A modeling framework to understand transient ocean climate change in large coupled ensembles

Yona Silvy[1], Clément Rousset[1], Eric Guilyardi[1,2], Jean-Baptiste Sallée[1], Juliette Mignot[1], Christian Ethé[3], and Gurvan Madec[1]

[1]LOCEAN-IPSL, Laboratoire d'Océanographie et du Climat: Expérimentation et Approches Numériques, Sorbonne Université/CNRS/IRD/MNHN, Paris, France
[2]NCAS-Climate, University of Reading, Reading, UK
[3]IPSL, Institut Pierre-Simon Laplace, Sorbonne Université, CNRS, Paris, France

Correspondence: Yona Silvy (yona.silvy@locean.ipsl.fr)

Abstract. The ocean responds to climate change through modifications of heat, freshwater and momentum fluxes at its boundaries. The role of these contributors in changing the thermohaline structure of the ocean and its circulation has been partly addressed by modeling studies using idealized CO_2 forcings. The question of timescales for these individual contributions during transient climate change is however lacking. Here, we propose a novel modeling framework to isolate these contributions during the entire historical period and projections of coupled climate models. We present the framework in the context of the IPSL-CM6A-LR model and its ocean component NEMO3.6. We start by reproducing a coupled pre-industrial control simulation with an ocean-only configuration, forced by fixed fluxes at its interface diagnosed from the coupled model. We then add a perturbation to each flux component, extracted from the historical+ssp ensemble of simulations of IPSL-CM6A-LR. With this configuration, we successfully replicate the ocean's response to transient climate change in the coupled model during 1850-2100. This full response is then decomposed in sensitivity experiments in which the perturbations are applied individually to the heat, freshwater and momentum fluxes. Passive tracers of temperature and salinity are implemented to discriminate the addition of heat and freshwater flux anomalies in the ocean from the redistribution of pre-industrial heat and salt content in response to ocean circulation changes. This framework brings new opportunities to precisely explore the mechanisms driving transient ocean changes within single climate models.

1 Introduction

The thermohaline structure of the ocean has been particularly affected by human-made climate change: observations of the upper ocean since the mid-20th century show an unabated warming, and large-scale salinity changes coherent with an intensification of the hydrological cycle (Fox-Kemper et al., 2021). These changes are consistent with the effects of rising greenhouse gas concentrations in the atmosphere, and are projected to amplify in response to continuing emissions (Fox-Kemper et al., 2021). The spatial patterns of ocean temperature and salinity modifications in response to this forcing are important to understand as they have widespread consequences, including regional sea level rise, one of the greatest threats associated to climate change for human societies and ecosystems living in coastal areas (IPCC, 2022). Importantly, these spatial patterns are set not

only via changes in local heat and freshwater fluxes, but also via changes in large-scale ocean circulation, which redistribute the heat and salt internally. However, the combination of physical drivers causing these changes and their spatial distribution is still unclear.

Ocean general circulation models or climate models are useful tools to investigate physical processes as they allow for different hypotheses to be tested, by, for example, decomposing potential physical drivers. Numerical experiments have thus been designed to explore the role of individual surface fluxes and/or ocean circulation changes in driving regional ocean heat content change in response to rising CO_2 atmospheric concentrations (e.g. Mikolajewicz and Voss (2000); Banks and Gregory (2006); Fyfe et al. (2007); Xie and Vallis (2012); Winton et al. (2013); Marshall et al. (2015); Gregory et al. (2016); Armour et al. (2016); Garuba and Klinger (2016, 2018); Liu et al. (2018); Zanna et al. (2019); Todd et al. (2020); Hu et al. (2020)). Other studies have also explored how salinity patterns were driven by a combination of changes in these surface fluxes (Lago et al., 2016; Zika et al., 2018; Shi et al., 2020). A gap however remains regarding the decomposition of mechanisms during the transient response to climate change. Indeed, most of the aforementioned studies looked at the response to idealized forcings and/or at multi-decadal to centennial scale. Therefore, the processes driving transient change, which cause the climate signal to emerge from background climate variations, remain unclear.

Here, we propose a novel modeling framework aiming at understanding the leading ocean processes responsible for climate change signals to emerge from climate variability in the ocean during the transient response to anthropogenic forcings. Specifically, we aim to delineate how changes in surface heat fluxes, freshwater fluxes and winds affect the patterns and timescales of ocean temperature and salinity changes. One of the issues will be to also track their relation to changes in the large-scale ocean circulation. The targeted time period for our investigation ranges from the pre-industrial period to the end of the projected 21st century, i.e. 1850-2100, under realistic emissions pathways. We designed numerical experiments in a stand-alone configuration of an ocean general circulation model. Air-sea flux perturbations during the period 1850-2100 are imposed based on the ensemble of historical+scenario simulations of the parent coupled model. The fluxes may be imposed together or separately. This allows the investigation of the individual role of surface flux changes in modifying ocean temperature and salinity. Furthermore, to separate patterns of change due to the perturbed circulation from those due to passive transport of heat and salt into the ocean, we implemented passive tracers of temperature and salinity forced by identical surface flux perturbations as their prognostic counterparts. In section 2, we describe the model and the experimental design. In sections 3, 4, and 5, we describe the implementation steps of these experiments, before concluding in section 6. Companion papers are underway to present the results of these simulations.

2 Model and experimental design

2.1 Coupled model and ocean component

We use the IPSL-CM6A-LR coupled model (Boucher et al., 2020) developed by the Institut Pierre-Simon Laplace modelling center for the 6th phase of the Coupled Model Intercomparison Project (CMIP6, Eyring et al. (2016)). It is composed of the LMDZ6A atmospheric model (Hourdin et al., 2020), the ORCHIDEE land surface model (Krinner et al., 2005) version 2.0

and the NEMO3.6 ocean model (Madec et al., 2017). The atmospheric component has a horizontal resolution of 2.5ºx1.3º on a regular latitude-longitude grid and 79 vertical layers, while the ocean component uses the eORCA1 tripolar grid with a nominal horizontal resolution of 1º refined to 1/3º at the equator, with 75 vertical levels with varying thicknesses (1m at surface to 200m at deepest levels). Indeed the vertical layers are time dependent, with a nonlinear evolving free surface using a variable volume

60 formulation. The ocean physics component (OPA) of NEMO3.6 is coupled to the LIM3 sea-ice model (Rousset et al., 2015) and to the PISCES-v2 biogeochemical model (Aumont et al., 2015). The oceanic equation of state is estimated with a polynomial representation of TEOS-10 (Roquet et al., 2015); the model prognostic fields are thus conservative temperature and absolute salinity. The different schemes and parameterizations employed in the eORCA1 configuration and used in IPSL-CM6A-LR are described in Boucher et al. (2020) and additional details can be found in Madec et al. (2017).

65 We also use the ocean physics component in stand-alone mode, without the sea-ice nor the biogeochemistry, with the same configuration eORCA1 as in the coupled model.

2.2 Experimental design

This work is set within the simulated climate of the IPSL-CM6A-LR coupled model. For the purpose of the CMIP6 exercise, multiple experiments were already conducted with this coupled model including: 2000 years of pre-industrial climate known

70 as "piControl" simulation (after a long spin-up), 32 members of the historical period (1850-2014) - extended to 2059 under the ssp245 scenario - and projections (2015-2100) with only 11 members. Our goal here is to investigate and isolate the mechanisms responsible for temperature and salinity emergence from natural variability in the IPSL-CM6A-LR historical+ssp245 ensemble. To do so, the target is to reproduce the 1850-2100 ocean response of the IPSL-CM6A-LR historical+ssp245 ensemble, and then run sensitivity experiments by applying the perturbations in surface heat, freshwater and wind stress fluxes individually

75 on the ocean. In these historical+ssp simulations, internal variability can play an important role in determining the timing of emergence of the anthropogenic signal, even within members of the same model ensemble (Lehner et al., 2017; Silvy et al., 2020). Yet in this study, it is crucial to compare simulations that have the same background internal variability, i.e. that have concurrent phases of unforced variability intrinsic to the climate system. Thus, we choose here to force the ocean with fixed fluxes instead of coupling it with an atmospheric model. It prevents retroactions between the ocean and the surface fluxes that

80 would generate different phases of internal variability. As a co-benefit, running the ocean model in a stand-alone configuration is computationally much less expensive than running a coupled model.

We first set up an ocean-only CTL experiment forced with fluxes from a pre-industrial control simulation (piControl, constant external forcings) of the IPSL-CM6A-LR coupled model (see the simplified schematic of the experiments in Fig. 1). The goal for this ocean-only CTL is to inherit the mean climate and internal variability of the coupled model piControl, thereby providing

85 a background climate with the same phases of variability for the sensitivity experiments. The piControl fluxes are thus imposed at high frequency (3-hourly) at the liquid ocean interface (below the atmosphere and the sea-ice, see Fig. 2), during 251 years. Since the fluxes are extracted below sea-ice and imposed on the ocean, the sea-ice component is excluded from the oceanic configuration. More details on the forcing methodology are given in the next sections. We then run perturbation experiments, forced with the same fluxes as CTL, plus a perturbation component that represents the effect of transient climate change on the

90   various components of the surface fluxes (Fig. 1). The perturbations are constructed from the monthly-mean anomalies (relative
      to 1850-1899) of the ensemble mean of the IPSL-CM6A-LR historical+ssp245 experiments (32 members over 1850-2059, 11
      members over 2060-2100). The latter is the best possible estimate of the externally-forced signal over the 1850-2100 period in
      the model. We admit that this reduction of ensemble members after 2060 could hamper this estimate after this date. However,
      no significant step in mean and variability of the forcing has been identified (not shown). The ALL experiment includes the
95   perturbation on all fluxes (heat, freshwater+salt, wind stress). It is designed to reproduce the historical+ssp245 forced response.
      It is initialized as any historical member from a parent piControl state, with its own internal variability. The HEAT, STRESS
      and WATER experiments are similar but only include the perturbation component on respectively the heat fluxes, wind stress,
      and freshwater+salt fluxes. The BUOY experiment includes both the heat flux and the freshwater+salt flux perturbations. All
      experiments are thus aligned on the piControl internal variability, with the additional perturbation (forced) components slowly
100   taking effect. The sensitivity experiments decompose the response of the ALL experiment while keeping identical phases of
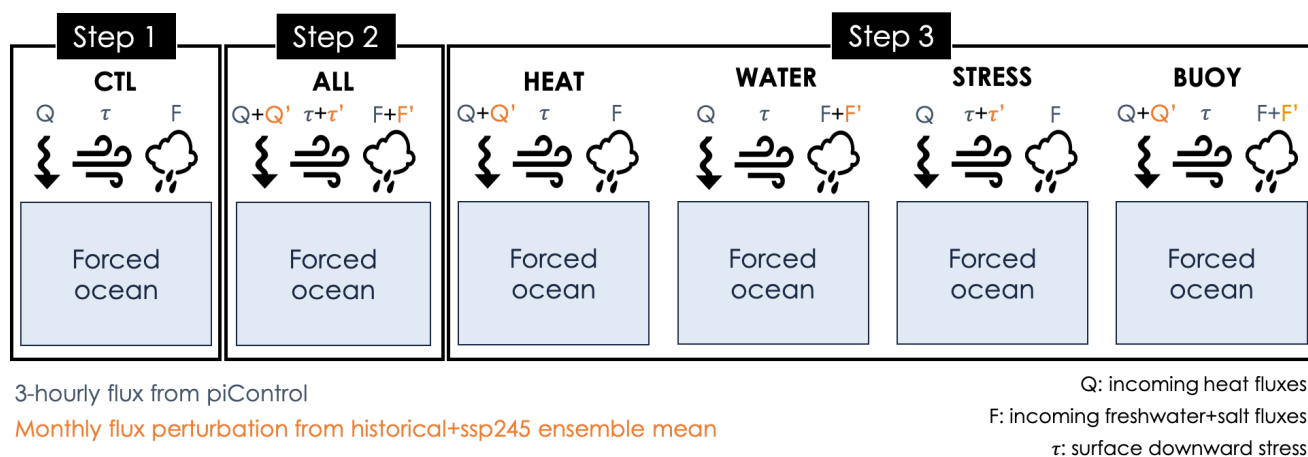   internal variability.

Figure 1. Schematic of the ocean-only simulations

       Configuring these experiments is undertaken in three main steps (Fig. 1):

Step 1. Setting up the CTL experiment from the piControl fluxes. This required listing all the components needed to force the
105   ocean from the coupled model, in order to reproduce as closely as possible the piControl variability for 251 years, and
         adapting the code as to enable the flux forcing. This step is described in section 3.

Step 2. Setting up the ALL experiment by adding the perturbation components on all surface fluxes, and ensuring the response
         is consistent with the historical+ssp245 large ensemble. This is described in section 4.2.

Step 3. Setting up the sensitivity experiments (HEAT, WATER, BUOY, STRESS), briefly described in section 4.3.

110   With these experiments, we thus aim to reproduce, with an ocean-only model, the internal variability and externally forced response of a coupled model. We also added two passive tracers in each of these experiments as described in section 5.

Because the piControl of the IPSL-CM6A-LR published for CMIP6 did not have all needed high frequency outputs, we also had to run before-hand another piControl simulation with the IPSL-CM6A-LR coupled model to save the fluxes at the liquid ocean interface at 3-hourly frequency as well as other components (Fig. 2). This piControl is initialized from the same spin-up
115   simulation used as restart state in the piControl r1i1p1f1 published for CMIP6, but ran on a different machine. Our piControl ran for 401 years, but only the 3-hourly fluxes from the last 251 years are used in the forced experiments presented here. The IPSL-CM6A-LR model has a systematic quasi-linear drift in global-mean ocean temperature (Mignot et al., 2021), passed on to this piControl and subsequently to the CTL and other ocean-only experiments, due to a negative incoming heat flux (see heat budget in Table 2). Unless noted otherwise, this drift is not removed in the figures presented here.

120   This work has been conducted with the IPSL-CM6A-LR coupled model and the NEMO3.6 ocean model, and this paper aims to describe how to apply our protocol with the NEMO3.6 model. Nevertheless, the numerical design could be applied to any coupled model and its ocean-only configuration, as long as the externally-forced historical and future response can be extracted. The historical+ssp large ensemble approach seems to be the most accurate way to isolate the forced response compared to e.g. fitting a 4th order polynomial to a single member (Lehner et al., 2020).

125   This protocol bares similarities with the ocean-only Flux-Anomaly-Forced Model Intercomparison Project (FAFMIP, Gregory et al. (2016); Todd et al. (2020), www.fafmip.org). However, significant differences exist since we don't aim at answering the same questions. This is addressed in the discussion section.

3   Step 1: reproducing a piControl coupled experiment with an ocean-only configuration

In this section, we describe how to set up the ocean-only CTL experiment (Fig. 1) to reproduce the ocean state of the coupled
130   piControl simulation. The ocean configuration set up for this study is based on the one used in IPSL-CM6A-LR, where we removed the sea-ice component (LIM3 (Rousset et al., 2015), as explained above) and the biogeochemistry component (PISCES (Aumont et al., 2015)), as we did not intend to look into the biogeochemical impacts. This configuration is forced with outputs from the coupled piControl simulation.

The first stage of the protocol is to select the correct fluxes needed to force the ocean. A synthetic schematic of the interactions
135   between the liquid ocean (under sea-ice) and all the other components of the IPSL-CM6A-LR model is presented in Fig. 2, and the list of needed fluxes is given in Table 1. All these fluxes (heat, freshwater, salt, wind stress), must be outputted from the coupled simulation at high frequency, and used in the forced experiments without time interpolation. We choose to output them at 3-hourly frequency, twice the coupling frequency in IPSL-CM6A-LR, in order to solve the diurnal cycle and keep close to piControl variability. The second stage of the protocol is to verify that the model is conservative, and the heat and freshwater
140   budgets are closed. Moreover, ocean global budgets must be quasi-identical between the coupled piControl and the ocean-only CTL experiments. A slight difference is however observed and is solely due to a slightly larger global ocean area in the coupled configuration which includes closed seas.
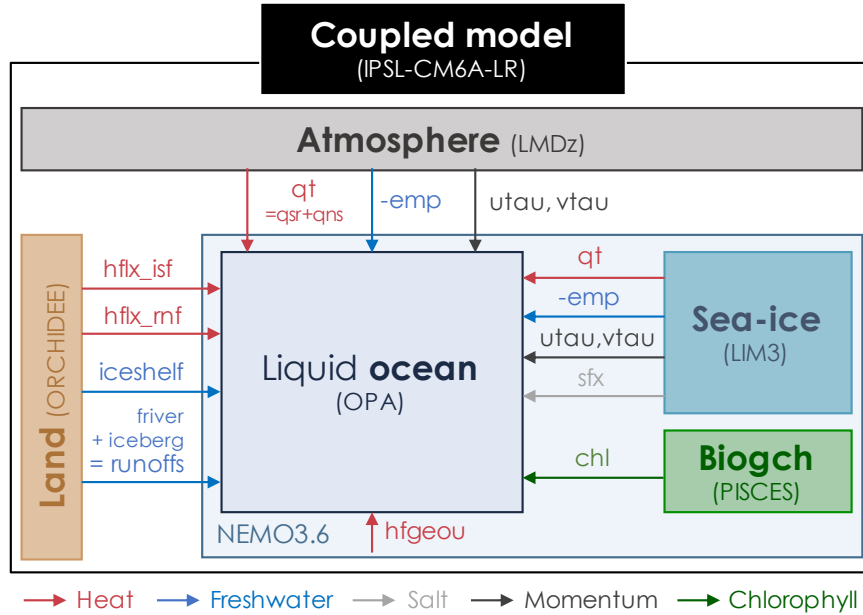
Figure 2. Schematic of the exchanges between the ocean and the other components in the IPSL-CM6A-LR coupled model affecting ocean physics. See Table 1 for the signification of these fluxes.

3.1 Freshwater fluxes

The ocean receives and loses water from liquid and solid precipitation, evaporation, sea ice melting and freezing, river runoffs,
145   iceberg melting and iceshelf melting. Under-iceshelf cavities are closed in the configuration of the ocean used here. Instead,
in the current version of the model, the mass of water contained in the icesheets is conserved, and all excess precipitation
falling on the icesheets eventually returns to the ocean through different melting terms. In the Southern Hemisphere, 50% of
the mass goes into iceshelf melting along the coast (based on the Depoorter et al. (2013) climatology), and 50% into iceberg
melting along the Merino et al. (2016) climatological map. In the Northern Hemisphere, all the mass goes into a calving term
150   uniformly distributed over the Northern Hemisphere ocean. The freshwater flux from rivers, icebergs and iceshelf melt are
vertically distributed over a prescribed depth. Table 1 presents how these terms are grouped and defined in the model.

The net amount of water entering the ocean liquid interface is given by -emp + runoffs + iceshelf (Fig. 2). If the model is
conservative, for a given time period Δt, the global ocean volume change (in m^3 s^{-1}) should verify:

$$\frac{\Delta V}{\Delta t} = \frac{1}{\rho_0} \times \overline{\sum_{i,j}((-\text{emp}(i,j,t) + \text{runoffs}(i,j,t) + \text{iceshelf}(i,j,t)) \times \text{areacello}(i,j))}^t \qquad (1)$$

155   Where ρ_0=1026 kg m^{-3} is the ocean volumic mass of reference, areacello is the ocean grid cell area, and i and j are the
horizontal coordinate indices. All three terms emp, runoffs and iceshelf outputted from the piControl are needed to force the

Short name	Signification	Unit	CMIP6 equivalent
emp	Evaporation-Precipitation, includes sea-ice formation and melt, and calving in the NH	kg m^{-2} s^{-1}	**-(wfo+friver+ficeberg)**
runoffs	River runoffs + iceberg melting (> 0 into ocean) = friver + iceberg	kg m^{-2} s^{-1}	**friver+ficeberg**
iceshelf	Iceshelf melting (> 0 into ocean)	kg m^{-2} s^{-1}	**flandice**
qt	Net downward heat flux = qns + qsr	W m^{-2}	**hfds**
qsr	Downward shortwave flux	W m^{-2}	**rsntds**
qns	Downward non solar heat flux (includes hflx_icb and hflx_cal)	W m^{-2}	nshfls
hflx_rnf	Sensible heat flux from river and iceberg runoffs (at SST)	W m^{-2}	**hfrunoffds**
hflx_icb	SH iceberg latent heat loss (<0), included in qns	W m^{-2}	
hflx_cal	NH calving latent heat loss (<0), included in qns	W m^{-2}	
hflx_isf	Heat flux from iceshelf melting (sensible+latent)	W m^{-2}	
hfgeou	Geothermal heat flux (constant in time)	W m^{-2}	
sfx	Downward salt flux into sea water	g m^{-2} s^{-1}	**sfdsi**
utau	Surface downward x stress	N m^{-2}	**tauuo**
vtau	Surface downward y stress	N m^{-2}	**tauvo**
DCHL	Mass concentration of diatoms expressed as chlorophyll	mg m^{-3}	
NCHL	Mass concentration of other phytoplankton component expressed as chlorophyll	mg m^{-3}	
CHL	Mass concentration of all phytoplankton expressed as chlorophyll = DCHL + NCHL	mg m^{-3}	
siconc	Sea-ice fraction	∅	
sithic	Sea-ice thickness	m	

Table 1. Exchanges between the ocean and the other components in the IPSL-CM6A-LR coupled model and their signification. The terms in bold are those read from the coupled model outputs to force the stand-alone ocean model in a fixed-flux configuration. The corresponding CMIP6 terms are given in the right column, indicating the terms needed to compute the flux perturbations.

ocean. Averaged over the global ocean (see Figs. 3a and Table 2), the main balance is between emp and friver (1718 mSv and 1636 mSv resp. in CTL, averaged over the entire simulation). The contribution from iceshelf melt and icebergs (the latter being included in runoffs) are very small (41 mSv each; 1 mSv = 10^3 m^3 s^{-1}). The net incoming freshwater flux for the simulation period is ∼1 mSv, equivalent to a total volume change of ∼9086 km^3 of water over the simulation. As compared to the mean volume of ocean water (1.33×10^{12} km^3) this change can be considered close to zero.

To check the closure of the freshwater budget over the period 1850-2100, the left hand term and right hand term in equation 1 are computed separately (Table 2). ΔV is computed as the difference in global ocean volume V between the last and first day of the simulation for CTL, and the last and first 3-hours in piControl. The right hand term in equation 1 is computed from annually-averaged freshwater fluxes. The ocean freshwater budget closes almost perfectly in piControl and CTL, with mismatches of only 0.002 mSv and 0.016 mSv respectively, which corresponds partly to precision errors introduced during
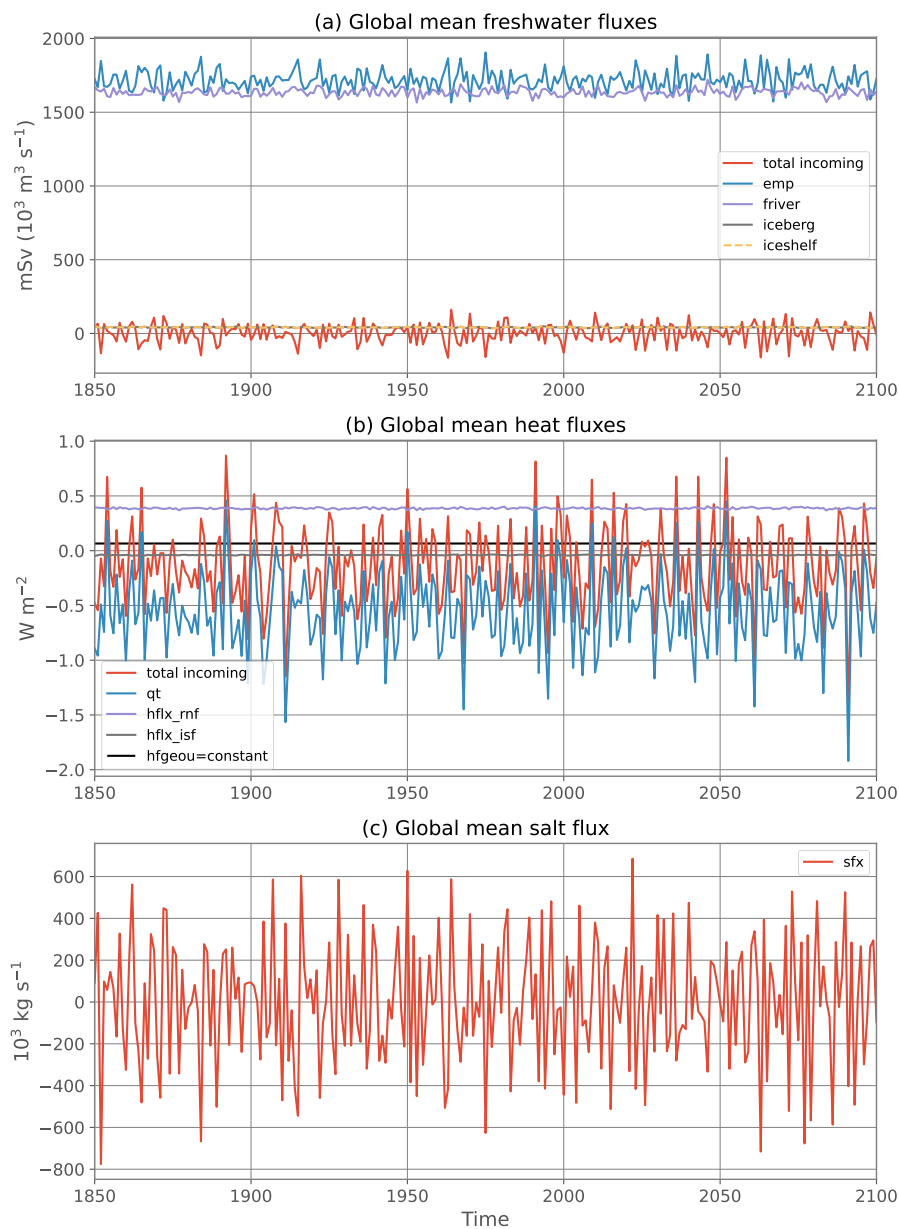
Figure 3. Evolution of the globally averaged, annual mean (a) freshwater fluxes, (b) heat fluxes and (c) salt flux in the piControl and CTL experiments. See Table 1 for the signification of each flux component.

data analysis. A larger mismatch in CTL is due to the use of daily means versus 3-hourly means in piControl to estimate ΔV. Indeed, we quickly lose precision in estimating the differences in global scalars when using larger time averages. For example, the use of monthly means instead of daily means to estimate ΔV in CTL leads to an equivalent freshwater flux of 0.990

170   mSv instead of 1.128 mSv, leading to an error of 0.154 mSv on the closure of the freshwater budget instead of 0.016 mSv.
Furthermore, piControl and CTL have quasi-identical net incoming freshwater fluxes (1.148 mSv vs 1.144 mSv respectively),
the difference being due to the global ocean area which includes closed seas in the coupled configuration while it does not in
the stand-alone ocean configuration, as stated before.

	piControl	CTL	ALL
emp	1734.78 mSv	1717.61 mSv	1808.35 mSv
runoffs	1694.57 mSv	1677.41 mSv	1769.50 mSv
iceshelf	41.34 mSv	41.35 mSv	44.27 mSv
Total incoming freshwater flux	**1.148 mSv**	**1.144 mSv**	**5.426 mSv**
= -emp + runoffs + iceshelf			
vs. ΔV	**1.146 mSv**	**1.128 mSv**	**5.417 mSv**
qt	-0.548 W m^{-2}	-0.546 W m^{-2}	0.166 W m^{-2}
qsr	172.125 W m^{-2}	172.114 W m^{-2}	171.703 W m^{-2}
qns	-172.673 W m^{-2}	-172.661 W m^{-2}	-171.537 W m^{-2}
hflx_rnf	0.388 W m^{-2}	0.386 W m^{-2}	0.424 W m^{-2}
hflx_isf	-0.040 W m^{-2}	-0.040 W m^{-2}	-0.043 W m^{-2}
hfgeou	0.066 W m^{-2}	0.066 W m^{-2}	0.066 W m^{-2}
Total incoming heat flux	**-0.1345 W m^{-2}**	**-0.1344 W m^{-2}**	**0.6125 W m^{-2}**
= qt + hflx_rnf + hflx_isf + hfgeou			
vs. ΔOHC	**-0.1343 W m^{-2}**	**-0.1346 W m^{-2}**	**0.6128 W m^{-2}**
sfx	**620 kg s^{-1}**	**622 kg s^{-1}**	**13 653 kg s^{-1}**
vs. ΔSC	**565 kg s^{-1}**	**111 kg s^{-1}**	**13 770 kg s^{-1}**

Table 2. Freshwater, heat and salt budgets over 1850-2100 in the piControl, CTL and ALL experiments (see Table 1 for the signification of
each flux component). All fluxes are globally-averaged. Global ocean changes (volume, heat and salt contents) are converted to flux units for
comparison.

3.2   Heat fluxes

175   The ocean exchanges heat at its upper boundary with the atmosphere and sea-ice components, through shortwave radiation (the
part not used to melt sea-ice), long-wave radiation and all other non-radiative fluxes (sensible and latent heat from evaporation,
precipitation and ice thermodynamics). It also receives sensible heat from river runoffs, sensible+latent heat from iceberg
melting in the Southern Hemisphere, latent heat from calving in the Northern Hemisphere and sensible+latent heat from
iceshelf melting. As for the freshwater fluxes, the heat fluxes associated with runoffs and iceshelf melting are distributed on the

9

180   vertical. Finally, the ocean bottom is warmed up by a constant, spatially-varying geothermal heating (Goutorbe et al., 2011).
These terms are synthetized in Table 1 together with notations used in the model.

The total heat input into the ocean is qt + hflx_rnf + hflx_isf + hfgeou (Fig. 2). For a given time period Δt, the global ocean heat content change (in Watts) should verify:

$$\frac{\Delta OHC}{\Delta t} = \overline{\sum_{i,j}((\text{qt}(i,j,t) + \text{hflx\_rnf}(i,j,t) + \text{hflx\_isf}(i,j,t) + \text{hfgeou(i,j)}) \times \text{areacello}(i,j))}^{t} \qquad (2)$$

185   With $OHC(t) = \rho_0 \times cp \times \sum_V \theta(t)dV$; $cp \approx 3981$ J K^{-1} kg^{-1} the ocean specific heat; θ the ocean conservative temperature with grid cell volume dV. The model needs to read qt, qsr, hflx_rnf and hfgeou to force the ocean. hflx_isf is reconstructed online from the freshwater flux term (iceshelf, see Table 1) and the freezing point temperature. In the NEMO code, the short-wave radiation qsr needs to be specified separately as it penetrates in the top hundred meters of the ocean depending on the chlorophyll concentration field (see subsection 3.5.2 on chlorophyll prescription).

190   Globally-averaged, the main balance is between solar (qsr=172.11 W m^{-2} in CTL) and non-solar (qns=-172.66 W m^{-2} in CTL) heat fluxes. They almost compensate and yield a net downward heat flux qt=-0.55 W m^{-2} on average (Table 2). The runoffs sensible heat flux (hflx_rnf) and iceshelf heat flux (hflx_isf) partly compensate qt with mean values of 0.39 W m^{-2} and -0.040 W m^{-2} respectively, though their interannual variability is much weaker than qt (see Fig. 3). The net incoming heat flux at the ocean interface sums up to -0.13 W m^{-2} during the 251 years (Table 2), illustrating the disequilibrium (or "drift") found in the piControl of this model, even after 2000 years of simulation (Silvy, 2022). By evaluating the left and right-hand side terms of the heat budget (equation 2) separately, we find a perfect closure of the budget, with an error of 0.002 W m^{-2} both in the piControl and CTL (Table 2).

3.3   Salt flux

When the ocean is coupled to a sea-ice model, there is a salt flux exchanged between the two components, as the ice salinity is
200   different from zero: when ice melts, there is a downward flux of both freshwater and salt. This salt flux sfx is thus needed to correctly reproduce the piControl with the ocean-only configuration. For a given time period Δt, the global ocean salt content change (in g of salt s^{-1}) should verify:

$$\frac{\Delta SC}{\Delta t} = \overline{\sum_{x,y}(\text{sfx}(x,y,t) \times \text{areacello}(x,y))}^{t} \qquad (3)$$

With $SC(t) = \rho_0 \times \sum_{V_{ocean}} S(x,y,z,t)dV$; S the ocean absolute salinity (in g kg^{-1}) with grid cell volume dV.

205   The globally-averaged salt flux oscillates around zero during the entire simulation (Fig. 3). It has a mean value of 1.7×10^{-9} g m^{-2} s^{-1}, equivalent to 622 kg s^{-1} integrated over the ocean surface, or to a total of 4.9×10^{12} kg of salt exchanged over the simulation period. Evaluating the left-hand term in equation 3 separately (see Table 2), we find an equivalent ocean salt content change rate equal to 565 kg s^{-1} in the piControl and 111 kg s^{-1} in the CTL. These numbers in fact correspond to

very small variations compared to the interannual standard deviation of the globally-averaged salt flux of around 300 000 kg

210 s^{-1}. Furthermore, the left-hand term is very sensitive to the computation method (e.g. frequency of the global scalar outputs), which can explain the small discrepancy between the left and right-hand terms, but otherwise the closure of the salt budget is respected.

3.4 Wind stress

Wind stress is prescribed to the ocean via its zonal (utau) and meridional (vtau) components (Table 1). In ice-covered region,

215 it is the stress below sea-ice that is read (Fig. 2).

3.5 Other inputs needed in the ocean-only configuration

3.5.1 Vertical mixing

The parameterization of ocean vertical mixing depends on sea-ice concentration and thickness in the IPSL-CM6A-LR configuration, in particular the mixing length scale. Thus, both the sea-ice fraction *siconc* and thickness *sithic* are extracted from

220 piControl at 3-hourly frequency and used in CTL experiment. Sensitivity tests using sea-ice outputs at lower frequency (typically monthly) led CTL to diverge from piControl in the first months of the simulations (not shown).

3.5.2 Chlorophyll field

The shortwave heat flux penetrates through the surface layers of the ocean. The penetration of that flux in the ocean is modulated by the concentration of total chlorophyll, the only biogeochemical component that has an effect on the ocean physics in NEMO

225 (Fig. 2). When the biogeochemistry model (PISCES, Aumont et al. (2015)) is activated (which is the case in the coupled piControl experiment), the chlorophyll concentration comes from PISCES. When PISCES is deactivated (which is the case for our forced experimental setup), the solar radiation vertical profile in the ocean depends on a prescribed chlorophyll field.

There are 3 options to prescribe a chlorophyll field in the ocean-only configuration of NEMO, plus a fourth one that we implemented specifically for this work:

230 – option 1: imposing a constant and uniform chlorophyll field at the surface (=0.05 mg.m^{-3}),

– option 2: reading a 2D surface file,

– option 3: reading a 2D surface file and reconstruct vertical chlorophyll profiles,

– option 4 (new option implemented): reading a full 3D file which gives chlorophyll concentration everywhere in the water column.

235 Two options give very satisfying results in terms of SST: reading the surface chlorophyll field from the piControl at monthly frequency with no reconstruction on the vertical (option 2, purple line in Fig. 4), and reading the full 3D chlorophyll field from piControl, also at monthly frequency, (option 4, red line in Fig. 4). The other options make CTL to diverge from piControl

(options 1 and 3: grey and blue lines respectively in Fig. 4). We choose to use option 4 (reading the 3D field) since it is the most accurate way to reproduce the reference piControl.
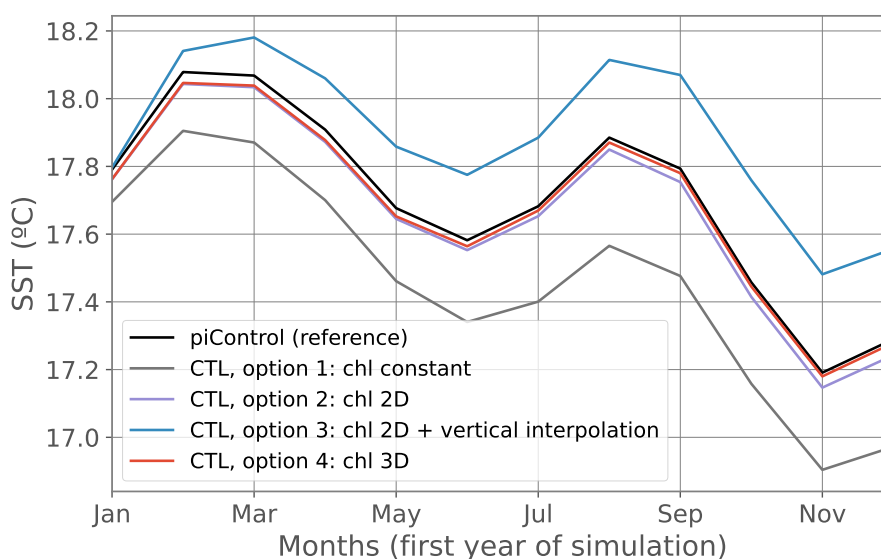

Figure 4. Monthly globally-averaged SST the first year of the simulation, for the: coupled piControl (black), CTL with the 3D chlorophyll read from the piControl (red), CTL with the 2D chlorophyll read from the piControl and interpolated vertically (blue), CTL with the 2D chlorophyll read from the piControl and the same values imposed in the subsurface (purple), and CTL with a constant and uniform chlorophyll value (grey).

3.5.3   Temperature below freezing point

Because of the ocean-only configuration, with no sea-ice, a flux formulation and no restoring term on the SST, it is possible for the temperature to locally fall below freezing point in polar regions, in all the ocean-only simulations including CTL. To deal with this issue, we implemented part of the solution proposed by Todd et al. (2020) in our experiments. Namely the ocean temperature is capped to the freezing point in the equation of state and in the calculation of the Brünt-Vaisala frequency. Hence, heat is conserved in the model while ocean density is capped so as to avoid spurious convection events. In that configuration, temperature can still fall below freezing point. In the ALL simulation for example, it can reach as low as -7°C in polar areas because of the negative heat flux perturbation especially in the Arctic, see section 4. This represents nevertheless only very small areas of the ocean and does not directly affect the circulation because it does not affect the equation of state, but heat transport can still be affected. We thus tested two alternative cases by constraining the temperature to the freezing point to confirm this choice (see Appendix A). The unconstrained case presented here turned out to be better suited for the scientific questions we aimed to answer, regarding in particular the timing of physical changes in the ocean interior and their emergence from background climate variability, as compared to the other methods presented in Appendix A.

3.6 Diagnostics and validation

The validation of the CTL experiment is an important step of this work, before going forward with adding the perturbation
255 components (section 4). Since the goal of the ocean-only CTL experiment is to reproduce the ocean state of the piControl, here
we compare these two simulations, i.e. their mean state and interannual variability, for a number of different diagnostics.

We keep in mind that CTL is forced at 3-hourly frequency using outputs from piControl. This frequency was chosen for
the physics in CTL to remain very close to the piControl. The coupling frequency in the coupled model between the ocean
and atmosphere is 1.5 hours (Boucher et al., 2020), so small errors are still introduced in the CTL all along the simulation
260 and can be amplified due to non-linearities of the system. Furthermore, we are also introducing errors by reading the monthly
chlorophyll field instead of an interactive chlorophyll and by reading the sea-ice fields at lower frequency than the coupling
frequency for the vertical mixing parameterization.

Figure 5. Annually-averaged global and overturning diagnostic time series for the coupled piControl (black) and the flux-forced ocean-only
CTL (red). The values in the bottom right corner correspond to the Pearson correlation coefficient (r) and to the root mean square error (rmse)
between the two time series.

Several 1-dimensional diagnostics are presented in Fig. 5 and 6 for piControl (black) and CTL (red) experiments for the
265 period 1850-2100. In terms of large scale circulation, we show the maximum in Atlantic meridional streamfunction (below
500m, 10°N-60°N, Fig. 5a), as well as the maximum of the global meridional streamfunction for the deacon cell (above 2000m,
30°S-80°S, Fig. 5b) and deep cell (below 2000m, 30°S-80°S, Fig. 5c) in the Southern Ocean. The similarity between the two
experiments is striking during the entire length of the simulations, both in terms of magnitude and of variability, especially for
the Southern Ocean cells where the two curves are almost perfectly superimposed (rmse two orders of magnitude weaker than
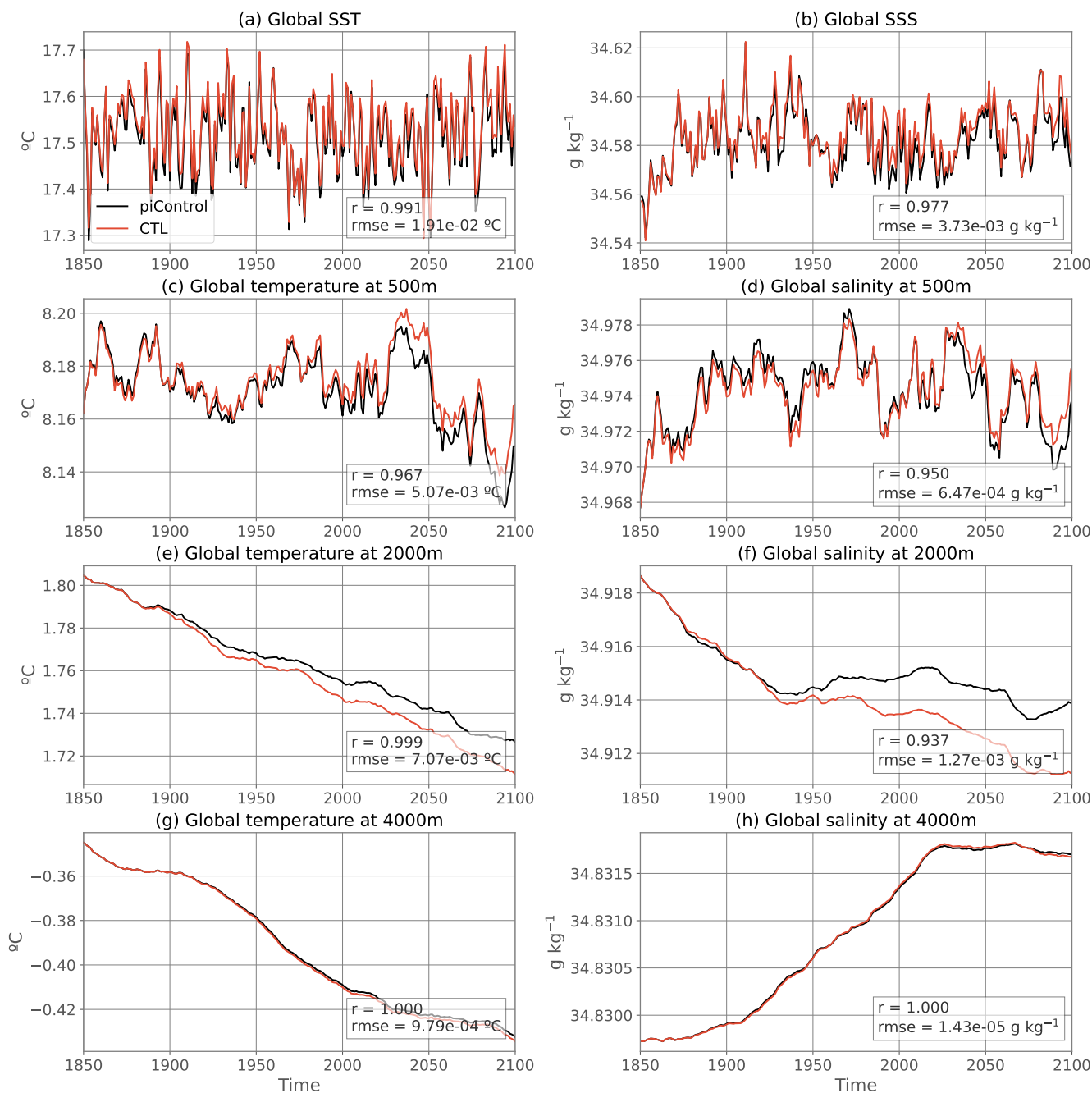
Figure 6. As in Fig. 5 but for global mean temperature and salinity at different vertical levels.

270   interannual variability). The AMOC time series is also very well reproduced, even though some extremes are not perfectly
replicated in terms of amplitude (rmse one order of magnitude smaller than interannual variability).

Global ocean heat, salt and volume are almost perfectly conserved between piControl and CTL (see the budgets in Table. 2) as illustrated by the respective superimposed time series in global mean conservative temperature (Fig. 5d), salinity (Fig. 5e) and total volume (Fig. 5f). In the surface (Fig. 6a,b) and bottom waters (Fig. 6g,h), temperature and salinity variability and
275　mean values in the CTL are the same as the piControl, with small differences in the peak values at the surface. The root mean square error (rmse) is an order of magnitude weaker than the interannual variability. Overall there is no drift away from the piControl even after 250 years of simulation in these variables. However, some small readjustment is slowly appearing at intermediate depths, with warmer and saltier waters at 500m (Fig. 6c,d) and colder and fresher at 2000m (Fig. 6e,f) in CTL than in piControl. The differences are nevertheless very small at the end of the simulation: rmse is about the same order of magnitude
280　as interannual variability. For all the time series presented in Fig. 5 and 6, the Pearson correlation coefficient between CTL and piControl is greater than 0.93, further validating the ocean-only CTL.

We now explore the climatological difference in temperature and salinity at the surface (Fig. 7a,c) and zonally averaged (Fig. 7b,d) between the two simulations. The main differences in SST (Fig. 7a) are localised in the subpolar North Atlantic. This is
285　where slightly below freezing temperatures locally occur in the CTL (not shown) due to the absence of a sea-ice model. The use of a 3-hourly forcing timescale rather than 1.5-hourly (i.e. the coupling frequency in the coupled model) may also induce such differences in this highly sensitive region. The strong SST differences are indeed located in a region of deep convection, where small differences in forcings without any retroaction can cause rapid divergence. These discrepancies however remain mostly limited to the subpolar North Atlantic and they are even smaller by the end of the simulation than at other time periods
290　(not shown). This could indicate strong internal variability governing the differences. In other parts of the globe, we note a slightly warmer Pacific in CTL compared to piControl and cooler Atlantic. Nonetheless, in most parts of the world except for the warmer patch in the subpolar North Atlantic, the differences between the ocean-only CTL and the coupled piControl are smaller than the piControl interannual variability (represented by the stipples in Fig. 7).

The largest SSS differences (Fig. 7c) are localised in the Arctic with very strong dipoles (>0.5 g kg^{-1} difference) which are
295　however smaller than the piControl interannual variability, are not constant in time (not shown) and seem to stay well within the corresponding ice-covered region of the coupled model. These differences might be due to the sensitivity of surface salinity to the formation and melting of sea-ice and to the fact that the forcing frequency in CTL is not the same as the coupling frequency in piControl, which induces propagating discrepancies. These differences in the Arctic do not impact other areas of the ocean where differences remain very small (<0.1 g kg^{-1} and smaller than piControl interannual variability everywhere at the surface)
300　without increasing much in time.

The differences in zonal mean temperature and salinity (Fig. 7b,d) between the CTL and the reference piControl confirm that the largest errors ($\sim$0.2-0.5 ℃ and $\sim$0.1 g kg^{-1} difference) are located in the Arctic and subpolar regions. The vertical dipolar structures suggest water-mass re-adjustment. There is some propagation at depth especially in the deep convection zone between 60-70 ℃N, but the differences don't increase in time after they are installed (not shown). This also confirms that the
305　CTL stays very close to the piControl in all other parts of the ocean, with very small differences between the two experiments (<0.05 ℃ and <0.01 g kg^{-1} outside the surface subtropical gyres). For temperature, these differences are smaller than the
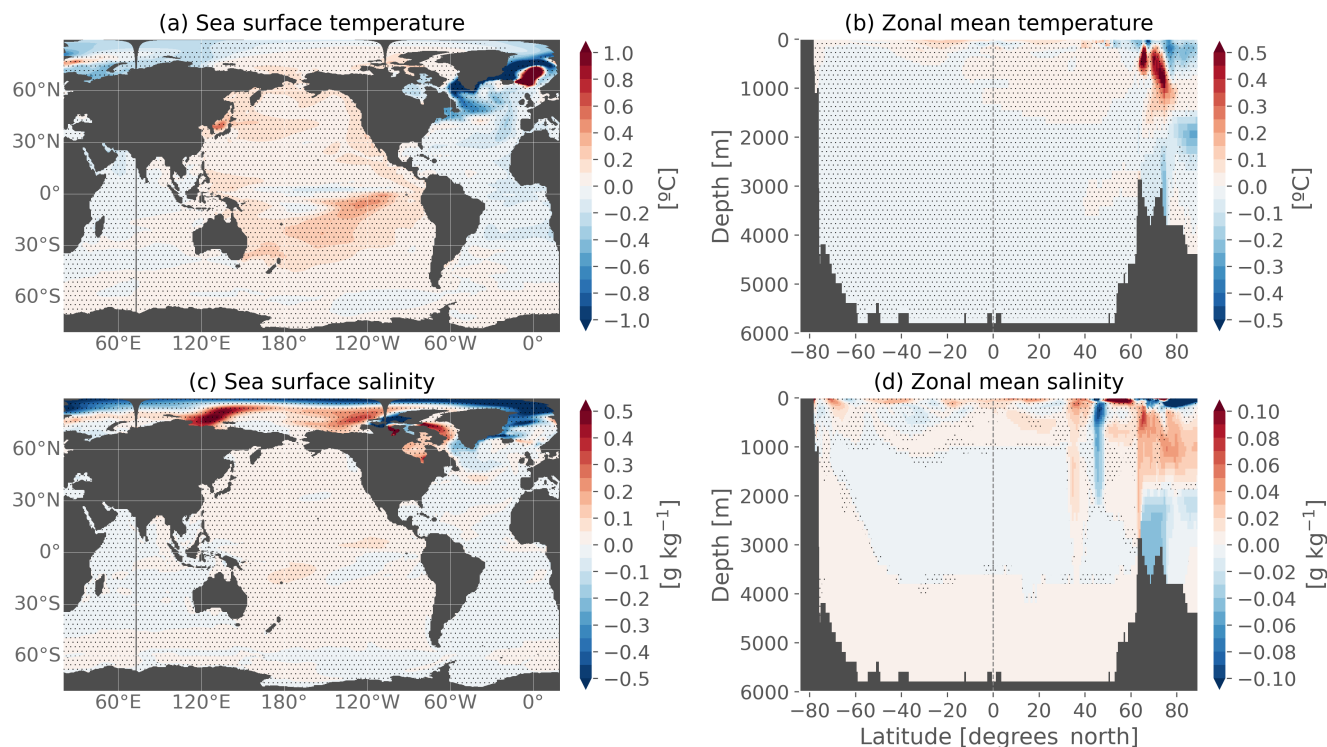
Figure 7. Annual-mean climatological differences between the ocean-only CTL and the coupled piControl experiments over the entire simulation 1850-2100 for (a) sea surface temperature, (b) zonal mean temperature, (c) sea surface salinity and (d) zonal mean salinity. Stipples indicate where the difference is lower than twice the interannual standard deviation of the piControl.

piControl interannual variability in all parts of the ocean except for some regions in the Northern high latitudes. However, for salinity, in the Northern high latitudes and in the global ocean interior, the differences are larger than the interannual variability, even though the amplitude remains quite small in absolute values.

310   Overall, these results are very satisfying for the purpose of the intended study and allow us to validate our CTL experiment.

4   Adding the perturbation components

The CTL experiment gives us the background state and internal variability of the ocean in the IPSL-CM6A-LR coupled model. To reproduce the oceanic response to climate change during 1850-2100 in the coupled model, we now set up the ALL experiment (see Fig. 1) using the same ocean-only configuration as CTL, and add an anomalous component to all surface fluxes
315   forcing the ocean (see Table 1). The anomalous component is constructed as follows:

$$Q'(i,j,t) = Q(i,j,t) - \overline{Q(i,j,t)}^{t=1850-1899}, \qquad (4)$$

where the overline denotes a temporal mean, and $Q(i,j,t) = <Q(i,j,t,k)>_{k=1..n}$ is the ensemble mean flux over all available members k of the IPSL-CM6A-LR large ensemble. There are 32 members for all variables over the historical-extended period 1850-2059, and 11 members over 2060-2100, following the ssp245 scenario from 2015; except for iceshelf which was outputted

320 in only 10 members over the full period. Furthermore, the net downward heat flux qt was missing 5 members over the 2015-2029 period. However, that did not affect the consistency of the ensemble means and no discontinuity was found between the different periods.

Also note that at the time of analysis, there were a few discrepancies in the published outputs of IPSL-CM6A-LR compared to the CMIP6 terminology, namely:

325 – hfrunoffds (sensible heat flux associated with river and iceberg runoffs) was not included in hfds (hfds=qt), unlike was was suggested in the heat budgets in Griffies et al. (2016).

– wfo (= water flux into sea water) was of opposite sign of what it should have been and was thus positive upward (=E-P-R)

– sfdsi (salt flux) is specified in kg m^{-2} s^{-1} in the CMIP6 requirements, however it was outputted in g m^{-2} s^{-1} in IPSL-CM6A-LR.

330 Here, we present the orders of magnitude, temporal evolution and spatial patterns of the flux anomalies, before presenting diagnostics to validate the ALL experiment.

4.1 Perturbations: budgets and spatial patterns

Table 2 compares the heat, freshwater and salt budgets in the ALL experiment compared to the CTL experiment, which differ only by the addition of the externally-forced perturbations (denoted by a prime sign in the text, e.g. emp').

335 The evolution of global mean freshwater flux anomalies (Fig. 8a) is a balance between two opposing terms increasing very rapidly at a similar rate (emp' and runoffs'=friver'+iceberg'; both reaching about 300 mSv in 2100), with an additional much smaller contribution from iceshelf' melting (reaching about 10 mSv in 2100). The latter is however significant in the total balance since the two large terms almost balance each other to yield an order of magnitude similar to the iceshelf' term. The balance between these three terms is similar to the freshwater flux budget in CTL (Table 2). Each individual freshwater flux

340 anomaly has a lower globally-averaged value than its CTL counterpart. After summation there is a net positive input of water in the ocean by the freshwater flux perturbations.

The increase in global mean surface heat flux anomaly (qt') is dominated by the increase in the non-solar heat flux term (qns') and damped by the decrease in the solar heat flux (qsr'), see Fig. 8b. There is a clear signature of the impact of volcanic eruptions (e.g. 1883, 1963, 1982, 1991) in the qsr' and qt' terms. The anomalous heat flux components from runoffs hflx_rnf'

345 and iceshelves hflx_isf' are more than one and two orders of magnitude smaller than qt', respectively (see inset plot in Fig. 8b). Still, there is a significant increase in hflx_rnf' owing to the large increase in river runoffs (Fig. 8a), but only a small decrease in hflx_isf'. The IPCC AR6 report assesses a global ocean heat content change of 8.42 [6.08-10.77] ZJ yr^{-1} during the period 1971-2018 (Gulev et al., 2021), equivalent to a 0.74 [0.53-0.94] W m^{-2} heat flux over the ocean area. As a comparison, the total incoming heat flux anomaly in this study during the same period is 0.68 W m^{-2}, consistent with the observed assessment.
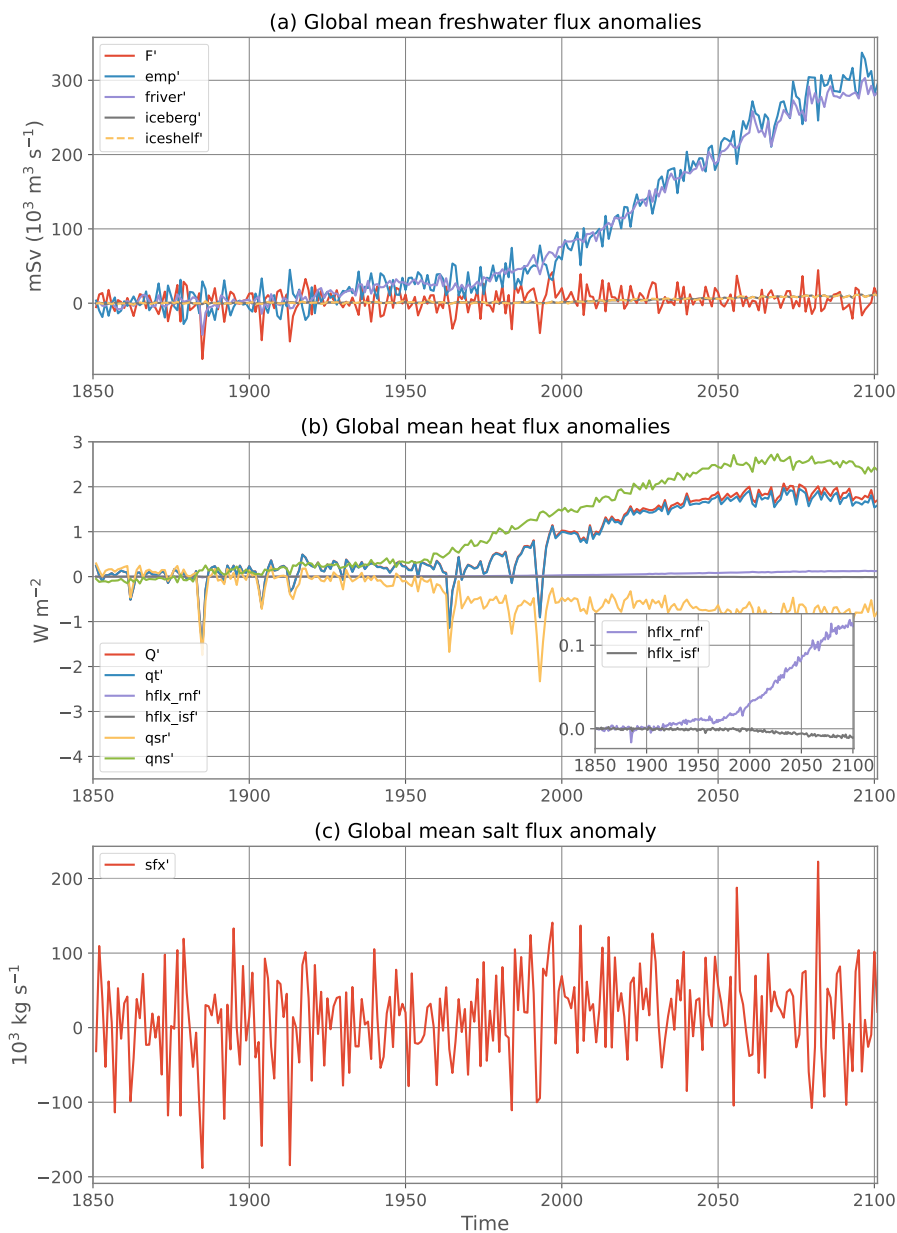
Figure 8. Evolution of the globally averaged, annual mean (a) freshwater flux, (b) heat flux and (c) salt flux anomalies as computed from the IPSL-CM6A-LR ensemble mean anomalies relative to 1850-1899. See Table 1 for the signification of each flux component.

350   Finally, globally integrated, the salt flux anomaly has a small positive trend with a much larger interannual variability (Table 2 and Fig. 8c), hiding marked spatial patterns of opposite signs (not shown).
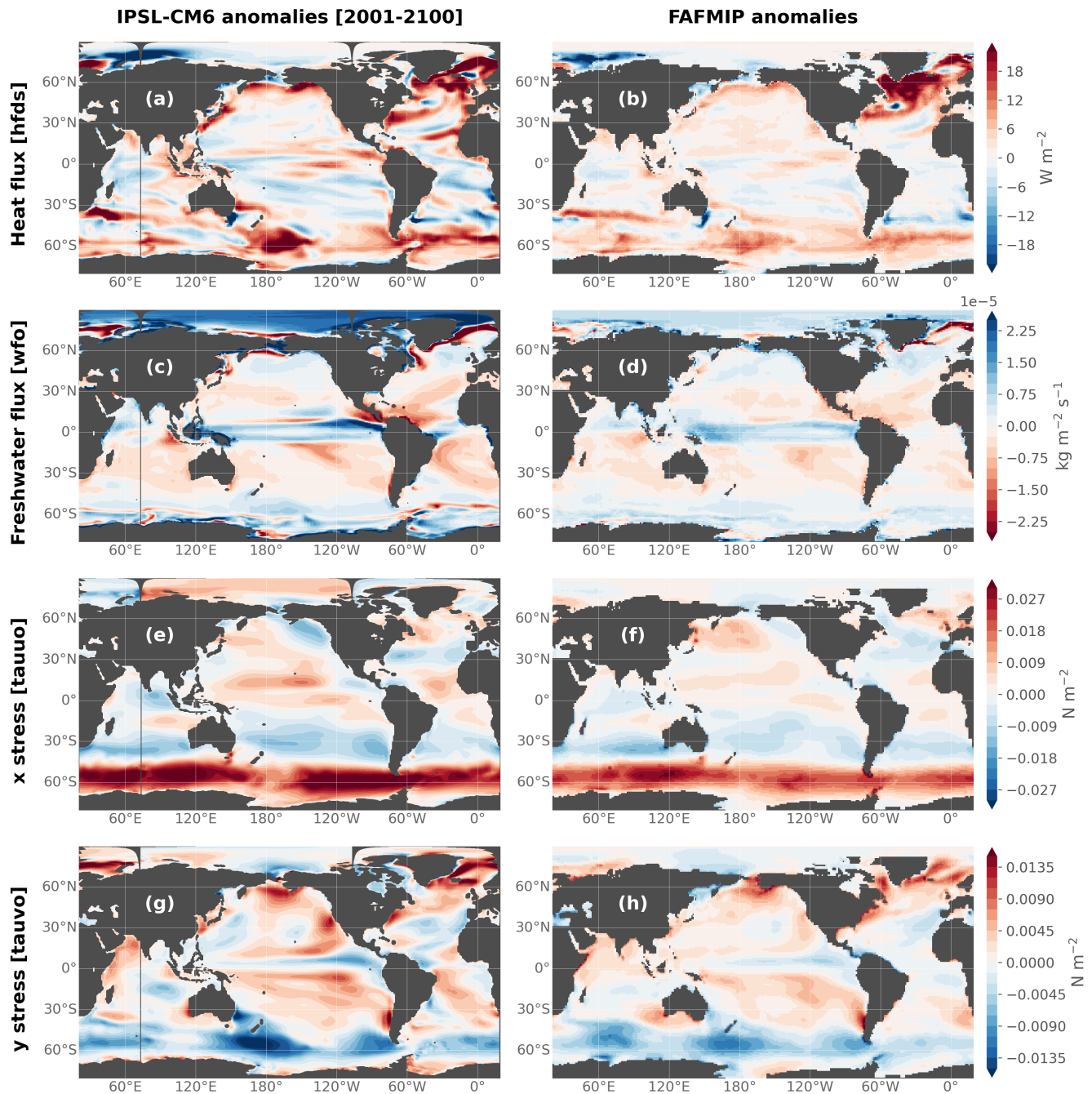
Figure 9. Heat flux, freshwater flux and surface downward x and y stress perturbations. (a,c,e,g) IPSL-CM6A-LR ensemble mean anomalies averaged over 2001-2100 relative to 1850-1899 compared to (b,d,f,h) FAFMIP anomalies downloaded from http://www.fafmip.org.

The spatial patterns of the anomalies averaged over the 21st century (Fig. 9a,c,e,g) qualitatively largely agree with those of the FAFMIP anomalies (Gregory et al., 2016). We recall that the latter are constructed from a multi-model mean of 1%CO2

355 idealized experiments. They are reproduced in Fig. 9b,d,f,h for comparison purposes. This similarity gives confidence in the response of the IPSL-CM6A-LR model to external forcings relatively to other coupled models in terms of surface fluxes. The most prominent features include large heat uptake over the subpolar North Atlantic and Southern Ocean; enhanced freshwater input over the tropics and high latitudes and weakening over the subtropical gyres, and intensifying and poleward-shifting westerly winds over the Southern Ocean.

360 ## 4.2 Step 2: reproducing the oceanic transient response to climate change

The goal for the ALL experiment is to simulate a climate evolution consistent with the one simulated in the IPSL-CM6A-LR historical+ssp245 ensemble. In other terms, the ALL experiment alone should have a comparable response to any member of this ensemble, i.e. a long-term trend on top of a specific phase of internal variability. It should thus be within or close to the ensemble envelope. Furthermore, the difference between the ALL and the CTL experiments (i.e. the estimate of the forced

365 response in our ocean-only setup) should, in theory, resemble that of the ensemble mean anomaly (i.e. the estimate of the forced response of the coupled model). Differences may however inevitably appear for several reasons. First, we impose the anomalous components at monthly frequencies. Furthermore, iceshelf and to a much less extent qt had a few members missing due to problems in the output files on different time periods (see above, section 4). This means that the ocean in the ALL experiment has seen slightly different forcings than those represented by the full ensemble mean oceanic results which we are trying

370 to reproduce. Second, in response to an anomalous surface cooling and in the absence of the sea-ice component and of any retroaction, the ocean cools instead of forming sea-ice, which leads to below-freezing temperatures in the Arctic (see section 3.5.3 and Appendix A). Third, the absence of sea-ice model also prevents the parameterization of vertical mixing under ice-covered areas to be fully exact. For all the sensitivity experiments, we choose to keep ice fraction and thickness as in the CTL experiment (section 3.5.1). Hence, mixing rates are not dependent on sea-ice melt resulting from climate change (anthropogenic

375 response). This choice is made for consistency reasons since the only external fields that must differ between the sensitivity experiments are the anomalous surface fluxes. Moreover, even though we could have provided the ice information from the IPSL-CM6A-LR historical+ssp245 ensemble for the vertical mixing of the ALL experiment, we cannot provide this information to the HEAT, WATER, STRESS and BUOY experiments since they are idealized cases with no reference for sea ice. The same choice is made for the chlorophyll field, which we keep as in CTL while it certainly changes under anthropogenic forcing.

380

Nevertheless, even with these limitations, the global response and patterns of change of the ALL experiment very closely resemble that of the IPSL-CM6A-LR historical+ssp245 ensemble. This is illustrated by several global and overturning diagnostics (Fig. 10) showing that the ALL experiment follows the response and stays within or close to the range of the large ensemble during the entire simulation (apart from a few peaks in interannual variability). It acts as an additional individual

385 member, that is, it has its own initial state and internal variability phased on the coupled piControl (and thus on the ocean-only CTL as well, cf Fig. 5 and 6). Additionally, it contains the response to external forcings inherited from the flux perturbations,

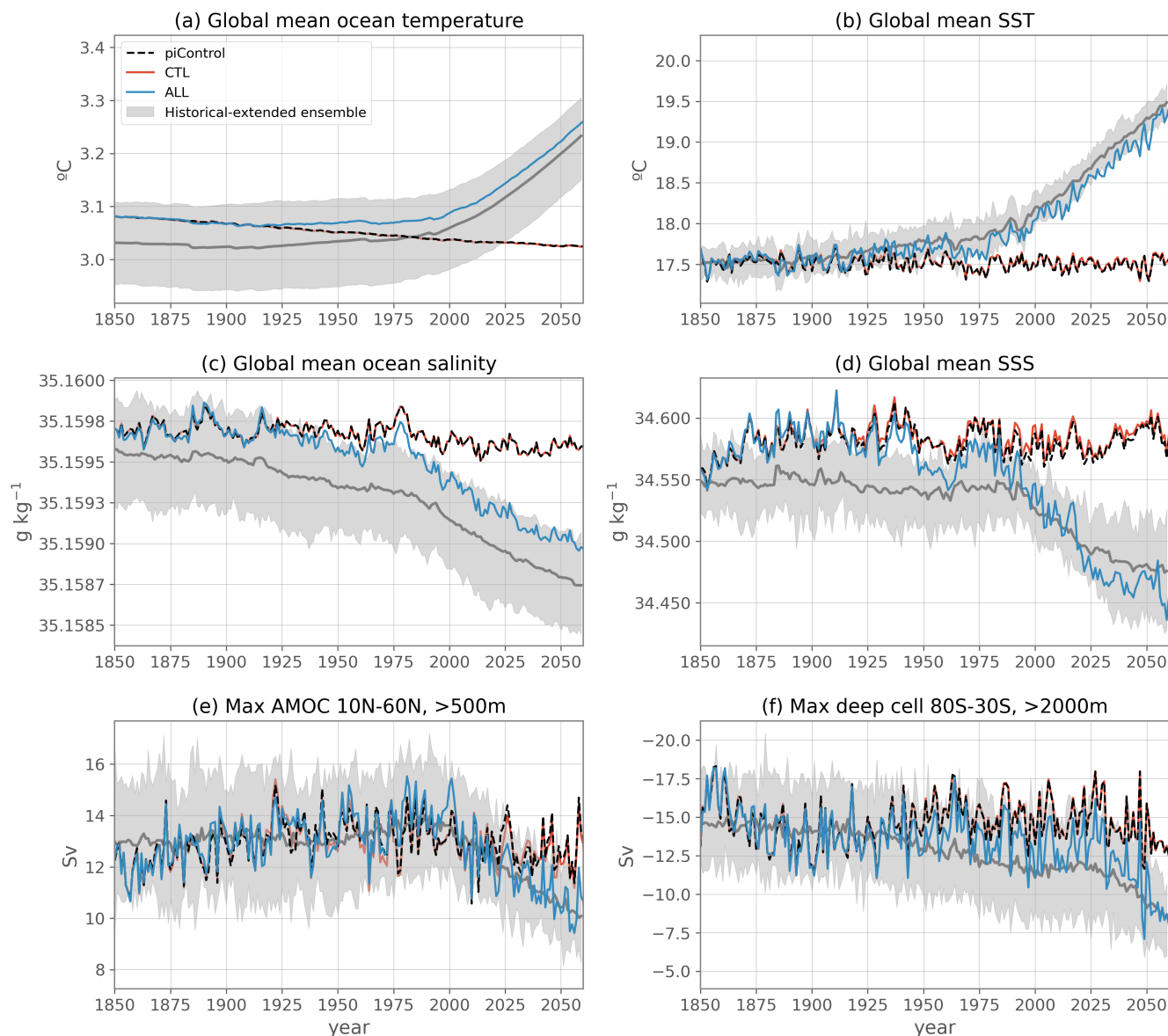**Figure 10.** Annually-averaged diagnostic time series for the coupled piControl (black), the flux-forced ocean-only CTL (red), ALL (blue) and the historical-extended large ensemble (grey shading = intermember spread; bold grey line = ensemble mean).

marked by long-term increasing anthropogenic emissions. Because of the ocean-only configuration forced with surface fluxes, there is no feedback to the atmosphere when adding a perturbation, thus all our experiments are aligned with the piControl internal variability and can be compared in time. Consequently, the forced response can be diagnosed precisely by taking the difference between the ALL and the CTL experiment. Note that since the internal variability in the ALL experiment is inherited

390

from the piControl simulation, it is not modified by the external forcings, unlike what occurs in simulations of the 21st century in which internal modes of variability can be affected by anthropogenic forcings (Bonnet et al., 2021).

We also compare the forced response in the ALL experiment to the ensemble mean response of the coupled model for zonal
395 mean temperature and salinity anomalies in the mid-21st century (Fig. 11a-f), as well as for vertically-integrated ocean heat and salt content (Fig. 11g-l) and sea surface temperature and salinity (Fig. 12).

Before doing that, the data represented in Fig. 11 has to be dedrifted for correct comparisons. Indeed, there is a cooling drift in the deep ocean in the IPSL-CM6A-LR model (see the heat budget of the piControl in Table 2, consistent with Mignot et al. (2021)). This drift is also visible in Fig. 10a. It leads to 1) different initial states between the historical members due to the fact
400 that they start from different dates of the pre-industrial control simulation and 2) contamination of the externally-forced trends if the data is not corrected for drift. Dedrifting is conducted for both temperature and salinity outputs by fitting a second-order polynomial to the 2000-year piControl at each gridpoint and removing the corresponding period of this fit from the historical members as well as from the ocean-only simulations which inherit an identical drift. The same mean state is then added to the historical members so that their initial conditions only differ because of internal modes of variability. Note that omitting this
405 dedrifting step leads to artificially very large intermember spread due to different initial states.

The response of the ALL experiment for zonal mean temperature, salinity, heat and salt contents is strikingly similar to the response of the IPSL-CM6A-LR ensemble mean, as displayed in Fig. 11. The difference between the two is smaller than the ensemble spread as measured by twice the intermember standard deviation almost everywhere, and as represented by the stippled areas in the right column of Fig. 11. This close reconstruction of the forced response validates the coherence of this
410 ocean-only configuration.

The only notable discrepancy seen for the SST anomaly is in the Arctic (Fig. 12) where surface temperatures fall below the freezing point in the ALL experiment (due to the absence of the sea-ice component), creating negative anomalies (Fig. 12a) unlike in the IPSL-CM6A-LR historical+ssp245 response where there is a small warming (Fig. 12b). This problem is addressed in Appendix A. We show that by constraining the temperature to not fall too much below the freezing point, we can solve these
415 negative anomalies and have better SST anomalies in the Arctic, but at the cost of degrading the temperature anomalies in other parts of the ocean.

The SSS patterns in the Arctic are also affected (Fig. 12c-f), with a larger freshening in the ALL experiment compared to the coupled model, particularly North of the Bering Strait. Nonetheless, these strong discrepancies in surface temperature and salinity are only located in the Arctic surface layers and don't seem to impact the rest of the ocean surface, nor the ocean
420 interior as illustrated by the previous diagnostics (Fig. 11).

More generally, we note the Arctic is the region of the world where we have the most difficulties reproducing the response of the coupled model, both in the CTL (Fig. 7) and the ALL (Fig. 12) experiments, due to the absence of a sea-ice model in our simulations. Interestingly, the absence of a sea-ice model is found to be much less problematic in the Southern Ocean. This work is thus not designed to study the mechanisms at play in the Arctic Ocean in the IPSL-CM6A-LR model, and any result
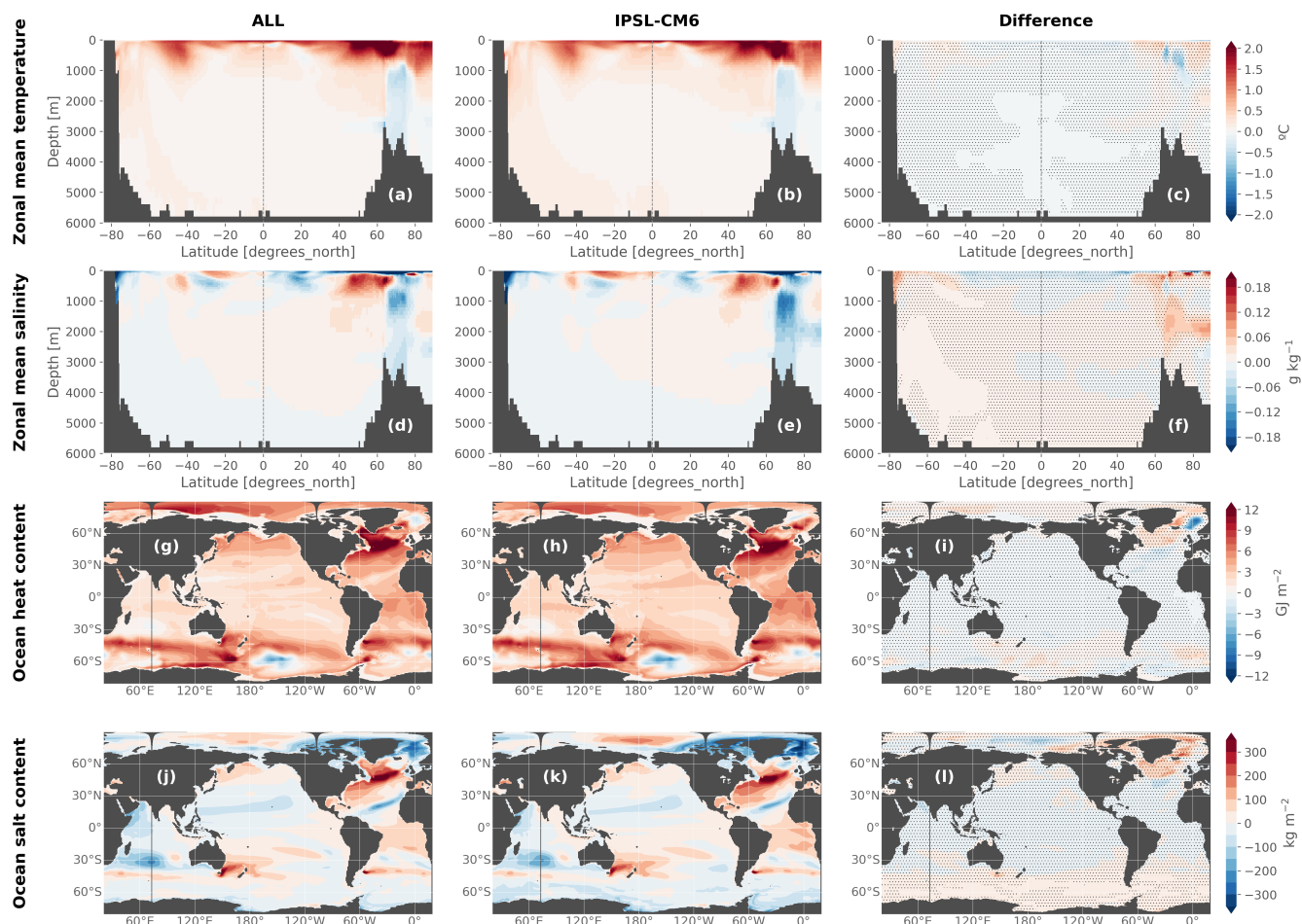425 in this region should be interpreted carefully.

Figure 11. Annual-mean anomalies in zonal mean temperature (a-c), zonal mean salinity (d-f), vertically-integrated ocean heat content (g-i) and ocean salt content (j-l), averaged over 2040-2059 relative to 1850-1899. The anomalies are calculated for the forced response in the ALL experiment (i.e. the difference between ALL and CTL; left column) compared to the forced response in the IPSL-CM6A-LR model as calculated from the ensemble mean (middle column). The difference between the ocean-only forced response in ALL and the coupled model ensemble mean are represented in the third column; the stipples indicate where the difference is lower than twice the intermember standard deviation. All the data has been corrected for ocean drift in this figure.

In all other regions, we have managed to reproduce very satisfactorily the IPSL-CM6A-LR large ensemble response with an ocean-only model and can coherently decompose the individual flux anomalies to investigate the different physical mechanisms within this framework.
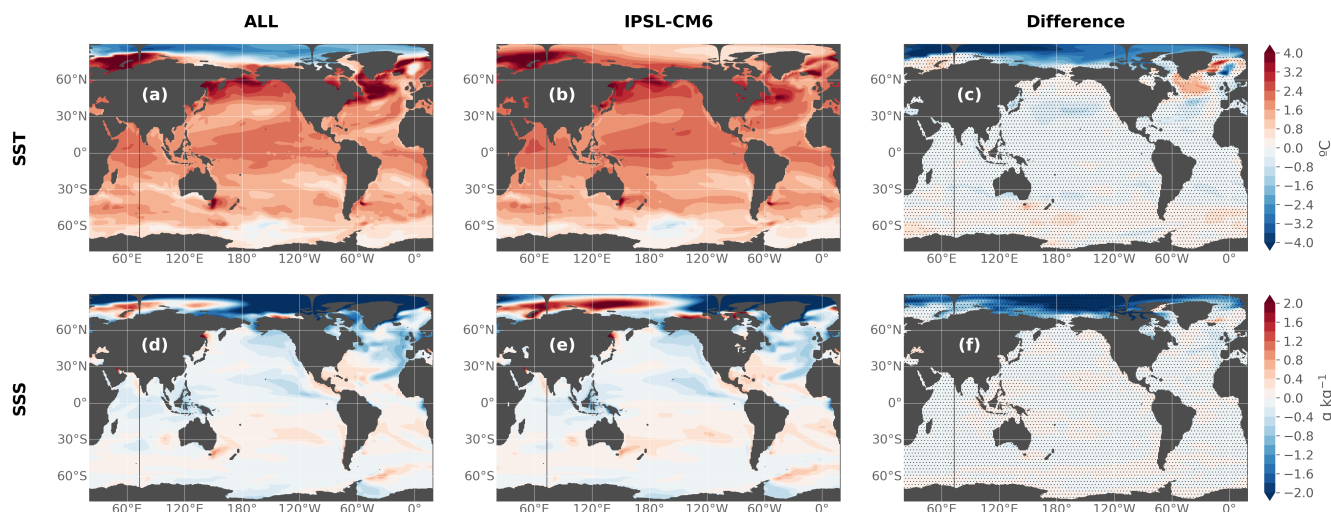
Figure 12. Same as Fig. 11 but for sea surface temperature (a-c) and salinity (d-f).

4.3 Step 3: decomposing the oceanic transient response with sensitivity experiments

430 The technical validation of the sensitivity experiments consisted in verifying whether the heat, freshwater and salt budgets were consistent with the CTL or the ALL experiment. In other words, in the simulations that have the heat flux anomalies activated, the heat budget should be identical as in the ALL experiment, while simulations that do not have the heat flux anomalies should have the same heat budget as CTL. These budgets were indeed verified and are rigorously equal to those shown in Table 2.

Then, we compared the long-term response of these sensitivity experiments to what was done previously in the scientific
435 literature in similar numerical designs, although most studies focused on heat storage and not many studies looked at salinity. Our results were found to be very coherent to previous studies in the spatial patterns of the temperature response, validating the long-term response, and enabling us to investigate the mechanistic decomposition of the transient response to anthropogenic climate change which was left unexplored. This is described in a companion paper (Silvy et al., in revision). This paper also presents the non-linear interactions between the surface fluxes.

440 ## 5 Implementing the passive tracers

In order to perform in-depth analysis of temperature and salinity evolution in response to climate change, two passive tracers are implemented using the TOP component in NEMO3.6 (Tracers in the Ocean Paradigm). These tracers will allow to diagnose the temperature and salinity anomalies in the perturbed experiments due solely to the addition of heat or freshwater in the ocean, without the contribution from the redistribution of pre-existing heat and salt in response to circulation changes. In the
445 following, we present their design based on Banks et al. (2002) and Banks and Gregory (2006), and their implementation in

the code.

5.1 Passive Anomaly Temperature

First, we write a simplified equation of the evolution of the prognostic temperature in the model by using a similar terminology
as Gregory et al. (2016) and Couldrey et al. (2021):

$$\frac{\partial T}{\partial t} = Q + \Phi(T). \tag{5}$$

Q is the net heat flux at the ocean boundaries (=qns+qsr+hflx_rnf+hflx_isf+hfgeou, we omit the constants ρ_0, cp and the grid
cell thickness by which Q needs to be divided for homogeneity), and the operator Φ represents advection and parameterizations
of sub-grid scale processes for temperature. For simplification, Φ is called transport in the following.

Second, we split T, Q and Φ into a CTL (unperturbed), and anomalous (perturbed) component, so that $T = T_{CTL} + T'$,
$\Phi = \Phi_{CTL} + \Phi'$, and $Q = Q_{CTL} + Q'$. The evolution of the temperature anomaly T' then writes as:

$$\frac{\partial T'}{\partial t} = Q' + \Phi_{CTL}(T') + \Phi'(T') + \Phi'(T_{CTL}). \tag{6}$$

The evolution of T' thus depends on the anomalous surface heat flux Q' (=qns'+qsr'+hflx_rnf'+hflx_isf'), also called "added
heat", the anomalous transport of CTL temperature ($\Phi'(T_{CTL})$), also called "redistributed heat", and the transport of T'.

Third, we further decompose T' into a passive uptake of added heat and a redistribution of pre-existing heat, so that $T' = T'_a + T'_r$. T'_a (for added heat), also called PAT (for Passive Anomaly Temperature), is implemented as a passive temperature
tracer representing the transport of added heat in the ocean without affecting the dynamics of the ocean. It is expressed in
ºC, initialized to 0, forced by the anomalous heat flux Q' (similarly as T') and transported in the ocean by the full circulation
$\Phi = \Phi_{CTL} + \Phi'$ as follows:

$$\frac{\partial T'_a}{\partial t} = Q' + \Phi_{CTL}(T'_a) + \Phi'(T'_a) \tag{7}$$

Since there is no feedback on the surface fluxes in our protocol, all the excess heat Q' that enters the ocean acts entirely to
change the global ocean heat content, that is, over a period of time Δt,

$$\iint \overline{Q'}^{\Delta t} dA = \frac{\rho_0 cp}{\Delta t} \iiint T' dV = \frac{\rho_0 cp}{\Delta t} \iiint T'_a dV = \frac{\Delta OHC}{\Delta t}, \tag{8}$$

with dA a surface grid cell area and dV a grid cell volume. The redistributed temperature, diagnosed from $T'_r = T' - T'_a$, has
no effect on the global ocean heat content and only changes temperature locally:

$$\frac{\rho_0 cp}{\Delta t} \iiint T'_r dV = 0, \tag{9}$$

25

All the forcing terms are applied to the passive tracer trend in the same way as for temperature: qns' acts on the top ocean
level; qsr' penetrates into the ocean with the same absorption coefficient as for qsr; hflx_rnf' and hflx_isf' are distributed ver-
tically.

Since T_a' does not affect the dynamics of the ocean, we can implement it in all the ocean-only simulations using the same
forcing term Q', irrespective of whether the perturbations are applied on the prognostic temperature (ALL, HEAT, BUOY) or
not (CTL, WATER, STRESS). The only difference in the evolution of T_a' between the ocean-only simulations is its transport
by the circulation which is specific to each experiment. In CTL, $\Phi' = 0$ by definition, so T_a' becomes the heat added from
externally-forced perturbations but passively transported by the CTL circulation, which corresponds to the passive tracer of
faf-passiveheat from the FAFMIP protocol (Gregory et al., 2016). It is also comparable to the temperature change in climate
change simulations for which the circulation is constrained to its climatological state (Winton et al., 2013; Bronselaer and
Zanna, 2020). Consequently, in CTL we have $\frac{\partial T_a'}{\partial t} = Q' + \Phi_{CTL}(T_a')$ and we can decompose the evolution of the added heat
T_a' in the ALL experiment into:

$$\left.\frac{\partial T_a'}{\partial t}\right|_{ALL} \approx \left.\frac{\partial T_a'}{\partial t}\right|_{CTL} + \Phi'|_{ALL}(T_a'|_{ALL}). \tag{10}$$

The difference among the T_a' tracers in-between experiments allows to diagnose the effect of the perturbed circulation on the
added heat, that is approximately the 2nd-order term $\Phi'(T_a')$.

In a companion paper (Silvy et al., in revision), we scientifically validate our PAT tracer by comparing its long-term spatial
distribution with previous studies. We find very similar patterns in most regions even though these previous studies mostly
focused on the response to idealized forcings.

5.2 Passive Anomaly Salinity

In NEMO, when using the variable volume formulation, the ocean salinity depends on a salt flux (sfx) and a concentra-
tion/dilution effect due to freshwater fluxes (emp, runoff, iceshelf), which makes the implementation of the Passive Anomaly
Salinity tracer (PAS) more complex than for PAT. For every simulation, PAS must only depend on the externally-induced
anomalies (sfx', emp', runoffs', iceshelf') and not on the background fluxes from the piControl. In particular, we do not want
PAS to be impacted by the background freshwater fluxes. Moreover, the freshwater flux perturbations are used to force only
some of the experiments (for instance ALL, WATER, BUOY), and not the others (CTL, HEAT, STRESS). It means that the
concentration/dilution effect is not the same in all the experiments. It complicates the formulation of PAS between experiments.
Hence:

1. to compensate the background concentration/dilution effect on PAS, we remove the effect of the piControl fluxes (emp,
   runoffs, iceshelf, without the anomalies) from the PAS trend;

505   2. to obtain the same effects of freshwater flux perturbations on PAS between experiments, we add freshwater flux anomalies in the PAS trend for CTL, HEAT and STRESS.

PAS is initialized to the approximate ocean global mean salinity (34.7 g kg^{-1}) because in the trend formulation, the salt flux anomaly is multiplied by the value of the tracer, which needs to have the same order of magnitude than salinity. However, this mean value is removed in all post-processing analyses to obtain an anomaly.

510   All the forcing terms are applied to the passive tracer trend in the same way as for salinity: sfx' and emp' act on the top ocean level while runoff' and iceshelf' are distributed vertically.

For both PAT and PAS, we had to locally, and at each time step, set them to 0 and 34.7 respectively (i.e. no variation) in a small and shallow area of the Arctic north of Canada which, unconstrained, was the source of unphysical values rapidly

515   propagating out of the area into the Arctic and North Atlantic basins. The reason for this error is yet unknown, although we suspect the shallowness of the Canadian Archipelago could be the source of these large anomalies in tracer values.

6   Discussion and conclusions

In this study, we propose a new modeling framework that aims at untangling the physical mechanisms driving temperature and salinity changes in the ocean interior, in the context of the IPSL-CM6A-LR simulations of human-induced global warming.

520   First, we reproduce a pre-industrial control experiment of the IPSL-CM6A-LR climate model, using an ocean-only configuration of the NEMO3.6 model. The model is forced with freshwater, salt, heat and wind stress fluxes at its boundaries. Second, in a sister simulation, we replicate the ocean's response of the IPSL-CM6A-LR historical+ssp245 ensemble by adding perturbations to these fluxes. Third, the role of each of the forcing terms is taken apart by building up sensitivity experiments. Fourth, two passive tracers are implemented to separate the effect of added heat and freshwater from the redistribution of pre-existing

525   heat and salt content by the anomalous circulation.

This paper describes the system, presents the approach, discusses the choices that are made and validates the simulations. Based on these simulations, a companion paper is underway to present scientific results on the transient ocean response to anthropogenic climate change (Silvy et al., in revision).

One limitation of the modeling framework presented here concerns changes at the surface of the Arctic Ocean. Choosing

530   another method to constrain temperature so that it does not fall too much below the freezing point might be better for questions tackling only the ocean surface, as presented in Appendix A, although it strongly deteriorates solutions over the rest of the ocean. Another limitation includes non-linear interactions in the system that lead to a mismatch between the all-forcing simulation (ALL) and the sum of individual-forcing simulations (e.g. HEAT+WATER+STRESS). However the mismatch is small in most of the ocean (wrt internal variability, Silvy et al., in revision), and the individual-forcing simulations are still very

535   helpful for in-depth understanding of the mechanisms leading to ocean changes.

Overall, this framework provides a new way to separate the internal variability from the externally-forced signal in ocean-only simulations. This is made possible, on the one hand, by a 3-hourly (i.e. high frequency) forcing at the ocean interfaces with background internally-driven fixed fluxes, and on the other hand by the extraction of the externally-forced signal from

540 a coupled model large ensemble, which is added to the background fluxes. To ensure that the externally-forced response is extracted correctly, the historical members need to sufficiently sample the various phases of internal variability, particularly in the ocean or coupled modes of variability, and thus enough different ocean states. Our framework is thus theoretically applicable to other large ensembles and their ocean model component, as long as the large ensembles are macro-initialized, i.e. historical members are branched from different times of a long pre-industrial control simulation. Indeed, these are not

545 necessarily sampled similarly in a micro-initialized (i.e. by applying small numerical perturbations in the atmospheric state) large ensemble (Hawkins et al., 2016). The CMIP6 archive could be particularly well adapted since its protocol ensures ensembles of historical simulations are macro-initialized, and several modeling centers have performed at least 30 historical members.

Our protocol is inspired by the ocean-only Flux-Anomaly-Forced Model Intercomparison Project (FAFMIP, Gregory et al.

550 (2016); Todd et al. (2020), www.fafmip.org) as well as other similar modeling experiments looking at the role of individual surface flux perturbations (e.g. Mikolajewicz and Voss (2000); Fyfe et al. (2007); Garuba and Klinger (2018); Liu et al. (2018); Shi et al. (2020)). Indeed, the shared goal is to understand the ocean response to the different forcings in response to climate change. However, the specific scientific questions that are tackled by FAFMIP differ from ours, which explains the important differences in experimental designs. Indeed, FAFMIP is by nature a multi-model study and aims at exploring the spread of

555 ocean model responses to the same surface flux perturbations. Their perturbation fields are constructed from the multi-model mean of idealized simulations in which CO_2 concentration increases at 1% per year, and taken at the time it has doubled in the atmosphere (wrt piControl). These perturbations are constant in time (step anomalies with a seasonal cycle but no interannual variations), allowing for a semi-equilibrium to be reached quickly, and long-term responses to be explored at a limited cpu cost (each simulation is run for 70 years). However the transient response to historical climate change cannot be assessed. FAFMIP

560 therefore does not investigate the balance of mechanisms causing ocean changes during the historical+scenario period in these climate models. This question of timescales is precisely the novel aspect we aim to tackle with this present study. The results obtained from our experiments can be broadly compared with those of FAFMIP simulations in the long-term response, and contain additional information on the transient response during the 1850-2100 period. Additionally, we start from a single climate model to decompose its ocean response. It allows to exactly implement the perturbed fluxes as they are seen in the

565 coupled model, as opposed to the FAFMIP perturbations.

As shown by some of the FAFMIP multi-model studies (Gregory et al., 2016; Todd et al., 2020; Couldrey et al., 2021), the spatial patterns of the ocean response to anthropogenic forcings can be model-dependent. Furthermore, there is also a large inter-model spread in their timescales of emergence from internal variability (Silvy et al., 2020). Consequently, a multi-model study would help to explore the uncertainties related to model responses. If other modeling centers were to gain interest in this

570 numerical design and the questions it aims to answer, this would be an interesting inter-comparison exercise, complementing the FAFMIP/ocean-only FAFMIP protocol.

All the simulations are done on the Jean-Zay supercomputer held at IDRIS center in France (http://www.idris.fr/eng/jean-zay/jean-zay-presentation-eng.html). The main cpu consumption of this protocol is the coupled piControl experiment (430 000 cpu hours), which is needed to output the surface fluxes at high-enough frequency. Running the ocean-only simula-
575 tions are much cheaper (220 000 cpu hours for the 6 ocean-only experiments). Because this was a novel framework and we had to run many sensitivity tests to the different parameters necessary to force the ocean model, the total cpu time consumed for this work is evaluated at 1 million cpu hours, accounting for 3-4 tCO_2e.

Code and data availability. The code modified for the development of the numerical experiments with NEMO3.6 is available here: https://github.com/ysilvy/simus_orca1_fluxforced. Forcing fields and simulation outputs can be made available upon reasonable request to Y.S.

580 **Appendix A: Treatment of the temperature below freezing point: sensitivity tests**

As mentioned in section 3.5.3, in both the unperturbed (CTL) and perturbed ocean-only experiments, the prognostic temperature in the ocean model can locally fall below the freezing point since there is no sea-ice component to physically constrain it. Here, we discuss two alternatives to this freely-evolving case by applying different treatments to the temperature when it falls below the freezing point. They are tested on the entire 251 years for both CTL and ALL:

585 Option 1. At each time step, we constrain the temperature to the freezing point temperature in the mixed layer when it falls below freezing, and we remove the equivalent heat flux in the non-solar heat flux qns distributed over the entire ocean surface, i.e. at each time step and every grid point: $qns(i,j,t) = qns(i,j,t) - qfrz(t)/S$ with qfrz the equivalent heat flux added by blocking the temperature to the freezing point over the mixed layer, integrated vertically and horizontally, and S the ocean area. This redistribution of the added heat flux allows for the global conservation of heat.

590 Option 2. At each time step, the temperature is relaxed to the freezing point temperature in the mixed layer when it falls below freezing, with a 30-day relaxation period. This is equivalent to adding a positive heat flux in the ocean locally, thus not conserving heat in the ocean globally, nor relative to the coupled model simulations. The question is to know whether this heat flux becomes large in time and if it has an impact on the rest of the ocean.

In the first alternative, the globally-averaged heat flux added and redistributed in qns due to temperatures falling below the
595 freezing point is qfrz=0.12 W m^{-2} in CTL compared to the total incoming heat flux -0.13 W.m^{-2}, over the entire simulation (251 years). In ALL, qfrz=0.16 W m^{-2} compared to the total incoming heat flux 0.61 W m^{-2}. This additional heat flux is thus globally non-negligible in both simulations and rises over time (not shown). This alternative conserves heat, as shown by the superposition of the red and blue lines in Fig. A1e). However there is a spurious cooling of the surface ocean layers (Fig. A1a) because of the removal and redistribution of qfrz in qns, that is balanced by a warming at depth (Fig. A1c). Salinity also
600 progressively deviates from the piControl at different depths (Fig. A1b and d). The AMOC and Southern Ocean deep cell suffer from the same problem (Fig. A1g and h). In terms of spatial patterns, this alternative prevents the SST from deviating too much

from the IPSL-CM6A-LR ensemble mean in the Arctic as expected (compare the freely-evolving simulation Fig. A2a with the alternative simulation Fig. A2d). But it clearly deteriorates the simulation at low and mid-latitudes (compare Fig. A2c and f). Overall this alternative is much less satisfactory than the freely-evolving simulations.

605

In the second alternative, the additional heat flux from relaxation amounts to 0.074 W m^{-2} in CTL and 0.11 W m^{-2} in ALL. It is less than the redistributed heat in the first alternative but still non-negligible considering over the ALL experiment, qfrz adds an additional 0.11 W m^{-2} to the ocean, on top of the total incoming heat flux of 0.61 W m^{-2}. There is a clear spurious warming of the ocean at every depth (Fig. A2a,c,e for CTL; Fig. A2g-i for ALL). Salinity deviates less from the piControl than

610 the first alternative but is still worst than in the freely-evolving simulation (Fig. A1b and d). The AMOC and Southern Ocean deep cell deviations are as strong as in the first alternative (Fig. A1g and h). In terms of spatial patterns, this alternative induces a better SST anomaly compared with the large ensemble in the Arctic as expected (compare Fig. A2a and g). But it clearly deteriorates the simulation below the surface layers in the polar regions (compare Fig. A2b and h, and Fig. A2c and i). Finally, the additional heat flux coming from relaxation makes the ocean warm faster and earlier than in the IPSL-CM6A-LR ensemble

615 mean (not shown). This is a big issue when investigating the timing of departure of a warming signal from background climate variability, which is what these simulations are designed for.

In summary, apart from better reproducing the SST pattern in the Arctic, and preventing the temperature from falling below the freezing point, these two alternatives do not better reproduce the response of the IPSL-CM6A-LR ensemble mean. Letting

620 the temperature evolve freely is found overall to not impact the response to anthropogenic climate change in other regions than the Arctic and it is without a doubt better at reproducing the temporal evolution of the piControl, which ensures minimal drift from our reference simulation. Furthermore, one of the goal of these experiments is to investigate the evolution of the added and redistributed heat components by implementing a passive anomaly tracer forced with identical surface flux perturbations as the prognostic temperature (see section 5). Modifying the prognostic temperature internally means creating discrepancies in

625 the relationship between temperature change and the passive tracer. We thus chose to run all the sensitivity experiments without any treatment of the temperature below freezing other than in the equation of state and Brunt-Vaisala frequency. The surface Arctic ocean is thus not a region of choice to analyse these experiments and should be considered with care. These technical choices are made in response to our scientific constrains; however, other scientific interests might have led to different choices.
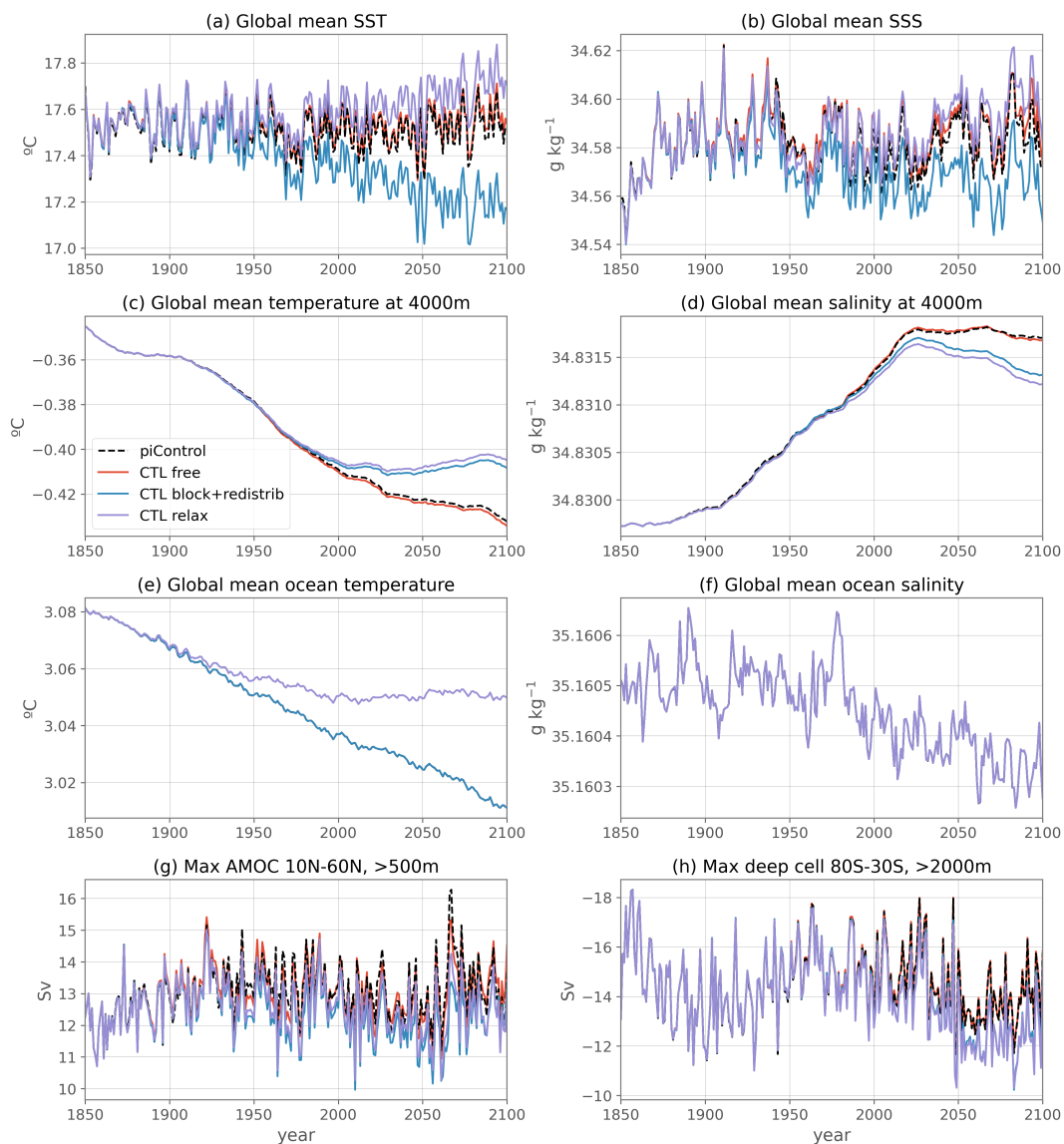
Figure A1. Annually-averaged diagnostic time series for the coupled piControl (black); the flux-forced ocean-only CTL as presented in the main text, i.e. in which the temperature evolves freely (red); the CTL in which the temperature is blocked to the freezing point if it falls below it, and the equivalent heat flux that has been added to block the temperature is redistributed over the globe (i.e. removed) in the non solar heat flux (blue, option 1 in the text); and the CTL in which the temperature is relaxed to the freezing point if it falls below it (purple, option 2 in the text). In all three CTL cases, in the equation of state and computation of the Brünt-Vaisala frequency, the temperature is maintained to the freezing point if it falls below it (see section 3.5.3)

.

Figure A2. Difference between the forced response in 2040-2059 relative to 1850-1899, in the ALL experiment compared to the IPSL-CM6A-LR ensemble mean. The forced response in the different ALL experiments are calculated by subtracting the corresponding CTL experiment, which ensures deep ocean drift is removed. In the IPSL-CM6 simulations, drift was removed prior to computations. These anomalies are calculated for sea surface temperature (left column), vertically-integrated ocean heat content (middle) and zonal mean temperature (right). The three rows compare the ocean-only simulations in which the temperature is allowed to evolve freely (top), in which it is blocked to the freezing point when it falls below it and the associated heat flux is redistributed over the ocean surface (middle, option 2 in the text), and in which it is relaxed to the freezing point when it falls below it (bottom, option 1 in the text).

630 *Author contributions.* Y.S and C.R developed the ocean-only configuration with the help of C.E and G.M. Y.S ran the simulations, conducted the post-processing analyses and wrote the first draft of the paper. E.G, J-B.S and J.M supervised the work and guided the choices made during the setup of the experimental design and its implementation, with the help of G.M. All authors contributed to the paper in its final form.

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

References

Armour, K. C., Marshall, J., Scott, J. R., Donohoe, A., and Newsom, E. R.: Southern Ocean warming delayed by circumpolar upwelling and equatorward transport, Nature Geoscience, 9, 549–554, https://doi.org/10.1038/ngeo2731, 2016.

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

Banks, H., Wood, R., and Gregory, J.: Changes to Indian Ocean Subantarctic Mode Water in a Coupled Climate Model as CO2 Forcing Increases, Journal of Physical Oceanography, 32, 12, 2002.

Banks, H. T. and Gregory, J. M.: Mechanisms of ocean heat uptake in a coupled climate model and the implications for tracer based

650   predictions of ocean heat uptake, Geophysical Research Letters, 33, https://doi.org/10.1029/2005GL025352, 2006.

Bonnet, R., Boucher, O., Deshayes, J., Gastineau, G., Hourdin, F., Mignot, J., Servonnat, J., and Swingedouw, D.: Presentation and Evaluation of the IPSL-CM6A-LR Ensemble of Extended Historical Simulations, Journal of Advances in Modeling Earth Systems, p. 25, https://doi.org/10.1029/2021MS002565, 2021.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot,

655   P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J., Guenet, B., Guez, E., L., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C.,

660   Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS002010, 2020.

Bronselaer, B. and Zanna, L.: Heat and carbon coupling reveals ocean warming due to circulation changes, Nature, 584, 227–233, https://doi.org/10.1038/s41586-020-2573-5, 2020.

665   Couldrey, M. P., Gregory, J. M., Boeira Dias, F., Dobrohotoff, P., Domingues, C. M., Garuba, O., Griffies, S. M., Haak, H., Hu, A., Ishii, M., Jungclaus, J., Köhl, A., Marsland, S. J., Ojha, S., Saenko, O. A., Savita, A., Shao, A., Stammer, D., Suzuki, T., Todd, A., and Zanna, L.: What causes the spread of model projections of ocean dynamic sea-level change in response to greenhouse gas forcing?, Climate Dynamics, 56, 155–187, https://doi.org/10.1007/s00382-020-05471-4, 2021.

Depoorter, M. A., Bamber, J. L., Griggs, J. A., Lenaerts, J. T. M., Ligtenberg, S. R. M., van den Broeke, M. R., and Moholdt, G.: Calving

670   fluxes and basal melt rates of Antarctic ice shelves, Nature, 502, 89–92, https://doi.org/10.1038/nature12567, 2013.

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

Fox-Kemper, B., Hewitt, H. T., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S. S., Edwards, T. L., Golledge, N. R., Hemer, M., Kopp, R. E.,

675   Krinner, G., Mix, A., Notz, D., Nowicki, S., Nurhati, I. S., Ruiz, L., Sallée, J.-B., Slangen, A. B. A., and Yu, Y.: Ocean, Cryosphere and Sea Level Change, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2021.

Fyfe, J. C., Saenko, O. A., Zickfeld, K., Eby, M., and Weaver, A. J.: The Role of Poleward-Intensifying Winds on Southern Ocean Warming, Journal of Climate, 20, 5391–5400, https://doi.org/10.1175/2007JCLI1764.1, 2007.

680 Garuba, O. A. and Klinger, B. A.: Ocean Heat Uptake and Interbasin Transport of the Passive and Redistributive Components of Surface Heating, Journal of Climate, 29, 7507–7527, https://doi.org/10.1175/JCLI-D-16-0138.1, 2016.

Garuba, O. A. and Klinger, B. A.: The Role of Individual Surface Flux Components in the Passive and Active Ocean Heat Uptake, Journal of Climate, 31, 6157–6173, https://doi.org/10.1175/JCLI-D-17-0452.1, 2018.

Goutorbe, B., Poort, J., Lucazeau, F., and Raillard, S.: Global heat flow trends resolved from multiple geological and geophysical proxies,
685 Geophysical Journal International, 187, 1405–1419, https://doi.org/10.1111/j.1365-246X.2011.05228.x, 2011.

Gregory, J. M., Bouttes, N., Griffies, S. M., Haak, H., Hurlin, W. J., Jungclaus, J., Kelley, M., Lee, W. G., Marshall, J., Romanou, A., Saenko, O. A., Stammer, D., and Winton, M.: The Flux-Anomaly-Forced Model Intercomparison Project (FAFMIP) contribution to CMIP6: investigation of sea-level and ocean climate change in response to CO2 forcing, Geoscientific Model Development, 9, 3993–4017, https://doi.org/10.5194/gmd-9-3993-2016, 2016.

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

Gulev, S. K., Thorne, P. W., Ahn, J., Dentener, F. J., Domingues, C. M., Gerland, S., Gong, D., Kaufman, D. S., Nnamchi, H. C., Quaas, J., Rivera, J. A., Sathyendranath, S., Smith, S. L., Trewin, B., von Shuckmann, K., and Vose, R. S.: Changing State of the Climate System, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2021.

700 Hawkins, E., Smith, R. S., Gregory, J. M., and Stainforth, D. A.: Irreducible uncertainty in near-term climate projections, Climate Dynamics, 46, 3807–3819, https://doi.org/10.1007/s00382-015-2806-8, 2016.

Hourdin, F., Rio, C., Grandpeix, J., Madeleine, J., Cheruy, F., Rochetin, N., Jam, A., Musat, I., Idelkadi, A., Fairhead, L., Foujols, M., Mellul, L., Traore, A., Dufresne, J., Boucher, O., Lefebvre, M., Millour, E., Vignon, E., Jouhaud, J., Diallo, F. B., Lott, F., Gastineau, G., Caubel, A., Meurdesoif, Y., and Ghattas, J.: LMDZ6A: The Atmospheric Component of the IPSL Climate Model With Improved and Better Tuned
705 Physics, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001892, 2020.

Hu, S., Xie, S.-P., and Liu, W.: Global Pattern Formation of Net Ocean Surface Heat Flux Response to Greenhouse Warming, Journal of Climate, 33, 7503–7522, https://doi.org/10.1175/JCLI-D-19-0642.1, 2020.

IPCC: Summary for Policymakers, in: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2022.

710 Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochemical Cycles, 19, https://doi.org/10.1029/2003GB002199, 2005.

Lago, V., Wijffels, S. E., Durack, P. J., Church, J. A., Bindoff, N. L., and Marsland, S. J.: Simulating the Role of Surface Forcing on Observed Multidecadal Upper-Ocean Salinity Changes, Journal of Climate, 29, 5575–5588, https://doi.org/10.1175/JCLI-D-15-0519.1, 2016.

715 Lehner, F., Deser, C., and Terray, L.: Toward a New Estimate of "Time of Emergence" of Anthropogenic Warming: Insights from Dynamical Adjustment and a Large Initial-Condition Model Ensemble, Journal of Climate, 30, 7739–7756, https://doi.org/10.1175/JCLI-D-16-0792.1, 2017.

Lehner, F., Deser, C., Maher, N., Marotzke, J., Fischer, E. M., Brunner, L., Knutti, R., and Hawkins, E.: Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6, Earth System Dynamics, 11, 491–508, https://doi.org/10.5194/esd-11-491-2020,
720 2020.

Liu, W., Lu, J., Xie, S.-P., and Fedorov, A.: Southern Ocean Heat Uptake, Redistribution, and Storage in a Warming Climate: The Role of Meridional Overturning Circulation, Journal of Climate, 31, 4727–4743, https://doi.org/10.1175/JCLI-D-17-0761.1, 2018.

Madec, G., Bourdallé-Badie, R., Bouttier, P.-A., Bricaud, C., Bruciaferri, D., Calvert, D., Chanut, J., Clementi, E., Coward, A., Delrosso, D., Ethé, C., Flavoni, S., Graham, T., Harle, J., Iovino, D., Lea, D., Lévy, C., Lovato, T., Martin, N., Masson, S., Mocavero, S., Paul, J.,
725 Rousset, C., Storkey, D., Storto, A., and Vancoppenolle, M.: NEMO ocean engine, https://doi.org/10.5281/ZENODO.3248739, publisher: Zenodo Version Number: v3.6-patch, 2017.

Marshall, J., Scott, J. R., Armour, K. C., Campin, J.-M., Kelley, M., and Romanou, A.: The ocean's role in the transient response of climate to abrupt greenhouse gas forcing, Climate Dynamics, 44, 2287–2299, https://doi.org/10.1007/s00382-014-2308-0, 2015.

Merino, N., Le Sommer, J., Durand, G., Jourdain, N. C., Madec, G., Mathiot, P., and Tournadre, J.: Antarctic icebergs melt over the Southern
730 Ocean: Climatology and impact on sea ice, Ocean Modelling, 104, 99–110, https://doi.org/10.1016/j.ocemod.2016.05.001, 2016.

Mignot, J., Hourdin, F., Deshayes, J., Boucher, O., Gastineau, G., Musat, I., Vancoppenolle, M., Servonnat, J., Caubel, A., Chéruy, F., Denvil, S., Dufresne, J., Ethé, C., Fairhead, L., Foujols, M., Grandpeix, J., Levavasseur, G., Marti, O., Menary, M., Rio, C., Rousset, C., and Silvy, Y.: The tuning strategy of IPSL-CM6A-LR, Journal of Advances in Modeling Earth Systems, https://doi.org/10.1029/2020MS002340, 2021.

735 Mikolajewicz, U. and Voss, R.: The role of the individual air-sea flux components in CO2-induced changes of the ocean's circulation and climate, Climate Dynamics, 16, 627–642, https://doi.org/10.1007/s003820000066, 2000.

Roquet, F., Madec, G., McDougall, T. J., and Barker, P. M.: Accurate polynomial expressions for the density and specific volume of seawater using the TEOS-10 standard, Ocean Modelling, 90, 29–43, https://doi.org/10.1016/j.ocemod.2015.04.002, 2015.

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

Shi, J.-R., Talley, L. D., Xie, S.-P., Liu, W., and Gille, S. T.: Effects of Buoyancy and Wind Forcing on Southern Ocean Climate Change, Journal of Climate, 33, 10 003–10 020, https://doi.org/10.1175/JCLI-D-19-0877.1, 2020.

Silvy, Y.: Emergence of temperature and salinity changes in the ocean interior in response to climate change: timescales and mechanisms,
745 Ph.D. thesis, Sorbonne Université, https://rgdoi.net/10.13140/RG.2.2.10983.93607, publisher: Unpublished, 2022.

Silvy, Y., Guilyardi, E., Sallée, J.-B., and Durack, P. J.: Human-induced changes to the global ocean water masses and their time of emergence, Nature Climate Change, https://doi.org/10.1038/s41558-020-0878-x, 2020.

Todd, A., Zanna, L., Couldrey, M., Gregory, J., Wu, Q., Church, J. A., Farneti, R., Navarro-Labastida, R., Lyu, K., Saenko, O., Yang, D., and Zhang, X.: Ocean-Only FAFMIP: Understanding Regional Patterns of Ocean Heat Content and Dynamic Sea Level Change, Journal of
750 Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS002027, 2020.

Winton, M., Griffies, S. M., Samuels, B. L., Sarmiento, J. L., and Frölicher, T. L.: Connecting Changing Ocean Circulation with Changing Climate, Journal of Climate, 26, 2268–2278, https://doi.org/10.1175/JCLI-D-12-00296.1, 2013.

Xie, P. and Vallis, G. K.: The passive and active nature of ocean heat uptake in idealized climate change experiments, Climate Dynamics, 38, 667–684, https://doi.org/10.1007/s00382-011-1063-8, 2012.

755 Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J., and Heimbach, P.: Global reconstruction of historical ocean heat storage and transport, Proceedings of the National Academy of Sciences, 116, 1126–1131, https://doi.org/10.1073/pnas.1808838115, 2019.

Zika, J. D., Skliris, N., Blaker, A. T., Marsh, R., Nurser, A. J. G., and Josey, S. A.: Improved estimates of water cycle change from ocean salinity: the key role of ocean warming, Environmental Research Letters, 13, 074 036, https://doi.org/10.1088/1748-9326/aace42, 2018.