The Tuning Strategy of IPSL-CM6A-LR

Juliette Mignot, Frédéric Hourdin, Julie Deshayes, Olivier Boucher, Guillaume Gastineau, Ionela Musat, Martin Vancoppenolle, Jérôme Servonnat, Arnaud Caubel, Frederique Cheruy, et al.

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2020MS002340

The Tuning Strategy of IPSL-CM6A-LR

Juliette Mignot[1], Frédéric Hourdin[2], Julie Deshayes[1], Olivier Boucher[3], Guillaume Gastineau[1], Ionela Musat[2], Martin Vancoppenolle[1], Jérôme Servonnat[4], Arnaud Caubel[4], Frédérique Chéruy[2], Sébastien Denvil[3,5], Jean-Louis Dufresne[2], Christian Ethé[3], Laurent Fairhead[2], Marie-Alice Foujols[3], Jean-Yves Grandpeix[2], Guillaume Levavasseur[3], Olivier Marti[4], Matthew Menary[1], Catherine Rio[6], Clément Rousset[1], and Yona Silvy[1]

[1]Laboratoire d'Océanographie et du Climat: Expérimentations et Approches Numériques, IRD, Sorbonne Université / CNRS, MNHN, Paris, France, [2]Laboratoire de Météorologie Dynamique, Sorbonne Université / CNRS / École Normale Supérieure, Paris, France, [3]Institut Pierre-Simon Laplace, Sorbonne Université / CNRS, Paris, France, [4]Laboratoire des Sciences du Climat et de l'Environnement, Institut Pierre-Simon Laplace, CEA / CNRS / UVSQ, Gif-sur-Yvette, France, [5]Now at the European Center for Medium-range Weather Forecasts, Reading, UK, [6]Centre National de Recherches Météorologiques, Météo-France / CNRS, Toulouse, France

Key Points:

- The tuning process of IPSL-CM6A-LR under present-day control conditions is described
- The associated continuous atmospheric energetics adjustment is presented
- Successes, lessons and prospects of the IPSL-CM6A-LR tuning strategy are discussed

Correspondence to:
J. Mignot,
juliette.mignot@locean.ipsl.fr

Citation:
Mignot, J., Hourdin, F., Deshayes, J., Boucher, O., Gastineau, G., Musat, I., et al. (2021). The tuning strategy of IPSL-CM6A-LR. *Journal of Advances in Modeling Earth Systems*, *13*, e2020MS002340. https://doi.org/10.1029/2020MS002340

Received 16 SEP 2020
Accepted 13 APR 2021

Author Contributions:
Conceptualization: Juliette Mignot, Frédéric Hourdin, Julie Deshayes, Olivier Boucher, Jean-Louis Dufresne
Data curation: Guillaume Gastineau, Sébastien Denvil
Formal analysis: Juliette Mignot, Frédéric Hourdin, Guillaume Gastineau, Martin Vancoppenolle, Jean-Louis Dufresne, Jean-Yves Grandpeix, Olivier Marti, Clément Rousset
Investigation: Juliette Mignot, Frédéric Hourdin, Julie Deshayes, Olivier Boucher, Guillaume Gastineau, Ionela Musat, Martin Vancoppenolle, Jérôme Servonnat, Arnaud Caubel, Frédérique Chéruy, Sébastien Denvil, Jean-Louis Dufresne, Jean-Yves Grandpeix

Abstract

The assessment of current and future risks for natural and human systems associated with climate change largely relies on numerical simulations performed with state-of-the-art climate models. Various steps are involved in the development of such models, from development of individual components of the climate system up to free parameter calibration of the fully coupled model. Here, we describe the final tuning phase for the IPSL-CM6A-LR climate model. This phase alone lasted more than 3 years and relied on several pillars: (i) the tuning against present-day conditions given a small adjustment of the ocean surface albedo to compensate for the current oceanic heat uptake, (ii) the release of successive versions after adjustments of the individual components, implying a systematic and recurrent adjustment of the atmospheric energetics, and (iii) the use of a few metrics based on large scale variables such as near-global mean temperature, summer Arctic sea-ice extent, as targets for the tuning. Successes, lessons and prospects of this tuning strategy are discussed.

Plain Language Summary

Evaluating current and future risks for natural and human systems associated with climate change is largely based on numerical simulations performed with models of the climate system, which includes the atmosphere, the land, the ocean, the cryosphere, and the oceanic and terrestrial biosphere. Various steps are involved in the development of such models. First, models for individual components are developed and tested. Second, many aspects are represented with parameterizations that summarize the effect of a missing process, such as those happening on scales that are smaller than the model grid sizes. The parameterizations in turn involve many parameters, sometimes poorly estimated from observations, that have to be calibrated. Here, we describe the final tuning phase of the IPSL-CM6A-LR climate model, which includes several novel aspects: first, the choice to calibrate the model against present-day observations, which implies taking into account the transient nature of the observed climate; second, the systematic and recurrent adjustment of the atmospheric radiative budget; third, the use of a few large scale observable variables as targets. Successes, lessons and prospects of this tuning strategy are discussed.

1. Introduction

Climate change is a serious issue for humanity with important ramifications for policy and decision making. Robust and cost-efficient policies on mitigation and adaptation require assessments of current and future risks for natural and human systems under a range of socio-economic scenarios. Those assessments rely on numerical simulations performed with state-of-the-art climate models. Simulations are coordinated at an international level within the Coupled Model Intercomparison Project (CMIP) which provides the bedrock for a substantial part of the publications synthesized in the Intergovernmental Panel on Climate Change (IPCC) reports. Such projects are fundamental in order to document the robust features as well as the relatively large uncertainties in the future climate projections. Among others, these uncertainties come from the various assumptions made by the ~30 teams that develop CMIP-class models. In particular, because of

Project Administration: Olivier Boucher
Resources: Jean-Louis Dufresne, Marie-Alice Foujols
Software: Ionela Musat, Jérôme Servonnat, Arnaud Caubel, Sébastien Denvil
Writing – original draft: Juliette Mignot, Frédéric Hourdin, Julie Deshayes, Olivier Boucher, Guillaume Gastineau, Martin Vancoppenolle, Jean-Louis Dufresne

the complex and multi-scale nature of a climate model, many aspects are represented with parameterizations that summarize the effect of processes too small-scale or complex to be explicitly resolved by models. The behavior of those parameterizations often depends crucially on the values of "parameters" (hence the naming parameterization) that enter in their formulation. Those parameters are often only weakly constrained by observations (a unique size of cloud ice particle usable from tropics to pole for example) or not even observable (such as the exchange coefficient in a mass flux convective scheme). The values retained in a model configuration can come either from theoretical, experimental and modeling expertize, or from a combination of these three elements. They can also be adjusted from the examination of certain metrics calculated in preliminary climate simulations.

One of the most examined results of the CMIPs is the rate of warming that one can expect given a certain perturbation of the atmospheric radiative budget, in particular at the top of the atmosphere. The Equilibrium Climate Sensitivity is a widely accepted measure of Earth's change to radiative forcing defined as the change in the global mean surface temperature after reaching equilibrium in response to a doubling of the atmospheric CO_2 concentration (Randall et al., 2007). Early analyses suggest that several models of the latest phase of CMIP, CMIP6 (Eyring et al., 2016), predict larger global mean surface temperature increase than previous versions with a correspondingly larger ECS (Zelinka et al., 2020). More generally Meehl et al. (2020) describe that the range of ECS in CMIP6 is the largest of any generation of intercomparison projects since the 1990s. While they do not manage to identify a single reason for this behavior, these authors point at cloud feedbacks and aerosol-cloud interactions as the most likely contributors to the high ECS values and its increased range in CMIP6. Beyond providing estimates of the increase in global temperature, global climate models are also key for anticipating the consequences of global warming in terms of impacts and consequences on societies. This represents an important driver for the modeling teams to improve the realism of the simulated climate.

The derivation of the Institut Pierre-Simon Laplace (IPSL) coupled model for CMIP6 (IPSL-CM6A-LR) was an unprecedented coordinated effort during which key processes and parameters for climate modeling were identified both in the atmosphere and in the ocean. This work is documented in a series of papers in the same Special collection. Boucher et al. (2020) presents the IPSL-CM6A-LR climate model and a preliminary evaluation of the *historical* simulations. Lurton et al. (2020) documents the implementation of the CMIP6 climate forcings. Individual components of the model are described in separate publications for the continental surfaces (Cheruy et al., 2020) and for the atmosphere (Hourdin, Rio, Jam, et al., 2020). Specific contributions are dedicated to the representation of clouds (Madeleine et al., 2020), the impact of sub-grid scale orography tuning on the simulation of northern high latitude climate (Gastineau, Mignot, Lott, & Hourdin, 2020) and the reduction of the Eastern Tropical oceans warm biases (Hourdin, Rio, Jam, et al., 2020). This study is devoted to the presentation of the 3-year collective work and of the tuning strategy that led to the final IPSL-CM6A-LR model configuration.

The word tuning, rather vague and specific to the climate modeling community, designates here the full phase of debugging and calibrating parameters, targeting some metrics or model behavior once the model's physical content has been fixed. Tuning of free parameters in particular is now recognized as a key step in the development of a climate model, in particular with the purpose of stabilizing the global mean temperature at a reasonable level (e.g., Hourdin et al., 2017; Mauritsen et al., 2012; Schmidt et al., 2017; Senior et al., 2020). Yet, it is particularly important to document aspects of the simulated climate which were explicitly targeted by the tuning from those which correspond to emerging properties of the simulated climate. In this respect, and in the framework of climate change simulations, one may clearly distinguish model groups that claim to tune the model ECS and/or the historical trajectory of global mean temperatures (e.g., Danabasoglu et al., 2020; Mauritsen & Roeckner, 2020) and those who claim not to (e.g., Dunne et al., 2020; Senior et al., 2020). The IPSL strategy should be classified in the second category, although it relies on the use of a present-day equilibrium setup. It also implies the systematic re-tuning of the energetics of the atmospheric component in stand-alone mode, with a few well-identified metrics each time a new configuration of the coupled model was available. Documenting the tuning targets that proved challenging may also point to model intrinsic deficiencies and possible error compensations. The aim of this study is to provide a summary of the performance metrics and tuning strategy as well as a high-level description of the

successes and challenges encountered during the tuning, complemented by insightful discussions on the implications of such an exercise.

The model description is summarized in Section 2. The tuning strategy is described in Section 3 and key lessons of this development phase are discussed in Section 4. Section 5 concludes this study and discusses implications for the ECS.

2. Building Up a New Coupled Model

2.1. The Components of IPSL-CM6A-LR

The IPSL-CM6A-LR climate model developed in view of the IPSL participation to CMIP6 is described in the Special collection by Boucher et al. (2020). It uses broadly the same components as the versions derived for CMIP5, each of which has nevertheless evolved separately since 2013. Most of the changes in the components are already published, in particular in a series of papers in the same Special collection. They are only briefly summarized in the following section that also provides a guide for the reading of the other articles of the Special collection.

Compared to the CMIP5 version, the version of the atmospheric model LMDZ designed for CMIP6, named LMDZ6A-LR (LR for Low Resolution) uses the same horizontal grid as the MR grid (Medium Resolution) used for CMIP5, with 144×143 points equally distributed in longitude and latitude. The vertical grid was changed from L39 to L79, with an increased vertical resolution in the first 3 km above the surface and in the stratosphere to improve the simulation of the atmospheric boundary layer and the troposphere-stratosphere coupling. In terms of atmospheric physics, the configuration is a continuation of the 5B version of LMDZ and IPSLCM since it uses the so-called "New Physics" version, by opposition to the 5A "standard physics" version (Hourdin, Grandpeix, et al., 2013). This 5B version encompasses about 10 years of research on the parameterization of convective and cloudy processes (Rio, Grandpeix, et al., 2013). Yet, this version, although presenting important advances on the representation of cloudy and convective processes, with a better representation of low clouds and the diurnal cycle of continental convection for example, suffered from obvious initial defects. For CMIP6, the LMDZ team has designed a more complete and better tuned version of this "New Physics" described in this Special collection (Hourdin, Rio, Jam, et al., 2020). The representation of very stable boundary layers was significantly improved (Vignon et al., 2017, 2018). A specific tuning of the effect of subgrid-scale orography was done targeting the northern mid and high latitude circulation (Gastineau, Mignot, Lott, & Hourdin, 2020). Thanks to parameterization improvements and a better tuning targeting the space and time distribution of Cloud Radiative Effect (CRE) at the top of the atmosphere, the clouds and radiative fluxes are much better represented in this new version (Madeleine et al., 2020). In particular, a modification of the boundary layer convective transport model to represent marine stratocumulus (Hourdin et al., 2019) allowed the reduction of Sea Surface Temperature (SST) warm biases in the Eastern part of tropical oceans (Hourdin, Rio, Jam, et al., 2020). The LMDZ6A version also includes a stochastic initiation of convection (Rochetin, Couvreux, et al., 2014; Rochetin, Grandpeix, et al., 2014). The density of cold pools which was set to a single constant over the globe in LMDZ5B takes two different values in the CM6A configuration (density of cold pools being smaller over land than over ocean). Parameters controlling the vertical velocity at the basis of deep convective towers were also used to adjust the intensity of precipitation over land. In general, those modifications tend to increase the variability of precipitation over the ocean, also reinforcing the contrast between suppressed and active phases of oceanic convection.

The coupling with the ORCHIDEE model for continental surfaces is described by Cheruy et al. (2020) and its impact on the West African climate documented by Diallo et al. (2017). Key improvements are a better representation of the surface soil moisture and the seasonal cycle of river discharge thanks to a multi-layer hydrology, improved surface temperature in polar regions thanks to the refined turbulent diffusion scheme, updated longwave radiation scheme in LMDZ and a novel snow scheme over land of intermediate complexity. Most of the tuning of land surface was based on sets of multidecadal (i.e., 10–30 years) simulations either with stand-alone land surface simulations forced by atmospheric analysis or with coupled land-surface-atmosphere simulations. The main targets were the seasonal near surface temperature biases at the regional scale and their response to the evaporative cooling. Particular attention has also been paid to the

seasonality of the river discharge in the northern basins in relation to the freshwater discharge as well as to land surface albedo.

The oceanic model, including sea ice, has evolved as well since CMIP5. IPSL-CM6A-LR uses the NEMOv3.6-stable version of NEMO on the eORCA1 grid, with typical horizontal resolution of 1° and an extension of the grid toward the South Pole in order to better represent ice-ocean interactions in the Southern Hemisphere. The number of vertical level was increased as well from 31 to 75, with a higher resolution close to the oceanic surface. The coupling frequency was also increased from 1 day to 90 min. These two changes were implemented in order to attempt to represent a diurnal layer. The sea ice component, based on LIM3 and part of the ocean model, has significantly evolved for CMIP6, now including, beyond ice thermodynamics and dynamics, brine physics and five ice thickness categories. As for the land surface model described above, the ocean and sea ice were developed, adjusted and tuned in stand-alone mode before being coupled to the atmosphere and before the tuning process presented here. The tuning of the oceanic component benefited as well from the fact that the same configuration was used in three other CMIP6 models: CNRM-CM6-1 developed at CNRM-Cerfacs (Voldoire et al., 2019), the EC-Earth3 model of the EC-Earth Consortium (Wyser et al., 2020) and HadGEM3-GC3.1-LL from the MOHC group (Kuhlbrodt et al., 2018; Menary, Kuhlbrodt, et al., 2018).

2.2. Building Up the Coupled Model

Building up the new coupled model has been a more than 3-year long process that started in mid-2015 when IPSL-CM6.0.1 was created. Over the 3 years that followed, 15 versions were released before IPSL-CM6A-LR was officially available. Each new release (from IPSL-CM6.0.1, named CM6.0.1 in the following to IPSL-CM6.0.15, named CM6.0.15) was characterized by a mix of implementation of recent model developments (during the first year or so mainly), bug removals (during the second year or so mainly) and parameter calibrations (in particular during the final stage), as summarized in Table 1. In this study we describe the last two phases of debugging and tuning, scanning the last 11 intermediate versions from CM6.0.5 until the final version CM6.0.15. The first four versions were relatively preliminary, mostly dealing with global conservation, quality control and numerical stability issues. More importantly, the use of present-day control simulations for model tuning began with version CM6.0.5 only. The last version CM6.0.15 was then frozen into CM6.1.0, described in Boucher et al. (2020) and named IPSL-CM6A-LR for the Earth System Grid Federation (ESGF).

Special attention was paid to water and energy budgets during the whole development procedure. Our objective was to track and minimize leakages that inevitably appear through the coupling. The full freshwater budget and the oceanic heat budget in the final configuration are presented in Appendices A and B.

Regarding the global water budget, Boucher et al. (2020) describe in their Section 2.5 the choices that were made to redistribute the water at the interface between land ice and the ocean. In addition, Appendix A of this study shows the diagnostics that we implemented during the model development to document freshwater fluxes within each of the model components, at the various interfaces (atmosphere-ocean, land-ocean, and land ice-ocean). We have largely improved water conservation between IPSL-CM5 and IPSL-CM6 and are left with a fairly small spurious source of freshwater of the order of 1 mSv (Appendix A).

At present, the energy is not conserved in the atmospheric model component, and no global energy fixer (Williamson et al., 2015) is implemented. There is a net radiative flux at the top of the atmosphere of about 0.7 W.m^{-2} in preindustrial control conditions, which is among the largest values found in CMIP5 (Hobbs et al., 2016). The ocean model has a net incoming heat flux of -0.011 W m^{-2}, while a drift of -0.114 W m^{-2} is diagnosed (see Appendix B). The discrepancy between the drift and net incoming flux requires further investigation, and may be due to diagnostic errors and/or to numerical approximations leading to nonconservation of heat. We also checked that this numerical imbalance does not change between preindustrial and present-day climate, and therefore does not introduce an artificial feedback. Additionally, it is likely the IPSL-CM6A-LR simulation has not reached equilibrium, which might also explain the drift. In any case, consistently with the results of (Hobbs et al., 2016) analyzing CMIP5 models, we find no simple relationship between the ocean heat content changes and net top-of-atmosphere radiation.

Table 1
Table Summarizing the Chronology of the 15 Successive Versions, Together With the Main Modifications, Improvements and Remaining Challenges in Each New Version as Compared to the Previous Ones

Version	Main modifications	Prominent features
CM6.0.1 (July 2015)	First assembled version of the coupled model with new grid. First long coupled simulation	Rain variability very weak
		Frequent crashes in the atmosphere
CM6.0.2 (February 2016)	Changes in atmospheric convective mixing	Continents much too cold
	Corrections in atmospheric dynamics	Increased rainfall variability
		Many fewer crashes
CM6.0.3 (April 2016)	New atmospheric radiation scheme	Centennial crashes, mostly over Himalayas
	Tuning of the atmospheric boundary layer	Warm SST biases (ETO, Southern Ocean)
		Cold continental bias reduced
CM6.0.4 (June 2016)	First reasonable tuning	Reduced warm SST biases,
		Good sea ice cover
		But vanishing AMOC
CM6.0.5 (July 2016)	New atmospheric tuning	
	Starting *pdControl* simulations with target metrics	
CM6.0.6 (October 2016)	Routing bugs corrected	
CM6.0.7 (December 2016)	Cloud parameterization (Madeleine et al., 2020)	Disappearing summer Arctic sea ice
	Freshwater fluxes around Antarctica	Atmospheric convection active everywhere
		Himalayas too cold, Siberia too hot
CM6.0.8 (February 2017)	First tests on sea ice parameters	Disappearing summer Arctic sea ice
CM6.0.9 (March 2017)	New SSO parameterization (Gastineau, Mignot, Lott, & Hourdin, 2020)	Improved atmospheric circulation over North Atlantic
	New sea ice parameters	Reduced winter warm bias over Siberia
	Optimization of running speed	Amplitude of ENSO overestimated
		Precipitation over the ocean too large
		More Arctic sea ice
		AMOC vanishes (bug in river discharges)
CM6.0.10 (April 2017)	Conditioning deep convection triggering (Ttop)	Overestimated SST in ETO
	New runoff distribution scheme	Better representation of ENSO
		Increased number of days without rainfall
		More intense AMOC
CM6.0.11 (July 2017	Spatial spread of the river outflow,	Lack of deep convection in the Labrador Sea (sea-ice covered)
	Vertical mixing under sea ice	
	Thermal plumes outside cold plumes only (split)	
CM6.0.12 (August 2017)	Vertical mixing under sea ice	Intensification of deep convection in the Labrador Sea
	Forcings, diagnostics, workflow	
CM6.0.13 (October 2017)	Preparing outputs and diagnostics	
	Forcings, diagnostics, workflow	
CM6.0.14 (November 2017)	Adjustment of aerosols indirect effect	
	Forcings, diagnostics, workflow	
CM6.0.15 (February 2018)	Adjustment of aerosols indirect effect	→ CM6.1.0 (March 2018)
	Diagnostics, workflow	

In the table and further in the text, ETO and SSO stand, respectively, for Eastern Tropical Ocean and Subgrid-Scale Orography respectively.

3. The Tuning Strategy

3.1. Specific Tuning Targets of the Coupled Model

Specific target metrics were explicitly identified for the coupled model allowing us to continuously adjust the deep ocean through successive versions of the coupled model.

The targets were chosen based on issues identified either in the pre-existing CM5B version or during the construction and testing phase of the CM6A version. In terms of SST, the CM5B version shows some classical biases among coupled models, such as a warm bias in the southern high latitudes and over the eastern sides of tropical oceanic basins. The reduction of those biases was an explicit target of the CM6A tuning. In the CM5B version and in preliminary versions of CM6A, summer Arctic sea ice extent was also systematically underestimated and this was associated with a warm winter bias of the near-surface atmosphere over the Arctic. The Southern Ocean tended to form a very large polynya, impacting the deep (1000 m) oceanic temperature while deep convection was too shallow in the Labrador Sea, as in many coarse resolution climate models (Menary, Hodson, et al., 2015) and the Atlantic overturning circulation was too weak. These identified issues led us to target three specific variables: (i) the pattern of SST bias in the tropical band, (ii) the Arctic sea ice volume and extent in summer, and (iii) the magnitude of the Atlantic meridional overturning circulation (AMOC). The latter is, however, difficult to precisely constrain given the lack of observations and the fact that it is tied to multiple processes. It was thus roughly compared to both CMIP5 ensembles and the RAPID estimations. The intensity and location of deep convection in the Southern Ocean and the North Atlantic Ocean were also considered. Yet, given the complex underlying processes they imply and the lack of direct observations, they could not always be used as quantitative tuning targets but simply inspected visually It should be noted that not much effort was put in to tuning the representation of the modes of rainfall variability, such as that associated with the Madden-Julian Oscillation.

More importantly, the ECS or the evolution of surface temperature (SST or Global mean surface temperature) during the historical period were not considered as targets for the tuning. Only the present-day global temperature is considered as a tuning target. This translates into a *piControl* climate relatively warm or cold for a respectively small or large ECS. This tuning strategy is detailed below.

3.2. Tuning Equilibrium Simulations With Present-Day Observations

The general idea of the tuning strategy of the climate model aims to tune the model in present-day (pd) conditions while running in parallel preindustrial (pi) and pd equilibrium simulations. The amount of climate observations of the last decades provides a strong incentive to favor present-day conditions for model tuning. Strictly speaking, however, present-day observations should be compared with transient historical simulations themselves initialized in preindustrial conditions. The transient increase of greenhouse gases concentration and other external forcings reduces the outgoing longwave radiation and leads to a radiative imbalance at the top-of-atmosphere, due primarily to the large thermal inertia of the ocean. The absolute value of this imbalance is not directly observed. Most recent estimates are derived from changes in ocean heat content with values of about 0.5–1 $W\,m^{-2}$ at the end of the twentieth century (Von Schuckmann et al., 2016), while atmospheric satellite measurements suggest 0.9 $W\,m^{-2}$ (Trenberth et al., 2009). Here, we chose a value of 0.6 $W\,m^{-2}$.

We propose here a simple strategy to tune multi-centennial (and thus in principle equilibrated) coupled simulations against present-day observations taking this imbalance into account. The present-day imbalance largely mirrors the heat taken up by the ocean as climate is warming. This imbalance disappears when equilibrium is reached after some warming. Performing a simulation that is both at equilibrium and close to the current climate therefore requires replacing this ocean heat uptake by a constant, artificially imposed heat flux at the atmosphere-ocean interface. A possible solution would be to add a flux divergence at the atmosphere-ocean interface, as was done for instance in the first coupled models to avoid large drifts or biases in surface temperature (Sausen & Hasselmann, 1988). We prefer another solution, easier to implement. In our approach, we choose to reduce the ocean model's incoming heat flux by increasing the ocean surface albedo. The two approaches are not exactly equivalent because the patterns of these two artificial fluxes differ and the additional reflected solar flux is partially absorbed by the atmosphere. However, we will see that this

simple solution still gives satisfactory results with respect to our objectives. The two approaches also differ when considering the global mean flux at the TOA. Assuming energy conservation is perfect in the model, the first approach implies that mean TOA flux is equal to the present day value of the ocean heat uptake whereas the second approach implies the mean TOA flux is zero. Indeed, an increase of the albedo leads to an increase of the solar radiation reflected at the surface and therefore a decrease of the temperature, but it does not change the global energy budget of the Earth that is zero at equilibrium.

Performing sensitivity runs with the atmospheric model with prescribed SST we find that an increase of 0.007 units of the ocean surface albedo (roughly 0.7%) reduces the mean TOA fluxes by 0.6 $\mathrm{W\,m^{-2}}$ (see end of Section 3.5). Then we run the atmosphere-ocean coupled model with constant present-day forcing taken from the last decades of the CMIP historical protocol (we choose here the forcing of the years 1990–2010) and with this small artificial increase of the oceanic surface albedo. By analogy with the *piControl* simulation with constant preindustrial conditions defined in CMIP6, we coined the term *pdControl* to name this present-day simulation at equilibrium, that is, with radiative forcing (effect of greenhouse gases, aerosols, land-use) prescribed at their present-day values and with an oceanic albedo increased by 0.7% to compensate the present-day ocean heat uptake.

In practice, when introducing the albedo offset in the coupled simulations, the global temperature and sea ice extent adjust to the modification of the surface albedo to compensate for the flux perturbation on a timescale of a few decades, with the global TOA imbalance coming back to its original value. This asymptotic value would be zero in a perfect model run until reaching equilibrium. It is not the case for the IPSL model, because of the energy numerical source of $\sim$0.7 $\mathrm{W\,m^{-2}}$ documented in *piControl* (Appendix B). Note that this value is at first order independent of the model tuning or simulation setup. It is in particular the same in the *piControl* and *pdControl* simulations. What happens in practice when introducing for instance an albedo perturbation of 0.007 in a coupled simulation that has reached equilibrium is that the imbalance decreases from 0.7 to 0.1 $\mathrm{W\,m^{-2}}$ (from 0 to −0.6 if the model was conserving energy) and then relaxes to 0.7 $\mathrm{W\,m^{-2}}$ (0 if the model was conserving energy) with a time constant of a few decades while the global temperature decreases. The asymptotic temperature change scales with the initial flux perturbation with a constant of about 1 $\mathrm{K\,W^{-1}\,m^{2}}$ which depends upon the sensitivity of the climate model considered.

3.3. From pd to pi Simulations

According to CMIP protocol, historical simulations and their prolongation into future projections must start from preindustrial conditions. Transitioning a simulation from *pdControl* to *piControl* conditions is simple: one simply has to change the radiative forcers from their 1990–2010 values to preindustrial ones, and remove the artificial albedo offset. However, it is important to take into account the time it takes to adjust the deep ocean, of the order of several hundred years, once the radiative forcers and albedo are fixed. Waiting for the end of the tuning process in *pdControl* conditions before starting *piControl* simulations would require another few months of preliminary computation before producing the actual CMIP simulations.

We thus decided from version CM6.0.11 onward to run in parallel pairs of simulations in present-day and preindustrial configurations. Each time a modification or re-tuning was done on the present-day simulation, it was also applied to the corresponding preindustrial simulation. The only difference between the *pdControl* and the corresponding *piControl* was thus the set of radiative forcings used and the surface ocean albedo correction switch. With this strategy, and because the *pdControl* was constantly retuned to a given target SST, the deep oceanic temperature kept on equilibrating during the tuning process. Of course, this does not remove the need for a final *piControl* spinup after the last modification in the model specifications. This final simulation spans 600 years, the last 100 of which corresponds to the preindustrial control spinup (*piControl-spinup*) simulation published on the Earth System Grid Federation (ESGF).

To summarize, the main goal of the methodology was to provide an easy and robust way to equilibrate *piControl* so that historical simulations reach the observed present-day mean global temperature in present-day conditions. Its success and the choice of the value of 0.007 for the albedo perturbation were demonstrated a posteriori by checking (Figure 1) that the global mean temperature of the *pdControl* simulation lies within the range of present-day global temperature as obtained in an ensemble of historical simulations, initialized from the corresponding *piControl* simulation. As mentioned above, both the *piControl* and

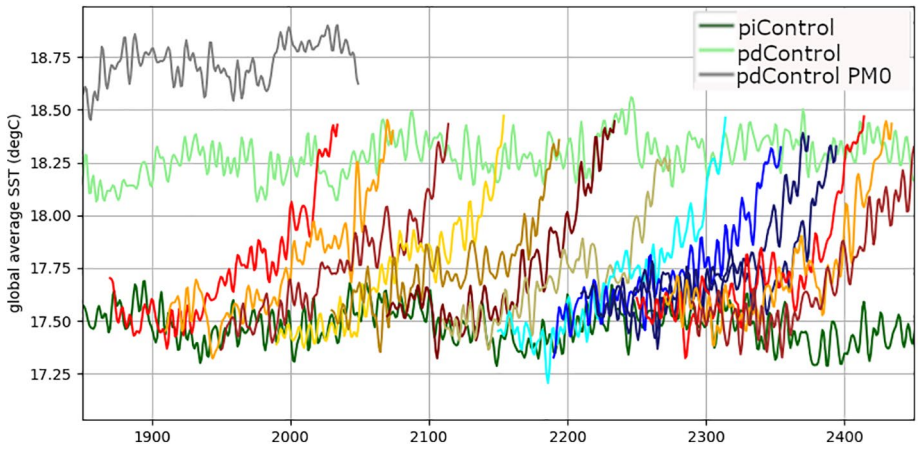

Figure 1. Time evolution of global mean SST in the preindustrial control (*piControl*) simulation (dark green), the present-day (*pdControl*) simulation used for the tuning process (light green) and 9 historical simulations starting from different initial conditions of the *piControl* simulation (color lines). The gray line shows the global mean SST in a present-day simulation in which the correction of surface ocean albedo described in the text was removed.

historical simulations are run without the albedo offset. Running the present-day simulation without the albedo offset, on the other hand, leads to a global SST about 0.5 K warmer (gray curve in Figure 1).

Note that this procedure, which was thought of as an indirect method to target the adjustment of the global mean temperature, also leads to present-day climate metrics relatively similar between the *pdControl* and historical ensembles when considering the Northern Hemisphere sea ice coverage and other variables (Section 4). However, the mean climate of the historical ensemble shows slightly more sea ice than the *pdControl* simulation (Figures 5 and 6). This may be related to some inertia in the melting of sea ice or in the land surface in general in the *historical* simulations. Of course, a slightly larger value of the albedo offset targeting the same present-day global temperature with the *pdControl* simulation (the light green curve in Figure 1) would produce a climate a bit warmer (a common positive shift of all the other curves) with less sea ice in the *historical* simulations.

A more sophisticated approach could be envisaged, using for instance a map rather than a constant value for the albedo offset (to account for the spatial distribution of the oceanic heat uptake, for example) so that the *pdControl* climate is even closer to that of the *historical* ensemble. We prefer however to keep it simple, having in mind that we are tuning here the global mean temperature. Nonetheless, we document for other metrics the difference between the *pdControl* and *historical* simulations of the present-day climate, as this information may be valuable for future tuning exercises.

There is nothing in this methodology that constrains or modifies the ECS, or uses the 20th century trend in global temperature as a tuning target. If the artificial albedo offset of the present-day simulations was chosen at a different and less appropriate value, both the preindustrial and present-day temperature from historical simulation would be either a bit colder or a bit warmer, but the 20th century trend would be the same. If two model configurations with a different sensitivity to greenhouse gases were tuned with the proposed methodology (which imposes the temperature at the end of the 20th century), the preindustrial simulation with the most sensitive configuration would be globally colder, resulting in a larger 20th century trend. Such a case is, for example, documented in the next section when two distinct branches of the model have been developed in parallel before one was selected and the other one stopped.

3.4. Systematic Retuning in Standalone Atmospheric Mode

The development and tuning was done in an iterative manner. Each time a new configuration of the coupled model was defined, a retuning of the model energetics was done with the stand-alone atmospheric model, including interactive continental surfaces but with prescribed SSTs and sea ice concentrations. This retuning was done by running series of typically 10 sensitivity experiments modifying the most uncertain

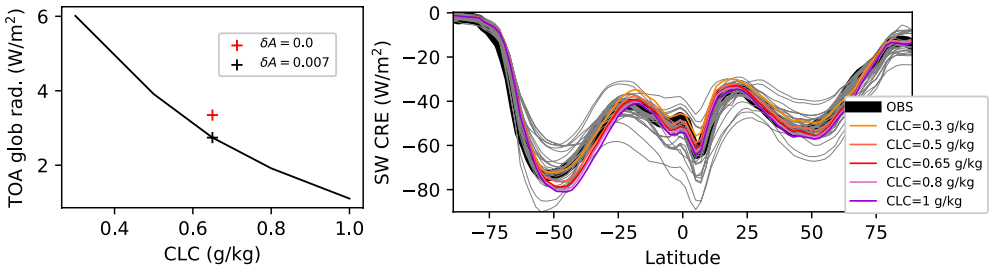

Figure 2. Sensitivity of the global top-of-atmosphere radiation imbalance (left panel) and of the latitudinal variation of the SW cloud radiative effect (right panel) to the value of the critical in-cloud liquid water content (CLC) parameter in the atmosphere-only (with AMIP SST) configuration of the model. The two crosses in the left panel show the TOA radiation imbalance for a specific choice of CLC and for the surface ocean albedo offset (black) and lack of offset (red). The colored curves on the right panel correspond to sensitivity experiments run with CLC = 0.3, 0.5, 0.8, and 1 $g\,kg^{-1}$ starting from the final CMIP6 configuration which uses CLC = 0.65 $g\,kg^{-1}$ (red curve). The gray curves correspond to the 15 intermediate versions of the coupled model.

parameters that control key aspects of the radiative budget. These parameters, that mostly concern the cloud and convection schemes, are listed in Table 3 of Hourdin, Rio, Grandpeix, et al. (2020). The difficulty of this retuning and the number of iterations required was highly dependent on the importance of the changes made on the new configuration.

The targeted metrics are listed in Hourdin, Rio, Grandpeix, et al. (2020). They concern mainly the top-of-atmosphere radiative balance, its decomposition into longwave (LW) and shortwave (SW) components, the cloud radiative effect (CRE), as well as key elements of their space and time distribution with the specific aim to reduce SST biases in the coupled model. The targets were more specifically the latitudinal variations of the zonally averaged quantities (targeting in particular a reduction of the Antarctic circumpolar warm bias) and the longitudinal variations over the tropical oceans to reduce the Eastern Tropical Ocean (ETO) warm bias. This last point is the central subject of Hourdin, Rio, Grandpeix, et al. (2020).

3.5. Readjusting Global SSTs on the Fly

When moving from one version of the coupled model to the other, particular care was given to reproducing an overall energy balance at the top of the atmosphere as close as possible to that of the previous configuration. By doing so, it was possible to extend the *pdControl* and *piControl* simulations performed with the previous version, avoiding jumps in the global-mean temperature, thus allowing the intermediate and deep ocean to continue its adjustment. Figure 3 (second panel) illustrates the success of this systematic retuning to allow changes in the model content (illustrated by the succession of orange and red colors for the *pdControl* and light and dark blue for *piControl* simulations) while keeping some continuity in the oceanic temperature evolution.

The calibration of the atmospheric model discussed above was most of the time sufficient to avoid jumps in the global-mean temperature larger than a few tenths of a degree. However, a single parameter, the Critical in-cloud Liquid water Content (CLC) that goes into the formulation of the conversion of cloud liquid water to rainwater, was used to ultimately adjust the global-mean SST of the running *pdControl* simulation to be typically within ±0.1 K of the previous version. Increasing the value of CLC increases the cloud water content and hence the cloud albedo. This parameter has a strong effect on the global radiation balance: a variation by ±50% of this parameter changes the top-of-atmosphere radiation balance by 2 or 3 $W\,m^{-2}$ (left panel in Figure 2), without affecting too much the latitudinal variation of the SW CRE (right panel). Similar curves were built on the successive versions of the coupled model and used as tools for the fine retuning of the TOA global radiation. Conversion to global-mean SST was done through an empirical process based on a correspondence ratio of about 1 K/($W\,m^{-2}$). This value corresponds approximately to the equilibrium climate sensitivity (ECS) to a radiative perturbation, which was not known at this stage for our climate model, and is model dependent (Sherwood et al., 2020). CLC modification proved to be very useful to bring back the global temperature to the predefined target if a small drift was observed after a change in model version, despite the atmospheric energetics tuning described above. Figure 2 (left panel) also

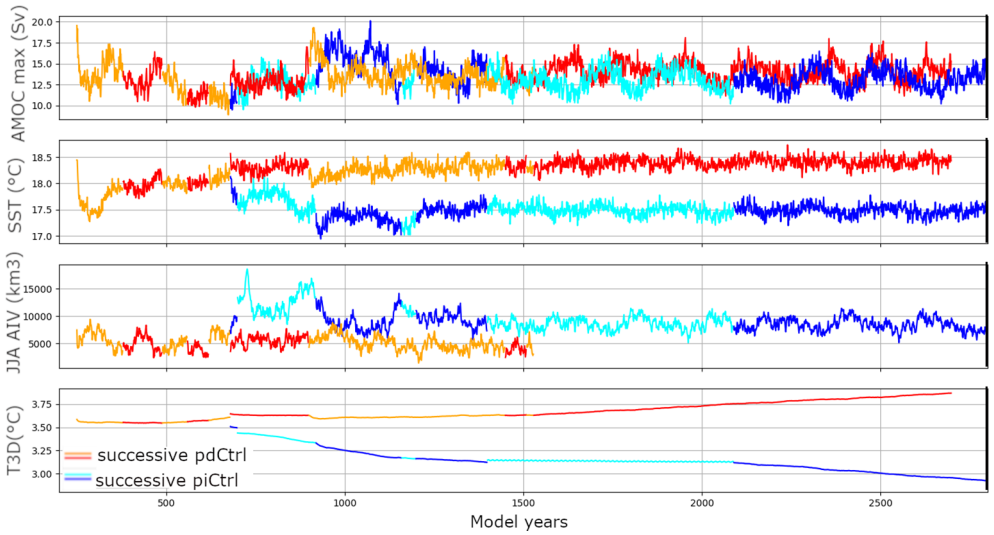

Figure 3. Time evolution of the AMOC maximum taken between 10 and 60°N (top panel), global mean SST (second panel), summer Arctic sea ice volume (third panel) and global mean averaged oceanic temperature (bottom panel) of the successive configurations starting from the last one (first on the left) that was restarted from the Levitus climatology (Locarnini et al., 2013) at rest. The orange and red curves are for the successive *pdControl* configurations. The first present-day simulation considered here is a tuned version of CM6.0.11-pd. The light and dark blue curves are for the successive *piControl* simulations. The first preindustrial control simulation was started from CM6.0.12-pd. Note that although, from this version onward, all changes applied to the *pdControl* configurations were systematically also applied to corresponding *piControl* simulations, there may have been some intermediate *pdControl* versions before changes were applied to *piControl* and all configurations were not integrated for rigorously the same length. This explains that there is not exact matching between the succession of *pdControl* and *piControl* versions.

shows for the nominal model version (CLC = 0.65 $g\,kg^{-1}$) the effect of removing the surface ocean albedo correction of $\delta A = 0.007$ for the standard model configuration. This produces a global change in radiative balance of 0.6 $W\,m^{-2}$ in stand-alone atmospheric simulations, consistent with the change in global equilibrium temperature of about 0.5 K in coupled simulations shown in Figure 1 between the *pdControl* and *pdControl PM0* configurations. Note to finish that, because our main goal was to keep the global mean SST of the *pdControl* simulations on the same global mean temperature target between version n and $n + 1$, the top-of-atmosphere global imbalance in the AMIP simulations was retuned not to a theoretical, observed or predefined value but to the value obtained in the same AMIP mode with the atmospheric tuning of coupled version n. Generally speaking, it is essential to think in terms of perturbations rather than absolute values for the global energy budget (or temperature in coupled mode), when considering both climate change simulations and tuning issues.

Figure 3 also shows that the deep water temperature is not fully stabilized at the end of the process because the *pdControl* and *piControl* curves continue to diverge from each other. The drift agrees with what is found in CMIP5 models (Hobbs et al., 2016). In the case of IPSL-CM6A-LR, this is related to the imbalance in heat flux received by the liquid ocean (see Figure B1), which is positive in *pdControl* conditions and negative in *piControl* conditions (not shown). To summarize, the tuning strategy presented in the previous sections is summarized in Figure 4.

4. Lessons Learned From the Development and Tuning Sequence

We now discuss the main modifications and parameter calibration done along the 15 successive versions of the coupled model, as well as the subset of metrics that were used to assess the improvements.

4.1. Sea Ice

The sea ice tuning was largely done offline, including many details of the sea ice state and processes, in particular sea ice drift. All this was done prior to the work presented here, through a long history of OMIP-type

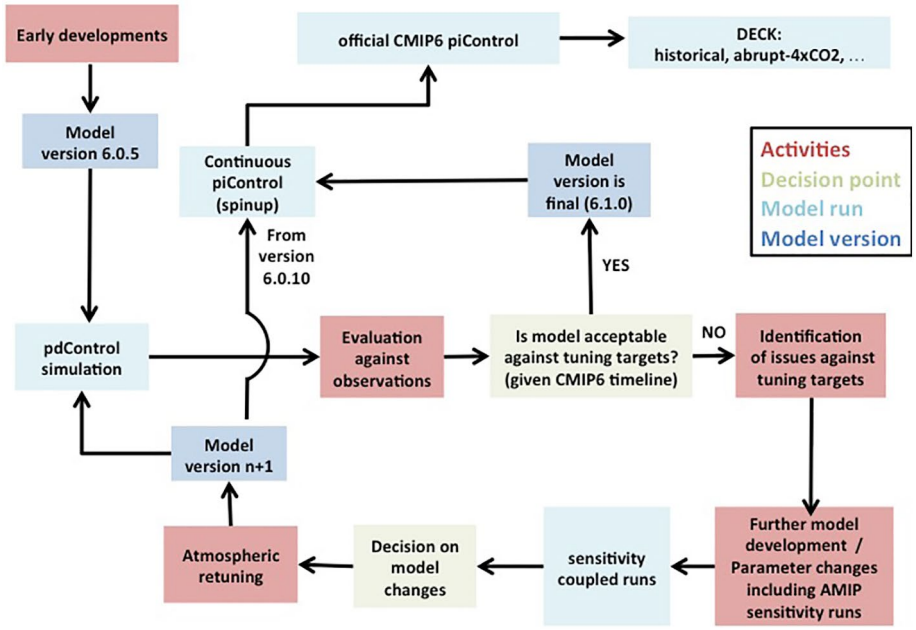

Figure 4. Schematic of the tuning strategy of the IPSL-CM6A-LR climate model highlighting the relative roles of *pdControl*, stand-alone atmosphere-only and *piControl* simulations. The color code corresponds to the nature of the boxes involved as per the legend in the top-right corner.

runs, only including the ocean and sea ice components, and forced alternatively by climatological or inter-annual atmospheric reanalyzes (Rousset et al., 2015; Uotila et al., 2017; Vancoppenolle et al., 2009).

In coupled mode, only a few tuning targets were retained for sea ice, including the observed seasonal cycle of Arctic and Antarctic sea ice area, and 20,000 km^3 for the annual mean Arctic sea ice volume (Notz et al., 2013). Because of the lack of multi-year observations of Antarctic sea ice, the Southern Ocean ice volume was not considered particularly important. The tuning was based on the visual inspection of times series. Ice concentration maps were also examined, in both hemispheres, to verify that the spatial distribution of sea ice looked consistent with observations.

Achieving a reasonable Arctic sea ice cover was arduous in fully coupled mode. The tuning of the energetics of the atmospheric component targeted the observed mean SST from 50°S to 50°N, in order to separate this tuning from that of sea ice. This SST tuning left a positive winter air temperature bias in the Arctic region. As a consequence, winter sea ice grew insufficiently and summer sea ice typically disappeared (CM6.0.7 and CM6.0.8 in Figure 6). To compensate for these deficiencies, we reduced surface melting by increasing the surface albedo for all ice surface types (in particular the diffuse broadband melting bare ice albedo from 0.53 to 0.58). Second, we increased winter basal growth by increasing the thermal conductivity of snow on sea ice to 0.5 $W m^{-1} K^{-1}$. For both albedo and snow thermal conductivity, we reached what we considered the largest acceptable values. The subgrid-scale orography parameterization of the atmospheric model was tuned as well in coupled mode leading to a reduction of the northward atmospheric heat and moisture transport at 60°N, decreasing the warm lower-tropospheric winter bias over the Arctic (Gastineau, Mignot, Lott, & Hourdin, 2020). These adjustments finally resulted in an increase of Arctic sea ice area of ~1.5–2 × 10^6 km^2 in June-July-August after version CM6.0.9 and a remaining low summer sea ice area bias of about 1–2 × 10^6 km^2 (depending on ensemble members).

Interestingly, the relationship between the global-mean surface temperature and the Arctic summer sea ice area (hereafter SITS for *sea ice temperature sensitivity*)—one of the key metrics used to describe the Arctic sea ice sensitivity in climate change scenarios (Mahlstein & Knutti, 2012; SIMIP Community, 2020)—was left nearly unaffected through the whole tuning process of the coupled model (Figure 7). For instance, the large increase in Arctic sea ice volume and area from versions CM6.0.8 to CM6.0.9, associated with surface albedo and subgrid-scale orography tuning, was not accompanied by changes in SITS. The SITS can also be

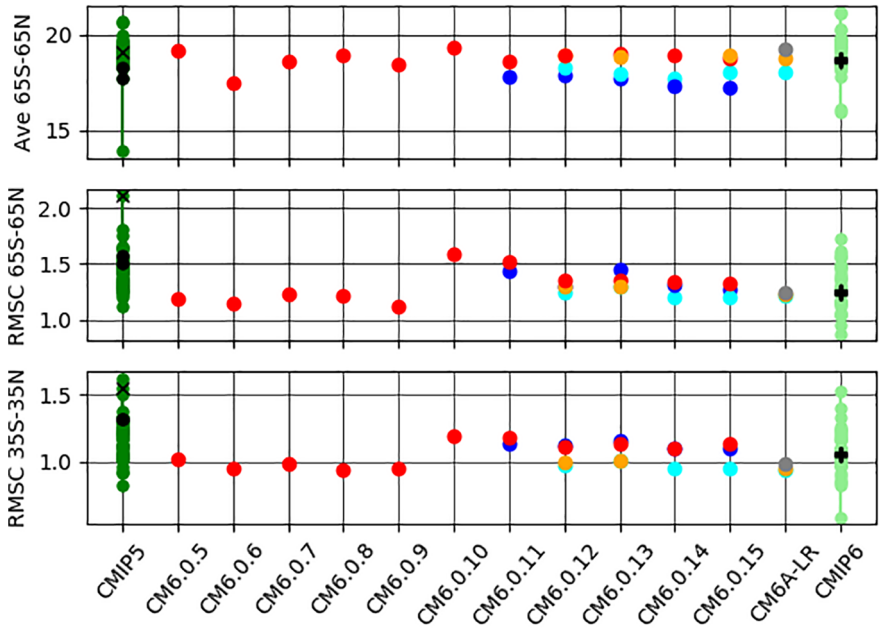

Figure 5. Evolution of three metrics related to SST biases along the course of the tuning process of IPSL-CM6A-LR. The three metrics are the sea surface temperature averaged between 65°S and 65°N and expressed in °C (top panel), the spatial and temporal root mean square error of the SST anomaly (called "centered root mean square error", RMSC on the figure), in °C, computed from the mean seasonal cycle over the same domain (middle panel) and over the 35°S–35°N tropical band (bottom panel). The red and orange circles stand for the successive *pdControl* simulations, given that two atmospheric sets of parameterizations have been kept after a while (see text for details). The blue and cyan circles stand for the equivalent *piControl* simulations that were systematically run in parallel from CM6.0.11 onward (see text for details). Orange and cyan circles stand for the "split" version run in parallel with the "Ttop" one from version CM6.0.12 onward (see Sections 4.3 and 4.4). The gray circle stands for the *pdControl* simulation running without compensation for oceanic heat uptake also shown in Figure 1. All these computations are done using the last 80 years of each simulation. The dark (light) green circles show the same diagnostics computed over the CMIP5 (CMIP6) multi-model historical ensembles (considering one member per model) over the period (1979–2005). The black dots in the CMIP5 stand for the two IPSL-CM5A configurations (LR and MR) and the black "x" mark corresponds to IPSL-CM5B-LR. The black crosses on the CMIP6 ensemble show the spread for the historical large ensemble (32 members) performed with IPSL-CM6A-LR.

diagnosed from >100-years equilibrium simulations, which suggests this relationship to be characteristic of the model centennial variability (Boucher et al., 2020).

Similar tendency to a lack of sea ice in the winter Southern Ocean was also seen in early versions (not shown). This was primarily linked to substantial warm surface temperature over the Antarctic ice sheet, which induced a weak and warm polar vortex and therefore a deficit of sea ice characterized by a very weak circumpolar circulation. The energy balance over the ice sheet was thus adjusted to cool the temperature at the surface of the ice sheet, leading to a recovery of the polar vortex and thus a more realistic ice distribution. Some attempts were done as well to adjust the atmospheric radiation and clouds there. In particular, this aspect was looked at when adjusting the fraction of liquid and ice in clouds (Madeleine et al., 2020). However, the lack of reliable observations of surface fluxes over the Antarctic made this exercise somewhat limited.

4.2. Overturning Circulations

In the Southern Hemisphere, the challenge was to ensure realistic deep water formation, which is crucial for biogeochemical cycles (in particular the injection of oxygen and nutrients in the deep ocean). Given the relative high vertical stratification in the Southern Ocean (Boucher et al., 2020) and the coarse spatial resolution over shelf around Antarctica, Antarctic Bottom Water (AABW) formation had to be sustained by opening large polynyas. This was done by reducing to 95% the maximal area covered by sea ice, which

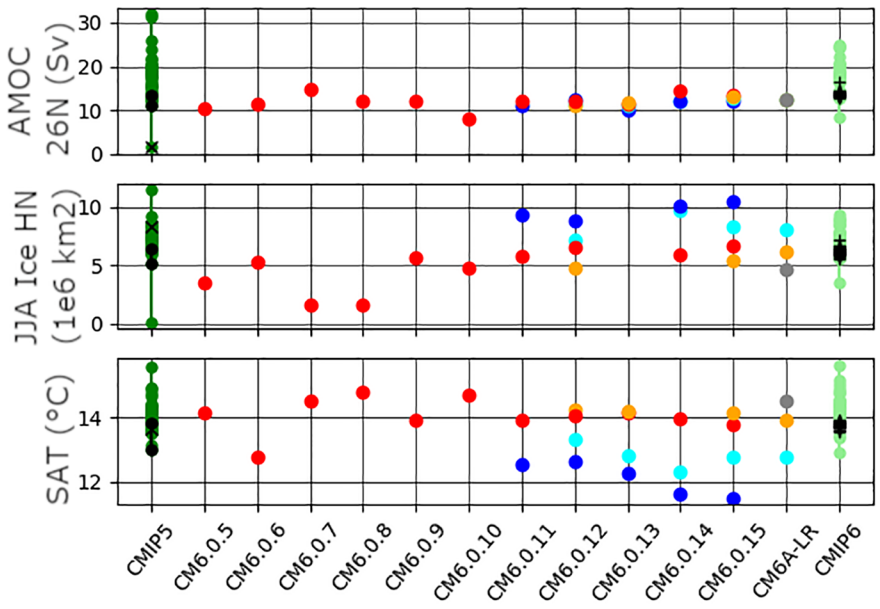

Figure 6. Same as Figure 5 but for the intensity of the AMOC at 26°N (top panel), the summertime (June-July-August) Arctic sea ice extent (middle panel), and the global-mean surface air temperature (bottom panel). Note that the differences in AMOC magnitude in the final IPSL-CM6A-LR versions are indistinguishable.

corresponds to a prescribed increase in the frequency of leads. AABW formation was also increased by injecting freshwater coming from the continental ice-sheet, at depth along the coastline. This strengthened the density gradient from the coast offshore, hence the geostrophic circulation along the coast, preconditioning for dense water formation offshore.

The AMOC was also too weak in the first versions of the model (Figure 6, top panel). Having a more realistic Atlantic overturning circulation was also considered as an important target. However, as already discussed,

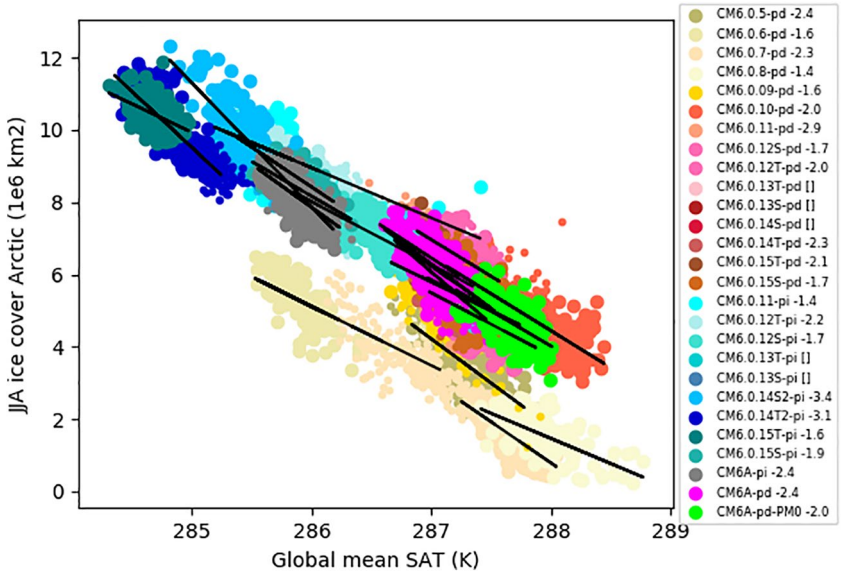

Figure 7. Summer Arctic sea ice area versus global-mean surface air temperature (1 dot per year) in the subsequent versions of IPSL-CM6A-LR throughout the tuning process. In the color code for the different model versions on the right, we used the following abbreviations: pd = "present-day" control (*pdControl*), pi = "preindustrial" control (*piControl*), T = "Ttop" and S = "split". "Ttop" and "split" refer to two different attempts to solve the atmospheric shallow-deep convection issue in the tropics, see text for details).

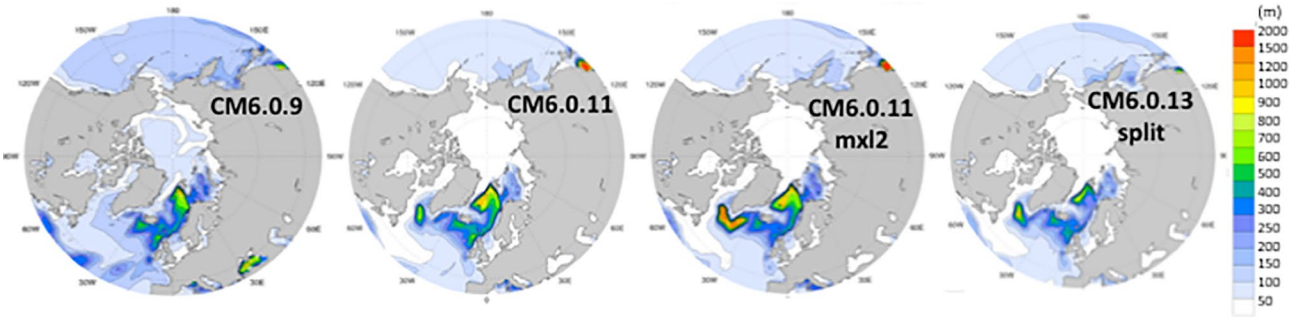

Figure 8. Annual monthly mean mixed layer in the northern part of the North Atlantic Ocean in three successive versions of the IPSL-CM6A-LR. The mixed layer is here defined at the depth where the surface density has decreased by 0.03 $kg\,m^{-3}$, consistently with the most variable defined in (Griffies et al., 2016) for CMIP6. CM6.0.9 and CM6.0.11 and CM6.0.13 are control versions while CM6.0.11 mxl2 is an intermediate version aimed at testing the effect of reducing the penetration length scale of mixing in sea-ice covered grid cells.

the AMOC is poorly constrained by observations. Furthermore, summer Arctic sea ice extent and AMOC strength were strongly tied in early versions of the model, with a strengthened AMOC yielding more poleward heat transport and systematically less Arctic sea ice. While this chain of interactions may have a physical basis (e.g., Zhang, 2015), we had to find ways around it so as to enable a tuning of both of these crucial elements of the climate system. This was partly achieved through the calibration of sea ice parameters discussed above, which helped to increase Arctic sea ice extent without changing global mean SAT and AMOC intensity (Figure 6, after roughly version CM6.0.11). Another clear issue limiting AMOC intensity was the lack of deep convection in the Labrador Sea until version CM6.0.11, because of an accumulation in sea ice. Reducing the penetration length scale of mixing in sea-ice covered grid cells allowed the acceleration of sea ice export, which contributed to resolving this issue (Figure 8).

4.3. Competition Between Shallow and Deep Atmospheric Convection

Overall, the improvement in rainfall and tropical variability was not pushed very far when setting up the IPSL-CM6A-LR configuration, particularly because priority was given to the radiation and SST tuning. It proved difficult to explore the parametric dependence of rainfall at the same time. However, the precipitation distribution and ENSO diagnostics have been reviewed on a regular basis. In particular, two problems were identified up to version CM6.0.9: (i) the amplitude of ENSO was strongly overestimated (Figure 9, left column second panel) and (ii) the frequency of days with precipitation was too large over the ocean (right column second panel). The second point highlighted the possible over-activation of the deep convective scheme. Indeed the competition between deep and shallow convection is key for the representation of precipitation in the tropics (Rio, Del Genio, & Hourdin, 2019).

In the LMDZ physics, shallow and deep convection are represented by two separate schemes. The thermal plume model (Hourdin, Couvreux, & Menut, 2002; Rio, Hourdin, et al., 2010) represents shallow clouds as the saturated part of thermals initiated at the surface and driven by buoyancy. Deep convection is handled by a modified version of the Emanuel (1991) mass-flux scheme. One originality of the LMDZ model is to include a parameterization of cold pools created by re-evaporation of rainfall below cumulonimbus and coupled with the deep convection scheme. The cold pool parameterization splits the grid cell horizontally into two parts: the cold pool area, in which convective precipitation falls and evaporates, and the exterior of cold pools (Grandpeix & Lafore, 2010), in which deep convective cells initiate. Both boundary-layer thermals and cold pools provide a lifting energy and a lifting power used for deep convection triggering and closure, respectively (Rio, Grandpeix, et al., 2013). As a consequence, cold pools stabilize low levels which inhibits turbulence and shallow convection while they also sustain deep convection via their lifting power. As already identified in the CM5B version of the model, deep convection tends to be too frequently active in the model, in turn inhibiting shallow convection.

A first solution to favor shallow convection with respect to deep convection was proposed in version CM6.0.10. It consisted in switching off the deep convection scheme when the altitude of the top of the convective updrafts was not high enough. In practice, this was done by controlling the temperature at the

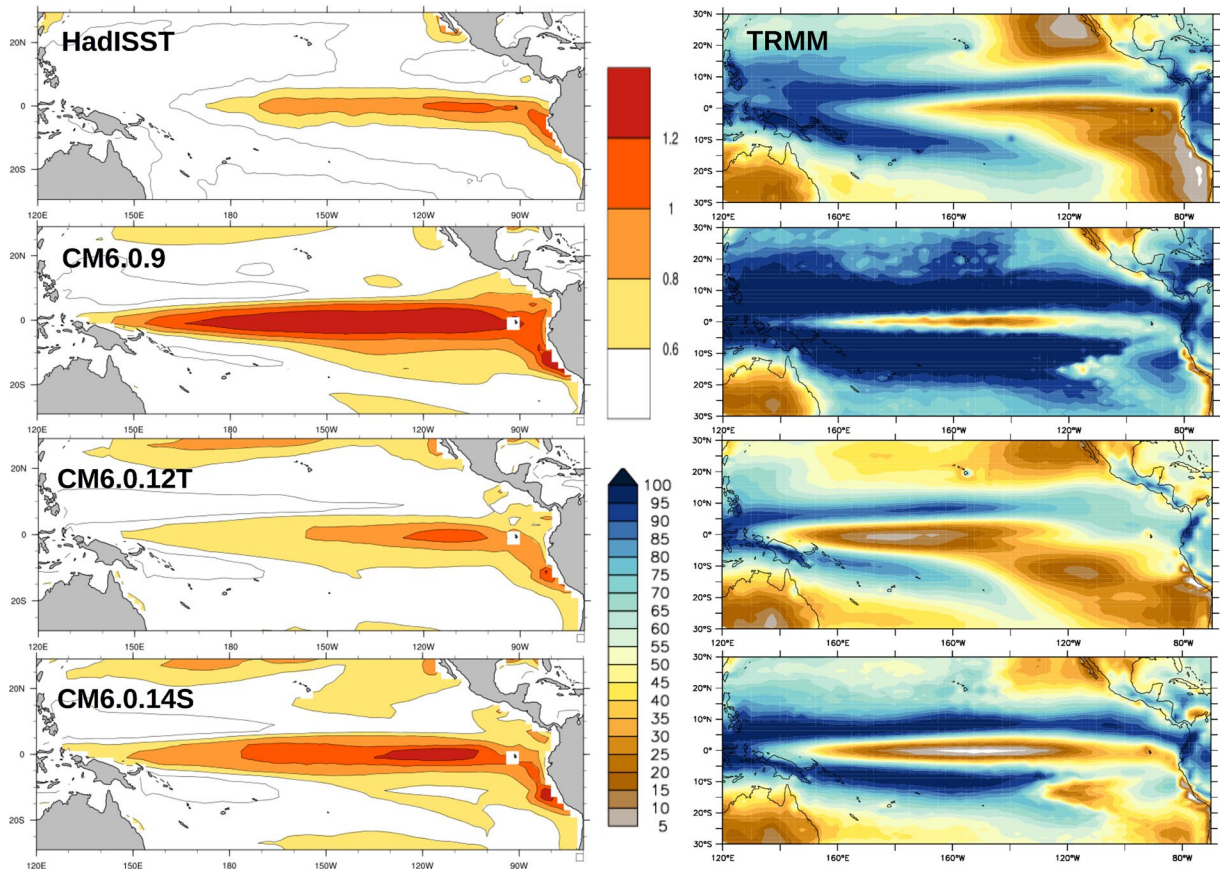

Figure 9. Amplitude of the SST interannual variability in °C computed as the standard deviation of SST annual anomalies (left) and fraction of days with rainfall larger than 0.1 mm. The observations (top panels) correspond to the HadISST dataset over the 1870–2010 period (left) and TRMM daily observations (Huffman et al., 2007) over the period 2000–2009 (right). *pdControl* simulations with three configurations of the IPSL-CM6A-LR model are compared over a period of 80 years for ENSO and 10 years for the Rainfall frequency.

top of the adiabatic ascent computed by the Emanuel deep convection parameterization. If this temperature "Ttop" was higher than 255 K (a threshold chosen after some tests), the deep convection scheme was deactivated. With this value of the temperature threshold, the deep convection scheme only activates for cases where the convection is deeper than typically 5 km in the Tropics. This change had a very positive effect on the two identified problems: it reduced drastically both the frequency of occurrence of deep convection (right column) and ENSO intensity (left column), as illustrated for the CM6.0.12 T version (label "T" standing for Ttop) in the third row of Figure 9. This is consistent with an enhancement of the ENSO amplitude through a larger convection-induced positive Bjerknes feedback. Nevertheless, the exact link between atmospheric convection and ENSO intensity deserves future analysis. Note also that the threshold of 255 K was fixed rather arbitrarily, and could be climate and/or configuration dependent.

4.4. Clouds and SST Biases

Another difficulty was yet to appear. By deactivating the deep convection most of the time over the trade wind regions, the cumulus clouds started to cover a much larger fraction of the tropics inducing a global cooling of the model before any retuning was done. This retuning resulted in lowering the critical value for conversion of cloud liquid water to rainfall, CLC, from 0.15 to 0.075 $g\,kg^{-1}$ to restore a global radiative balance acceptable for the coupled model. By doing so, the brightness of stratocumulus was lowered as well, which in turn reduced the contrast in planetary albedo between cumulus and stratocumulus regions, thus resulting in an increase of the ETO warm bias. This increased bias is responsible for the sudden increase of the SST RMS error between version CM6.0.9 and CM6.0.10, in Figure 5 (bottom panel). The consequence of

this retuning was thus to cancel the improvements of the parameterization of stratocumulus and transition from cumulus to stratocumulus clouds (Hourdin et al., 2019, 2020).

Rather than arbitrarily switching off deep convection using an ad-hoc threshold on Ttop, a more satisfactory solution was found, consisting in activating the thermal plumes outside the cold pools only. The corresponding model versions were labeled "S" as the modification takes into account the "split" of the grid cell between the cold pools and their environment for thermal plumes. In order to give physical ground to this change, we ran specific large eddy simulations (LES) in radiative-convective equilibrium that confirmed that thermal plumes preferentially develop outside cold pools (not shown). In practice, the modification consists of providing the thermal plume model with the profiles of the exterior of cold pools rather than the profiles of the mean grid cell. The exterior of cold pools being generally more unstable than the grid average profile, this modification enables the coexistence of deep and shallow convection and thus reinforces the role of shallow convection. This change was first tested in version CM6.0.11. It produced results intermediate between the CM6.0.9 and Ttop versions in terms of both ENSO amplitude and over-activation of deep convection over oceans as shown for version CM6.0.14S in the fourth row of Figure 9. Retuning the global radiation resulted in a larger value of the CLC parameter of 0.16 $g\,kg^{-1}$, in turn reinforcing the contrast between stratocumulus and cumulus clouds and thus reducing the ETO warm bias (Figure 5 bottom panel, light blue and orange dots from version CM6.0.12 onward).

From version CM6.0.12 to version CM6.0.14, both "Ttop" (red and deep blue circles in 5 and Figure 6) and "split" simulations (orange and light blue circles) were pursued in parallel in both *pdControl* and *piControl* configurations. The reduction of the ETO warm bias in the split versions is clearly visible in the RMS error of the SST pattern between 35°S and 35°N (Figure 5, bottom panel). After retuning, this metric actually improved compared to versions before CM6.0.10, where the Ttop modification was introduced. The near-global SST pattern RMS error decreased as well after version CM6.0.10, and more so for the split than for the Ttop configuration. It was nevertheless not as good as before in version CM6.0.9 for this particular metric, due to the persistence of a classical warm bias around Antarctica and, to a lesser extent, in the North Pacific Ocean.

Because of the reduction of SST biases, a high priority of the model tuning, the "split" configuration was finally preferred to the "Ttop" one. It can be noticed that this choice led to a reduced difference between *piControl* and *pdControl* global-mean surface temperature and SST, as seen by comparing in both the lower panel of Figure 6 and upper panel of Figure 5, the distance between cyan and orange circles (*piControl* and *pdControl* simulations with the "Ttop" version) to the distance between deep blue and red circles (same for the "split" simulations). As discussed above, this probably reflects a larger ECS in the "Ttop" configuration. However, this did not govern the final choice of the "split" configuration.

4.5. Indirect Effect of Aerosols on Global SST

As the configuration was almost frozen, we realized by looking carefully at the model results that with the final tuning of the cloud parameters, the indirect effect of aerosols had almost completely vanished.

Parameters of the relationship between the accumulation-mode soluble aerosol concentration and cloud droplet number concentration (CDNC) in liquid clouds were then used as tuning parameters for the global SST. These parameters were kept within some "reasonable" limits, keeping in mind that they also control the radiative forcing due to aerosol-cloud interactions. The relationship is a revised formulation of Boucher and Lohmann (1995):

$$CDNC = 10^{b_0 + b_1 \log_{10} m_{\mathrm{aer}}} \quad (1)$$

where CDNC is in $particles\,cm^{-3}$ and m_{aer} stands for the mass of soluble accumulation mode aerosols per unit volume given in $\mu g\,m^{-3}$.

Some leeway was deemed possible in the choice of b_0 and b_1 parameters because (i) there is some scatter in this empirical relationship, (ii) the total mass of soluble accumulation mode aerosols is used in the model instead of the sulfate accumulation-mode aerosols only as in the original parameterization, and (iii) we prescribe monthly averages of aerosol concentrations in the model rather than interactive aerosols. The

b_0 parameter controls the average CDNC value, and lower values of b_0 result in larger droplets and thus less bright clouds. Points ii) and iii) provide some justification for a smaller value of b_0, which has indeed been lowered from 2.0 to 1.7 and then 1.3 in the CM6.0.14 and CM6.0.15 configurations, respectively. The radiative forcing due to aerosol-cloud interactions is essentially controlled by the b_1 parameter which has been left unchanged at a value of 0.2. Note however that decreasing b_0 resulted in CDNC reaching its lower default limit of 20 cm^{-3} more often, which has suppressed the aerosol indirect effect because CDNC would be set to its lower limit both with and without anthropogenic aerosols in case of rather pristine clouds. To overcome this issue, the lower limit of CDNC was changed from 20 to 10 cm^{-3} in the final CM6A configuration, a value which is still within observational limits and avoids the anthropogenic effect saturating in clean environments.

This retuning of the aerosol indirect effect between version CM6.0.12 and CM6.0.15 resulted in a colder atmosphere which was compensated by a retuning of CLC from ∼0.2 in 6.0.12 to ∼0.4 $g kg^{-1}$ in CM6.0.14. The final value of CLC was still increased to 0.65 $g kg^{-1}$ in CM6.0.15, as a compromise, to have tropical temperatures a bit too cold compared to present-day observations in order to keep more Arctic sea ice. The magnitude of the instantaneous radiative forcing for aerosol-cloud interactions was thus adjusted to be more negative than −0.3 $W m^{-2}$ (see Lurton et al., 2020), while it was almost zero in the previous versions. This translates to a cooling of roughly 0.45 K.

5. Discussion and Conclusion

We have discussed here the tuning strategy developed at IPSL in order to achieve the development of IPSL-CM6A-LR. This strategy was based on several pillars.

First, each model component had been developed, improved and tuned to some extent separately. Components were then assembled, and final tuning could begin. This “final” process in fact lasted three years, implying back and forth adjustments in (mainly atmospheric) forced and coupled modes.

Second, the strategy was based on tuning the coupled model system in a present-day equilibrium configuration referred to as *pdControl*, in which the sea surface albedo was increased by about 0.7% to offset the need of oceanic heat uptake by about 0.6 $W m^{-2}$. This approach proved its success a posteriori as transient *historical* simulations starting from equivalent preindustrial conditions indeed reached the global mean temperature targeted in the *pdControl* simulations. It allowed the use of reliable and directly comparable observations to assess the model mean state. Nonetheless, while useful and to large extent successful, this protocol is not perfect. It could be refined by imposing the albedo offset as a map rather than a constant, in order to account for the real spatial structure of the oceanic heat uptake, potentially important for the Southern Ocean or ice-covered regions. On the other hand, this may partly neglect horizontal and isopycnal propagation of the heat uptake and may therefore lead to inconsistent heat redistribution within the ocean. The comparison to present-day observations is thus limited by these inaccuracies which we consider nevertheless of the second order. Furthermore, the risk in such refinement is an overfitting to the present-day climate hampering the confidence in the simulation of the future climate. The potential role of internal climate variability for the ocean heat uptake also needs to be properly removed (Gastineau, Mignot, Arzel, & Huck, 2018). Finally, the tuning protocol presented here does not account either for other possible source of inertia in the system, such as that associated with the land surfaces and the melting of the permanent sea ice. Nevertheless, this simple setup provides an easy way to indirectly tune *piControl* and *historical* simulations on present-day observations to a large extent. One advantage of its simplicity is that it could be very easily shared among modeling groups in order to define a common present-day protocol for the CMIP DECK (Diagnostic, Evaluation and Characterization of Klima).

Third, the tuning strategy is based on the release of a relatively large (15) number of versions of the coupled model, each of which being derived from the previous one after parameter calibration and/or code improvement based on the performance of the previous version. This strategy has proven efficiency in reducing some major biases of the model, as detailed in Section 4. The use of global mean metrics all along the evolution of the model versions was also crucial to keep track of the main target variables that were considered as priorities for our modeling group. However, several large scale biases remain (Boucher et al., 2020), such as the strong cold anomaly in the North Atlantic and a relatively warm Southern Ocean at

the surface, associated with an over-stratified thermocline. There are also important biases in precipitation: excess in the eastern part of the Atlantic and Pacific ITCZ as well as over the land of the maritime continent, and not enough around it. These remaining biases result from a failure to fix them with simple parameter calibration, a lack of dedicated effort on a specific topic (global precipitations) and/or the necessity to start the CMIP6 coordinated experiments in time.

The need for such compromises is also due to the fact that tuning strategy described here is still relatively rudimentary. Some of the limits of this iterative strategy appeared through the development of the atmospheric convection scheme, showing that it is not possible to change only one parameter in the model: the global budget always has to be re-equilibrated. This illustrates that optimized components in forced mode cannot be simply plugged together to create an optimal coupled model: Final parameter calibration in coupled mode cannot be skipped. Future studies at IPSL will consist of quantifying the uncertainties associated with the choice of free parameters in a more systematic way using the latest version of the IPSL climate model with three major targets: (i) speed up and improve the calibration of future higher resolution versions, (ii) estimate the errors associated with these parameters in the present-day climate representation (i.e., mean state and variability), and (iii) estimate the associated uncertainty in the ECS. The first point was already partly addressed as we tested the behavior of the final configuration with enhanced atmospheric resolution. The resulting climate mean state was satisfactory, suggesting robustness in the tuning choices (not shown).

Regarding point (iii), we note again that the tuning targeted the present-day temperature taking heat uptake into account and not the historical trend. This clearly differs from the strategy of other modeling groups such as, e.g., Mauritsen and Roeckner (2020) and Danabasoglu et al. (2020). In fact, during the development and tuning phase of the model, and because of an *abrupt-4xCO2* simulation performed with an intermediate version of the model (not shown), it was suspected that the model ECS was in the upper part of the previously estimated range. However, this was never seen as a target for tuning, favoring the idea that ECS should be an emerging property of the model. The amplitude of the historical warming was only somewhat adjusted when it was discovered that the overestimation of the 20th century warming was due in part to the suppression of the aerosol indirect radiative effect after changes in the cloud tuning procedure. The acceptable ranges for the parameters of the aerosol-cloud droplet number concentration relationship were then restricted, as explained in Section 4.5 and in Lurton et al. (2020). However, we did not seek to get a perfect match between the simulated and observed 20th century temperature evolution and indeed most ensemble members tend to warm more than observations during the historical period (not shown).

As discussed in the text, the final choice of the "split" rather than "Ttop" version was primarily motivated by (1) the reduction of the ETO warm bias, (2) the reluctance to introduce a new threshold in the model and, (3) the fact that LES demonstrated that boundary layer convection is indeed much more active outside cold pools. However, the reduction of the difference between the global-mean surface temperature of the corresponding *piControl* and *pdControl* simulations as compared to "Ttop" was seen as good news at that time since the model ECS was suspected to be quite large. Thus it cannot be ruled out that ECS played an indirect role in the collective decision process. It must at least be kept in mind that the "Ttop" version had an even larger ECS than the "split" version which was finally retained.

Another obvious limit of the current tuning procedure is its cost. If the future of climate modeling lies in part in increased resolution (Haarsma et al., 2016), neglecting research in improvements of modeling and tuning efficiency may hamper the potential gain of increased resolution (Stevens & Bony, 2013). Indeed, the long and expensive tuning phase described here cannot be carried on heavier configurations. The IPSL modeling group is developing state-of-the-art machine learning approaches coming from the Uncertainty Quantification community for both tuning and uncertainty quantification future purposes. The idea is to replace the long calibration process described here for which 15 model configurations were re-adjusted based on several dozens of sensitivity experiments performed by varying one parameter at a time, by an automatic random sampling of the full parameter domain. Preliminary tests are promising. Such an approach will enable a more effective and rapid tuning of new versions of the climate model, thus allowing a more continuous development and more intermediate versions than was previously the case.

Appendix A: Global Freshwater Budget

Closure of the water budget has been a constant concern during the development of IPSL-CM6A-LR. First, we have activated or implemented a range of diagnostics in the atmosphere, the ocean and sea ice, and the land surface models in order to diagnose precisely the water reservoirs and their evolution, and the freshwater fluxes between the model components. For each process exchanging freshwater between two model components, we have diagnosed the flux in each of the components so as to diagnose potential errors in the coupling process. This revealed a number of issues with the water conservation. The main corrections during the model development involved the correct positioning of the river mouths onto the ocean grid to ensure a better mass conservation in the interpolation involved in the atmosphere-ocean coupling (version CM6.0.10), the redistribution of the freshwater flux from endorheic basins into the ocean (version CM6.0.11), and a proper accounting of the water vapor condensation flux over land ice (version CM6.0.13). These technical issues are described in Boucher et al. (2020).

Figure A1 summarizes the freshwater fluxes between model components for 200 years of preindustrial control simulation, referred to as *piControl*, of the final configuration of the IPSL-CM6A-LR model. The simulation has been run for a long period, hence the changes in water reservoirs are fairly small albeit not zero as the climate system is affected by internal variability. The net atmospheric flux (precipitation minus evaporation, labeled $P - E$) is computed in the atmospheric model for the different surface types and transferred to the other components. The atmosphere, the land surface and land ice models share the same grid so the freshwater fluxes are perfectly conserved between the atmosphere and the land. The water fluxes between the atmosphere and ocean are dealt with through the Oasis coupler. It should be noted that a part of the snow flux that falls onto the sea ice fraction of an oceanic gridbox is assumed to be blown to the ice-free fraction of that oceanic gridbox (which is always nonzero for numerical reason). Hence the conservation of $P - E$ between the atmosphere and the ocean should not be assessed separately for sea-ice and ice-free fractions of the ocean. Coupling weights between the oceanic and atmospheric grids have been designed such as to ensure flux per unit surface conservation at the ocean-atmosphere interface. Yet, given that the total area of the ocean is not exactly the same in the two models, slight differences remain at the

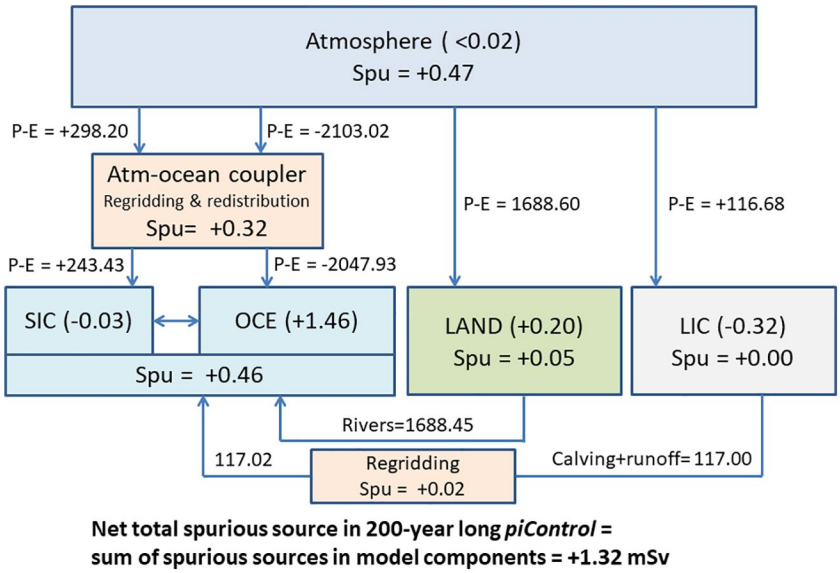

Figure A1. Freshwater fluxes in IPSL-CM6A-LR in mSv (10^3 m^{3}s^{-1}) averaged over 200 years of the model *piControl* simulation close to equilibrium. The fluxes include the net $P - E$ (precipitation minus evaporation) flux, the river flow and the sum of calving and runoff from land ice. Values in parenthesis indicate the rates of change in the various reservoirs over the 200 years simulation. These rates of change are computed as the difference between the final and initial states divided by the length of the simulation and expressed in mSv. We estimate the contribution of each component to the nonconservation of the freshwater budget as the difference between the rate of change and the net flux in a component or in the coupler. These terms are labeled Spu on the diagram, which stands for "spurious source", and are also expressed in mSv. SIC = sea ice, OCE = ice-free ocean, LIC = land ice, LAND = land except land ice.

global scale. The coupler is therefore responsible for a small spurious source of freshwater of 0.32 mSv in its current setup. The land ice model, given its simplicity, has been designed to perfectly conserve water. Again there is a very small spurious source of freshwater of 0.02 mSv associated with the regridding of the calving and runoff terms onto the ocean grid. In contrast to the land ice model, the other model components (the land surface, atmospheric and oceanic models) also suffer from small spurious sources of freshwater, as indicated in Figure A1. For the atmosphere, this appears to be due to the large-scale advection of water species in LMDZ which is not perfectly conservative (not shown). For the ocean, the apparent spurious source may be an artifact due to the fact that we approximate the freshwater reservoirs from the first and last months of the simulation (instead of from the initial and final states). The sum of all spurious sources of freshwater in the coupled model is of the order of 1 mSv, which corresponds to about 9 mm of sea level rise per century. Such a drift is considered acceptable in particular as compared to the observed 18 cm of sea level rise since 1900, even considering the only ∼4 cm attributed to thermal expansion (e.g., Frederikse et al., 2020). Even though it is weak, this drift should be reduced in the future.

Appendix B: Global Oceanic Heat Budget

Diagnosing the sum of all heat fluxes received by the ocean component (Figure B1) enables to evaluate the closure of the energy budget.

Heat content of the ocean is calculated by integrating the conservative temperature of ocean water over the entire volume. Evolution of the latter over the 200 years of the *piControl* sample is computed as the difference between the last 3 h and the first 3 h of the period. It is denoted as Δ(OHC), and converted into $\mathrm{W\,m^{-2}}$ (dividing by the area of the ocean surface). Total incoming fluxes is the sum of all heat fluxes reaching the ocean, averaged every year and integrated over the whole oceanic domain. These fluxes reach the ocean (i) from the atmosphere, which includes radiative and turbulent fluxes, and heat fluxes associated to precipitation heat content; this term also contains the heat exchanged between the sea ice and the ocean in ice-covered areas; (ii) through runoffs; (iii) through ice shelf melting, via icebergs (denoted as calving) and directly along the coastline of Antarctica (denoted as Antarctica ice shelf melting); (iv) through geothermal heating. All fluxes mentioned above are calculated online (see, Boucher et al., 2020 for details).

In IPSL-CM6A-LR, after more than 3,000 years of integration of *piControl* simulation, the sum of all heat fluxes diagnosed to reach the ocean is negative (mean = $-0.136\ \mathrm{W\,m^{-2}}$) The fact that this sum does not equal 0, as should be the case in a perfectly stationary simulation may have several explanations (uncertainty in the diagnostics and/or incomplete spinup) not investigated here. This imbalance is consistent with the decrease in ocean heat content (Δ[OHC]) and translates into an almost linear drift in ocean heat content during the standard CMIP6 *piControl* simulation. Note that the same computation including the sea ice component yields practically the same heat content drift (mean = $-0.135\ \mathrm{W\,m^{-2}}$), as the ice heat content is practically constant. However, the sum of incoming fluxes to ocean and sea ice together amounts

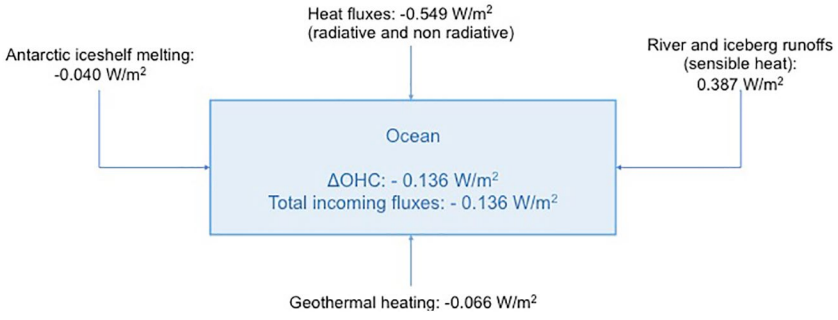

Figure B1. Heat fluxes (in $\mathrm{W\,m^{-2}}$) into the oceanic component of IPSL-CM6A-LR averaged over 200 years of the model *piControl* simulation, after more than 3,000 years of integration from initial conditions. Δ(OHC) corresponds to the trend in heat content in ocean only, computed as the difference between the last 3 h minus the first 3 h of the 200 years, while Total incoming fluxes sums all heat fluxes into ocean, yearly averaged. Heat fluxes associated to volume flux of water (runoff and precipitation) are calculated using a temperature reference of 0°C and water flowing from the Antarctic ice shelf is supposed to reach the ocean at −1.9°C.

to $-0.032\ \mathrm{W\,m^{-2}}$. This discrepancy remains unexplained. Note also that the diagnostic performed in Irving et al. (2021) is erroneous, certainly due to errors committed in the CMIP6 diagnostics of heat fluxes, as the liquid oceanic heat budget in IPSL-CM6A-LR is closed.

The total anomalous heat sink in the ocean is primarily dominated by the anomalous upward heat flux due to turbulent and radiative fluxes (mean = $-0.549\ \mathrm{W\,m^{-2}}$), while anomalous heat fluxes associated with runoffs are downward (mean = $0.387\ \mathrm{W\,m^{-2}}$). The other terms are smaller, with the contribution of ice-shelf melting cooling the ocean ($-0.040\ \mathrm{W.m^{-2}}$), and that of the geothermal heat flux ($0.066\ \mathrm{W.m^{-2}}$). This reflects the variety of processes associated with the heat flux exchanged between the ocean, sea ice and the atmosphere, while the heat flux due to runoff and land ice depend on the average freshwater cycle. The large variability of heat exchanges with atmosphere (std = $0.339\ \mathrm{W\,m^{-2}}$), not shown, while all other heat fluxes have much reduced variability (2 orders of magnitude smaller, not shown).

Conflict of Interest

The authors do not declare any competing interests.

Data Availability Statement

LMDZ, XIOS, NEMO, and ORCHIDEE are released under the terms of the CeCILL license. OASIS -MCT is released under the terms of the Lesser GNU General Public License (LGPL). IPSL -CM6A-LR code is publicly available through svn, with the following command lines: svn co https://forge.ipsl.jussieu.fr/igcmg/browser/modipsl/branches/publications/IPSLCM6.1.11-LR_05012021 modipsl cd modipsl/util; ./model IPSLCM6.1.11-LR. The mod.def file provides information regarding the different revisions used, namely: (1) NEMOGCM branch nemov36STABLE revision 9455; (2) XIOS2 branchs/xios-2.5 revision 1873; (3) IOIPSL/src svn tags/v224; (4) LMDZ6 branches/IPSLCM6.0.15 rev 3643; (5) tags/ORCHIDEE20/ORCHIDEE revision 6592; (6) OASIS3-MCT 2.0branch (rev 4775 IPSL server). The login/password combination requested at first use to download the ORCHIDEE component is anonymous/anonymous. We recommend to refer to the project website: http://forge.ipsl.jussieu.fr/igcmg_doc/wiki/Doc/Config/IPSLCM6 for a proper installation and compilation of the environment. The python code developed for Figure 1 and 3 and 5–7 is available under https://zenodo.org/record/4415818. The preprocessed data sets used for Figures 5–7 and partly Figure 1 are available https://zenodo.org/record/4415805. Figure 1 uses CMIP6 outputs available for all simulations except *pdControl* and *pdControl PM0* which are available here: https://zenodo.org/record/4415805. Figures 8 and 9 were made using the CliMAF Python library (Climate Model Assessment Framework, https://github.com/rigoudyg/climaf).

Acknowledgments
The authors are grateful to the full IPSL Climate Modeling Community (ICMC) for discussions, interactions and feedbacks. This work was undertaken in the framework of the LABEX L-IPSL and EUR Climate Graduate School. It benefited from the ANR funded project Convergence (grant ANR-13-MONU-008) and the "Investissements d'avenir" program with the reference ANR-11-IDEX-0004 - 17-EURE-0006. The CMIP6 project at IPSL used the HPC resources of TGCC under the allocations 2016-A0030107732, 2017-R0040110492, and 2018-R0040110492 (project gencmip6) provided by GENCI (Grand Équipement National de Calcul Intensif). This study benefited from the ESPRI (Ensemble de Services Pour la Recherche à l'IPSL) computing and data center (https://mesocentre.ipsl.fr) which was supported by CNRS, Sorbonne Université, École Polytechnique and CNES and through national and international grants. Support from the European Commission's Horizon 2020 Framework Program is acknowledged, under Grant Agreement number 641816 for the "Coordinated Research in Earth Systems and Climate: Experiments, kNowledge, Dissemination and Outreach (CRESCENDO)" project (11/2015-10/2020). J.M. and G.G. were also funded by the EU-H2020 Blue Action (Grant Agreement no. 727852) and EUCP (Grant Agreement no 776613) Research Programmes. J.M. was also funded by ARCHANGE ANR-18-MP-GA-0001. M.M. was supported by the EPICE project funded by the European Union's Horizon 2020 program, Grant Agreement 789445.

References

Bastrikov, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. https://doi.org/10.1029/2019MS001916

Boucher, O., & Lohmann, U. (1995). The sulfate-CCN-cloud albedo effect. *Tellus B: Chemical and Physical Meteorology*, *47*(3), 281–300. https://doi.org/10.3402/tellusb.v47i3.1604810.1034/j.1600-0889.47.issue3.1.x

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., et al. (2020). Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS002010. https://doi.org/10.1029/2019MS002010

Brient, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., et al. (2017). The art and science of climate model tuning. *Bulletin of the American Meteorological Society*, *98*(3), 589–602. https://doi.org/10.1175/BAMS-D-15-00135.1

Cheruy, F., Ducharne, A., Hourdin, F., Musat, I., Vignon, É., Gastineau, G., et al. (2020). Improved Near-Surface Continental Climate in IPSL-CM6A-LR by combined evolutions of atmospheric and land surface physics. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS002005. https://doi.org/10.1029/2019MS002005

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The Community Earth System Model952 Version 2 (CESM2). *Journal of Advance Modeling Earth Systems*, *12*(2), e2019MS001916. https://doi.org/10.1029/2019MS001916954

Diallo, F. B., Hourdin, F., Rio, C., Traore, A.-K., Mellul, L., Guichard, F., & Kergoat, L. (2017). The surface energy budget computed at the grid-scale of a climate model challenged by station data in West Africa. *Journal of Advances in Modeling Earth Systems*, *9*(7), 2710–2738. https://doi.org/10.1002/2017MS001081

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., et al. (2020). The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2019MS002015. https://doi.org/10.1029/2019MS002015

Emanuel, K. A. (1991). A scheme for representing cumulus convection in large-scale models. *Journal of the Atmospheric Sciences*, *48*(21), 2313–2329. https://doi.org/10.1175/1520-0469(1991)048⟨2313:ASFRCC⟩2.0.CO;2

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

Frederikse, T., Landerer, F., Caron, L., Adhikari, S., Parkes, D., Humphrey, V. W., et al. (2020). The causes of sea-level rise since 1900. *Nature*, *584*(7821), 393–397. https://doi.org/10.1038/s41586-020-2591-3

Gastineau, G., Mignot, J., Arzel, O., & Huck, T. (2018). North Atlantic Ocean internal decadal variability: Role of the mean state and ocean-atmosphere coupling. *Journal of Geophysical Research: Oceans*, *123*, 5949–5970. https://doi.org/10.1029/2018JC014074

Gastineau, G., Mignot, J., Lott, F., & Hourdin, F. (2020). Impact of unresolved orography for the Northern Hemisphere climate in a coupled model. *Journal of Advances in Modeling Earth Systems*. submitted.

Grandpeix, J.-Y., & Lafore, J.-P. (2010).A density current parameterization coupled with Emanuel's convection scheme. Part I: The models. *Journal of the Atmospheric Sciences*, *67*, 881–897. https://doi.org/10.1175/2009JAS3044.1

Griffies, S. M., Danabasoglu, G., Durack, P. J., Adcroft, A. J., Balaji, V., Böning, C. W., et al. (2016). Omip contribution to CMIP6: Experimental and diagnostic protocol for the physical component of the ocean model intercomparison project. *Geoscientific Model Development*, *9*(9), 3231–3296. https://doi.org/10.5194/gmd-9-3231-2016

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

Hobbs, W., Palmer, M. D., & Monselesan, D. (2016). An energy conservation analysis of ocean drift in the CMIP5 Global Coupled Models*. *Journal of Climate*, *29*(5), 1639–1653. https://doi.org/10.1175/JCLI-D-15-0477.1

Hourdin, F., Couvreux, F., & Menut, L. (2002). Parameterization of the dry convective boundary layer based on a mass flux representation of thermals. *Journal of the Atmospheric Sciences*, *59*, 1105–1123. https://doi.org/10.1175/1520-0469(2002)059⟨1105:POTDCB⟩2.0.CO;2

Hourdin, F., Grandpeix, J.-Y., Rio, C., Bony, S., Jam, A., Cheruy, F., et al. (2013). LMDZ5B: The atmospheric component of the IPSL climate model with revisited parameterizations for clouds and convection. *Climate Dynamics*, *40*, 2193–2222. https://doi.org/10.1007/s00382-012-1343-y

Hourdin, F., Rio, C., Grandpeix, J. Y., Madeleine, J. B., Cheruy, F., Rochetin, N., et al. (2020). LMDZ6A: The atmospheric component of the IPSL climate model with improved and better tuned physics. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001892. https://doi.org/10.1029/2019MS001892

Hourdin, F., Rio, C., Jam, A., Traore, A. K., & Musat, I. (2020). Convective boundary layer control of the sea surface temperature in the tropics. *Journal of Advances in Modeling Earth Systems*, *12*(6), e01988. https://doi.org/10.1029/2019MS001988

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., et al. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology*, *8*(1), 38–55. https://doi.org/10.1175/JHM560.1

Irving, D., Hobbs, W., Church, J., & Zika, J. (2020). A mass and energy conservation analysis of drift in the CMIP6 ensemble. *Journal of Climate*, *34*(8), 1–43. https://doi.org/10.1175/JCLI-D-20-0281.1

Kuhlbrodt, T., Jones, C. G., Sellar, A., Storkey, D., Blockley, E., Stringer, M., et al. (2018). The Low-Resolution Version of HadGEM3 GC3.1: Development and evaluation for global climate. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2865–2888. https://doi.org/10.1029/2018MS001370

Locarnini, R., Mishonov, A., Antonov, J., Boyer, T., Garcia, H., Baranova, O., et al. (2013). *World ocean atlas 2013, volume 1: Temperature*. Technical Report (Vol. 73). NOAA Atlas NESDIS.

Lurton, T., Balkanski, Y., Bastrikov, V., Bekki, S., Bopp, L., Braconnot, P., et al. (2020). Implementation of the CMIP6 Forcing Data in the IPSL-CM6A-LR Model. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001940. https://doi.org/10.1029/2019MS001940

Madeleine, J. B., Hourdin, F., Grandpeix, J. Y., Rio, C., Dufresne, J. L., Vignon, E., et al. (2020). Improved representation of clouds in the atmospheric component LMDZ6A of the IPSL-CM6A Earth System Model. *Journal of Advances in Modeling Earth Systems*, *12*, e2020MS002046. https://doi.org/10.1029/2020MS002046

Mahlstein, I., & Knutti, R. (2012). September Arctic sea ice predicted to disappear near 2°C global warming above present. *Journal of Geophysical Research*, *117*, D06104. https://doi.org/10.1029/2011JD016709

Mauritsen, T., & Roeckner, E. (2020). Tuning the MPI-ESM1.2 global climate model to improve the match with instrumental record warming by lowering its climate sensitivity. *Journal of Advances in Modeling Earth Systems*, *12*(5), e2019MS002037. https://doi.org/10.1029/2019MS002037

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

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., et al. (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances*, *6*(26), eaba1981. https://doi.org/10.1126/sciadv.aba1981

Menary, M. B., Hodson, D. L. R., Robson, J. I., Sutton, R. T., Wood, R. A., & Hunt, J. A. (2015). Exploring the impact of CMIP5 model biases on the simulation of North Atlantic decadal variability. *Geophysical Research Letters*, *42*, 5926–5934. https://doi.org/10.1002/2015GL064360

Menary, M. B., Kuhlbrodt, T., Ridley, J., Andrews, M. B., Dimdore-Miles, O. B., Deshayes, J., et al. (2018). Preindustrial Control Simulations With HadGEM3-GC3.1 for CMIP6. *Journal of Advances in Modeling Earth Systems*, *10*(12), 3049–3075. https://doi.org/10.1029/2018MS001495

Notz, D., Haumann, F. A., Haak, H., Jungclaus, J. H., & Marotzke, J. (2013). Arctic sea-ice evolution as modeled by Max Planck Institute for meteorology's Earth system model. *Journal of Advances in Modeling Earth Systems*, *5*(2), 173–194. https://doi.org/10.1002/jame.20016

Randall, D., Wood, R., Bony, S., Colman, R., Fichefet, T., Fyfe, J., et al. (2007). *Climate models and their evaluation. In Climate change 2007: The physical science basis. contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change*. Cambridge University Press.

Rio, C., Del Genio, A. D., & Hourdin, F. (2019). Ongoing breakthroughs in convective parameterization. *Current Climate Change Reports*, *5*(2), 95–111. https://doi.org/10.1007/s40641-019-00127-w

Rio, C., Grandpeix, J.-Y., Hourdin, F., Guichard, F., Couvreux, F., Lafore, J.-P., et al. (2013). Control of deep convection by sub-cloud lifting processes: The ALP closure in the LMDZ5B general circulation model. *Climate Dynamics*, *40*, 2271–2292. https://doi.org/10.1007/s00382-012-1506-x

Rio, C., Hourdin, F., Couvreux, F., & Jam, A. (2010). Resolved versus parametrized boundary-layer plumes. Part II: Continuous formulations of mixing rates for mass-flux schemes. *Boundary-Layer Meteorology*, *135*, 469–483. https://doi.org/10.1007/s10546-010-9478-z

Rochetin, F., Jam, A., Rio, C., Couvreux, F., Sandu, I., Lefebvre, M. P., et al. (2019). Unified parameterization of convective boundary layer transport and clouds with the thermal plume model. *Journal of Advances in Modeling Earth Systems*, *11*(9), 2910–2933. https://doi.org/10.1029/2019MS001666

Rochetin, N., Couvreux, F., Grandpeix, J.-Y., & Rio, C. (2014). Deep convection triggering by boundary layer thermals. Part I: LES analysis and stochastic triggering formulation. *Journal of the Atmospheric Sciences*, *71*, 496–514. https://doi.org/10.1175/JAS-D-12-0336.1

Rochetin, N., Grandpeix, J.-Y., Rio, C., & Couvreux, F. (2014). Deep convection triggering by boundary layer thermals. Part II: Stochastic triggering parameterization for the LMDZ GCM. *Journal of the Atmospheric Sciences*, *71*, 515–538. https://doi.org/10.1175/JAS-D-12-0337.1

Rousset, C., Vancoppenolle, M., Madec, G., Fichefet, T., Flavoni, S., Barthélemy, A., et al. (2015). The Louvain-La-Neuve sea ice model LIM3.6: Global and regional capabilities. *Geoscientific Model Development*, *8*(10), 2991–3005. https://doi.org/10.5194/gmd-8-2991-2015

Sausen, R., Hasselmann, K., & Hasselmann, K. (1988). Coupled ocean-atmosphere models with flux correction. *Climate Dynamics*, *2*, 145–163. https://doi.org/10.1007/BF01053472

Schmidt, G. A., Bader, D., Donner, L. J., Elsaesser, G. S., Golaz, J.-C., Hannay, C., et al. (2017). Practice and philosophy of climate model tuning across six US modeling centers. *Geoscientific Model Development*, *10*(9), 3207–3223. https://doi.org/10.5194/gmd-10-3207-2017

Senior, C. A., Jones, C. G., Wood, R. A., Sellar, A., Belcher, S., Klein-Tank, A., et al. (2020). U.K. Community Earth System Modeling for CMIP6. *Journal of Advances in Modeling Earth Systems*, *12*(9), e2019MS002004. https://doi.org/10.1029/2019MS002004

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., et al. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, *58*(4), e2019RG000678.

SIMIP Community (2020). Arctic sea ice in CMIP6. *Geophysical Research Letters*, *47*, e2019GL086749. https://doi.org/10.1029/2019GL086749

Stevens, B., & Bony, S. (2013). What are climate models missing? *Science*, *340*(6136), 1053–1054.

Trenberth, K. E., Fasullo, J. T., & Kiehl, J. (2009). Earth's Global Energy Budget. *Bulletin of the American Meteorological Society*, *90*(3), 311–324. https://doi.org/10.1175/2008BAMS2634.1

Uotila, P., Iovino, D., Vancoppenolle, M., Lensu, M., & Rousset, C. (2017). Comparing sea ice, hydrography and circulation between NEMO3.6 LIM3 and LIM2. *Geoscientific Model Development*, *10*, 1009–1031. https://doi.org/10.5194/gmd-10-1009-2017

Vancoppenolle, M., Fichefet, T., Goosse, H., Bouillon, S., Madec, G., & Morales Maqueda, M. A. (2009). Simulating the mass balance and salinity of Arctic and Antarctic sea ice. 2. Importance of sea ice salinity variations. *Ocean Modelling*, *27*, 54–69. https://doi.org/10.1016/j.ocemod.2008.11.003

Vignon, E., Hourdin, F., Genthon, C., Gallée, H., Bazile, E., Lefebvre, M.-P., et al. (2017). Antarctic boundary layer parametrization in a general circulation model: 1-D simulations facing summer observations at Dome C. *Journal of Geophysical Research: Atmospheres*, *122*, 6818–6843. https://doi.org/10.1002/2017JD026802

Vignon, E., Hourdin, F., Genthon, C., Van de Wiel, B. J. H., Gallée, H., Madeleine, J. B., & Beaumet, J. (2018). Modeling the dynamics of the atmospheric boundary layer over the Antarctic plateau with a general circulation model. *Journal of Advances in Modeling Earth Systems*, *10*, 98–125. https://doi.org/10.1002/2017MS001184

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., et al. (2019). Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2177–2213. https://doi.org/10.1029/2019MS001683

Von Schuckmann, K., Palmer, M. D., Trenberth, K. E., Cazenave, A., Chambers, D., Champollion, N., et al. (2016). An imperative to monitor Earth's energy imbalance. *Nature Climate Change*, *6*(2), 138–144. https://doi.org/10.1038/nclimate2876

Williamson, D. L., Olson, J. G., Hannay, C., Toniazzo, T., Taylor, M., & Yudin, V. (2015). Energy considerations in the Community Atmosphere Model (CAM). *Journal of Advances in Modeling Earth Systems*, *7*(3), 1178–1188. https://doi.org/10.1002/2015MS000448

Wyser, K., van Noije, T., Yang, S., von Hardenberg, J., O'Donnell, D., & Döscher, R. (2020). On the increased climate sensitivity in the EC-earth model from cmip5 to cmip6. *Geoscientific Model Development*, *13*(8), 3465–3474. https://doi.org/10.5194/gmd-13-3465-2020

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

Zhang, R. (2015). Mechanisms for low-frequency variability of summer Arctic sea ice extent. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(15), 4570–4575. https://doi.org/10.1073/pnas.1422296112