Role of Ocean Model Formulation in Climate Response Uncertainty

JOHN P. KRASTING, RONALD J. STOUFFER,[a] STEPHEN M. GRIFFIES, ROBERT W. HALLBERG, SERGEY L. MALYSHEV, BONITA L. SAMUELS, AND LORI T. SENTMAN

NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey

(Manuscript received 19 January 2018, in final form 5 September 2018)

ABSTRACT

Oceanic heat uptake (OHU) is a significant source of uncertainty in both the transient and equilibrium responses to increasing the planetary radiative forcing. OHU differs among climate models and is related in part to their representation of vertical and lateral mixing. This study examines the role of ocean model formulation—specifically the choice of the vertical coordinate and the strength of the background diapycnal diffusivity K_d—in the millennial-scale near-equilibrium climate response to a quadrupling of atmospheric CO_2. Using two fully coupled Earth system models (ESMs) with nearly identical atmosphere, land, sea ice, and biogeochemical components, it is possible to independently configure their ocean model components with different formulations and produce similar near-equilibrium climate responses. The SST responses are similar between the two models (r^2 = 0.75, global average ~4.3°C) despite their initial preindustrial climate mean states differing by 0.4°C globally. The surface and interior responses of temperature and salinity are also similar between the two models. However, the Atlantic meridional overturning circulation (AMOC) responses are different between the two models, and the associated differences in ventilation and deep-water formation have an impact on the accumulation of dissolved inorganic carbon in the ocean interior. A parameter sensitivity analysis demonstrates that increasing the amount of K_d produces very different near-equilibrium climate responses within a given model. These results suggest that the impact of the ocean vertical coordinate on the climate response is small relative to the representation of subgrid-scale mixing.

1. Introduction

Model-based estimates of the equilibrium climate sensitivity to a doubling of atmospheric CO_2 presented in the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (IPCC 2013) fall in a likely range of 1.5°–4.5°C. Understanding the sources of uncertainty in climate models' response to changes in radiative forcing is an important part of climate research. Model differences in their representation of key climate processes contribute to this uncertainty range. Among the processes are clouds, aerosol effects (direct and indirect), and the ocean response to radiative forcing (Bindoff et al. 2013; Bony et al. 2006). In particular, the deep ocean responses on centennial to millennial time scales are relevant to the equilibrium climate response (Stouffer and Manabe 1999; Danabasoglu and Gent 2009; Li et al. 2013; Flato et al. 2013).

Deep and abyssal oceanic heat uptake (OHU) and sea surface temperature (SST) warming are inversely related (Gregory et al. 2004; Hansen et al. 2005; Meehl et al. 2011). For a given change in radiative forcing, more OHU at depth produces smaller increases in SST and vice versa. One can conceptually split the uncertainty in OHU into two parts: 1) ocean heat uptake *efficiency*, which directly relates ocean warming to radiative forcing, and 2) ocean heat uptake *efficacy* (Winton et al. 2010), which quantifies the feedback of warming-induced changes to the ocean circulation on heat uptake.

The amount of OHU in ocean models can be directly related to their representation of mesoscale eddies (Griffies et al. 2015) and other mixing processes. Many of the current-generation coupled models have horizontal resolutions that are insufficient to fully resolve eddies (>0.5°), and these models employ eddy parameterizations such as

[a] Current affiliation: Department of Geosciences, The University of Arizona, Tucson, Arizona.

Corresponding author: John P. Krasting, john.krasting@noaa.gov

DOI: 10.1175/JCLI-D-18-0035.1

those in Redi (1982) and Gent and McWilliams (1990). Vertical and diapycnal mixing is also often parameterized in ocean models, such as through the K-profile parameterization (Large et al. 1994). The choice and configuration of these subgrid-scale mixing parameterizations varies greatly among ocean models and has significant effects on the mean climate state (Weaver and Wiebe 1999; Schmittner and Weaver 2001; Danabasoglu and Marshall 2007; Melet et al. 2013, 2016) and the response to forcing (Hansen et al. 1985; Manabe and Stouffer 1999; Wiebe and Weaver 1999; Ehlert et al. 2017). It is an area of active study in both improving these parameterizations in more coarse-resolution models while leveraging recent advancements in high-performance computing to develop and run higher-resolution eddy-permitting ($\leq 0.25°$) and eddy-active ocean models ($\leq 0.1°$).

Ocean models also differ in their vertical coordinate. Level based (i.e., depth z) and density based (ρ) are among the two most-common vertical coordinates used for global climate studies and each has its strengths and weaknesses. Level-based coordinates more naturally represent the surface boundary layer and thus are frequently used for coupled climate models. Density-based coordinates excel at representing interior tracer transport along potential density surfaces and have better representations of flow interactions with the topography (Hallberg 1995; Griffies et al. 2000b). Depth-based coordinate models are more susceptible to warming in the abyssal ocean than ρ-coordinate models (Griffies et al. 2000a). This drift results from excessive entrainment in dense-water overflows that fill the abyssal ocean (Winton et al. 1998) and numerically induced spurious diapycnal mixing (Griffies et al. 2000a; Ilicak et al. 2012; Megann 2018). Density-based coordinates tend to have better water mass formation and preservation and less spurious mixing in the ocean interior than their z-coordinate counterparts (Delworth et al. 2006; Megann et al. 2010; Dunne et al. 2012).

The results of Danabasoglu and Gent (2009) demonstrate that the equilibrium climate sensitivity (ECS) derived from a 50-yr slab model simulation agrees well with the ECS obtained from a 3000-yr simulation of the fully coupled NCAR CCSM3 model. The global-mean SST response between these two NCAR simulations agree within 0.14°C. As the ocean equilibrates in both the slab and coupled model configurations, increases in sensible and latent heat fluxes with the atmosphere balance the increased downward longwave heat flux that increases the SST. Thus, the authors argue that the equilibrium response of the coupled system depends on only the atmospheric and SST responses to the increased planetary radiative forcing and is ''independent of the ocean model formulation.'' This argument is rooted in the way many coupled models exchange heat and momentum across the air–sea interface, including the National Center for Atmospheric Research (NCAR) and Geophysical Fluid Dynamics Laboratory (GFDL) models.

This work seeks to quantify how much and in what aspects the choice of ocean model formulation influences the response to climate forcing. In support of phase 5 of the Coupled Model Intercomparison Project (CMIP5), GFDL developed two coupled carbon cycle–climate Earth system models (ESMs), which differ mainly in their ocean model formulation and, principally, in their vertical coordinate. The first model, GFDL-ESM2Mb, is based on the Modular Ocean Model, version 4.1 (Griffies et al. 2009), where the advective terms and physical parameterizations are computed with a z^* vertical coordinate, where the asterisk denotes the inclusion of a nonlinear free surface discussed in the following section. The second model, GFDL-ESM2G, is based on the Generalized Ocean Layer Dynamics model (Hallberg and Adcroft 2009), which employs an isopycnal ρ-coordinate. In this study, we use these two models to explore the role of ocean formulation—both the choice of the vertical coordinate and the implementation of subgrid-scale mixing—on the near-equilibrium climate response in the models. To date, there are few studies that examine the millennial-scale differences between z- and ρ-based ocean models.

Evaluation of the models' near-equilibrium response to forcing is based on a series of idealized simulations with increasing atmospheric CO_2 concentrations ($1\%\,yr^{-1}$ to quadrupling of the preindustrial value of 286 ppmv) performed with both of the GFDL ESMs. Dunne et al. (2012, 2013) document and evaluate the physical climate and carbon cycle simulations for both models. Prior studies comparing ocean heat uptake across opportunistic ensembles of climate models (e.g., CMIP; Taylor et al. 2012) often suffer from large and important differences in other components of the atmosphere–ocean general circulation models (AOGCMs). The present analysis eliminates those sources of uncertainty by using virtually identical atmosphere, land, and sea ice components in both of the ESMs. The terrestrial ecosystem and ocean biogeochemistry components are also very similar between the two ESMs, allowing for comparison of their carbon cycle responses.

During the development of ESM2Mb and ESM2G, both models were configured with equivalent physics choices and parameterization settings that are constrained by both observations and theory. Nevertheless, some differences remain between the physics of both models and are discussed in more detail in the next section. Both ESMs produce plausible simulations of the preindustrial climate (Dunne et al. 2012) and do not

exhibit excessive climate drift over the 1000+ years of model simulation required to spin up the physical climate and biogeochemical tracers in the ocean to a near-equilibrium state. A key result of this study is that despite having slightly different spunup preindustrial states, both ESMs produce remarkably similar climate *responses* across a range of climate diagnostics.

Following a short model and experimental design description, this paper presents a large-scale overview of those responses obtained from ESM2Mb and ESM2G. This overview includes comparisons of ocean temperature, salinity, sea ice, meridional circulation and heat transport, and dissolved inorganic carbon. The paper concludes with an analysis of varying background diapycnal diffusivity K_d in the ESM2G model. This analysis aims to demonstrate that physical parameterization configurations influence ocean model simulations more than the choice of vertical coordinates (Griffies and Treguier 2013). While the near-equilibrium responses are broadly consistent between the two standard versions of the models, varying the amount of subgrid-scale mixing has a much larger effect on the millennial-scale response of the model to increased amounts of atmospheric CO_2.

2. Model description and experimental design

The ESMs used in this study, GFDL-ESM2Mb and GFDL-ESM2G, feature an AOGCM coupled to a fully interactive carbon cycle. The atmospheric component is based on AM2 (Anderson et al. 2004) with a resolution of ~2° and is very similar to the version used in the GFDL CM2.1 coupled climate model (Delworth et al. 2006). The ESMs use an upgraded land model (LM3.0; Milly et al. 2014), which features improved hydrology and includes terrestrial ecosystem dynamics (Shevliakova et al. 2009). Both ESMs use GFDL's Sea Ice Simulator (SIS), version 1 (Winton 2000; Delworth et al. 2006).

The carbon cycle in both models consists of a prognostic, radiatively active CO_2 tracer shared among the atmosphere, land, and ocean biogeochemistry components of the model. The ESM2Mb model differs slightly from the version submitted to CMIP5 (named GFDL-ESM2M) in that by adopting the same land model parameters as in ESM2G, the aboveground terrestrial biomass was reduced so that both ESMs are consistent at ~850 GtC. The ESMs use the Tracers of Ocean Phytoplankton and Allometric Zooplankton (TOPAZ) biogeochemistry model (Dunne et al. 2013) coupled to each of the ESM's physical ocean components.

The primary difference between the two ESMs lies in their physical ocean model. ESM2Mb uses the Modular Ocean Model, version 4.1 (MOM4p1; Griffies et al. 2009), configured using a z^* coordinate with 50 vertical levels and includes a nonlinear free surface. The z^* coordinate rescales the vertical coordinate with the free surface height so that surface undulations are spread throughout the vertical column (Stacey et al. 1995; Adcroft and Campin 2004). ESM2G uses the Generalized Ocean Layer Dynamics model (GOLD; Hallberg and Adcroft 2009) based on an isopycnal model developed by Hallberg (1995). GOLD features a density-based vertical coordinate with 63 vertical layers. Both models' grids have a nominal resolution of 1° that increases to 1/3° in the tropics. The grids transition from spherical to tripolar north of 65°N, which results in relatively fine resolution across the Arctic.

Both MOM4p1 and GOLD use a mesoscale eddy parameterization based on the work of Gent and McWilliams (1990). The strength of the parameterized mixing is dependent on the local flow and isopycnal slope. The allowable ranges for the local coefficients span 100–800 and 10–900 $m^2 s^{-1}$ for MOM4p1 and GOLD, respectively (Dunne et al. 2012), with these ranges being dependent on the grid resolution. MOM4p1 includes vertical mixing parameterizations based on the K-profile parameterization (Large et al. 1994; Danabasoglu et al. 2006) while GOLD uses a bulk mixed layer scheme based on Hallberg (2003) with a shear mixing parameterization based on Jackson et al. (2008). In both MOM4p1 and GOLD, mixed layer restratification by eddies is based on the scheme of Fox-Kemper et al. (2011) and the models allow for chlorophyll attenuation of surface shortwave radiation. Both models use the tidal mixing scheme of Simmons et al. (2004) and implement a geothermal heating flux (Adcroft et al. 2001).

The most notable difference between the two ocean models is in their treatment of background diapycnal diffusivity K_d. The seminal estimates of K_d described in Munk (1966) suggest a value of $1 \times 10^{-4} m^2 s^{-1}$ while several recent observational estimates suggest K_d that is an order of magnitude smaller (e.g., Gregg 1987; Ledwell et al. 1993, 1998, 2011). In this configuration of MOM4p1, K_d has values ranging between $1 \times 10^{-5} m^2 s^{-1}$ equatorward of 30° latitude and $1.5 \times 10^{-5} m^2 s^{-1}$ poleward of this latitude. In GOLD, background K_d has a larger pole-to-equator gradient ranging from a value of $2 \times 10^{-6} m^2 s^{-1}$ at the equator to $2 \times 10^{-5} m^2 s^{-1}$ poleward of 30° latitude. The increased K_d used in GOLD was necessary to counteract an extremely unrealistic cold drift that developed in the abyssal ocean when the model used the same values as MOM4p1. Another notable difference between the models is that GOLD includes the bottom-drag mixing scheme from Legg et al. (2006), which is important in representing overflows (Legg et al. 2009). We refer the reader to Dunne et al. (2012) for a

more complete discussion of the similarities and differences of both ocean model components.

The ocean components of both coupled models were initialized from present-day temperature and salinity fields and were spun up with fixed preindustrial CO_2 concentrations (286 ppmv) until reaching a near-equilibrium climate state (1600 years for ESM2G and 2400 years for ESM2Mb). Following the spinup simulations, each model was integrated further for another 5000 years to obtain a long, continuous control simulation. Starting from the 1860 control integration and following the CMIP5 protocol (Taylor et al. 2012), atmospheric CO_2 increased in the model at a rate of 1% yr^{-1} compounded until reaching 4 times the control integration's value (1144 ppmv). The models achieve this $4 \times CO_2$ concentration near model year 140 and remain fixed at this increased value throughout the rest of the $4 \times CO_2$ integration. All other forcing (i.e., well-mixed greenhouse gases, aerosols, and solar) were held at their preindustrial values and land-use forcing was not applied in any of the described simulations. The effects of model drift are accounted for by referencing a corresponding 5000-yr piControl simulation performed with each ESM. The first 500 years of the control simulations and the first 300 years of the $4 \times CO_2$ simulations are available for download through the Earth System Grid Federation (http://esgdata.gfdl.noaa.gov). The remaining lengths of the simulations are available upon request.

To explore the role of subgrid-scale mixing on the near-equilibrium response, we evaluate a three-member ensemble of 5000-yr $4 \times CO_2$ simulations with different values for the diapycnal diffusivity in the ESM2G model. Although the horizontal and vertical distributions of mixing have an impact on the ocean state (Melet et al. 2016), we apply a spatially uniform increase to the background K_d in this sensitivity analysis for simplicity and its ability to directly influence the main thermocline region. The ensemble consists of simulations where the background K_d (2×10^{-5} m^2 s^{-1}) increases to 4×10^{-5}, 6×10^{-5}, and 1×10^{-4} m^2 s^{-1}. Prior studies have noted that increasing the value of K_d tends to increase the ocean heat uptake and reduce surface air temperature as the climate warms (Dalan et al. 2005). This effect is evaluated in the ESM2G model and is presented as a means of contextualizing the results obtained from varying the ocean vertical coordinate.

3. Results and discussion

a. Large-scale surface response

ESM2Mb and ESM2G have remarkably stable global annual average SSTs in their piControl simulations (Fig. 1a, Table 1). Based on a linear least squares regression over the 5000-yr control simulation, ESM2Mb has a slight warming trend (0.002°C $century^{-1}$) and ESM2G has a slight cooling trend (−0.001°C $century^{-1}$). The detrended time series of global average SSTs from the control runs are shown in Fig. 1c. The global average SST difference between the two models in the piControl simulations is 0.43°C.

As concentrations of atmospheric CO_2 increase, the SSTs warm in both models. This warming continues throughout the 5000-yr integration even after the CO_2 concentration stabilizes at 4 times the preindustrial control value and the model adjusts to the increase in radiative forcing. The near-equilibrium warmings after 5000 years in area mean SST are 4.92° and 5.04°C for ESM2Mb and ESM2G, respectively. In both ESMs, nearly half of this warming (43%) occurs during the first 200 years when CO_2 concentrations are increasing. The rest of the warming occurs over the remainder of the simulation, similar to the results of Wetherald et al. (2001). The responses of the two models are remarkably similar in both timing and magnitude. When the responses are normalized by the mean difference between the two models' control run SSTs (0.43°C), the response curves from the two models lie nearly on top of each other (Fig. 1e). The similarity persists across centennial to millennial time scales as both models slowly adjust to the increased radiative forcing.

The spatial patterns of SST response early in the experiment (averaged over years 1–100; Figs. 2a,b) are also similar between the two models with a pattern correlation r^2 of 0.85. The spatial patterns are also similar to prior studies (Manabe et al. 1991; Andrews et al. 2015; Marshall et al. 2015; Armour et al. 2016; Gregory et al. 2016; Garuba et al. 2018). Warming over the Southern Ocean is largely limited to 0.5°C and there is cooling by as much as 1°C over the convective regions of the North Atlantic. In the tropical Atlantic and western Pacific Ocean basins, SSTs warm by about 1°C along the equator. Both models also have a warming maximum of about 1°C located in the central Pacific and Atlantic Oceans near 20°N and again near 50°N. While the magnitude of the warming is similar between the two models, the amplitudes of the local minima and maxima are generally larger in ESM2G than in ESM2Mb.

As the atmosphere warms and the saturation vapor pressure increases, the hydrologic cycle intensifies and the atmosphere transports more water poleward (Held and Soden 2006). In the tropics, enhanced evaporation causes surface waters to become saltier in these regions while increased precipitation contributes to freshening at high latitudes (Manabe et al. 1991; Durack et al. 2012). Both models exhibit this salinity response pattern,

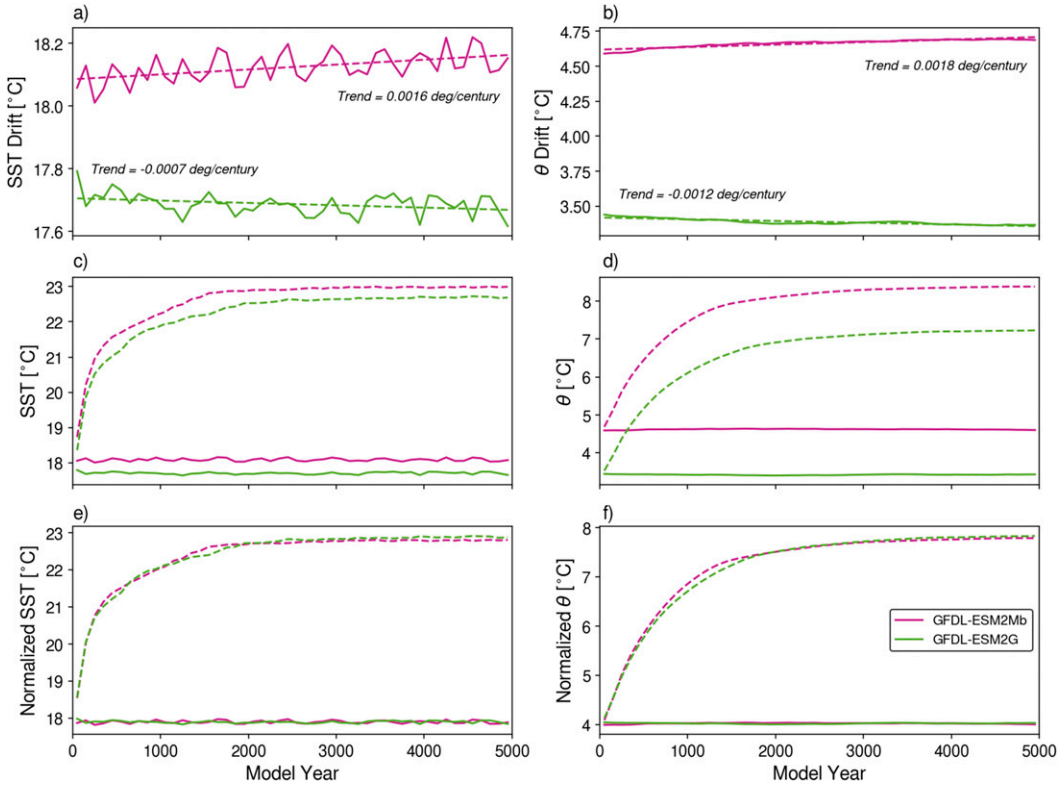

FIG. 1. (left) Area-mean SST (°C) and (right) volume-mean ocean potential temperature θ (°C) for GFDL-ESM2Mb (magenta) and GFDL-ESM2G (green). (a),(b) The drift over the 5000-yr piControl simulations with the dashed line illustrating the linear trend. (c),(d) The detrended time series of SST and θ, where the solid lines are the piControl simulations and dashed lines are the $4 \times CO_2$ simulations. (e),(f) The normalized time series in (c) and (d) where the time-mean difference between the two models' piControl simulations is removed. These differences are 0.43° and 1.27°C for SST and θ, respectively.

but with less agreement between them than when compared to the SST response pattern (Figs. 2e,f; r^2 = 0.68).

Similar to SST, ESM2G has a larger-amplitude sea surface salinity (SSS) response than ESM2Mb in the Atlantic. As part of the basinwide overturning circulation, the Gulf Stream transports warmer, saltier surface water from the tropics to the high latitudes where it cools to form North Atlantic Deep Water (NADW). While the atmosphere primarily drives the SSS response pattern in the first 100 years of forcing, the larger amplitude of the response pattern in ESM2G (i.e., saltier tropics and fresher high latitudes) is consistent with a weaker Atlantic meridional overturning circulation (AMOC) and less poleward mass transport of water from the tropics to the high latitude that will be discussed in more detail. ESM2G also shows an increase in SSS in the Barents Sea region of 0.4 psu compared to ESM2Mb, likely reflecting an increase in ESM2G's ocean convection in this region as the atmospheric CO_2 increases.

The SST response by the time of near equilibrium in the experiment (averaged over years 4901–5000; Figs. 2c,d) is also similar between the two models (r^2 = 0.75). Warming in excess of 6°C occurs in the high-latitude

TABLE 1. Global-mean SST (°C) and volume-mean ocean potential temperature θ (°C) for years 4901–5000 from the piControl and $4 \times CO_2$ simulations. Responses relative to the piControl simulation are shown in parentheses.

	GFDL-ESM2Mb	GFDL-ESM2G
SST, control mean	18.12	17.69
SST, $4 \times CO_2$, years 1–100	18.73 (0.61)	18.36 (0.67)
SST, $4 \times CO_2$, years 101–200	20.23 (2.11)	19.87 (2.18)
SST, $4 \times CO_2$, years 4901–5000	23.07 (4.95)	22.65 (4.96)
θ control mean	4.66	3.39
θ, $4 \times CO_2$, years 1–100	4.68 (0.02)	3.53 (0.14)
θ, $4 \times CO_2$, years 4901–5000	8.47 (3.81)	7.17 (3.78)

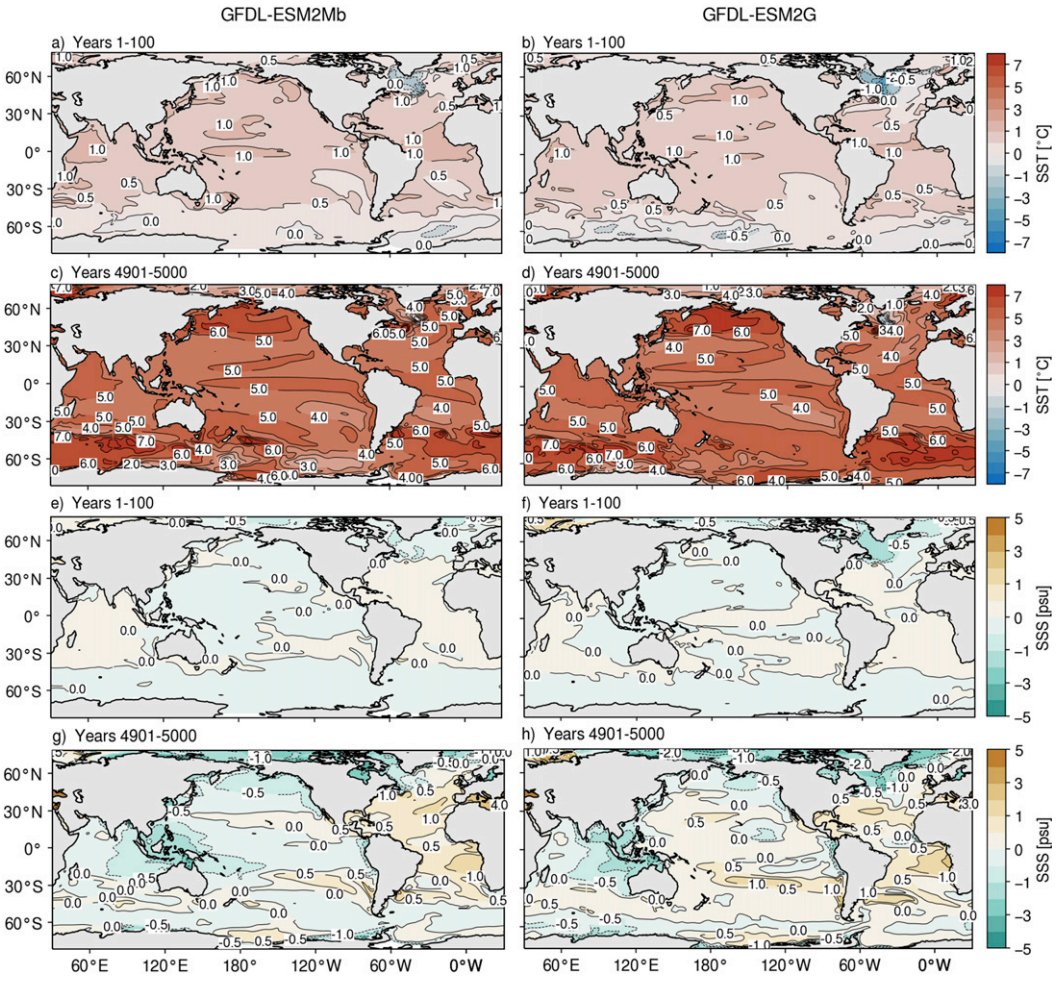

FIG. 2. Spatial patterns of (a)–(d) SST response (°C) and (e)–(h) sea surface salinity (SSS; psu) response for years 1–100 and 4901–5000 of the 4 × CO_2 simulations. Differences are relative to the corresponding years of the piControl simulations.

regions that are equatorward of the sea ice edge, with the area of warming extending across the entire North Pacific in ESM2G. The Southern Ocean also warms in excess of 6°C, especially off the coast of East Antarctica in ESM2Mb and north of the Weddell Sea in ESM2G. These results are broadly consistent with those found in previous studies (e.g., Manabe and Stouffer 1994). The SST response pattern here is also similar to earlier atmosphere–mixed layer ocean results (Manabe and Stouffer 1980).

ESM2G and ESM2Mb have similar SSS responses to a quadrupling of atmospheric CO_2 in both their magnitude and spatial pattern ($r^2 = 0.83$). Both models freshen in the Arctic Ocean by about 2 psu (Figs. 2g,h) in their near-equilibrium state. The combination of increased net water flux into the Arctic basin and sea ice melt contributes to this freshening (Duffy et al. 2001). Both model results have a saltier Atlantic Ocean from 60°N to 30°S. Both models also have a freshening of approximately 1 psu in the Indian Ocean. In the Pacific, the comparison is more complex. In the eastern part of the Pacific basin, both models have two areas of increased salinity around 30°N and 30°S. In the western part of the Pacific basin, ESM2G becomes more saline while ESM2Mb slightly freshens.

Both models differ in their simulation of preindustrial Arctic sea ice, but their responses are similar (Fig. 3 and Table 2). March sea ice extent, defined as the total area of grid cells with at least 15% coverage of sea ice, is 18.48 million km^2 in ESM2Mb and 22.75 million km^2 in ESM2G averaged over the 5000-yr preindustrial control run. This represents about a 20% greater sea ice extent in ESM2G compared to ESM2Mb. The relative sea ice extent difference between the two models is greater

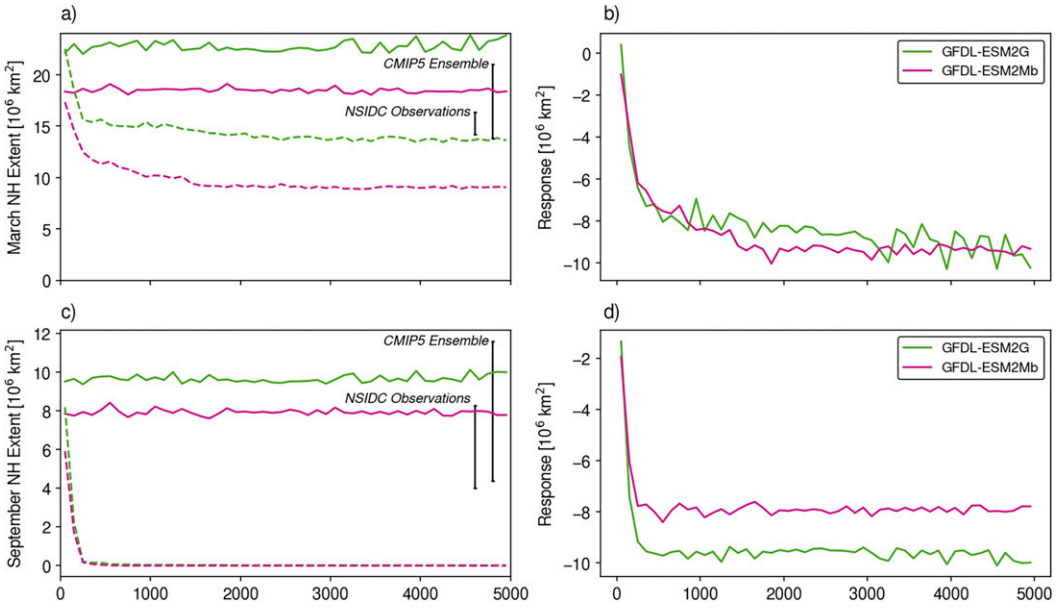

FIG. 3. Northern Hemisphere sea ice extent (million km^2) for (a),(b) March and (c),(d) September for GFDL-ESM2Mb (magenta) and GFDL-ESM2G (green). Solid lines are the piControl simulations and dashed lines are the $4 \times CO_2$ simulations. The black bars denote the CMIP5 historical simulation ensemble range corresponding to year 1900 (Stroeve et al. 2012) and present-day sea ice extent (mean plus and minus two standard deviations) averaged over years 1979–2018 from NSIDC (Fetterer et al. 2017).

during September (~30%), with ESM2Mb averaging 7.93 million km^2 and ESM2G averaging 9.66 million km^2. The larger sea ice extent in the ESM2G control run is consistent with the cooler climatological SSTs in this model as well as with the model's weaker AMOC circulation.

An analysis of Arctic sea ice extent across an ensemble of CMIP5 models from Stroeve et al. (2012) is provided here for context. The ensemble range from the historical simulations corresponding to year 1900 is shown for comparison against the piControl simulations of ESM2Mb and ESM2G. The Stroeve et al. (2012) analysis includes output from 20 different AOGCMs, but does not include the two GFDL ESMs. ESM2Mb's March extent is within the range of CMIP5 models while ESM2G is outside of the ensemble range. Both models have September sea ice extent that is within the CMIP5 ensemble range. Observed sea ice extent (climatological mean $\pm 2\sigma$) from the National Snow and Ice Data Center Sea Ice Index, version 3 (Fetterer et al. 2017), averaged over years 1979–2018 is also shown for reference.

Despite different preindustrial climatologies in Arctic sea ice, the two models' responses to $4 \times CO_2$ forcing are remarkably similar. In March, both models show a sea ice extent reduction greater than 9 million km^2, and the sea ice extent stabilizes past year 3000. The March sea ice becomes more variable in ESM2G over the course of the $4 \times CO_2$ simulation, which is likely the result of variability changes in AMOC and associated high-latitude oceanic convection. September sea ice is also similar between the models, with the Arctic becoming ice free after the first few centuries. While the extent in September is less than that of March, it is much more important in terms of albedo given that this is the season of greatest insolation in the Arctic. In this case, the difference in the response curves

TABLE 2. Arctic sea ice extent (million km^2) for March and September from the piControl simulations, the $4 \times CO_2$ simulations, and the response ($4 \times CO_2$ simulations minus piControl) in parentheses.

	GFDL-ESM2Mb	GFDL-ESM2G
March extent, piControl	18.48	22.75
March extent, $4 \times CO_2$, years 1–100	17.34 (−1.14)	22.52 (−0.23)
March extent, $4 \times CO_2$, years 4901–5000	9.06 (−9.42)	13.64 (−9.11)
September extent, piControl	7.93	9.66
September extent, $4 \times CO_2$, years 1–100	5.92 (−2.01)	8.20 (−1.46)
September extent, $4 \times CO_2$, years 4901–5000	0.0 (−7.93)	0.01 (−9.65)

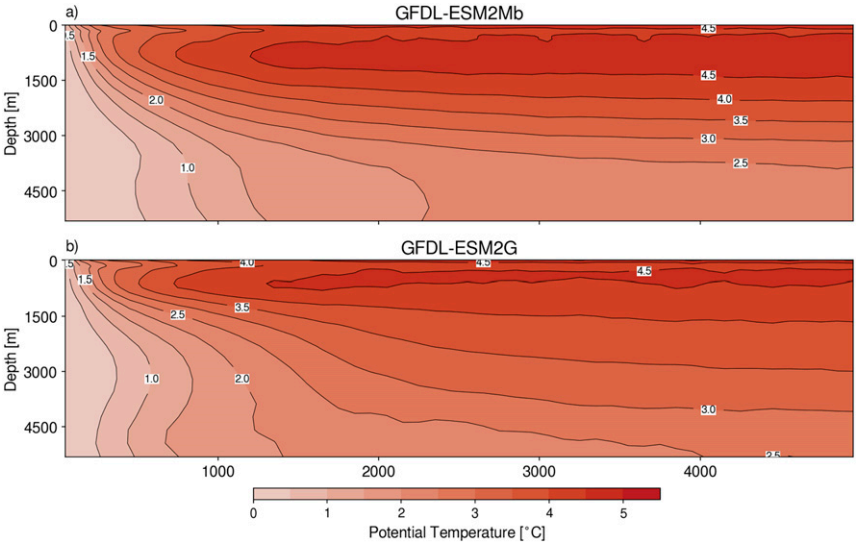

FIG. 4. Depth vs time contour plots of potential temperature response relative to the piControl simulation for the 5000-yr 4 × CO_2 simulations for (a) GFDL-ESM2Mb and (b) GFDL-ESM2G. Contour and shading interval is 1°C.

mainly reflects differences in the models' initial preindustrial states.

b. *Ocean interior response*

1) DEPTH VERSUS TIME TEMPERATURE RESPONSE

As was the case with SST, the detrended global volume-mean ocean potential temperature in the piControl simulation is offset between the models, with ESM2G approximately 1.3°C colder than ESM2Mb (Fig. 1d). This difference is associated with a warming trend in ESM2Mb and a small cooling trend in the ESM2G over the course of the spinup simulations prior to the start of the piControl integrations. During the piControl integration, very small temperature trends are evident in both models (Fig. 1b). The exact causes for these trends are unclear, but they are related in part to differences in the models' ability to simulate dense water overflows. Unless the lateral and vertical resolution are both fine enough to resolve the bottom boundary layer, z-coordinate models are subject to excessive entrainment (Winton et al. 1998) while this problem does not impact isopycnal-coordinate models.

The responses of the volume-mean ocean potential temperature from both models are normalized by their mean offset from the control run (1.2°C; Fig. 1f). The rates and magnitudes of ocean warming over the course of the 5000-yr simulation are similar between the two models although ESM2Mb warms at a slightly faster rate than in ESM2G around year 1000. By the end of the 4 × CO_2 integrations, the rate of warming becomes very small and both models are nearly in equilibrium with the increased radiative forcing.

There are differences in the internal warming responses between the two models. The near-surface layers warm faster than the deeper layers in both models during the 4 × CO_2 simulation, with the largest warming occurring at 750-m depth. Compared to ESM2G, subthermocline warming is stronger in ESM2Mb as evidenced by the 4.5°C contour extending downward to almost 1500 m in Fig. 4. However, the magnitude of deep warming below 2000 m is also larger in ESM2G than in ESM2Mb by the end of the experiment. The difference is related to the models' treatment of Antarctic Bottom Water (AABW); AABW in ESM2G is systematically colder than ESM2Mb and there is more warming in the ocean interior when AABW formation slows in the transient simulation. Although the volume-mean ocean heat uptake is similar between these two models, the different warming patterns at depth will have an impact on spatial sea level rise, particularly during the transient phase of the simulation and during the peak AMOC decline (Krasting et al. 2016).

The magnitude of the warming near the bottom of the ocean in both models is smaller than the warming near the surface, pointing to changes in the origin of water masses at and near the surface. Bottom waters are formed near the ocean surface in high latitudes of the North Atlantic and Southern Oceans, where the warming is smaller than in midlatitudes (see Figs. 2c,d). Surface and intermediate water masses form in midlatitudes and tend to flow equatorward at depth and contribute

to lower thermocline water (near 750 m). These water masses experience a larger magnitude of warming than the water masses that fill the bottom of the World Ocean.

One difference in the depth-versus-time temperature response between the two models is seen near the bottom of the ocean in the first few centuries of the integrations. ESM2G warms at a faster rate than ESM2Mb. During the spinup, cooling of the abyssal waters dominates ESM2G's global volume-mean temperature cold drift. The initial cooling in the $4 \times CO_2$ simulation reflects a continuation of these processes during the first several centuries. Once the warming signal propagates from the surface to the bottom of the ocean, the cooling trend reverses and the warming rate is similar in magnitude to that in ESM2Mb.

By the end of the experiment, the depth of the isotherms does not change much in time, indicating that both models are approaching an equilibrium state. Manabe and Stouffer (1994) found that it took about 1500 model years for the ocean temperatures near the bottom of the ocean to reach 70% of their equilibrium value. Inspection of Fig. 4 indicates that both ESM2Mb and ESM2G give very similar response time scales to that found in Manabe and Stouffer (1994).

2) Depth-versus-latitude temperature and salinity response

During the first 100 years of the $4 \times CO_2$ simulations, the greatest warming occurs at the surface in both models (Figs. 5a,b). The 0.5°C contour is relatively shallow in the tropics and extends downward to about 500 m at 40° latitude in each hemisphere. Both models show zonal-mean surface cooling north of 50°N, but it is larger in magnitude in ESM2G where the cooling exceeds 1°C. Since 100 years is short relative to the ventilation time scales of the deep ocean, much of the ocean interior shows little response to forcing. After 5000 years (Figs. 5c,d), however, the ocean interior has warmed in response to the $4 \times CO_2$ forcing. The temperature gradient as a function of depth is stronger in ESM2Mb compared to ESM2G. The most pronounced warming in both models occurs in the Southern Ocean between 40° and 55°S.

In the Southern Ocean, a maximum area of warming exceeding (>6°C) extends equatorward and downward to about 1500 m in ESM2Mb. A similar area also occurs in ESM2G, but is more confined to the surface, and the latitudinal temperature gradients are less sharp. Heat uptake in this region occurs as water upwelling near the Antarctic sea ice edge is warmed by the atmosphere and transported downward and equatorward along isopycnals. In the piControl simulation, this downward heat flux balances an upward heat flux associated with eddy diffusion. The warming of the subsurface ocean in this region reflects either a weakening of the eddy diffusion heat flux or an increase in the Ekman-driven upwelling and advective heat flux (Morrison et al. 2016; Gregory 2000; Marshall and Zanna 2014). The atmospheric wind stress responses are similar between the two models, with a slight increase in magnitude and a shift poleward. Therefore, much of the Southern Ocean heat uptake differences are likely related to their representation of along-isopycnal mixing in this region.

Temperature differences associated with changes in high-latitude deep-water formation are also evident after 5000 years. Along the Antarctic coast, both models show a similar magnitude response extending downward from the surface. In ESM2G, this warming is stronger and it extends all the way to the bottom of the ocean in this region, exceeding 3°C along the shelf. This warming is limited to ~2°C in ESM2Mb. Similarly in the North Atlantic, ESM2G's warming in the NADW formation regions is approximately 5°C in ESM2G and 4°C in ESM2Mb. The signature of a reduced AMOC is also evident in the slope of the temperature contours, with a pronounced shoaling in the tropics.

The changes in zonal-averaged salinity are consistent with the mechanisms responsible for the temperature changes. After 100 years (Figs. 5e,f), most of the changes are confined to the surface. The North Atlantic freshening (>0.2 psu) responsible for the reduction in convection in ESM2G is particularly evident. After 5000 years, the bottom third of the ocean in both models freshens relative to the piControl simulation. This freshening also extends upward to the surface along the coast of Antarctica. Both models also exhibit a broad area of salinification of 0.1–0.2 psu that occurs in the tropics and midlatitudes. This region extends downward to about 3000 m in ESM2Mb, while only extending downward to ~1500 m in the tropics in ESM2G.

The magnitude of the response in ESM2G is generally larger than in ESM2Mb. Both models strengthen the intense halocline found in the Arctic Ocean and have an increase in salinity from about 250 m to the ocean bottom at that latitude. Both models also have a broad region of increased salinity from 500 to 2500 m, between 60°S and 40°N. Within this region, the largest increases occur between 20° and 40°N and above 1500 m. This region of increased salinity is slightly deeper in ESM2Mb than in ESM2G. There is an increase in salinity at depth near 30°N that is associated with the Mediterranean water mass changes. Model projections of future climate (Manabe and Stouffer 1980) suggest that the Mediterranean is a region of intense evaporation that further intensifies as the climate warms.

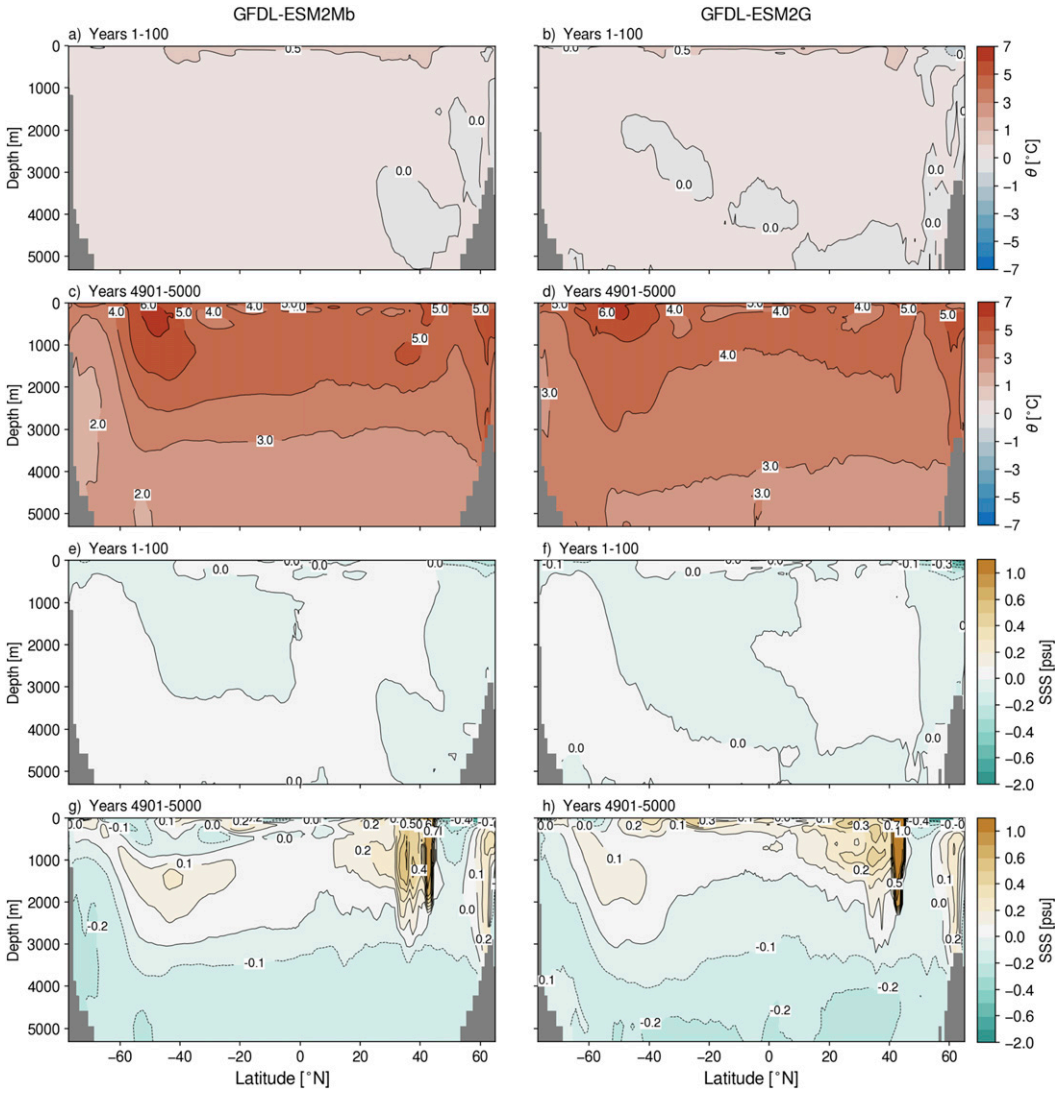

FIG. 5. Century time-average depth vs latitude sections of (a)–(d) potential temperature response (°C) and (e)–(h) salinity response (psu). Sections are shown for years 1–100 and 4901–5000 from the 4 × CO_2 simulations relative to the piControl simulation for (left) GFDL-ESM2Mb and (right) GFDL-ESM2G.

There are several regional difference between the two ESMs. Subsurface salinity increases of 0.4 psu occur in the tropics in ESM2G while they are less pronounced in ESM2Mb. A tongue of fresher water (0.1–0.2 psu) also extends downward to 1000 m in ESM2Mb that accompanies the region of maximum heat uptake in the Southern Ocean while this feature is less pronounced in ESM2G.

c. *AMOC, heat transport, and ventilation response*

In both models, the time series of the AMOC strength at 26.5°N, or the latitude of the Rapid Climate Change monitoring array (RAPID; Johns et al. 2011), is stable throughout the piControl simulations (Fig. 6a, Table 3). AMOC is slightly stronger in ESM2Mb (23.1 Sv, where 1 Sv ≡ $10^6\,m^3\,s^{-1}$) compared to ESM2G (20.3 Sv). Both models are within the observed range of variability from the RAPID array, or 18.5 ± 4.9 Sv.

As the atmospheric CO_2 concentration increases and the climate warms, the AMOC weakens in both models (Flato et al. 2013). The maximum weakening occurs during the second century of the 4 × CO_2 simulations, with AMOC strength decreasing by 7.3 Sv in ESM2Mb and 9.2 Sv in ESM2G. As the atmospheric CO_2 concentration stabilizes, the AMOC begins to recover in subsequent centuries. ESM2G recovers about half of the initial weakening relative to the piControl and remains stable after year 3000. The AMOC recovers fully by year

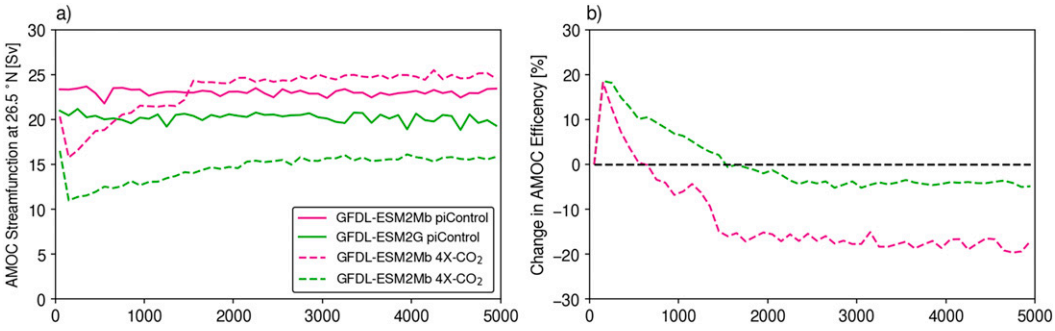

FIG. 6. (a) Time series of the century-averaged maximum value of the AMOC (Sv) at the latitude of the RAPID observing array (26.5°N) for GFDL-ESM2Mb (magenta) and GFDL-ESM2G (green). Solid lines represent the piControl simulations and dashed lines represent the $4 \times CO_2$ simulations. (b) Percentage change in AMOC efficiency, where the efficiency is defined as the ratio of the annual maximum Atlantic total poleward heat transport divided by the annual maximum overturning streamfunction.

1300 in ESM2Mb, similar to Stouffer and Manabe (1999). Unlike ESM2G, the AMOC in ESM2Mb increases relative to the piControl run by 1.5 Sv by the end of the $4 \times CO_2$ simulation.

The structure of AMOC, while showing broad large-scale agreement, highlights some of the key differences between the two ocean models (Figs. 7a,b). In both models, water enters the Atlantic basin in the Southern Hemisphere and flows northward within the top 500 m on the ocean. As the water reaches the high latitudes and cools, NADW forms and sinks to near the bottom of the ocean around 50°N, as denoted by the 0-Sv contour. AMOC extends deeper into the Atlantic by several hundred meters in the ESM2G model compared with ESM2Mb and these differences between the models are related to their treatment of overflows (Wang et al. 2018). At the southern limit of the Atlantic basin, NADW is found mainly above 3000 m, again as denoted by the 0-Sv contour. The highest streamfunction values are found between 1000 and 1500 m. As noted above, the overall magnitude of the AMOC in ESM2G is smaller than in ESM2Mb. The deepest circulation in the model oceans is made up of AABW, whose boundary is also denoted by the 0-Sv contour. The spatial structures of the AABW overturning cell in the control simulations are similar in pattern and magnitude between the two models.

The North Atlantic freshens in both models as the climate warms in response to the CO_2 forcing. There are several factors that contribute to this freshening. The effects of an enhanced hydrologic cycle and increases in high-latitude precipitation and runoff are evident in observations (Durack et al. 2012), while sea ice melt (Duffy et al. 2001) and a strengthening Atlantic subpolar gyre (Tesdal et al. 2018) also contribute to freshening. This freshening reduces convection at high latitudes, reduces NADW formation, and is responsible for the initial reduction in AMOC in both models. Simultaneously, enhanced evaporation over the tropical Atlantic enhances salinity (Durack et al. 2012). This higher-salinity water is transported northward by the Gulf Stream and a portion of this water becomes part of the southward-flowing horizontal gyre circulation. In ESM2G, much of this high-salinity water remains in the gyre circulation, enabling the model to maintain the fresh cap at the surface and its reduced convection. In ESM2Mb, however, more of the high-salinity surface water from the tropics is able to ''leak'' out of the gyre

TABLE 3. AMOC (Sv) at 26.5°N averaged over years 1–100 and 4901–5000 from the piControl and $4 \times CO_2$ simulations. Values for the maximum Atlantic poleward heat transport (total, overturning, and gyre contributions; PW) correspond to years 4901–5000 from both the piControl and $4 \times CO_2$ simulations. Responses relative to piControl are in parentheses in all cases.

	GFDL-ESM2Mb	GFDL-ESM2G
AMOC, piControl	23.07	20.25
AMOC, $4 \times CO_2$, years 101–200	15.79 (−7.28)	11.01 (−9.23)
AMOC, $4 \times CO_2$, years 4901–5000	24.6 (1.52)	15.87 (−4.37)
Total heat transport, piControl	1.11	1.11
Total heat transport, $4 \times CO_2$	1.0 (−0.11)	0.92 (−0.19)
Overturning heat transport, piControl	1.32	1.26
Overturning heat transport, $4 \times CO_2$	1.23 (−0.09)	1.14 (−0.12)
Gyre heat transport, piControl	0.44	0.36
Gyre heat transport, $4 \times CO_2$	0.43 (0.01)	0.34 (0.02)

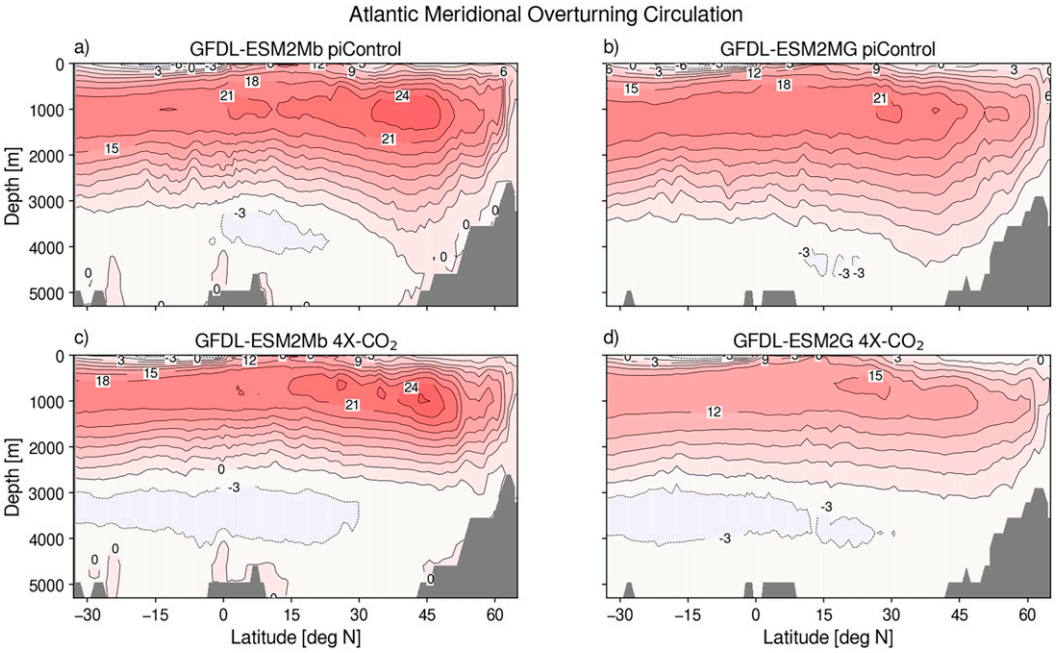

FIG. 7. Century time-averaged depth vs latitude sections of the AMOC at model years 4901–5000 in the (a),(b) control and (c),(d) 4 × CO_2 simulations for (left) GFDL-ESM2Mb and (right) GFDL-ESM2G. Contour and shading interval is 3 Sv.

into the North Atlantic. After 1000 years, this water is able to erode the fresh cap that is hindering convection in ESM2Mb, particularly over the Labrador Sea region (Fig. 8). The high-salinity water eventually enhances NADW formation and strengthens the AMOC relative to the piControl simulation in ESM2Mb.

After 5000 years of the 4 × CO_2 integration, the depth of the 0-Sv streamfunction contour rises in the water column by ~500 m in both models (Figs. 7c,d). The largest weakening in the overturning streamfunction occurs at a depth of approximately 1000 m in both models. The shoaling of NADW water allows a stronger Antarctic bottom circulation in both models. The increase is more pronounced in ESM2Mb but is also present in ESM2G. A surprising aspect of the different near-equilibrium AMOC responses is their relatively small impact on the evolution, pattern, and magnitude of change in the surface climate, as shown in the previous sections.

The meridional overturning circulation plays an important role in the equator-to-pole heat transport. The differences in poleward heat transport between ESM2Mb and ESM2G are consistent with their changes in AMOC, although their individual heat transport efficiencies also need to be considered. The total basin-integrated Atlantic poleward heat transport at the end of the 4 × CO_2 simulations shows a reduction relative to the control run in both models (Figs. 9a,b; Table 3). Both models have a maximum poleward Atlantic heat transport of 1.11 PW in

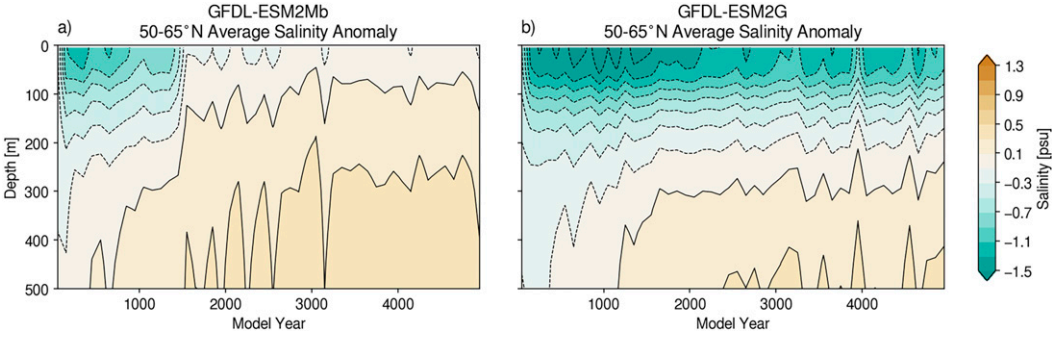

FIG. 8. North Atlantic (50°–65°N) upper-ocean salinity response (piControl minus 4 × CO_2; psu) for (a) GFDL-ESM2Mb and (b) GFDL-ESM2G.

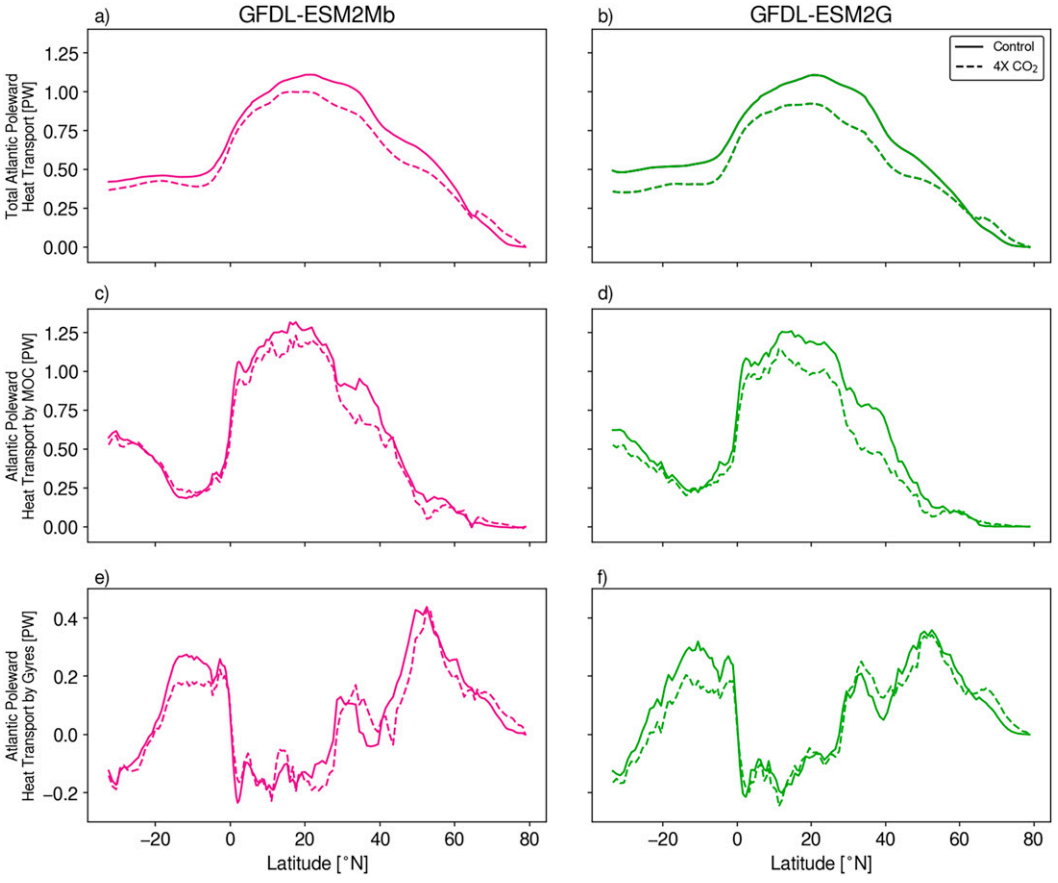

FIG. 9. Atlantic basin-integrated poleward ocean heat transport (PW; 1 PW = 10^{15} W) vs latitude averaged over years 4901–5000 for the piControl and 4 × CO_2 simulations. (a),(b) The total heat transport is decomposed into contributions from the (c),(d) meridional overturning circulation and (e),(f) gyre circulation for (left) GFDL-ESM2Mb and (right) GFDL-ESM2G. Solid lines correspond to the piControl simulations and the dashed lines correspond to the 4 × CO_2 simulations.

their control simulation, which is reduced to 1.0 and 0.92 in the 4 × CO_2 simulations for ESM2Mb and ESM2G, respectively. These values correspond to a 10% reduction in poleward heat transport in ESM2Mb and a 17% reduction in ESM2G.

The contributions of the total poleward heat transport can be decomposed into parts attributable to the meridional overturning circulation and the gyre circulation (Vallis and Farneti 2009):

$$\rho_0 C_p L \int_{-H}^{\eta} [v\theta]\, dz = \rho_0 C_p L \int_{-H}^{\eta} [v][\theta]\, dz + \rho_0 C_p L \int_{-H}^{\eta} [v^*\theta^*]\, dz. \quad (1)$$

In Eq. (1), ρ_0 corresponds to the model's reference density, C_p is the heat capacity of seawater, L is the meridionally varying basin width, H is the ocean depth, and η is the surface height. Zonal-mean quantities are denoted by square brackets while deviations from this zonal mean are denoted by asterisks. The first term on the rhs of Eq. (1) is the basin poleward heat transport contribution associated with the meridional overturning circulation and the second term is the basin gyre circulation contribution to the total heat transport. The decomposition analysis was performed for years 4901–5000 from both the piControl simulation and the 4 × CO_2 simulation.

Most of the change in the total poleward heat transport is explained by changes in the heat transport associated with the overturning circulation (Figs. 9c,d; Table 3). In ESM2Mb, the maximum heat transport associated with the overturning circulation declines by 0.09 PW (7%) in ESM2Mb and by 0.11 PW (9%) in ESM2G. The changes in heat transport associated with gyres (0.01–0.02 PW; Figs. 9e,f) are small compared to the changes associated with the overturning circulation, except in the Southern Hemisphere tropics between

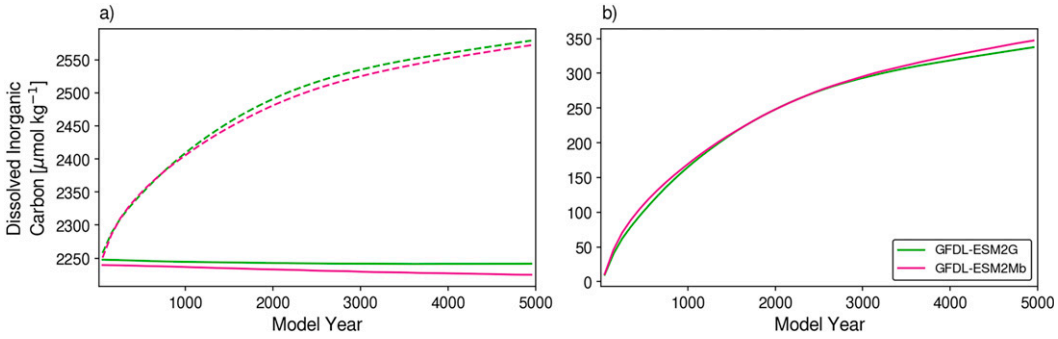

FIG. 10. (a) Time series of dissolved inorganic carbon inventory (μmol kg^{-1}) for the piControl simulations (solid) and the 4 × CO_2 simulations (dashed) for GFDL-ESM2Mb (magenta) and GFDL-ESM2G (green). (b) Time series of the DIC response (4 × CO_2 simulations minus the piControl simulations).

20°S and the equator, where the gyre changes are dominant. There is also a poleward shift in the gyre transport that is evident in ESM2Mb and less pronounced in ESM2G. The similarity of the Atlantic poleward heat transport responses are even more remarkable given that the mass transport in ESM2Mb recovers and strengthens relative to the piControl simulation during the 4 × CO_2 simulation.

Over the first 1000 years, the efficiency of the AMOC circulation's ability to transport heat poleward is increased in both models (Fig. 6b). The efficiency is defined here as the ratio of the maximum overturning streamfunction divided by the maximum total poleward heat transport, expressed as a percentage change. Both models increase their efficiency by almost 20% initially as the AMOC declines but the ocean maintains its poleward heat transport. Over time, the efficiency relaxes and decreases relative to the control run as the poleward heat transport decreases. The efficiency is closely related to the temperature gradient between the surface waters flowing northward and the deep waters flowing southward in the Atlantic. In ESM2Mb, the warming that occurs in the deep and abyssal waters reduces the vertical temperature gradient in the ocean, thus leading to a reduction in the AMOC efficiency. In ESM2G, there is less warming at depth in the Atlantic and the vertical temperature gradient is stronger. Despite the near −50% decrease in the strength of AMOC in ESM2G, there is only a ~5% decrease in the efficiency in transporting heat poleward.

d. Dissolved inorganic carbon response

As atmospheric CO_2 concentrations increase in the atmosphere, the disequilibrium of heat and carbon across the air–sea interface leads to a net flux of both quantities into the ocean. Unlike heat fluxes across the upper boundary of the ocean, which directly induces an "active" dynamical response in the ocean, carbon behaves more like a "passive" tracer in these experiments (Winton et al. 2013).

As atmospheric CO_2 concentrations rise in the model, the ocean sequesters approximately one-quarter of the additional carbon in the Earth system (Ciais et al. 2013; Jones et al. 2013). As both models slowly come into equilibrium with the 4 × CO_2 concentration in the atmosphere, the rate of dissolved inorganic carbon (DIC) uptake by the ocean slows over the 5000 years of model simulation but never fully ceases. When the 4 × CO_2 simulations are subtracted from their respective piControl simulations (Fig. 10b), both model responses are very similar in terms of both their timing and magnitude of DIC uptake (Fig. 11). After 5000 years, there is a maximum increase of DIC exceeding 400 μmol kg^{-1} between 500 m and the surface in both models. DIC changes are fairly uniform between 1000- and 4000-m depths, ranging between 300 and 350 μmol kg^{-1}.

Below 3000-m depth, ESM2Mb shows more accumulation of DIC compared to ESM2G. Part of the difference between the models' representation of DIC accumulation is related to their ventilation responses to the 4 × CO_2 forcing. As noted earlier, AMOC mass transport declines markedly in the ESM2G model by year 5000, while it increases relative to the piControl value in ESM2Mb. The high-latitude regions that drive deep-water formation are also regions of large ocean carbon uptake. As processes such as AMOC weaken in response to the warmer climate, their ability to mechanically transport carbon from the surface to the deep ocean is more limited.

e. Response sensitivity to diapycnal diffusivity

The results of the preceding sections show similarities between ESM2Mb and ESM2G in their responses to 4 × CO_2 forcing. These similarities suggest that the contributions to the climate response uncertainty arising from the formulation, vertical coordinate, and numerics

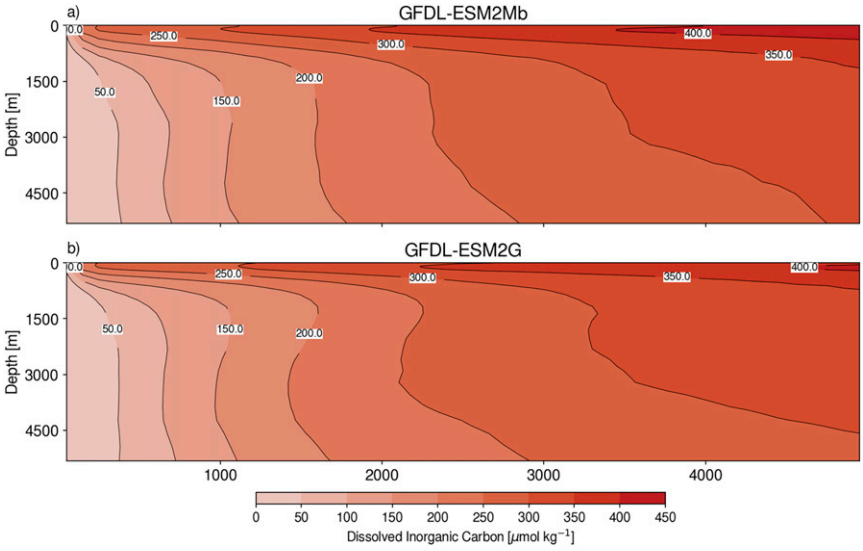

FIG. 11. Depth vs time contour plot of the DIC response relative to the piControl simulation for the 5000-yr $4 \times CO_2$ simulations for (a) GFDL-ESM2Mb and (b) GFDL-ESM2G. Contour and shading interval is 50 μmol kg^{-1}.

are small in these two ocean model configurations. To illustrate the relative impact of mixing compared to model formulation, a series of 5000-yr piControl and $4 \times CO_2$ experiments were run with the GFDL-CM2G model, where the spatially uniform background diapycnal diffusivity K_d was increased in the ocean. The diffusivity K_d was chosen as it is a parameter that targets the main thermocline region, thus having an impact on the time scale and magnitude of heat exchange with the deep ocean. We ran the non-ESM version, GFDL-CM2G, to reduce computational resource needs for these sensitivity experiments. CM2G is similar in every respect to ESM2G except that it does not include the interactive ocean biogeochemistry component of the model. The background value for K_d used in ESM2G is $2 \times 10^{-5}\,m^2\,s^{-1}$ and the sensitivity tests with CM2G used values of 4×10^{-5}, 6×10^{-5}, and $1 \times 10^{-4}\,m^2\,s^{-1}$ for K_d for the piControl and $4 \times CO_2$ simulations. Furthermore, we did not run sensitivity tests with ESM2Mb due to limited computational resources.

Increasing K_d cools SSTs in the tropics and midlatitudes and warms them in the high latitudes (Fig. 12, left). Increasing K_d also leads to weaker stratification near the surface and enhanced stratification at depth in the water column, leading to greater subsurface warming and increased ocean heat uptake (Marshall and Zanna 2014). After 5000 years of $4 \times CO_2$ forcing, the tropical Pacific thermocline in all of the experiments becomes sharper and shoals slightly (Fig. 13). However, the initial thermocline structure in the piControl simulations varies considerably with K_d. Increasing K_d in the GFDL-CM2G model leads to a weaker, more diffuse thermocline compared to the ESM2G and ESM2Mb simulations.

Increasing K_d warms the global-mean SST, which varies by more than 2°C (Fig. 14a) among this suite of sensitivity tests. After 1000 yr of $4 \times CO_2$ forcing, the ensemble spread of global-mean SSTs among all of the model configurations is 1°C. After 3000 years, however, all of the configurations, including ESM2Mb and ESM2G, converge on the same global-mean SST response after 3000 years despite different initial states. The similarity in the global-mean SST response with varying K_d masks important regional differences (Fig. 12). With the lower values of K_d, the response in the Arctic basin remains closer to zero, or even negative, in the convective sinking regions near the Labrador Sea. With higher values of K_d, spatial heterogeneity in the warming response is lost and there is a more uniform warming that occurs, even in the Arctic basin.

Ventilation of the deep ocean, deep-water formation processes, and thus the volume-mean ocean temperature, is sensitive to changes that are occurring in the high latitudes of the simulations with varying K_d. In the piControl simulations with differing values of K_d, this volume-mean temperature varies by almost a factor of 2 (Fig. 14c). The responses of the global volume-mean temperature to the $4 \times CO_2$ forcing also differ with respect to changing the value of K_d. A quadrupling of atmospheric CO_2 produces a more global volume-mean temperature rise with higher values of the background diffusivity (Fig. 14d). Therefore, these results demonstrate that significant

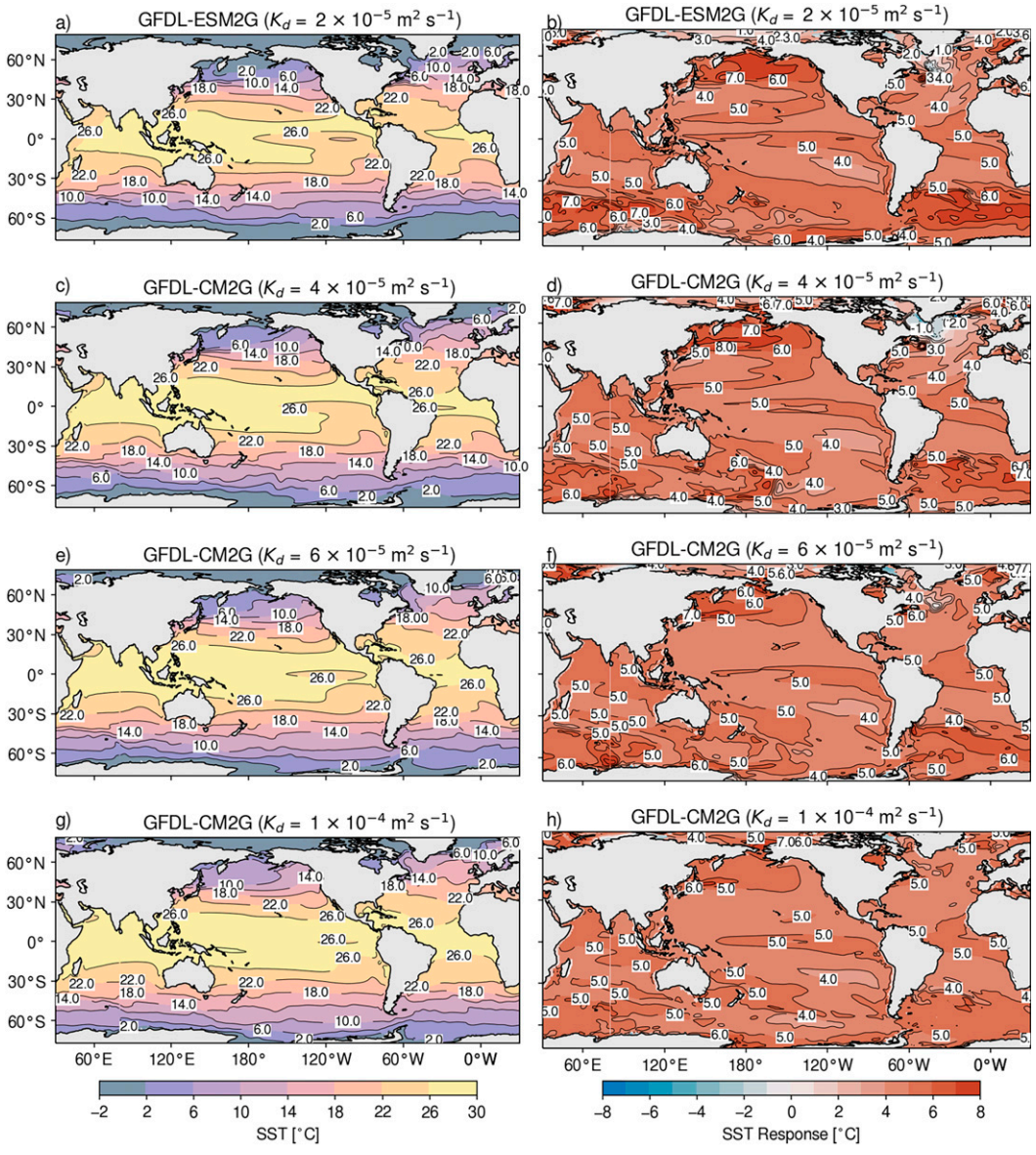

FIG. 12. Spatial patterns of (left) SST (°C) averaged over the piControl simulation and (right) for warming (°C) at near equilibrium (years 4901–5000) relative to the piControl simulations under varying amounts of background diapycnal diffusivity K_d for (a),(b) GFDL-ESM2G ($K_d = 2 \times 10^{-5}\,\mathrm{m^2\,s^{-1}}$), (c),(d) GFDL-CM2G ($K_d = 4 \times 10^{-5}\,\mathrm{m^2\,s^{-1}}$), (e),(f) GFDL-CM2G ($K_d = 6 \times 10^{-5}\,\mathrm{m^2\,s^{-1}}$), and (g),(h) GFDL-CM2G ($K_d = 1 \times 10^{-4}\,\mathrm{m^2\,s^{-1}}$).

differences in the response to the 4 × CO_2 forcing in the ocean interior can occur by localized surface differences that have a relatively small impact on the global-mean SST response.

This suite of K_d simulations with the GFDL-CM2G model supports the results of Danabasoglu and Gent (2009) in that the global-mean SST response is insensitive to the ocean model response. This similarity is remarkable, since the suite of simulations produces very different spatial patterns of SST and patterns of warming at depth. This response is governed by energetics, as the radiative forcing from increased CO_2 in the atmosphere is the same among these simulations. The results of the K_d sensitivity study extend the findings of Danabasoglu and Gent (2009) by highlighting some important limits of this insensitivity of global-mean SST to ocean model formulation. First, the spatial patterns of SST and warming at depth are crucial for diagnosing and understanding regional responses to climate warming, specifically sea level rise, and will vary depending on the ocean model formulation. Second, although the SST *responses* are similar among all of the models in

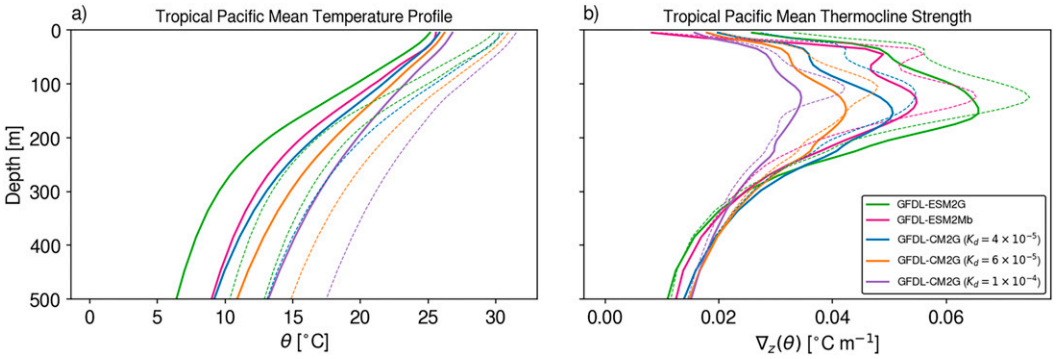

FIG. 13. (a) Upper 500-m profiles of potential temperature (°C) in the tropical Pacific averaged between 30°S and 30°N. (b) Vertical gradient of potential temperature (°C m^{-1}). The solid lines are averaged over 5000 years of the piControl simulation and the dashed lines are averaged over years 4901–5000 at the end of the 4 × CO_2 simulations.

this study, this is not the same as the ECS, which is defined as the atmospheric near-surface (2 m) temperature response to forcing. The ECS varies by 0.6°C among the simulations considered in this study (Table 4), and the differences between the ECS and the SST response reflect the ocean's influence on key feedback processes, such as clouds and sea ice.

4. Summary

Using two Earth system models that differ primarily in their ocean component, and principally their implementation of the vertical coordinate, their near-equilibrium ocean response to a quadrupling of atmospheric CO_2 concentrations is evaluated via a "brute force" approach of running the control and perturbation experiments for 5000 years. The GFDL-ESM2Mb model used in this study employs a depth-based vertical coordinate while the GFDL-ESM2G model uses an isopycnal vertical coordinate. Both ocean models were configured with comparable settings for their physical parameterizations and the remaining components of the coupled model remain virtually identical.

The near-equilibrium response to quadrupling atmospheric CO_2 is remarkably similar on millennial time scales between these models. For most of the climate

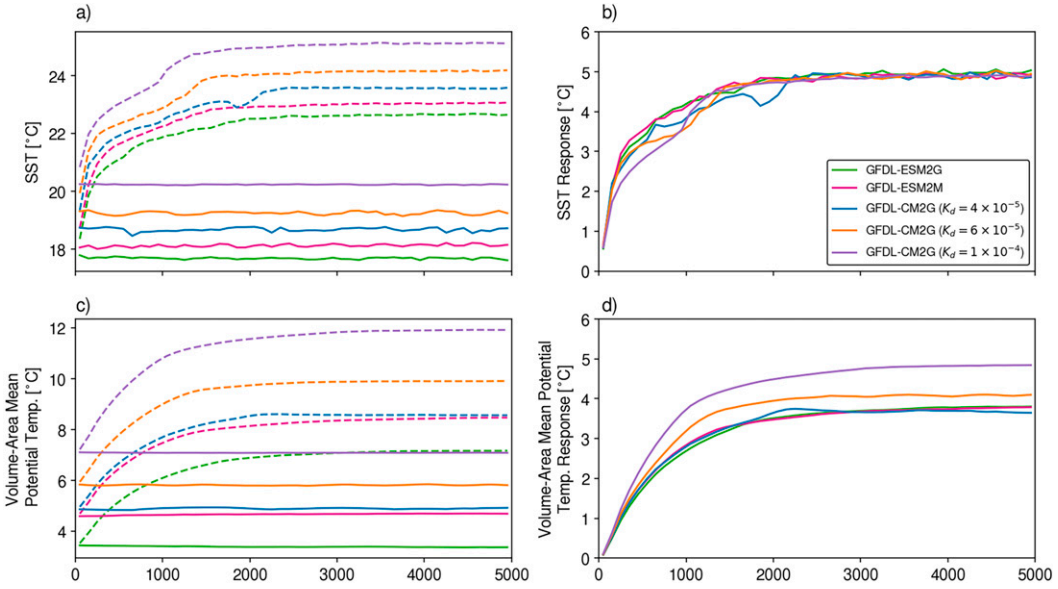

FIG. 14. Time series of (a) SST (°C) and (c) volume-mean ocean potential temperature (°C) for GFDL-ESM2Mb (magenta), GFDL-ESM2G (green), and three GFDL-CM2G simulations with varying background values of diapycnal diffusivity (blue, orange, and purple). Solid lines represent the piControl simulations and dashed lines represent the 4 × CO_2 simulations. (b),(d) The responses are defined as the 4 × CO_2 simulations minus their respective piControl simulations.

TABLE 4. Global-mean SST (°C), volume-mean ocean potential temperature θ (°C), and ECS based on atmospheric 2-m temperature for years 4901–5000 from the piControl simulations and their response ($4 \times CO_2$ simulations minus piControl).

Model	SST (piControl)	SST response	θ (piControl)	θ response	ECS
GFDL-ESM2Mb	18.12	4.92	4.66	3.79	6.30
GFDL-ESM2G	17.69	5.04	3.39	3.80	6.54
GFDL-CM2G ($K_d = 4 \times 10^{-5}\,m^2\,s^{-1}$)	18.68	4.86	4.89	3.65	6.18
GFDL-CM2G ($K_d = 6 \times 10^{-5}\,m^2\,s^{-1}$)	19.25	4.95	5.82	4.10	6.14
GFDL-CM2G ($K_d = 1 \times 10^{-4}\,m^2\,s^{-1}$)	20.24	4.89	7.09	4.84	5.93

features considered in this analysis, the patterns and magnitudes of the responses are similar despite both models having different climate-mean states in their preindustrial control simulations (piControl). Under $4 \times CO_2$ forcing, both models produce a global mean SST response near 5°C with similar patterns ($r^2 = 0.75$). The volume-mean ocean temperature increases by 3.8°C in both ESMs. Despite the ESM2G model having more Arctic sea ice in the piControl simulation, both models have similar responses in March (~9 million km^2) and become ice free after several centuries during September. The AMOC strength declines in ESM2G (−9.2 Sv) and to a lesser degree in ESM2Mb (−7.2 Sv), consistent with the freshening of the North Atlantic (Manabe and Stouffer 1993; Sévellec and Fedorov 2016). While AMOC remains weakened in ESM2G, AMOC recovers and strengthens as ESM2Mb is able to erode the North Atlantic fresh capping. The overturning circulation contribution to the poleward heat transport is reduced in both models at near-equilibrium, despite AMOC recovering in ESM2Mb. Both models show a decline in the AMOC's efficiency in transporting heat poleward, but the magnitude of this reduction is smaller in ESM2G (~5%). The differing AMOC responses are also evident in the depth structure of DIC over time.

These results suggest that the climate response uncertainty on millennial time scales related to the ocean model vertical coordinate is smaller than the uncertainty arising from the model physics. Both models were carefully constructed using observations and theory as guides to configure their physical parameterizations appropriately for their different formulations and vertical coordinates. This study demonstrates that it is possible to construct two ocean models that produce very similar climate responses to forcing despite them having very different formulations and numerics. Our results support the conjecture by Griffies and Treguier (2013) that ocean simulations are more highly affected by physical parameterizations than by vertical coordinate choices. Furthermore, we demonstrate that changing the amount of background diapycnal mixing in one of the models, ESM2G, produces different patterns and magnitudes of warming in response to the same $4 \times CO_2$ forcing.

In this study, we only examine results from non-eddying ocean climate simulations. It remains to be demonstrated if similar responses could be obtained from higher-resolution eddy-permitting and eddy-active models where the dependence on mesoscale eddy parameterizations is greatly reduced or eliminated. Ocean models also employ "hybrid" vertical coordinates (Bleck 2002) that are depth based near the surface and transition to density based in the ocean interior. More work is needed to evaluate the long-term (i.e., millennial scale) response to forcing in this latest generation of models.

The computational costs of running long multimillennial simulations with fully coupled ESMs are high and will increase with higher resolution and more complexity (e.g., more biogeochemical tracers) in future-generation models. Nevertheless, such experiments remain an important tool in understanding the long-term behavior of models and provide additional verification to extrapolate the effective climate sensitivity from shorter transient simulations.

In summary, it is possible to construct coupled climate models that produce the same equilibrium climate response, despite their ocean model components having very different formulations (i.e., vertical coordinate and numerics). The response uncertainty related to model physics is larger than the uncertainty arising from model formulation or dynamics alone. These results highlight the importance of constructing ocean model components where lateral and vertical mixing parameterizations are well constrained by observations and theoretical understanding.

Acknowledgments. The authors are grateful to John Dunne and Sonya Legg for their comments on this paper. We also thank three anonymous reviewers whose comments greatly helped to clarify many aspects of this paper. We thank Catherine Raphael for comments and guidance regarding the figures in this paper. Several of the color maps used in the figures were obtained from the cmocean package (Thyng et al. 2016) and line color selections based in part on the ColorBrewer palettes (Brewer et al. 2003).

REFERENCES

Adcroft, A., and J. M. Campin, 2004: Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models. *Ocean Modell.*, **7**, 269–284, https://doi.org/10.1016/j.ocemod.2003.09.003.

——, J. R. Scott, and J. Marotzke, 2001: Impact of geothermal heating on the global ocean circulation. *Geophys. Res. Lett.*, **28**, 1735–1738, https://doi.org/10.1029/2000GL012182.

Anderson, J. L., and Coauthors, 2004: The new GFDL global atmosphere and land model AM2–LM2: Evaluation with prescribed SST simulations. *J. Climate*, **17**, 4641–4673, https://doi.org/10.1175/JCLI-3223.1.

Andrews, T., J. M. Gregory, and M. J. Webb, 2015: The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *J. Climate*, **28**, 1630–1648, https://doi.org/10.1175/JCLI-D-14-00545.1.

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

Bindoff, N. L., and Coauthors, 2013: Detection and attribution of climate change: From global to regional. *Climate Change 2013: The Physical Science Basis*, T. Stocker et al., Eds., Cambridge University Press, 867–952.

Bleck, R., 2002: An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Modell.*, **4**, 55–88, https://doi.org/10.1016/S1463-5003(01)00012-9.

Bony, S., and Coauthors, 2006: How well do we understand and evaluate climate change feedback processes? *J. Climate*, **19**, 3445–3482, https://doi.org/10.1175/JCLI3819.1.

Brewer, C. A., G. W. Hatchard, and M. A. Harrower, 2003: Colorbrewer in print: A catalog of color schemes for maps. *Cartogr. Geogr. Inf. Sci.*, **30**, 5–32, https://doi.org/10.1559/152304003100010929.

Ciais, P., and Coauthors, 2013: Carbon and other biogeochemical cycles. *Climate Change 2013: The Physical Science Basis*, T. Stocker et al., Eds., Cambridge University Press, 465–570.

Dalan, F., P. H. Stone, and A. P. Sokolov, 2005: Sensitivity of the ocean's climate to diapycnal diffusivity in an EMIC. Part II: Global warming scenario. *J. Climate*, **18**, 2482–2496, https://doi.org/10.1175/JCLI3412.1.

Danabasoglu, G., and J. Marshall, 2007: Effects of vertical variations of thickness diffusivity in an ocean general circulation model. *Ocean Modell.*, **18**, 122–141, https://doi.org/10.1016/j.ocemod.2007.03.006.

——, and P. R. Gent, 2009: Equilibrium climate sensitivity: Is it accurate to use a slab ocean model? *J. Climate*, **22**, 2494–2499, https://doi.org/10.1175/2008JCLI2596.1.

——, W. G. Large, J. J. Tribbia, P. R. Gent, B. P. Briegleb, and J. C. McWilliams, 2006: Diurnal coupling in the tropical oceans of CCSM3. *J. Climate*, **19**, 2347–2365, https://doi.org/10.1175/JCLI3739.1.

Delworth, T. L., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Climate*, **19**, 643–674, https://doi.org/10.1175/JCLI3629.1.

Duffy, P. B., M. Eby, and A. J. Weaver, 2001: Climate model simulations of effects of increased atmospheric CO_2 and loss of sea ice on ocean salinity and tracer uptake. *J. Climate*, **14**, 520–532, https://doi.org/10.1175/1520-0442(2001)014<0520:CMSOEO>2.0.CO;2.

Dunne, J. P., and Coauthors, 2012: GFDL's ESM2 global coupled climate–carbon Earth system models. Part I: Physical formulation and baseline simulation characteristics. *J. Climate*, **25**, 6646–6665, https://doi.org/10.1175/JCLI-D-11-00560.1.

——, and Coauthors, 2013: GFDL's ESM2 global coupled climate–carbon Earth system models. Part II: Carbon system formulation and baseline simulation characteristics. *J. Climate*, **26**, 2247–2267, https://doi.org/10.1175/JCLI-D-12-00150.1.

Durack, P. J., S. E. Wijffels, and R. J. Matear, 2012: Ocean salinities reveal strong global water cycle intensification during 1950 to 2000. *Science*, **336**, 455–458, https://doi.org/10.1126/science.1212222.

Ehlert, D., K. Zickfeld, M. Eby, and N. Gillett, 2017: The sensitivity of the proportionality between temperature change and cumulative CO_2 emissions to ocean mixing. *J. Climate*, **30**, 2921–2935, https://doi.org/10.1175/JCLI-D-16-0247.1.

Fetterer, F., K. Knowles, W. Meier, M. Savoie, and A. K. Windnagel, 2017: Sea Ice Index, version 3: Years 1979–2018. National Snow and Ice Data Center, Boulder, CO, accessed May 2018, https://doi.org/10.7265/N5K072F8.

Flato, G. M., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. Stocker et al., Eds., Cambridge University Press, 741–866.

Fox-Kemper, B., and Coauthors, 2011: Parameterization of mixed layer eddies. III: Implementation and impact in global ocean climate simulations. *Ocean Modell.*, **39**, 61–78, https://doi.org/10.1016/j.ocemod.2010.09.002.

Garuba, O. A., J. Lu, F. Liu, and H. A. Singh, 2018: The active role of the ocean in the temporal evolution of climate sensitivity. *Geophys. Res. Lett.*, **45**, 306–315, https://doi.org/10.1002/2017GL075633.

Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, **20**, 150–155, https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2.

Gregg, M. C., 1987: Diapycnal mixing in the thermocline: A review. *J. Geophys. Res.*, **92**, 5249–5286, https://doi.org/10.1029/JC092iC05p05249.

Gregory, J. M., 2000: Vertical heat transports in the ocean and their effect on time-dependent climate change. *Climate Dyn.*, **16**, 501–515, https://doi.org/10.1007/s003820000059.

——, H. T. Banks, P. A. Stott, J. A. Lowe, and M. D. Palmer, 2004: Simulated and observed decadal variability in ocean heat content. *Geophys. Res. Lett.*, **31**, L15312, https://doi.org/10.1029/2004GL020258.

——, and Coauthors, 2016: The Flux-Anomaly Forced Model Intercomparison Project (FAFMIP) contribution to CMIP6: Investigation of sea-level and ocean climate change in response to CO_2 forcing. *Geosci. Model Dev.*, **9**, 3993–4017, https://doi.org/10.5194/gmd-9-3993-2016.

Griffies, S. M., and A. M. Treguier, 2013: Ocean circulation models and modeling. *Ocean Circulation and Climate*, G. Siedler et al., Eds., International Geophysics, Vol. 103, Academic Press, 521–551, https://doi.org/10.1016/B978-0-12-391851-2.00020-9.

——, R. C. Pacanowski, and R. W. Hallberg, 2000a: Spurious diapycnal mixing associated with advection in a z-coordinate ocean model. *Mon. Wea. Rev.*, **128**, 538–564, https://doi.org/10.1175/1520-0493(2000)128<0538:SDMAWA>2.0.CO;2.

——, and Coauthors, 2000b: Developments in ocean climate modelling. *Ocean Modell.*, **2**, 123–192, https://doi.org/10.1016/S1463-5003(00)00014-7.

——, M. Schmidt, and M. Herzfeld, 2009: Elements of mom4p1. GFDL Ocean Group Tech. Rep. 6, 371 pp., https://data1.gfdl.noaa.gov/~arl/pubrel/o/old/doc/mom4p1_guide.pdf.

——, and Coauthors, 2015: Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate

models. *J. Cl0069mate*, **28**, 952–977, https://doi.org/10.1175/JCLI-D-14-00353.1.

Hallberg, R., 1995: Some aspects of the circulation in ocean basins with isopycnals intersecting sloping boundaries. Ph.D. thesis, University of Washington, 244 pp. [Available from University Microfilms, 1490 Eisenhower Place, P.O. Box 975, Ann Arbor, MI 48106.]

——, 2003: The ability of large-scale ocean models to accept parameterizations of boundary mixing, and a description of a refined bulk mixed-layer model. *Proc. 'Aha Huliko'a Hawaiian Winter Workshop*, Honolulu, HI, University of Hawai'i at Mānoa, 187–203.

——, and A. Adcroft, 2009: Reconciling estimates of the free surface height in Lagrangian vertical coordinate ocean models with mode-split time stepping. *Ocean Modell.*, **29**, 15–26, https://doi.org/10.1016/j.ocemod.2009.02.008.

Hansen, J., G. Russell, A. Lacis, I. Fung, D. Rind, and P. Stone, 1985: Climate response times: Dependence on climate sensitivity and ocean mixing. *Science*, **229**, 857–859, https://doi.org/10.1126/science.229.4716.857.

——, and Coauthors, 2005: Earth's energy imbalance: Confirmation and implications. *Science*, **308**, 1431–1435, https://doi.org/10.1126/science.1110252.

Held, I. M., and B. J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *J. Climate*, **19**, 5686–5699, https://doi.org/10.1175/JCLI3990.1.

Ilicak, M., A. J. Adcroft, S. M. Griffies, and R. W. Hallberg, 2012: Spurious dianeutral mixing and the role of momentum closure. *Ocean Modell.*, **45–46**, 37–58, https://doi.org/10.1016/j.ocemod.2011.10.003.

IPCC, 2013: Summary for policymakers. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 3–29.

Jackson, L., R. Hallberg, and S. Legg, 2008: A parameterization of shear-driven turbulence for ocean climate models. *J. Phys. Oceanogr.*, **38**, 1033–1053, https://doi.org/10.1175/2007JPO3779.1.

Jones, C., and Coauthors, 2013: Twenty-first-century compatible CO_2 emissions and airborne fraction simulated by CMIP5 earth system models under four representative concentration pathways. *J. Climate*, **26**, 4398–4413, https://doi.org/10.1175/JCLI-D-12-00554.1.

Johns, W. E., and Coauthors, 2011: Continuous, array-based estimates of Atlantic Ocean heat transport at 26.5°N. *J. Climate*, **24**, 2429–2449, https://doi.org/10.1175/2010JCLI3997.1.

Krasting, J. P., J. P. Dunne, R. J. Stouffer, and R. W. Hallberg, 2016: Enhanced Atlantic sea-level rise relative to the Pacific under high carbon emission rates. *Nat. Geosci.*, **9**, 210–214, https://doi.org/10.1038/ngeo2641.

Large, W. G., J. C. McWilliams, and S. C. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.*, **32**, 363–403, https://doi.org/10.1029/94RG01872.

Ledwell, J. R., A. J. Watson, and C. S. Law, 1993: Evidence for slow mixing across the pycnocline from an open-ocean tracer-release experiment. *Nature*, **364**, 701–703, https://doi.org/10.1038/364701a0.

——, ——, and ——, 1998: Mixing of a tracer in the pycnocline. *J. Geophys. Res.*, **103**, 21 499–21 529, https://doi.org/10.1029/98JC01738.

——, L. C. St. Laurent, J. B. Girton, and J. M. Toole, 2011: Diapycnal mixing in the Antarctic Circumpolar Current. *J. Phys. Oceanogr.*, **41**, 241–246, https://doi.org/10.1175/2010JPO4557.1.

Legg, S., R. W. Hallberg, and J. B. Girton, 2006: Comparison of entrainment in overflows simulated by z-coordinate, isopycnal and non-hydrostatic models. *Ocean Modell.*, **11**, 69–97, https://doi.org/10.1016/j.ocemod.2004.11.006.

——, and Coauthors, 2009: Improving oceanic overflow representation in climate models: The Gravity Current Entrainment Climate Process Team. *Bull. Amer. Meteor. Soc.*, **90**, 657–670, https://doi.org/10.1175/2008BAMS2667.1.

Li, C., J.-S. Von Storch, and J. Marotzke, 2013: Deep-ocean heat uptake and equilibrium climate response. *Climate Dyn.*, **40**, 1071–1086, https://doi.org/10.1007/s00382-012-1350-z.

Manabe, S., and R. J. Stouffer, 1980: Sensitivity of a global climate model to an increase of CO_2 concentration in the atmosphere. *J. Geophys. Res.*, **85**, 5529–5554, https://doi.org/10.1029/JC085iC10p05529.

——, and ——, 1993: Century-scale effects of increased atmospheric CO_2 on the ocean–atmosphere system. *Nature*, **364**, 215–218, https://doi.org/10.1038/364215a0.

——, and ——, 1994: Multiple-century response of a coupled ocean–atmosphere model to an increase of atmospheric carbon dioxide. *J. Climate*, **7**, 5–23, https://doi.org/10.1175/1520-0442(1994)007<0005:MCROAC>2.0.CO;2.

——, and ——, 1999: Are two modes of thermohaline circulation stable? *Tellus*, **51A**, 400–411, https://doi.org/10.3402/tellusa.v51i3.13461.

——, ——, M. J. Spelman, and K. Bryan, 1991: Transient responses of a coupled ocean–atmosphere model to gradual changes of atmospheric CO_2. Part I: Annual mean response. *J. Climate*, **4**, 785–818, https://doi.org/10.1175/1520-0442(1991)004<0785:TROACO>2.0.CO;2.

Marshall, D. P., and L. Zanna, 2014: A conceptual model of ocean heat uptake under climate change. *J. Climate*, **27**, 8444–8465, https://doi.org/10.1175/JCLI-D-13-00344.1.

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

Meehl, G. A., J. M. Arblaster, J. T. Fasullo, A. Hu, and K. E. Trenberth, 2011: Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nat. Climate Change*, **1**, 360–364, https://doi.org/10.1038/nclimate1229.

Megann, A., 2018: Estimating the numerical diapycnal mixing in an eddy-permitting ocean model. *Ocean Modell.*, **121**, 19–33, https://doi.org/10.1016/j.ocemod.2017.11.001.

——, A. L. New, A. T. Blaker, and B. Sinha, 2010: The sensitivity of a coupled climate model to its ocean component. *J. Climate*, **23**, 5126–5150, https://doi.org/10.1175/2010JCLI3394.1.

Melet, A., R. Hallberg, S. Legg, and K. Polzin, 2013: Sensitivity of the ocean state to the vertical distribution of internal-tide-driven mixing. *J. Phys. Oceanogr.*, **43**, 602–615, https://doi.org/10.1175/JPO-D-12-055.1.

——, S. Legg, and R. Hallberg, 2016: Climatic impacts of parameterized local and remote tidal mixing. *J. Climate*, **29**, 3473–3500, https://doi.org/10.1175/JCLI-D-15-0153.1.

Milly, P. C. D., and Coauthors, 2014: An enhanced model of land water and energy for global hydrologic and Earth-system studies. *J. Hydrometeor.*, **15**, 1739–1761, https://doi.org/10.1175/JHM-D-13-0162.1.

Morrison, A. K., S. M. Griffies, M. Winton, W. G. Anderson, and J. L. Sarmiento, 2016: Mechanisms of Southern Ocean heat uptake and transport in a global eddying climate model. *J. Climate*, **29**, 2059–2075, https://doi.org/10.1175/JCLI-D-15-0579.1.

Munk, W. H., 1966: Abyssal recipes. *Deep-Sea Res.*, **13**, 707–730, https://doi.org/10.1016/0011-7471(66)90602-4.

Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.*, **12**, 1154–1158, https://doi.org/10.1175/1520-0485(1982)012<1154:OIMBCR>2.0.CO;2.

Schmittner, A., and A. J. Weaver, 2001: Dependence of multiple climate states on ocean mixing parameters. *Geophys. Res. Lett.*, **28**, 1027–1030, https://doi.org/10.1029/2000GL012410.

Sévellec, F., and A. V. Fedorov, 2016: AMOC sensitivity to surface buoyancy fluxes: Stronger ocean meridional heat transport with a weaker volume transport? *Climate Dyn.*, **47**, 1497–1513, https://doi.org/10.1007/s00382-015-2915-4.

Shevliakova, E., and Coauthors, 2009: Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink. *Global Biogeochem. Cycles*, **23**, GB2022, https://doi.org/10.1029/2007GB003176.

Simmons, H. L., S. R. Jayne, L. C. Laurent, and A. J. Weaver, 2004: Tidally driven mixing in a numerical model of the ocean general circulation. *Ocean Modell.*, **6**, 245–263, https://doi.org/10.1016/S1463-5003(03)00011-8.

Stacey, M. W., S. Pond, and Z. P. Nowak, 1995: A numerical model of the circulation in Knight Inlet, British Columbia, Canada. *J. Phys. Oceanogr.*, **25**, 1037–1062, https://doi.org/10.1175/1520-0485(1995)025<1037:ANMOTC>2.0.CO;2.

Stouffer, R. J., and S. Manabe, 1999: Response of a coupled ocean–atmosphere model to increasing atmospheric carbon dioxide: Sensitivity to the rate of increase. *J. Climate*, **12**, 2224–2237, https://doi.org/10.1175/1520-0442(1999)012<2224:ROACOA>2.0.CO;2.

Stroeve, J. C., V. Kattsov, A. Barrett, M. Serreze, T. Pavlova, M. Holland, and W. N. Meier, 2012: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophys. Res. Lett.*, **39**, L16502, https://doi.org/10.1029/2012GL052676.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1.

Tesdal, J.-E., R. P. Abernathey, J. I. Goes, A. L. Gordon, and T. W. N. Haine, 2018: Salinity trends within the upper layers of the subpolar North Atlantic. *J. Climate*, **31**, 2675–2698, https://doi.org/10.1175/JCLI-D-17-0532.1.

Thyng, K. M., C. A. Greene, R. D. Hetland, H. M. Zimmerle, and S. F. DiMarco, 2016: True colors of oceanography: Guidelines for effective and accurate colormap selection. *Oceanography*, **29**, 9–13, https://doi.org/10.5670/oceanog.2016.66.

Vallis, G. K., and R. Farneti, 2009: Meridional energy transport in the coupled atmosphere–ocean system: Scaling and numerical experiments. *Quart. J. Roy. Meteor. Soc.*, **135**, 1643–1660, https://doi.org/10.1002/qj.498.

Wang, H., S. Legg, and R. Hallberg, 2018: The effect of Arctic freshwater pathways on North Atlantic convection and the Atlantic meridional overturning circulation. *J. Climate*, **31**, 5165–5188, https://doi.org/10.1175/JCLI-D-17-0629.1.

Weaver, A. J., and E. C. Wiebe, 1999: On the sensitivity of projected oceanic thermal expansion to the parameterisation of sub-grid scale ocean mixing. *Geophys. Res. Lett.*, **26**, 3461–3464, https://doi.org/10.1029/1999GL002362.

Wetherald, R. T., R. J. Stouffer, and K. W. Dixon, 2001: Committed warming and its implications for climate change. *Geophys. Res. Lett.*, **28**, 1535–1538, https://doi.org/10.1029/2000GL011786.

Wiebe, E., and A. J. Weaver, 1999: On the sensitivity of global warming experiments to the parametrisation of sub-grid scale ocean mixing. *Climate Dyn.*, **15**, 875–893, https://doi.org/10.1007/s003820050319.

Winton, M., 2000: A reformulated three-layer sea ice model. *J. Atmos. Oceanic Technol.*, **17**, 525–531, https://doi.org/10.1175/1520-0426(2000)017<0525:ARTLSI>2.0.CO;2.

——, R. Hallberg, and A. Gnanadesikan, 1998: Simulation of density-driven frictional downslope flow in *Z*-coordinate ocean models. *J. Phys. Oceanogr.*, **28**, 2163–2174, https://doi.org/10.1175/1520-0485(1998)028<2163:SODDFD>2.0.CO;2.

——, K. Takahashi, and I. M. Held, 2010: Importance of ocean heat uptake efficacy to transient climate change. *J. Climate*, **23**, 2333–2344, https://doi.org/10.1175/2009JCLI3139.1.

——, A. Adcroft, S. M. Griffies, R. W. Hallberg, L. W. Horowitz, and R. J. Stouffer, 2013: Influence of ocean and atmosphere components on simulated climate sensitivities. *J. Climate*, **26**, 231–245, https://doi.org/10.1175/JCLI-D-12-00121.1.