The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6

(Article begins on next page)

01 September 2025

Geoscientific
Model Development

The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6

Ralf Döscher[1], Mario Acosta[2], Andrea Alessandri[3], Peter Anthoni[4], Thomas Arsouze[2], Tommi Bergman[5], Raffaele Bernardello[2], Souhail Boussetta[6], Louis-Philippe Caron[2], Glenn Carver[6], Miguel Castrillo[2], Franco Catalano[7], Ivana Cvijanovic[2], Paolo Davini[8], Evelien Dekker[1], Francisco J. Doblas-Reyes[2], David Docquier[1], Pablo Echevarria[2], Uwe Fladrich[1], Ramon Fuentes-Franco[1], Matthias Gröger[1], Jost v. Hardenberg[9,8], Jenny Hieronymus[1], M. Pasha Karami[1], Jukka-Pekka Keskinen[10], Torben Koenigk[1], Risto Makkonen[11], François Massonnet[12], Martin Ménégoz[13], Paul A. Miller[14], Eduardo Moreno-Chamarro[2], Lars Nieradzik[14], Twan van Noije[5], Paul Nolan[15], Declan O'Donnell[11], Pirkka Ollinaho[11], Gijs van den Oord[5], Pablo Ortega[2], Oriol Tintó Prims[2], Arthur Ramos[2], Thomas Reerink[5], Clement Rousset[16], Yohan Ruprich-Robert[2], Philippe Le Sager[5], Torben Schmith[16], Roland Schrödner[14], Federico Serva[17], Valentina Sicardi[2], Marianne Sloth Madsen[18], Benjamin Smith[14], Tian Tian[18], Etienne Tourigny[2], Petteri Uotila[10], Martin Vancoppenolle[19], Shiyu Wang[1], David Wårlind[14], Ulrika Willén[1], Klaus Wyser[1], Shuting Yang[18], Xavier Yepes-Arbós[2], and Qiong Zhang[20]

[1]Swedish Meteorological and Hydrological Institute (SMHI), Norrköping, 60176, Sweden
[2]Barcelona Supercomputing Center, Barcelona, 08034, Spain
[3]Institute of Atmospheric Sciences and Climate, Consiglio Nazionale delle Ricerche, ISAC-CNR, 40129 Bologna, Italy
[4]Karlsruhe Institute of Technology (KIT), 82467 Garmisch-Partenkirchen, Germany
[5]Royal Netherlands Meteorological Institute (KNMI), De Bilt, 3731, the Netherlands
[6]European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, RG2-9AX, United Kingdom
[7]Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), 00196 Rome, Italy
[8]Institute of Atmospheric Sciences and Climate, Consiglio Nazionale delle Ricerche, ISAC-CNR, 10133 Turin, Italy
[9]Politecnico di Torino, 10129 Turin, Italy
[10]University of Helsinki, Helsinki, 00014, Finland
[11]Finnish Meteorological Institute, FMI, Helsinki, 00560, Finland
[12]Université Catholique de Louvain UCLouvain, Ottignies-Louvain-la-Neuve, 1348, Belgium
[13]IGE, University of Grenoble, Grenoble, 38400, France
[14]Lund University, Lund, 22100, Sweden
[15]Irish Centre for High End Computing, ICHECK, Dublin, Ireland
[16]University Pierre and Marie Curie (UPMC), Paris, 75005, France
[17]Istituto di Scienze Marine CNR-ISMAR, 30122 Venice, Italy
[18]Danish Meteorological Institute, Copenhagen, 2100, Denmark
[19]Institut Pierre Simon Laplace (IPSL), Paris, 75005, France
[20]Department of Physical Geography, Stockholm University, Stockholm, 106 91, Sweden

Correspondence: Ralf Döscher (ralf.doescher@smhi.se)

Abstract. The Earth system model EC-Earth3 for contributions to CMIP6 is documented here, with its flexible coupling framework, major model configurations, a methodology for ensuring the simulations are comparable across different high-performance computing (HPC) systems, and with the physical performance of base configurations over the historical period. The variety of possible configurations and submodels reflects the broad interests in the EC-Earth community. EC-Earth3 key performance metrics demonstrate physical behavior and biases well within the frame known from recent CMIP models. With improved physical and dynamic features, new Earth system model (ESM) components, community tools, and largely improved physical performance compared to the CMIP5 version, EC-Earth3 represents a clear step forward for the only European community ESM. We demonstrate here that EC-Earth3 is suited for a range of tasks in CMIP6 and beyond.

1 Introduction

The latest challenges in climate research have evolved to include biophysical and biogeochemical processes (WCRP Strategic Plan, 2019–2028, https://www.wcrp-climate.org/wcrp-sp, last access: 18 March 2022) contributing to the exchange of energy, mass, aerosols, trace and greenhouse gases, and nutrients between the atmosphere, land, and ocean, allowing the description of various feedback processes. This challenge resulted in a need for the next generation of climate models – namely, the Earth system models (ESMs); see, e.g., Flato (2011).

The Paris Climate Accord is calling for limiting climate change "well below 2 °C and to pursue efforts to limit the increase to 1.5 °C". ESMs represent our most relevant tools available for exploring the emission pathways necessary for achieving this goal (Kawamiya et al., 2020), as well as for understanding the consequences of not making this target. The Paris Agreement requires firm measures of mitigation, including carbon dioxide removal. Given the complexity of the climate system, alternative emission pathways towards this goal can be carefully explored only with Earth system models (ESMs), which describe the most relevant feedback mechanisms and provide methods for assessments of uncertainty. ESMs are the primary source of information for understanding the Earth's climate feedbacks, for attributing changes to specific drivers, for future climate projections and predictions, and for the development of mitigation policies.

While the exact definition of ESM varies, in general, it refers to a complex model that besides the classical climate model core (consisting of physical models of the atmosphere, sea ice, ocean, and land) combines additional optional components covering biophysical and biogeochemical processes as well as more sophisticated treatment of aerosols. A flexible coupling framework facilitates a range of ESM config-urations with or without certain model components or processes. Given the important role of ESMs, these models need to be developed together with use cases for science, climate services, and decision-making that control the priorities of development.

This article describes EC-Earth3, an Earth system model with the flexibility of different configurations that allow users to consider (or exclude) various climate feedbacks and processes. It has been developed collaboratively by the European research consortium EC-Earth to provide a community of European research institutes and universities with an integrated state-of-the-art tool for Earth system studies. While its development goals were largely motivated by the Coupled Model Intercomparison Project phase 6 (CMIP6; Eyring et al., 2016), its suite of ESM configurations allows exploration of a broad range of climate science questions.

The predecessor system EC-Earth2 (Hazeleger et al., 2012) approached the concept of "seamless prediction" to forge models for weather forecasting and climate change studies into a joint system. EC-Earth version 2.2 was based on an adapted version of the atmosphere model IFS 31r1, the Integrated Forecasting System of the European Centre for Medium-Range Weather Forecasts (ECMWF), as used in their seasonal prediction system 3. In addition, a configuration including the atmospheric composition model TM5 was developed (van Noije et al., 2014) and released as EC-Earth version 2.4. EC-Earth2 has been used for simulations under CMIP5 and in a range of climate studies (e.g., Koenigk et al., 2013; Seneviratne et al., 2013).

The current version EC-Earth3 for CMIP6 still leans on the original idea of a climate model system based on the seasonal prediction system of ECMWF. Development started in 2012 by redesigning the software infrastructure and updating the atmosphere model to IFS 36r4, corresponding to the ECMWF seasonal prediction system 4. Since then, various updates, improvements, and forcings have been implemented; the model has been tuned for several intermediate versions and finally for the CMIP6 version, EC-Earth3.

Adaptation of IFS for EC-Earth follows up on the strategy of mutual benefits between short- and medium-range weather prediction on the one hand and longer-timescale climate prediction and projection on the other. While short-term processes and feedbacks are expected to be covered well in the seasonal prediction system, longer-term conservation and trends are the focus of climate model development. During the development process, EC-Earth has been able to feed back valuable information to ECMWF. Examples are a stochastic physics tendency conservation fix for humidity and energy (Leutbecher et al., 2017) forcing (tropospheric and stratospheric aerosol, ozone) and an implementation of aerosol forcing as used in CMIP6 ("MACv2-SP").

The EC-Earth ESM exists in different coupled configurations that reflect a variety of study options and science interests. The system comes with a pure physical core configuration in the form of a global climate model (GCM)

with a range of options: a GCM with prescribed or interactively coupled dynamic vegetation, a dynamical Greenland Ice Sheet, and a closed carbon cycle. Also, a configuration with interactive aerosols and atmospheric chemistry is available, and GCM configurations have been established in different resolutions for the atmosphere and ocean.

As a community model, EC-Earth3 is run on several different high-performance computing (HPC) platforms. While expecting the same simulated climate on each machine, we cannot expect binary identical results. To ensure consistency between different machines, a test protocol and statistical test procedure have been designed.

This paper describes the EC-Earth3 model concept and provides an overview of its component models and the range of available coupled configurations. Specific configurations will be described in more detail in forthcoming papers. The model's physical performance is illustrated based on the core GCM configurations, with a focus on results from historical simulations performed under the CMIP6 protocol. The EC-Earth consortium consists of 27 partners in 10 European countries.

2 Configurations

2.1 The model architecture and coupling framework

EC-Earth is a modular Earth system model (ESM) that is collaboratively developed by the European consortium with the same name. The current generation of the model, EC-Earth3, was developed after CMIP5, and version 3.3 is used for CMIP6 experiments.

EC-Earth3 comprises model components for various physical domains and system components describing atmosphere, ocean, sea ice, land surface, dynamic vegetation, atmospheric composition, ocean biogeochemistry, and the Greenland Ice Sheet. The component models are described in Sect. 3. The atmosphere and land domains are covered by ECMWF's IFS cycle 36r4 (based on IFS system 4, https://www.ecmwf.int/sites/default/files/elibrary/2011/11209-new-ecmwf-seasonal-forecast-system-system-4.pdf, last access: 18 March 2022), which is supplemented with a coupling interface to allow boundary data exchange with other components (ocean, dynamic vegetation, aerosols, and atmospheric chemistry). The NEMO3.6 (Madec and the NEMO team, 2008; Madec et al., 2015) and LIM3 (Vancoppenolle et al., 2009; Rousset et al., 2015) models are the ocean and sea ice components, respectively. Biogeochemical processes in the ocean are simulated by the PISCES model (Aumont et al., 2015). Both LIM3 and PISCES are codewise integrated in NEMO. Dynamical vegetation, land use, and terrestrial biogeochemistry are provided by LPJ-GUESS (Smith et al., 2014; Lindeskog et al., 2013). Aerosols and chemical processes in the atmosphere are described by TM5 (van Noije et al., 2014). The ice sheet model PISM (Bueler

and Brown, 2009; Winkelmann et al., 2011; The PISM Team, 2019) is optionally utilized to model the Greenland Ice Sheet.

An overview of five ESM model configurations is given in this section. Descriptions are schematic, and more detailed specifications will be given in forthcoming publications. Table 1 lists the configurations and their composition, while Table 2 shows the commonly used resolutions for CMIP6.

Most of the model components are coupled through the OASIS3-MCT coupling library (Craig et al., 2017), while some software components include more than one model component, e.g., the sea ice model being a part of the ocean model. A new coupling interface has been developed and implemented to allow a flexible exchange between the model components (see Sect. 3). The OASIS3-MCT coupler provides a technical means of exchanging (sending and receiving) two- and three-dimensional coupling fields between different model components on their different grids. Of the above-named model components, NEMO, LIM3, and PISCES exchange data directly via shared data structures. Thus, EC-Earth3 is implemented following a multi-executable MPMD (multiple programs, multiple data) approach. The model components run concurrently, and a message-passing interface (MPI) is used for parallelization within the components. A potential configuration of all components is illustrated in Fig. 1, which also shows coupling links and frequencies. Note that a configuration including all possible components is not implemented in practice.

In order to manage different configurations at both build time and runtime, EC-Earth3 includes tools to store and retrieve configuration parameters for different model configurations, computational platforms, and experiment types. This allows consistent control of the build and run environments and improves reproducibility across platforms and use cases.

Initial and forcing data (see Table 13), in the form of data files, are provided centrally for the EC-Earth community, and the data are versioned and check-summed for reproducibility.

For EC-Earth3 a tool was developed to convert the native model output to CF-compliant ("Climate and Forecast" standard) NetCDF format (i.e., Climate Model Output Rewriter, CMOR), thus fulfilling the CMIP6 Data Requests for the model intercomparison projects (MIPs) that the community is contributing to (van den Oord et al., 2022, 2017, https://github.com/EC-Earth/ece2cmor3/ (last access: 18 March 2022), https://doi.org/10.5281/zenodo.1051094, van den Oord, 2017).

2.2 Basic configurations EC-Earth3 and EC-Earth3-Veg

EC-Earth3 is the standard configuration consisting of the atmosphere model IFS (Sect. 3.1) including the land surface module HTESSEL (Sect. 3.2) and the ocean model NEMO3.6 with the sea ice module LIM3 (Sect. 3.5). Coupling variables are communicated between the different com-

Table 1. Configurations of the EC-Earth model for CMIP6; the name of the configuration is used as source_id in the CMIP6 context.

Configuration	Atmosphere + land surface	Ocean + sea ice	Dynamic vegetation	Atmospheric composition	Ocean biogeo-chemistry	Greenland Ice Sheet
	IFS 36r4 + HT-ESSEL	NEMO3.6 + LIM3	LPJ-GUESS	TM5	PISCES	PISM
EC-Earth3	x	x				
EC-Earth3-Veg	x	x	x			
EC-Earth3-AerChem	x	x		x		
EC-Earth3-CC	x	x	x	x (CO_2 only)	x (CO_2 fluxes and mixing ratio only)	
EC-Earth3-GrIS	x	x	x			x

Table 2. Commonly used resolutions for CMIP6. The suffixes LR and HR are added to the name of the model configuration where applicable (e.g., EC-Earth3-Veg-LR).

	Resolution atmosphere	Resolution ocean	Time step
Standard resolution	T255L91 ($\sim 80\,$km)	ORCA1L75 (1°)	2700 s (atm) 2700 s (oce) 2700 s (coupling)
Low resolution (EC-Earth3-LR and EC-Earth3-Veg-LR)	T159L62 ($\sim 125\,$km)	ORCA1L75 (1°)	3600 s (atm) 2700 s (oce) 10 800 s (coupling)
High resolution (EC-Earth3P-HR and EC-Earth3-HR)	T511L91 ($\sim 40\,$km)	ORCA025L75 (0.25°)	900 s (atm) 900 s (oce) 2700 s (coupling)

ponent models (see Sect. 3) via the OASIS3-MCT coupler. The physical interfaces are defined by specifying the variables exchanged and the algorithms used.

At the atmosphere–ocean interface, we follow the principle that the ocean provides state variables and the atmosphere sends fluxes (Table 3). Flux formulations correspond to the documentation of IFS CY36R1, Sect. 3, at https://www.ecmwf.int/en/publications/ifs-documentation (last access: 18 March 2022). Atmosphere fluxes are remapped onto the ocean grid by a nearest-neighbor distance-based Gauss-weighted interpolation. The energy (solar and non-solar radiation) and mass (evaporation and precipitation) fluxes are treated with a conservation post-processing method during coupling, in which the residual (target minus source grid integrals) is distributed over the target grid proportional to the original interpolated value. This does not constitute a locally conservative method, but it does conserve mass and energy of the coupling fields. The "sensitivity of non-solar heat flux" in Table 3 refers to the sensitivity with respect to sea ice surface temperature. The variable is used by the sea ice model to distribute the non-solar heat fluxes over different ice categories.

The freshwater runoff from land to ocean is derived from a runoff mapper (Table 4). It uses OASIS3-MCT to interpolate local runoff and ice-shelf calving (from Greenland and Antarctica) to the ocean. The runoff and calving received from the atmosphere and from the surface model HTESSEL are interpolated onto 66 hydrological drainage basins, remapped onto an intermediate grid by the same method and same post-processing as described above for the mass flux. The resulting runoff to the ocean is evenly and instantaneously distributed along several ocean coastal points connected to each hydrological basin in the vicinity of the major river outlet. The runoff is even distributed vertically. The distribution depths are taken (read in from a file) from an ocean-only simulation using a feature of NEMO to save these depths when the NEMO input parameter ln_rnf_depth_ini is set to true in the namelist. For a detailed description of the method we refer to the NEMO documentation (https://www.nemo-ocean.eu/doc/node53.html, last access: 18 March 2022).

In order to avoid a significant long-term sea surface height reduction in coupled model runs due to a net precipitation–
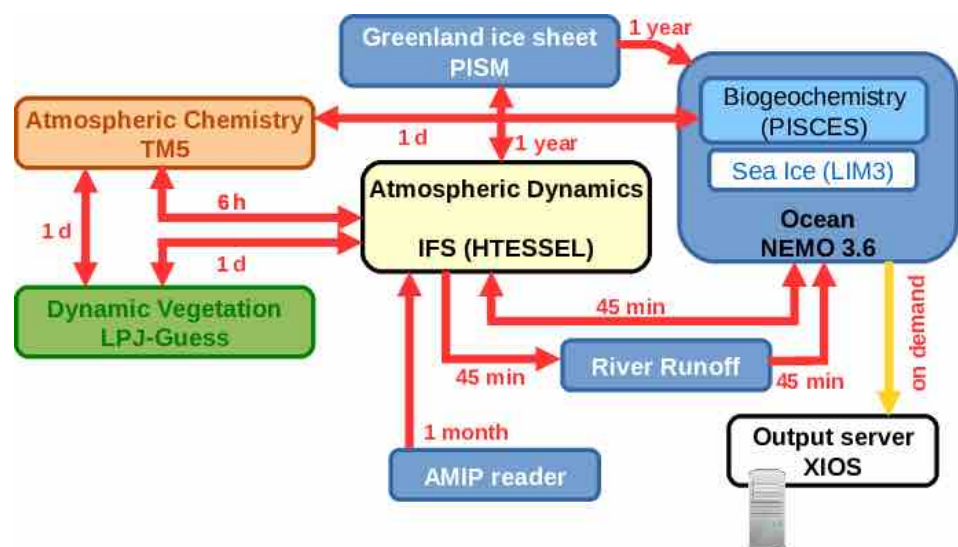
Figure 1. Coupling links and typical frequencies at standard resolution between all components that potentially can be coupled. Existing configurations include subsets of component models and associated couplings.

Table 3. Variables and fluxes exchanged at the ocean–atmosphere interfaces.

Atmosphere → Ocean	Ocean → Atmosphere
Momentum flux	Sea surface temperature
Heat flux solar + non-solar	Sea ice concentration
Evaporation	Sea ice temperature
Precipitation liquid + solid	Sea ice albedo
Sensitivity of non-solar heat flux over ice	Sea ice thickness (not used)
	Snow thickness on sea ice

Table 4. Variables and fluxes provided to the ocean via the runoff mapper.

Atmosphere → Runoff mapper	Runoff mapper → Ocean
Runoff	Runoff
Excess snow	Calving

evaporation ($P-E$) imbalance in the EC-Earth3 atmosphere of about $-0.016\,\mathrm{mm\,d^{-1}}$ in the historical period, the coupled model implements a runoff flux corrector, which amplifies river runoff by 7.95 % in order to compensate for this effect. The compensating flux by the corrector is calculated separately for different resolutions, since different resolutions give different results. The effects are also described in the section "Low-resolution configuration". Correctors are derived for observed climate and applied throughout future scenario periods without change. Comparing the $P-E$ imbalance in CMIP6 historical runs with $4\times CO_2$ experiments, we find that the $P-E$ imbalance does not change significantly.

EC-Earth3-Veg is a configuration extending EC-Earth3 by the interactively coupled second-generation dynamic global vegetation model LPJ-GUESS, which is described together with the coupling principles in Sect. 3.3. Here we provide the variables exchanged through the coupler.

The coupling interface between the atmosphere and vegetation (Table 5) is characterized by the atmospheric model sending the driving variables, as well as selected biogeophysical soil parameters computed within HTESSEL. LPJ-GUESS returns vegetation parameters for both high and low vegetation categories needed for computing surface energy and water exchange in HTESSEL. This ensures that EC-Earth makes best use of both the advanced biophysics in the HTESSEL land surface model and the state-of-the-art vegetation dynamics, land use functionality, and terrestrial biogeochemistry (carbon and nitrogen) in LPJ-GUESS. Since HTESSEL and LPJ-GUESS have very different soil water schemes (LPJ-GUESS updates soil moisture separately in each patch and stand type for each grid cell – see Sect. 3.3, whereas HTESSEL simulates soil moisture per grid cell), the water cycle is discontinuous and each model operates its own water cycle. The water cycle of LPJ-GUESS is thus loosely coupled to the rest of EC-Earth by means of the driving variables sent by HTESSEL/IFS (Boysen et al., 2021). However,

the conservation of moisture in the climate system is not affected by coupling to LPJ-GUESS.

The land mask for the atmosphere is binary and is derived from GTOPO30 (see Sect. 11.2 and 11.4 in the IFS documentation (https://www.ecmwf.int/en/publications/ifs-documentation, last access: 18 March 2022). The ocean mask is binary as well, and a remapping of coupling fields between the atmosphere and ocean grid is carried out by the coupler OASIS_MCT.

2.2.1 Atmospheric tuning of EC-Earth3 and EC-Earth3-Veg

The atmospheric component of EC-Earth has been tuned with the goal of achieving a reasonably small radiative imbalance at the top of the atmosphere (TOA) at standard resolution (T255L91 – to which we refer in the following) in present-day atmosphere stand-alone (AMIP) runs using the CERES_EBAF_Ed4.0 dataset as a reference (Loeb et al., 2018). In particular, the goal was to minimize the mean weighted absolute error in the global means of the net radiative flux at the surface, the TOA longwave flux, longwave cloud forcing, and shortwave cloud forcing, with the first two fluxes considered the most important. The net radiative flux at the surface included the latent heat contribution associated with snowfall, which is not included in the latent heat flux stored by IFS. A series of convective and microphysical atmospheric tuning parameters was identified that are listed in Table 6. Similar parameters have also been commonly used for the tuning of other climate models (e.g., Mauritsen et al., 2012). An additional critical radius for the autoconversion process of liquid cloud droplets, added in EC-Earth3, was considered for tuning (see Rotstayn, 2000, for a discussion on the use of such parameters for model tuning). Changes in the tuning parameters have been adopted to avoid values too different from the original IFS CY36R4 values. In order to proceed with tuning, the sensitivity of the model radiative fluxes to changes in these parameters was determined through a series of short (6 years) AMIP runs for present-day conditions. The resulting linear sensitivities considerably accelerate the tuning process and reduce the number of simulations needed, allowing for the construction of a linear "tuning simulator" used to predict the impact of different combinations of tuning parameter changes on the target radiative fluxes and the determination of combinations providing an optimal score. An iterative process was followed, alternating the construction of new sets of tuning parameters using the known sensitivities, AMIP tuning runs for present-day conditions (20 years, from 1990 to 2010), and the following construction of a new set of tuning parameters to correct the residual biases, allowing for rapid convergence to a desired radiative balance. During this process model biases in other fields were monitored using a Reichler and Kim (2008) metric. Following a suggestion by ECMWF, we reintroduced a condensation limiter for clouds

in the code, which had been removed in CY36R4 but then reintroduced in later cycles starting from CY37R2. Apart from improving the upper-tropospheric distribution of humidity in IFS, this change has an important impact on radiative fluxes (more than $+1.6\,\mathrm{W\,m^2}$ in net flux at TOA), making it a useful tool for tuning the global radiative balance. The atmospheric tuning process showed that energy conservation in IFS is severely dependent on the time step used. For example, at standard resolution, reducing the time step from 2700 to 900 s changes net surface fluxes by $-2\,\mathrm{W\,m^2}$, mainly due to an increase in low clouds, possibly due to resolution-dependent parameterizations. This issue has been improved in later operational versions at ECMWF. In final model configurations time steps ranging from 900 s (high resolution) to 3600 s (low resolution) have been used; see Table 2.

A similar tuning procedure was also used to find alternative tuning parameter sets for other configurations (EC-Earth3-AerChem, EC-Earth3-LR, and EC-Earth3-Veg-LR). The atmospheric tuning for EC-Earth3 and EC-Earth3-Veg is the same, as is the case for EC-Earth3-LR and EC-Earth3-Veg-LR. This is because the vegetation fields used for EC-Earth3 were derived from dynamic vegetation model runs. Therefore, there are only very small differences between the two configurations (with and without dynamic vegetation) for each resolution in terms of the impact of vegetation on the global energy balance.

2.2.2 Coupled tuning of EC-Earth3 and EC-Earth3-Veg

In parallel to forced ocean tuning experiments, the tuned atmosphere was used in coupled present-day experiments researching optimal ocean parameters that allow for a realistic ocean circulation. See Sect. 3.5 for details on some of the changes developed in this phase.

Tuning the final coupled model was aimed primarily at obtaining a realistic global climate at equilibrium in CMIP6 pre-industrial experiments, focusing in particular on the sea ice distribution and extent, the near-surface air temperature distribution, atmospheric variability, the sea surface temperature (SST) distribution (in particular the Southern Ocean temperature bias), and ocean transport due to the Atlantic Meridional Overturning Circulation (AMOC), while at the same time reaching a realistic average global temperature at equilibrium (286.7 to 286.9 K following IPCC; Hoegh-Guldberg et al., 2018; Hawkins et al., 2017; Brohan et al., 2006). The goal was to maintain the same atmospheric tuning as much as possible and only modify the ocean and sea ice parameters. A common set of tuning parameters suitable for both EC-Earth3 and EC-Earth3-Veg experiments was sought. To this end we performed both a range of pre-industrial simulations and, for comparison, corresponding present-day simulations (using fixed 1990 forcing fields and compared to 2010 observations). Gregory plots (Gregory et al., 2004) were used to compare different coupled experi-

Table 5. Variables exchanged between the atmosphere and the vegetation model, with a coupling frequency of 1 d (in standard and low resolution).

Atmosphere → Vegetation	Vegetation → Atmosphere
2 m temperature	Dominant vegetation type low + high
Precipitation	Leaf area index low + high
Soil temperature (four layers)	Vegetation cover low + high
Shortwave radiation	
Longwave radiation	

Table 6. Atmospheric tuning parameters changed in EC-Earth compared to IFS CY36R4. The table reports the new values adopted for T255L91 EC-Earth3 and EC-Earth3-Veg tuning.

IFS parameter name	Description	EC-Earth3	IFS CY36R4
RPRCON	Rate of conversion of cloud water to rain	1.34×10^{-3}	1.4×10^{-3}
ENTRORG	Entrainment in deep convection	1.7×10^{-4}	1.8×10^{-4}
RVICE	Fall speed of ice particles	0.137	0.15
RLCRITSNOW	Critical autoconversion threshold for snow in large-scale precipitation	4.2×10^{-5}	5.0×10^{-5}
RSNOWLIN2	Constant governing of the temperature dependence of the autoconversion of ice crystals to snow in large-scale precipitation	0.035	0.025
ENTRDD	Average entrainment rate for downdrafts	3.0×10^{-4}	2.0×10^{-4}
RMFDEPS	Fractional mass flux for downdrafts	0.3	0.35
RCLDIFF	Mixing coefficient for turbulence, controls cloud cover	3.6×10^{-6}	3.0×10^{-6}
RLCRIT_UPHYS	Critical droplet radius for autoconversion in large-scale precipitation	0.875×10^{-5}	0.93×10^{-5}

ments, to anticipate their approximate equilibrium temperatures even when only partial results were available, and to derive suggested corrections to the global net radiative forcing. The main change which was adopted during this stage was an improved pre-industrial aerosol climatology produced with a different calculation of the sea spray source, characterized by a stronger dependence on surface wind speed (reverting from the formulation of Salisbury et al., 2013, to that of Monahan et al., 1986) and by a dependence on sea surface temperature, following Salter et al. (2015). These changes increased sea spray production over the Southern Ocean and helped to reduce the Southern Ocean SST bias. Details about the revised parameterization are given by van Noije et al. (2021). Finally, a further minor change was a small reduction of thermal conductivity of snow in LIM3 (rn_cdsn = 0.27).

An interesting observation for pre-industrial equilibrium simulations is that at equilibrium we expect radiative balance at TOA and at the surface on average, but we have to take into account two additional effects: (1) while NEMO takes into account the temperature of incoming and outgoing mass fluxes (rainfall, snowfall, evaporation, and runoff fluxes) to represent dilution effects, IFS does not account for the heat content of the moisture field and of precipitation, leading to a missing closure of the global heat budget, corre-

sponding to a heat sink in the ocean. (2) NEMO includes a representation of geothermal energy sources. Estimating the total heat imbalance in the ocean by comparing the ocean heating rate of increase with the net flux at the surface in a pre-industrial experiment leads to a total estimate of about $-0.2\,\mathrm{W\,m^2}$ (as a global average). This energy sink compensates to a large extent for internal energy production observed in IFS (as the difference between the net TOA and net surface radiative fluxes) of about $0.25\,\mathrm{W\,m^2}$, explaining the TOA net flux close to 0 of EC-Earth3 in pre-industrial experiments.

The spin-up of the coupled model prior to the final tuning was a continuous process during the years of development. Long runs were started as soon as a promising candidate version was available. After updating tuning parameters new runs were continued from the end of the previous run, assuming the changes to the model have only an incremental effect. This restart–stop–evaluation cycle was repeated before a final spun-up version was available that allowed for the start of the piControl experiment. The entire length of all simulations is 1100 years, with the last chunk – done with the same model configuration as the CMIP6 experiments – stretching over 250 years.

2.2.3 Low-resolution configurations

EC-Earth3-Veg-LR is a configuration with interactive vegetation (using LPJ-GUESS) feedback at low resolution (T159 for IFS and 1° for ORCA/NEMO). This configuration is applied in the Paleoclimate Modelling Intercomparison Project (PMIP; Kageyama et al., 2018). The major aim of PMIP is to understand the response of the climate system to different climate forcings and feedbacks in the last millennium and in earlier periods. This requires substantial computational resources for multiple multi-centennial simulations. EC-Earth3-Veg-LR makes this possible with reduced resolution. In addition to resolution differences, new physical parameterizations are also included, and tuning parameters are further modified following the same strategy described in the previous paragraph.

Compared to the corresponding configuration with the standard resolution (EC-Earth3-Veg), additional parameter adjustments are introduced to allow for paleoclimate simulations. The adjustments mainly include two parts. Most importantly, orbital forcing parameters are made variable in time. In other configurations used for centennial-scale simulations, these parameters are treated as constants representing present-day climate. That approximation does not hold for multi-centennial to millennial timescales. The new variable calculation for the orbital parameters is taken and modified from CAM3.0 (2004) using the method of Berger (1978). The annual and diurnal cycles of solar insolation are calculated with a repeatable solar year of 365 d and with a mean solar day of exactly 24 h, respectively. This adjusted formulation facilitates paleoclimate simulations for any time within 10^6 years of 1950 CE. A more detailed description of the implemented variable orbital parameters is provided in Sect. 3.1.

Another adjustment is related to the description of glaciers and the Greenland Ice Sheet. In the standard-resolution configuration EC-Earth3-Veg, the physics of land ice are not accounted for. This is not appropriate for paleoclimate simulations. Therefore, a land ice physics package is implemented describing surface physics and time-varying snow albedo over land ice (except for Antarctica) without including a dynamic ice sheet model. More details are provided in the description of EC-Earth3-GrIS below in Sect. 2.6.

Due to the revised parameterizations and reduced resolution (including the different time step), key quantities and model biases are different from the standard configuration EC-Earth3-Veg. Therefore, the EC-Earth3-Veg-LR configuration requires a separate tuning. The difference between net TOA and net surface radiative fluxes is almost independent of the tuning and only depends on the resolution. In the standard resolution, the difference is of the order of $-0.25\,\mathrm{W\,m^{-2}}$, while the difference increases to about $0.3\,\mathrm{W\,m^{-2}}$ for the low resolution.

Rather than tuning towards the currently observed transient climate state with a global mean imbalance of the order of $0.5\,\mathrm{W\,m^2}$ at the TOA (Hansen et al., 2011), we aimed at a tuning of a climate in radiative equilibrium to prevent the global mean surface temperature from drifting too much under the conditions of a stable climate. This approach is necessary for millennium-scale simulations. We aimed at a net surface energy balance close to $0\,\mathrm{W\,m^2}$ under pre-industrial-level forcing (1850) after hundreds of years of spin-up. Thereby, we mainly focused on the net surface energy (SFC) balance rather than the TOA energy budget as we know that the atmospheric model is not fully conservative. The resulting parameter combination, together with historical simulations, will be described in a forthcoming paper in conjunction with partners in the EU-Crescendo project.

In order to avoid a significant long-term sea surface height reduction in coupled model runs due to a net precipitation–evaporation (P–E) imbalance in the EC-Earth3 atmosphere of about $-0.0174\,\mathrm{mm\,d^{-1}}$ in the historical period, the coupled model implements a runoff flux corrector, which amplifies river runoff by 8.65 % in order to compensate for this effect.

In addition to the EC-Earth3-Veg-LR configuration, there is also a configuration without interactive vegetation, EC-Earth3-LR. In this configuration, vegetation is prescribed by the Paleo-MIP (PMIP). These two configurations produce very similar results when EC-Earth3-LR is forced by the vegetation from a corresponding EC-Earth3-Veg-LR simulation. The tuning parameters are identical in both configurations.

2.2.4 High-resolution configurations

Earlier studies with EC-Earth at high resolution using EC-Earth 3.1 have shown improvements with resolution, e.g., in North Atlantic blocking (Davini et al., 2017b) and in the representation of tropical rainfall extremes (Davini et al., 2017a). This motivated further development of the EC-Earth3 configuration in high resolution, with increased atmospheric and oceanic resolution, derived from an earlier state of development. It features a T511 spectral resolution for IFS and 0.25° resolution for ORCA/NEMO. A preliminary tuned version, EC-Earth3P-HR, is used in current projects and in CMIP6 MIPs. Another high-resolution configuration, EC-Earth3-HR, closer to the EC-Earth3 base configuration, is still under development. Here we focus on the configuration EC-Earth3P-HR, which so far has been better documented.

At an early stage of development, EC-Earth3P-HR was branched off from the main line in order to apply it for the EU project PRIMAVERA and the HighResMIP endorsed by CMIP6. PRIMAVERA and HighResMIP are focusing on the impact of horizontal resolution on the simulation of climate and its variability. The HighResMIP protocol requires modifications of the standard configuration to allow for a clean assessment of the impact of horizontal resolution. The motivation and a detailed description of those deviations from the base version, EC-Earth3, can be found in Haarsma et

al. (2020). Below we give a short summary of the most important deviations of EC-Earth3P-HR.

–   The stratospheric aerosol forcing is handled in a simplified way that neglects the details of the vertical distribution and only takes into account the total aerosol optical depth in the stratosphere, which is then evenly distributed across the stratosphere. No indirect aerosol effect has been implemented.

–   A SST and sea ice forcing dataset specially developed for HighResMIP is used for AMIP experiments (Kennedy et al., 2017). The major differences compared to the standard SST forcing datasets for CMIP6 are the higher spatial (0.25° vs. 1°) and temporal (daily vs. monthly) resolution.

–   The vegetation and its albedo are prescribed as present-day climatologies that are constant in time.

Under HighResMIP, simulations are performed with EC-Earth3P-HR in high resolution and in the standard-resolution EC-Earth3P (T255 for IFS and 1.0° for ORCA/NEMO). A full description of EC-Earth3P-HR including technical implementation and post-processing can be found in Haarsma et al. (2020). EC-Earth3P-HR was not tuned differently compared to the standard resolution at the time due to very high computational demands. This approach is consistent with most other models in Europe, as represented in the H2020 PRIMAVERA project (Roberts et al., 2018).

Based on results of Haarsma et al. (2020), increasing horizontal resolution does not result in a general reduction of biases and overall improvement of the climate variability. Deteriorating impacts can be detected for specific regions and phenomena such as some Euro-Atlantic weather regimes, whereas others such as the El Niño–Southern Oscillation show a clear improvement in their spatial structure. Analysis of the kinetic energy spectrum indicates that the sub-synoptic scales are better resolved at higher resolution (Klaver et al., 2020) in EC-Earth.

Despite a lack of clear improvement with respect to biases and synoptic-scale variability for the high-resolution version of EC-Earth, the better representation of sub-synoptic scales results in better representation of phenomena and processes on these scales such as tropical cyclones (Roberts et al., 2020) and ocean–atmosphere interaction along western boundary currents (Belluci et al., 2021). The impact of resolution for EC-Earth and other climate models participating in HighResMIP will be analyzed more in-depth in upcoming publications.

2.3   EC-Earth3-AerChem

EC-Earth3-AerChem (van Noije et al., 2021) is the configuration with interactive aerosols and atmospheric chemistry used in the Aerosol and Chemistry Model Intercomparison

Project (AerChemMIP; Collins et al., 2017). In this configuration, TM5 is used to simulate tropospheric aerosols and chemistry based on the CMIP6 emission pathways for aerosols and chemically reactive gases. The resolution of TM5 is $3 \times 2°$ (longitude × latitude) with 34 vertical levels and a top at 0.1 hPa. IFS and NEMO have the same resolutions as in the standard configuration. TM5 and IFS exchange fields with a 6 h frequency. TM5 receives a large set of 2D and 3D meteorological fields from IFS and provides 3D distributions of aerosols, ozone (O_3), and methane (CH_4) in return. Table 7 lists the fields exchanged between IFS and TM5 through the coupler.

2.4   EC-Earth3-CC

EC-Earth3-CC is the configuration that includes a description of the carbon cycle, which is used for the Coupled Climate–Carbon Cycle Model Intercomparison Project (C4MIP; Jones et al., 2016). EC-Earth3-CC allows simulations with emissions forcing rather than with prescribed concentrations only as in the ScenarioMIP. This configuration uses a single carbon tracer in the atmosphere, advected by a version of TM5 with a reduced number of vertical levels (10 instead of 34), to simulate the transport of CO_2 through the atmosphere. The resolution and coupling frequency for the exchange between IFS and TM5 are the same as for the interactive aerosols and chemistry version of TM5 (EC-Earth3-AerChem) described in the preceding section. In effect the data transfer in both directions is much reduced. The CO_2 exchange with the ocean and terrestrial biosphere is calculated in PISCES and LPJ-GUESS, respectively, based on surface mixing ratios from the previous day received from TM5.

PISCES calculates the air–sea CO_2 flux at every time step after solving for carbon chemistry in seawater. This flux is proportional to the difference in pCO$_2$ between the atmosphere and the surface of the ocean. The exchange of CO_2 between the ocean and TM5 is realized once a day after accumulating the flux over each grid cell over 24 h. Furthermore, physical transport of passive tracers in the ocean presents a slight artificial mass imbalance. To prevent it from becoming significant for carbon during the spin-up we applied a uniform correction to dissolved inorganic carbon at the end of each year, after taking into account all sources and sinks.

A variant of EC-Earth-CC can also be run concentration-driven by excluding TM5. PISCES and LPJ-GUESS then read a uniform global atmospheric CO_2 concentration.

Tables 8 and 9 list the fields exchanged between the CTM on the one hand and the vegetation and ocean biogeochemistry models on the other hand through the coupler.

2.5   EC-Earth3-GrIS

EC-Earth3-GrIS is a configuration that couples the EC-Earth3-Veg to the Parallel Ice Sheet Model v1.1 (PISM, Sect. 3.8). It is used to model the Greenland Ice Sheet (GrIS)

Table 7. Variables exchanged with a 6 h frequency between the atmosphere and the chemical transport model (CTM) TM5 in EC-Earth3-AerChem.

Atmosphere → CTM	CTM → Atmosphere
Logarithm of surface pressure	Ozone mixing ratio
Vorticity (3D)	Methane mixing ratio
Divergence (3D)	Aerosol number and component mass mixing ratios (25 fields in total)
Surface orography	Aerosol extinction (14 wavelengths)
Surface pressure	Aerosol single-scattering albedo (14 wavelengths)
Air temperature (3D)	Aerosol asymmetry factor (14 wavelengths)
Specific humidity (3D)	
Cloud liquid/ice water content (3D)	
Cloud area fraction (3D)	
Overhead/underfoot cloud area fraction (3D)	
Updraft/downdraft convective air mass flux (3D)	
Updraft/downdraft convective air mass detrainment rate (3D)	
Land–sea mask	
Surface albedo	
Surface roughness length	
Sea ice fraction	
Sea surface temperature	
10 m wind speed	
Skin reservoir water content	
2 m temperature	
2 m dew-point temperature	
Surface latent heat flux	
Surface sensible heat flux	
Eastward/northward surface stress	
Large-scale precipitation	
Convective precipitation	
Surface shortwave radiation	
Snow depth	
Soil wetness in topsoil layer	
Vegetation type fraction (15 categories)	
High/low vegetation cover	

evolution and its feedback with the climate system in the Ice Sheet Model Intercomparison project (ISMIP6; Nowicki et al., 2016). GrIS handles the ice sheet dynamical and thermodynamical processes, including ice flow, subglacial hydrology, bed deformation, and the basal ice melt.

In the configurations EC-Earth3 and EC-Earth3-Veg, ice sheets are represented by a perennial snow layer of 9 m water equivalent. Snowfall on these areas is immediately redistributed into the ocean as ice to prevent excessive snow accumulation. Perennial snow albedo and snow density are fixed at 0.8 and $300 \, \mathrm{kg} \, \mathrm{m}^{-3}$, respectively, and the snowpack is in thermal contact with the underlying soil. In EC-Earth3-GrIS, the surface parameterization in EC-Earth3 is adjusted in order to better account for the presence of the ice sheet. The modifications include introduction of an explicit ice sheet mask obtained from PISM into HTESSEL and application of values representative of an ice sheet to calculate the surface energy balance and subsurface heat and energy transfer for glacierized grid points. In addition, if a grid cell is with an ice sheet but no snow cover (i.e., bare ice), the ice

can melt and contribute to surface runoff if the energy flux at the surface is positive. Furthermore, a time-varying snow albedo parameterization is introduced for snow on ice sheets (Helsen et al., 2017) in EC-Earth3-GrIS. The parameterization allows the dependence of snow albedo on snow aging, melt, and refreezing. For fresh snow a maximum value of 0.85 is used. Under dry non-melting conditions, aging may reduce the snow albedo to 0.75, and during snowmelt the albedo decreases to a lower limit of 0.6. The albedo of refrozen meltwater is set to 0.65.

The new land ice physics described above are used for EC-Earth3 low-resolution configurations, in particular for PMIP experiments. For other resolutions, there is no coupling to the ice sheet model. Instead, the ice sheet mask can either be read in as boundary conditions or defined by snow depth exceeding a certain threshold (9 m).

The fields exchanged between EC-Earth and PISM are listed in Table 10. Information is exchanged once a year with monthly variations. IFS provides forcing fields of surface mass balance (SMB) and subsurface temperature to PISM.

Table 8. Variables exchanged with a 24 h frequency between the vegetation model LPJ-GUESS and the chemical transport model TM5 in EC-Earth3-CC. * Fluxes occur once a year and are distributed evenly over the following year.

Vegetation → CTM (CC only)	CTM → Vegetation (CC only)	Atmosphere → CTM (every 6 h)	CTM → Atmosphere
Net primary production	CO_2 mixing ratio	Logarithm of surface pressure	CO_2 mixing ratio
Heterotrophic respiration		Vorticity (3D)	
Establishment*		Divergence (3D)	
Reproduction*		Surface orography	
Burned vegetation and litter*		Surface pressure	
Sowing*		Air temperature (3D)	
Fast and slow harvested products*		Specific humidity (3D)	
Land cover change*		Updraft/downdraft convective air mass flux (3D)	
		Updraft/downdraft convective air mass detrainment rate (3D)	
Carbon leaching		Land–sea mask	
		Surface roughness length	
		10 m wind speed	
		Surface latent heat flux	
		Surface sensible heat flux	
		Eastward/northward surface stress	

Table 9. Variables exchanged with a 24 h frequency between the chemical transport model TM5 and the ocean biogeochemistry model PISCES.

Ocean BGC → CTM (CC only)	CTM → Ocean BGC (CC only)
CO_2 flux	CO_2 mixing ratio

The SMB is calculated from precipitation, evaporation, and runoff. PISM returns the ice topography and ice mask to IFS and the calving (mass and energy) and basal melt (mass) fluxes to NEMO.

2.6 HPC of different configurations

The increasing capabilities of ESMs, such as EC-Earth3, and the ability to perform large community experiments, such as CMIP6, are strongly linked to the amount of HPC capacity available and to the efficient use of these resources. As such, CMIP6 is an excellent opportunity to study the computational performance of ESMs, in particular for models such as EC-Earth3 that are developed and used by a wide range of institutions and integrated on different computational platforms.

The computational performance of EC-Earth3 has been evaluated in order to achieve different goals:

– to detect performance bottlenecks for future improvements,

– to compare the performance of different computational platforms used by the consortium and evaluate how different hardware can affect the performance of EC-Earth, and

– to compare different model configurations to analyze which components or calculations represent bottlenecks in the execution.

A first optimization and performance analysis of a preliminary version of EC-Earth3 (EC-Earth3P-HR) was presented in Haarsma et al. (2020). This particular version, which was used in the context of the H2020 PRIMAVERA project (Roberts et al., 2018), was integrated at both standard and high resolutions following the HighResMIP protocol (Haarsma et al., 2016). In Haarsma et al. (2020), the high-resolution configuration was analyzed in order to detect performance bottlenecks and to provide solutions for these. The high resolution was used for this purpose because of easier detectability of problems related to the scalability and computational efficiency.

The rest of this section will focus on the performance of the standard-resolution version of EC-Earth3 in order to fulfill the second and third goals presented.

The evaluation was done through a set of metrics independent of the platform and of the underlying parallel programming models. To make this possible, the EC-Earth standard-resolution configuration discussed hereafter was analyzed through CPMIP, a computational performance model intercomparison project (MIP) presented by Balaji et al. (2017).

Table 10. Variables exchanged between the atmosphere model IFS and the ice sheet model PISM, as well as between the ocean model NEMO and ice sheet model PISM. Information is exchanged once a year with monthly variations.

Atmosphere → Ice sheet	Ice sheet → Atmosphere
Subsurface temperature	Ice topography
Surface mass balance ($P-E-R$)	Ice extent
Ocean → Ice sheet	Ice sheet → Ocean
	Calving fluxes (mass and energy)
	Basal melt flux (mass)

This analysis is done in two levels. The first level (Table 11) includes basic performance metrics for four different platforms (Rhino, RN; Marenostrum4, MN4; ECMWF-CCA, CCA; and Beskow, BK) in order to compare the performance of two configurations (EC-Earth3 and EC-Earth-Veg) on those different platforms. The second level (Table 12) includes the complete set of CPMIP metrics collected on Marenostrum4 for EC-Earth3.

In Table 11 SYPD measures the model speed by counting the number of years the model could simulate within a 24 h period, given a certain configuration and computational platform. ASYPD is measured for a long-running experiment and includes queueing time, taking into account the sharing of HPC resources. CHSY measures the computational cost of the model for the given configuration and computational platform. Finally, parallelization represents the number of MPI processes used.

Comparing EC-Earth3 on two platforms (MN4 and RN) with similar parallelization, the BullX B500 experiment is slightly faster than the LENOVO SD530 experiment, but it also uses more resources and as a consequence the CHSY is slightly higher. On the other hand, we can obtain similar performance in BullX B500 and LENOVO SD530 experiments, even though the BullX B500 experiment is run on a platform with technology 5 years older, proving that configurations without an expensive computation can be simulated efficiently in more commodity clusters. Obviously, as shown in Haarsma et al. (2020), the performance of more demanding configurations will be affected by several issues, such as the MPI communications overhead, and a better network will ensure that better hardware will obtain better performance too. Finally, the experiment with EC-Earth3 on CCA proves that the user can achieve a similar efficiency using a setup with fewer processes and obtaining a similar CHSY, though the results will need more time to be executed.

It is important to note that the workflow of these experiments comprises different steps, with dependencies between them. This is especially true when the storage is a constraint and simulation steps need data from prior steps before post-processing. In such cases, the way these dependencies are handled may have an impact on the overall throughput.

LENOVO SD530 experiments at BSC were run using a workflow management tool called Autosubmit. This tool handles dependencies in an automatic way and is able to pack multiple tasks or simulation steps in the same job execution, which may reduce the amount of job queuing and thus have an impact on the ASYPD. This does not necessarily explain the differences between the three platforms in the study given the different use policies, load on the machine from other users, scheduling parameters, and usage existing among them.

The BullX B500 and CRAY XC40 (on BK platform) experiments are not directly comparable because the CRAY XC40 experiment includes LPJ-GUESS as a vegetation component. Both simulations use the same parallel resources, and the performance of the CRAY XC40 experiment on BK is lower. The results suggest that LPJ-GUESS is less efficient than the other components present in the standard configuration of EC-Earth and that this difference in performance is largely due to the way the output is performed. The problem is to be studied to improve it in the future. A new approach is under development to improve the computational efficiency of LPJ-GUESS. On the other hand, the EC-Earth-Veg configuration run in the CRAY XC40 experiment on CCA suggests that when the execution time of IFS and NEMO components is long enough (since we are using fewer parallel resources for their execution), the LPJ-GUESS component is not a bottleneck anymore, achieving a CHSY only slightly higher. However, the single point to take into account in this case is that the user will need more time to finish the simulations, since the SYPD is lower compared to the setup used on the BK platform.

These results will be used to compare the computational performance of EC-Earth with other models running the same CMIP6 configuration or with a similar complexity. However, preliminary results from the collection (provided by other institutions) prove that the efficiency of EC-Earth (comparing CHSY among models with a similar complexity or number of grid points) seems to show good results on average from the computational performance side. The cost of indirect processes such as coupling or output costs is also similar to the results obtained by other models.

Table 11. Basic CPMIP metrics of EC-Earth3 and EC-Earth3-Veg with standard resolution for four different architectures and platforms: Marenostrum4 (MN4, LENOVO SD530), Rhino (RN, BullX B500), ECMWF-CCA (CCA, CRAY XC40), and Beskow (BK, CRAY XC40). The basic metrics are SYPD (simulated years per day), ASYPD (actual simulated years per day), CHSY (core hours per simulated year), and parallelization (number of MPI processes used).

Architecture	Configuration	Platform	SYPD	ASYPD	CHSY	Parallelization
LENOVO	EC-Earth3	MN4	15.2	9.87	1119	768
SD530	EC-Earth3	RN	16.2	16.2	1276	864
BullX B500	EC-Earth3	CCA	6.03	4.84	1289	324
CRAY XC40	EC-Earth-Veg	BK	12.4	6.65	1676	864
CRAY XC40	EC-Earth-Veg	CCA	6.67	5.32	1332	342
CRAY XC40						

3 The component models

3.1 Atmosphere

The atmosphere component of the EC-Earth model is based on the Integrated Forecast System (IFS) CY36R4 of the European Centre for Medium-Range Weather Forecasts (ECMWF). This specific cycle of the IFS was part of ECMWF's operational seasonal forecast system S4 (https://www.ecmwf.int/sites/default/files/elibrary/2011/11209-new-ecmwf-seasonal-forecast-system-system-4.pdf, last access: 18 March 2022). IFS solves the hydrostatic primitive equations using a two-time-level, semi-implicit, semi-Lagrangian discretization. Horizontal derivatives are computed in spectral space, while the computation of advection, the physical parameterizations, and in particular the nonlinear terms is conducted on the linear reduced Gaussian grid. The IFS is documented extensively at https://www.ecmwf.int/en/publications/ifs-documentation (last access: 18 March 2022) (for example https://www.ecmwf.int/sites/default/files/elibrary/2010/9232-part-iii-dynamics-and-numerical-procedures.pdf (last access: 18 March 2022) for the dynamics and https://www.ecmwf.int/sites/default/files/elibrary/2010/9233-part-iv-physical-processes.pdf (last access: 18 March 2022) for the physical processes). Here we only document the updates to the original IFS that were necessary for making long climate simulations.

The physical aspects of the atmosphere model in EC-Earth needed some adjustments and updates compared to the original IFS CY36R4. Most of these modifications are not necessary for numerical weather prediction (NWP) or even seasonal forecasts but are crucial when running long climate simulations (decadal, centennial, or longer) or simulations under different climate conditions (e.g., future scenarios or paleo-simulations).

The semi-Lagrangian advection scheme of IFS does not conserve mass or energy in the NWP version. A dry air mass conservation fixer has been available in IFS since CY25R1 and is active in EC-Earth to correct global pressure for the gain or loss of atmospheric mass. Similarly, to conserve humidity during transport we backported a simple proportional fixer from IFS cycle CY38R1 (Rasch and Williamson, 1990; Diamantakis and Flemming, 2014). This significantly reduced the bias of the average global precipitation–evaporation balance in the model from about $+0.030$ to $-0.017\,\mathrm{mm\,d^{-1}}$ and consistently (due to the associated latent heat of condensation) in the radiative balance in the atmosphere from about $-1.65\,\mathrm{W\,m^2}$ (a source of energy) to about $-0.25\,\mathrm{W\,m^2}$.

The IFS CY36R4 version adopted for EC-Earth3 produces a reasonable quasi-biennial oscillation (QBO) in the tropical stratosphere when running at the standard resolution (T255L91), but not for any other available horizontal or vertical resolutions. Therefore, we substituted the original version-dependent latitudinal profile of the momentum flux in the non-orographic gravity wave scheme (which was originally developed ad hoc for the ECMWF system 4 seasonal forecast system) with a resolution-dependent parameterization of non-orographic gravity wave drag by backporting changes later introduced in IFS CY40R1 (see Davini et al., 2017a, for more details). This change allowed EC-Earth to recover a realistic QBO at all resolutions considered without deteriorating the jet streams.

Convection in the NWP version of IFS CY36R4 reaches its maximum around local noon in contrast to observations that peak later in the afternoon. A closure described by Bechtold et al. (2014) improving the diurnal cycle of convection has been implemented in EC-Earth3. For EC-Earth3, Rayleigh friction was activated in EC-Earth IFS for all resolutions to avoid unphysically large wind speeds at higher resolution.

In atmosphere-only simulations, the sea ice albedo is taken from a look-up table with climatological monthly values for sea ice albedo (Ebert and Curry, 1993) that take into account the annual cycle of highly reflective snow cover during winter and spring and the darker surface of melting sea ice during summer. In the coupled model, the sea ice albedo is computed in the sea ice model LIM3, and the updated values are used by the atmospheric component. The broadband sea ice albedo from LIM3 is then mapped on six shortwave bands with a mapping function.

The time stepping scheme needed technical adjustments to avoid an overflow of integer time step counters in order to allow making simulations beyond 32 768 time steps. The IFS output is saved in the GRIB1 data format, which also has a limit in the number of time steps that can be saved. This limit was overcome in EC-Earth3 by setting the time step to 0 and updating the GRIB-encoded reference time instead each time that output is written.

CMIP6 requires transient climate forcings to account for the change in atmospheric composition and other external drivers of the climate (e.g., insolation). The necessary interfaces to read the prescribed greenhouse gas concentrations, aerosol optical properties, stratospheric aerosols, stratospheric ozone, and insolation have been implemented in the IFS code in EC-Earth. Table 13 lists the sources and versions of the CMIP6 forcing datasets.

Well-mixed greenhouse gases (WMGHGs) explicitly included in EC-Earth's radiation scheme are CO_2, CH_4, nitrous oxide (N_2O), CFC-12, and CFC-11. Together these are responsible for about 98 % of the total radiative forcing by WMGHGs in 2014 compared to 1850 (Meinshausen et al., 2017). The radiative effects of the remaining WMGHGs (HCFC-22, CFC-113, CCl4) are accounted for in terms of CFC-11 equivalents (Meinshausen et al., 2017). The mixing ratios of each of the WMGHGs that are explicitly included and not provided by TM5 are prescribed by scaling their monthly zonal mean climatologies as used in IFS by a single time-dependent global factor. In this way, the global mean surface mixing ratios are forced to their CMIP6 pathways (Meinshausen et al., 2017). To reduce discontinuities, the scale factors are calculated on a monthly basis by interpolation of the time series of annual values provided by CMIP6. Any delays due to transport from the surface to the upper parts of the atmosphere are ignored in this approach.

Tropospheric aerosols are either simulated interactively in TM5 (in the EC-Earth3-AerChem configuration) or prescribed as a pre-industrial climatology plus an anthropogenic contribution (all other configurations). The pre-industrial aerosol background is specified using a monthly climatology based on TM5. This climatology was obtained from an offline TM5 simulation driven by ERA-Interim meteorology for the years 1981–1985 using CMIP6 anthropogenic emissions for the year 1850. The radiative and cloud effects of the pre-industrial aerosols are calculated based on the ERA-Interim reanalysis and the same set of variables as when aerosols are interactively simulated by TM5. The anthropogenic contribution is specified following the simple plume approach of MACv2-SP (Stevens et al., 2017), which provides a simplified parametric representation of the optical properties (extinction, single-scattering albedo, and asymmetry factor) of the anthropogenic contribution to the tropospheric aerosol burden (relative to 1850 levels), consistent with the CMIP6 time series of historical (Stevens et al., 2017) and future (Fiedler et al., 2019) anthropogenic emissions. In EC-Earth, MACv2-SP is coupled with the IFS

radiation scheme to compute the optical properties for the 14 wavelength bands of the SW radiation. More precisely, the optical properties are calculated at the band mean wavelengths weighted by the incoming solar radiation. In addition, MACv2-SP provides a simple way to account for the effect of anthropogenic aerosols on clouds. Specifically, it provides a scale factor for the cloud droplet number concentration (CDNC) in each column based on the vertically integrated optical depth at 550 nm.

In the EC-Earth3-AerChem, aerosol impacts on clouds are included by calculating CDNC depending on the modal number and mass concentrations from TM5, following Abdul-Razzak and Ghan (2000). For all other model configurations the CDNC corresponds to pre-industrial aerosol conditions and an additional scaling factor from MACv2-SP that is included to account for the cloud forcing by anthropogenic aerosols. The resulting forcing includes contributions due to both cloud reflectivity and cloud lifetime effects, as the lifetime of clouds explicitly depends on CDNC. Currently only the activation and autoconversion of liquid cloud droplets are linked explicitly to ambient aerosol concentrations. For ice clouds the EC-Earth3 model still retains the parameterization from the original IFS CY36R4.

As EC-Earth3 uses MACv2-SP in combination with a pre-industrial aerosol climatology, natural aerosol variability is only accounted for via the prescribed seasonal cycle of the climatology. Furthermore, MACv2-SP only captures the seasonal cycle and long-term changes in the optical properties and the derived CDNC impact factor of anthropogenic aerosols. Diurnal variability in aerosol amounts or properties is not explicitly described. Day-to-day variability is only included to the extent captured by the seasonal cycles of the pre-industrial climatology and MACv2-SP. Of the interannual variability in the amount and properties of anthropogenic aerosols, only the long-term changes in plume strengths, which are assumed to covary with the 11-year averaged emissions of SO_x plus NH_3 in the associated countries, are accounted for. Changes in the spectral distribution of the optical properties, the single-scattering albedo, and asymmetry factor of anthropogenic aerosols due to long-term changes in their size distribution and composition are ignored by MACv2-SP.

Stratospheric aerosols are prescribed using the CMIP6 dataset of aerosol radiative properties, which covers the period 1850 to 2014 and for the more recent period is based on satellite data assembled by Thomason et al. (2018). The dataset consists of monthly resolved zonal mean fields, which are provided at the 14 shortwave (SW) and 16 longwave (LW) bands of the IFS's radiation schemes. For the SW scheme, the extinction, single-scattering albedo, and asymmetry factor are specified, whereas only the absorption is taken into account for the LW scheme, since aerosol scattering in the LW is neglected in the atmospheric component of EC-Earth. Forcing data are vertically interpolated beforehand for the 62- and 91-level configurations, taking

into account the seasonality of model level heights, whereas horizontal and monthly to daily interpolation is done on-line. When interpolating or averaging the radiative property fields, they are first made extensive by including the appropriate weighting factors (e.g., extinction is converted to optical depth, single-scattering albedo to absorption optical depth, and likewise for the asymmetry factor). The forcing located below the online-diagnosed thermal tropopause level is excluded. This implementation is used in all current EC-Earth3 configurations with the exception of the EC-Earth3P-HR configuration, which uses a simplified implementation based on a monthly vertically integrated, latitude-dependent aerosol optical depth (AOD) forcing at 550 nm, which is then vertically distributed across the stratosphere. In both implementations, it is possible to set the forcing fields to a constant background distribution computed as the time average over 1850 to 2014. This background forcing is applied in pre-industrial control and future simulations, as recommended in the CMIP6 protocol.

The land use forcing dataset (LUH2) from CMIP6 (Hurtt et al., 2020) cannot be used directly as input to IFS because it does not provide the same vegetation cover or type categories as those used by the land surface scheme in IFS (HTESSEL; van den Hurk et al., 2000; Balsamo et al., 2009; Dutra et al., 2010; Boussetta et al., 2013) but instead provides agricultural management information and land use transitions that are annually updated. The vegetation cover, leaf area index (LAI), and vegetation type that are needed for the land surface scheme and albedo parameterization in IFS can be simulated by the dynamic vegetation model LPJ-GUESS (Smith et al., 2014). This happens automatically in the EC-Earth3-Veg configuration wherein the dynamic vegetation model, which uses the LUH2 dataset as an input, is active, but for all other configurations the required vegetation cover and type need to be precomputed. This is done by first making all CMIP6 experiments with the EC-Earth3-Veg configuration and saving the vegetation variables that can then be reused when making the same experiment with other model configurations.

The orbital parameters of the original IFS CY36R4 are fixed for present-day conditions following the recommendations of the International Astronomical Union (ARPEGE-Climate Version 5.1, 2008), which is sufficient for simulations of the recent past or near future. However, for paleo-simulations in PMIP the orbital parameters need to be variable or set fixed for a different time period. Orbital parameters and insolation are computed using the method of Berger (1978). Using this formulation, the insolation can be determined for any year within 10^6 years of 1950 CE. The formulation determines the Earth–Sun distance factor and solar zenith angle. The annual cycle and diurnal cycle of solar insolation are represented with a repeatable solar year of exactly 365 d and with a mean solar day of exactly 24 h, respectively. The repeatable solar year does not allow for leap years. The orbital state may be specified in one of two ways.

The first method is to specify a year, which is held constant during the integration for an equilibrium simulation or varies yearly for a transient simulation. The second method is to specify the orbital parameters: eccentricity, longitude of perihelion, and obliquity. This set of values is sufficient to specify the complete orbital state. For example, settings for pi-Control integrations under 1850 CE conditions are obliquity of 23.549, eccentricity of 0.016764, and longitude of perihelion of 100.33.

3.2 Land surface and vegetation

The Hydrology Tiled ECMWF Scheme of Surface Exchanges over Land (HTESSEL; van den Hurk et al., 2000; Balsamo et al., 2009; Dutra et al., 2010; Boussetta et al., 2013) is the land surface model interfacing with the atmospheric boundary layer and solving the energy and water balance at the land surface in EC-Earth. HTESSEL discretization, for each grid point, solves for up to six different land surface tiles that may be present over land (bare ground, low and high vegetation, intercepted water by vegetation, and vegetation-shaded and exposed snow). Surface radiative, latent heat, and sensible heat fluxes are calculated as a weighted average of the values over each tile.

The discretization in HTESSEL is such that coexistence in each grid point of more than one type of low and high vegetation, respectively, is not allowed. Therefore, for each grid point and for both low and high vegetation cover, a dominant type (dominant meaning the type with the higher relative area fraction for either high or low vegetation) is identified, T_l and T_h, and a vegetation coverage for high and low vegetation types, C_h and C_l, is specified.

Vegetation types and vegetation coverage can be

1. prescribed from a static land use map from the Global Land Cover Characteristics (GLCC, standard HTESSEL configuration; van den Hurk et al., 2000; Balsamo et al., 2009; Dutra et al., 2010; Boussetta et al., 2013),

2. interactively provided when coupled with LPJ-GUESS, or

3. prescribed from a previous simulation with LPJ-GUESS.

When the tile fractions are prescribed from GLCC, vegetation density is parameterized according to the Lambert–Beer law of extinction of light under a vegetation canopy and is therefore allowed to change as a function of leaf area index (LAI) for both low and high vegetation as described in Alessandri et al. (2017). Otherwise, LPJ-GUESS provides its own consistently simulated background tile fractions and vegetation densities.

The coupling of biophysical parameters in HTESSEL has been enhanced since CMIP5 (Weiss et al., 2014), for which only the surface resistance to evapotranspiration and water intercepted and directly evaporated from vegetation canopies

Table 12. CPMIP metrics analysis for EC-Earth3 on Marenostrum4. Complete CPMIP metrics are shown in this table: resolution complexity (Cmpx), SYPD (simulated years per day), ASYPD (actual simulated years per day), CHSY (core hours per simulated year), parallelization, JPSY (Joules per year simulated), Coup. C. (coupling cost), Mem. B.(memory bloat), DO (data output cost), and DI (data intensity). From left to right we have resolution (Resol) as the total number of grid points for all the components used (ocean, atmosphere, and so on). Cmpx includes all prognostic variables of a model. JPSY quantifies the energy cost of the execution. Coup. C. represents the cost associated with the coupling among components (including interpolation and communication calculations; 8 % in this case with respect to the total execution time). Mem. B. is the division between the theoretical memory and the real one. DO is the cost of the output process (12 % in this case with respect to the total execution time). DI is the output volume in GB per day of simulation.

Configuration	Resolution (no. grid points)	Cmpx	SYPD	ASYPD	CHSY	Paral.	JPSY (Joules)	Coup. C. (%)	Mem. B.	DO (%)	DI (GB)
EC-Earth3	1.60×10^7	0.31	15.2	9.87	1119	768	4.41×10^7	8	11	12	3

Table 13. CMIP6 forcing datasets used by EC-Earth3 and EC-Earth3-Veg for DECK (diagnostic, evaluation, and characterization of klima) and historical experiments. All datasets are available from https://esgf-node.llnl.gov/search/input4mips/ (last access: 18 March 2022). A more detailed description of the CMIP6 forcing datasets is available at http://goo.gl/r8up31 (last access: 18 March 2022).

Forcing dataset	Version	Further info	Comments
Greenhouse gas concentration	1.2.0	Meinshausen et al. (2017)	
Stratospheric aerosols	3.0.0	Thomason et al. (2018)	
Ozone volume mixing ratio	1.0	http://blogs.reading.ac.uk/ccmi/ forcing-databases-in-support-of-cmip6/, (last access: 18 March 2022; Hegglin et al., 2021)	
Solar	3.2	Matthes et al. (2017)	
Aerosol optical properties and relative change in cloud droplet number concentration	MACv2-SP	Stevens et al. (2017)	
Land use	v2.1 (h)istoric and (f)uture	Hurtt et al. (2019a, b, 2020)	Used only in combination with dynamic vegetation model
Nitrogen deposition	v2.0	http://blogs.reading.ac.uk/ccmi/ forcing-databases-in-support-of-cmip6/ (last access: 18 March 2022; Hegglin et al., 2021)	Used only in combination with dynamic vegetation model

were made to depend on LPJ-GUESS vegetation dynamics. In the version for CMIP6, as used in EC-Earth3-Veg, the surface albedo (including the shading effect of high vegetation), surface roughness length, and soil water exploitable by roots for evapotranspiration also vary following the variability of the effective vegetation cover. The improved representation of the effective vegetation cover variability brought a significant enhancement of the EC-Earth performance over regions where the land–atmosphere coupling is strong, in particular over boreal winter middle to high latitudes (Alessandri et al., 2017).

To represent time-dependent albedo for each grid point, a new scheme has been adopted that computes the total surface albedo (A_{tot}) as a weighted combination of contributions

from the albedo of the low and high vegetation types present in each grid point (a_{v}(type), which is a function of the low or high vegetation type), plus a time-constant background soil albedo (a_{s}, a function of space):

$$A_{\text{tot}} = a_{\text{v}}(T_{\text{l}}) C_{\text{low}}^{\text{eff}} + a_{\text{v}}(T_{\text{h}}) C_{\text{high}}^{\text{eff}} + a_s \left[1 - C_{\text{low}}^{\text{eff}} - C_{\text{high}}^{\text{eff}}\right], \tag{1}$$

where $C_{\text{low}}^{\text{eff}}$ and $C_{\text{high}}^{\text{eff}}$ are the effective fractional coverages for low and high vegetation, and T_{l} and T_{h} are the low and high vegetation types, respectively, at each grid point. The background soil albedo was adopted from the map from Rechid et al. (2009), and a look-up table of the albedo values a_{v} for each vegetation type was estimated using least square

minimization of errors against available monthly climatology of snow-free monthly MODIS albedo (Morcrette et al., 2008).

3.3 Dynamic vegetation and terrestrial biogeochemistry

LPJ-GUESS (Smith et al., 2001, 2014; Lindeskog et al., 2013; Olin et al., 2015a, b), a process-based second-generation dynamic vegetation and biogeochemistry model, is the terrestrial biosphere component of EC-Earth globally simulating vegetation dynamics, land use and land management following the LUH2 dataset (Hurtt et al., 2020), and both carbon (C) and nitrogen (N) cycling in terrestrial ecosystems. LPJ-GUESS has been evaluated in numerous studies (Smith et al., 2014; Wårlind et al., 2014) and reproduces vegetation patterns, dynamics, and productivity, C and N fluxes and pools, and hydrological cycling from global to regional scales, in line with independent datasets and comparable models (e.g., Piao et al., 2013; Zaehle et al., 2014; Sitch et al., 2015; Peters et al., 2018).

LPJ-GUESS is a new component in EC-Earth3 (Boysen et al., 2021), though it has previously been coupled to EC-Earth v2.3 (Weiss et al., 2012; Alessandri et al., 2017) using a simplified coupling scheme in which updates to leaf area index (LAI) alone were transferred between the sub-models.

LPJ-GUESS is one of the first vegetation sub-models interactively coupled to an atmospheric model, in which the size, age structure, temporal dynamics, and spatial heterogeneity of the vegetated landscape are represented and simulated dynamically. Such functionality has been argued to be essential for correctly capturing biogeochemical and biophysical land–atmosphere interactions on longer timescales (Purves and Pacala, 2008; Fisher et al., 2018) and has been shown to improve realism compared with more common area-based vegetation schemes (Wolf et al., 2011; Pugh et al., 2018). Different plant functional types (PFTs) co-occur in natural and managed stands governed by climate, atmospheric CO_2 (Meinshausen et al., 2017; Riahi et al., 2017), and N deposition (Hegglin et al., 2021) forcings. Evolving stand structure impacts growth, survivorship, and the outcome of competition by affecting the availability of the key resources: light, space, water, and nitrogen. Disturbances due to management actions such as forest clearing, prognostic wildfires, and a stochastic generic disturbance regime affect patches at random, inducing biomass loss and resetting vegetation succession (Hickler et al., 2004). N-cycle-induced limitations on natural vegetation and crop growth, C–N dynamics in soil biogeochemistry, and N trace gas emissions are included (e.g., Smith et al., 2014; Olin et al., 2015a, b) as are biogenic volatile organic compound (VOC) emissions (Hantson et al., 2017).

Meteorological inputs imposed on LPJ-GUESS are daily fields of surface air temperature and 25 cm soil temperatures, precipitation, and net shortwave and net longwave radiation from IFS/HTESSEL (Table 5). LPJ-GUESS calculates its own soil moisture for potential plant uptake in all patches in each of the six simulated stands independently of the single grid-cell-averaged hydrology scheme used in HTESSEL.

Vegetation dynamics are simulated on six stand types in the land portion of the grid cell (excluding large water bodies based on the static LUH2 ice and water fraction information), five stands having dynamic grid cell fractions consistent with the LUH2 dataset, namely natural, pasture, urban, crop, and irrigated crop, and one, peatland, having a fixed grid cell fraction derived from the GLCC global map used in the standard HTESSEL configuration – see Sect. 3.2. The LUH2 dataset, including land cover fractions, management options (N fertilization in this case), and land cover transitions, are read in yearly after aggregation to the atmospheric and land surface model resolution in a preprocessing step. A total of 10 woody and two herbaceous PFTs compete in the natural stand (Smith et al., 2014), whereas two herbaceous species, one each conforming to the C_3 and C_4 photosynthetic pathways, are simulated on pasture, urban, and peatland fractions. The crop stands each have five crop functional types (CFTs) representing the properties of global crop types and encompassing the classes found in the LUH2 database, namely both annual and perennial C_3 and C_4 crops, as well as C_3 N fixers (Lindeskog et al., 2013).

At the end of each day, LPJ-GUESS calculates the effective cover for low (high) vegetation, C_l (C_h), and LAI for low (high) vegetation, LAI_{low} (LAI_{high}), taking into account phenology and stand fractions in the grid cell. Dominant high and low vegetation types corresponding to the standard HTESSEL types are calculated and sent by LPJ-GUESS to IFS/HTESSEL on 31 December each year. These six fields link the vegetation dynamics and land use in LPJ-GUESS to the biophysical processes simulated at the land surface in HTESSEL, namely albedo, latent and sensible heat exchange, runoff, and momentum exchange.

In the EC-Earth-CC configuration, LPJ-GUESS is coupled to TM5 in addition to IFS and exchanges additional fields to enable prognostic global C cycle calculations. Spatiotemporally variable surface CO_2 concentrations are sent by TM5 to LPJ-GUESS (and PISCES) to replace the annual and global mean CO_2 concentrations used in the EC-Earth-Veg configuration. LPJ-GUESS sends daily averaged fields of net ecosystem C exchange (i.e., uptake or release) to TM5 to complement the surface C exchange with the ocean calculated in PISCES (see below), thereby completing the carbon cycle in EC-Earth-CC. This daily flux includes contributions from net primary production (NPP), heterotrophic respiration (Rh), wildfires, land use (including crop and pasture harvest), and natural disturbances on non-managed land. Since some processes in LPJ-GUESS are simulated with a yearly time step (e.g., wildfires, disturbance, establishment of new individuals and mortality, land use change), these annual fluxes are distributed evenly throughout the year and added to the daily NPP and Rh fluxes the following year to conserve car-

bon mass. Negative NPP fluxes account for CO_2 uptake by vegetation.

3.4 Atmospheric chemistry

The Tracer Model version 5 (TM5) is the atmospheric composition model of EC-Earth (van Noije et al., 2014) used in the EC-Earth3-AerChem and EC-Earth3-CC configuration. It can be used for the interactive simulation of carbon dioxide (CO_2), methane (CH_4), ozone (O_3), tropospheric aerosols, and other trace gases. These components are prescribed in IFS from forcing datasets (see Sect. 3.1) if not provided interactively by TM5. Other well-mixed greenhouse gases and stratospheric aerosols are prescribed in all configurations. This section briefly describes how the various components are configured.

As an alternative to the scaling approach for WMGHGs presented in Sect. 3.1, the 3D distributions of CO_2 and CH_4 can be calculated online by TM5. In the EC-Earth-CC configuration a single-tracer version of TM5 is used for simulating the transport of CO_2 through the atmosphere. Anthropogenic emissions of CO_2 are prescribed following the CMIP6 historical inventory (Hoesly et al., 2018) or future scenarios (Gidden et al., 2019). Exchange of CO_2 with the ocean and terrestrial biosphere is included by coupling TM5 to PISCES and LPJ-GUESS, respectively (see Sect. 3.3). An important feature of the model is that the transport in TM5 is mass-conserving (Krol et al., 2005). For the simulation of CH_4, a version of TM5 that includes atmospheric chemistry and aerosols is used (van Noije et al., 2021). A recent description of the chemistry scheme applied in EC-Earth has been presented by Williams et al. (2017). Emissions of aerosols and chemically reactive gases are taken from the CMIP6 historical datasets for anthropogenic sources (Hoesly et al., 2018) and biomass burning (van Marle et al., 2017) or the corresponding CMIP6 scenario datasets (Gidden et al., 2019). To force the CH_4 simulation to follow the pathway provided by CMIP6, its surface mixing ratios are nudged towards the monthly zonal means from CMIP6 interpolated to daily values. Moreover, because TM5 lacks a comprehensive stratospheric chemistry scheme, the CH_4 mixing ratios in the stratosphere are nudged towards a monthly zonal mean observational climatology representative for the 1990s (interpolated to daily values), scaled by a global factor based on the CMIP6 time series of global annual mean surface values. To calculate the scale factor, we assume a 1-year delay between the mixing ratios at the surface and in the stratosphere (Meinshausen et al., 2017) and a reference value based on a 10-year average.

The chemical production of water vapor (H_2O) by oxidation of methane in the stratosphere is included in IFS in a similar way as in the standard version of IFS. The assumption made in the standard version of IFS is that

$$2 \times [CH_4] + [H_2O] = Co, \qquad (2)$$

where square brackets denote local mixing ratios (in ppmv) and the constant is set to 6.8 ppmv based on observations for the present day. To account for long-term variations in CH_4, in EC-Earth it is assumed instead that

$$2[CH_4] + [H_2O] = C(t), \qquad (3)$$

where

$$C(t) = Co + 2([CH_4]S(t) - [CH_4]0S). \qquad (4)$$

Here $[CH_4]S(t)$ is the monthly varying global mean surface mixing ratio obtained by linear interpolation from the CMIP6 time series of annual values, and $[CH_4]0S$ is a reference value for the present day, which is set to 1.78 ppmv.

Ozone is simulated by TM5. As for CH_4, TM5 applies a nudging scheme for O_3 in the stratosphere. In EC-Earth3, the mixing ratios are nudged towards daily zonal means obtained from the CMIP6 dataset.

For aqueous-phase chemistry in the troposphere, the acidity of cloud droplets is calculated assuming a uniform CO_2 mixing ratio following the CMIP6 time series of annual global mean surface values.

TM5 simulates tropospheric aerosols, namely sulfate, black carbon, primary and secondary organic aerosol, sea salt, and mineral dust, in four size ranges describing nucleation, Aitken, accumulation, and coarse modes using the M7 aerosol microphysical model (Vignati et al., 2004). In addition, it simulates the total mass of ammonium, nitrate, and methane sulfonic acid (MSA). Optical properties of the aerosol mixture are calculated based on Mie theory in combination with the mixing assumptions described by van Noije et al. (2014).

For calculation of the SW radiative effects of the aerosol mixture, TM5 provides the extinction, single-scattering albedo, and asymmetry factor at the 14 wavelength bands of the SW radiation scheme (using the same wavelength values as in MACv2-SP). In addition, TM5 provides the particle number and component mass mixing ratios for each of the M7 modes, plus the total mass mixing ratios of nitrate and MSA. LW absorption is calculated based on the mass mixing ratios of the M7 components using absorption efficiencies from IFS. The contribution of the aerosol mixture described by MACv2-SP and/or TM5 to SW extinction and LW absorption is removed above the tropopause, where the stratospheric aerosol dataset from CMIP6 is applied. The tropopause level is diagnosed online following the thermal tropopause definition of the World Meteorological Organization (WMO, 1957), as detailed by Reichler et al. (2003). Where the thermal tropopause does not exist according to this definition, tropospheric and stratospheric aerosols are merged at the 100 hPa level.

3.5 Ocean

The ocean component of the EC-Earth model is the Nucleus for European Modelling of the Ocean (NEMO; Madec

and the NEMO team, 2008; Madec et al., 2015) that includes the ocean model OPA (Océan Parallélisé), the LIM3 sea ice model (see Sect. 3.6), and the PISCES biogeochemistry model (see Sect. 3.7). The CMIP6 version of the EC-Earth model uses NEMO3.6 (revision r9466) in combination with the ORCA1 shared configuration (regular ORCA1, not eORCA1).

OPA is a primitive equation model of ocean circulation. Prognostic variables are velocity, hydrostatic pressure, sea surface height, and thermohaline variables (potential temperature and salinity). The distribution of variables is given by a three-dimensional Arakawa-C-type grid (Arakawa and Lamb, 1977). OPA uses a partial step implementation for the geopotential z^* coordinate (grid boxes do not continue below topography) and a diffusive bottom boundary layer scheme (similar to that of Beckmann and Döscher, 1997) with implicit bottom friction to mix dense water down a slope.

NEMO allows for various choices for the physical subgrid-scale parameterizations as well as the numerical algorithms. EC-Earth uses the turbulent kinetic energy (TKE) scheme for vertical mixing. The vertical eddy viscosity and diffusivity coefficients are computed from a 1.5 turbulent closure model based on a prognostic equation for the turbulent kinetic energy and a closure assumption for the turbulent length scales. This turbulence closure model has been developed by Bougeault and Lacarrère (1989) in atmospheric cases, adapted by Gaspar et al. (1990) for oceanic cases, and embedded in OPA by Blanke and Delecluse (1993).

Since the CMIP5 version of EC-Earth, major changes in the TKE schemes have been implemented: it now includes a Langmuir cell parameterization (Axell, 2002), the Mellor and Blumberg (2004) surface wave breaking parameterization, and a time discretization which is energetically consistent with the ocean model equations (Burchard, 2002; Marsaleix et al., 2008). A mixed layer eddy parameterization following Fox-Kemper et al. (2008) has been newly implemented in NEMO3.6. An enhanced vertical diffusion and a double diffusive mixing parameterization are part of the OPA code in EC-Earth. Since CMIP5, a tidal mixing parameterization has been added to OPA (de Lavergne et al., 2020).

Horizontal tracer diffusion is described by the Gent–McWilliams (Gent and McWilliams, 1990) parametrization of mesoscale eddy-induced turbulence.

The ORCA family is a series of global ocean grid configurations. The ORCA grid is a tripolar grid based on the semi-analytical method of Madec and Imbard (1996). ORCA1 with a resolution of about 1° is used for standard- or low-resolution simulations, and ORCA025 (resolution 0.25°) is used for high-resolution simulations with EC-Earth3-HR. A meridional grid refinement of 1/3° in the tropics allows a partial representation of tropical instability waves. There are 75 vertical levels in the ocean with an upper level of about 1 m and 24 levels distributed over the uppermost 100 m.

The main difference of the OPA version used in EC-Earth compared to the reference OPA version of NEMO3.6 is that the parameterization of the penetration of TKE below the mixed layer due to internal and inertial waves is switched off (nn_etau = 0). This has been done because the penetration of TKE below the mixed layer caused a surface layer of warm summer water masses in the North Atlantic convection areas that is too deep, which leads to a breakdown of the Labrador Sea convection within a few years and a strongly underestimated Atlantic Meridional Overturning Circulation (AMOC) in EC-Earth. A minor modification compared to the standard NEMO setup from the ORCA1-shared configuration for NEMO (ShacoNemo) is an increased tuning parameter rn_lc (0.2) in the TKE turbulent closure scheme that directly relates to the vertical velocity profile of the Langmuir cell circulation. Consequently, the Langmuir cell circulation is strengthened.

3.6 Sea ice

The sea ice component is version 3.6 of the Louvain-la-Neuve Ice Model (LIM; Vancoppenolle et al., 2009; Rousset et al., 2015), which works directly on the NEMO environment, including the ORCA grid. LIM3.6 is based on the Arctic Ice Dynamics Joint EXperiment (AIDJEX) framework (Coon et al., 1974), combining the ice thickness distribution (ITD) framework, the conservation of horizontal momentum, an elastic–viscous–plastic rheology, and energy-conserving halo-thermodynamics (Vancoppenolle et al., 2009). All of these components of the sea ice model have been introduced or revised since CMIP5.

The ice thickness distribution framework was introduced (Thorndike et al., 1975) to deal with meter-scale variations in ice thickness, which cannot be resolved explicitly but should preferably be accounted for, as many sea ice processes, in particular growth and melt, depend nonlinearly on thickness h. In practice, this is achieved by treating h as an independent variable, leading to the introduction in discrete form of $L = 5$ thickness categories, each characterized by a specific set of state variables (namely ice concentration, ice volume per unit area, snow volume per unit area, ice enthalpy, snow enthalpy, sea ice salt content). Ice and snow enthalpy also depend on vertical depth in the ice (z). All sea ice state variables $X_{ijl}, l = 1, \ldots, L$ are updated due to transport and thermodynamic processes. The default choice of five categories, with the upper category above 4 m, has been shown to provide reasonable results at an acceptable computing cost (Massonnet et al., 2019).

Vertical sea ice motions are irrelevantly small and hence neglected, and the sea ice velocity field reduces to its horizontal components. The 2D ice velocity vector is considered the same for all categories and stems from the horizontal momentum conservation equation. The internal stress term is formulated assuming that sea ice is a viscous–plastic material, i.e., assuming viscous ice flow at very small deformation and plastic flow (stress independent of deformation) above a plastic failure threshold. This threshold lies on

an elliptical yield curve in the principal stress component space, whose size can be changed by tuning the classical ice strength parameter $P^* = 20\,000$ following the classical formulation of Hibler (1979). The horizontal momentum equation is resolved using the elastic–viscous–plastic (EVP) C-grid formulation of Bouillon et al. (2009) using 120 sub-time steps. Once the velocity field is computed, the sea ice state variables are transported horizontally using the second-order moment-conserving scheme of Prather (1986).

Ice thermodynamics are based on the Bitz and Lipscomb (1999) enthalpy formulation and account for dynamic changes in ice salinity through temperature- and salinity-dependent thermal properties (Ono, 1968; Pringle et al., 2007). The salt entrapment and drainage parameterizations follow from Vancoppenolle et al. (2009): each category is characterized by a dynamic mean salinity, from which a profile shape is derived for the computation of the vertical diffusion of heat. The broadband surface albedo of each ice category empirically depends on ice thickness, snow depth, surface temperature, and cloud fraction based on a reformulation of the Shine and Henderson-Sellers (1985) parameterization that solves a few inconsistencies associated with state transitions (e.g., snow/no snow), following Grenfell and Perovich (2004), and is tuned to match observations of Brandt et al. (2005). The impact of melt ponds is implicitly accounted for through imposed changes on the albedo activated when the surface temperature is 0 °C. Energy, salt, and mass conservations have been carefully checked within the ice component and its interfaces with the atmosphere and ocean (Rousset et al., 2015).

All surface fluxes are computed in the atmosphere, and the IFS atmospheric model has only one ice thickness category. The solar and non-solar heat fluxes are therefore distributed on the different sea ice categories in LIM3, taking into account the differences in albedo and temperature among the sea ice categories in each grid point.

During the tuning phase it was found that the Arctic sea ice volume grew to unrealistically high values, especially during phases with reduced AMOC. An analysis showed that the thermal conductivity of snow needed a slight reduction (rn_cdsn = 0.27) to reduce basal growth and increase bottom melt (see also Sect. 2).

3.7   Ocean biogeochemistry

PISCES-v2 (Pelagic Interactions Scheme for Carbon and Ecosystem Studies volume 2) is a biogeochemical model that simulates the nutrient cycle and the inorganic and organic carbon cycle, and it comprises lower trophic phytoplankton and zooplankton (Aumont et al., 2015). It has two functional groups for phytoplankton (nanophytoplankton, including calcite producers, and diatoms that can produce siliceous shells) and two size classes for zooplankton (mesozooplankton and microzooplankton). The growth rate of phytoplankton depends on photosynthetically available radiation (PAR)

intensity and temperature. A limitation for primary production is computed based on the availability of the main nutrients (P, N, Si, Fe). In the case of low nitrate concentrations, nitrogen fixation by diazotrophic cyanobacteria is parameterized in waters warmer than 20 °C (Aumont et al., 2015). PISCES uses a constant P / N / C ratio of 1/16/122 for primary production. Organic particulate matter produced by food-web processes in the euphotic layer is represented by two size classes. These sink throughout the water column with different velocities while being decomposed into dissolved inorganic nutrients (DIN, DOP) and dissolved inorganic carbon (DIC). A further pool for dissolved organic matter (DOM) is fed by phytoplanktonic exudation and excretion by zooplankton. DOM in PISCES represents only the semi-labile fraction, with turnover times ranging from months to years, and it is further remineralized at a constant rate. PISCES includes two different chemistry models to describe iron pool interactions. In EC-Earth3 we use the complex model by Tagliabue and Arrigo (2006). The global river and atmospheric deposition input of nutrients are not balanced to match the fraction lost by sediment burial. For this reason, PISCES allows for a homogeneous correction towards global mean values for alkalinity, nitrate, phosphate, and silicate. Furthermore, physical transport of passive tracers presented a slight artificial mass imbalance. To prevent it from becoming significant for carbon during the spin-up we applied a uniform correction to dissolved inorganic carbon at the end of each year after taking into account all sources and sinks.

With respect to climate studies PISCES is capable of simulating the relevant processes of the marine carbon cycle; i.e., it comprises the soft-tissue carbon pump and the carbonate counter-pump to realistically simulate the feedback of the marine carbon cycle to the climate.

The air–sea gas exchange for carbon dioxide and oxygen is parameterized according to Wanninkhof (1992). The interface to the seafloor is given by basic assumptions for the exchange between the active sediment layer and the water bottom layer where different assumptions are made for the burial efficiency for silicate, calcite, and particular organic matter (see Aumont et al., 2015, for further details).

PISCES is part of the community model NEMO and runs on the same model grid. In EC-Earth3 the horizontal resolution is about 1° (ORCA1) with 75 vertical levels. Advection and diffusion of the 24 biogeochemical tracers are done in the hydrodynamic ocean model. A detailed description of the PISCES reference version is given in Aumont et al. (2015). In EC-Earth3 PISCES can be run in passive mode or with feedback to the atmosphere by prognostically simulating air–sea carbon fluxes and contributing to determining atmospheric $p\mathrm{CO}_2$ when the global carbon cycle is fully closed in the case that the atmospheric chemistry model TM5 and the terrestrial biosphere model LPJ-GUESS are also enabled. A feedback to the ocean physics is not foreseen; i.e., the thermal effect of

light absorption by chlorophyll on water temperature is not communicated to NEMO (although possible).

3.8 Greenland Ice Sheet

The Parallel Ice Sheet Model v1.1 (PISM; Bueler and Brown, 2009; Winkelmann et al., 2011; The PISM Team, 2019) is used in the EC-Earth3-GrIS configuration to model the Greenland Ice Sheet (GrIS) evolution in the climate system. PISM is an open-source model jointly developed by a group of universities available from http://www.pism-docs.org (last access: 18 March 2022). While all surface processes over an ice sheet (such as the snow layer) are modeled in EC-Earth3, PISM handles the ice sheet dynamical and thermodynamical processes, including ice flow, subglacial hydrology, bed deformation, and the basal ice melt.

The spatial domain of PISM is built on a three-dimensional, equidistant polar stereographic grid. The equations are solved with an adaptive time stepping procedure. Boundary conditions include subsurface temperature and mass balance on the ice surface (provided by EC-Earth3), bedrock elevation, and bed geothermal heat flux (considered to be invariant geographic conditions).

PISM considers the ice sheets as a slow, nonlinearly viscous isotropic fluid characterized by a creeping flow induced by gravitational forces and constrained by the conservation laws of momentum, mass, and energy for ice. A combination of two shallow ice approximations, the non-sliding shallow ice (SIA) and the shallow shelf (SSA) approximations, is applied, depending on the ice regime (Bueler and Brown, 2009). The former is applied to bed-frozen parts of the ice sheets, while the latter is applied to ice shelves and also used as a sliding law in areas with low basal resistance. This hybrid formulation enables the modeling of fast-flowing ice streams and outlet glaciers and is commonly used for simulations of whole ice sheets for which it is too expensive to solve the full set of stress balance equations.

The ice velocities are determined from geometry (i.e., ice thickness and ice surface elevation), ice temperature, and basal strength using momentum–stress balance equations. The ice thermodynamics are formulated as the energy balance based on enthalpy that enables solutions for polythermal ice masses (Ashwanden et al., 2012) and the Glen–Paterson–Budd–Lliboutry–Duval flow law (Lliboutry et al., 1985) that accounts for softening of the ice as the liquid water fraction increases. The ice flow law is a single power law in which the exponent can be selected independently for the SIA and SSA. Furthermore, an enhancement factor is used to account for the anisotropic nature of the ice.

The subglacial processes are resolved by the sliding law that relates the basal sliding velocity to the basal shear stress. PISM uses a pseudo-plastic sliding law and the Mohr–Coulomb model for yield stress that depends on the till friction angle and the effective pressure of the saturated till. The latter is based on a subglacial routing scheme, and the basal

melt rate is calculated from energy conservation across the ice–bedrock layer. A geothermal heat flux map is applied at the basal boundary to account for the heat entering the ice sheet from below. Ice bed deformation is approximated by the viscoelastic deformable Earth model formulated in Bueler et al. (2007).

Calving of marine-terminating glaciers at the ocean boundary is parameterized in PISM as the model does not have a good representation of the narrow fjord systems of the marine outlet at the considered resolution. Several calving schemes are implemented in PISM to cope with different conditions, including eigencalving, von Mises calving, thickness calving, and flow-kill calving. A commonly used calving scheme for GrIS is adapted from the von Mises yield criterion, which is suited for ice flows confined in narrow valleys and fjords (Morlighem et al., 2016). The parameterization assesses the calving speed from the amount of ice fracturing.

3.9 Lake treatment

EC-Earth3 has no explicit lake model. All grid points in the atmosphere model that are covered with less than 50 % water in the land–sea mask are assumed to be land and use the HT-ESSEL scheme as described above. Grid points with more than 50 % water are considered water, and the water temperature, sea ice cover, and sea ice temperature in these points are updated by the coupler. This process is straightforward for grid points over the ocean, but for inland water bodies such as large lakes this implies an extrapolation of temperature, for example, from the nearest unmasked grid point in the ocean model. This method potentially leads to problems in lakes where the closest ocean point is much further north such as the Great Lakes for which the closest ocean point is the Hudson Bay. This constellation implies colder lakes and longer sea ice cover, which has an impact on the local climate (e.g., lake breeze) around these lakes.

4 Procedures to test possible model climate dependencies on computing platform, compilers, and domain decomposition

EC-Earth3 runs under different high-performance computing (HPC) environments in the EC-Earth partner supercomputing centers. This has several advantages, from allowing different groups to work with the same tool in parallel to leveraging the burden of ensemble climate simulations. However, for obvious scientific reasons, it is critical to ensure that ESMs provide replicable results under changes in the computing environment. While bit-for-bit replicability is in general infeasible because of the existence of hardware and software constraints that are beyond the control of climate researchers, it must be expected that results obtained under

one computing environment are statistically indistinguishable from those obtained under another environment.

EC-Earth as a community has developed a protocol (Massonnet et al., 2018, 2020) to assess replicability for the EC-Earth3 model system. This protocol is based on a statistical comparison of standard climate metrics derived from control integrations executed in different HPC environments. This protocol has been tested with EC-Earth 3.2, allowing judgment of replicability of EC-Earth 3.2 results. It was also shown that the interim version of the model, EC-Earth 3.1 (developed between CMIP5 and CMIP6), was not fully replicable. The experience gained (Massonnet et al., 2020) suggested that codes (especially when they are bugged) interfere with computing environments in sometimes unpredictable ways. The default assumption in EC-Earth is that a given model simulation is *not* replicable when the HPC environment is changed until proven otherwise, i.e., that the model executed in the two computing environments gives results that cannot be deemed incompatible. The protocol developed within EC-Earth fulfills this goal, and it is now required to check the replicability of EC-Earth each time it is ported to a new machine before new production runs can be started in this machine.

Replicability is a cornerstone of science and, by extension, climate research. Testing whether climate models provide replicable output under changes in HPC environments is a prerequisite to ensure trusted distribution of large ensembles across multiple platforms, which were foreseen within the EC-Earth community. The workload of CMIP6 experiments was shared and run in collaboration by different institutions of the consortium across different platforms.

Here we give an example for an application of the replicability test. It consists of the executions of the same code in coupled (CMIP) and atmosphere-only configurations (AMIP), with five ensemble members, carried out by different institutions using their respective platforms. In essence, this test (see Box 1) assesses the differences due to the varying HPC environment in 20-year simulations after accounting for internal variability.

The performance indices by Reichler and Kim (2008) (RandK), which compare model performance against the reanalysis product ERA40, were calculated using the EC-Earth analysis tool EC-mean, which is available in the model source tree. The tests included six different machines (Tetralith, Rhino, Marenostrum4, CCA, Sisu – Finnish IT Center for Science, Hpcdev – CRAY XT-5, and Kay), two different compilers (Intel and Cray), different domain decomposition (number of parallel resources used), and different versions of libraries and compilers used (subject to the availability of each platform).

The results of this test showed that the climates simulated by EC-Earth were statistically indistinguishable in most of the cases. The differences were evaluated for 13 different variables (Table 14) with similar results for all variables and vertical levels. An example for 2 m temperature (t2m) is

Table 14. Difference in the 20-year mean and ensemble mean near-surface values for different variables between the Rhino and CCA platform experiments.

Variable	Description	Spatial mean differences (%)
t2m	2 m air temperature	1.2
msl	Mean sea level pressure	0.7
qnet	Net thermal radiation	0.8
tp	Total precipitation	1.4
ewss	Zonal wind stresses	1.1
nsss	Meridional wind stress	1.3
SST	Sea surface temperature	0.7
SSS	Sea surface salinity	0.8
SICE	Sea ice concentration	1.1
T	3D air temperature	1.3
U	3D zonal wind	1.5
V	3D meridional wind	1.2
Q	Specific humidity	0.7

shown in Fig. 2 for the HPC systems Rhino and CCA. Both the AMIP and CMIP experiments return similar mean values when averaged spatially (Fig. 2a and b, respectively). Figure 2c also shows that the results at the end of the execution (20 years) are statistically similar in spite of the internal variability; only 1.2 % of the total results could be considered significantly different (under the null hypothesis of no difference, significant differences are expected to occur 5 % of the time). Similar results were obtained for the remaining comparison of HPC systems, except for two integrations which failed this test. These failures were traced back to the specific compiler versions used. Using a different setup (with a different version of the compilers) removed the differences, thus proving that the replicability test can be used to highlight incorrect configurations.

5 Climate conditions as simulated in historical ensembles

To illustrate the performance of the EC-Earth3 model we present results from the ensemble of historical experiments. The forcing of the historical experiments follows the CMIP6 protocol and has been described in detail in Sect. 3. After tuning the coupled configurations EC-Earth3 and EC-Earth3-Veg, spin-up and piControl (pre-industrial control) simulations have been carried out. PiControl simulations for other model configurations have been carried out but will not be used in this paper.

Each ensemble member of the historical experiment was branched off from the corresponding piControl experiment at a different time, with branching times separated by intervals of 20 years (Fig. 3). The complete information about the branch time of each member is included in the metadata on all variables. Separate piControl experiments have been

The protocol is designed to test whether EC-Earth3 gives replicable results under two computing environments, named "A" and "B" hereinafter.

1. **Deterministic check.** This initial step is performed to confirm that EC-Earth is fully deterministic. Two 1-year integrations are conducted consecutively under the same computing environment (same executable, same machine, same domain decomposition). The results are required to be bit-for-bit identical.

2. **Accounting for internal variability.** Ensemble simulations are conducted so that the imprint of internal variability on climate indices like time means can be gauged. In an attempt to find the right balance between statistical power and limited computational resources, we run a five-member, 20-year integration for both "A" and "B" computing environments. In the subsequent steps, these simulations are referred to as "ensembles".

3. **Generation of simulations**. The five members of ensembles A and B always start from unique atmospheric and sea ice restarts, obtained from a long equilibrium simulation conducted on one computer. An oceanic restart is also obtained from this equilibrium simulation, and five random but deterministic perturbations are added to the sea surface temperature of this restart (Gaussian perturbation, standard deviation: 10^{-4} K). The introduction of these tiny perturbations allows ensemble spread to develop in ensembles A and B. Note that by the deterministic nature of the perturbations, pairs of members always start from the same triplet of atmospheric, oceanic and sea ice restarts: the first member of ensemble A and the first member of ensemble B start from identical initial conditions, and so on. Ensembles A and B are conducted under an annually repeating pre-industrial constant forcing. The ensembles start in 1850 and extend to 1869. The method of perturbation does not allow sampling the entire distribution of climatic states, as simulations are only 20 years long. If differences in computing environments cause differences in model output that are further enhanced by differences in climatic states, then there is a chance that our approach misses the detection of cases of replicability. However, differences in surface properties like SSS and SST develop rather rapidly in the ensemble, with inter-model spread comparable to interannual variability well before 20 years.

4. Calculation of standard indices. Due to the large amount of output produced by each simulation, the outputs from ensembles A and B are first post-processed in an identical way. Based on the list of standard metrics proposed by Reichler and Kim (2008), we record for each ensemble standard ocean, atmosphere, and sea ice parameters: 3D air temperature, humidity, and components of the wind; 2D total precipitation, mean sea level pressure, air surface temperature, wind stress, and surface thermal radiation; 2D sea surface temperature and salinity; and sea ice concentration. These fields are averaged monthly (240 time steps over 20 years).

5. **Calculation of standard metrics.** Then, the model fields are compared to the same reference datasets as those used in Reichler and Kim (2008), which consist of observational and reanalysis datasets. For each field, a grid cell area-weighted average of the model departure from the corresponding reference is evaluated and then normalized by the variance of that field in the reference dataset. Thus, for each field, one number is retained that describes the mismatch between that field in the ensemble and the reference. Five such numbers are available for each ensemble, since each ensemble uses five ensemble members.

6. **Statistical testing.** For each index we compare two five-member ensembles and determine whether the two ensembles are statistically indistinguishable from one another. Since no prior assumption can be made on the underlying statistical distribution of the samples, we use a two-sample Kolmogorov-Smirnov (KS) test. KS tests are non-parametric, which makes them suitable for our application. A Monte Carlo analysis reveals that for a prescribed level of significance of 5%, the power of the two-sample KS test exceeds 80% (a standard in research) when the means from the two samples are separated by at least 2 standard deviations (Figure 2). Stated otherwise, there is a non-negligible probability (>20%) that small but actual differences (less than 2 standard deviations) are not detected by the test. In this test, the null hypothesis is that the two samples are statistically indistinguishable from each other with a confidence of 95%. This means that under this null hypothesis of no difference, significant differences are expected to occur 5 % of the time. Because the tests are done on global metrics, there is a possibility of false negatives when the same performance is obtained from spatially varying biases. However, the use of multiple variables for the assessment of maps aims at minimizing this issue.

Box 1. The protocol for testing replicability.
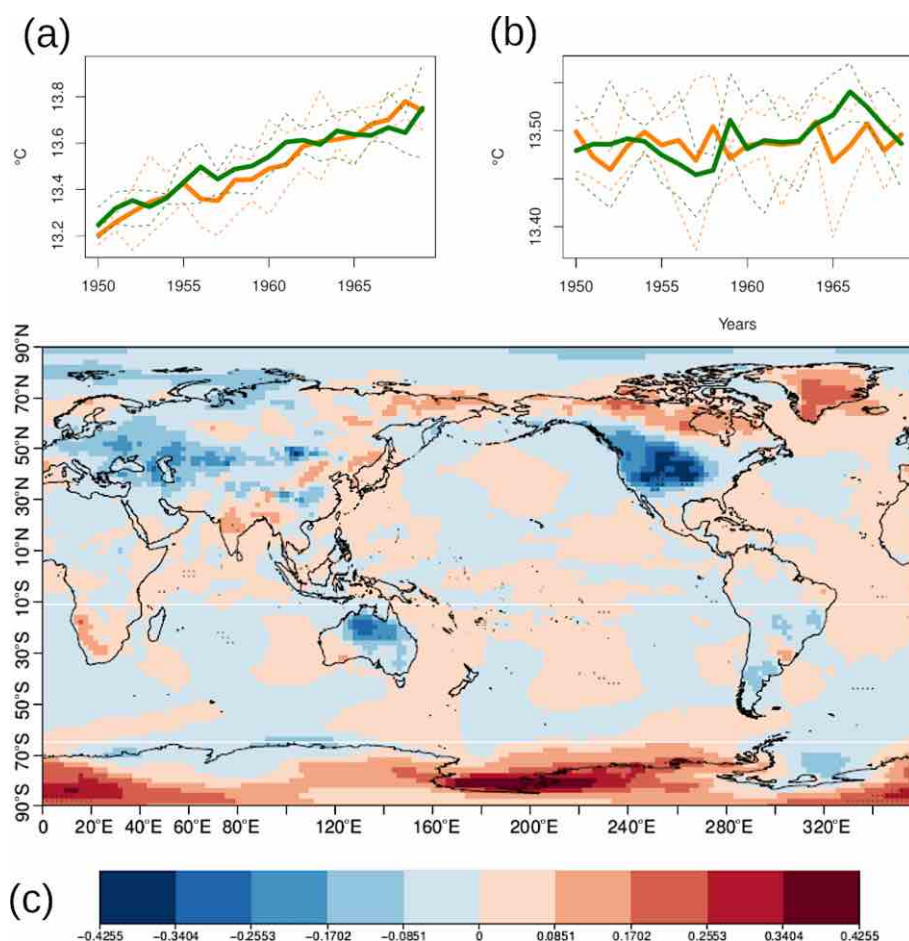
Figure 2. Comparison between Rhino and CCA platforms for AMIP experiments and historical CMIP experiments for 2 m temperature (t2m). Panels **(a)** and **(b)** show mean (solid line) as well as maximum and minimum values (dashed lines) between the ensemble members for AMIP **(a)** and CMIP **(b)** experiments, respectively. Green and orange refer to Rhino and CCA platforms, respectively. Panel **(c)** shows the t2m difference between the CMIP ensemble runs. Black dotted regions indicate statistical significance according to the Kolmogorov–Smirnov test (1.2 % of grid points show a significant difference in this case).

produced for EC-Earth3 and EC-Earth3-Veg, and more will follow for other model configurations. Most of the results presented here are based on the 20-member large ensemble of the EC-Earth3 configuration that were available from the ESGF at the time of writing. The results from four available EC-Earth3-Veg members were found to be very similar, and therefore we focus on the analysis of results from EC-Earth3.

5.1 Temperature

The time series of the global annual mean near-surface air temperature (TAS) is shown in Fig. 4. Reanalysis data from ERA5 and ERA-20C are shown for comparison. Compared to ERA5 the model ensemble has a warm bias of about 0.5 K in the global mean. ERA5 can be considered more relevant for the global mean TAS in comparison with ERA-20C because ERA5 has stronger observational constraints and updated physical parameterization in the underlying global

model, while ERA-20C is limited to assimilation of daily surface pressure and surface winds over the ocean (Poli et al., 2016).

The bias in the EC-Earth3 global mean TAS is mainly due to a warmer ocean and especially due to a strong warm bias in the Southern Ocean, as we will show below. Another feature is the large ensemble spread shown as a shaded area around the ensemble mean. The difference between the coldest and warmest ensemble member is 0.8 K on average. For further comparison, the CMIP5 ensemble of EC-Earth2.3 is also shown, which is about 1.2 K colder and indicates a smaller spread among the 10 members in comparison to the current 20-member ensemble. When analyzing the long piControl run we found that the EC-Earth3 model oscillates between two states that are characterized by low or high values of the AMOC, cold or warm North Atlantic temperatures, and more or less sea ice in the Arctic, with a period of about 200 years. The temperature difference over the North Atlantic is large
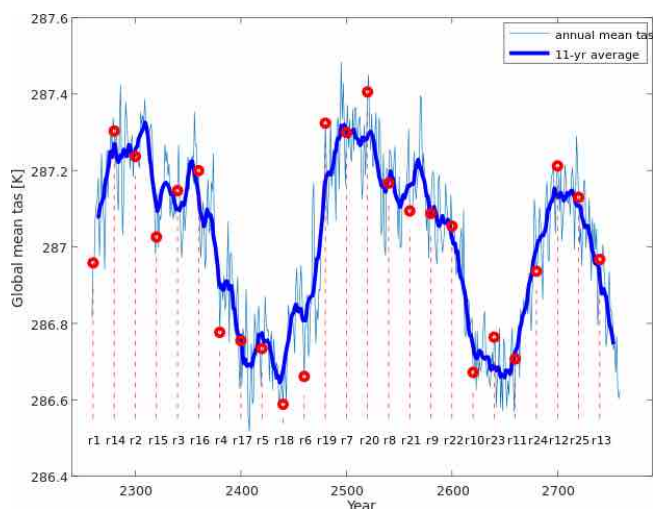
Figure 3. Time series of the global mean of annual mean near-surface temperature (TAS) in the 500-year-long EC-Earth3 piControl experiment. The thick blue line is an 11-year running average of the annual means. The time axis is arbitrary because of constant forcing. Red circles mark the initial states from which the members of the historical experiment are started. The realization_id of the members of the historical ensemble is shown at the bottom of the figure.

enough to have a discernible impact on the global mean TAS, resulting in a warm and a cold state. The model can remain in either of these states for several decades before turning to the other state, with transitions occurring at irregular intervals. We speculate here that the oscillation decreases with the warming climate. This is indicated by a larger ensemble of historical and scenario simulations than shown here. The exact processes as well as the triggers for the oscillation are still under investigation on the basis of the larger ensemble and will be presented in a later study. However, it is already clear that the large differences between the warmest and coldest member of the ensemble of historical simulations are related to the two states of the model climate, i.e., whether the initial states for the historical run stem from the cold or warm phase of the piControl run, and that transitions between different states occur during the historical simulation.

The TAS trend after 1980 is found to be stronger in the EC-Earth3 ensemble mean (0.25 K per decade) than in the ERA5 reanalysis (0.18 K per decade). However, when looking at the warming during the entire historical period by comparing the mean TAS from 1851–1880 against 1981–2010 we find that the model warmed by 0.7 K, which is only slightly higher than the 0.63 K estimate for the observed warming (IPCC; Hoegh-Guldberg et al., 2018)

To study the EC-Earth3 recent past climate we focus on the period 1980–2010. The ensemble mean spatial TAS is compared to ERA5 for this period in Fig. 5. We find cold TAS biases over the land regions and the Arctic and warm biases over the Southern Ocean and Antarctica as well as

for the stratocumulus regions of the continents (Fig. 5b). The individual member biases have a large spread in the Northern Hemisphere as previously mentioned, while the Southern Hemisphere biases are similar for all members as seen in the zonal mean bias plot (Fig. 5d). The CMIP5 ensemble EC-Earth2.3 is colder than ERA5 except over the Southern Ocean (Fig. 5c), and the spread is much smaller than for EC-Earth3 (Fig. 5d).

Most coupled climate models suffer from a warm Southern Ocean (SO) bias (Hyder et al., 2018). In EC-Earth3 configurations, the warm bias is found in all seasons. Large parts of the bias have been attributed to biases in shortwave cloud radiative effects. Modifications in the cloud scheme and the representation of supercooled liquid water made in more recent versions of IFS, including cycle 45r1 (Forbes and Ahlgrimm, 2014; Forbes et al., 2016), together with the introduction of the new ecRad radiation scheme in cycle 43r3 (Hogan et al., 2017), have been shown to substantially reduce these biases.

6 Precipitation

Global mean precipitation (pr) patterns are well represented in EC-Earth3 in comparison with ERA5 for present-day conditions (Fig. 6), except for the double Intertropical Convergence Zone (ITCZ) pattern noticeable as a distinct peak in the zonal mean pr at 9° S. This is a persistent model bias common to most global climate models. The largest pr biases relative to the observational dataset Global Precipitation Climatology Project (GPCP; Gehne et al., 2016) are found in the tropics, in particular in the Southern Hemisphere (SH) and to a lesser extent in the Northern Hemisphere (NH). EC-Earth3 precipitation is closer to ERA5 but has less precipitation at the Equator. The overestimation of pr by the model in the SH and the double ITCZ are likely consequences of the strong warm temperature bias over the Southern Ocean that leads to a more southward-displaced tropical rainfall belt and more vigorous hydrological cycle (Hwang and Frierson, 2013).

6.1 Sea level pressure

EC-Earth3 mean sea level pressure (PSL) fields are close to ERA5 as shown in Fig. 7. The PSL bias of individual members and the ensemble values are within 1.5 hPa compared to ERA5 except over Antarctica where the pressure is between 0.5 and 2 hPa too high, while over the Southern Ocean the pressure is 1 hPa too low. The variability, measured as the annual standard deviation, of EC-Earth3 PSL is very close to ERA5, indicating realistic surface winds (not shown) and modes of variability as shown later in this section.

6.2 Zonal wind

An overview of the atmospheric circulation is shown in Fig. 8, where the yearly averaged biases of the zonal com-
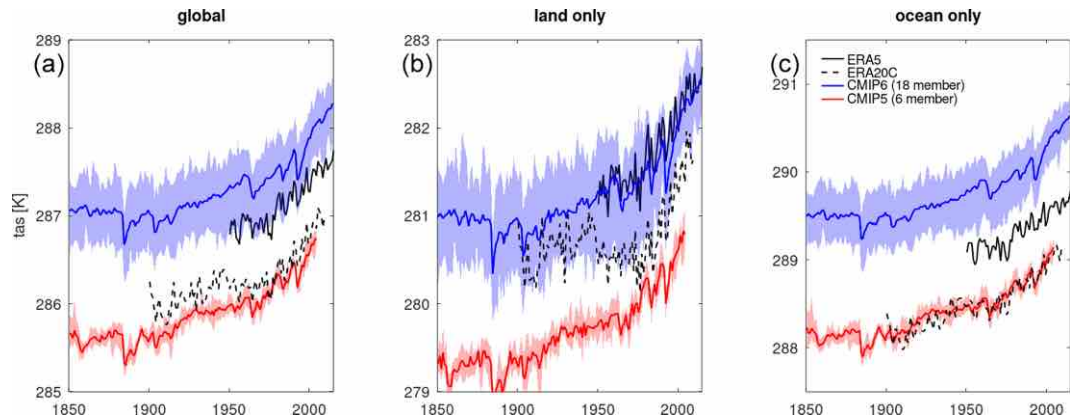
Figure 4. Global annual mean TAS (in K) from the EC-Earth2.3 (red, for CMIP5) and EC-Earth3 (blue, for CMIP6) ensembles for **(a)** global mean, **(b)** land only, and **(c)** ocean only. Ensemble means are shown as thick lines, and the ensemble spread is shown as a shaded area. Global annual mean TAS from the ERA5 (black, solid) and ERA-20C (black, dashed) reanalyses is shown for comparison.
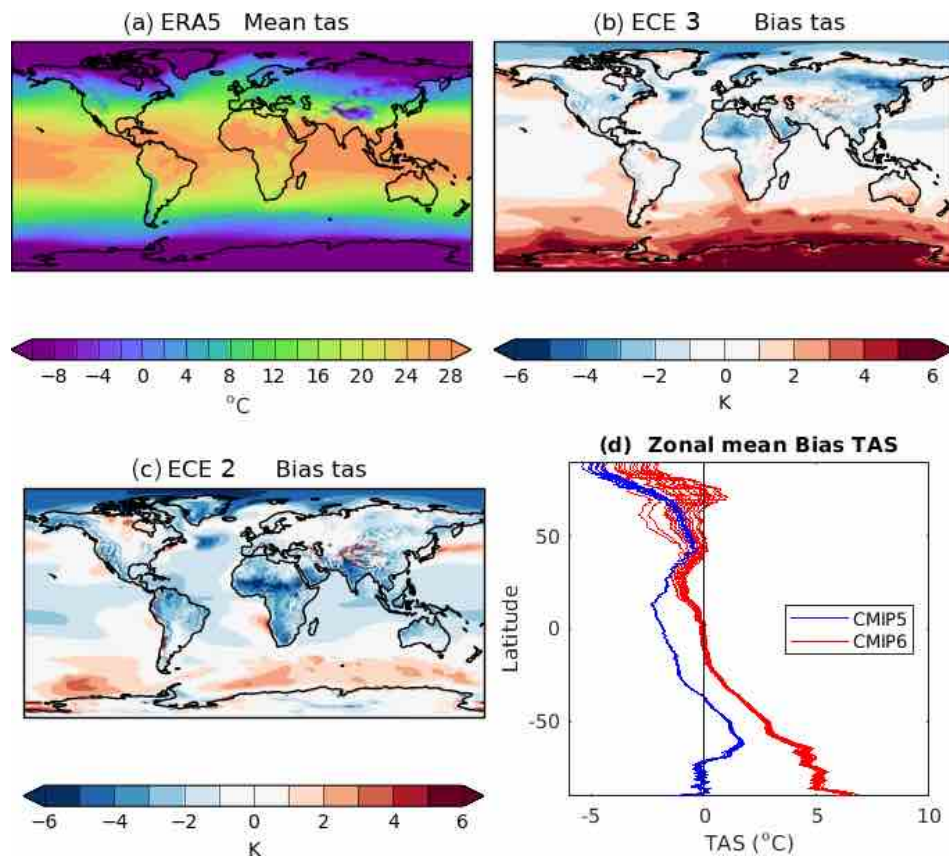
Figure 5. (a) ERA5 1980–2010 mean TAS (in °C); **(b)** EC-Earth3 and **(c)** EC-Earth2.3 ensemble mean biases compared to ERA5 (°); **(d)** zonal annual mean TAS from the EC-Earth2.3 (blue, CMIP5) and EC-Earth3 (red, CMIP6) ensembles.

ponent of wind of EC-Earth2.3 and EC-Earth3 are assessed against ERA5. Zonal averages (Fig. 8a, b) show that both models are characterized by an underestimation of the Southern Hemisphere jet, which is larger in the upper troposphere. However, EC-Earth3 shows a reduction of this negative bias compared to EC-Earth2.3.

Similar improvements are seen in the Northern Hemisphere: here EC-Earth2.3 was characterized by a negative bias on the order of $-5\,\mathrm{m\,s^{-1}}$ in the core of the upper-tropospheric jet stream, while EC-Earth3 shows limited bias with a slight overestimation on its equatorward side at the tropopause level.
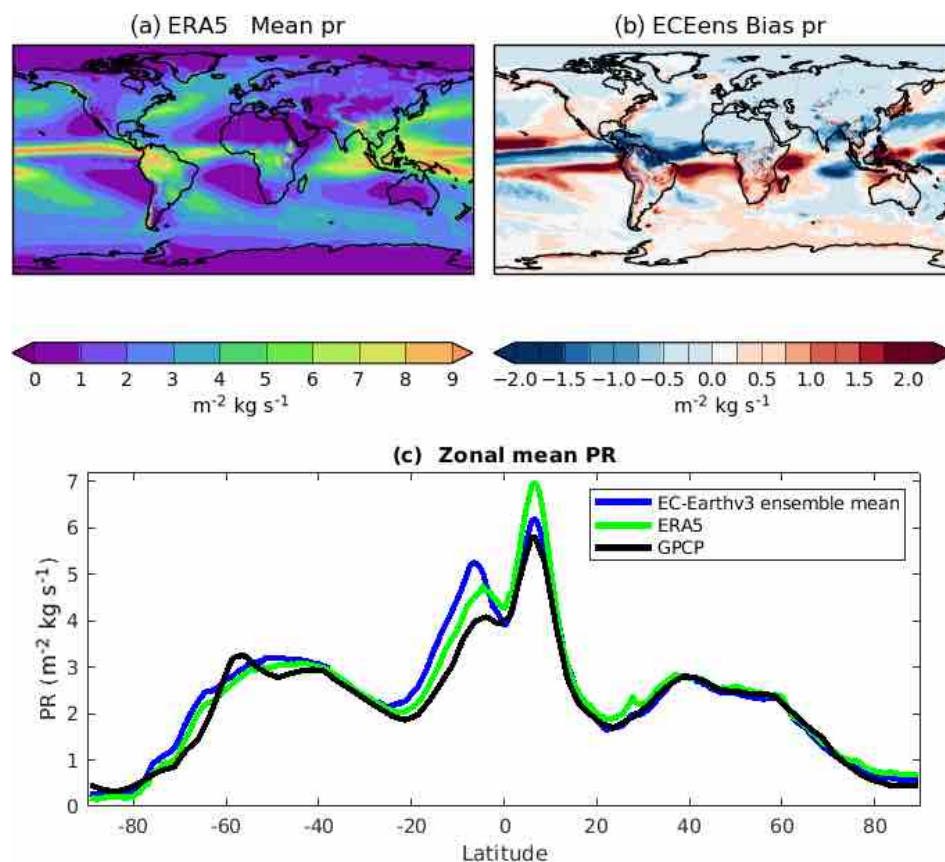
Figure 6. Mean precipitation for the period 1980–2010 for ERA5 **(a)**, precipitation anomaly with respect to ERA5 for EC-Earth3veg **(b)**, and zonal mean precipitation for ERA5 **(c)** (green), GPCPv2.2 (black), and EC-Earth3veg ensemble mean (blue).

Looking at upper-level zonal wind patterns (Fig. 8c, d) it can be seen that considerable improvements have been obtained for the Pacific jet, which is strongly underestimated in EC-Earth2.3 and shows a minor southward bias in EC-Earth3. Conversely, the North Atlantic jet still shows a poleward displacement – extending from the western Atlantic up to eastern Europe – of the same magnitude in both configurations. A southward displacement of the subtropical jet over Africa also emerges in EC-Earth3.

Overall, CMIP6 EC-Earth3 shows a reduction of the bias when compared with CMIP5 EC-Earth2.3. This is confirmed by smaller figures for both root mean squared error and mean bias shown in the top of each panel in Fig. 8. Larger biases in both models are observed when looking at specific seasons (not shown): EC-Earth3 shows overall smaller biases in DJF and larger biases in JJA than its CMIP5 predecessor. In general, EC-Earth3 is characterized by an underestimation of the winter jet and by an overestimation of the equatorward component of the tropical jet in the summer hemisphere. An underestimation of the winter stratospheric polar vortex – stronger in the Southern Hemisphere – is also found.

6.3 Blocking

Atmospheric blocking – the recurrent long-lasting quasi-stationary high-pressure system developing at the exit of the jet streams in midlatitudes (Woolings et al., 2018) – is assessed in Fig. 9 for both EC-Earth2.3 and EC-Earth3 against ERA5 reanalysis. Analysis of blocking is relevant for climate models considering both (1) the large impact on regional weather that blocking has in both summer and winter seasons (e.g., Sillmann et al., 2011; Schaller et al., 2018) and (2) the constant struggle that GCMs experience in correctly simulating the observed blocking frequencies (e.g., Masato et al., 2013; Davini and D'Andrea, 2020). Blocking is shown here as percentage of blocked days per season following the definition of Davini and D'Andrea (2020), according to which a bidimensional blocking index based on the reversal of the geopotential height gradient at 500 hPa is used.

Small differences arise from the comparison between EC-Earth2.3 and EC-Earth3, supported by the negligible changes in root mean squared error and mean bias reported in each panel. Both models are in line with the CMIP5 and CMIP6 multi-model mean (see Davini and D'Andrea, 2020). The most evident improvement is the reduced bias in winter
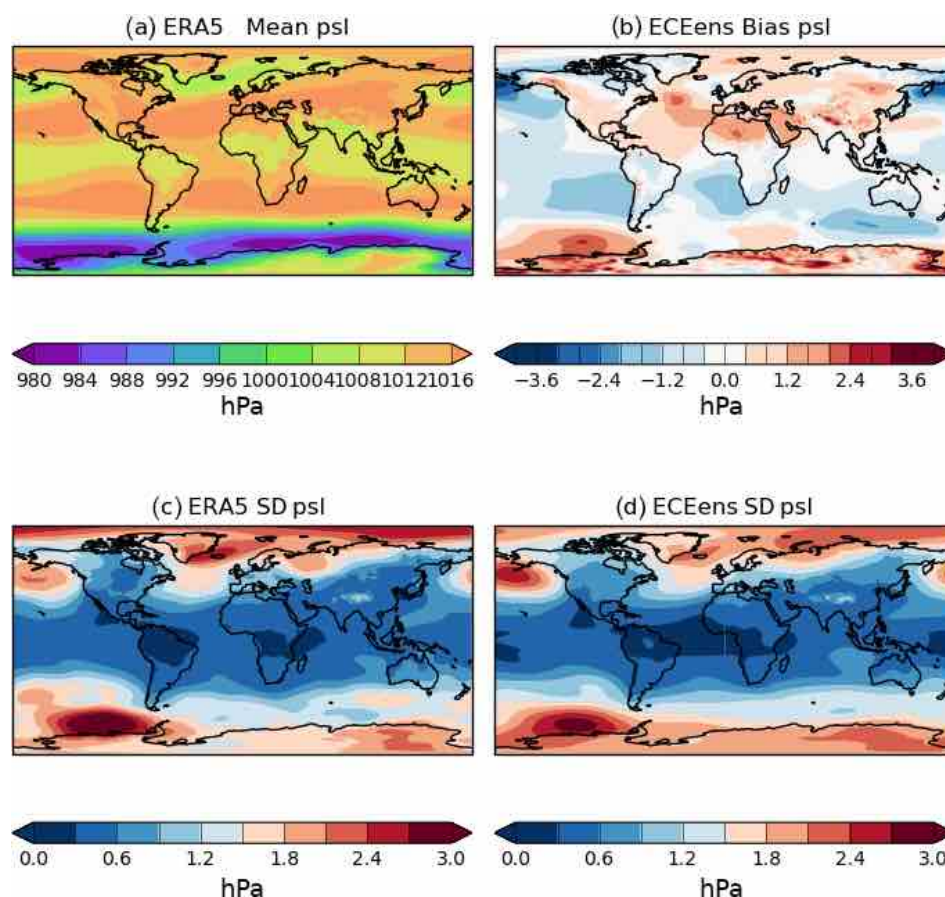
Figure 7. Upper panel: mean PSL 1980–2010 (hPa) for ERA5 **(a)** and for the EC-Earth3 ensemble bias compared to ERA5 **(b)**. Lower panel: interannual standard deviation (hPa) for ERA5 **(c)** and for the EC-Earth3 ensemble **(d)**.

North Pacific blocking in EC-Earth3 compared to its predecessor. However, both EC-Earth2.3 and EC-Earth3 show the common underestimation of winter European blocking with similar magnitude. It is interesting to notice that while in EC-Earth2.3 the winter European bias is characterized by a north–south dipole probably associated with an equatorward displacement of the Atlantic eddy-driven jet, in EC-Earth3 the dipole is located on the east–west axis, suggesting the presence of a jet that is too penetrative over the European continent (Fig. 9a, b).

Larger biases are seen in summer: negligible differences between EC-Earth2.3 and EC-Earth3 are also found, with slightly larger North Pacific and European blocking biases observed in EC-Earth3.

6.4 Sea ice

We compare EC-Earth3 historical ensemble mean sea ice variables and spread to observations and reanalysis datasets.

The ensemble mean of EC-Earth3 slightly overestimates the Arctic sea ice area (Fig. 10), whereby the ensemble spread encases the OSI SAF satellite observations (Lavergne et al., 2019). The trend in the Arctic sea ice area over

1980–2014 is captured well by the model during March and September (Fig. 10).

Interestingly, the modeled summer Arctic sea ice area minimum occurs in August instead of September as in observations. This result is in agreement with previous studies using NEMO-LIM forced by atmospheric forcing (Rousset et al., 2015; Docquier et al., 2017) and several coupled CMIP6 models including NEMO as an ocean component (Keen et al., 2021). The exact reason for the minimum sea ice area occurring in August is not clear (Keen et al., 2021).

While the total Arctic sea ice area is captured well by the model, there are large regional differences in the sea ice concentration between EC-Earth3 and satellite observations (Fig. 11). In March, the model overestimates the concentration near the ice margins in the Atlantic sector, including the Labrador, Greenland–Iceland–Norwegian (GIN), and Barents seas, while the concentration is underestimated by the model in the Bering Sea. In September, the sea ice concentration is generally overestimated by the model at the ice margins, with exceptions in the Kara, Laptev, and Chukchi seas, where the concentration is underestimated.
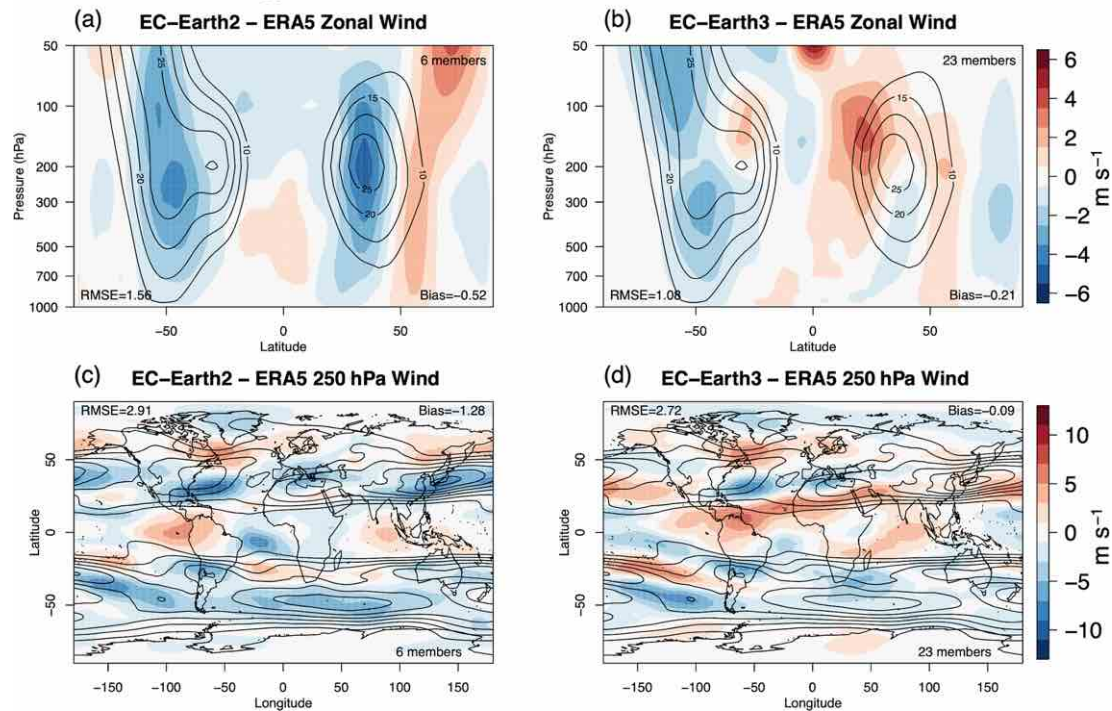
Figure 8. (a, b) Biases (colored shading) of the zonally averaged zonal component of wind for EC-Earth2.3 **(a)** and EC-Earth3 against ERA5 reanalysis over the 1980–2010 period. Contours show the ERA5 climatology. **(c, d)** Biases (colored shading) of the 250 hPa zonal component of wind for EC-Earth2.3 **(a)** and EC-Earth3 against ERA5 reanalysis over the 1980–2010 period. Contours show the ERA5 climatology. Root mean square error, mean bias, and the number of ensemble members used are reported in each panel.

The total Arctic sea ice volume is higher in the ensemble mean of EC-Earth3 compared to the PIOMAS reanalysis (Fig. 12). The PIOMAS volume is close to the lower edge of the EC-Earth3 ensemble spread, which is considerably large, of about $30 \times 10^3 \, \text{km}^3$ in the 1980s and 1990s. This is consistent with the Arctic sea ice being generally thicker in the model compared to PIOMAS (Fig. 13). In September, the Arctic sea ice is clearly too thick in the model with a bias up to 2 m compared to PIOMAS. In March, the Arctic sea ice thickness is overestimated by EC-Earth3 in the central Arctic, while in the Bering and Karas seas the thickness is lower compared to PIOMAS.

We chose PIOMAS as a reference product for sea ice thickness and volume because of the relatively long available time frame (i.e., from 1979 to now) compared to observational products, which cover much shorter periods. Uncertainties in PIOMAS are related to the underlying ocean and sea ice models, to the atmospheric forcing, and to the observational data available to the assimilation scheme. PIOMAS ice thickness estimates agree well with the ICESat ice thickness retrievals for the central Arctic, the area for which submarine data are available, with a mean difference smaller than 0.1 m, while differences outside this area are larger (Schweiger et al., 2011). Also, PIOMAS spatial thickness patterns agree well with ICESat thickness. PIOMAS ap-

pears to overestimate thin ice thickness and underestimate thick ice. The latter feature partly explains the higher ice thickness in EC-Earth3 in the central Arctic compared to PIOMAS (Fig. 13).

While the EC-Earth3 historical ensemble represents the Arctic sea ice area relatively well, it clearly underestimates the Antarctic sea ice area (Fig. 10) by ~ 5 million km^2 in September and ~ 2 million km^2 in March. This underestimation is linked to the warm bias in the Southern Ocean (Fig. 4) and is visible for all areas around Antarctica during the Southern Hemisphere summer (Fig. 14). The underestimation is more pronounced close to the ice edge than in the central pack ice during the Southern Hemisphere winter (Fig. 14). Also, the modeled trend in Antarctic sea ice area in both September and March is slightly negative, while observations show a slightly positive trend over 1979–2014 (Fig. 10). However, including the most recent years (after 2014), the observed Antarctic sea ice area does not exhibit any significant trend (Meredith et al., 2019).

Due to the absence of reliable long-term reanalysis and/or observational products for Antarctica, we do not show maps of sea ice thickness in the Southern Hemisphere.
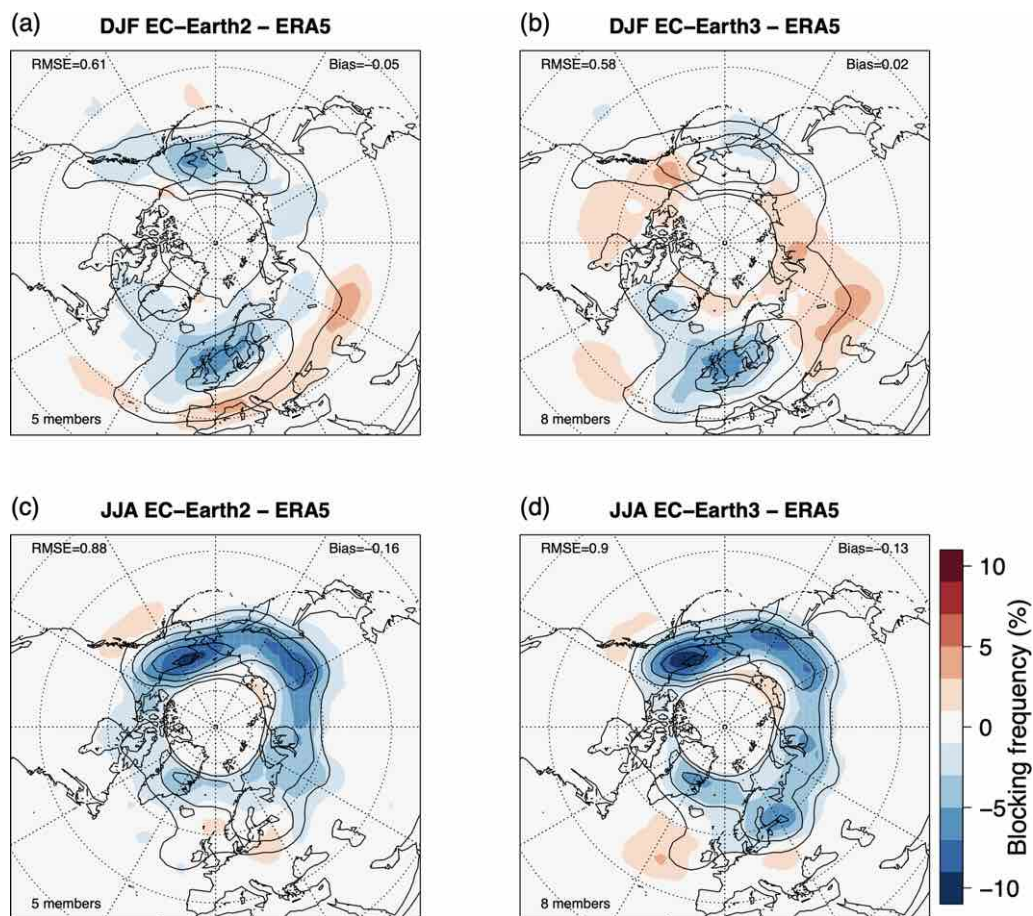
Figure 9. Biases for the blocking frequency according to the Davini and D'Andrea (2020) index for EC-Earth2 **(a, c)** and EC-Earth3 **(b, d)** against ERA5 reanalysis over the 1980–2010 period. Contours show the ERA5 climatology. Panels **(a)** and **(b)** are winter, and panels **(c)** and **(d)** are summer. Root mean square error, mean bias, and the number of ensemble members used are reported in each panel.
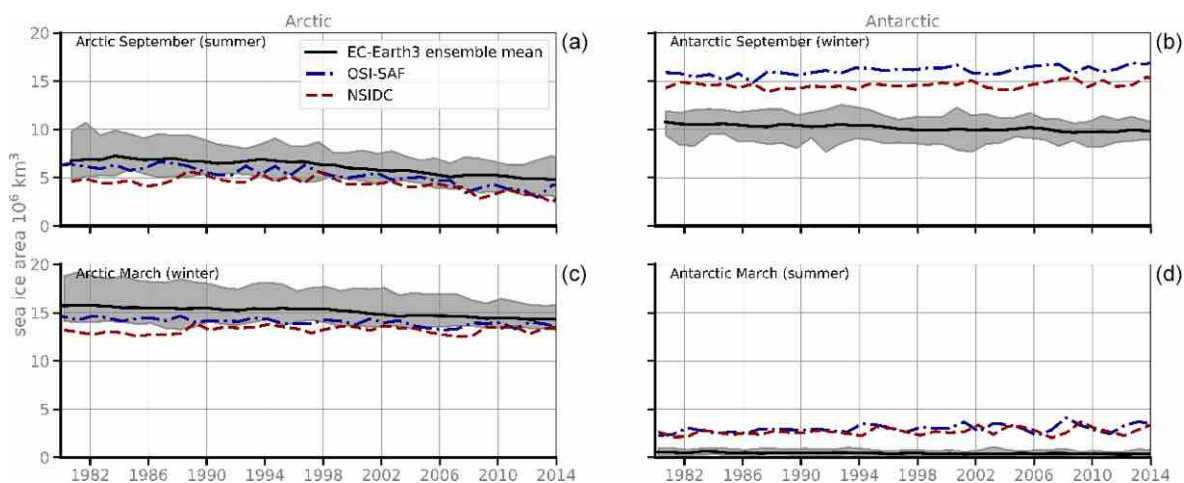
Figure 10. Time series of Arctic **(a, c)** and Antarctic **(b, d)** sea ice area for both EC-Earth3 (ensemble mean as thick solid lines) and satellite observations (OSI SAF as blue dash-dotted lines and NSIDC as dashed lines). The EC-Earth3 ensemble minimum up to maximum value is represented by the shading around the ensemble mean.
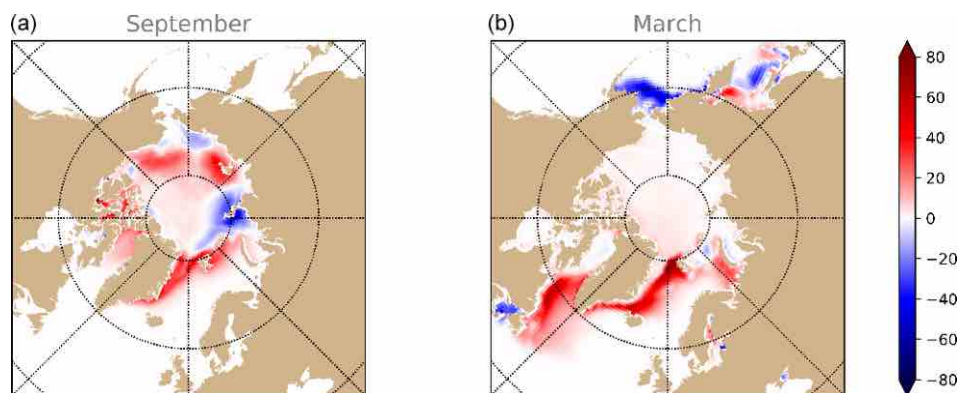
Figure 11. Difference in Arctic sea ice concentration in percent between the ensemble mean of EC-Earth3 and OSI SAF observations in September (**a**) and March (**b**), averaged over 1980–2010.
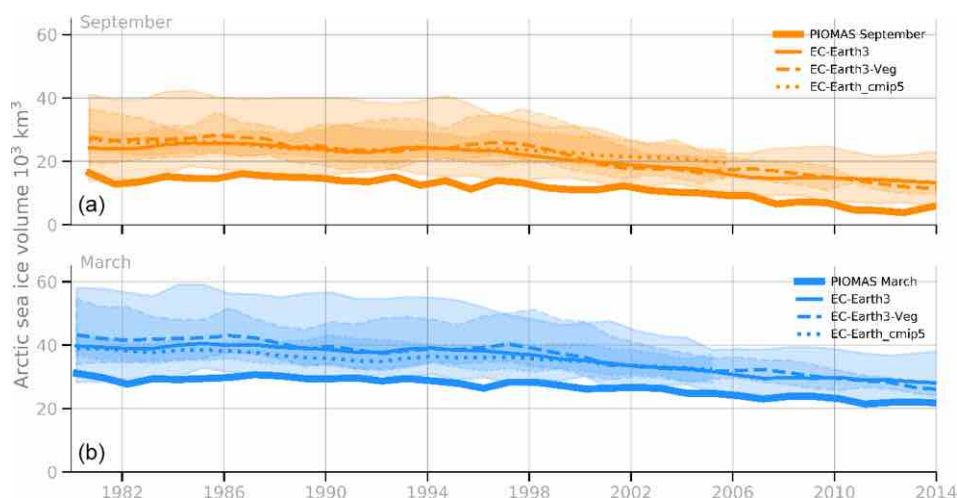
Figure 12. Time series of September (orange, upper panel) and March (blue, lower panel) Arctic sea ice volume for EC-Earth3 (thin solid lines representing the ensemble mean), EC-Earth3-Veg (dashed lines representing the ensemble mean), the CMIP5 version of EC-Earth (dotted lines), and PIOMAS reanalysis (thick solid lines). The EC-Earth3 and EC-Earth3-Veg ensemble minimum and maximum are represented by the same line style as their means, but with transparent shading added around the ensemble means.

6.5 AMOC

The Atlantic Meridional Overturning Circulation (AMOC) is connected with a northward flow of warm and salty water in the upper layers of the Atlantic Ocean and exports of cold and dense water southward in the deeper layers (Buckley and Marshall, 2016). The ensemble mean of the AMOC stream function obtained from the EC-Earth3 ensemble simulations (Fig. 15a), after being averaged over 1980–2010, features the expected overturning clockwise circulation cell with a maximum transport of 18 Sv centered at around 35° N and a depth of 1000 m. Compared to the 12-member ensemble mean of EC-Earth 2.3 (Brodeau and Koenigk, 2016) used for CMIP5 (no figure), the CMIP6 version of EC-Earth presented here has a stronger AMOC closer to observations (Smeed et al., 2018).

The ensemble mean time series of the AMOC index, defined as the maximum volume transport stream function between 24.5 and 27.5° N, covers values from well within the range of the RAPID-MOCHA array observations (Smeed et al., 2017). The ensemble mean shows a weak decrease of about 0.5 Sv from the year 1850 to 1876 with a relatively steady period until 1931 around 17.5 Sv, followed by an increase of around 2 Sv until 1980 and a decrease afterwards. Individual members of the ensemble vary between 2 and 5 Sv, with the upper range matching the RAPID observational variability well. It has to be noted that the RAPID data are available only for the last 20 years. Several other studies with ocean models forced by atmospheric reanalysis data (e.g., Yeager and Danabasoglu, 2014; Huang et al., 2012) show a later increase in AMOC between 1980 and the 1990s and a decrease after the mid-1990s. Most CMIP5 models also
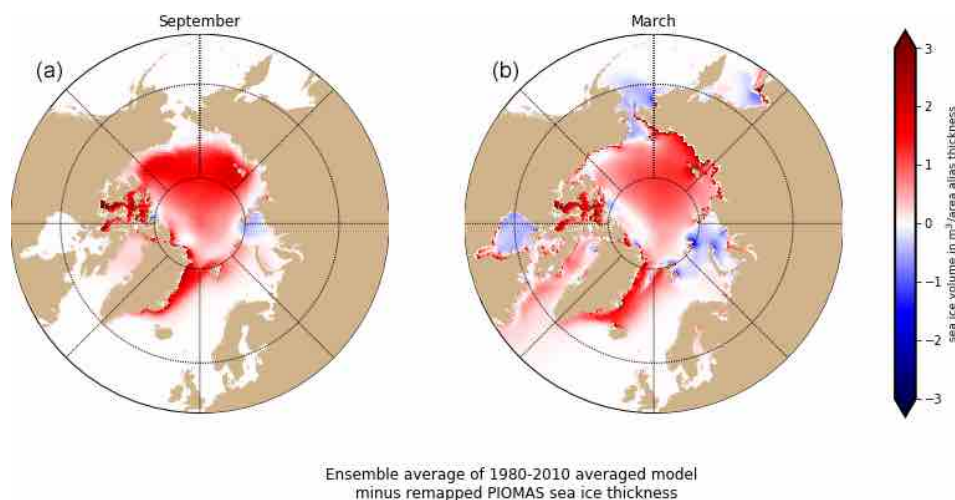
Figure 13. Difference in Arctic sea ice thickness between the ensemble mean of EC-Earth3 and PIOMAS in September **(a)** and March **(b)**, averaged over 1980–2010.
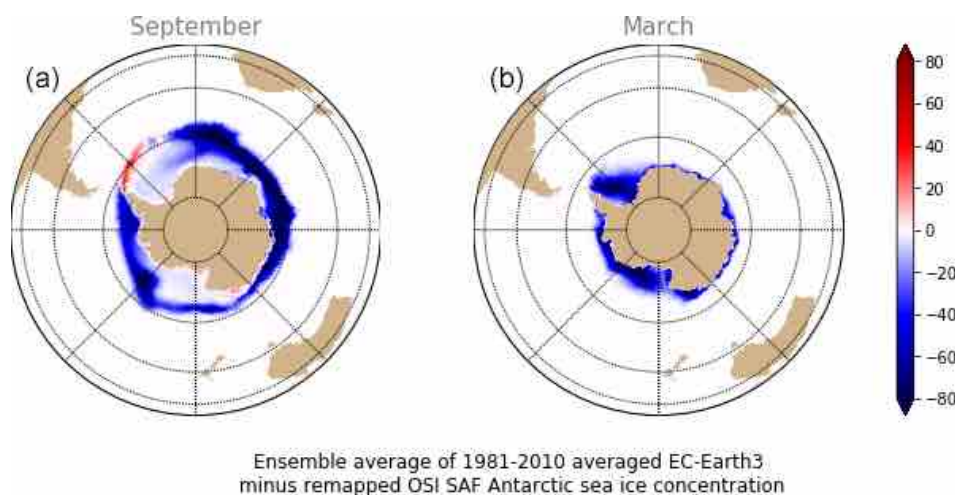
Figure 14. Difference in Antarctic sea ice concentration between the ensemble mean of EC-Earth3 and OSI SAF observations in September **(a)** and March **(b)**, averaged over 1980–2010.

show a later decrease in AMOC, mainly after the year 2000 (e.g., Collins et al., 2013; Cheng et al., 2013).

There is a wide range of variability (Fig. 15b) between ensemble members, possibly because each member starts from a different initial condition that evolves differently depending on the state of the model's internal variability. The variation between members occurs mainly at depth of 1000–2000 m and between 0 and 40° N, and it has a pattern and magnitude similar to the AMOC multidecadal variability (Fig. 2 in Boulton et al., 2014). A first analysis suggests that the lowest AMOC strength values correspond to extended periods with absent deepwater formation and expanded sea ice in the Labrador Sea. Our preliminary analysis suggests that those ensemble members with similar initial conditions (in terms of SST and SAT) have similar curves in their AMOC index

time series. This further suggests that those members with a more realistic initial state might be better in capturing the AMOC variability. Nevertheless, time series of 12 individual members (out of 20 members) show a roughly similar trend after 1950.

The AMOC (Fig. 16) shows a variability of 70 to 100 years as in ECHAM5 and the Max Planck Institute Ocean Model (Jungclaus et al., 2005) but a narrower range than the variability of 50–200 years in CCSM4 (Danabasoglu et al., 2012). One reason that could explain the somewhat long period of variability for our EC-Earth simulations is that the modeled sea ice is extending too far into the Labrador Sea, which keeps AMOC in a weaker state for a longer period of time before recovering. In CMIP5, most models without the
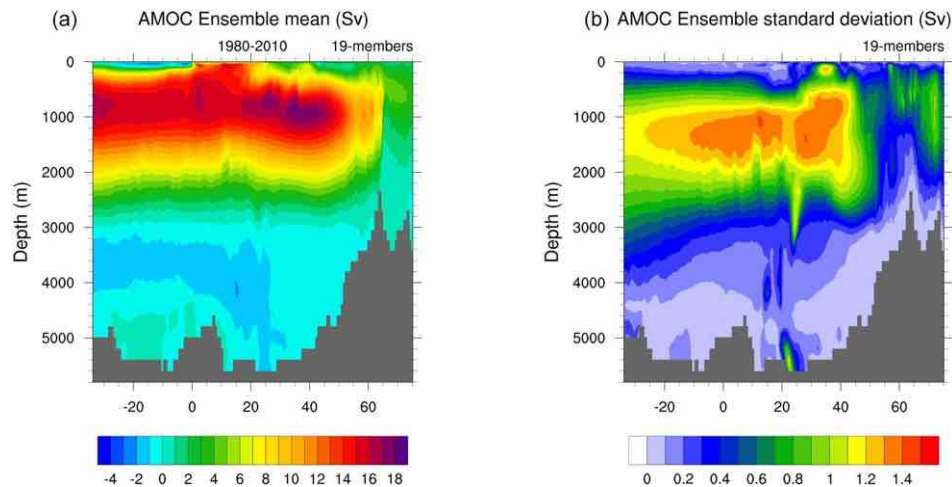
Figure 15. (a) AMOC stream function in the depth and latitude plane for the EC-Earth3 ensemble mean, averaged over 1980–2010 (in Sv). **(b)** Standard deviation of AMOC between the members. The *x* axis denotes degrees latitude.
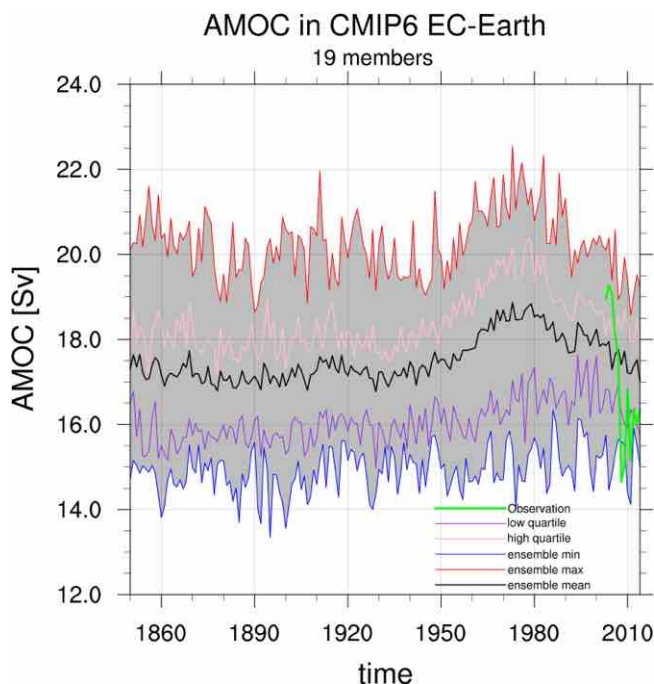
Figure 16. Time series of maximum AMOC (maximum between 24.5–27.5° N) for ensemble mean (black), low quartile (purple), high quartile (pink), ensemble minimum (blue), and ensemble maximum (red) are shown for 19 members of EC-Earth3. Observation data are shown with green from RAPID-MOCHA array observations (Smeed et al., 2017).

Figure 17. Ocean northward heat transport: the blue curve is the ensemble mean, and the light blue shading is delimited by the ensemble minimum and maximum. The black curve is taken from Trenberth et al. (2019). The hydrographic measurements at 26° N (RAPID) and 57° N (OSNAP) are shown as vertical bars.

tion estimates from Trenberth et al. (2019) that cover the period from 2000–2014.

6.6 Modes of variability

Since the North Atlantic Oscillation (NAO) is one of the most relevant modes of variability for the European climate, we assess its representation separately. The NAO is characterized by atmospheric oscillation between the Arctic and the subtropical Atlantic and can also be defined through changes in surface pressure (Hurrell et al., 2003). During winter, the positive phase of NAO has been defined when the difference between the Icelandic low-pressure center and the Azores high-pressure center is intensified: conversely, a negative NAO phase is when this difference is weaker than usual. The NAO pattern is estimated by calculating the leading empirical orthogonal function (EOF) of the detrended December to February (DJF) monthly sea level pressure (SLP) anomalies

convection were also covered by sea ice in the Labrador Sea in winter (Heuzé, 2017).

The ocean heat transport (Fig. 17) is related to the AMOC. North of 20° N it shows values slightly lower than observa-
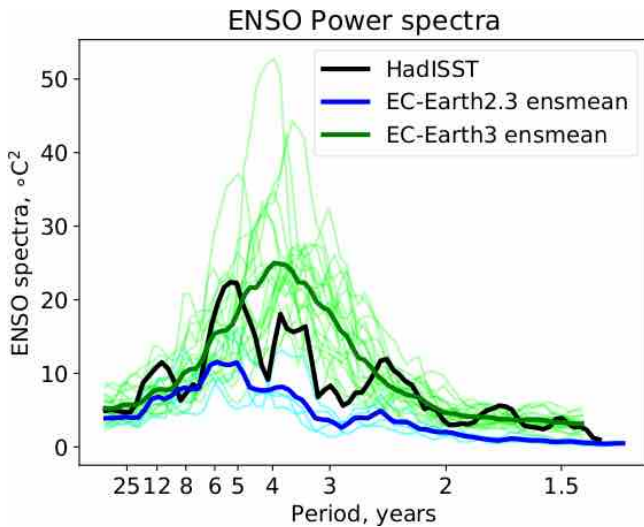
Figure 18. Power spectra of Niño3.4 sea surface temperature (SST) time series for HadISST observations (black thick line), 20 EC-Earth3 ensemble members (light green lines), and 4 EC-Earth 2.3 ensemble members (cyan lines). The EC-Earth3 and EC-Earth2.3 ensemble means are shown as green and blue thick lines, respectively. The Niño3.4 is calculated between latitudes 5° N–5° S and longitudes 170–120° W. The power spectra were calculated for the 1900–2009 period.

over the Atlantic sector (20–80° N, 90° W–40° E). The ERA-20C and both EC-Earth ensemble means show very similar spatial patterns (Fig. 20), with EC-Earth2.3 showing more intense negative values over Iceland and less intense positive values over the Bay of Biscay and northern Spain when compared to EC-Earth3 and observations.

Overall the representation of the NAO in EC-Earth3 shows minor improvements with respect to EC-Earth2.3: the root mean square error (RMSE) of EC-Earth3 is 0.64 compared with 0.52 for EC-Earth2.3.

Previous assessments of coupled models have found that the CMIP5 and CMIP3 models represent NAO very similarly as in reanalyses, with no general improvements of CMIP5 compared to CMIP3 models (Davini and Cagnazzo, 2014).

In EC-Earth3, the ENSO spatial patterns and amplitude are well represented (not shown). In order to assess the variability of ENSO in the EC-Earth3 ensemble we calculated the Niño3.4 SST power spectra (Fig. 18) for the model and for the HadISST observational dataset (Rayner et al., 2003). HadISST shows the highest peak at 5.5 years and a range of periods between 3.3 and 3.7 forming the second highest peak. The mean of the EC-Earth3 ensemble spectra shows the highest peak at around 3.7 years. Some individual members show their main peak at about 5.5 years, similar to the HadISST dataset. Most of the ensemble members (two with the highest peaks) show their maximum in the 3–4-year period.

EC-Earth3 shows larger ENSO spectral power than EC-Earth2.3, with an ensemble mean closer to the HadISST observations, especially in frequencies with high energy. Thus, we see a clear improvement in the representation of ENSO in EC-Earth3. Similar improvements were seen earlier for CMIP5 models (Jha et al., 2014). While the frequency distribution in EC-Earth3 is distinctly improved compared to EC-Earth2.3, a challenge is still seen in representing distinct peaks in the power spectrum.

The range among the ENSO spectra of different members is considerable. Climatologically, most ensemble members show only small differences over the tropics when compared with the whole ensemble mean. Although the members with the most energetic ENSO share some climatological features (cold Arctic and Labrador seas) compared to the ensemble mean, the reason why they have developed a more energetic ENSO remains unclear.

Previous studies have shown that winters of El Niño have an impact on the circulation over the North Atlantic, favoring a negative phase of the NAO in the late winter (e.g., García-Serrano et al., 2011) teleconnected via the extratropical Pacific–North American (PNA) pattern (Enfield and Mayer 1997; Giannini et al., 2000). This nonstationary relationship between ENSO and the atmospheric variability over the North Atlantic and European sector is already reproduced in the internal variability behavior of coupled ocean–atmosphere systems in CMIP5 simulations (López-Parages et al., 2016)

To better assess this relationship between ENSO and NAO, we calculated the regression of the Niño3.4 index on sea level pressure (SLP) for the winter season (December to February, DJF) for ERA-20C reanalysis and for both EC-Earth ensembles (EC-Earth3 and EC_Earth2.3) (Fig. 19). As reported by previous studies (Wallace and Gutzler, 1981), a positive PNA-like pattern arises with an intensified lower Aleutian low and lower pressure over the eastern United States. This low pressure extends along the North American east coast into the central North Atlantic Ocean, while there is higher pressure over and between Iceland and Scandinavia, a pattern that is similar to a negative phase of NAO (Fig. 19). The EC-Earth3 ensemble demonstrates a high resemblance to ERA-20C (Fig. 19), which is much improved compared to the EC-Earth2.3 ensemble (Fig. 19). The latter shows weaker SLP anomalies, which can be associated with a weaker ENSO intensity in EC-Earth2.3 compared to the newer EC-Earth3 and to observations. Thus, EC-Earth3 gives a clear improvement of the ENSO–NAO link.

Previous research has linked La Niña and El Niño events to the positive and negative NAO patterns (Fereday at al., 2020). Although this link is relatively weak due to the fact that internal atmospheric variability is large in the North Atlantic European (NAE) region (Brönnimann, 2007), it depends on ENSO strength (Jiménez-Esteve and Domeisen, 2019; Toniazzo and Scaife, 2006). Therefore, a more energetic ENSO (more comparable in scale with observations) such as in the
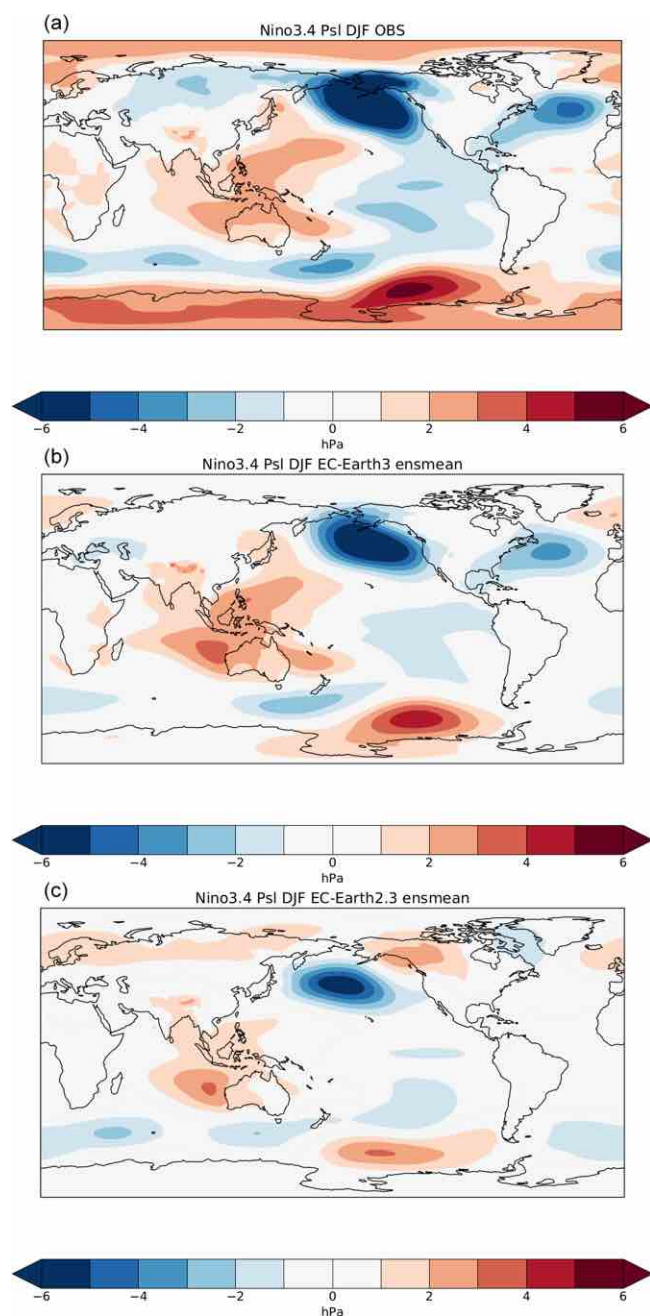
Figure 19. Regression of Niño3.4 SST index onto sea level pressure during winter (DJF) for the ERA-20C reanalysis **(a)**, 20-member EC-Earth3 ensemble mean **(b)**, and 4-member EC-Earth2.3 ensemble mean **(c)**. Analyses are based on 3-month means for December–February.

current EC-Earth3 could impact the intensity and sign of NAO.

The PNA pattern in EC-Earth3 shows strong spatial similarities with ERA-20C, which is an improvement compared to EC-Earth2.3 (not shown). RMSEs are respectively 0.9 and 0.54.

EC-Earth3 also shows a more realistic behavior than EC-Earth2.3 with respect to typical features of the PNA patterns, i.e., an above-average surface pressure over the subtropical Pacific (west of Hawaii) and over western Canada, together with below-average surface pressure over the North Pacific Ocean and along the southeastern United States.

The quasi-biennial oscillation (QBO) dominates the inter-annual variability in the tropical stratosphere (Baldwin et al., 2001). As in most climate models, the vertical discretization and the parameterization of gravity waves are crucial (e.g., Bushell et al., 2020) for a realistic representation of the QBO in EC-Earth.

In EC-Earth 2.3, with 62 vertical levels and without an ad hoc tuning, the zonal wind in the equatorial band is easterly on average, reaching $-35 \, \mathrm{m \, s^{-1}}$ close to the 5 hPa level (Fig. 21a). Note that results in this panel are based on model level outputs interpolated onto equivalent pressure levels. CMIP5 pressure level outputs are available only up to 20 hPa for other ensemble members.

Like other models participating in CMIP6 (Richter et al., 2020), the realism of the modeled QBO is notably improved in EC-Earth3, with 91 vertical levels and a revised gravity wave scheme (see Sect. 3.1). A marked alternation of easterly and westerly zonal wind shears can be seen between approximately 100 and 5 hPa in Fig. 21b. While a realistic QBO-like oscillation has been obtained, there is no improvement in the zonal wind above 5 hPa, as it remains easterly year-round without a clear sign of any stratospheric semi-annual oscillation. This phenomenon, linked with the QBO, is also driven by vertically propagating waves (Smith et al., 2019).

From the comparison of the amplitude profiles with ERA5 (Fig. 21c), it is clear that, even if still underestimated, the zonal wind variability in the EC-Earth3 ensemble members is improved with respect to those of EC-Earth2.3. As with other climate models (Bushell et al., 2020), the maximum of the oscillation peaks somewhat higher (at 10 hPa) in EC-Earth3 than in the reanalysis (closer to 20 hPa).

In Fig. 21d we also report the Fourier spectra at the 30 hPa (25 km) level, again including ERA5 and ensemble members from EC-Earth2.3 and EC-Earth 3. For the CMIP5 model only the annual harmonic stands out, whereas the EC-Earth 3 spectrum at 30 hPa resembles that of ERA5 much more closely, despite the modeled QBO being slightly faster.

6.7 Carbon cycle components

As an outlook to the EC-Earth3-CC configuration, we present first results here from the carbon cycle components of one member of the historical experiment. In this configuration, the physical climate is equivalent to that of the EC-Earth3-Veg configuration. However, besides the land vegetation and biogeochemistry model, the ocean biogeochemistry model has been activated. The latter describes the evolution of biogeochemical tracers in the ocean, but it does
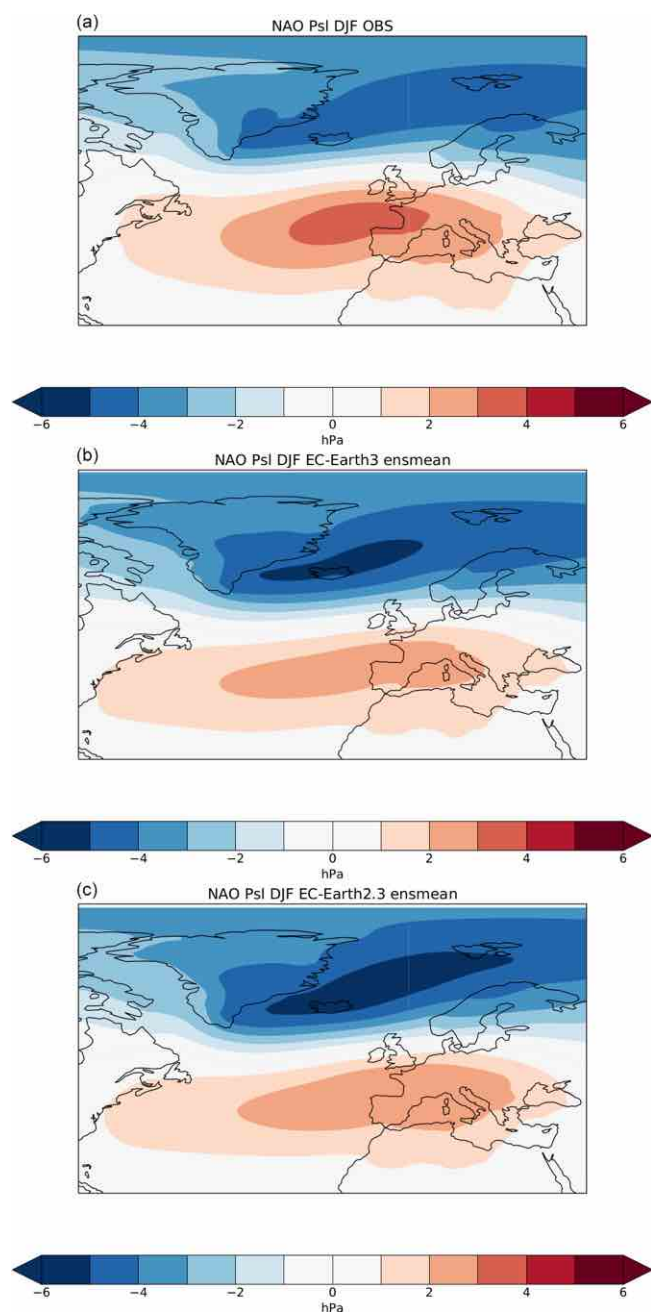
Figure 20. Regression of sea level pressure (SLP) anomalies to the NAO index during winter (December to February) for the ERA-20C reanalysis **(a)**, 20-member EC-Earth ensemble mean **(b)**, and 4-member EC-Earth2.3 ensemble mean **(c)**.

not have any feedback on climate. Both land vegetation and ocean biogeochemistry were equilibrated with a 1600-year-long spin-up keeping atmospheric CO_2 at pre-industrial level (284.32 ppm). The completion of the spin-up was assessed following the recommendation of C4MIP (Jones et al., 2016), with both the ocean and land C stocks having to drift by less than 10 Pg C per century. Once this condition was

met, a piControl simulation and historical simulation were started from the same initial conditions.

The ocean's cumulative C uptake between 1870 and 2014 is 150.78 Pg C, in good agreement with the estimate of 155 ± 20 Pg C by the Global Carbon Project (GCP) 2015 for the same period (Le Queré et al., 2015). The average air–sea CO_2 exchange between the atmosphere and the ocean for the period 1981–2014 is compared here to the observation-based reconstruction of Landschutzer et al. (2016) for the same period (Fig. 22). The overall spatial distribution of the air–sea CO_2 flux is in broad agreement with observations, with the largest differences occurring at high latitudes. In particular, in the Southern Ocean EC-Earth3-CC exhibits a large negative bias (i.e., more outgassing) in some regions due to the presence of extended periods of unrealistic open-sea convection. In the North Pacific and North Atlantic, EC-Earth3-CC has a stronger CO_2 uptake than observations, likely due to a deeper mixed layer and frequent active convection in the Labrador Sea, respectively.

In Fig. 23 we show CO_2 fluxes from the land and ocean components of EC-Earth3 for the period 1960–2014. The ocean CO_2 sink (SOCEAN; Fig. 23c) of the EC-Earth3-CC historical simulation closely follows the multi-model mean of the GCP 2019 (Friedlingstein et al., 2019), with differences in variability due to the mismatch between climate in EC-Earth-CC and the atmospheric forcing products used in the GCP. Land fluxes (computed from two historical EC-Earth3-Veg runs, one with dynamic land use change and management and one with land use fixed at 1850 levels) have increased following increased atmospheric CO_2 and temperature, but these show a stronger interannual variability in both emissions from land use change (ELUC; Fig. 23a) and terrestrial CO_2 sink (SLAND; Fig. 23b) compared to the estimates from the DGVMs in GCP 2019. This interannual variability originates from both climate differences between EC-Earth3-Veg and the observation-based forcing, which includes the right timing of weather events like El Niño in the offline simulations used for the GCP (Harris et al., 2014) and the mismatch between the dynamic vegetation in LPJ-GUESS compared to the land use dataset LUH2 (Hurtt et al., 2020) that dictates where and when land use change occurs.

7   Summary and conclusions

The EC-Earth research consortium represents a community of European institutes developing and utilizing the EC-Earth Earth system model. In this paper we document the overall concept of EC-Earth3, the model version used for contributions to CMIP6, and its flexible coupling framework, major model configurations, a methodology for ensuring the simulations are comparable across different HPC systems, and the physical performance of base configurations over the historical period. Simulations described in this paper have been carried out under the CMIP6 framework conditions (Eyring
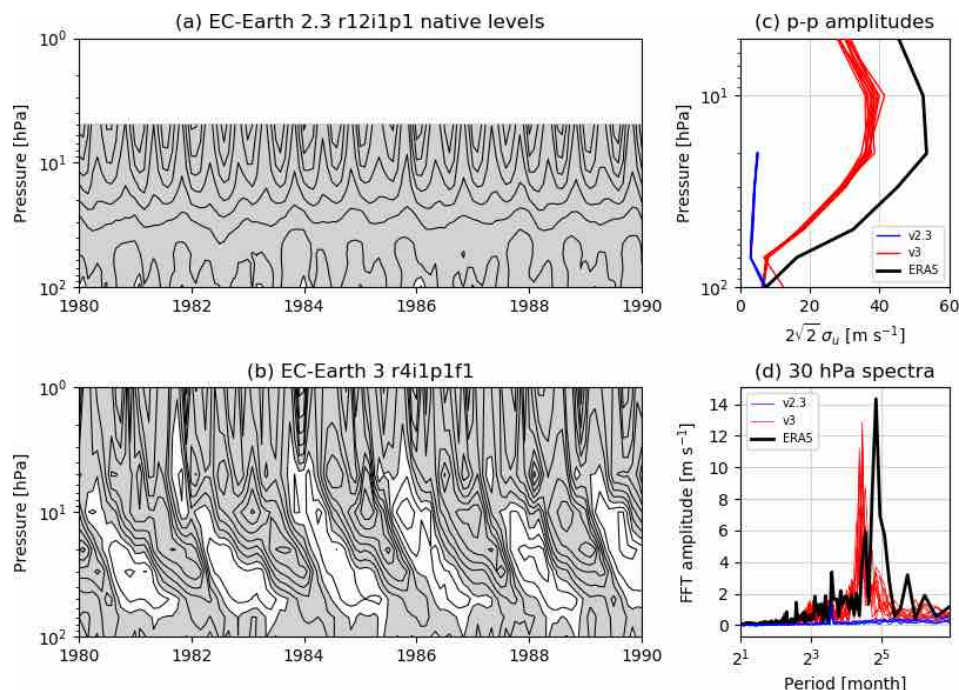
Figure 21. 10 years of the equatorially (within 5° from the Equator) averaged monthly mean zonal wind in EC-Earth2.3 (r12i1p1, **a**) and EC-Earth3 (r4i1p1f1, **b**) as a function of time and altitude. Easterly winds are shaded in gray, and the contour interval is 5 m s^{-1}. The peak-to-peak amplitude profiles of the equatorial zonal wind, estimated from the temporal standard deviation for the period 1980–2010, are reported in (**c**) for ERA5 (black), 12 CMIP6 (red), and 6 CMIP5 (blue) ensemble members. In panel (**d**) the Fourier spectra at the 30 hPa level for ERA5 (black) and the same experiments in panel (**c**) are shown. Only for panel (**a**) are the results for EC-Earth2.3 interpolated onto pressure levels from native model levels.

et al., 2016). Wherever possible, we also compare the CMIP6 results to the previous model version EC-Earth2 under the conditions of CMIP5.

The different configurations of EC-Earth3 described in Sect. 2 are enabled by a flexible coupling framework. A traditional GCM configuration, comprising a coupled atmosphere and ocean model, in different spatial resolutions is accompanied by configurations with an interactive vegetation module, active atmospheric composition and aerosols, full carbon cycle configuration, and a configuration with a Greenland Ice Sheet model. The variety of possible configurations and submodels reflects the broad interests in the EC-Earth community. The releases of the different configurations were staggered in time to efficiently gain from preceding coupling and tuning efforts.

The different component models linked for the EC-Earth3 framework, described in Sect. 3, are partly community models, arise from developments of partner institutes, or arise from centralized efforts such as for the forecast model IFS developed by ECMWF.

Model tuning has been carried out for different configurations. It was feasible to share identical tuning for both the GCM (base configuration EC-Earth3) and the configuration coupled to dynamic vegetation for a given resolution. Different tuning outcomes had to be applied to varying resolutions and for the configuration with interactive aerosols and atmospheric composition.

For standard resolutions, the goal of atmosphere tuning was to minimize the error in the global means of the net radiative flux at the surface, the top-of-atmosphere (TOA) fluxes, and longwave and shortwave cloud forcing without compromising the radiative imbalance at the top of the atmosphere. In the configurations with the coupled ocean, it was possible to adjust only ocean and sea ice parameters while maintaining atmospheric tuning to prevent the model from drifting under constant forcing.

For low-resolution configurations largely used for millennium-scale simulations, we aimed at a tuning for a climate in a radiative equilibrium to prevent the global mean surface temperature from drifting under the conditions of a stable climate.

Running the ESM on different HPC systems across various partner institutions constitutes a challenge for comparability of simulations. Therefore, the EC-Earth consortium has chosen to implement a protocol that judges compatibility based on statistical differences between ensembles carried out in different computing environments. The protocol is applied in all cases when simulations are shared between partners with their respective HPC systems.
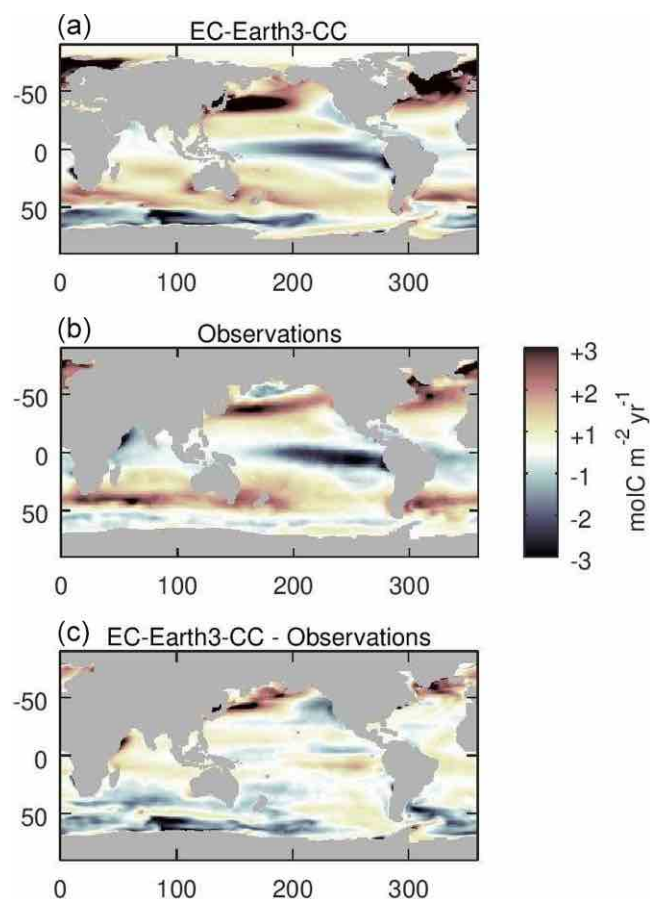
Figure 22. Air–sea CO_2 flux averaged over the period 1982–2014 for EC-Earth3-CC **(a)**, the observation-based reconstruction by Landschutzer et al. (2016; **b**), and their difference **(c)**.

The basic physical performance of EC-Earth3 is presented by a number of key indicators and quantities, with a focus on CMIP6 historical simulations carried out for 1850–2014. EC-Earth3 represents a large step forward compared to previous versions. As the basic configuration we chose EC-Earth3, the classic GCM configuration, because of the larger ensemble of simulations compared to the EC-Earth3-Veg configuration, which is the GCM coupled to the dynamic vegetation model. As performance metrics we chose global means of key variables, their geographical patterns, behavior of oscillation patterns, and circulation features.

We find that the global mean temperature in the historical ensemble has a warm bias of about 0.5 K in comparison with ERA5, which is mainly due to a strong warm bias in the Southern Ocean area. We find an oscillatory behavior between two states that are characterized by low or high values of the AMOC, cold or warm North Atlantic temperatures, and more or less sea ice in the Arctic. There is an indication that the oscillation might be reduced as the climate warms. The global warming over the historical period, given as the near-surface air temperature difference between the periods
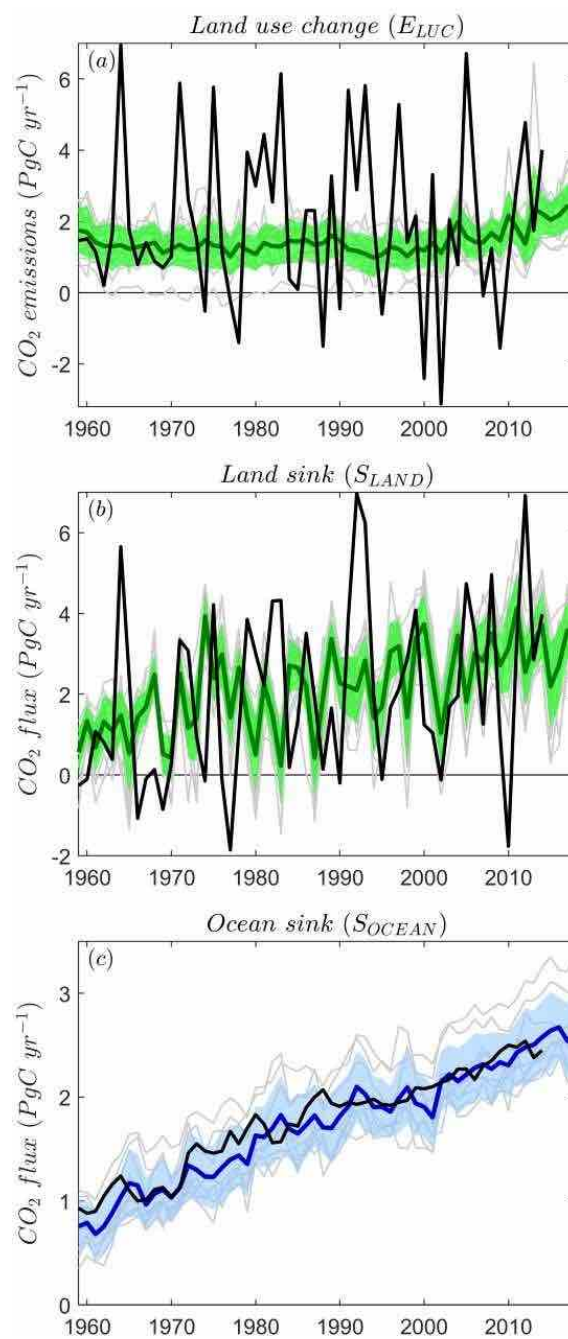
Figure 23. CO_2 exchanges between the atmosphere and the terrestrial biosphere for **(a)** CO_2 emissions from land use change (ELUC) and **(b)** land CO_2 sink (SLAND) with individual DGVMs (gray), the multi-model mean (dark green), and its range ($\pm 1\sigma$, light green shading) from GCP overlaid with EC-Earth3-Veg fluxes (black). Panel **(c)** shows the CO_2 exchange between the atmosphere and ocean (SOCEAN) with individual ocean models (gray), the multi-model mean (dark blue), and its range ($\pm 1\sigma$, light blue shading) from GCP overlaid with EC-Earth3-CC fluxes (black).

1981–2010 and 1851–1880, is 0.7 K, which is only slightly higher than the 0.63 K estimate for the observed warming (IPCC; Hoegh-Guldberg et al., 2018).

The global mean precipitation patterns are well represented in EC-Earth3, while the amplitude is overestimated, pointing to an overestimation of the hydrological cycle. Mean sea level pressures are close to ERA5 for most geographical areas, as is the interannual variability.

The ensemble mean Arctic sea ice area is slightly overestimated in several regions, while the trend since 1980 is captured well by the model during all months. Furthermore, the total Arctic sea ice volume is overestimated compared to reanalysis data. EC-Earth3 clearly underestimates the Antarctic sea ice area as a consequence of a warm bias in the Southern Ocean.

Zonal winds in EC-Earth3 are characterized by an underestimation of the winter westerly jet and by an overestimation of the equatorward component of the tropical jet in the summer hemisphere. An underestimation of the winter stratospheric polar vortex – stronger in the Southern Hemisphere – is identified.

Atmospheric blocking shows a typical bias over the winter North Pacific and the common underestimation of winter European blocking, with indications of a jet that is too penetrative over the European continent.

The ensemble AMOC index covers values from well within the range of existing observations. Variability in the individual members of the ensemble varies between 2 and 5 Sv, which is in line with observed decadal-scale variability.

The NAO and PNA patterns in the EC-Earth3 ensemble show a spatially high resemblance to the ERA-20C reanalysis. A small improvement compared to previous versions is in line with generally minor improvements between model generations since CMIP3. A clear improvement of the ENSO–NAO link is seen.

The ENSO power spectrum for EC-Earth3 shows peak amplitudes close to the observations, which is a pronounced improvement compared to older versions. Also, the frequency distribution in EC-Earth3 is markedly improved, though representing distinct peaks in the power spectrum is still a challenge.

The realism of the modeled QBO is notably improved in EC-Earth3 thanks to increased vertical resolution and a revised non-orographic gravity wave scheme.

Parallel papers extend the current analysis to improve our understanding of the EC-Earth3 climate sensitivity (Wyser et al., 2020a), the impact of new emission scenarios (Wyser et al., 2020b), results from the high-resolution configuration (Haarsma et al., 2020), and a platform comparison study (Massonnet et al., 2020). Forthcoming papers will, for example, explore dynamic oscillations during the PI control, assess the skill of climate forecasts (Bilbao et al., 2021), and highlight future climate projections.

In summary, the EC-Earth3 key performance metrics demonstrate physical behavior and biases well within the

frame known from CMIP5 models, with improved physical and dynamic features (Sects. 2 and 3), new ESM components, a much more flexible system framework, community tools, and largely improved indicators compared to the CMIP5 version. In short, EC-Earth3 represents a large step forward for the only European community ESM. We show here that EC-Earth3 is suited for a range of tasks in CMIP6 and beyond.

Author contributions. All authors contributed to the development of the code or to the analysis of results. All authors contributed to the discussion of the results and the final paper. The majority of authors provided contributions across several sections of the paper.

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

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Abdul-Razzak, H. and Ghan, S. J.: A parameterization of aerosol activation: 2. Multiple aerosol types, J. Geophys. Res., 105, 6837–6844, https://doi.org/10.1029/1999JD901161, 2000.

Alessandri, A., Catalano, F., De Felice, M., Van Den Hurk, B., Doblas Reyes, F., Boussetta, S., Balsamo G., and Miller, P. A.: Multi-scale enhancement of climate prediction over land by increasing the model sensitivity to vegetation variability in EC-Earth Clim. Dynam., 49, 1215, https://doi.org/10.1007/s00382-016-3372-4, 2017.

Arakawa, A. and Lamb, V. R.: Computational design of the basic dynamical processes of the UCLA general circulation model, General circulation models of the atmosphere, 17, Supplement C: 173–265, 1977.

ARPEGE-Climate Version 5.1: Algorithmic Documentation, Technical report by Meteo France, 2008

Ashwanden, A., Bueler, E., Khroulev, C., and Blatter, H.: An enthalpy formulation for glaciers and ice sheets, J. Glaciol., 58, 441–457, https://doi.org/10.3189/2012JoG11J088, 2012.

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., and Gehlen, M.: PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies, Geosci. Model Dev., 8, 2465–2513, https://doi.org/10.5194/gmd-8-2465-2015, 2015.

Axell, L. B.: Wind-driven internal waves and Langmuir circulations in a numerical ocean model of the southern Baltic Sea, J. Geophys. Res., 107, 3204, https://doi.org/10.1029/2001JC000922, 2002.

Balaji, V., Maisonnave, E., Zadeh, N., Lawrence, B. N., Biercamp, J., Fladrich, U., Aloisio, G., Benson, R., Caubel, A., Durachta, J., Foujols, M.-A., Lister, G., Mocavero, S., Underwood, S., and Wright, G.: CPMIP: measurements of real computational performance of Earth system models in CMIP6, Geosci. Model Dev., 10, 19–34, https://doi.org/10.5194/gmd-10-19-2017, 2017.

Baldwin, M. P., Gray, L.P., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., Holton, J. R., Alexander, M. J., Hirota, I., Horinouchi, T., Jones, D. B. A., Kinnersley, J. S., Marquardt, C., Sato, K., and Takahashi, M.: The quasi-biennial oscillation, Rev. Geophys., 39, 179–229, https://doi.org/10.1029/1999RG000073, 2001.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K.: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System, J. Hydrometeorol., 10.3, 623–643, 2009.

Bechtold, P., Semane, N., Lopez, P., Chaboureau, J., Beljaars, A., and Bormann, N.: Representing Equilibrium and Nonequilibrium Convection in Large-Scale Models, J. Atmos. Sci., 71, 734–753, https://doi.org/10.1175/JAS-D-13-0163.1, 2014

Beckmann, A. and Döscher, R.: A Method for Improved Representation of Dense Water Spreading over Topography in Geopotential-Coordinate Models, J. Phys. Oceanogr., 27, 581–591, https://doi.org/10.1175/1520-0485(1997)027<0581:AMFIRO>2.0.CO;2, 1997.

Bellucci, A., Athanasiadis, P.J., Scoccimarro, E., Ruggieri, P., Gualdi, S., Fedele, G., Haarsma, R. J., Garcia-Serrano, J., Castrillo, M., Putrahasan, D., and Sanchez-Gomez, E.: Air-Sea interaction over the Gulf Stream in an ensemble of HighResMIP present simulations, Clim. Dynam., 56, 2093–2111, 2021.

Berger, A.: Long-Term Variations of Daily Insolation and Quaternary Climatic Changes, J. Atmos. Sci., 35, 2362–2367, https://doi.org/10.1175/1520-0469(1978)035<2362:LTVODI>2.0.CO;2, 1978.

Bilbao, R., Wild, S., Ortega, P., Acosta-Navarro, J., Arsouze, T., Bretonnière, P.-A., Caron, L.-P., Castrillo, M., Cruz-García, R., Cvijanovic, I., Doblas-Reyes, F. J., Donat, M., Dutra, E., Echevarría, P., Ho, A.-C., Loosveldt-Tomas, S., Moreno-Chamarro, E., Pérez-Zanon, N., Ramos, A., Ruprich-Robert, Y., Sicardi, V., Tourigny, E., and Vegas-Regidor, J.: Assessment of a full-field initialized decadal climate prediction system with the CMIP6 version of EC-Earth, Earth Syst. Dynam., 12, 173–196, https://doi.org/10.5194/esd-12-173-2021, 2021.

Bitz, C. M. and Lipscomb, W. H.: An energy-conserving thermodynamic model of sea ice, J. Geophys. Res.-Oceans, 104, 15669–15677, 1999.

Blanke, B. and Delecluse, P.: Variability of the Tropical Atlantic Ocean Simulated by a General Circulation Model with Two Different Mixed-Layer Physics. J. Phys. Oceanogr., 23, 1363–1388, https://doi.org/10.1175/1520-0485(1993)023<1363:VOTTAO>2.0.CO;2, 1993.

Bougeault, P. and Lacarrere, P.: Parameterization of Orography-Induced Turbulence in a Mesobeta–Scale Model, Mon. Weather Rev., 117, 1872–1890, https://doi.org/10.1175/1520-0493(1989)117<1872:POOITI>2.0.CO;2. 1989.

Bouillon, S., Maqueda, M. A. M., Legat, V., and Fichefet, T.: Sea ice model formulated on Arakawa B and C grids, Ocean Model., 27, 174–184, 2009.

Boulton, C., Allison, L., and Lenton, T.: Early warning signals of Atlantic Meridional Overturning Circulation collapse in a fully coupled climate model, Nat. Commun., 5, 5752, https://doi.org/10.1038/ncomms6752, 2014.

Boussetta, S., Balsamo, G., Beljaars, A., Kral, T., and Jarlan, L.: Impact of a satellite-derived leaf area index

monthly climatology in a global numerical weather prediction modelm Int. J. Remote Sens., 34, 3520–3542, https://doi.org/10.1080/01431161.2012.716543, 2013.

Boysen, L. R., Brovkin, V., Pongratz, J., Lawrence, D. M., Lawrence, P., Vuichard, N., Peylin, P., Liddicoat, S., Hajima, T., Zhang, Y., Rocher, M., Delire, C., Séférian, R., Arora, V. K., Nieradzik, L., Anthoni, P., Thiery, W., Laguë, M. M., Lawrence, D., and Lo, M.-H.: Global climate response to idealized deforestation in CMIP6 models, Biogeosciences, 17, 5615–5638, https://doi.org/10.5194/bg-17-5615-2020, 2020.

Brandt, R. E., Warren, S. G., Worby, A. P., and Grenfell, T. C.: Surface Albedo of the Antarctic Sea Ice Zone, J. Climate, 18, 3606–3622, https://doi.org/10.1175/JCLI3489.1, 2005.

Brodeau, L. and Koenigk, T.: Extinction of the northern oceanic deep convection in an ensemble of climate model simulations of the 20th and 21st centuries, Clim. Dynam., 46, 2863, https://doi.org/10.1007/s00382-015-2736-5, 2016.

Brohan, P., Kennedy, J. J., Harris, I., Tett, S. F.m and Jones, P. D.: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850, J. Geophys. Res.-Atmos., 111, D12106. https://doi.org/10.1029/2005JD006548, 2006.

Buckley, M. W. and Marshall, J.: Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review, Rev. Geophys., 54, 5–63, 2016.

Brönnimann, S.: Impact of El Niño–Southern Oscillation on European climate, Rev. Geophys., 45, RG3003, https://doi.org/10.1029/2006RG000199, 2007.

Bueler, E. and Brown, J.: "Shallow Shelf Approximation as a "Sliding Law" in a Thermomechanically Coupled Ice Sheet Model", J. Geophy. Res., 114 F03008, https://doi.org/10.1029/2008JF001179, 2009.

Bueler, E., Lingle, C. S., and Brown, J.: Fast computation of a viscoelastic deformable Earth model for ice-sheet simulations, Ann. Glaciol., 46, 97–105, https://doi.org/10.3189/172756407782871567, 2007.

Burchard, H.: Applied turbulence modelling in marine waters, Springer Science and Business Media, Germany, 2002.

Bushell, A. C., Anstey, J. A., Butchart, N., Kawatani, Y., Osprey, S. M., Richter, J. H., Serva, F., Braesicke, P., Cagnazzo, C., Chen, C.-C., Chun, H.-Y., Garcia, R. R., Gray, L. J., Hamilton, K., Kerzenmacher, T., Kim, Y.-H., Lott, F., McLandress, C, Naoe, H., Scinocca, J., Smith, A. K., Stockdale, T. N., Versick, S., Watanabe, S., Yoshida, K., and Yukimoto, S.: Evaluation of the Quasi-Biennial Oscillation in global climate models for the SPARC QBO-initiative, Q. J. Roy. Meteor. Soc., 1–31, https://doi.org/10.1002/qj.3765, 2020.

CAM3.0: Description of the NCAR community Atmosphere Model (CAM3.0), NCAR technical note, NCAR/TN-464+STR, 102–104, June 2004.

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

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A. J., and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, in: Climate Change 2013: The Physical Sci-

ence Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1029–1136, https://doi.org/10.1017/CBO9781107415324.024, 2013.

Collins, W. J., Lamarque, J.-F., Schulz, M., Boucher, O., Eyring, V., Hegglin, M. I., Maycock, A., Myhre, G., Prather, M., Shindell, D., and Smith, S. J.: AerChemMIP: quantifying the effects of chemistry and aerosols in CMIP6, Geosci. Model Dev., 10, 585–607, https://doi.org/10.5194/gmd-10-585-2017, 2017.

Coon, M. D.: Oceanic and Atmospheric Boundary Layer Study, Washington Univ Seattle Arctic Ice Dynamics Joint Experiment Office, 1974.

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

Danabasoglu, G., Yeager, S. G., Kwon, Y.-O., Tribbia, J. J., Phillips, A. S., and Hurrell, J. W.: Variability of the Atlantic meridional overturning circulation in CCSM4, J. Climate, 25, 5153–5172, https://doi.org/10.1175/JCLI-D-11-00463.1, 2012.

Davini, P. and Cagnazzo, C.: On the misinterpretation of the North Atlantic Oscillation in CMIP5 models, Clim. Dynam., 43, 1497–1511, https://doi.org/10.1007/s00382-013-1970-y, 2014.

Davini, P. and D'Andrea, F.: From CMIP-3 to CMIP-6: Northern Hemisphere atmospheric blocking simulation in present and future climate, J. Climate, 33, 10021–10038, https://doi.org/10.1175/JCLI-D-19-0862.1, 2020.

Davini, P., von Hardenberg, J., Corti, S., Christensen, H. M., Juricke, S., Subramanian, A., Watson, P. A. G., Weisheimer, A., and Palmer, T. N.: Climate SPHINX: evaluating the impact of resolution and stochastic physics parameterisations in the EC-Earth global climate model, Geosci. Model Dev., 10, 1383–1402, https://doi.org/10.5194/gmd-10-1383-2017, 2017a.

Davini, P., Corti, S., D'Andrea, F., Rivière, G., and von Hardenberg, J.: Improved Winter European Atmospheric Blocking Frequencies in High-Resolution Global Climate Simulations, J. Adv. Model Earth. Sy., 9, 2615–2634, https://doi.org/10.1002/2017MS001082, 2017b.

de Lavergne, C., Vic, C., Madec, G., Roquet, F., Waterhouse, A. F., Whalen, C. B. Cuypers, Y., Bouruet-Aubertot, P., Ferron, B., and Hibiya, T.: A Parameterization of Local and Remote Tidal Mixing, J. Adv. Model Earth Sy., 12, e2020MS002065, https://doi.org/10.1029/2020MS002065, 2020.

Diamantakis, M. and Flemming, J.: Global mass fixer algorithms for conservative tracer transport in the ECMWF model, Geosci. Model Dev., 7, 965–979, https://doi.org/10.5194/gmd-7-965-2014, 2014.

Docquier, D., Massonnet, F., Barthélemy, A., Tandon, N. F., Lecomte, O., and Fichefet, T.: Relationships between Arctic sea ice drift and strength modelled by NEMO-LIM3.6, The Cryosphere, 11, 2829–2846, https://doi.org/10.5194/tc-11-2829-2017, 2017.

Dutra, E., Balsamo, G., Viterbo, P., Miranda, P. M., Beljaars, A., Schär, C., and Elder, K.: An improved snow scheme for the ECMWF land surface model: descrip-

tion and offline validation, J. Hydrometeorol., 11, 899–916, https://doi.org/10.1175/2010JHM1249.1, 2010.

Ebert, E. and Curry, J. A.: An intermediate one-dimensional thermodynamic sea ice model for investigating ice-atmosphere interactions, J. Geophys. Res.-Oceans, 98, 10085–10109, 1993.

EC-Earth consortium: EC-Earth – A European community Earth-System Model, http://ecearth.org, last access: December 2019a.

EC-Earth Consortium (EC-Earth): EC-Earth-Consortium EC-Earth3-Veg model output prepared for CMIP6 ScenarioMIP ssp119, Earth System Grid Federation [data set], https://doi.org/10.22033/ESGF/CMIP6.4872, 2019b.

Enfield, D. B. and Mayer, D. A.: Tropical Atlantic sea surface temperature variability and its relation to El Niño-Southern Oscillation, J. Geophys. Res., 102, 929–945, 1997.

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

Fereday, D., Maidens, A., Arribas, A., Scaife, A., and Knight, J.: Seasonal forecasts of Northern Hemisphere winter 2009/10, Environ. Res. Lett., 7, 034031, https://doi.org/10.1088/1748-9326/7/3/034031, 2012.

Fiedler, S., Stevens, B., Gidden, M., Smith, S. J., Riahi, K., and van Vuuren, D.: First forcing estimates from the future CMIP6 scenarios of anthropogenic aerosol optical properties and an associated Twomey effect, Geosci. Model Dev., 12, 989–1007, https://doi.org/10.5194/gmd-12-989-2019, 2019.

Fisher, R., Koven, C. D., Anderegg, W. R. L., Christoffersen, B. O., Dietze, M. C., Farrior, C., Holm, J. A., Hurtt, G., Knox, R. G., Lawrence, P. J., Longo, M., Matheny, A. M., Medvigy, D., Muller-Landau, H. C., Powell, T. L., Serbin, S. P., Sato, H., Shuman, J., Smith, B., Trugman, A. T., Viskari, T., Verbeeck, H., Weng, E., Xu, C., Xu, X., Zhang, T., and Moorcroft, P.: Vegetation demographics in Earth system models: a review of progress and priorities, Glob. Change Biol., 24, 35–54, 2018.

Flato, G. M.: Earth system models: an overview, Wires Clim. Change, 2, 783–800, 2011.

Forbes, R. M. and Ahlgrimm, M.: On the representation of high-latitude boundary layer mixed-phase cloud in the ECMWF global model, Mon. Weather Rev., 142, 3425–3445, https://doi.org/10.1175/MWR-D-13-00325.1, 2014.

Forbes, R., Geer, A., Lonitz, K., and Ahlgrimm, M.: Reducing systematic error in cold-air outbreaks, ECMWF Newsletter, 146, 17–22, 2016.

Fox-Kemper, B., Ferrari, R., and Hallberg, R.: Parameterization of Mixed Layer Eddies. Part I: Theory and Diagnosis, J. Phys. Oceanogr., 38, 1145–1165, https://doi.org/10.1175/2007JPO3792.1, 2008.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzjer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N.,

Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf, G. R., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2019, Earth Syst. Sci. Data, 11, 1783–1838, https://doi.org/10.5194/essd-11-1783-2019, 2019.

García-Serrano, J., Rodríguez-Fonseca, B., Bladé, I., Zurita-Gotor, P., and de La Cámara, A.: Rotational atmospheric circulation during North Atlantic-European winter: the influence of ENSO, Clim. Dynam., 37, 1727–1743, 2011.

Gaspar, P., Grégoris, Y., and Lefevre, J.-M.: A simple eddy kinetic energy model for simulations of the oceanic vertical mixing: Tests at station Papa and long-term upper ocean study site, J. Geophys. Res., 95, 16179–16193, https://doi.org/10.1029/JC095iC09p16179, 1990.

Gehne, M., Hamill, T. M., Kiladis, G. N., and Trenberth, K. E.: Comparison of Global Precipitation Estimates across a Range of Temporal and Spatial Scales, J. Climate, 29, 7773–7795, https://doi.org/10.1175/JCLI-D-15-0618.1, 2016.

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

Giannini, A., Kushnir, Y., and Cane, M. A.: Interannual variability of Caribbean rainfall, ENSO, and the Atlantic Ocean, J. Climate, 13, 297–311, 2000.

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., van Vuuren, D. P., van den Berg, M., Feng, L., Klein, D., Calvin, K., Doelman, J. C., Frank, S., Fricko, O., Harmsen, M., Hasegawa, T., Havlik, P., Hilaire, J., Hoesly, R., Horing, J., Popp, A., Stehfest, E., and Takahashi, K.: Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century, Geosci. Model Dev., 12, 1443–1475, https://doi.org/10.5194/gmd-12-1443-2019, 2019.

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

Grenfell, T. C. and Perovich, D. K.: Seasonal and spatial evolution of albedo in a snow-ice-land-ocean environment, J. Geophys. Res-Oceans, 109, C01001, https://doi.org/10.1029/2003JC001866, 2004.

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

Haarsma, R., Acosta, M., Bakhshi, R., Bretonnière, P.-A., Caron, L.-P., Castrillo, M., Corti, S., Davini, P., Exarchou, E., Fabiano, F., Fladrich, U., Fuentes Franco, R., García-Serrano, J., von Hardenberg, J., Koenigk, T., Levine, X., Meccia, V. L., van Noije,

T., van den Oord, G., Palmeiro, F. M., Rodrigo, M., Ruprich-Robert, Y., Le Sager, P., Tourigny, E., Wang, S., van Weele, M., and Wyser, K.: HighResMIP versions of EC-Earth: EC-Earth3P and EC-Earth3P-HR – description, model computational performance and basic validation, Geosci. Model Dev., 13, 3507–3527, https://doi.org/10.5194/gmd-13-3507-2020, 2020.

Hansen, J., Sato, M., Kharecha, P., and von Schuckmann, K.: Earth's energy imbalance and implications, Atmos. Chem. Phys., 11, 13421–13449, https://doi.org/10.5194/acp-11-13421-2011, 2011.

Hantson, S., Knorr, W., Schurgers, G., Pugh, T. A., and Arneth, A.: Global isoprene and monoterpene emissions under changing climate, vegetation, CO_2 and land use, Atmos. Environ, 155, 35–45, 2017.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, Int. J. Climatol., 34, 623–642, https://doi.org/10.1002/joc.3711, 2014.

Hazeleger, W., Wang, X., Severijns, C., Ştefănescu, S., Bintanja, R., Sterl, A., Wyser, K., Semmler, T., Yang, S., van den Hurk, B., van Noije, T., van der Linden, E., and van der Wiel, K.: EC-Earth V2. 2: description and validation of a new seamless earth system prediction model, Clim. Dynam., 39, 2611–2629, https://doi.org/10.1007/s00382-011-1228-5, 2012.

Hawkins, E., Ortega, P., Suckling, E., Schurer, A., Hegerl, G., Jones, P., Joshi, M., Osborn, T. J., Masson-Delmotte, V., Mignot, J., Thorne, P., and van Oldenborgh G. J.: Estimating Changes in Global Temperature since the Preindustrial Period, B. Am. Meteorol. Soc., 98, 1841–1856, https://doi.org/10.1175/BAMS-D-16-0007.1, 2017.

Hegglin, M. I., Kinnison, D., Plummer, D., et al.: Historical and future ozone database (1850–2100) in support of CMIP6, Geosci. Model. Dev. Discuss., in preparation, 2022.

Helsen, M. M., van de Wal, R. S. W., Reerink, T. J., Bintanja, R., Madsen, M. S., Yang, S., Li, Q., and Zhang, Q.: On the importance of the albedo parameterization for the mass balance of the Greenland ice sheet in EC-Earth, The Cryosphere, 11, 1949–1965, https://doi.org/10.5194/tc-11-1949-2017, 2017.

Heuzé, C.: North Atlantic deep water formation and AMOC in CMIP5 models, Ocean Sci., 13, 609–622, https://doi.org/10.5194/os-13-609-2017, 2017.

Hibler III, W. D.: A dynamic thermodynamic sea ice model, J. Phys. Oceanogr., 9, 815–846, 1979.

Hickler, T., Smith, B., Sykes, M. T., Davis, M. B., Sugita, S., and Walker, K.: Using a generalized vegetation model to simulate vegetation dynamics in the western Great Lakes region, USA, under alternative disturbance regimes, Ecology, 85, 519–530, 2004.

Hoegh-Guldberg, O., Jacob, D., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K., Engelbrecht, F., Guiot, J., Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S. I., Thomas, A., Warren, R., Zhou, G., Halim, S. A., Achlatis, M., Alexander, L. V., Allen, M., Berry, P., Boyer, C., Byers, E., Brilli, L., Buckeridge, M., Cheung, W., Craig, M., Ellis, N., Evans, J., Fischer, H., Fraedrich, K., Fuss, S., Ganase, A., Gattuso, J. P., Greve, P., Bolaños, T. G., Hanasak, N., Hasegawa, T., Hayes, K., Hirsch, A., Jones, C., Jung, T., Kanninen, M., Krinner, G., Lawrence, D., Lenton, T., Ley, D., Liverman, D., Mahowald, N., McInnes, K., Meissner, K. J., Millar, R., Mintenbeck, K., Mitchell, D., Mix, A. C., Notz, D., Nurse, L., Okem, A., Olsson, L., Oppenheimer, M.,

Paz, S., Petersen, J., Petzold, J., Preuschmann, S., Rahman, M. F., Rogelj, J., Scheuffele, H,. Schleussner. C.-F., Scott, D., Séférian, R., Sillmann, J., Singh, C., Slade, R., Stephenson, K., Stephenson, T., Sylla, M. B., Tebboth, M., Tschakert, P., Vautard, R., Wartenburger, R., Wehner, M., Weyer, N. M., Whyte, F., Yohe, G., Zhang, X., and Zougmoré, R. B.: Impacts of 1.5 °C Global Warming on Natural and Human Systems, in: Global warming of 1.5 °C.: An IPCC Special Report, edited by: Masson-Delmotte, V., Zhai, P., Pörtner, H. O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J. B. R., Chen, Y., Zhou, X., Gomis, M. I., Lonnoy, E., Maycock, T., Tignor, M., and Waterfield, T., IPCC Secretariat, 175–311, 2018.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-I., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geosci. Model Dev., 11, 369–408, https://doi.org/10.5194/gmd-11-369-2018, 2018.

Hogan, R., Ahlgrimm, M., Balsamo, G., Beljaars, A., Berrisford, P., Bozzo, A., Di Giuseppe, F., Forbes, R. M., Haiden, T., Lang, S., Mayer, M., Polichtchouk, I., Sandu, I., Vitart, F., and Wedi, N.: Radiation in numerical weather prediction, ECMWF Technical Memorandum No. 816, 49 pp., https://doi.org/10.21957/2bd5dkj8x, 2017.

Huang, B., Xue, Y., Kumar, A., and Behringer, D. W.: AMOC variations in 1979–2008 simulated by NCEP operational ocean data assimilation system, Clim. Dynam., 38 513, https://doi.org/10.1007/s00382-011-1035-z, 2012.

Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M.: An overview of the North Atlantic oscillation, Geophys. Monogr., American Geophysical Union, 134, 1–36, 2003.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J., Fisk, J., Fujimori, S., Goldewijk, K. K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J., Krisztin, T., Lawrence, D., Lawrence, P., Mertz, O., Pongratz, J., Popp, A., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., van Vuuren, D., and Zhang, X.: Harmonization of Global Land Use Change and Management for the Period 850-2015, Version 20190529, Earth System Grid Federation, https://doi.org/10.22033/ESGF/input4MIPs.10454, 2019a.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J., Fisk, J., Fujimori, S., Goldewijk, K. K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J., Krisztin, T., Lawrence, D., Lawrence, P., Mertz, O., Pongratz, J., Popp, A., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., van Vuuren, D., and Zhang, X.: Harmonization of Global Land Use Change and Management for the Period 2015–2300, Version 20190529, Earth System Grid Federation, https://doi.org/10.22033/ESGF/input4MIPs.10468, 2019b.

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., and

Zhang, X.: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6, Geosci. Model Dev., 13, 5425–5464, https://doi.org/10.5194/gmd-13-5425-2020, 2020.

Hwang, Y. T. and Frierson, D. M.: Link between the double-Intertropical Convergence Zone problem and cloud biases over the Southern Ocean, P. Natl. Acad. Sci. USA, 110, 4935–4940, 2013.

Hyder, P., Edwards, J. M., Allan, R. P., Hewitt, H. T., Bracegirdle, T. J., Gregory, J. M., Wood, R. A., Meijers, A. J., Mulcahy, J., Field, P., Furtado, K., Bodas-Salcedo, A., Williams, K. D., Copsey,D., Josey, S. A., Liu, C., Roberts, C. D., Sanchez, C., Ridley, J., Thorpe, L., Hardiman, S. C., Mayer, M., Berry, D. I., ansd Belcher, S. E.: Critical Southern Ocean climate model biases traced to atmospheric model cloud errors, Nat. Commun., 9, 3625, https://doi.org/10.1038/s41467-018-05634-2, 2018.

Jha, B., Hu, Z., and Kumar, A.: SST and ENSO variability and change simulated in historical experiments of CMIP5 models, Clim. Dynam., 42, 2113–2124, https://doi.org/10.1007/s00382-013-1803-z, 2014.

Jiménez-Esteve, B. and Domeisen, D.: Nonlinearity in the North Pacific atmospheric response to a linear ENSO forcing, Geophys. Res. Lett., 46, 2271–2281, 2019.

Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., Graven, H., Hoffman, F., Ilyina, T., John, J. G., Jung, M., Kawamiya, M., Koven, C., Pongratz, J., Raddatz, T., Randerson, J. T., and Zaehle, S.: C4MIP – The Coupled Climate–Carbon Cycle Model Intercomparison Project: experimental protocol for CMIP6, Geosci. Model Dev., 9, 2853–2880, https://doi.org/10.5194/gmd-9-2853-2016, 2016.

Jungclaus, J. H., Haak, H., Latif, M., and Mikolajewicz, U.: Arctic–North Atlantic Interactions and Multidecadal Variability of the Meridional Overturning Circulation, J. Climate, 18, 4013–4031, https://doi.org/10.1175/JCLI3462.1, 2005.

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

Kawamiya, M., Hajima, T., Tachiiri, K., Watanabe, S., and Yokohata, T.: Two decades of Earth system modeling with an emphasis on Model for Interdisciplinary Research on Climate (MIROC), Progr. Earth Planet. Sci., 7, 1–13, 2020.

Keen, A., Blockley, E., Bailey, D. A., Boldingh Debernard, J., Bushuk, M., Delhaye, S., Docquier, D., Feltham, D., Massonnet, F., O'Farrell, S., Ponsoni, L., Rodriguez, J. M., Schroeder, D., Swart, N., Toyoda, T., Tsujino, H., Vancoppenolle, M., and Wyser, K.: An inter-comparison of the mass budget of the Arctic sea ice in CMIP6 models, The Cryosphere, 15, 951–982, https://doi.org/10.5194/tc-15-951-2021, 2021.

Kennedy, J., Titchner, H., Rayner, N., and Roberts, M.: Input4mips. MOHC. SSTsandseaice. Highresmip MOHC-hadisst-2-2-0-0-0, Earth System Grid Federation, https://doi.org/10.22033/ESGF/input4mips.1221, 2017.

Klaver, R., Haarsma, R., Vidale, P. L., and Hazeleger, W.: Effective resolution in high resolution global atmospheric models for climate studies, Atmos. Sci. Lett., 21, 1–8, https://doi.org/10.1002/asl.952, 2020.

Koenigk, T., Brodeau, L., Graversen, R. G., Karlsson, J., Svensson, G., Tjernström, M., Willén, U., and Wyser, K.: Arctic climate change in 21st century CMIP5 simulations with EC-Earth, Clim. Dynam., 40, 2719–2743, 2013.

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, Atmos. Chem. Phys., 5, 417–432, https://doi.org/10.5194/acp-5-417-2005, 2005.

Landschützer, P., Gruber, N. and Bakker, D. C. E.: Decadal variations and trends of the global ocean carbon sink, Global Biogeochem. Cy., 30, 1396–1417, https://doi.org/10.1002/2015GB005359, 2016.

Lavergne, T., Sørensen, A. M., Kern, S., Tonboe, R., Notz, D., Aaboe, S., Bell, L., Dybkjær, G., Eastwood, S., Gabarro, C., Heygster, G., Killie, M. A., Brandt Kreiner, M., Lavelle, J., Saldo, R., Sandven, S., and Pedersen, L. T.: Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records, The Cryosphere, 13, 49–78, https://doi.org/10.5194/tc-13-49-2019, 2019.

Le Quéré, C., Moriarty, R., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Friedlingstein, P., Peters, G. P., Andres, R. J., Boden, T. A., Houghton, R. A., House, J. I., Keeling, R. F., Tans, P., Arneth, A., Bakker, D. C. E., Barbero, L., Bopp, L., Chang, J., Chevallier, F., Chini, L. P., Ciais, P., Fader, M., Feely, R. A., Gkritzalis, T., Harris, I., Hauck, J., Ilyina, T., Jain, A. K., Kato, E., Kitidis, V., Klein Goldewijk, K., Koven, C., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lima, I. D., Metzl, N., Millero, F., Munro, D. R., Murata, A., Nabel, J. E. M. S., Nakaoka, S., Nojiri, Y., O'Brien, K., Olsen, A., Ono, T., Pérez, F. F., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Rödenbeck, C., Saito, S., Schuster, U., Schwinger, J., Séférian, R., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Vandemark, D., Viovy, N., Wiltshire, A., Zaehle, S., and Zeng, N.: Global Carbon Budget 2015, Earth Syst. Sci. Data, 7, 349–396, https://doi.org/10.5194/essd-7-349-2015, 2015.

Leutbecher, M., Lock, S. J., Ollinaho, P., Lang, S. T., Balsamo, G., Bechtold, P., Bonavita, M., Christensen, H. M., Diamantakis, M., Dutra, E., and English, S.: Stochastic representations of model uncertainties at ECMWF: state of the art and future vision, Q. J. Roy. Meteor. Soc., 143, 2315–2339, 2017.

Lindeskog, M., Arneth, A., Bondeau, A., Waha, K., Seaquist, J., Olin, S., and Smith, B.: Implications of accounting for land use in simulations of ecosystem carbon cycling in Africa, Earth Syst. Dynam., 4, 385–407, https://doi.org/10.5194/esd-4-385-2013, 2013.

Lliboutry, L. A. and Duval, P.: Various isotropic and anisotropic ices found in glaciers and polar ice caps and their corresponding rheologies, Ann. Geophys., 3, 207–224, 1985.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Liang, L., Mitrescu, C., Rose, F. G., and Kato, S.: Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product, J. Climate, 31, 895–918, https://doi.org/10.1175/JCLI-D-17-0208.1, 2018.

López-Parages, J., Rodríguez-Fonseca, B., Dommenget, D. and Frauen, C.: ENSO influence on the North Atlantic European climate: a non-linear and non-stationary approach, Clim. Dynam., 47, 2071–2084, 2016.

Madec, G.: NEMO ocean engine, Note du Pole de modelisation de l'Institut Pierre-Simon Laplace No 27, ISSN No 1288-1619, 2015.

Madec, G. and Imbard, M.: A global ocean mesh to overcome the North Pole singularity, Clim. Dynam., 12, 381–388, https://doi.org/10.1007/BF00211684, 1996.

Madec, G. and the NEMO team: NEMO ocean engine, Note du Pôle de modélisation, Institut Pierre-Simon Laplace (IPSL), France, No 27, ISSN No 1288-1619, 2008.

Marsaleix, P., Auclair, F., Floor, J. W., Herrmann, M. J., Estournel, C., Pairaud, I., and Ulses, C.: Energy conservation issues in sigma-coordinate free-surface ocean models, Ocean Model., 20, 61–89, 2008.

Masato, G., Hoskins, B. J., and Woollings, T.: Winter and summer Northern Hemisphere blocking in CMIP5 models, J. Climate, 26, 7044–7059, 2013.

Massonnet, F., Ménégoz, M., Acosta, M. C., Yepes-Arbós, X., Exarchou, E., and Doblas-Reyes, F. J.: Reproducibility of an Earth System Model under a change in computing environment, Technical Report, Barcelona Supercomputing Center, Spain, 2018.

Massonnet, F., Reid, P., Lieser, J. L., Bitz, C. M., Fyfe, J., and Hobbs, W. R.: Assessment of summer 2018–2019 sea-ice forecasts for the Southern Ocean, Antarctic Climate & Ecosystems Cooperative Research Centre, University of Tasmania, Hobart (Australia), https://doi.org/10.25959/100.00029984, 2019.

Massonnet, F., Ménégoz, M., Acosta, M., Yepes-Arbós, X., Exarchou, E., and Doblas-Reyes, F. J.: Replicability of the EC-Earth3 Earth system model under a change in computing environment, Geosci. Model Dev., 13, 1165–1178, https://doi.org/10.5194/gmd-13-1165-2020, 2020.

Matthes, K., Funke, B., Andersson, M. E., Barnard, L., Beer, J., Charbonneau, P., Clilverd, M. A., Dudok de Wit, T., Haberreiter, M., Hendry, A., Jackman, C. H., Kretzschmar, M., Kruschke, T., Kunze, M., Langematz, U., Marsh, D. R., Maycock, A. C., Misios, S., Rodger, C. J., Scaife, A. A., Seppälä, A., Shangguan, M., Sinnhuber, M., Tourpali, K., Usoskin, I., van de Kamp, M., Verronen, P. T., and Versick, S.: Solar forcing for CMIP6 (v3.2), Geosci. Model Dev., 10, 2247–2302, https://doi.org/10.5194/gmd-10-2247-2017, 2017.

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., Haak, H., Jungclaus, J., Klocke, D., Matei, D., Mikolajewicz, U., Notz, D., Pincus, R., Schmidt, H., and Tomassini, L.: Tuning the climate of a global model, J. Adv. Model. Earth Sy., 4, M00A01, https://doi.org/10.1029/2012MS000154, 2012.

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., and Weiss, R.: Historical greenhouse gas concentrations for cli-

mate modelling (CMIP6), Geosci. Model Dev., 10, 2057–2116, https://doi.org/10.5194/gmd-10-2057-2017, 2017.

Mellor, G. and Blumberg, A.: Wave Breaking and Ocean Surface Layer Thermal Response, J. Phys. Oceanogr., 34, 693–698, https://doi.org/10.1175/2517.1, 2004.

Meredith, M., Sommerkorn, M., Cassotta, S., Derksen, C., Ekaykin, A., Hollowed, A., Kofinas, G., Mackintosh, A., Melbourne-Thomas, J., Muelbert, M. M. C., Ottersen, G., Pritchard, H., and Schuur, E. A. G.: Polar Regions. Chapter 3, IPCC Special Report on the Ocean and Cryosphere in a Changing Climate, 2019.

Monahan, E. C., Spiel, D. E., and Davidson, K. L.: A Model of Marine Aerosol Generation Via Whitecaps and Wave Disruption, in: Oceanic Whitecaps, edited by: Monahan, E. C. and Niocaill, G. M., Oceanographic Sciences Library, vol 2. Springer, Dordrecht, https://doi.org/10.1007/978-94-009-4668-2_16, 1986.

Morcrette, J.-J., Barker, H. W., Cole, J. N. S., Iacono, M. J., and Pincus, R.: Impact of a New Radiation Package, McRad, in the ECMWF Integrated Forecasting System, Mon. Weather Rev., 136, 4773–4798, 2008.

Morlighem, M., Bondzio, J., Seroussi, H., Rignot, E., Larour, E., Humbert, A., and Rebuffi, S.: Modelling of Store Gletscher's calving dynamics, West Greenland, in response to ocean thermal forcing, Geophys. Res. Lett., 43, 2659–2666, https://doi.org/10.1002/2016GL067695, 2016.

Nowicki, S. M. J., Payne, A., Larour, E., Seroussi, H., Goelzer, H., Lipscomb, W., Gregory, J., Abe-Ouchi, A., and Shepherd, A.: Ice Sheet Model Intercomparison Project (ISMIP6) contribution to CMIP6, Geosci. Model Dev., 9, 4521–4545, https://doi.org/10.5194/gmd-9-4521-2016, 2016.

Olin, S., Lindeskog, M., Pugh, T. A. M., Schurgers, G., Wårlind, D., Mishurov, M., Zaehle, S., Stocker, B. D., Smith, B., and Arneth, A.: Soil carbon management in large-scale Earth system modelling: implications for crop yields and nitrogen leaching, Earth Syst. Dynam., 6, 745–768, https://doi.org/10.5194/esd-6-745-2015, 2015a.

Olin, S., Schurgers, G., Lindeskog, M., Wårlind, D., Smith, B., Bodin, P., Holmér, J., and Arneth, A.: Modelling the response of yields and tissue C : N to changes in atmospheric CO_2 and N management in the main wheat regions of western Europe, Biogeosciences, 12, 2489–2515, https://doi.org/10.5194/bg-12-2489-2015, 2015b.

Ono, K.: Temperature dependence of dispersed barrier hardening, J. Appl., Phys., 39, 1803–1806, 1968.

Peters, W., van der Velde, I. R., Van Schaik, E., Miller, J. B., Ciais, P., Duarte, H. F., van der Laan-Luijkx, I. T., van der Molen, M. K., Scholze, M., Schaefer, K., Vidale, P. L., Verhoef, A., Wårlind, D., Zhu, D., Tans, P. P., Vaughn, B., and White J. W. C.: Increased water-use efficiency and reduced CO_2 uptake by plants during droughts at a continental scale, Nat Geosci., 11, 744–748, https://doi.org/10.1038/s41561-018-0212-7, 2018.

Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., Ahlström, A., Anav, A., Canadell, J. G., Cong, N., Huntingford, C., Jung, M., Levis, S., Levy, P. E., Li, J., Lin, X., Lomas, M. R., Lu, M., Luo, Y., Ma, Y., Myneni, R. B., Poulter, B., Sun, Z., Wang, T., Viovy, N., Zaehle, S., and Zeng, N.: Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO_2 trends, Glob. Change Biol., 19, 2117–2132, 2013.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G., Peubey, C., Thépaut, J., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An Atmospheric Reanalysis of the Twentieth Century, J. Climate, 29, 4083–4097, https://doi.org/10.1175/JCLI-D-15-0556.1, 2016.

Prather, M. J.: Numerical advection by conservation of second-order moments, J. Geophys. Res.-Atmos., 91, 6671–6681, 1986.

Pringle, D. J., Eicken, H., Trodahl, H. J., and Backstrom, L. G. E.: Thermal conductivity of landfast Antarctic and Arctic sea ice, J. Geophys. Res.-Oceans., 112, C04017, https://doi.org/10.1029/2006JC003641, 2007.

Pugh, T. A. M., Jones, C. D., Huntingford, C., Burton, C., Arneth, A., Brovkin, V., Ciais, P., Lomas, M., Robertson, E., Piao, S. L., and Sitch, S.: A large committed long-term sink of carbon due to vegetation dynamics, Earth's Future, 6, 1413–1432, 2018.

Purves, D. and Pacala, S.: Predictive models of forest dynamics, Science, 320, 1452–1453, 2008.

Rasch, P. J. and Williamson, D. L.: Computational aspects of moisture transport in global models of the atmosphere, Q. J. Roy. Meteor. Soc., 116, 1071–1090, https://doi.org/10.1002/qj.49711649504, 1990.

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

Rechid, D., Raddatz, T. J., and Jacob, D.: Parameterization of snow-free land surface albedo as a function of vegetation phenology based on MODIS data and applied in climate modelling, Theor. Appl. Climatol., 95, 245–255, https://doi.org/10.1007/s00704-008-0003-y, 2009.

Reichler, T. and Kim, J.: How Well Do Coupled Models Simulate Today's Climate?, B. Am. Meteorol. Soc., 89, 303–312, https://doi.org/10.1175/BAMS-89-3-303, 2008.

Reichler, T., Dameris, M., and Sausen, R.: Determining the tropopause height from gridded data, Geophys. Res. Lett., 30, 2042, https://doi.org/10.1029/2003GL01824, 2003.

Riahi, K., Van Vuuren, D. P., Kriegler, E., Edmonds, J., O'neill, B. C., Fujimori, S., Bauer, N., Calvin, K., Dellink, R., Fricko, O., Lutz, W., Popp, A., Cuaresma, J. C., KC, S., Leimbach, M., Jiang, L., Kram, T., Rao, S., Emmerling, J., Ebi, K., Hasegawa, T., Havlik, P., Humpenöder, F., Aleluia, Da Silva, L. A., Smith, S., Stehfest, E., Bosetti, V., Eom, J., Gernaat, D., Masui, T., Rogelj, J., Strefler, J., Drouet, L., Kery, V., Luderer, G., Harmsen, M., Takahashi, K., Baumstark, L., Doelman, J. C., Kainuma, M., Klimont, Z., Marangoni, G., Lotze-Campen, H., Obersteiner, M., Tabeau, A., and Tavoni, M.: The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview, Global Environ. Chang., 42, 153–168, 2017.

Richter, J. H., Anstey, J. A., Butchart, N., Kawatani, Y., Meehl, G. A., Osprey, S. and Simpson I. R.: Richter, J. H., Anstey, J. A., Butchart, N., Kawatani, Y., Meehl, G. A., Osprey, S., and Simpson, I. R.: Progress in simulating the quasi-biennial oscillation in CMIP models, J. Geophys. Res.-Atmos., 125, e2019JD032362, https://doi.org/10.1029/2019JD032362, 2020.

Roberts, M. J., Vidale, P. L., Senior, C., Hewitt, H. T., Bates, C., Berthou, S., Chang, P., Christensen, H. M., Danilov, S., Demory, M. E., and Griffies, S. M.: The benefits of global high resolution for climate simulation: process understanding and the enabling of stakeholder decisions at the regional scale, B. Am. Meteorol. Soc., 99, 2341–2359, 2018.

Roberts, M. J., Camp, J., Seddon, J., Vidale, P. L., Hodges, K., Vanniere, B., Mecking, J., Haarsma, R., Bellucci, A., Scoccimarro, E., Caron, L., Chauvin, F., Terray, L., Valcke, S., Moine, M., Putrasahan, D., Roberts, C., Senan, R., Zarzycki, C., and Ullrich, P.: Impact of Model Resolution on Tropical Cyclone Simulation Using the HighResMIP–PRIMAVERA Multimodel Ensemble, J. Climate, 33, 2557–2583, https://doi.org/10.1175/JCLI-D-19-0639.1, 2020.

Rotstayn, L. D.: On the "tuning" of autoconversion parameterizations in climate models, J. Geophys. Res., 105, 15495–15507, https://doi.org/10.1029/2000JD900129, 2000.

Rousset, C., Vancoppenolle, M., Madec, G., Fichefet, T., Flavoni, S., Barthélemy, A., Benshila, R., Chanut, J., Levy, C., Masson, S., and Vivier, F.: The Louvain-La-Neuve sea ice model LIM3.6: global and regional capabilities, Geosci. Model Dev., 8, 2991–3005, https://doi.org/10.5194/gmd-8-2991-2015, 2015.

Salisbury, D. J., Anguelova, M. D., and Brooks, I. M.: On the variability of whitecap fraction using satellite-based observations, J. Geophys. Res.-Oceans, 118, 6201–6222, 2013.

Salter, M. E., Zieger, P., Acosta Navarro, J. C., Grythe, H., Kirkevåg, A., Rosati, B., Riipinen, I., and Nilsson, E. D.: An empirically derived inorganic sea spray source function incoporating sea surface temperature, Atmos. Chem. Phys., 15, 11047–11066, https://doi.org/10.5194/acp-15-11047-2015, 2015.

Schaller, N., Sillmann, J., Anstey, J., Fischer, E. M., Grams, C. M., and Russo, S.: Influence of blocking on northern european and western russian heatwaves in large climate model ensembles, Environ. Res. Lett., 13, 054015, https://doi.org/10.1088/1748-9326/aaba55, 2018.

Schweiger, A., Lindsay, R., Zhang, J., Steele, M., Stern, H., and Kwok, R.: Uncertainty in modeled Arctic sea ice volume, J. Geophys. Res., 116, C00D06, https://doi.org/10.1029/2011JC007084, 2011.

Seneviratne, S. I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A., Cheruy, F., Higgins, M. E., Meier, A., Brovkin, V., Claussen, M., Ducharne, A., Dufrene, J.-L., Findell, K. L., Ghattas, J., Lawrence, D. M., Malyshev, S., Rummukainen, M., and Smith, B.: Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment, Geophys. Res. Lett., 40, 5212–5217, 2013.

Shine, K. P. and Henderson-Sellers, A.: The sensitivity of a thermodynamic sea ice model to changes in surface albedo parameterization, J. Geophys. Res.-Atmos., 90, 2243–2250, 1985.

Sillmann, J., Croci-Maspoli, M., Kallache, M., and Katz, R. W.: Extreme Cold Winter Temperatures in Europe under the Influence of North Atlantic Atmospheric Blocking, J. Climate, 24, 5899–5913, https://doi.org/10.1175/2011JCLI4075.1, 2011.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J. G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S. L., Le Quéré, C., Smith, B.,

Zhu, Z., and Myneni, R.: Recent trends and drivers of regional sources and sinks of carbon dioxide, Biogeosciences, 12, 653–679, https://doi.org/10.5194/bg-12-653-2015, 2015.

Smeed, D. A., Josey, S. A., Beaulieu, C., Johns, W. E., Moat, B. I., Frajka-Williams, E., and McCarthy, G. D.: The North Atlantic Ocean is in a state of reduced overturning, Geophys. Res. Lett., 45, 1527–1533, 2018.

Smith, B., Prentice, I. C., and Sykes, M. T.: Representation of vegetation dynamics in modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space, Global. Ecol. Biogeogr., 10, 621–637, 2001.

Smith, A. K, Holt, L. A., Garcia, R. R., Anstey, J. A., Serva, F., Buthard, N., Osprey, S., Bushell, A. C., Kawatani, Y., Kim, Y.-H., Lott, F., Braesicke, P., Cagnazzo, C., Chen, C.-C. Chun, H.-Y., Gray, L., Kerzenmacher, T., Naoe, H., Richter, J., Versick, S., Schenzinger, V., Watanabe, S., and Yoshida, K.: The equatorial stratospheric semiannual oscillation and time-mean winds in QBOi models, Q. J. Roy. Meteor. Soc., 2020, 1–17, https://doi.org/10.1002/qj.3690, 2019.

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., and Zaehle, S.: Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model, Biogeosciences, 11, 2027–2054, https://doi.org/10.5194/bg-11-2027-2014, 2014.

Stevens, B., Fiedler, S., Kinne, S., Peters, K., Rast, S., Müsse, J., Smith, S. J., and Mauritsen, T.: MACv2-SP: a parameterization of anthropogenic aerosol optical properties and an associated Twomey effect for use in CMIP6, Geosci. Model Dev., 10, 433–452, https://doi.org/10.5194/gmd-10-433-2017, 2017.

Tagliabue, A., and Arrigo, K. R.: Processes governing the supply of iron to phytoplankton in stratified seas, J. Geophys. Res.-Oceans., 111, C06019, https://doi.org/10.1029/2005JC003363, 2006.

The PISM Team: PISM, A Parallel Ice Sheet Model, available at: http://www.pism-docs.org (last access: 30 December 2020), 2019.

Thomason, L. W., Ernest, N., Millán, L., Rieger, L., Bourassa, A., Vernier, J.-P., Manney, G., Luo, B., Arfeuille, F., and Peter, T.: A global space-based stratospheric aerosol climatology: 1979–2016, Earth Syst. Sci. Data, 10, 469–492, https://doi.org/10.5194/essd-10-469-2018, 2018.

Thorndike, A. S., Rothrock, D. A., Maykut, G. A., and Colony, R.: The thickness distribution of sea ice, J. Geophys. Res., 80, 4501–4513, https://doi.org/10.1029/JC080i033p04501, 1975.

Toniazzo, T. and Scaife, A.: The effect of non-linearity on winter ENSO teleconnections over Europe, Geophys. Res. Lett., 33, L24704, https://doi.org/10.1029/2006GL027881, 2006.

Trenberth, K. E., Zhang, Y., Fasullo, J. T., and Cheng, L.: Observation-Based Estimates of Global and Basin Ocean Meridional Heat Transport Time Series, J. Climate, 32, 4567–4583, 2019.

Vancoppenolle, M., Fichefet, T., Goosse, H., Bouillon, S., Madec, G., and Maqueda, M. A. M.: Simulating the mass balance and salinity of Arctic and Antarctic sea ice. 1. Model description and validation, Ocean Model., 27, 33–53, https://doi.org/10.1016/j.ocemod.2008.10.005, 2009.

van den Hurk, B. J., Viterbo, P., Beljaars, A. C. M., and Betts, A. K.: Offline validation of the ERA40 surface scheme, Reading, UK, ECMWF, p. 43, https://doi.org/10.21957/9aoaspz8, 2000.

van den Oord, G.: ece2cmor3 (0.9), Zenodo [code], https://doi.org/10.5281/zenodo.1051094, 2017.

van Den Oord, A. and Oriol, V.: Neural discrete representation learning, Adv. Neur. In., 30, 6306–6315, 2017.

van den Oord, G., Reerink, T., Bergman, T., Anthoni, T., Nieradzik, L., Tourigny, T., Fladrich, U., Le Sager, P., Wyser, K., and van Noije, T.: The ece2cmor3 package for post-processing EC-Earth3 climate model output for CMIP6, Geosci. Model. Dev. Discuss., in preparation, 2022.

van Marle, M. J. E., Kloster, S., Magi, B. I., Marlon, J. R., Daniau, A.-L., Field, R. D., Arneth, A., Forrest, M., Hantson, S., Kehrwald, N. M., Knorr, W., Lasslop, G., Li, F., Mangeon, S., Yue, C., Kaiser, J. W., and van der Werf, G. R.: Historic global biomass burning emissions for CMIP6 (BB4CMIP) based on merging satellite observations with proxies and fire models (1750–2015), Geosci. Model Dev., 10, 3329–3357, https://doi.org/10.5194/gmd-10-3329-2017, 2017.

van Noije, T. P. C., Le Sager, P., Segers, A. J., van Velthoven, P. F. J., Krol, M. C., Hazeleger, W., Williams, A. G., and Chambers, S. D.: Simulation of tropospheric chemistry and aerosols with the climate model EC-Earth, Geosci. Model Dev., 7, 2435–2475, https://doi.org/10.5194/gmd-7-2435-2014, 2014.

van Noije, T., Bergman, T., Le Sager, P., O'Donnell, D., Makkonen, R., Gonçalves-Ageitos, M., Döscher, R., Fladrich, U., von Hardenberg, J., Keskinen, J.-P., Korhonen, H., Laakso, A., Myriokefalitakis, S., Ollinaho, P., Pérez García-Pando, C., Reerink, T., Schrödner, R., Wyser, K., and Yang, S.: EC-Earth3-AerChem: a global climate model with interactive aerosols and atmospheric chemistry participating in CMIP6 , Geosci. Model Dev., 14, 5637–5668, https://doi.org/10.5194/gmd-14-5637-2021, 2021.

Vignati, E., Wilson, J., and Stier, P.: M7: An efficient size-resolved aerosol microphysics module for large-scale aerosol transport models, J. Geophys. Res., 109, D22202, https://doi.org/10.1029/2003JD004485, 2004.

Wårlind, D., Smith, B., Hickler, T., and Arneth, A.: Nitrogen feedbacks increase future terrestrial ecosystem carbon uptake in an individual-based dynamic vegetation model, Biogeosciences, 11, 6131–6146, https://doi.org/10.5194/bg-11-6131-2014, 2014.

Wallace, J. M. and Gutzler, D. S.: Teleconnections in the geopotential height field during the Northern Hemisphere winter, Mon. Weather Rev., 109, 784–812, 1981.

Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean, J. Geophys. Res., 97, 7373–7382, 1992.

Weiss, M., van den Hurk, B., Haarsma, R. and Hazeleger, W.: Impact of vegetation variability on potential predictability and skill of EC-Earth simulations, Clim. Dynam., 39, 2733–2746, https://doi.org/10.1007/s00382-012-1572-0, 2012.

Weiss, M., Miller, P. A., van den Hurk, B. J. J. M., van Noije, T., Ştefănescu, S., Haarsma, R., van Ulft, L. H., Hazeleger, W., Le Sager, P., Smith, B., and Schurgers, G.: Contribution of Dynamic Vegetation Phenology to Decadal Climate Predictability, J. Climate, 27, 8563–8577, 2014.

Williams, J. E., Boersma, K. F., Le Sager, P., and Verstraeten, W. W.: The high-resolution version of TM5-MP for optimized satellite retrievals: description and validation, Geosci. Model Dev., 10, 721–750, https://doi.org/10.5194/gmd-10-721-2017, 2017.

Winkelmann, R., Martin, M. A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK) – Part 1: Model description, The Cryosphere, 5, 715–726, https://doi.org/10.5194/tc-5-715-2011, 2011.

WMO: World Meteorological Organization (WMO), Meteorology – A three-dimensional Science: Second Session of the Commission for Aerology, WMO, Bulletin VI(4), Geneva, 134–138, 1957.

Wolf, A., Ciais, P., Bellassen, V., Delbart, N., Field, C. B., and Berry, J. A.: Forest biomass allometry in global land surface models, Global Biogeochem. Cy., 25, GB3015, https://doi.org/10.1029/2010GB003917, 2011.

Woollings, T., Barriopedro, D., Methven, J., Son, S.-W., Martius, O., Harvey, B., Sillmann, J., Lupo, A. R., and Seneviratne, S.: Blocking and its Response to Climate Change, Curr. Clim. Change Rep., 4, 287–300, https://doi.org/10.1007/s40641-018-0108-z, 2018.

Wyser, K., van Noije, T., Yang, S., von Hardenberg, J., O'Donnell, D., and Döscher, R.: On the increased climate sensitivity in the EC-Earth model from CMIP5 to CMIP6, Geosci. Model Dev., 13, 3465–3474, https://doi.org/10.5194/gmd-13-3465-2020, 2020a.

Wyser, K., Kjellström, E., Koenigk, T., Martins, H., and Döscher, R.: Warmer climate projections in EC-Earth3-Veg: the role of changes in the greenhouse gas concentrations from CMIP5 to CMIP6, Environ. Res. Lett., 15, 054020, https://doi.org/10.1088/1748-9326/ab81c2, 2020b.

Yeager, S. and Danabasoglu, G.: The Origins of Late-Twentieth-Century Variations in the Large-Scale North Atlantic Circulation, J. Climate, 27, 3222–3247, https://doi.org/10.1175/JCLI-D-13-00125.1, 2014.

Zaehle, S., Medlyn, B. E., De Kauwe, M. G., Walker, A. P., Dietze, M. C., Hickler, T., Luo, Y., Wang, Y.-P., El-Masri, B., Thornton, P., Jain, A., Wang, S., Wårlind, D., Weng, E., Parton, W., Iversen, C. M., Gallet-Budynek, A., McCarthy, H., Finzi, A., Hanson, P. J., Prentice, I. C., Oren, R., and Norby, R. J.: Evaluation of 11 terrestrial carbon-nitrogen cycle models against observations from two temperate Free-Air CO_2 Enrichment studies, New Phytol., 202, 803–822, 2014.