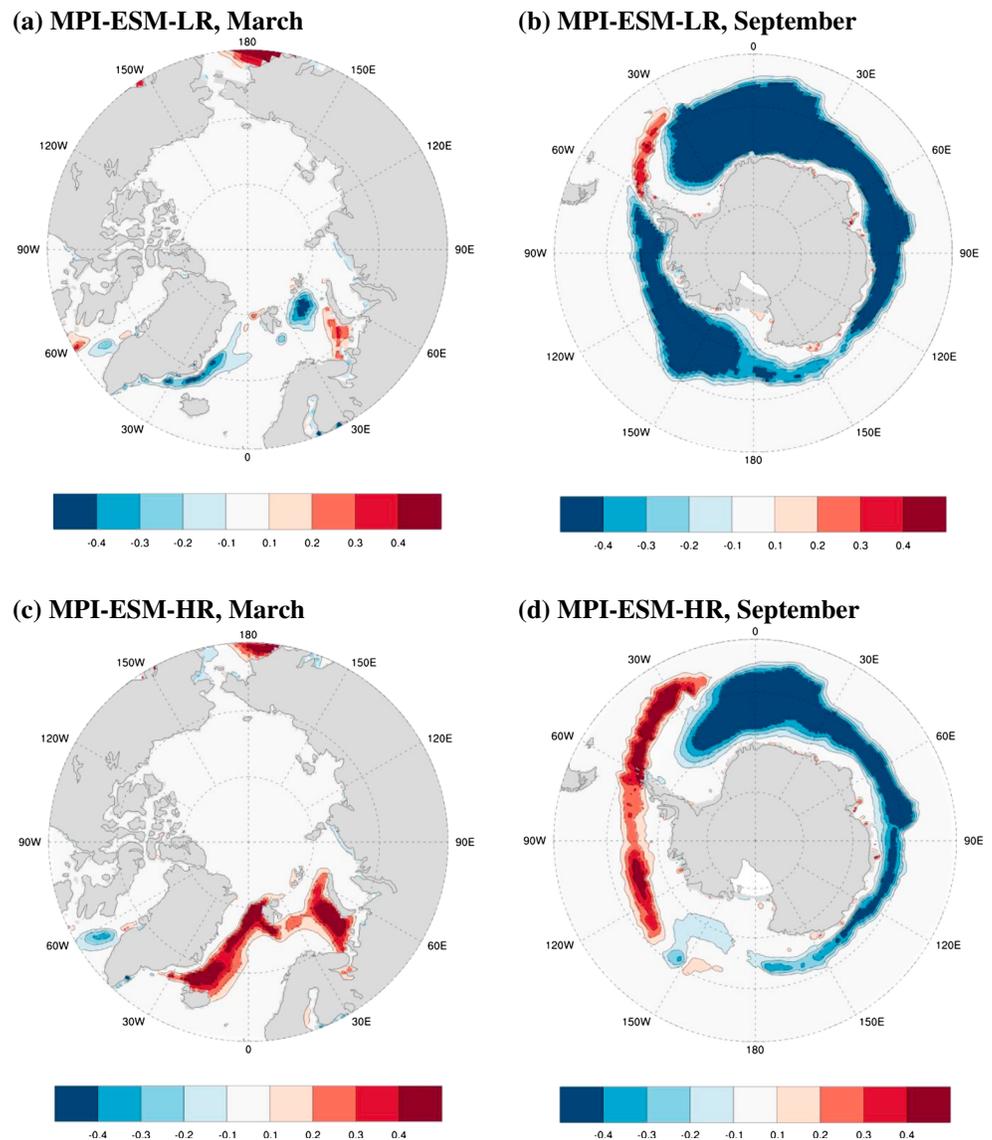
Figure 4. Bias of sea ice fraction for (a, c) Arctic in March and (b, d) Antarctic in September for (a, b) MPI-ESM1.2-LR and (c, d) MPI-ESM1.2-HR. Ensemble means of historical experiments for the period 1979–2005 are shown. Shown are the models with respect to OSI SAF satellite data. Units are scaled to 1.

AMOC in MPI-ESM-LR is about 19 Sv. The amplitude of the mean seasonal cycle is represented in both MPI-ESM-HR and MPI-ESM-LR, but the minimum in March is more pronounced in MPI-ESM-HR than in MPI-ESM-LR. While we find for the AMOC an overall agreement between the simulated and observed seasonal cycles in phase and magnitude, both MPI-ESM-HR and MPI-ESM-LR overestimate the minimum in spring and underestimate the maximum for fall in AMOC minus Ekman (Figure 5d).

A small but notable improvement in the mean seasonal cycle can be seen in the UMO (Figure 5e), where MPI-ESM-HR reproduces the occurrence of the maximum of the UMO in November (Figure 5e). This improvement of MPI-ESM-HR over MPI-ESM-LR is partly due to the separate resolution of the western boundary current (Figure 5e) and also due to a better resolution of the eastern boundary (Figure 5e), whose density variability is thought to be responsible for UMO's seasonal cycle (Chidichimo et al., 2010). Note that the discrepancy between the observed and simulated UMO seasonal cycle summer value depends on the number of years the RAPID time series is averaged over (Kanzow et al., 2010; Smeed et al., 2016). The overall differences
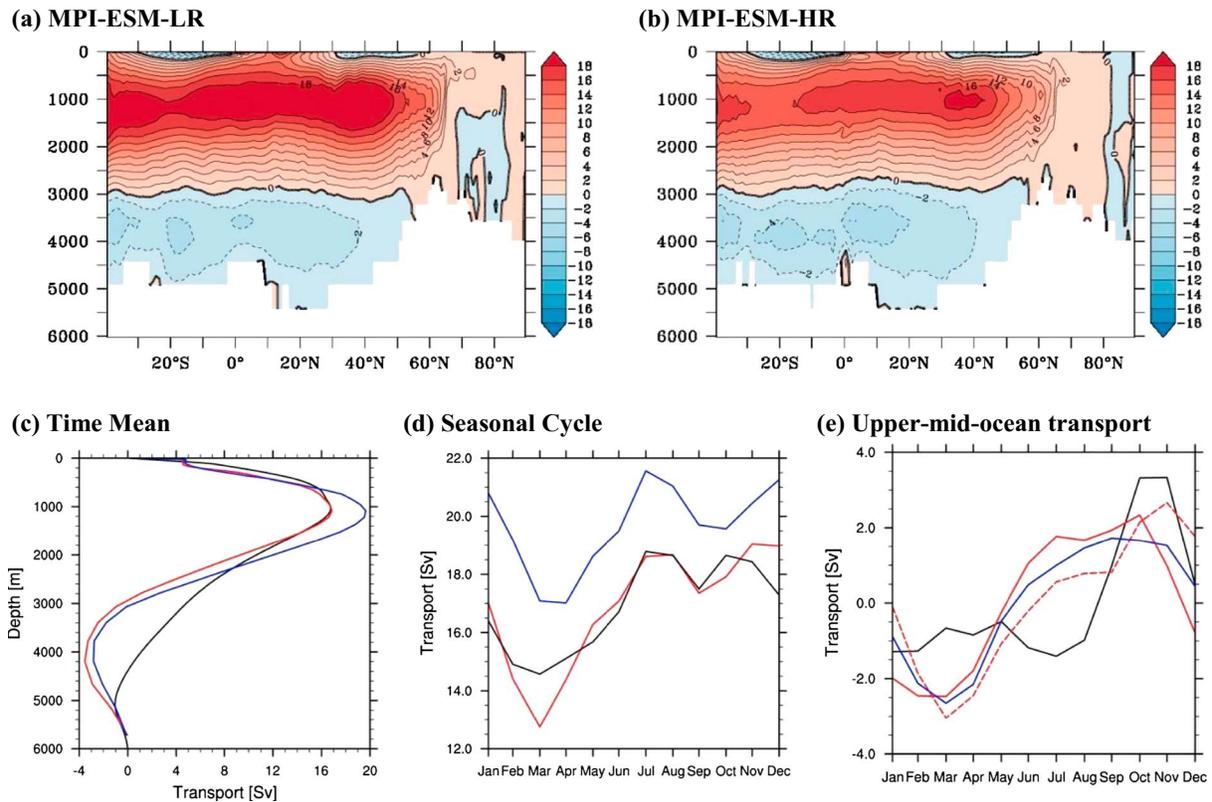
Figure 5. Time mean of AMOC for (a) MPI-ESM1.2-LR and (b) MPI-ESM1.2-HR piControl experiments. Further shown are (c) the time-mean of (black) RAPID AMOC time series (2004–2015), (red) MPI-ESM-HR, and (blue) MPI-ESM-LR control simulations, both using the AMOC at 26°N at 1,000 m. In (d) 26°N AMOC mean annual cycle is shown. Units are in Sverdrup (Sv). (e) Anomalies of upper-mid-ocean geostrophic transport. Shown is the transport for MPI-ESM-LR, for which the entire northward flowing western boundary current is integrated to calculate the Florida Current transport (solid blue line, and see methods). For comparison, a similar procedure is applied in MPI-ESM1.2-HR (solid red line). For MPI-ESM1.2-HR, the Strait of Florida is geographically resolved, and the upper-mid-ocean geostrophic transport is also shown for calculating the Florida Current directly (dashed red line). AMOC = Atlantic meridional overturning circulation.

between the observed and simulated UMO has been suggested to be a major limitation of the AMOCs multiyear predictability (Müller et al., 2016).

3.3. Ocean Biochemistry

Next we compare climatological mean states of ocean biochemistry in terms of surface $p\text{CO}_2$, phosphate, and TA in both model versions and observations (Figure 6). Ocean surface $p\text{CO}_2$ matches quite well with observations in both model versions. Maximum values are found in the tropical Pacific and minimum values are shown in the extratropical regions. Both models show, however, a high $p\text{CO}_2$ in the equatorial Pacific extending too far to the western coastal region, with positive biases in the western equatorial Pacific and negative biases in the eastern equatorial Pacific. The bias of $p\text{CO}_2$ is reduced at the western boundary of the tropical Pacific Ocean, and its maximum shifted eastward in MPI-ESM-HR compared to MPI-ESM-LR. This is consistent for all model members and related to the reduction and eastward shift of the cold-tongue SST bias in the MPI-ESM-HR version as shown in Figure 2. In the high-latitude Southern Ocean both models underestimate $p\text{CO}_2$, which is similar to the CMIP5 model (Ilyina et al., 2013). As shown in Ilyina et al. (2013), the bias of $p\text{CO}_2$ can be attributed to the model's deficiency in producing correct water masses along the Antarctic Circumpolar Current.

Both model versions reproduce the spatial pattern of observed phosphate concentration, with high concentration in the North Pacific, Southern Ocean, and eastern equatorial Pacific; and low concentration in the subtropical Pacific, Atlantic, and Indian Oceans. Both models underestimate the amplitude of the spatial gradient of phosphate concentration compared to observations. Underestimation is found in high-concentration
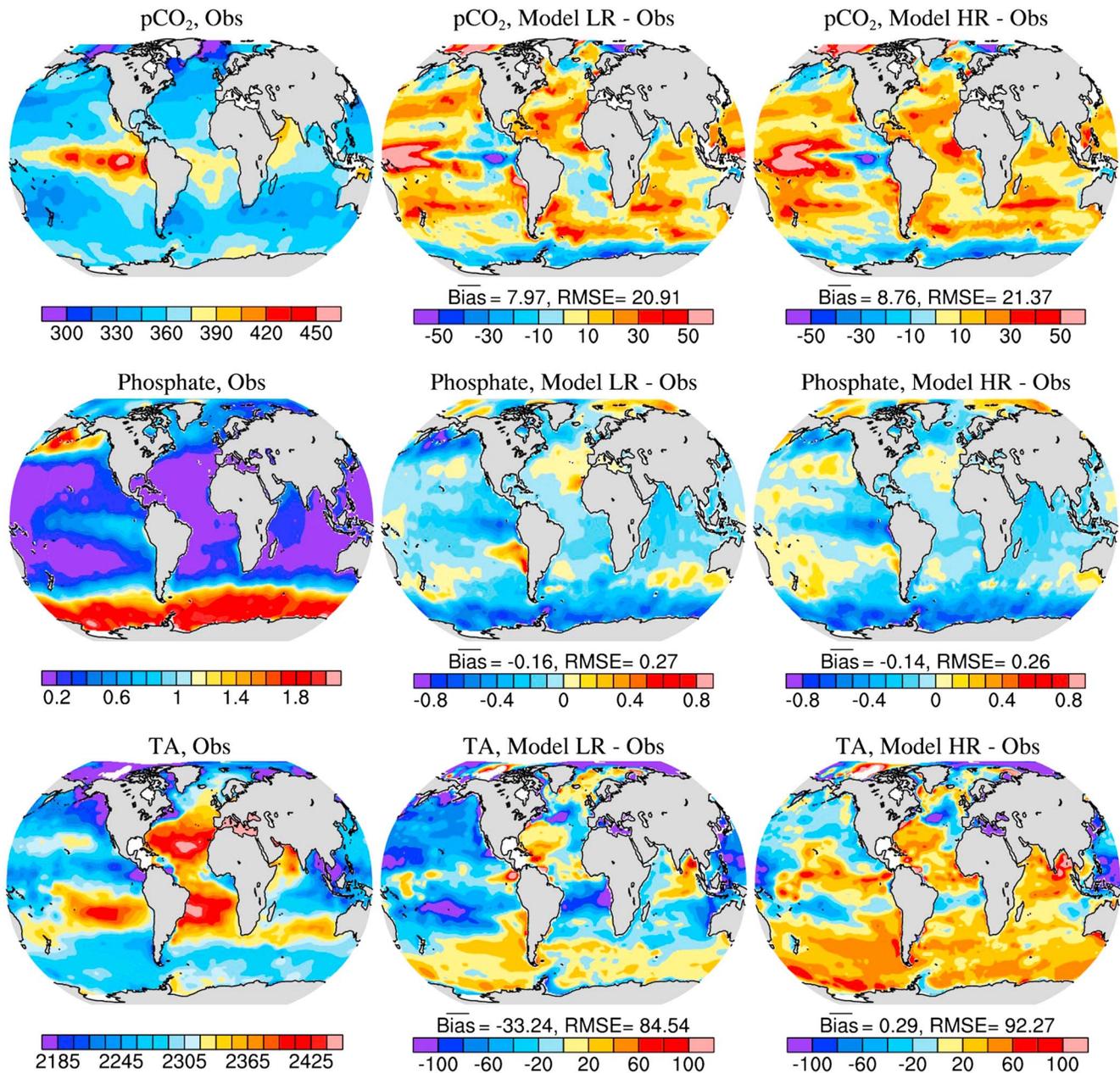
Figure 6. Climatological mean state of ocean biogeochemical variables from observation (left column), MPI-ESM1.2-LR minus observation (middle column), and MPI-ESM1.2-HR minus observations (right column). The shown variables are ocean surface pCO_2 (top row), phosphate concentration (middle row), and total alkalinity (TA, bottom row). The observed pCO_2 data are from Takahashi et al. (2009), phosphate data are from World Ocean Atlas, and TA data are from the Global Ocean Data Analysis Project version 2 database. Modeled pCO_2, phosphate, and TA are averaged over time periods corresponding to their respective observations, that is, 1999–2001, 1933–2005, and 1972–2005, respectively. Five-member ensemble mean results of MPI-ESM1.2-HR and MPI-ESM1.2-LR historical simulations are presented here. Units are ppm for pCO_2, mmol/m^3 for phosphate and μmol/kg for TA.

areas and overestimation in low-concentration areas. In some regions such as the North Pacific and the Southern Ocean, where the phosphate concentration is relatively high, the magnitude of the bias is up to 50% of the values of the observations. MPI-ESM-HR regionally improves phosphate concentration, such as in the North Pacific and in the eastern tropical Pacific Ocean. However, the bias patterns are overall comparable in MPI-ESM-LR and MPI-ESM-HR.

The models also capture the spatial distribution of the TA, with high values in the subtropical Atlantic and Pacific regions and lower values in the extratropics. The bias structure of MPI-ESM-HR reveals a strong

positive bias over the tropical regions and Southern Oceans and is larger compared to MPI-ESM-LR. This is found in all ensemble members. Further, the negative biases in the North Pacific are reduced in MPI-ESM-HR compared to MPI-ESM-LR and are also seen in all ensemble members.

4. Evaluation of Atmospheric Mean State

Increasing the horizontal resolution of ECHAM in AMIP-type experiments causes a general reduction of the bias in the atmospheric mean state and variance (Hertwig et al., 2015; Roeckner et al., 2006). Further, it leads to a warming of the troposphere and a poleward shift of the westerlies, and the temperature bias in the upper-level troposphere is substantially reduced (Figure 3 in Hertwig et al., 2015). As also documented by Roeckner et al. (2006), increasing the vertical resolution leads to a cooling of the atmosphere in the middle-to-low latitudes, surface warming, and equatorward shift of the westerlies. We now investigate the impact of increased model resolution on the atmospheric mean state in MPI-ESM-HR compared with MPI-ESM-LR.

4.1. Surface Air Temperature, Sea Level Pressure, Precipitation, and Top-of-the-Atmosphere Radiation Fluxes

We start with the description of the biases of surface-related quantities. The general structure of the bias of near-surface air temperature (SAT) over land is similar in the two model versions (Figures 7a and 7b). Large biases of SAT with more than 5°C remain over continental areas. Over the Rocky Mountains, the SAT bias is reduced in MPI-ESM-HR compared to MPI-ESM-LR. Further, a reduction of SAT bias is found over Brazil. This is consistent for all ensemble members. Over the northern part of Eurasia, the negative SAT bias is largely increased in MPI-ESM-HR compared to MPI-ESM-LR, which is possibly related to the negative bias over the high-latitude North Atlantic (Figure 2). Generally, the magnitudes of the SAT errors averaged globally decrease for MPI-ESM-HR, with an ensemble mean root-mean-square error (RMSE) of 2.24°C compared to MPI-ESM-LR with RMSE = 2.48°C (note that the ensemble spread is ±0.02°C for both model versions).

The bias of SLP also reveals similar structures in both model versions (Figures 7c and 7d). Over the North Atlantic and the Eurasian continent, alternating bands of bias with magnitudes greater than 2 hPa are found. This is coherent with a stronger-than-observed zonal structure of storm activity shown in the next subsections. The magnitude of the bias is slightly lower over the North Atlantic in MPI-ESM-HR compared to MPI-ESM-LR for all ensemble members. Over the North Pacific, the bias is consistently reduced in MPI-ESM-HR, compared to MPI-ESM-LR. This indicates a better representation of the Aleutian low in MPI-ESM-HR. Also, the SLP bias in the southern Atlantic is reduced in all ensemble members in MPI-ESM-HR compared to MPI-ESM-LR. The improvement of SLP bias in MPI-ESM-HR is likely due to improved SST (Figure 2; Grodsky et al., 2012; Milinski et al., 2016). Overall the global SLP error decreases from RMSE = 2.05 Pa for MPI-ESM-LR to RMSE = 1.91 hPa for MPI-ESM-HR (note that the ensemble spread is ±0.04 hPa for MPI-ESM-LR and ±0.03 hPa for MPI-ESM-HR). The errors in MPI-ESM-HR are comparable to those of AMIP-type experiments in Hertwig et al. (2015).

The largest precipitation biases occur in the tropical regions in both model versions (Figures 7e and 7f). Both versions show the prominent doubled Intertropical Convergence Zone (ITCZ) in the Pacific (e.g., Lin, 2007) and the Atlantic ITCZ bias (e.g., Siongco et al., 2015). The double ITCZ is a common feature in coupled GCMs (e.g., Lin, 2007) and is characterized by excessive precipitation over the southern Pacific convergence zone. This is also prominent in MPI-ESM-LR and MPI-ESM-HR, for which precipitation is too high by 4–5 mm/day, compared to Global Precipitation Climatology Project. The MPI-ESM-HR bias is slightly higher than the bias in MPI-ESM-LR. Over the western equatorial Pacific, the precipitation dry bias in MPI-ESM-HR is substantially decreased compared to MPI-ESM-LR and is closely associated with the reduction of the SST cold-tongue bias (Figure 2). This is found in all ensemble members. Over the tropical Atlantic, the precipitation bias is reduced in MPI-ESM-HR compared to MPI-ESM-LR, particular over the Gulf of Guinea. Here the precipitation bias is reduced from >5 mm/day in MPI-ESM-LR to 3-4 mm/day in MPI-ESM-HR. This results from a decrease of the SST bias over this region and is related to the improved representation of orography and wind stress (Milinski et al., 2016) and is found in all ensemble members. The precipitation error averaged globally and over all ensemble members decreases from RMSE = 1.63 mm/day in MPI-ESM-LR to RMSE = 1.46 mm/day in MPI-ESM-HR (note that the ensemble spread is ±0.01 mm/day for both model versions).
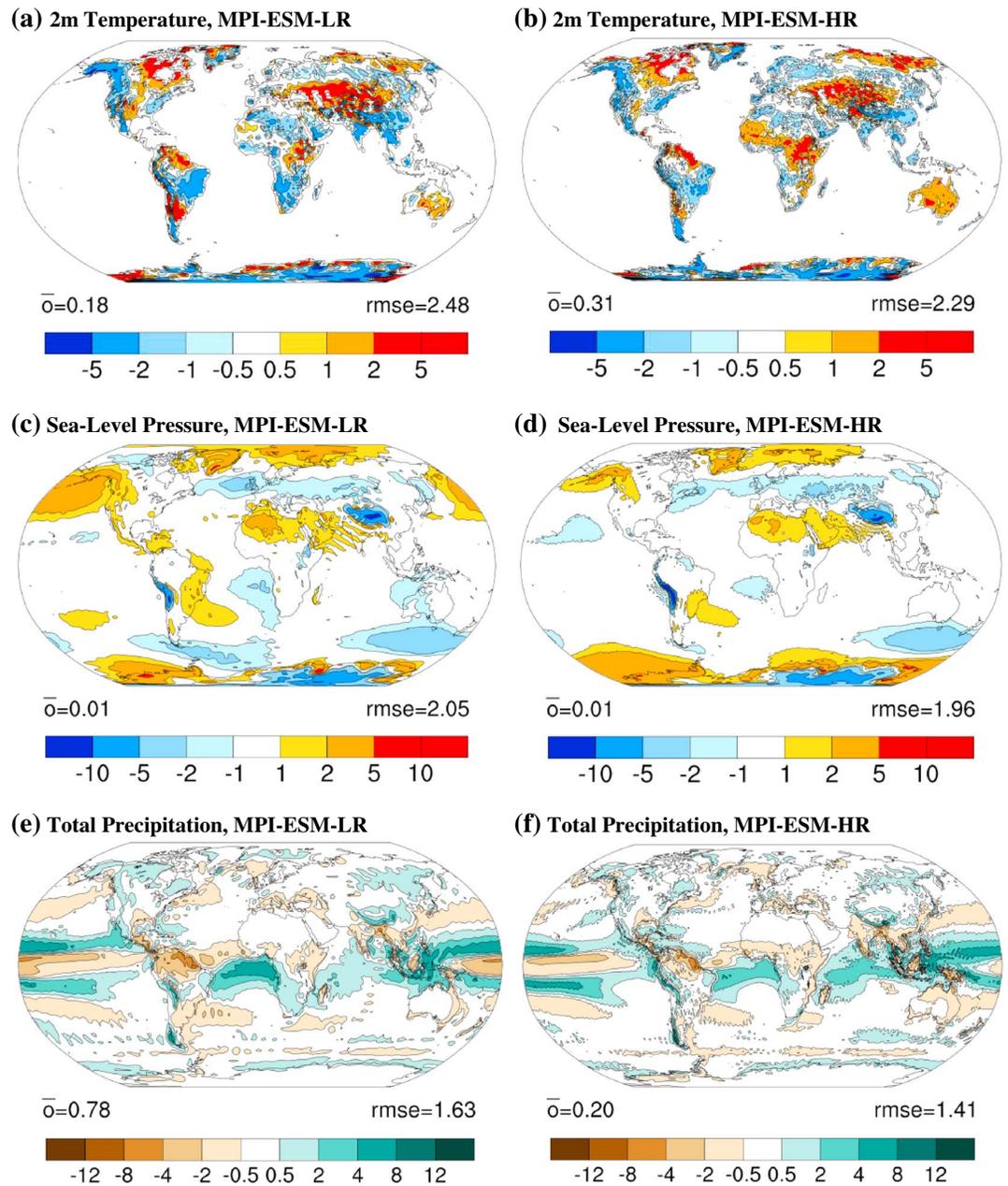
Figure 7. Bias of (a, b) 2-m temperature, (c, d) sea level pressure, and (e, f) total precipitation for (a, c, and e) MPI-ESM1.2-LR and (b, d, and e) MPI-ESM1.2-HR. Shown are results for ensemble mean of 5 historical experiments for the period 1979–2005. For 2-m temperature and sea level pressure the reference is ERA-Interim. For the total precipitation the reference is Global Precipitation Climatology Project. Units are degree Celsius (°C), hPa, and mm/day. Numbers show mean biases and root-mean-squared errors averaged over all grid points.

Related error patterns of shortwave and longwave radiation with respect to CERES are found at the top-of-the-atmosphere. Stronger incoming shortwave and low reflected longwave radiation is shown in the West Pacific in MPI-ESM-LR (Figures 8a and 8c). This is associated with a lack of middle-high level clouds and relative dryness in this region. For MPI-ESM-HR, the lower reflected radiation is shifted to the central Pacific (Figure 8c) and is associated with a negative precipitation bias in this region. In contrast, higher reflected radiation is found over the sub-tropical southern Atlantic. This error is larger in MPI-ESM-LR compared to MPI-ESM-HR (Figures 8b and 8d), and is associated with the SST and precipitation biases in this region (Figures 2 and 7).
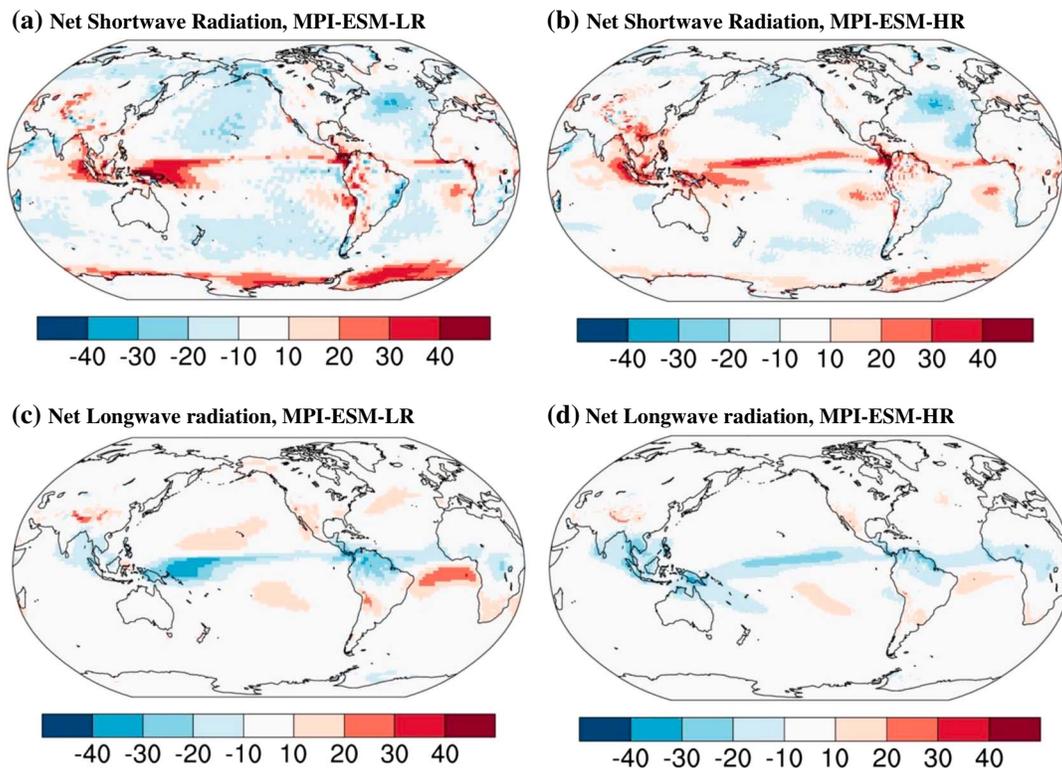
Figure 8. Bias of annually averaged (a, b) shortwave and (c, d) longwave radiation at the top-of-the-atmosphere of (a, c) MPI-ESM1.2-LR and (b, d) MPI-ESM1.2-HR with respect to CERES. Units are in W/m^2.

Overall, the biases in near-surface climate are similar in structure, but the doubled resolution does lead to reduced RMSEs of about 10% across the board.

4.2. Vertical Structure of Temperature and Zonal Wind

The ERA-Interim climatology and model biases in the vertical atmospheric structure of temperature and zonal wind are shown in Figure 9. The temperature bias patterns indicate cold zonal-mean temperature bias in the upper troposphere for both model versions (Figures 9c and 9e). For MPI-ESM-HR, the temperature cross section reveals a general warming in the upper troposphere and thereby a local reduction of bias, compared to MPI-ESM-LR. This is found in all ensemble members and is in line with previous AMIP-type experiments with varying resolutions of ECHAM (Hertwig et al., 2015; Roeckner et al., 2006) and earlier studies (e.g., Stratton, 1999), which generally show a bias reduction of zonal-mean temperatures in this region with increasing horizontal atmospheric resolution.

The zonal-mean zonal wind reveals the typical subtropical jet structure with maximum wind at ~200 hPa similar to ERA-Interim (Figures 9b, 9d, and 9f). The model biases indicate that both model versions have stronger jets with an extra contribution in the upper troposphere at 40–45°S and 40–45°N, compared to ERA-Interim (Figures 9d and 9f). At the location of the jet strength maxima of ERA-Interim, the bias of the zonal-mean zonal wind in MPI-ESM-HR is smaller compared to MPI-ESM-LR. This is consistent for all ensemble members. Below 200 hPa the bias is slightly higher in MPI-ESM-HR than in MPI-ESM-LR. A negative bias is seen in both model versions in the upper tropical troposphere, whereas in the middle tropical troposphere both models exhibit stronger zonal winds. The bias of the zonal-mean zonal wind in MPI-ESM-HR in this region is smaller compared to MPI-ESM-LR.

Previous AMIP-type experiments diagnosed a poleward shift of the subtropical jets and thereby a decrease of the bias with increasing atmospheric horizontal resolution. A poleward shift with increasing resolution is not seen here. This may be due to the increase in both horizontal and vertical resolution in MPI-ESM-HR compared to MPI-ESM-LR. Roeckner et al. (2006) point out that with the increase in the horizontal resolution
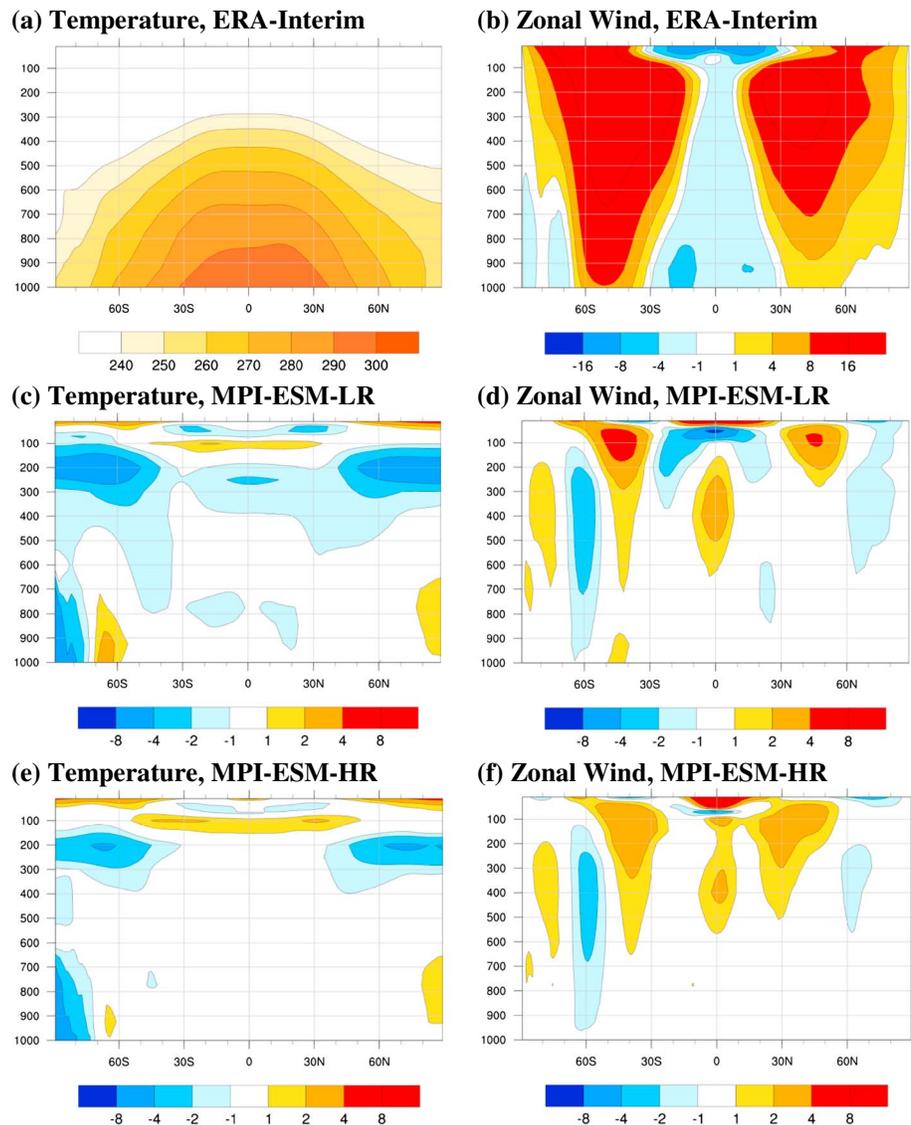
Figure 9. Time mean zonal mean (a, c, and e) temperature and (b, d, and f) zonal wind as a function of latitude and height for (a, b) ERA-Interim, (c, d) MPI-ESM1.2-LR minus ERA-Interim, and (e, f) MPI-ESM1.2-HR minus ERA-Interim for the period 1979–2005. Model results are based on the ensemble mean of historical experiments. Units are kelvin (°K) and m/s.

the midlatitude westerlies intensify and shift poleward. The increase in vertical resolution results in a decrease and equatorward shift of the westerlies. However, the underlying causes are not well known.

4.3. Storm Track Climatology and Atmospheric Blocking Frequency

A reduction of the bias of the jet strength as that shown in MPI-ESM-HR may have consequences for regional dynamical systems such as the storm tracks and atmospheric blocking. The development of synoptic storms is confined to the baroclinically unstable jet stream. In the Northern Hemisphere these storm tracks are located over preferred regions with strong temperature gradients in the North Atlantic and North Pacific, with larger amplitudes in winter than in summer. Within the CMIP5 models, the winter North Atlantic storm tracks tend to be either too zonal or displaced southward, resulting in too many storms entering into southern or central Europe (Zappa et al., 2013). A good association is found between the too zonal structure of storm track density and upper-level zonal wind bias at the jet exit. Models with higher resolution tend to have better resolved storm track density.
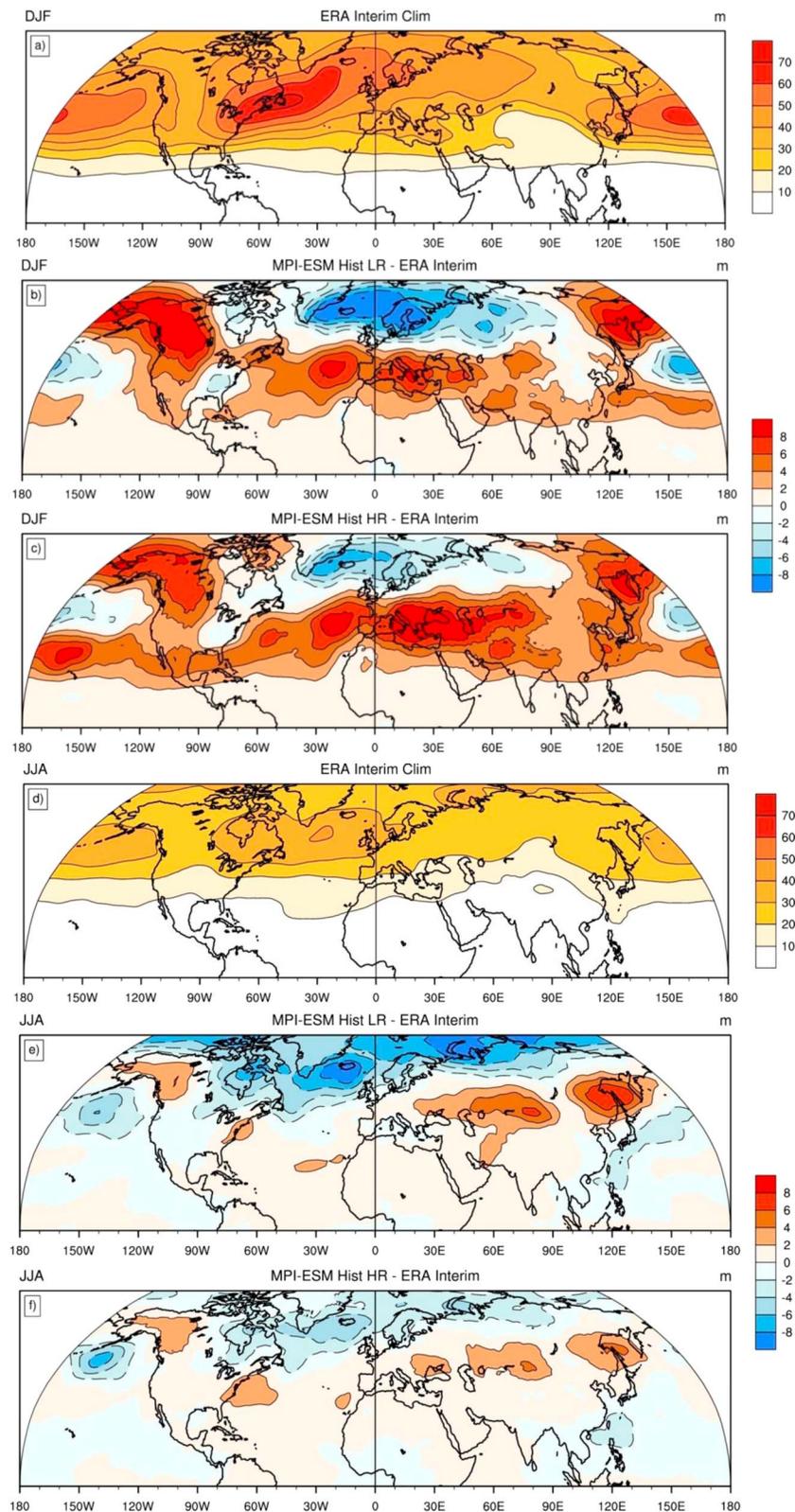
Figure 10. The square root of 2- to 6-day band-pass-filtered geopotential height variance at 500 hPa for (a–c) winter December–February (DJF) means and (d–f) summer June–August (JJA) means for the period 1979–2005. Shown are (a, d) ERA-Interim, (b, e) MPI-ESM1.2-LR minus ERA-Interim, and (c, f) MPI-ESM1.2-HR minus ERA-Interim. For the models, the ensemble average of the squared root variances is calculated. Units are in meters (m).

The storm tracks in both model versions are shown for winter and summer means (Figure 10). Shown is the square-root of the variances of 2- to 6-day band-pass-filtered geopotential height at 500 hPa of the respective seasons. For winter, the ERA-Interim climatology shows the two distinct Northern Hemispheric storm track regions over the North Pacific and North Atlantic. A higher value of the square root variance is also seen over the Mediterranean region (Figure 10a). Over the North Atlantic, both model versions reveal a southward shift of the storm track activity at their exit regions, with increasing values of the square root variance over the Mediterranean region compared to ERA-Interim (Figures 10b and 10c). The differences to ERA-Interim are ~ −8 m for both model versions and are about 10–20% of the total values. This is closely associated with the horizontal distributions of the zonal wind biases at 200 hPa for the North Atlantic region (Figure 11). A strong positive zonal wind bias is apparent at the southern branch of the jet exit, with values up to 6 m for MPI-ESM-HR and 8 m for MPI-ESM-LR. This systematic bias is found in all ensemble members. The square root variance is larger in MPI-ESM-HR compared to MPI-ESM-LR, indicating more storms entering the Mediterranean region. Over the northern part of the North Atlantic, the square root variances reveal an improvement of MPI-ESM-HR compared to MPI-ESM-LR. Here differences with respect to ERA-Interim amount to −6 m for the MPI-ESM-HR and −8 m for MPI-ESM-LR, indicating more storms entering the northern European region in MPI-ESM-HR compared to MPI-ESM-LR. This is found in all ensemble members and is in line with the reduced bias in the upper-level zonal wind in MPI-ESM-HR in this region (Figure 11). Over the North Pacific, winter storm tracks are shifted northward and have a stronger than observed extension over North America. A southward shift of storm tracks is evident in the subtropical North Pacific, and increased variance north of Hawaii is stronger in MPI-ESM-HR compared to MPI-ESM-LR.

During summer the square root variances are generally reduced compared to winter (Figure 10d). Similar to winter means, however, the storm tracks over the North Atlantic are shifted southward in both model versions (Figures 10e and 10f). The bias of the square root variance is about 4 m in both versions over the subtropical region. Over the northern North Atlantic, however, the bias is ~ −4 m in MPI-ESM-HR and −8 m in MPI-ESM-LR and therefore reduced in the higher resolution. This is found in all ensemble members and indicates that more storms enter the northern European region during summer in MPI-ESM-HR compared to MPI-ESM-LR. Furthermore, over central and East Asia there is a notable reduction of the bias in MPI-ESM-HR with respect to MPI-ESM-LR. Here the bias is substantially reduced by 2 m in MPI-ESM-HR compared to MPI-ESM-LR. This is associated with the subtropical jet over Asia, which appears too strong and too much shifted to the north in both versions, but whose bias is substantially reduced in MPI-ESM-HR (~4 m/s) compared to MPI-ESM-LR (8 m/s) as indicated by the upper tropospheric zonal wind (Figure 11). This can be related to the higher-resolution topography of the Himalayan region in MPI-ESM-HR. Over central to eastern Europe the zonal wind bias is also substantially reduced, with the magnitudes of the bias of 4 m/s in MPI-ESM-HR compared to 8–10 m/s in MPI-ESM-LR. This part of the bias reduction may be related to the adjustment of the subtropical jet and less related to the improvement of the North Atlantic storm tracks.

Similar results can be found for storm tracks in the Southern Hemisphere (not shown). Storm tracks are shifted toward the equator for both seasons. Similar to the Northern Hemisphere, the biases in the higher latitudes are decreased in the MPI-ESM-HR compared to MPI-ESM-LR, in both seasons. The bias closer to the equator is increased in MPI-ESM-HR.

The reduction of the bias of the zonal wind over the central European region further affects other atmospheric processes such as atmospheric blocking. Global climate models generally tend to underestimate the blocking frequency (Anstey et al., 2013; D'Andrea et al., 1998; Masato et al., 2013), and the increase in horizontal resolution of the atmospheric component has been shown to improve the blocking frequency (e.g., Berckmans et al., 2013). Other studies point toward the improvement of the North Atlantic current in higher-resolution coupled models that improve the winter blocking statistics (Scaife et al., 2011). The biases of the winter storm track densities have been related to biases in the blocking frequencies (Zappa et al., 2014, 2013). By using the CMIP5 models, Zappa et al. (2013) show that models with more frequent winter blocking over Europe have a higher cyclone density over the Norwegian and Mediterranean Seas and lower cyclone density over Europe, and better match with observations. In summer, only a weak relationship is found between European blocking and cyclone density. In the following the blocking frequencies are calculated for both model versions following Scherrer et al. (2006). The frequencies are calculated as the mean of the frequencies over all ensemble members.
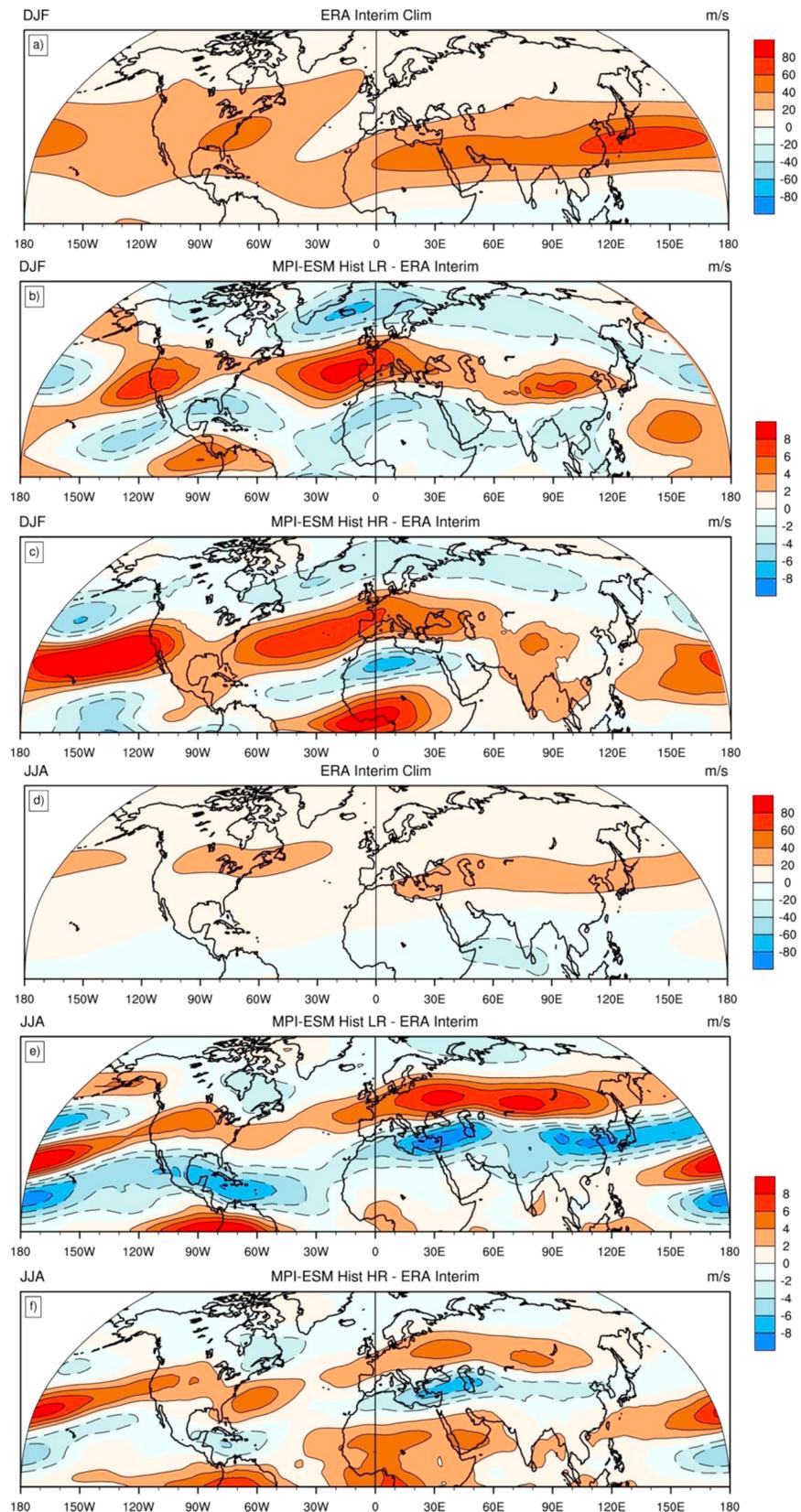
Figure 11. Same as Figure 10 but for zonal wind at 200 hPa for (a–c) December–February (DJF) and (d–f) June–August (JJA). For the models, the ensemble mean of the historical experiments are used. Units are in m/s.
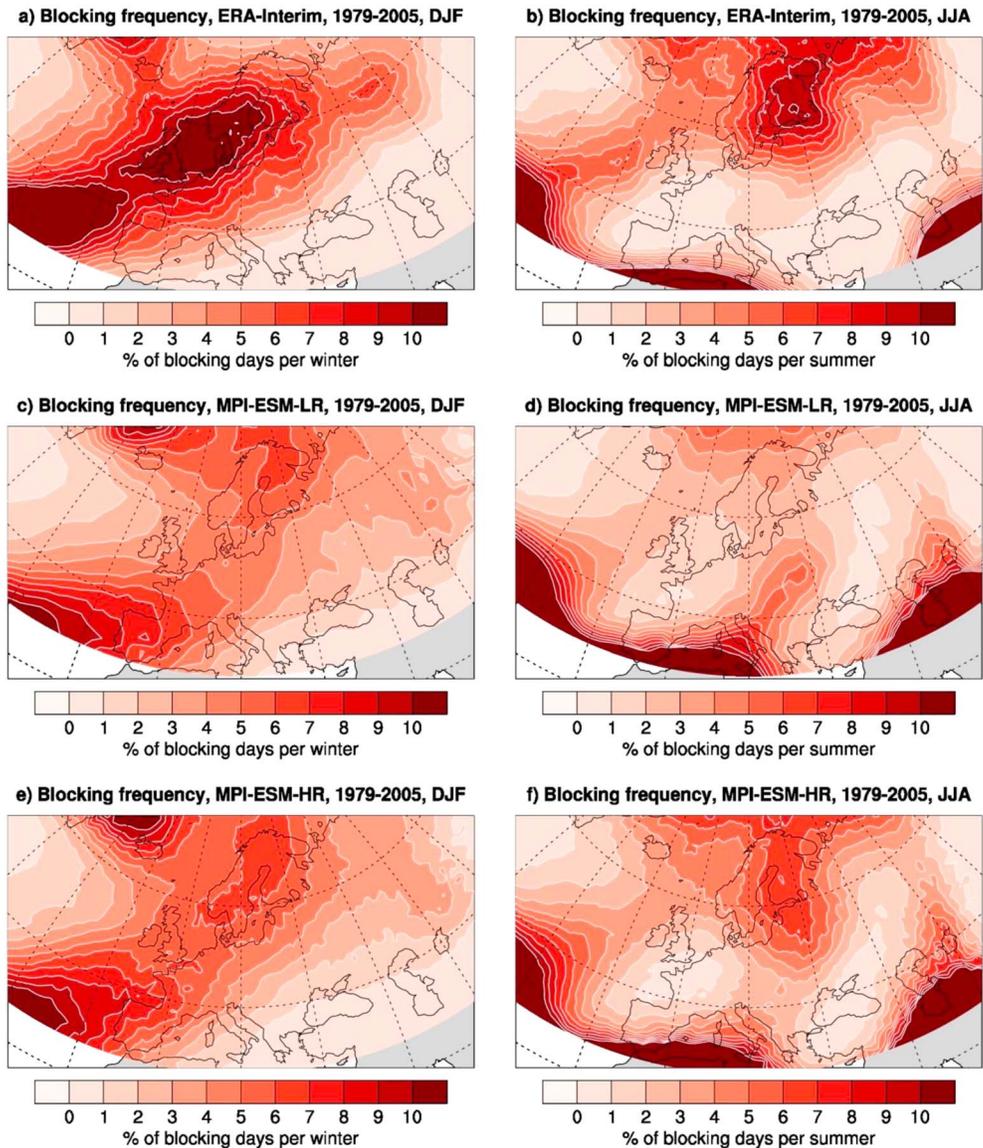
Figure 12. Blocking frequencies for (a, c, and e) winter (December–February, DJF) and (b, d, and f) summer (June–August, JJA) for (a, b) ERA-Interim, (c, d) MPI-ESM1.2-LR, and (e, f) MPI-ESM1.2-HR. Shown are the percentage of blocking days per season. The results for MPI-ESM1.2-LR and MPI-ESM1.2-HR are based on an ensemble mean for the period 1979–2005. Units are in percentage (%).

For winter means, both model versions tend to underestimate the number of blocked days with respect to ERA-Interim (Figures 12a, 12c, and 12e). Over the North Atlantic European region, two centers of action appear in the reanalysis with more than 10% of blocked days per season: over Greenland and over northern Europe. The location of the Greenland blocking region is captured by both model versions. Over northern Europe centers of blocking actions are shifted toward northeast Europe in both model versions compared to ERA-Interim. However, for MPI-ESM-HR, the blocking frequency is increased up to 6% of blocked days per season.

For the summer mean, both model versions underestimate the blocking frequencies; however, there is a substantial improvement in MPI-ESM-HR over the northern to eastern European region. ERA-Interim shows a local maximum of frequency over this region with up to 10% of blocked days per summer (Figure 12b). The MPI-ESM-LR version does not show this pattern, with a frequency of 2% of blocking days per season
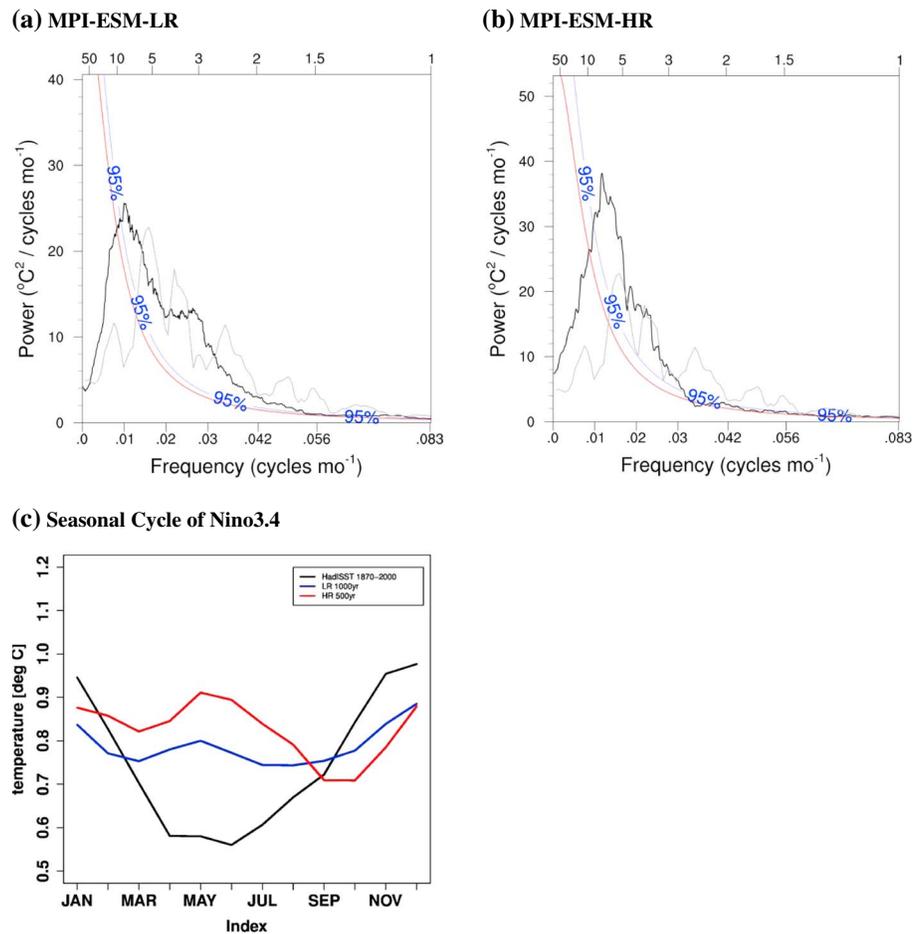
Figure 13. Nino3.4 power spectrum for (a) piControl of MPI-ESM1.2-LR and (b) piControl of MPI-ESM1.2-HR. Thin black lines in upper panels shows spectra of HadISST1.1 for the period 1870–2010 and thick black lines the spectra of the models. Also shown are the 95% significance levels based on an auto-regressive AR(1)-process for (blue) HadISST1.1 and (red) MPI-ESM. The y axis denote the power, and the x axis denote frequency (cycle/month) and years. (c) Standard deviations of sea surface temperature anomalies in Nino3.4 region as a function of calendar months for (black) HadISST1.1, (blue) MPI-ESM1.2-LR, (red) MPI-ESM1.2-HR. Units are degree Celsius (°C).

(Figure 12d). The MPI-ESM-HR version exhibits a structure similar to the reanalysis, with a frequency of 6–7% of blocking days per summer (Figure 12f). The increase in the blocking frequencies in summer is coherent with the reduction of bias of the zonal wind in the MPI-ESM-HR version and is found in all ensemble members. During summer, the wind bias over Europe in MPI-ESM-HR is only half of that in MPI-ESM-LR, and weaker winds seem to be associated with atmospheric blocking conditions.

5. Interannual Variability in the Coupled System

Next we describe two key processes of seasonal to interannual variability: the ENSO and the NAO.

5.1. ENSO

The power spectra of ENSO exhibit significant values at the prominent frequencies of 2–10 years in both model versions (Figures 13a and 13b). For frequencies 5–10 years, MPI-ESM-HR shows larger power than HadISST.

The amplitudes of indices for the different ENSO regions for both versions are also within the range of HadISST (Figure 13c, for Niño3.4). A maximum amplitude in winter with standard deviations of 0.85 and 0.9 is found for MPI-ESM-LR and MPI-ESM-HR, respectively, compared to 0.95 in HadISST. During the tuning
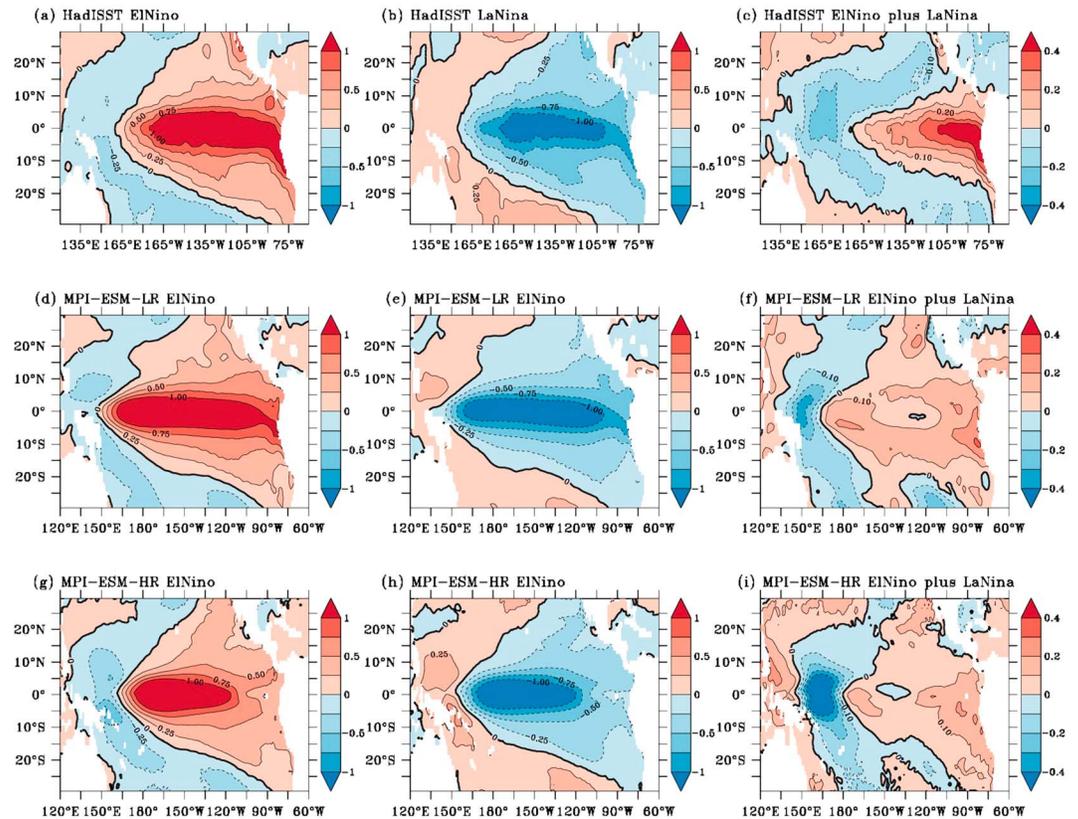
Figure 14. Composites of sea surface temperature (SST) anomalies for (left) El Niño phases, (middle) La Niña phases, and (right) El Niño phases plus La Niña phases. Composites are based on mean SST anomalies for Nino 3.4 index greater than +1 standard deviation and Nino 3.4 less than −1 standard deviation. Shown are results for (upper) HadISST1.1 (1870–2010), (middle) MPI-ESM1.2-LR piControl, and (lower) MPI-ESM1.2-HR piControl. Units are degree Celsius (°C). Please note that the composites are based on a representative sample. For example, given the +1 standard deviation 34 events are used for El Niño phases for HadISST1.1.

process, the amplitude of Niño3.4 has been increased in MPI-ESM-HR, as a result of decreasing the convective updraft parameter, the rate at which condensate is converted to precipitation. A lower conversion rate leads to more detrainment into anvil clouds and hence strengthens the longwave cloud feedback on the atmospheric circulation, thereby strengthening ENSO variability (Mauritsen et al., 2012; Rädel et al., 2016). Lowering the conversion rate has been a deliberate decision because the MPI-ESM-HR version had too little variance compared to the observed standard deviation of Niño3.4 using the MPI-ESM-LR setting. The seasonal cycle of the amplitudes, however, shows a mismatch with respect to the phase locking that appears in March to May in HadISST. This phase locking is important for teleconnections such as the Indian Monsoon (Webster et al., 1998). Figure 13c shows that standard deviations in both model versions decrease from the peak winter months; however, in contrast to the single minimum that occurs in springtime in observations, a second maximum can be seen in MPI-ESM-LR and MPI-ESM-HR, though this peak is more pronounced in MPI-ESM-HR.

A further characteristic of ENSO is the asymmetry between its positive and negative phases: the amplitude of SST anomalies is generally stronger during El Niño than during La Niña (Burgers & Stephenson, 1999; Hoerling et al., 1997). Even though ENSO asymmetry is suggested to be an important character of ENSO, it is generally underestimated by many CGCM models in CMIP3 and CMIP5 (Zhang & Sun, 2014; Zhang et al., 2009). Figure 14 shows the SST composites relative to the El Niño and La Niña phases and their sums. The asymmetry is pronounced in the eastern Pacific, the upwelling region, and the western Pacific, as shown in the observations. In MPI-ESM-LR the typical structure of the asymmetry is apparent but with different magnitude and extension. The modeled positive values are stretched too far westward compared to HadISST. This can also be

Table 6

Skewness of El Niño/Southern Oscillation for HadISST (1870–2010), MPI-ESM1.2-LR (piControl 500 Years), MPI-ESM1.2-HR (piControl 500 Years), MPI-ESM1.0-LR (piControl 500 Years), and MPI-ESM1.0-MR (piControl 500 Years) for the Nino3, Nino3.4, and Nino4 Regions

Skewness	Nino3	Nino3.4	Nino4
HadISST	0.76	0.38	−0.16
MPI-ESM1.2-LR	0.16	0.23	0.17
MPI-ESM1.2-HR	0.24	0.12	−0.37
MPI-ESM1.0-LR (CMIP5)	0.08	0.02	0.17
MPI-ESM1.0-MR (CMIP5)	−0.04	−0.01	0.01

Note. Units are in degree Celsius (°C). CMIP = Coupled Model Intercomparison Project.

seen in the composites inferring a westward shift of the centers of anomalies. This asymmetry pattern is similar to the MPI-ESM-LR version in CMIP5 (not shown). As a consequence the positive values of the asymmetry are shifted westward by 5–10°. In MPI-ESM-HR, the composite anomalies show that the bulk of ENSO activity is constrained to the central Pacific. The western edges are near the dateline and comparable to HadISST. In the West Pacific the asymmetry is improved. In MPI-ESM-HR, however, the ENSO activity is much reduced in the upwelling and Niño1.2 regions. Therefore, the asymmetry in the eastern Pacific is reduced compared to MPI-ESM-LR. Skewness estimates such as those of Burgers and Stephenson (1999) for the different ENSO regions are shown in Table 6. Both model resolutions show an improvement of the skewness with respect to their precursor from CMIP5 in the Niño3 and Niño3.4 region. In the Niño4 region MPI-ESM-HR shows negative values, but larger than those from observations.

It is noteworthy that the coupling frequency between ocean and atmosphere is different between MPI-ESM-LR, where daily coupling is used, and MPI-ESM-HR, where the frequency is increased to hourly. Tian et al. (2017) examine the effect of changing the coupling from daily to hourly frequency in both versions. They show that the diurnal cycle of SST is well resolved in the hourly coupled MPI-ESM. This affects the cooling of SST after sunset, which is slower than the cooling of the atmosphere and causes further convection at nighttime. During El Niño events, the enhanced nighttime convection is amplified by the background SST in the central tropical Pacific. Anomalies of the Walker circulation, which are directly related to convection, also appear stronger with the hourly coupling. For La Niña events, the effect of hourly coupling is small. A direct interpretation of this effect in Figure 14, however, is difficult, since this analysis is based on two model configurations and two different coupling frequency, and we refer to Tian et al. (2017).

5.2. The NAO

For the North Atlantic-European region the NAO is an important climate diagnostic with distinct signatures in the atmosphere-ocean coupling and impacts on continental-scale key parameters. The winter NAO is the principal mode of variability in both model versions in all ensemble members and in ERA-Interim (Figures 15a–15c; the figure shows the principal mode only for one ensemble member). The NAO explains 46% of the total variability in MPI-ESM-HR and 49% of the total variability in MPI-ESM-LR (mean values of explained variance of the ensemble members). The dipole structure is present in both versions and comparable with ERA-Interim. The principal components are standardized, showing the EOFs (Empirical Orthogonal Functions) multiplied with their eigenvalues in physical units. From Figure 15, the maximum amplitudes of MPI-ESM-HR are closer to ERA-Interim compared to MPI-ESM-LR. ERA-Interim shows maximum amplitudes larger than ±3 hPa. For MPI-ESM-HR, the maximum amplitudes of the negative anomalies are also greater than 3 hPa and extend far more toward Europe than in MPI-ESM-LR. The amplitudes of the positive anomalies in MPI-ESM-HR are also similar to ERA-Interim, with values of −3 hPa. These maximum values, however, are shifted westward by 10–20° compared to ERA-Interim. The NAO in MPI-ESM-LR exhibits smaller amplitudes of up to −2.5 hPa for the negative anomalies. For the positive anomalies, the maximum values are of similar size than MPI-ESM-HR and observations. These results are consistent throughout the ensemble members.

On seasonal to interannual timescales, the NAO is associated with transient eddy momentum fluxes and their interaction with the mean flow and variation of storm tracks (e.g., Woollings et al., 2015). Composites of the NAO on band-pass-filtered geopotential height reveal the meridional shift and intensification of storm activity with respect to the different NAO phases (Figures 16a–16c). During positive NAO phases, an increase in storm activity is found in ERA-Interim over the northern North Atlantic. During the negative NAO phases, more storms enter the Mediterranean region. This behavior is present in both model versions, though their magnitudes are much smaller than inERA-Interim. MPI-ESM-HR improves the NAO-related storm track activity. An increase in storm activity appears over northern Europe, indicative of an eastward extension of the jet during the positive NAO phase.

Variations of the storm tracks with the different phases of the NAO are ultimately linked to the jet position, the transient eddy momentum fluxes, and their interaction with the mean flow. On seasonal to interannual
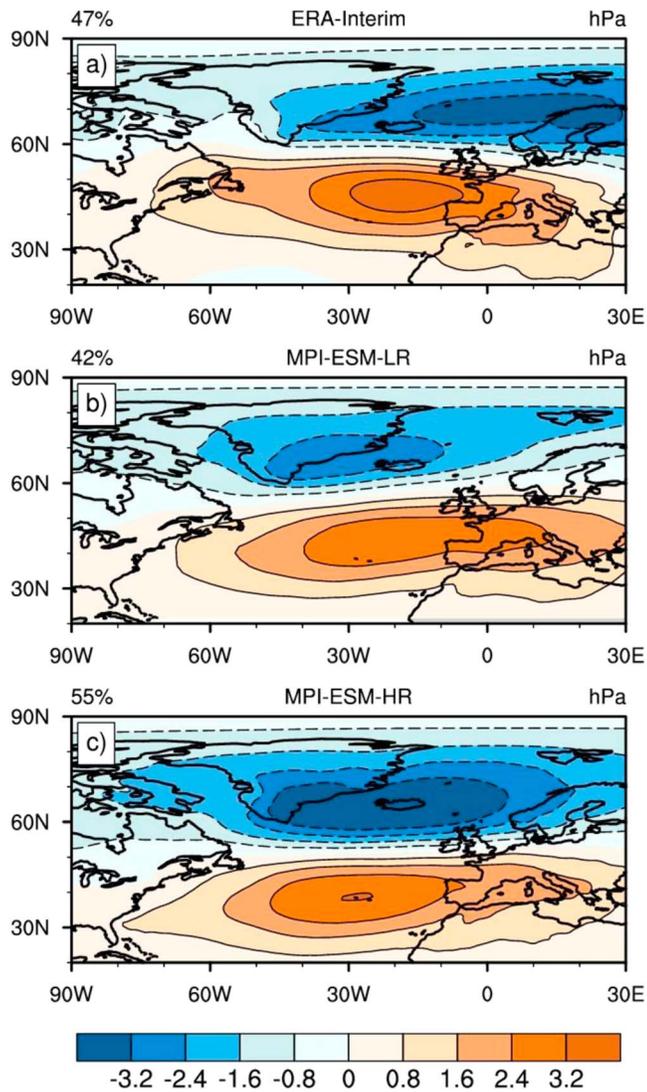
Figure 15. Winter (December–February) North Atlantic Oscillation (NAO) for (a) ERA-Interim, (b) MPI-ESM1.2-LR, and (c) MPI-ESM1.2-HR for the period 1980–2005. The NAO is defined as the first mode of variability of an EOF analysis based on sea level pressure. One ensemble member of the historical runs is shown for MPI-ESM1.2-LR and MPI-ESM1.2-HR as an example. However, the amplitudes look similar for all ensemble members. The NAO explains 47% of the total variance for ERA-Interim, 42% (51%, 45%, 54%, and 53%) for MPI-ESM1.2-LR and 55% (50%, 42%, 43%, and 43%) for MPI-ESM1.2-HR. In brackets the explained variances for the other ensemble members. Units are in hectopascals.

timescales, the transient eddies act to reinforce the mean jet during the positive phases of the NAO and broaden the jet during negative NAO phases (Woollings et al., 2015, cf. their Figure 8). This is illustrated with the eddy-kinetic energy for ERA-Interim (Figure 16d). In both model versions, the eddy kinetic energy is transformed to the mean flow energy over the North Atlantic during positive phases of the NAO. However, maximum values of eddy-kinetic energy are located near the climatological minimum, which indicates that only the amplitude of the jet is altered but not its position. In MPI-ESM-HR, a second maximum appears over Europe. This shows that much of the eddy-kinetic energy transformation takes place further downstream. The amplitudes, however, are smaller than in ERA-Interim. This strengthening of the jet position is associated with the increasing storm track activity over the northern North Atlantic.

In summary, the winter NAO variability is improved in the MPI-ESM-HR compared to MPI-ESM-LR, and amplitudes in MPI-ESM-HR are close to ERA-Interim. The NAO-related storm track activity further shows improvement in MPI-ESM-HR, with more storms entering the northern European region. This is related to enhanced transient eddy momentum fluxes and eddy-mean flow interaction in MPI-ESM-HR compared to MPI-ESM-LR. This should lead to further improvements in quantities influenced by the NAO, such as precipitation and continental surface air temperature, but those quantities are not considered here.

6. Summary

We have presented a higher-resolution version of the coupled MPI-ESM1.2-HR (hereafter MPI-ESM-HR) and compared it with the model's lower-resolved version MPI-ESM1.2-LR (hereafter MPI-ESM-LR). Both are baseline versions for CMIP6, and standardized sets of experiments similar to the CMIP6 design (ongoing Diagnosis, Evaluation, and Characterization of Klima, Eyring et al., 2016) are performed. The experiments are done with the CMIP5 forcing, because CMIP6 forcing was not available when these simulations were carried out. We analyze oceanic and atmospheric mean states and important modes of climate variability, the ENSO and NAO. A documentation of the development of the atmospheric component ECHAM6.3 and code changes relative to the predecessor model is presented in a forthcoming manuscript on MPI-ESM1.2.

MPI-ESM-HR is well tuned, and drifts of the ocean hydrography are relatively weak. The mean ocean circulation, such as ocean gyres and the AMOC, is stable with magnitudes close to observed estimates. For example, the AMOC at 26°N is 16 Sv in MPI-ESM-HR compared to 17 Sv in the RAPID array (McCarthy et al., 2015; Smeed et al., 2016). Similarly, a stable AMOC is present in MPI-ESM-LR with magnitudes of ~19 Sv at 26°N. In the atmosphere, vertical cross sections of the mean temperature and zonal wind show that upper level biases are reduced with increased resolution, and confirm previous studies (Hertwig et al., 2015; Roeckner et al., 2006). This has an effect on regional atmospheric dynamics as discussed below. Biases in near-surface climate are similar in structure in both model versions, but the doubled resolution leads to reduced RMSEs of about 10% across the board.

The North Atlantic climate system and adjacent continental regions suffer from a large SST bias, strong zonal elongation of the North Atlantic current, and too zonal storm tracks. The SST bias is prominent in the earlier generations of MPI-ESM, and increasing the ocean resolution to eddy permitting has not yet led to desired results. The SST bias constitutes a drawback for certain applications, when remote atmospheric impacts

Figure 16. (a–c) Composites of storm tracks (2- to 6-day band-pass-filtered geopotential height at 500 hPa) with respect to positive minus negative North Atlantic Oscillation phases for (a) ERA-Interim, (b) MPI-ESM1.2-LR, and (c) MPI-ESM1.2-HR. Shown are the averages of composites for all ensemble members. Contours show ensemble mean climatology. Units are in meter (m). (d–f) Composites of transient eddy momentum flux on to the mean flow $E \cdot D$ with respect to the positive minus negative North Atlantic Oscillation phases (here $E = ((v'^2 - u'^2)/2, u'v')$ is similar to the E vector, and $D = (U_x - V_y, V_x + U_y)$ is the deformation vector of the time mean flow; u' and v' are the 2- to 6-day band-pass-filtered transient zonal and meridional wind components at 200 hPa; U and V are the wind components averaged over the winter seasons; for details see Raible et al., 2010.) Positive values denote eddy growth on the expense of the background state. Negative values show the loss of eddy-kinetic energy to the mean flow. Shown are the averages of composites for all ensemble members. Contours show ensemble mean climatology.

are considered. Examples are the relative role of North Atlantic heat fluxes for decadal summer climate variations over Europe and the Eurasian continent (Ghosh et al., 2016; Wu et al., 2016), or the prediction of storm intensity for winter over Europe (Kruschke et al., 2015). However, the increase in the horizontal resolution results in a reduction of the bias of the mean upper-level winds over the North Atlantic-European region. For winter means, this results in an improvement of the storm tracks over the northern North Atlantic and an increase in winter storms entering the northern European region. For summer means, the higher-resolution atmospheric model component substantially decrease the bias of the mean zonal wind in the upper layers over the Eurasian continent. This results in an improved position of the Asian jet and an increase in summer blocking frequency over Europe.

We have also shown that the winter NAO and NAO-related storm track variability are closer to ERA-Interim in MPI-ESM-HR than in MPI-ESM-LR. On interannual timescale the NAO is dominated by latitudinal variations of

the jet position (Woollings et al., 2015). Moreover, transient eddy momentum fluxes act to reinforce the mean flow and strengthen the jet further downstream during positive NAO phases. These processes are enhanced in the higher-resolution atmospheric model components. However, there is also evidence that other climate components such as the stratosphere, sea ice, and local and remote air-sea interactions affect the NAO variability. Recently, it has been shown that these subcomponents can be used to enhance the NAO prediction (Dobrynin et al., 2018; Hall et al., 2017; Wang et al., 2017). To examine the total effect of the higher-resolution model version, the contributions of these climate subsystems on the NAO variability need to be considered; this is not done here.

The tropical Pacific climate system is characterized by the cold-tongue SST bias and the double ITCZ, two typical features in many coupled climate models (Li & Xie, 2014; Lin, 2007). The cold-tongue SST bias is reduced in the MPI-ESM-HR version compared to MPI-ESM-LR. This is also reflected in the bias of other variables, such as precipitation, column water vapor, and cloud cover. In addition, ENSO variability is reasonably simulated in both model versions, and significant peaks occur at a 2- to 10-year time range. The amplitudes of ENSO in their peak phases are captured by both model versions. A drawback is a secondary maximum during summer, which is not visible in HadISST. Such a secondary maximum may interfere with applications such as seasonal forecasts, for which forecasts that are initialized in early spring drift toward higher amplitudes than observed.

The tropical Atlantic climate system reveals large biases in both model versions, a warm SST bias in the eastern tropical Atlantic, and a cold bias in the western part. MPI-ESM-HR, however, improves the mean state by reducing the SST bias by ~1°K in the western Atlantic and also the costal area of the eastern tropical Atlantic. This is associated with a decrease of the precipitation bias in the southern tropical Atlantic and Gulf of Guinea. The SAT of nearby continental regions and the SLP are also improved. The reduction of the bias is due to a better representation of the coastal orography in the atmospheric component together with higher resolved winds, which result in improved surface wind stress and oceanic upwelling (Milinski et al., 2016). This is in line with studies, which proposed that coarse resolutions of the coastal orography in coupled climate models result in biases in coastal winds and errors in the surface fluxes (Large & Danabasoglu, 2006). Others have shown that increase in model resolution associated with more realistic wind stress curls at the eastern boundaries are required to simulate the upwelling system (Small et al., 2015).

7. Discussion

The reductions of the regional biases in MPI-ESM-HR are expected to improve short and near-term climate prediction systems and climate projections and thus provide more reliable regional climate information. As an example, decadal climate predictions are currently employed in many modeling centers with the goal of operational use (Marotzke et al., 2016; Smith et al., 2013) and scientific contributions such as for CMIP6 (Boer et al., 2016). Typically, the North Atlantic is a key region for decadal climate predictions, and forecast skill is found for various quantities such as heat content and SST (e.g., Kröger et al., 2012; Matei et al., 2012; Pohlmann et al., 2013; Yeager et al., 2012). Prediction skill is further assessed for climate impacts, such as for Atlantic tropical cyclones (Dunstone et al., 2011; Smith et al., 2010). However, forecast skill is strongly reduced over continents. Based on MPI-ESM, the predictability of winter storms has been assessed for the North Atlantic-European region, and forecast skill is found beyond the first season (Kruschke et al., 2014, 2015). Further, wind speed and associated wind energy potential over central Europe have been evaluated and forecast skill is achieved for the first years (Moemken et al., 2016). However, the regional climate information for applications is limited. An improvement of storm frequencies and associated wind speed is expected to further improve the reliability of such predictions.

Further, for summer means an improvement of the blocking statistics in higher resolution is expected to increase the seasonal and decadal summer climate predictability over central Europe in MPI-ESM (Bunzel et al., 2017; Müller et al., 2012; Sienz et al., 2016). Recent work has established an atmospheric pathway from North Atlantic multidecadal SST variations to the European summer (Ghosh et al., 2016). By using the twentieth century reanalysis, Ghosh et al. (2016) showed that heat flux variations in the North Atlantic result in a linear baroclinic atmospheric response with alternating low pressure systems over the Atlantic and blocking-like events over central Europe. The SST bias is crucial in this framework and in the coupled system leads to a displacement of the baroclinic response. However, an improvement of the blocking frequency, such as

shown in the MPI-ESM-HR experiments, might also result in a more realistic occurrence of blocking events in the prediction systems.

In summary, the increase in atmospheric resolution employed in MPI-ESM-HR leads to an improvement of atmospheric mean state and regional atmospheric dynamics compared to MPI-ESM-LR. Though some of the long-standing biases are only modestly reduced, the improvements of atmospheric dynamics in the higher resolution make this model well suited for prediction and impact studies.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P., Janowiak, J., et al. (2003). The version 2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *Journal of Hydrometeorology*, 4, 1147–1167.

Anstey, J. A., Davini, P., Gray, L. J., Woollings, T. J., Butchart, N., Cagnazzo, C., et al. (2013). Multi-model analysis of northern Hemisphere winter blocking: Model biases and the role of resolution. *Journal of Geophysical Research: Atmospheres*, 118, 3956–3971. https://doi.org/10.1002/jgrd.50231

Bacon, S. (1997). Circulation and fluxes in the North Atlantic between Greenland and Ireland. *Journal of Physical Oceanography*, 27, 1420–1435.

Baehr, J., Fröhlich, K., Botzet, M., Domeisen, D. I. V., Kornblueh, L., Notz, D., et al. (2015). The prediction of surface temperature in the new seasonal prediction system based on the MPI-ESM coupled climate model. *Climate Dynamics*, 44, 2723–2735. https://doi.org/10.1007/s00382-014-2399-7

Bellenger, H., Guilyardi, E., Leloup, J., Lengaigne, M., & Vialard, J. (2014). ENSO representation in climate models: From CMIP3 to CMIP5. *Climate Dynamics*, 42(7-8), 1999–2018. https://doi.org/10.1007/s00382-013-1783-z

Berckmans, J., Woollings, T., Demory, M.-E., Vidale, P.-L., & Roberts, M. (2013). Atmospheric blocking in a high resolution climate model: Influences of mean state, orography and eddy forcing. *Atmospheric Science Letters*, 14, 34–40.

Bersch, M. (1995). On the circulation of the northeastern North Atlantic. *Deep-Sea Research Part I*, 42, 1583–1607.

Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirtman, B., et al. (2016). The decadal climate prediction project. *Geoscientic Model Development Discussions*, 2016, 1–32. https://doi.org/10.5194/gmd-2016-78

Buehler, T., Raible, C. C., & Stocker, T. F. (2011). The relationship of winter season North Atlantic blocking frequencies to extreme cold or dry spells in the ERA-40. *Tellus*, 63(A), 212–222.

Bunzel, F., Müller, W. A., Stacke, T., Hagemann, S., Dobrynin, M., Baehr, J., & Froehlich, K. (2017). Improved seasonal prediction of European summer temperatures with a new 5-layer soil-hydrology scheme. *Geophysical Research Letters*, 45, 346–353. https://doi.org/10.1002/2017GL076204

Burgers, G., & Stephenson, D. B. (1999). The normality of El Niño. *Geophysical Research Letters*, 26, 1027–1039.

Butler, A., Alves, O., Arribas, A., Athanassiadou, M., Baehr, J., Calvo, N., et al. (2016). The Climate-system Historical Forecast Project: Do stratosphere-resolving models make better seasonal climate predictions in boreal winter? *Quarterly Journal of the Royal Meteorological Society*, 142, 1413–1427. https://doi.org/10.1002/qj.2743

Chidichimo, M. P., Kanzow, T., Cunningham, S. A., Johns, W. E., & Marotzke, J. (2010). The contribution of eastern-boundary density variations to the Atlantic meridional overturning circulation at 26.5°N. *Ocean Science*, 6, 475–490. https://doi.org/10.5194/os-6-475

Clarke, R. A. (1984). Transport through the Cape Farewell-Flemish Cap section. *Rapports et Procès-Verbaux des Réunions*, 185, 120–130.

Cunningham, S. A., Alderson, S. G., King, B. A., & Brandon, M. A. (2003). Transport and variability of the Antarctic circumpolar current in Drake Passage. *Journal of Geophysical Research*, 108(C5), 8084. https://doi.org/10.1029/2001JC001147

Cunningham, S. A., & Marsh, R. (2010). Observing and modeling changes in the Atlantic MOC. *Wiley Interdisciplinary Reviews: Climate Change*, 1, 180–191.

D'Andrea, F., Tibaldi, S., Blackburn, M., Boer, G., Déqué, M., Dix, M. R., et al. (1998). Northern Hemisphere atmospheric blocking as simulated by 15 atmospheric general circulation models in the period 1979–1998. *Climate Dynamics*, 14(6), 385–407. https://doi.org/10.1007/s003820050230

Dawson, A., & Palmer, T. N. (2015). Simulating weather regimes: Impact of model resolution and stochastic parameterization. *Climate Dynamics*, 44. https://doi.org/10.1007/s00382-014-2238-x

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. https://doi.org/10.1002/qj.828

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

DiMarco, S. F., Chapman, P., Nowlin, W. D. Jr., Hacker, P., Donohue, K., Luther, M., et al. (2002). Volume transport and property distribution of the Mozambique Channel. *Deep-Sea Research Part II*, 49, 1481–1511.

Dobrynin, M., Domeisen, D. I. V., Müller, W. A., Bell, L., Brune, S., Bunzel, F., et al. (2018). Improved teleconnection-based dynamical seasonal prediction of boreal winter. *Geophysical Research Letters*, 45. https://doi.org/10.1002/2018GL077209

Donohue, K. A., Tracey, K. L., Watts, D. R., Chidichimo, M. P., & Chereskin, T. K. (2016). Mean Antarctic Circumpolar Current transport measured in Drake Passage. *Geophysical Research Letters*, 43, 11,760–11,767. https://doi.org/10.1002/2016GL070319

Dunstone, N. J., Smith, D. M., & Eade, R. (2011). Multi-year predictability of the tropical Atlantic atmosphere driven by the high latitude North Atlantic Ocean. *Geophysical Research Letters*, 38, L14701. https://doi.org/10.1029/2011GL047949

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

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., et al. (2013). Evaluation of climate models. In T. F. Stocker, et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 741–766). Cambridge, United Kingdom and New York, NY: Cambridge University Press.

Ganachaud, A., & Wunsch, C. (2000). Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. *Nature*, 408, 453–457.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O. K., Zweng, M. M., et al. (2014). In S. Levitus & A. Mishonov (Eds.), *World Ocean Atlas 2013, Volume 4: Dissolved inorganic nutrients (phosphate, nitrate, silicate), NOAA Atlas NESDIS* (Vol. 76, p. 25).

Gent, P. R., Yeager, S. G., Neale, R. B., Levis, S., & Bailey, D. A. (2010). Improvements in a half degree atmosphere/land version of the CCSM. *Climate Dynamics*, *34*(6), 819–833. https://doi.org/10.1007/s00382-009-0614-8

Ghosh, R., Müller, W. A., Baehr, J., & Bader, J. (2016). Impact of observed North Atlantic multidecadal variations to European summer climate: A linear baroclinic response to surface heating. *Climate Dynamics*. https://doi.org/10.1007/s00382-016-3283-4

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

Goll, D., Brovkin, V., Liski, J., Raddatz, T., Thum, T., & Todd-Brown, K. (2015). Strong dependence of CO_2 emissions from anthropogenic land cover change on soil carbon parametrization and initial land cover. *Global Biogeochemical Cycles*, *29*, 1511–1523. https://doi.org/10.1002/2014GB004988

Gordon, A. L., Sprintall, J., Van Aken, H. M., Susanto, D., Wijffels, S., Molcard, R., et al. (2010). The Indonesian throughflow during 2004–2006 as observed by the INSTANT program. *Dynamics of Atmospheres and Oceans*, *50*, 115–128. https://doi.org/10.1016/j.dynatnoce.2009.12.002

Griffies, S., Winton, M., Anderson, W. G., Benson, R., Delworth, T. L., Dufour, C. O., et al. (2015). Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *Journal of Climate*, *28*(3), 952–977. https://doi.org/10.1175/JCLI-D-14-00353.1

Grodsky, S. A., Carton, J. A., Nigam, S., & Okumura, Y. M. (2012). Tropical Atlantic biases in CCSM4. *Journal of Climate*, *25*, 3684–3701.

Hagemann, S., & Dumenil-Gates, L. (2003). An improved sub grid runoff parameterization scheme for climate models. *Climate Dynamics*, *21*, 349–359.

Hagemann, S., & Stacke, T. (2015). Impact of the soil hydrology scheme on simulated soil moisture memory. *Climate Dynamics*, *44*, 1731–1750. https://doi.org/10.1007/s00382-014-2221-6

Hall, R., Scaife, A., Hanna, E., Jones, J., & Erdélyi, R. (2017). Simple statistical probabilistic forecasts of the winter NAO. *Weather and Forecasting*. https://doi.org/10.1175/WAF-D-16-0124.1

Hansen, B., Østerhus, S., Turrell, W. R., Jonsson, S., Valdiarsson, H., Hatun, H., & Olsen, S. M. (2008). The inflow of Atlantic water, heat, and salt to the Nordic Seas across the Greenland-Scotland Ridge. In R. R. Dickson, et al. (Eds.), *Arctic-Subarctic Ocean Fluxes* (pp. 15–43). Dordrecht, Netherlands: Springer.

Hertwig, E., von Storch, J.-S., Handorf, D., Dethloff, K., Fast, I., & Krismer, T. (2015). Effect of horizontal resolution on ECHAM6-AMIP performance. *Climate Dynamics*, *45*(1), 185–211.

Hoerling, M. P., Kumar, A., & Zhong, M. (1997). El Niño, La Niña, and the nonlinearity of their teleconnections. *Journal of Climate*, *10*, 1769–1786.

Holliday, N. P., Bacon, S., Allen, J., & McDonagh, E. I. (2009). Circulation and transport in the western boundary currents at Cape Farewell, Greenland. *Journal of Physical Oceanography*, *39*, 1854–1870.

Hoskins, B. J., James, I. N., & White, G. H. (1983). The shape, propagation and mean-flow interaction of large scale weather systems. *Journal of Atmospheric Sciences*, *40*, 1595–1612.

Hurrell, J. W., Kushnir, Y., Visbeck, M., & Ottersen, G. (2003). An overview of the North Atlantic Oscillation. In J. W. Hurrell, et al. (Eds.), *The North Atlantic Oscillation: Climatic significance and environmental impact, Geophys. Monogr. Ser.* (Vol. 134, pp. 1–35). Washington, D. C.: American Geophysical Union.

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H., & Nunez-Riboni, I. (2013). The global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth System Model in different CMIP5 experimental realizations. *Journal of Advances in Modeling Earth Systems*, *5*, 287–315. https://doi.org/10.1029/2012MS000178

Imawaki, S., Uchida, H., Ichikawa, H., Fukasawa, M., Umatani, S., & the ASUKA Group (2001). Satellite altimeter monitoring of the Kuroshio transport south of Japan. *Geophysical Research Letters*, *28*, 17–20.

Jochumsen, K., Moritz, M., Nunes, N., Quadfasel, D., Larsen, K. M. H., Hansen, B., et al. (2017). Revised transport estimates of the Denmark Strait overflow. *Journal of Geophysical Research, Oceans*, *122*, 3434–3450. https://doi.org/10.1002/2017JC012803

Jochumsen, K., Quadfasel, D., Valdimarsson, H., & Jónsson, S. (2012). Variability of the Denmark Strait overflow: Moored time series from 1996–2011. *Journal of Geophysical Research*, *117*, C12003. https://doi.org/10.1029/2012JC008244

Johns, W., Shay, T., Bane, J., & Watts, D. (1995). Gulf Stream structure, transport, and recirculation near 68 W. *Journal of Geophysical Research*, *100*, 817–838.

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

Kanzow, T., Cunningham, S. A., Johns, W. E., Hirschi, J. J. M., Marotzke, J., Baringer, M. O., et al. (2010). Seasonal variability of the Atlantic meridional overturning circulation at 26.5°N. *Journal of Climate*, *23*, 5678–5698.

Kirtman, B. P., Bitz, C., Bryan, F., Collins, W., Dennis, J., Hearn, N., et al. (2012). Impact of ocean model resolution on CCSM climate simulations. *Climate Dynamics*, *39*(6), 1303–1328. https://doi.org/10.1007/s00382-012-1500-3

Krismer, T. R., Giorgetta, M. A., & Esch, M. (2013). Seasonal aspects of the quasi-biennial oscillation in MPI-ESM and ERA-40. *Journal of Advances in Modeling Earth Systems*, *5*, 1–16. https://doi.org/10.1002/jame.20024

Kröger, J., Müller, W. A., & von Storch, J. S. (2012). Impact of different ocean reanalyses on decadal climate prediction. *Climate Dynamics*. https://doi.org/10.1007/s00382-012-1310-7

Kruschke, T., Rust, H. W., Kadow, C., Leckebusch, G. C., & Ulbrich, U. (2014). Evaluating decadal predictions of Northern Hemispheric cyclone frequencies. *Tellus A*, *66*. https://doi.org/10.3402/tellusa.v66.22830

Kruschke, T., Rust, H. W., Kadow, C., Müller, W. A., Pohlmann, H., Leckebusch, G. C., & Ulbrich, U. (2015). Probabilistic evaluation of decadal predictions of Northern Hemisphere winter storms. *Meteorologische Zeitschrift*. https://doi.org/10.1127/metz/2015/0641

Large, W. G., & Danabasoglu, G. (2006). Attribution and impacts of upper-ocean biases in CCSM3. *Journal of Climate*, *19*, 2325–2346.

Lauvset, S. K., Key, R. M., Olsen, A., van Heuven, S., Velo, A., Lin, X., et al. (2016). A new global interior ocean mapped climatology: The 1° × 1° GLODAP version 2. *Earth System Science Data*, *8*, 325–340. https://doi.org/10.5194/essd-8-325-2016

Levitus, S., Boyer, T. P., Conkright, M. E., Johnson, D., O'Brien, T., Antonov, J., et al. (1998). NOAA Atlas NESDIS 19,World Ocean database 1998 volume 2: Temporal distribution of mechanical bathythermograph profiles, U.S. Gov. Printing Office, Wash., D.C., 286 pp.

Lherminier, P., Mercier, H., Gourcuff, C., Alvarez, M., Bacon, S., & Kermabon, C. (2007). Transports across the 2002 Greenland–Portugal Ovide section and comparison with 1997. *Journal of Geophysical Research*, *112*, C07003. https://doi.org/10.1029/2006JC003716

Li, G., & Xie, S.-P. (2014). Tropical biases in the CMIP5 multi-model ensemble: The excessive equatorial Pacific cold-tongue and double ITCZ problems. *Journal of Climate*, *27*, 1765–1780. https://doi.org/10.1175/JCLI-D-13-00337.1

Lin, J.-L. (2007). The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean–atmosphere feedback analysis. *Journal of Climate*, *20*, 4497–4525. https://doi.org/10.1175/JCLI4272.1

Mak, M., & Cai, M. (1989). Local barotropic instability. *Journal of the Atmospheric Sciences*, *46*(21), 3289–3311.

Marotzke, J., Müller, W. A., Vamborg, F. S. E., Becker, P., Cubasch, U., Feldmann, H., et al. (2016). MiKlip—A national research project on secadal climate prediction. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-15-00184.1

Martin, J. H., Knauer, G. A., Karl, D. M., & Broenkow, W. W. (1987). VERTEX: Carbon cycling in the northeast Pacific. *Deep Sea Research Part A. Oceanographic Research Papers*, *34*(2), 267–285.

Masato, G., Hoskins, B. J., & Woollings, T. J. (2013). Winter and summer Northern Hemisphere blocking in CMIP5 models. *Journal of Climate*, *26*, 7044–7059. https://doi.org/10.1175/JCLI-D-12-00466.1

Matei, D., Pohlmann, H., Jungclaus, J. H., Müller, W. A., Haak, H., & Marotzke, J. (2012). Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model. *Journal of Climate*. https://doi.org/10.1175/JCLI-D-11-00633.1

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., et al. (2012). Tuning the climate of a global model. *Journal of Advances in Modeling Earth Systems*, *4*, M00A01. https://doi.org/10.1029/2012MS000154

McCarthy, G. D., Smeed, D. A., Johns, W. E., Frajka-Williams, E., Moat, B. I., Rayner, D., et al. (2015). Measuring the Atlantic Meridional Overturning Circulation at 26g°N. *Progress in Oceanography*, *130*, 91–111. https://doi.org/10.1016/j.pocean.2014.10.006

McDonagh, E. L., King, B. A., Bryden, H. L., Courtois, P., Szuts, Z., Baringer, M., et al. (2015). Continuous estimate of Atlantic Oceanic freshwater flux at 26.5 N. *Journal of Climate*, *28*, 8888–8906.

Meehl, G. A., Gent, P. R., Arblaster, J. M., Otto-Bleisner, B. L., Brady, E. C., & Craig, A. (2001). Factors that affect the amplitude of El Nino in global coupled models. *Climate Dynamics*, *17*, 515–526.

Meredith, M., Woodworth, P. L., Chereskin, T. K., Marshall, D. P., Allison, L. C., Bigg, G. R., et al. (2011). Sustained monitoring of the Southern Ocean at Drake Passage: Past achievements and future priorities. *Reviews of Geophysics*, *49*, L05603. https://doi.org/10.1029/2010RG000348

Milinski, S., Bader, J., Haak, H., Siongco, A. C., & Jungclaus, J. H. (2016). High atmospheric horizontal resolution eliminates the wind-driven coastal warm bias in the southeastern tropical Atlantic. *Geophysical Research Letters*, *43*, 10,455–10,462. https://doi.org/10.1002/2016GL070530

Moemken, J., Reyers, M., Buldmann, B., & Pinto, J. G. (2016). Decadal predictability of regional scale wind speed and wind energy potentials over Central Europe. *Tellus A*, *68*, 29199. https://doi.org/10.3402/tellusa.v68.29199

Müller, W. A., Baehr, J., Haak, H., Jungclaus, J. H., Kröger, J., Matei, D., et al. (2012). Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology. *Geophysical Research Letters*, *39*, L22707. https://doi.org/10.1029/2012GL053326

Müller, V., Pohlmann, H., Matei, D., Marotzke, J., Müller, W. A., & Baehr, J. (2016). Hindcast skill for the Atlantic meridional overturning circulation at 26.5°N within two MPI-ESM decadal climate prediction systems. *Climate Dynamics*. https://doi.org/10.1007/s00382-016-3482-z

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

Nowlin, W. D. Jr., & Klinck, J. M. (1986). The physics of the Antarctic Circumpolar Current. *Reviews of Geophysics*, *24*(3), 469–491. https://doi.org/10.1029/RG024i003p00469

Olsen, A., Key, R. M., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., et al. (2016). The Global Ocean Data Analysis Project version 2 (GLODAPv2)—An internally consistent data product for the world ocean. *Earth System Science Data*, *8*, 297–323. https://doi.org/10.5194/essd-8-297-2016

Paulsen, H., Ilyina, T., Six, K. D., & Stemmler, I. (2017). Incorporating a prognostic representation of marine nitrogen fixers into the global ocean biogeochemical model HAMOCC. *Journal of Advances in Modeling Earth Systems*, *9*, 438–464. https://doi.org/10.1002/2016MS000737

Pincus, R., & Stevens, B. (2013). Paths to accuracy for radiation parameterizations in atmospheric models. *Journal of Advances in Modeling Earth Systems*, *5*, 225–233. https://doi.org/10.1002/jame.20027

Pinto, J. G., & Raible, C. C. (2012). Past and recent changes in the North Atlantic Oscillation. *Climatic Change*, *3*, 79–90. https://doi.org/10.1002/wcc.150

Pithan, F., Shepherd, T. G., Zappa, G., & Sandu, I. (2016). Climate model biases in jet streams, blocking and storm tracks resulting from missing orographic drag. *Geophysical Research Letters*, *343*, 7231–7240. https://doi.org/10.1002/2016GL069551

Pohlmann, H., Smith, D. M., Balmaseda, M. A., Keenlyside, N. S., Masina, S., Matei, D., et al. (2013). Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system. *Climate Dynamics*, *41*, 775–785. https://doi.org/10.1007/s00382-013-1663-6

Polyakov, I. V., Alexeev, V. A., Bhatt, U. S., Polyakova, E. I., & Zhang, X. (2010). North Atlantic warming: pattern of long-term trend and multi-decadal variability. *Climate Dynamics*, *34*, 439–457.

Rädel, G., Mauritsen, T., Stevens, B., Domenget, D., Matei, D., Bellomo, K., & Clement, A. (2016). Amplification of El Nino by cloud longwave coupling to atmospheric circulation. *Nature Geoscience*, *9*(2), 106–110.

Raible, C. C., Ziv, B., Saaroni, H., & Wild, M. (2010). Winter synoptic-scale variability over the Mediterranean basin under future climate conditions simulated by the ECHAM5. *Climate Dynamics*, *35*, 473–488.

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

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

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

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., et al. (2006). Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. *Journal of Climate*, *19*, 3771–3791.

Rossby, T., & Flagg, C. (2012). Direct measurement of volume flux in the Faroe-Shetland Channel and over the Iceland-Faroe Ridge. *Geophysical Research Letters*, *39*, L07602. https://doi.org/10.1029/2012GL051269

Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., et al. (2014). Skillful long-range prediction of European and North American winters. *Geophysical Research Letters*, *41*, 2514–2519. https://doi.org/10.1002/2014GL059637

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

Scherrer, S. C., Croci-Maspoli, M., Schwierz, C. B., & Appenzeller, C. (2006). Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region. *International Journal of Climatology*, *26*, 233–249.

Schweiger, A., Lindsay, R., Zhang, J., Steele, M., Stern, H., & Kwok, R. (2011). Uncertainty in modeled Arctic sea ice volume. *Journal of Geophysical Research*, *116*, C00D06. https://doi.org/10.1029/2011JC007084

Shaffrey, L. C., Stevens, I., Norton, W. A., Roberts, M. J., Vidale, P. L., Harle, J. D., et al. (2009). U.K. HiGEM: The new U.K. High-Resolution Global Environment Model – Model description and basic evaluation. *Journal of Climate*, *22*(8), 1861–1896. https://doi.org/10.1175/2008JCLI2508.1

Sienz, F., Müller, W. A., & Pohlmann, H. (2016). Ensemble size impact on the decadal predictive skill assessment. *Meteorologische Zeitschrift*. https://doi.org/10.1127/metz/2016/0670

Siongco, A. C., Hohenegger, C., & Stevens, B. (2015). The Atlantic ITCZ bias in CMIP5 models. *Climate Dynamics*, *45*, 1169. https://doi.org/10.1007/s00382-014-2366-3

Small, R. J., Curchister, E., Hedstrom, K., Kauffman, B., & Large, W. G. (2015). The Benguela Upwelling System: Quantifying the sensitivity to resolution and coastal wind representation in a global climate model. *Journal of Climate*, *28*, 9409–9432.

Smeed, D., McCarthy, G., Rayner, D., Moat, B. I., Johns, W. E., Baringer, M. O., & Meinen, C. S. (2016). Atlantic meridional overturning circulation observed by the RAPID-MOCHA-WBTS (RAPID-Meridional Overturning Circulation and Heatflux Array-Western Boundary Time Series) array at 26 N from 2004 to 2015. British Oceanographic Data Centre - Natural Environment Research Council, UK.

Smith, D., Eade, R., Dunstone, N. J., Fereday, D., Murphy, J. M., Pohlmann, H., & Scaife, A. A. (2010). Skillful multi-year predictions of Atlantic hurricane frequency. *Nature Geoscience*, *3*(12), 846–849. https://doi.org/10.1038/ngeo1004

Smith, D. M., Scaife, A. A., Boer, G. J., Caian, M., Doblas-Reyes, F. J., Guemas, V., et al. (2013). Real-time multi-model decadal climate predictions. *Climate Dynamics*, *41*, 2875–2888. https://doi.org/10.1007/s00382-012-1600-0

Stacke, T., & Hagemann, S. (2016). Life time of soil moisture perturbations in a coupled land-atmosphere simulation. *Earth System Dynamics*, *7*, 1–19. https://doi.org/10.5194/esd-7-1-2016

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

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

Stratton, R. A. (1999). A high resolution AMIP integration using the Hadley Centre model HadAM2b. *Climate Dynamics*, *15*, 9–28.

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., et al. (2009). Climatological mean and decadal change in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*(8–10), 554–577.

Taylor, K., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and experimental design. *Bulletin of the American Meteorological Society*, *93*, 485–498.

Tian, F., von Storch, J.-S., & Hertwig, E. (2017). Air–sea fluxes in a climate model using hourly coupling between the atmospheric and the oceanic components. *Climate Dynamics*, *48*, 2819–2836. https://doi.org/10.1007/s00382-016-3228-y

Tyrilis, E., & Hoskins, B. J. (2007). Aspects of a northern hemisphere atmospheric blocking climatology. *Journal of the Atmospheric Sciences*, *65*, 1638–1652.

von Storch, J.-S., Eden, C., Fast, I., Haak, H., Hernandez-Deckers, D., Maier-Reimer, E., et al. (2012). An estimate of the Lorenz energy cycle for the world ocean based on the 1/10° STORM/NCEP simulation. *Journal of Physical Oceanography*, *42*, 2185–2205.

Wang, L., Ting, M., & Kushner, P. J. (2017). A robust empirical seasonal prediction of winter NAO and surface climate. *Scientific Reports*, *279*. https://doi.org/10.1038/s41598-017-00353-y

Webster, P. J., Magana, V. O., Palmer, T. N., Shukla, J., Tomas, R. A., Yanai, M., & Yasunari, T. (1998). Monsoons: Processes, predictability, and the prospects for prediction. *Journal of Geophysical Research*, *103*, 14,451–14,510. https://doi.org/10.1029/97JC02719

Wolff, J. O., Maier-Reimer, E., & Legutke, S. (1997). The Hamburg Ocean Primitive Equation model HOPE, Tech. Rep. 13, 98 pp., German Climate Computer Center (DKRZ), Hamburg, Germany.

Woodgate, R. A., Aagaard, K., & Weingartner, T. J. (2006). Interannual changes in the Bering Strait fluxes of volume, heat, and freshwater between 1991 and 2004. *Geophysical Research Letters*, *33*, L15609. https://doi.org/10.1029/2006GL02693

Woodgate, R., Weingartner, T., & Lindsa, R. (2012). Observed increases in Bering Strait oceanic fluxes from the Pacific to the Arctic from 2001 to 2011 and their impacts on the Arctic Ocean water column. *Geophysical Research Letters*, *39*, L24603. https://doi.org/10.1029/2012GL054092

Woollings, T., Franzke, C., Hodson, D. L. R., Dong, B., Barnes, E. A., Raible, C. C., & Pinto, J. G. (2015). Contrasting interannual and multidecadal NAO variability. *Climate Dynamics*, *45*, 539–556. https://doi.org/10.1007/s00382-014-2237-y

Wu, B., Ling, J., & Zhou, T. (2016). Interdecadal circumglobal teleconnection pattern during boreal summer. *Atmospheric Science Letters*, *17*, 446–452.

Yeager, S., Karspeck, A., Danabasoglu, G., Tribbia, J., & Teng, H. (2012). A decadal prediction case study: Late twentieth-century North Atlantic Ocean heat content. *Journal of Climate*, *25*, 5173–5189.

Zappa, G., Masato, G., Shaffrey, L., Woollings, T., & Hodges, K. (2014). Linking northern Hemisphere blocking and storm track biases in the CMIP5 climate models. *Geophysical Research Letters*, *41*, 135–139. https://doi.org/10.1002/2013GL058480

Zappa, G., Shaffrey, L. C., & Hodges, K. I. (2013). The ability of CMIP5 models to simulate North Atlantic extratropical cyclones. *Journal of Climate*, *26*, 5379–5396.

Zhang, T., & Sun, D.-Z. (2014). ENSO asymmetry in CMIP5 models. *Journal of Climate*, *27*, 4070–4093.

Zhang, T., Sun, D. Z., Neale, R., & Rasch, P. J. (2009). An evaluation of ENSO asymmetry in the Community Climate System Models: A view from the subsurface. *Journal of Climate*, *22*(22), 5933–5961. https://doi.org/10.1175/2009JCLI2933.1