

Flux Correction and Overturning Stability: Insights from a Dynamical Box Model[Ⓞ]

ANAND GNANADESIKAN, RICHARD KELSON, AND MICHAELA STEN

Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland

(Manuscript received 25 June 2018, in final form 28 August 2018)

ABSTRACT

A recent paper suggested that the global climate models used to project future climate changes may significantly overestimate the stability of the Atlantic meridional overturning circulation under anthropogenic global warming. It further asserted that “flux adjusting” the models (adding offsets to heat and freshwater fluxes to reproduce the observed density field) may reduce model stability. However, temperature, salinity, and density fields in climate models likely deviate from observations because of biases in model physics as well as inaccuracies in fluxes. In such cases it is unclear whether adjusting the fluxes to produce a more realistic density field will result in a model with more realistic stability properties, as flux correction may be compensating for other inaccuracies in model formulation. We investigate this question using a simplified dynamical box model, in which we can flux correct one version of the model to match density gradients within another version of the model. We show that flux adjustment can realistically compensate for biases in stability associated with some processes, such as uncertainty in the value of the vertical diffusivity, but not other processes, such as inaccurate simulation of the relationship between density structure and overturning or the isopycnal tracer diffusivity coefficient. An ability to compensate for biases when the overturning shuts off does not imply an ability to compensate for biases in when it reestablishes itself, and vice versa.

1. Introduction

The Atlantic meridional overturning circulation (AMOC) plays a critical role in global climate, transporting heat to polar regions and warming the high northern latitudes (Zhang and Delworth 2005). The resulting impact on climate is felt at global scales, with changes in the overturning causing shifts in tropical rainfall and asymmetries in tropical cyclone formation (Goldenberg et al. 2001; Zhang and Delworth 2006) as well as impacting ecosystems around the North Atlantic (Shuman et al. 2002). Because the high-latitude North Atlantic is both colder and fresher than subtropical waters, it has been recognized for many years that changes in freshwater fluxes could cause a collapse of the AMOC similar to that seen during the last deglaciation (Broecker et al. 1988). Stommel (1961) used a

simple box model to develop an analytic solution for overturning collapse, suggesting that it could occur if the density difference associated with salinity approached half of the density difference associated with temperature. His model has two potential states, one in which the overturning is strong and flushes the high latitudes with salty waters, and the reversed case where the overturning is weak and high latitudes are very fresh. Manabe and Stouffer (1988) found similar states within a fully coupled atmosphere–ocean general circulation model. However, their model was “flux adjusted” (i.e., the fluxes produced by the atmospheric and oceanic states were corrected so as to produce a realistic density field). Flux adjustment has been abandoned in the modeling community in recent years, and the resulting models have generally failed to show a strong instability in the overturning circulation under global warming (Weaver et al. 2012) or bistability under applied freshwater flux (see, however, Mecking et al. 2017).

Recent work by Liu et al. (2017) suggested that the current class of models is biased toward being too stable. Taking a version of the NCAR CESM (Collins et al. 2006; Yeager et al. 2006), Liu et al. flux adjust the density fields and find that the AMOC in the adjusted model

[Ⓞ] Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-18-0388.s1>.

Corresponding author: Anand Gnanadesikan, gnanades@jhu.edu

declines much more sharply under global warming than in the unadjusted model. They link this to a failure to simulate the salinity structure in the South Atlantic. As noted by [Rahmstorf \(1996\)](#), if the overturning imports freshwater into the Atlantic, a slower overturning will cause the Atlantic to become saltier and thus stabilize the overturning.

However, it is not clear whether the failure of models such as this particular version of NCAR CESM to simulate the observed salinity and temperature fields ultimately stems from incorrect surface fluxes or from biases in model physics. If biases in density result from lack of atmospheric water vapor transport into (or out of) the Atlantic basin, correcting surface water fluxes will address the root cause of the problem. On the other hand, if the problem lies in the physics of the climate model itself it is far from clear that flux adjustment would produce a more accurate estimate of model stability.

One way of answering this question is to flux correct one climate model such that it looks like another, and then seeing if the flux-corrected model reproduces the stability of the target model. Unfortunately, climate models are extremely complicated and differ greatly in resolution, physical parameterizations, and model numerics. As a result, even models with identical atmospheres and sea ice codes can give startlingly different overturning behavior depending on the ocean model ([Gnanadesikan et al. 2014](#)). This means that even if flux correction could make one model look like another, we still would have little sense of why.

In this paper we turn to a simpler system, a dynamical box model with three surface boxes (southern, low-latitude, and northern) and one deep box similar to that studied by [Johnson et al. \(2007\)](#). The volume of the low-latitude deep box is allowed to adjust to changes in forcing following the mass balance set out by [Gnanadesikan \(1999\)](#). In contrast to that paper, however, the temperature and salinity in the four boxes are allowed to vary. The model has a relatively small number of parameters, some of which are known to vary widely across climate models. This enables an experimental design where one parameter is varied, and the sensitivity of the AMOC to the freshwater flux into the North Atlantic box is examined.

This paper is complementary to previous work of [Marotzke and Stone \(1995\)](#) who considered the impacts of flux adjustment in a simple box model. Their model consisted of two boxes in the atmosphere and two in the ocean, representing high and low latitudes. The atmospheric heat and moisture transport between low and high latitudes were a function of the temperature gradient. The overturning circulation in the ocean was modeled as in [Stommel \(1961\)](#) with the circulation

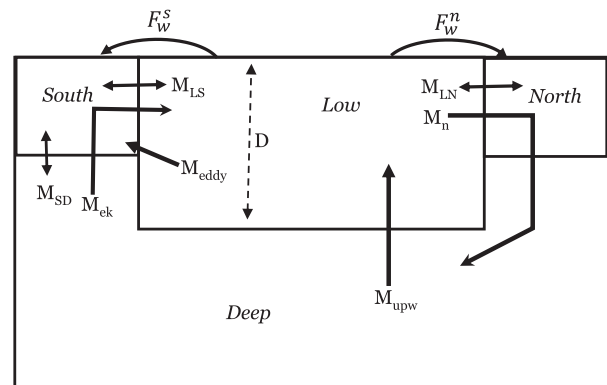


FIG. 1. Schematic of the dynamical box model used in this paper. Italics denote the individual boxes. Single-headed bold arrows denote net fluxes of water. Double-headed arrows denote mixing fluxes. The dashed line denotes the depth of the low- to midlatitude pycnocline D .

depending on the density difference between the two well-mixed, fixed-depth ocean boxes. Marotzke and Stone found that flux adjustment was not able to reproduce differences in stability arising from biases in the atmospheric transports—highlighting the importance of properly modeling atmospheric physics and dynamics. In contrast to their work, our paper considers the impact of biases in oceanic physics.

The model is introduced in [section 2](#), and the uncertainties in a number of parameters are considered in [section 3](#). Results for changing these parameters are presented in [section 4](#), and the implications and discussions are summarized in [section 5](#). We show that flux adjustment is able to adjust for biases in how much freshwater is needed to collapse the overturning due to certain differences in physical parameterization, but not for others. However, even in the cases where flux adjustment does provide a more realistic estimate of the marginal stability for a collapse of the overturning, it may not be able to correctly estimate the freshwater flux at which the overturning reestablishes itself.

2. Model description

The model used in this paper is a dynamical box model with four boxes: three surface boxes in the southern high latitudes ($<30^{\circ}\text{S}$), northern high latitudes ($>45^{\circ}\text{N}$), and mid- to low latitudes, which for conciseness we will refer to as the low-latitude box; and a deep box that lies beneath all of the surface boxes. A schematic of this model is shown in [Fig. 1](#), and the formulation of the different fluxes between the boxes, described below in more detail, is summarized in [Table 1](#).

The physics governing the model evolution is as follows. The volumes of the high-latitude surface boxes are

TABLE 1. Summary of the formulation of the fluxes given used to transport tracers and mass between the different boxes and the default values of these fluxes.

	Description	Formula	Default value
M_{eddy}	Advective eddy flux in the Southern Ocean	$M_{\text{eddy}} = \frac{A_{\text{GM}} D L_x^s}{L_y^s}$	10 Sv
M_n	Northern Hemisphere overturning	$M_n = \frac{g' D^2}{\epsilon}$	20 Sv
M_{upw}	Upwelling from the deep ocean	$M_{\text{upw}} = \frac{K_v A}{D}$	5 Sv
$M_{\text{LS,LN}}$	Mixing fluxes between low and high latitudes to simulate along along-isopycnal mixing of heat and salt	$M_{\text{LS,LN}} = A_{\text{Redi}} D \frac{L_x^{s,n}}{L_y^{s,n}}$	10 and 2 Sv for south and north, respectively
M_{ek}	Ekman flux from the Southern Ocean	$M_{\text{ek}} = \frac{\tau_s L_x^s}{\rho f_s}$	25 Sv
M_{SD}	Antarctic Bottom Water formation	$M_{\text{SD}} = \text{const}$	15 Sv

fixed, but the volume of the low-latitude box and deep box is a function of the pycnocline depth D , which is allowed to vary with time. The volume of the low-latitude box V_{Low} is controlled by a mass balance following Gnanadesikan (1999) in which there are two inputs—diapycnal upwelling from the deep ocean M_{upw} and Ekman flux from the Southern Ocean M_{ek} —and four outputs—the Northern Hemisphere overturning M_n , the advective eddy flux in the Southern Ocean M_{eddy} , which returns a significant fraction of the Ekman flux to the Southern Ocean, and two freshwater fluxes transfer mass between the tropics and high-latitude boxes, so that the equation for volume tendency is as follows:

$$\frac{\partial V_{\text{Low}}}{\partial t} = -\frac{\partial V_{\text{Deep}}}{\partial t} = M_{\text{upw}} + M_{\text{ek}} - M_{\text{eddy}} - M_n - F_w^n - F_w^s. \quad (1)$$

We use the same closures as Gnanadesikan (1999) for each of the fluxes denoted by M in Eq. (1). The upwelling flux M_{upw} depends on the magnitude of the vertical diffusion coefficient K_v , the depth of the low-latitude box D , and the surface area of the low-latitude box (Area_{Low} ; set to $2 \times 10^{14} \text{ m}^2$ to cover the low to midlatitudes):

$$M_{\text{upw}} = \frac{K_v \text{Area}_{\text{Low}}}{D}. \quad (2)$$

The Southern Ocean Ekman flux is calculated as

$$M_{\text{ek}} = \frac{\tau_s L_x^s}{\rho f_s}, \quad (3)$$

where τ_s is a wind stress, L_x^s is the length around the circumpolar region (set to 25 000 km), and f_s is a Coriolis parameter. In our code we set M_{ek} as a model parameter. The streamfunction associated with the eddy flux M_{eddy} depends on the interface height diffusion coefficient A_{GM}

(Gent and McWilliams 1990) and the slope of the pycnocline in the Southern Hemisphere D/L_y^s , where L_y^s (set to 1000 km here) is the length scale over which the pycnocline shallows in the Southern Ocean. Integrating around the circumpolar region and using the same length scale as the Ekman flux L_x^s gives us the eddy flux:

$$M_{\text{eddy}} = \frac{A_{\text{GM}} D L_x^s}{L_y^s}. \quad (4)$$

Finally, the Northern Hemisphere overturning M_n depends on the density difference between the northern high and low latitudes, which we use to define a reduced gravity

$$g' = g \frac{[\rho(S_N, T_N, p=0) - \rho_L(S_{\text{low}}, T_{\text{low}}, p=0)]}{\rho(S_N, T_N, p=0)}, \quad (5)$$

where the density is computed using the full nonlinear equation of state and $T_{\text{low},N}$, $S_{\text{low},N}$ refer to the temperatures and salinities in the low and northern boxes, respectively. Then

$$M_n = \frac{g' D^2}{\epsilon}, \quad (6)$$

where ϵ is a resistance parameter that governs how efficiently the depth-integrated pressure gradient (which goes as $g' D^2$) drives overturning. The parameter ϵ implicitly accounts for a variety of processes including the location of convection, the shape of the pycnocline, and the physics of entrainment in overflows that can potentially affect the rate of North Atlantic Deep Water formation.

We also add mixing fluxes between the southern surface box and the deep ocean to simulate the formation of Antarctic Deep Waters and add mixing fluxes between the low- and high-latitude boxes in both hemispheres to simulate the along-isopycnal mixing of heat and salt resulting from gyre circulation and mesoscale eddy diffusion.

The exchange of water associated with the latter flux is defined as follows:

$$M_{LS,LN} = A_{\text{Redi}} D \frac{L_x^{s,n}}{L_y^{s,n}}, \quad (7)$$

by analogy with the Gent–McWilliams flux, but with $L_x^n = 5000$ km (as it only applies within the North Atlantic Ocean rather than around the Circumpolar Current) and $L_y^n = 1000$ km. The vertical mixing flux in the Southern Ocean is set to 15 Sv ($1 \text{ Sv} \equiv 10^6 \text{ m}^3 \text{ s}^{-1}$), to account for the fact that the deep ocean is affected by both northern and southern source waters. As noted by Gent (2018), because slowing down the overturning increases the salinity of the tropics it also results in more salt export from the tropics by gyres and eddies—our mixing parameterization can be thought of as a coarse representation of this.

The four net mass fluxes, M_{upw} , M_{ek} , M_{eddy} , M_n ; three mixing fluxes $M_{LS,LN,SD}$; and the freshwater fluxes $F_w^{n,s}$ are then used to advect temperature and salinity between the different boxes. Heat fluxes are handled by restoring the top 100 m of each surface box to a target temperature (in the cases described here 17°C for the low-latitude box, 2°C for the high-latitude Northern Hemisphere box, and 4°C—to account for both Antarctic intermediate and bottom water sources—in the Southern Hemisphere box). The full equations for the “on” state (in which the overturning generates deep water in the North Atlantic) and the “off” state (in which case the northern surface box is lighter than the low-latitude box) are presented in the appendix.

Below we vary a subset of these parameters (K_v , M_{ek} , A_{GM} , ϵ , and A_{Redi}) in ways that mimic differences seen across modern climate models in order to determine to what degree they effect the mean overturning circulation and its stability to changes in freshwater flux. We initialize each model with high- and low-latitude salinities and temperatures that match observations, but with two possible pycnocline depths D . The first is the observed value of around 400 m. The second is a much smaller value of around 100 m, which results in a very weak initial overturning circulation, which as we will show can result in a second stable state with a reversed overturning.

3. Sources of parameter uncertainty in climate models

a. Northern Hemisphere (NH) resistance to overturning

The relationship between overturning and NH density gradients may differ across models. Levermann and Fürst (2010) examined a suite of simulations using the

simplified Climate–Biosphere model (Climber-3a), which couples a coarse-resolution, simplified AGCM to a full ocean model in which the atmospheric carbon dioxide concentration, Southern Ocean wind stress, pycnocline diffusivity coefficient, and Northern Hemisphere freshwater flux were varied. Using a definition of D , which is twice the value we use here, they found that a constant value of the resistance parameter $\epsilon \approx 2 \times 10^{-4} \text{ s}^{-1}$ produced a reasonable fit to this suite but that the actual parameter in individual simulations differed by up to 15%. De Boer et al. (2010) used an idealized coupled model and found that changes in overturning geometry could yield substantially different values of ϵ . In our model we started by assuming a pycnocline depth (defined as the first moment of the density perturbation relative to 2000 m in the 2013 *World Ocean Atlas*) of 400 m and an NH overturning of 20 Sv. The difference between the average density within the top 400 m of the Atlantic (the average observed density over this depth range at 20°–45°N and 80°–40°W is 1026.3 kg m^{-3}) and the North Atlantic Deep Waters (average observed density at 45°–55°N, 80°–40°W from 1500 to 2500 m is 1027.8 kg m^{-3}) results in a density difference of 1.5 kg m^{-3} . Then letting $M_n = 20$ Sv, the resistance parameter must be $1.2 \times 10^{-4} \text{ s}^{-1}$. We then examined the impact of doubling and halving ϵ .

b. Vertical mixing

It is generally recognized that the value $K_v = 1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$ measured by Ledwell et al. (1993) is typical of the low-latitude thermocline. Given the e -folding scale of 400 m found in observed hydrography and the previously stated low-latitude area of $2 \times 10^{14} \text{ m}^2$, this then gives an upwelling flux of 5 Sv, or about $1/4$ of Northern Hemisphere overturning. Such a low value of low-latitude upwelling implies that much of the transformation of light to dense water in the North Atlantic is primarily balanced by upwelling in the Southern Ocean. As discussed in Gnanadesikan et al. (2004), higher values of the mixing coefficient produce much greater amounts of tropical upwelling, which results in lower radiocarbon concentrations in the tropics than observed. However, the effective mixing coefficient in models may be higher than $10^{-5} \text{ m}^2 \text{ s}^{-1}$ because of implicit mixing from the advection scheme. We thus consider three values of K_v : a higher value at $0.5 \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$, our baseline of $1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$, and a limiting lower value of 0.

c. Southern Hemisphere flux

The net transformation of dense to light water in the Southern Hemisphere is the sum of a northward Ekman flux and southward eddy flux. Given a Northern Hemisphere overturning of 20 Sv, and a tropical upwelling

of 5 Sv, our model would then demand that the sum of the two Southern Hemisphere fluxes be 15 Sv. This is broadly consistent with Meijers (2014) who found the net inflow of Circumpolar Deep Water to the Southern Ocean in the CMIP5 models to be 13 ± 9 Sv.

Climate models have long shown significant differences in how well they simulate wind stress over the Southern Ocean, with clear implications for the resulting Ekman flux. Russell et al. (2006) found a twofold range in the maximum wind stress over the Southern Ocean across 23 models that were part of the CMIP3 intercomparison. Bracegirdle et al. (2013) demonstrated that the CMIP5 models showed a smaller range of maximum wind stresses from 0.18 to 0.22 Pa but also found a significant range in where this maximum was found, with one model placing it as far north at 45°S, and one placing it as far south at 53°S. The ensemble average of the CMIP5 models places the mean position at 50°S, about 2° equatorward of the observed mean position. While such differences may seem minor, moving the jet equatorward biases L_x^s high and f_s low, so getting the peak wind stress right but shifting the peak latitude equatorward by 2°, results in a 7% bias in Ekman flux. Across the entire range of CMIP5 models the potential bias is closer to 25%.

Exactly how such biases translate into the range of water mass transformation from 4 to 22 Sv seen by Meijers (2014) is unclear. Allison et al. (2010) and Marshall et al. (2016) argue that the average Ekman flux over the latitudes of the Circumpolar Current streamlines provides the best measure of the relevant flux. They note that the center of the current will shift as the wind stress shifts, and so the exact position of the winds may not matter. Using the ECMWF winds and the Maximenko (2009) surface dynamic topography yields an Ekman flux of about 35 Sv. On the other hand, we are concerned with the winds that accomplish water mass transformation between the dense deep waters and lighter surface waters. It is less clear what the critical isopycnal is for this process. If we assume it is the Ekman flux crossing the $\sigma_0 = 26.3$ surface corresponding to the average density of the North Atlantic waters above 400 m, we get a flux of about 35 Sv. However, if instead we use $\sigma_0 = 27.1$, a density that is close to the base of the intermediate water and possibly more appropriate for looking at how much dense deep water is transformed to light water, we get a value closer to 25 Sv. Ultimately, it will be the mapping of this flux onto the density fields produced by a coupled model which will determine this effective rate of transformation. Below, we consider three values for the Ekman flux, 15, 25, and 35 Sv, with the central value as the control case.

The Gent–McWilliams coefficient in the Southern Ocean also varies across models. In many modern climate

models, the coefficient varies spatially with a value that is a function of the growth rate of baroclinic eddies and thus depends on the slope of isopycnals in the main pycnocline. The peak value, however, is generally capped to avoid instability in regions where there is deep convection and the slope of isopycnals is infinite. A quick survey of the literature shows values of $800 \text{ m}^2 \text{ s}^{-1}$ in the GFDL global coupled carbon–climate Earth System Model (ESM2M; Dunne et al. 2012), $1800 \text{ m}^2 \text{ s}^{-1}$ in the U.S. Department of Energy’s Energy Exascale Earth System Model (E3SM; M. Petersen et al. 2018, personal communication), and $2000 \text{ m}^2 \text{ s}^{-1}$ in the HadGEM model (Johns et al. 2006). Below we consider three values of A_{GM} , 500, 1000, and $2000 \text{ m}^2 \text{ s}^{-1}$, roughly spanning the range of potential values used in modern models.

In full coupled models it is likely that changes in wind stress and eddy coefficients might also be associated with changes in the meridional length scale L_y^s (Levermann and Fürst 2010), but currently no theory exists to predict such changes. Moreover, it has long been recognized that eddy fluxes are not uniform but instead are concentrated in the lee of topographic features (MacCready and Rhines 2001; Hallberg and Gnanadesikan 2001). Our results should thus be taken as illustrative of the kinds of variability associated with changes in the geometry of Southern Ocean flows and strength of Southern Ocean eddies rather than strictly predictive of the impact of changing an eddy coefficient—much less the impact of a changing eddy field. Given $L_y^s = 1000 \text{ km}$, a value of $1000 \text{ m}^2 \text{ s}^{-1}$ with a pycnocline depth of 400 m will give an eddy return flux of 10 Sv, closing our balance when the Ekman flux is set to 25 Sv.

d. Along-isopycnal mixing coefficient

As described in Gnanadesikan et al. (2018, manuscript submitted to *J. Geophys. Res. Oceans*) the lateral mixing coefficient A_{Redi} (Redi 1982) shows even more variability across models than does A_{GM} . Some modern climate models [e.g., GFDL’s ESM2G model of Dunne et al. (2012) and NCAR’s CESM4 of Danabasoglu et al. (2012)] prescribe values of $A_{\text{Redi}} = A_{GM}$ with a spatially varying pattern that depends on the local baroclinicity. Others use constant values ranging from $500 \text{ m}^2 \text{ s}^{-1}$ [the HadGEM model of Johns et al. (2006)] to $2000 \text{ m}^2 \text{ s}^{-1}$ [the CMCC model of Fogli et al. (2009)]. Still others scale the diffusion coefficient to the grid spacing [the MPI model of Marsland et al. (2003) has a maximum value of $400 \text{ m}^2 \text{ s}^{-1}$]. As described in Pradal and Gnanadesikan (2014), the value of A_{Redi} can have important impacts on high-latitude convection, as higher values transport more salt into high-latitude surface layers, destabilizing the surface layer and allowing for more vigorous deep convection. In this study we choose a

TABLE 2. Key parameters varied in our experiments with base values shown.

Parameter	Description	Default value/units	Trials
K_v	Vertical diffusion coefficient	$0.1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$	$0, 1 \times 10^{-5}, 5 \times 10^{-5}$
M_{ek}	Southern Ocean Ekman flux	25 Sv	15, 25, 35
A_{GM}	Gent–McWilliams coefficient	$1000 \text{ m}^2 \text{ s}^{-1}$	500, 1000, 2000
ϵ	Resistance parameter	$1.2 \times 10^{-4} \text{ s}^{-1}$	$0.6 \times 10^{-4}, 1.2 \times 10^{-4}, 2.4 \times 10^{-4}$
A_{Redi}	Lateral diffusion coefficient	$1000 \text{ m}^2 \text{ s}^{-1}$	500, 1000, 2000

baseline value of $A_{Redi} = 1000 \text{ m}^2 \text{ s}^{-1}$, equivalent to our baseline A_{GM} . We also consider the impact of doubling this coefficient (yielding a value identical to that used in CMCC) and halving it (yielding a value identical to that used in HadGEM). We do this independently of A_{GM} as is done in a subset of climate models. Such divergence may be justified by the fact that A_{GM} involves advection of potential vorticity. In the quasigeostrophic limit, potential vorticity anomalies are only advected by the ageostrophic flow, not the geostrophic flow, whereas no such inherent limitation exists for passive tracers. Smith and Marshall (2009) note that in this physical limit over a flat-bottomed ocean, the two mixing coefficients have quite different vertical distributions.

4. Impact of parameter changes on model stability

We now step through different coefficients, though in a slightly different order than they were presented above. We start with the vertical diffusion, then move to the Ekman flux and A_{GM} coefficients, and then finish with the resistance parameter ϵ and the lateral tracer diffusion coefficient A_{Redi} . As will be seen, this allows us to contrast the parameters where flux adjustment does a reasonable job in estimating the critical stability with the ones where it can yield serious errors. A summary of the experiments, recapping the base values and perturbations outlined in the previous section, can be found in Table 2.

a. Vertical diffusion

We begin by looking at how the vertical diffusive coefficient affects the stability of the overturning circulation as freshwater flux is increased. As shown in Fig. 2a, all three of our simulations show that the overturning circulation decreases and then collapses as the NH freshwater flux increases. When the mixing coefficient is zero, the collapse occurs when the freshwater flux is above 0.65 Sv and corresponds to an overturning of 13.5 Sv. For the baseline case, the freshwater flux must be greater than 0.75 Sv and the corresponding overturning is 16.9 Sv, while for the high mixing case the collapse occurs at a flux of 0.95 Sv and an overturning exceeding 28 Sv. The collapse is associated with a deepening of the pycnocline (Fig. 2c). This deepening sharply

increases the eddy-induced flux of light water into the Southern Ocean, and represents a state where this flux essentially balances the sum of the Ekman flux and the low-latitude diffusive upwelling ($M_{eddy} \approx M_{ek} + M_{upw}$). We also see that this “off mode” of the overturning can be accessed from an initial condition in which the pycnocline depth is initially very small (even though the pycnocline in the off-mode equilibrium state is very deep). Thus, both the on and off modes are stable solutions of the equations. There is a range of values of freshwater fluxes where both states can coexist. Such two-state behavior of this type of model was described by Johnson et al. (2007) and Cimadoribus et al. (2014). As described in Rahmstorf et al. (2005) similar behavior is found in fully coupled general circulation models as well.

However, if we look at the density difference at which the overturning collapses, it is very similar across all three simulations. The common threshold density for overturning collapse has interesting implications for what flux correction would do to our model, because flux correction aims to reproduce the surface density distribution. We illustrate this schematically in Fig. 3. Suppose that we have a “true model” given by the blue line and an initial freshwater flux, giving a true solution at point A. The model we wish to flux adjust (shown by the red line) will have a density difference that is too large (shown by point B). Flux correcting this model requires adding a large enough water flux to move our model to the solution at C, where it will have a density difference identical to the true model. Note that while point B is a long way from the point at which the model will collapse, point C is much closer to the collapse point.

Returning to Fig. 2d, we can see how this works for models with errors in the diffusion coefficient. If we assume that the correct diffusive coefficient is $1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$, the observed density difference between the pycnocline waters of the Atlantic and North Atlantic Deep Waters of 1.5 kg m^{-3} would correspond to a NH freshwater flux of $\sim 0.5 \text{ Sv}$ in the control simulation, which is about 0.25 Sv below the critical transition point. This would imply that an additional freshwater flux of 0.25 Sv would be required shut off the overturning. If our model had too much diffusion, but the correct

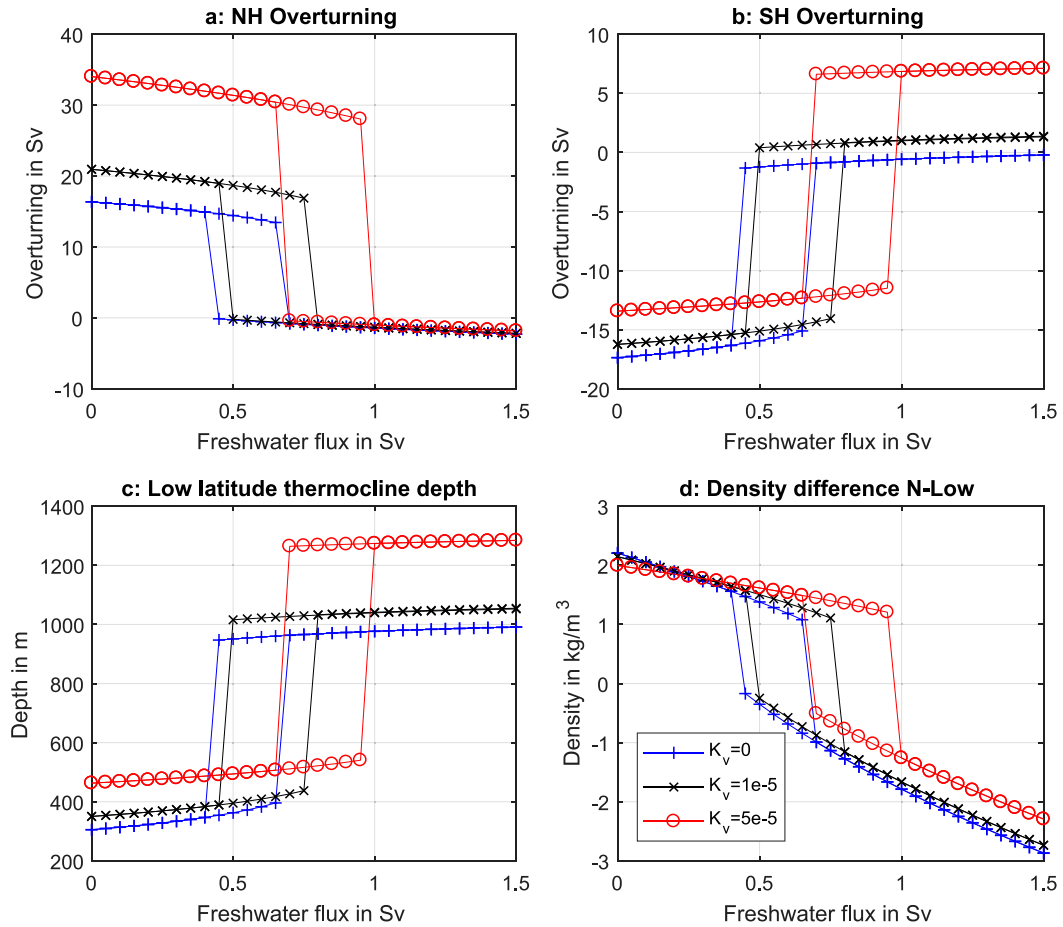


FIG. 2. Sensitivity of the model solution to the changing Northern Hemisphere freshwater flux at different values of the vertical eddy mixing coefficient K_v that affects the flux M_{upw} in Fig. 1. Blue lines with plus signs show the results for $K_v = 0 \text{ m}^2 \text{ s}^{-1}$, black lines with \times symbols show the results for $K_v = 1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$, (default for the other simulations), and red lines with open circles show results for $K_v = 1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$. (a) NH overturning M_n ; (b) SH overturning ($A_{GM}DL_x^s/L_y^s + \tau_s L_x^s/\rho f_s$), positive values correspond to net transformation of light to dense water as for M_n ; (c) low-latitude pycnocline depth; and (d) density difference between the high-latitude NH surface box and low-latitude surface box.

freshwater flux, the density gradient would be slightly too large (1.61 kg m^{-3}), and such a model would predict that the required additional freshwater flux to shut off the overturning would be 0.45 Sv , so that the overturning would be biased stable. However, if we flux correct the high mixing model to get the right density gradient, we would add an additional 0.15 Sv of freshwater flux. Thus, the marginal increase of the freshwater flux required to switch off the overturning would then be only 0.3 Sv (much closer to the baseline value of 0.25 Sv), and the density criterion to switch off the overturning would be nearly correct (1.21 vs 1.11 kg m^{-3}). Flux adjustment in this case would not result in a model that reproduced the solution of the baseline model (the overturning would still be too high and the pycnocline too shallow), but the estimate for the size of the

perturbation required to shut off the overturning would be substantially improved.

We can visualize the impact of flux adjustment by adjusting the relationship between freshwater flux and the density difference between northern and low latitudes seen in Fig. 2d. If we flux adjust the high mixing model by attributing 0.15 Sv of the freshwater flux to “flux adjustment,” we can slide the red curve in Fig. 2d to the left. Similarly, if we flux adjust the low-mixing model by attributing -0.075 Sv of the freshwater flux to flux adjustment we can slide the blue curve in Fig. 2d to the right. The resulting set of curves is shown in Fig. 4a, with the green circle showing the point toward which the different models are adjusted. It can be seen that the three curves are relatively congruent, with both the collapse and reestablishment of the overturning occurring

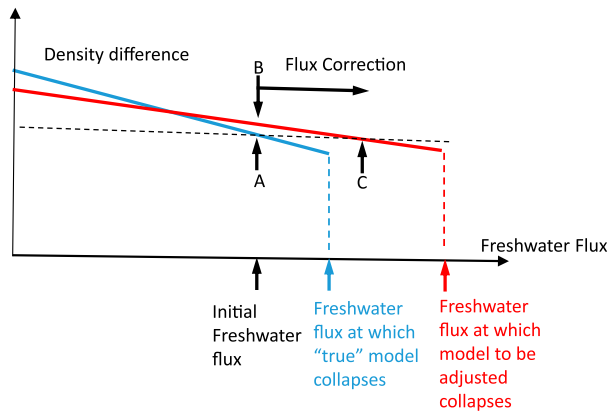


FIG. 3. Schematic of how to think about flux correction when the model solutions are presented in terms of the density difference between high and low latitude as a function of the imposed Northern Hemisphere freshwater flux. Point A represents the solution assuming the correct model (whose dependence on the freshwater flux is shown by the blue line) and correct freshwater flux. Point B represents a solution for a different model (whose dependence on freshwater flux is shown by the red line) when the correct freshwater flux is applied. Point C is the solution for this incorrect model when a flux correction is applied to ensure the correct surface density difference. Note that in this case, point C is closer to the collapse point for the incorrect model.

within a range of 0.1 Sv and at relatively similar densities. This suggests that flux adjustment does provide a means of correcting for the impact of biases in vertical mixing both for the collapse and reestablishment of the overturning.

b. Ekman flux

We now turn to the impact of changing the Southern Ocean Ekman flux. As seen in Fig. 5a, the same picture of collapse seen in the previous subsection holds as freshwater fluxes increase, with a concomitant deepening of the pycnocline (Fig. 5c) and shift in the Southern Ocean circulation from net negative to near-zero/net positive (Fig. 5b). Similar to vertical mixing, increasing the magnitude of the Ekman flux increases both the overturning circulation and the freshwater flux required to shut it off.

Focusing on the point of collapse reveals a subtler picture. The overturning at the point of collapse is quite different across the three values of Ekman flux, 10.2 Sv for a 15-Sv Ekman flux but 28 Sv for a 35-Sv Ekman flux—an increase of almost a factor of 3. However, the required freshwater flux is 0.55 Sv in the low-Ekman flux case and 0.85 Sv in the high Ekman flux case. Increasing the “pumping” of water from the Ekman flux by a factor of 2.3 thus requires only a 55% increase in freshwater flux to shut off the overturning.

If we were to assume the real world were like our base-case model, flux-adjusting the high Ekman flux

model flux to match the true density gradient would give us a NH freshwater flux of around 0.6 Sv, again ~ 0.25 Sv below the transition point. For the low Ekman flux model, matching the density gradient gives us a freshwater flux of ~ 0.37 Sv, about 0.18 Sv below the transition point. In both cases, matching the density gradient moves the perturbed model closer to the “true” distance from the transition point and thus improves the estimate of how large a perturbation freshwater flux is required to shut off the overturning.

This is illustrated graphically in Fig. 4b. The flux-adjusted model with high Ekman flux looks remarkably similar to our control. Although the density difference between the northern and low-latitude box is systematically different in the off state (in contrast to Fig. 4a), the overturning is reestablished at the same freshwater flux. The low Ekman flux case, however, looks systematically different, with a noticeably smaller range over which it exhibits bistability and systematically different behavior of density in the off state.

c. Gent–McWilliams eddy coefficient

A broadly similar picture to Fig. 5 is seen for the Gent–McWilliams coefficient (Fig. 6), where lower values of the coefficient give a stronger transformation of dense to light water in the Southern Ocean as less of the Ekman flux is returned. This then produces a stronger NH overturning, requiring a higher freshwater flux to shut it off. For $A_{GM} = 2000 \text{ m}^2 \text{ s}^{-1}$ the overturning collapses between freshwater fluxes of 0.55 and 0.6 Sv and a strength of ~ 10.9 Sv. These values are very similar to the $M_{ek} = 15$ Sv case reported in the previous section. For $A_{GM} = 500 \text{ m}^2 \text{ s}^{-1}$ the collapse occurs between freshwater fluxes of 0.8 and 0.85 Sv at a value of ~ 21.8 Sv. However, as was the case for the Ekman flux, the value of the critical density gradient has greater similarity between the two extreme cases (Fig. 3c), occurring near a value of 1.24 kg m^{-3} in the low mixing case and 1.09 kg m^{-3} in the high mixing case.

The performance of flux adjustment at predicting the collapse of the overturning in the perturbed models when the bias in A_{GM} is similar to that for M_{ek} . The high mixing case gets the right density for a freshwater flux of 0.38 Sv and the transition occurs at 0.55 Sv, just as it does in the $M_{ek} = 15$ Sv run, so that a perturbation of ~ 0.17 Sv is required to shut off the overturning. This is still larger than the 0.05 increase we would predict if we got the freshwater flux right but the GM coefficient wrong, and so just as with Ekman flux, flux adjustment helps to improve the estimate of stability in our physically biased model.

However, we now see that the width of the bistable zone is a nonlinear function of mixing, shrinking at both

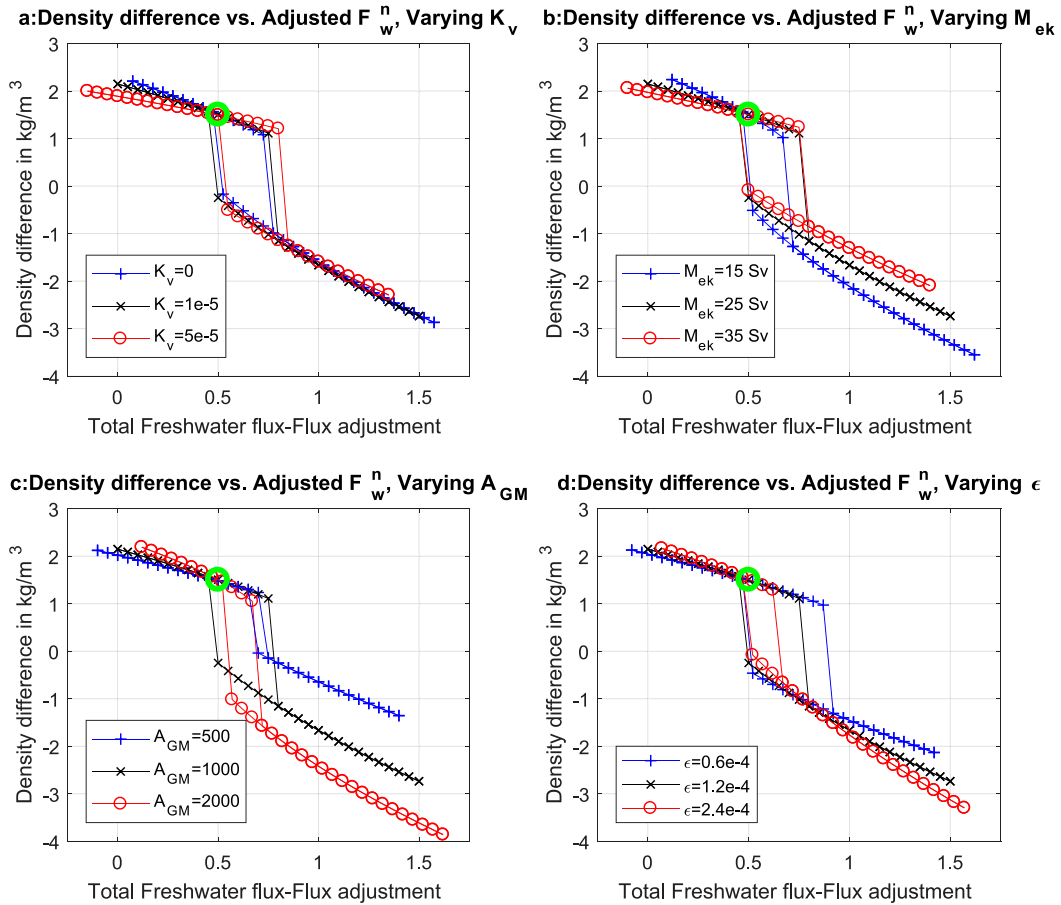


FIG. 4. Relationship between the density gradient and freshwater flux across the different “flux adjusted” model suites. Black lines show the control case, blue lines show the case for the low value of the parameter being varied, and red lines show the case for the high value of the parameter being varied. Thick green circles show the points toward which we are flux-adjusting the suites. (a) Effect of varying vertical diffusion coefficient K_v . (b) Effect of varying Southern Ocean Ekman flux M_{ek} . (c) Effect of varying Gent–McWilliams coefficient, which affects Southern Ocean eddy flux M_{eddy} . (d) Effect of varying the resistance coefficient ϵ , which affects the proportionality between depth-integrated pressure gradient and NH overturning M_n .

high and low values of A_{GM} . This means that it is impossible to slide the curves in Fig. 5d horizontally so that they will overlay each other. Although we may get some qualitative improvement in estimating what level of freshwater perturbation is required to shut off the overturning, we do not necessarily get a better prediction of the critical value of freshwater flux at which the overturning reestablishes itself. This can be seen quite clearly in Fig. 4c. Such variation of the size of the bistable zone is also seen across climate models, as described by Rahmstorf et al. (2005).

d. NH resistance parameter

We now turn to the resistance parameter ϵ that governs the relationship between the large-scale density structure and the overturning. Increasing this parameter results in an overturning that is weaker and that collapses at lower

values of freshwater flux (Fig. 7a). At $\epsilon = 2.4 \times 10^{-4} s^{-1}$ the overturning collapses at a value of $\sim 14.4 Sv$ and a freshwater flux of 0.55 Sv, while at $\epsilon = 0.6 \times 10^{-4} s^{-1}$ it collapses at a value of $\sim 19.7 Sv$ and a freshwater flux of 0.95 Sv, almost twice as high. The increased resistance results in a deeper pycnocline depth (Fig. 7c), which then reduces the net transformation of dense to light water in the Southern Ocean (Fig. 7b). In the $\epsilon = 2.4 \times 10^{-4} s^{-1}$ case, the target density of $1.5 kg m^{-3}$ is found at a freshwater flux value of 0.43 Sv, close to the 0.5 Sv in which it is found in the baseline case. Similarly, the model with low resistance produces the right density gradient at 0.58 Sv. The resulting flux adjustments are thus small in comparison with the differences in freshwater flux required to shut off the overturning.

As a result, when the model is flux adjusted (Fig. 4d) the different curves do not lie on top of each other. The

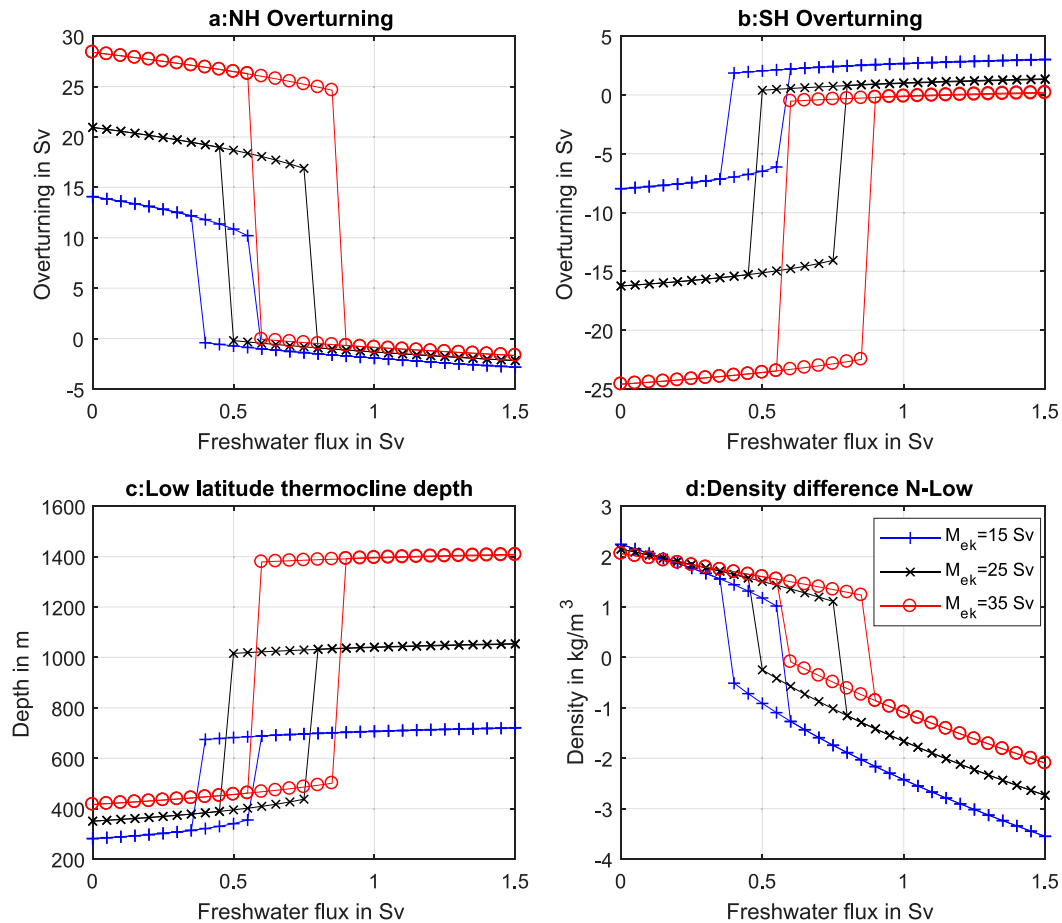


FIG. 5. Sensitivity of the model solution to the changing Northern Hemisphere freshwater flux at different values of the Southern Hemisphere Ekman flux. Blue lines with plus signs show the results for an Ekman flux of 15 Sv, black lines with \times symbols show the results for an Ekman flux of 25 Sv (default for the other simulations), and red lines with open circles show results for an Ekman flux of 35 Sv. (a) NH overturning M_n ; (b) SH overturning ($A_{GM}DL_x^S/L_y^S + \tau_s L_x^S/\rho_f s$), positive values correspond to net transformation of light to dense water as for M_n ; (c) low-latitude pycnocline depth; and (d) density difference between the high-latitude NH surface box and low-latitude surface box.

high-resistance model still collapses at a lower freshwater flux than the control, while the low-resistance model still collapses at a higher freshwater flux than the control. As the models reestablish their overturning at the same freshwater flux, it is the width of the bistable region that changes with resistance. Flux adjusting the model to account for biases in resistance may move the models in the right direction but will not correct for most of the bias. This is particularly worrisome as the resistance parameter is the parameter within our box model that is the least well justified from a physical point of view.

e. Along-isopycnal diffusion coefficient

The along-isopycnal diffusion coefficient A_{Redi} has a weaker impact on overturning and its sensitivity than A_{GM} does, but its impacts are in the opposite direction.

As shown in Fig. 8, increasing A_{Redi} from 500 to $2000\text{ m}^2\text{ s}^{-1}$ causes the freshwater flux required to produce overturning collapse to increase from 0.7 to 0.85 Sv and the value of overturning at which the collapse occurs to decrease slightly from 17.5 to 16.3 Sv. This makes sense as high diffusion coefficients reduce the salinity difference between high northern and low latitudes, acting in the opposite direction to the freshwater flux. The A_{Redi} has a big impact on the range of freshwater fluxes over which the model has two stable states, with higher values sharply reducing the range. This is potentially important—as noted by Cimattoribus et al. (2014) and Gent (2018), mixing of salinity out of the southern Atlantic acts to prevent two stable states from developing, and it may well be the case that the real ocean behaves more like this high-mixing regime.

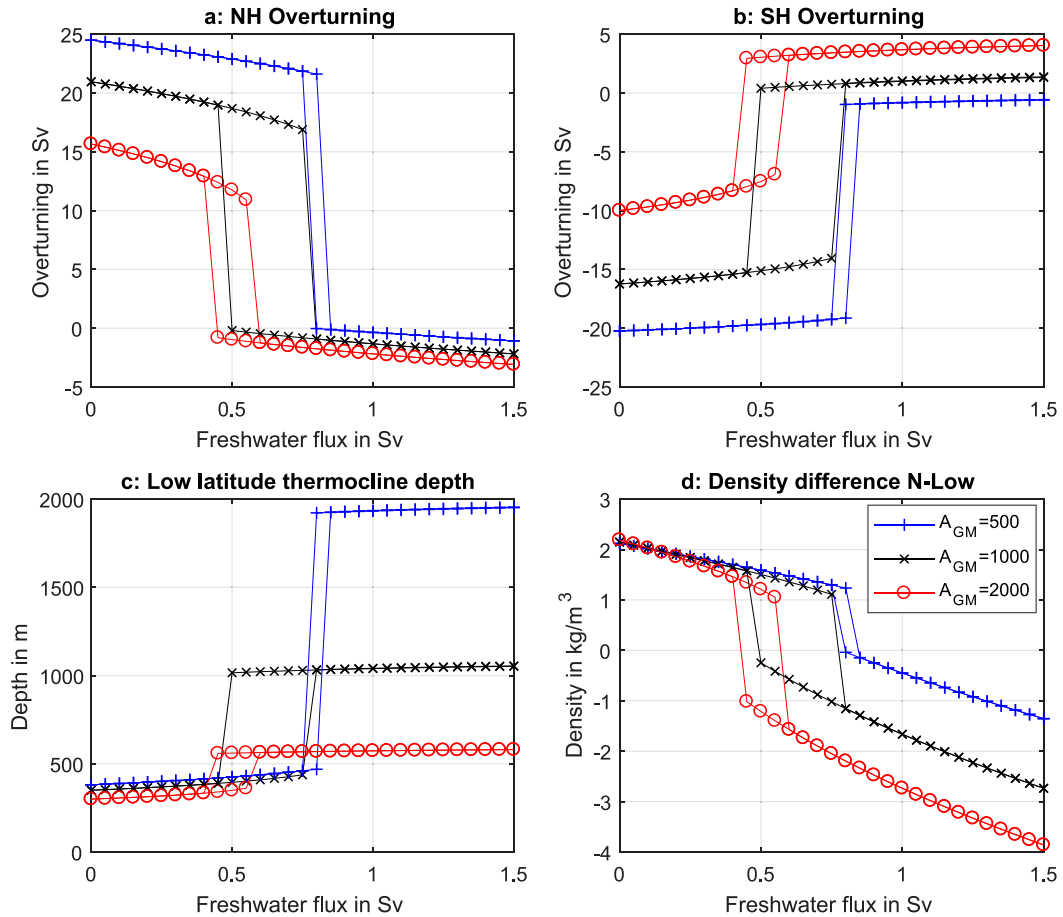


FIG. 6. Sensitivity of the model solution to the changing Northern Hemisphere freshwater flux at different values of the Gent–McWilliams eddy mixing coefficient A_{GM} that affects the flux M_{eddy} in Fig. 1. Blue lines with plus signs show the results for $A_{GM} = 500 \text{ m}^2 \text{ s}^{-1}$, black lines with \times symbols show the results for $A_{GM} = 1000 \text{ m}^2 \text{ s}^{-1}$ (default for the other simulations), and red lines with open circles show results for $A_{GM} = 2000 \text{ m}^2 \text{ s}^{-1}$. (a) NH overturning M_H ; (b) SH overturning ($A_{GM}DL_x^s/L_y^s + \tau_s L_x^s/\rho f_s$), positive values correspond to net transformation of light to dense water as for M_H ; (c) low-latitude pycnocline depth; and (d) density difference between the high-latitude NH surface box and low-latitude surface box.

Flux adjustment will do little to improve our estimates of when the overturning collapses and when it is reestablished. Along the branch of the solution where the NH overturning is on, the solutions for fluxes below about 0.7 Sv are essentially identical for the three values of A_{Redi} . This means that if our true model had a density difference of 1.5 kg m^{-3} between the low latitude and northern surface boxes, but the wrong A_{Redi} , we would calculate a flux adjustment of zero. As a result, we would do nothing to actually correct the stability biases introduced by the incorrect coefficient. In contrast to all the examples shown in Fig. 4, the flux-adjusted version of Fig. 8d would be unchanged.

5. Conclusions

We have shown that flux correction may offer useful insight into the stability of the ocean overturning beyond

what would be expected from its ability to correct for biases in fluxes. In our simple model, while flux adjustment does not correct for biases in circulation because of incorrect values of vertical mixing, Southern Ocean Ekman flux, or Southern Ocean eddy overturning, it does substantially allow the incorrect model to better estimate the size of the perturbation of freshwater flux needed to shut off the overturning and the density difference between low- to midlatitudes and high northern latitudes at which the collapse occurs. However, this is not necessarily true for all parameters—weak dependence of the solution on the value of the lateral mixing coefficient A_{Redi} means that it cannot correct for the differences in stability that biases in this coefficient (which are remarkably poorly known) can produce. Similarly, flux correction does not capture the changes in stability associated with biases

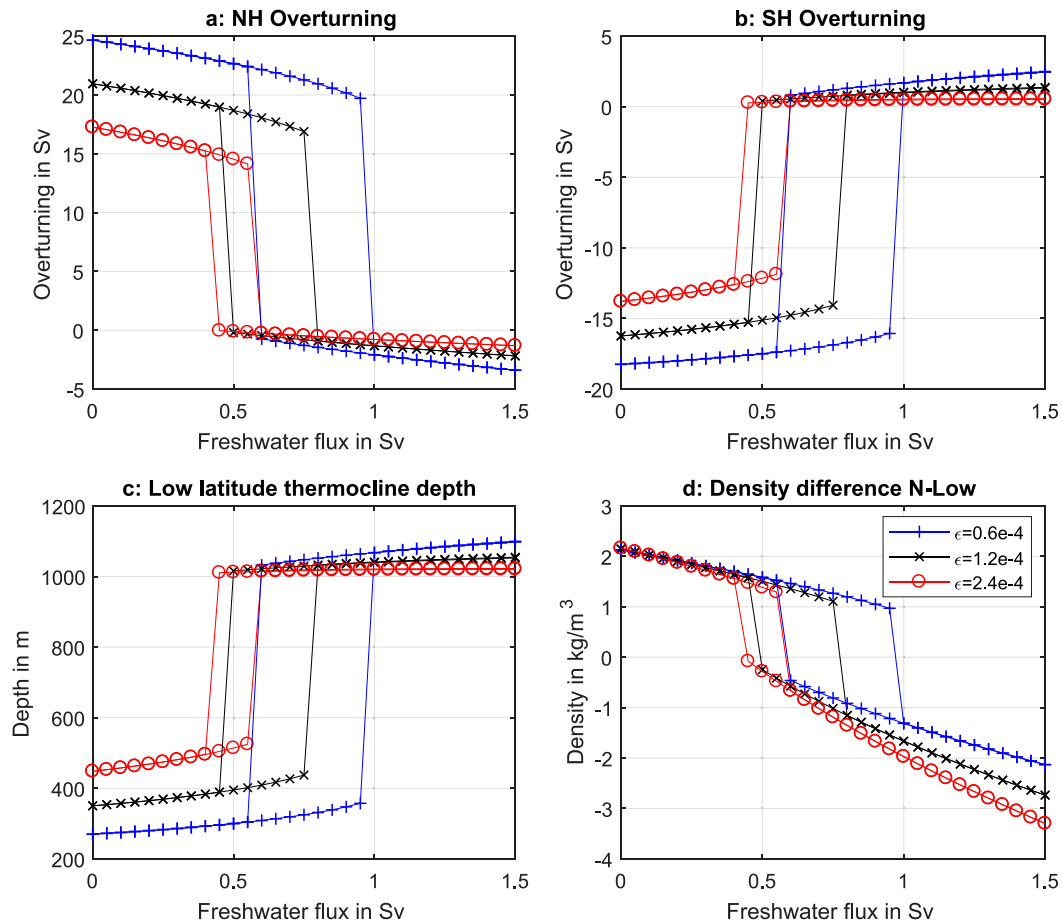


FIG. 7. Sensitivity of the model solution to the changing Northern Hemisphere freshwater flux at different values of the resistance parameter ϵ , which governs the conversion between the large-scale pressure gradients and the overturning. Blue lines with plus symbols show the results for $\epsilon = 6 \times 10^{-5} \text{ s}^{-1}$, black lines with \times symbols show the results for $\epsilon = 1.2 \times 10^{-4} \text{ s}^{-1}$ (default for the other simulations), and red lines with open circles show results for $\epsilon = 2.4 \times 10^{-4} \text{ s}^{-1}$. (a) NH overturning M_n ; (b) SH overturning ($A_{GM}DL_x^2/L_y^2 + \tau_s L_x^2/\rho f_s$), positive values correspond to net transformation of light to dense water as for M_n ; (c) low-latitude pycnocline depth; and (d) density difference between the high-latitude NH surface box and low-latitude surface box.

in the parameter ϵ governing the relationship between large-scale pressure gradients and NH overturning. With the exception of biases in vertical mixing (where the range of freshwater fluxes for which the model is bistable is relatively similar across different values), flux adjusting the model to fix biases in the base state also fails to fix biases in when the overturning will be reestablished.

This result has some interesting implications for the strategy employed by Liu et al. (2017). First it suggests that there may be qualitative value in understanding how a flux-corrected version of a climate model behaves. If the resulting model is more stable or less stable than the base model, this result may indicate that the base model has either too little or too much stability. However, the reverse is not necessarily true—

a model that gets the surface densities right does not necessarily always get the stability right. Our results also suggest that a detailed process understanding of North Atlantic overturning, in particular the ways in which large-scale density fields are converted into an overturning circulation, continues to be a highly worthwhile endeavor.

Extensions of this work to more complicated box models may also yield useful insight. For example, recent work by Cessi and Jones (2017) extended a similar model to a two-basin configuration that was found to qualitatively predict pathways of interbasin exchange. Similarly, Cimattoribus et al. (2014) added a subtropical “supergyre” box in order to look at advective removal of salinity. A key question that such models address is whether the water masses that feed

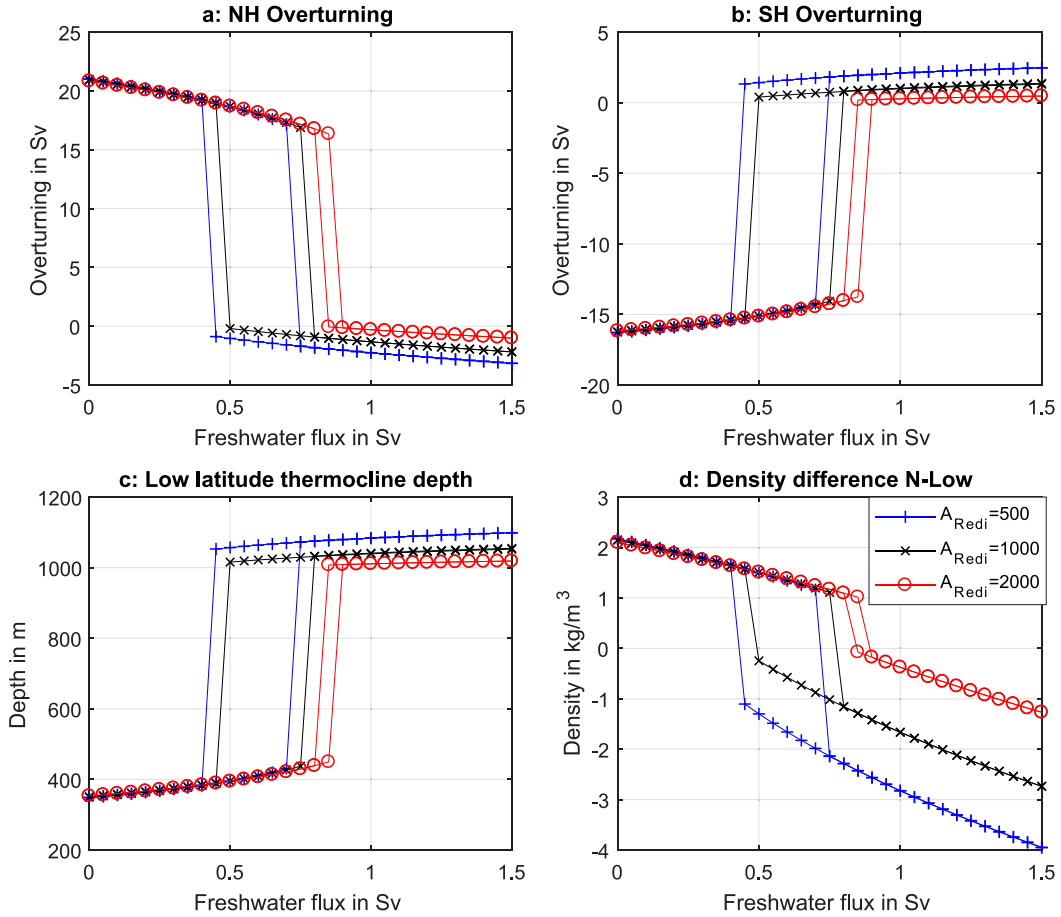


FIG. 8. Sensitivity of the model solution to the changing Northern Hemisphere freshwater flux at different values of the lateral eddy mixing coefficient A_{Redi} that affects the flux $M_{\text{LS,LN}}$ in Fig. 1. Blue lines with plus signs show the results for $A_{\text{Redi}} = 500 \text{ m}^2 \text{ s}^{-1}$, black lines with \times symbols show the results for $A_{\text{Redi}} = 1000 \text{ m}^2 \text{ s}^{-1}$ (default for the other simulations), and red lines with open circles show results for $A_{\text{Redi}} = 2000 \text{ m}^2 \text{ s}^{-1}$. (a) NH overturning M_n ; (b) SH overturning ($A_{\text{GM}}DL_x^2/L_y^2 + \tau_s L_x^2/\rho f_s$), positive values correspond to net transformation of light to dense water as for M_n ; (c) low-latitude pycnocline depth; and (d) density difference between the high-latitude NH surface box and low-latitude surface box.

the NH overturning enter the Atlantic basin as relatively fresh Antarctic waters or relatively salty Indian Ocean surface waters. The Liu et al. (2017) paper that initially motivated this work argued that the current generation of models gets this balance wrong. We plan to investigate such questions further in future work.

Acknowledgments. Anand Gnanadesikan acknowledges support from the U.S. AMOC program under NOAA Grant NA16OAR4310174. Kevin Huang and Alyssa Sooklal also participated in the class assignment that led to this manuscript. The model code used here is provided as supplemental material. Our thanks to Peter Gent, David Marshall, and two anonymous reviewers for comments that greatly improved this manuscript.

APPENDIX

Tracer Balance Equations

When the reduced gravity [Eq. (4)] is greater than zero, combining Eqs. (1)–(7) gives us

$$\begin{aligned} \frac{\partial V_{\text{low}}}{\partial t} &= \text{Area}_{\text{Low}} \frac{\partial D}{\partial t} \\ &= M_{\text{upw}} + M_{\text{ek}} - M_{\text{eddy}} - M_n - F_w^n - F_w^s \\ &= \frac{K_v \text{Area}_{\text{Low}}}{D} + M_{\text{ek}} - \frac{A_{\text{GM}} DL_x^s}{L_y^s} - \frac{g'D^2}{\epsilon} - F_w^s - F_w^n, \end{aligned} \tag{A1}$$

which is time-stepped forward to solve for the depth. We solve for the evolution of the heat and salt content

(rather than temperature and salinity) in the low-latitude box as

$$\begin{aligned} \frac{\partial}{\partial t}(V_{\text{low}}S_{\text{low}}) &= M_{\text{upw}}S_{\text{deep}} + (M_{\text{ek}} + M_{\text{LS}})S_S \\ &\quad + M_{\text{LN}}S_N - (M_n + M_{\text{eddy}} \\ &\quad + M_{\text{LS}} + M_{\text{LN}})S_{\text{low}}, \end{aligned} \quad (\text{A2a})$$

$$\begin{aligned} \frac{\partial}{\partial t}(V_{\text{low}}T_{\text{low}}) &= M_{\text{upw}}T_{\text{deep}} + (M_{\text{ek}} + M_{\text{LS}})T_S + M_{\text{LN}}T_N \\ &\quad - (M_n + M_{\text{eddy}} + M_{\text{LS}} + M_{\text{LN}} + F_w^S \\ &\quad + F_w^N)T_{\text{low}} + v_T\text{Area}_{\text{Low}}(T_{\text{low}}^{\text{rad}} - T_{\text{low}}), \end{aligned} \quad (\text{A2b})$$

where v_T represents an air–sea heat transfer velocity of 100 m yr^{-1} . The equations for the salt content in the remaining three boxes are as follows:

$$\begin{aligned} \text{Area}_N D_{\text{high}} \frac{\partial S_N}{\partial t} &= (M_{\text{LN}} + M_n)S_{\text{low}} \\ &\quad - (M_{\text{LN}} + M_n + F_w^n)S_N, \end{aligned} \quad (\text{A3a})$$

$$\begin{aligned} \text{Area}_S D_{\text{high}} \frac{\partial S_S}{\partial t} &= (M_{\text{LS}} + M_{\text{eddy}})S_{\text{low}} \\ &\quad + (M_{\text{ek}} + M_{\text{SD}})S_{\text{deep}} - (M_{\text{LN}} \\ &\quad + M_{\text{ek}} + F_w^S + M_{\text{SD}} + M_{\text{eddy}})S_S, \end{aligned} \quad (\text{A3b})$$

$$\begin{aligned} \frac{\partial}{\partial t}(V_{\text{deep}}S_{\text{deep}}) &= (F_w^S + M_{\text{SD}} + M_{\text{eddy}})S_S \\ &\quad + (F_w^N + M_n)S_N - (M_{\text{LN}} + M_{\text{ek}} \\ &\quad + M_{\text{upw}} + M_{\text{SD}})S_{\text{deep}}, \end{aligned} \quad (\text{A3c})$$

where $\text{Area}_{S,N}$ and $D_{\text{high}} = 100\text{ m}$ represent the areas and depth (fixed in time) of the northern and southern surface layer boxes. The equation for the evolution of temperature content in the remaining three boxes is

$$\begin{aligned} \text{Area}_N D_{\text{high}} \frac{\partial T_N}{\partial t} &= (M_{\text{LN}} + M_n)T_{\text{low}} - (M_{\text{LN}} + M_n)T_N \\ &\quad + v_T\text{Area}_N(T_{\text{rad}}^N - T_N), \end{aligned} \quad (\text{A4a})$$

$$\begin{aligned} \text{Area}_S D_{\text{high}} \frac{\partial T_S}{\partial t} &= (M_{\text{LS}} + M_{\text{eddy}})T_{\text{low}} + (M_{\text{ek}} + M_{\text{SD}})T_{\text{deep}} \\ &\quad - (M_{\text{LN}} + M_{\text{ek}} + M_{\text{SD}} + M_{\text{eddy}})T_S \\ &\quad + v_T\text{Area}_S(T_{\text{rad}}^S - T_S), \end{aligned} \quad (\text{A4b})$$

$$\begin{aligned} \frac{\partial}{\partial t}(V_{\text{deep}}T_{\text{deep}}) &= (F_w^S + M_{\text{SD}} + M_{\text{eddy}})T_S \\ &\quad + (F_w^N + M_n)T_N - (M_{\text{LN}} + M_{\text{ek}} \\ &\quad + M_{\text{upw}} + M_{\text{SD}})T_{\text{deep}}. \end{aligned} \quad (\text{A4c})$$

Note that we assume (in common with most climate models) that freshwater enters and leaves the ocean surface at the temperature of that surface. After each time step, the salt and heat content in each box are divided by the volume of each box to compute the temperature and salinity.

When the reduced gravity [Eq. (4)] is less than zero, the overturning circulation is set to

$$M_n = \frac{g'D_{\text{high}}^2}{\epsilon}. \quad (\text{A5})$$

As this flow now moves from the deep box through the north box, it alters the low, north, and deep salt content equations:

$$\begin{aligned} \frac{\partial}{\partial t}(V_{\text{low}}S_{\text{low}}) &= M_{\text{upw}}S_{\text{deep}} + (M_{\text{ek}} + M_{\text{LS}})S_S \\ &\quad + (M_{\text{LN}} - M_n)S_N - (M_{\text{eddy}} \\ &\quad + M_{\text{LS}} + M_{\text{LN}})S_{\text{low}}, \end{aligned} \quad (\text{A6a})$$

$$\begin{aligned} \text{Area}_N D_{\text{high}} \frac{\partial S_N}{\partial t} &= M_{\text{LN}}S_{\text{low}} - M_nS_{\text{deep}} \\ &\quad - (M_{\text{LN}} - M_n + F_w^n)S_N, \end{aligned} \quad (\text{A6b})$$

$$\begin{aligned} \frac{\partial}{\partial t}(V_{\text{deep}}S_{\text{deep}}) &= (F_w^S + M_{\text{SD}} + M_{\text{eddy}})S_S + F_w^N S_N \\ &\quad - (M_{\text{LN}} + M_{\text{ek}} + M_{\text{upw}} + M_{\text{SD}} - M_n)S_{\text{deep}}, \end{aligned} \quad (\text{A6c})$$

as well as the heat content equations:

$$\begin{aligned} \frac{\partial}{\partial t}(V_{\text{low}}T_{\text{low}}) &= M_{\text{upw}}T_{\text{deep}} + (M_{\text{ek}} + M_{\text{LS}})T_S + (M_{\text{LN}} \\ &\quad - M_n)T_N - (M_{\text{eddy}} + M_{\text{LS}} + M_{\text{LN}} + F_w^S \\ &\quad + F_w^N)T_{\text{low}} + v_T\text{Area}_{\text{Low}}(T_{\text{low}}^{\text{rad}} - T_{\text{low}}), \end{aligned} \quad (\text{A7a})$$

$$\begin{aligned} \text{Area}_N D_{\text{high}} \frac{\partial T_N}{\partial t} &= M_{\text{LN}}T_{\text{low}} - M_nT_{\text{deep}} - (M_{\text{LN}} - M_n)T_N \\ &\quad + v_T\text{Area}_N(T_{\text{rad}}^N - T_N), \end{aligned} \quad (\text{A7b})$$

$$\begin{aligned} \frac{\partial}{\partial t}(V_{\text{deep}} T_{\text{deep}}) = & (F_w^S + M_{\text{SD}} + M_{\text{eddy}})T_S + F_w^N T_N \\ & - (M_{\text{LN}} + M_{\text{ck}} + M_{\text{upw}} \\ & + M_{\text{SD}} - M_n)T_{\text{deep}}. \end{aligned} \quad (\text{A7c})$$

REFERENCES

- Allison, L. C., H. L. Johnson, D. P. Marshall, and D. R. Munday, 2010: Where do winds drive the Antarctic Circumpolar Current? *Geophys. Res. Lett.*, **37**, L12605, <https://doi.org/10.1029/2010GL043355>.
- Bracegirdle, T. J., E. Shuckburgh, J.-B. Sallee, Z. Wang, A. J. S. Meijers, N. Bruneau, T. Phillips, and L. J. Wilcox, 2013: Assessment of surface winds over the Atlantic, Indian, and Pacific Ocean sectors of the Southern Ocean in CMIP5 models: Historical bias, forcing response, and state dependence. *J. Geophys. Res. Atmos.*, **118**, 547–562, <https://doi.org/10.1002/jgrd.50153>.
- Broecker, W. S., M. Andreev, W. Wolfli, H. Oeschger, G. Bonani, J. Kennett, and D. Peteet, 1988: The chronology of the last Deglaciation: Implications to the cause of the Younger Dryas Event. *Paleoceanography*, **3**, 1–19, <https://doi.org/10.1029/PA003i001p00001>.
- Cessi, P., and C. S. Jones, 2017: Warm-route versus cold-route interbasin exchange in the meridional overturning circulation. *J. Phys. Oceanogr.*, **47**, 1981–1997, <https://doi.org/10.1175/JPO-D-16-0249.1>.
- Cimatoribus, A. A., S. S. Drijfhout, and H. A. Dijkstra, 2014: Meridional overturning circulation: Stability and ocean feedbacks in a box model. *Climate Dyn.*, **42**, 311–328, <https://doi.org/10.1007/s00382-012-1576-9>.
- Collins, W. D., and Coauthors, 2006: The Community Climate System Model version 3 (CCSM3). *J. Climate*, **19**, 2122–2143, <https://doi.org/10.1175/JCLI3761.1>.
- Danabasoglu, G., S. C. Bates, B. P. Briegleb, S. R. Jayne, M. Jochum, W. G. Large, S. Peacock, and S. G. Yeager, 2012: The CCSM4 ocean component. *J. Climate*, **25**, 1361–1389, <https://doi.org/10.1175/JCLI-D-11-00091.1>.
- de Boer, A., A. Gnanadesikan, N. R. Edwards, and A. J. Watson, 2010: Meridional density gradients do not control the Atlantic overturning circulation. *J. Phys. Oceanogr.*, **40**, 368–380, <https://doi.org/10.1175/2009JPO4200.1>.
- 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>.
- Fogli, P. G., and Coauthors, 2009: INGV-CMCC Carbon (ICC): A Carbon Cycle Earth System Model. CMCC Research Paper 61, 31 pp., <https://doi.org/10.2139/ssrn.1517282>.
- Gent, P. R., 2018: A commentary on the Atlantic meridional overturning stability in climate models. *Ocean Modell.*, **122**, 57–66, <https://doi.org/10.1016/j.ocemod.2017.12.006>.
- , 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](https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2).
- Gnanadesikan, A., 1999: A simple predictive model for the structure of the oceanic pycnocline. *Science*, **283**, 2077–2079, <https://doi.org/10.1126/science.283.5410.2077>.
- , J. P. Dunne, R. M. Key, K. Matsumoto, J. L. Sarmiento, R. D. Slater, and P. S. Swathi, 2004: Oceanic ventilation and biogeochemical cycling: Understanding the physical mechanisms that produce realistic distributions of tracers and productivity. *Global Biogeochem. Cycles*, **18**, GB4010, <https://doi.org/10.1029/2003GB002097>.
- , —, and R. Msadek, 2014: Connecting Atlantic temperature variability and biological cycling in two Earth System Models. *J. Mar. Syst.*, **133**, 39–54, <https://doi.org/10.1016/j.jmarsys.2013.10.003>.
- Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Núñez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**, 474–479, <https://doi.org/10.1126/science.1060040>.
- Hallberg, R., and A. Gnanadesikan, 2001: An exploration of the role of transient eddies in determining the transport of a zonally reentrant current. *J. Phys. Oceanogr.*, **31**, 3312–3330, [https://doi.org/10.1175/1520-0485\(2001\)031<3312:AEOTRO>2.0.CO;2](https://doi.org/10.1175/1520-0485(2001)031<3312:AEOTRO>2.0.CO;2).
- Johns, T. C., and Coauthors, 2006: The New Hadley Centre Climate Model (HadGEM1): Evaluation of coupled simulations. *J. Climate*, **19**, 1327–1353, <https://doi.org/10.1175/JCLI3712.1>.
- Johnson, H. L., D. P. Marshall, and D. A. J. Sproson, 2007: Reconciling theories of a mechanically driven meridional overturning circulation with thermohaline forcing and multiple equilibria. *Climate Dyn.*, **29**, 821–836, <https://doi.org/10.1007/s00382-007-0262-9>.
- 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>.
- Levermann, A., and J. J. Fürst, 2010: Atlantic pycnocline theory scrutinized using a coupled climate model. *Geophys. Res. Lett.*, **37**, L14602, <https://doi.org/10.1029/2010GL044180>.
- Liu, W., S.-P. Xie, Z. Liu, and J. Zhu, 2017: Overlooked possibility of a collapsed Atlantic Meridional Overturning Circulation in warming climate. *Sci. Adv.*, **3**, e1601666, <https://doi.org/10.1126/sciadv.1601666>.
- MacCready, P., and P. B. Rhines, 2001: Meridional transport across a zonal channel: Topographic localization. *J. Phys. Oceanogr.*, **31**, 1427–1439, [https://doi.org/10.1175/1520-0485\(2001\)031<1427:MTAAZC>2.0.CO;2](https://doi.org/10.1175/1520-0485(2001)031<1427:MTAAZC>2.0.CO;2).
- Manabe, S., and R. J. Stouffer, 1988: Two stable equilibria of a coupled ocean-atmosphere model. *J. Climate*, **1**, 841–866, [https://doi.org/10.1175/1520-0442\(1988\)001<0841:TSEOAC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1988)001<0841:TSEOAC>2.0.CO;2).
- Marotzke, J., and P. H. Stone, 1995: Atmospheric transports, the thermohaline circulation, and flux adjustments in a simple coupled model. *J. Phys. Oceanogr.*, **25**, 1350–1364, [https://doi.org/10.1175/1520-0485\(1995\)025<1350:ATTTCA>2.0.CO;2](https://doi.org/10.1175/1520-0485(1995)025<1350:ATTTCA>2.0.CO;2).
- Marshall, D. P., D. R. Munday, L. C. Allison, R. J. Hay, and H. L. Johnson, 2016: Gill's model of the Antarctic Circumpolar Current, revisited: The role of latitudinal variations in wind stress. *Ocean Modell.*, **97**, 37–51, <https://doi.org/10.1016/j.ocemod.2015.11.010>.
- Marsland, S. J., H. Haak, J. H. Jungclaus, M. Latif, and F. Röske, 2003: The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates. *Ocean Modell.*, **5**, 91–127, [https://doi.org/10.1016/S1463-5003\(02\)00015-X](https://doi.org/10.1016/S1463-5003(02)00015-X).
- Maximenko, N., 2009: Mean dynamic topography of the ocean derived from satellite and drifting buoy data using three different techniques. *J. Atmos. Oceanic Technol.*, **26**, 1910–1919, <https://doi.org/10.1175/2009JTECHO672.1>.
- Mecking, J. V., S. S. Drijfhout, L. C. Jackson, and M. B. Andrews, 2017: The effect of model bias on Atlantic freshwater transport and implications for AMOC bi-stability. *Tellus*, **69A**, 1299910, <https://doi.org/10.1080/16000870.2017.1299910>.

- Meijers, A. J. S., 2014: The Southern Ocean in the Coupled Model Intercomparison Project phase 5. *Philos. Trans. Roy. Soc. London*, **372A**, 2019, <https://doi.org/10.1098/rsta.2013.0296>.
- Pradal, M. A., and A. Gnanadesikan, 2014: How does the Redi parameter for mesoscale mixing impact climate in an Earth System Model? *J. Adv. Model. Earth Syst.*, **6**, 586–601, <https://doi.org/10.1002/2013MS000273>.
- Rahmstorf, S., 1996: On the freshwater forcing and transport of the Atlantic thermohaline circulation. *Climate Dyn.*, **12**, 799–811, <https://doi.org/10.1007/s003820050144>.
- , and Coauthors, 2005: Thermohaline circulation hysteresis: A model intercomparison. *Geophys. Res. Lett.*, **32**, L23605, <https://doi.org/10.1029/2005GL023655>.
- Redi, M., 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](https://doi.org/10.1175/1520-0485(1982)012<1154:OIMBCR>2.0.CO;2).
- Russell, J. L., R. J. Stouffer, and K. W. Dixon, 2006: Intercomparison of the Southern Ocean circulations in IPCC coupled model control simulations. *J. Climate*, **19**, 4560–4575, <https://doi.org/10.1175/JCLI3869.1>.
- Shuman, B., T. Webb III, P. Bartlein, and J. W. Williams, 2002: The anatomy of a climatic oscillation: Vegetation change in eastern North America during the Younger Dryas chronozone. *Quat. Sci. Rev.*, **21**, 1777–1791, [https://doi.org/10.1016/S0277-3791\(02\)00030-6](https://doi.org/10.1016/S0277-3791(02)00030-6).
- Smith, K. S., and J. Marshall, 2009: Evidence for enhanced eddy mixing at middepth in the Southern Ocean. *J. Phys. Oceanogr.*, **39**, 50–69, <https://doi.org/10.1175/2008JPO3880.1>.
- Stommel, H. M., 1961: Thermohaline convection with two stable regimes of flow. *Tellus*, **13**, 224–230, <https://doi.org/10.1111/j.2153-3490.1961.tb00079.x>.
- Weaver, A. J., and Coauthors, 2012: Stability of the Atlantic meridional overturning circulation: A model intercomparison. *Geophys. Res. Lett.*, **39**, L20709, <https://doi.org/10.1029/2012GL053763>.
- Yeager, S. G., C. A. Shields, W. G. Large, and J. J. Hack, 2006: The low-resolution CCSM3. *J. Climate*, **19**, 2545–2566, <https://doi.org/10.1175/JCLI3744.1>.
- Zhang, R., and T. L. Delworth, 2005: Simulated tropical response to a substantial weakening of the Atlantic thermohaline circulation. *J. Climate*, **18**, 1853–1860, <https://doi.org/10.1175/JCLI3460.1>.
- , and —, 2006: Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophys. Res. Lett.*, **33**, L17712, <https://doi.org/10.1029/2006GL026267>.