Flux correction and overturning stability: Insights from a dynamical box model

Anand Gnanadesikan[1], Richard Kelson[1] and Michaela Sten[1]

1: Department of Earth and Planetary Sciences, Johns Hopkins University

Corresponding Author: Anand Gnanadesikan

327 Olin Hall

Johns Hopkins University

Baltimore, MD 21218

gnanades@jhu.edu

Abstract

Recent work has suggested that the non-flux adjusted global climate models used to project future climate changes may significantly overestimate the stability Atlantic Meridional Overturning Circulation under anthropogenic global warming. Recent work has suggested that correcting the temperature and salinity fields via flux adjustment reduces model stability. However temperature, salinity and density fields in flux-adjusted 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 ~~look like~~match density gradients within~~imitate the characteristics of~~ 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 eddy stirring coefficient A_{GM} and wind stress biases in the Southern Ocean) but not others (such as inaccurate simulation of the relationship between density structure and overturning or the eddy stirring coefficient A_{Redi}).

1. Introduction

The Atlantic Meridional Overturning Circulation plays a critical role in global climate, moving heat into polar regions and warming the high Northern l~~L~~atitudes (Zhang and Delworth, 2005). The resulting impact on climate is felt at global scales, with changes in the overturning linked causing shifts in tropical rainfall, 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 play cause a collapse of the AMOC similar to that seen during in the wake of the last deglaciation. 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 that associated with temperature. This model has two potential states, one in which the overturning is strong and flushes the high latitudes of salty waters, and a reversed case where the high latitudes are very fresh and the overturning is weak and reversed. 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).

Recent work by Liu et al. (2017) has suggested that the current class of models is biased towards being too stable. Taking a version of the NCAR CESM (Collins et al., 2006; Yeager et

al., 2006~~REF~~) they flux-adjust the density fields and find that the AMOC in the adjusted model declines much more sharply under global warming than in the unadjusted model. They link this to a failure to simulate the salinity structure and thus the freshwater flux associated with overturning.

However, it is not clear whether the failure of models such as the NCAR CESM to simulate the observed salinity and temperature field ultimately stems from incorrect surface fluxes or from biases in model physics. Correcting the density field will give the correct result if biases result from lack of water vapor transport into (or out of) the Atlantic basin as it addresses 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 would be to try to flux-correct one climate model to look like another and see whether such a correction reproduced the change in stability. However, climate models are extremely complicated and differ 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 it did so.

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 vary following the mass balance set out by Gnanadesikan (1999) and 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. We show that flux-adjustment is able to adjust for biases in stability due to certain differences in physical parameterization, but not for others. 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.

2. Model description

The model used in this paper is a dynamical box model with four boxes, surface boxes in the southern high, low and northern high latitudes and a deep box. The volumes of the high latitude surface boxes are fixed but the volume of the low latitude box and deep box vary as the pycnocline depth is allowed to vary. The volume of the low-latitude box V_{Low} is controlled by a mass balance following Gnanadesikan (1999) in which there are two inputs – upwelling from the deep ocean M_{upw} and Ekman flux from the Southern Ocean M_{ek} – and two outputs – the Northern Hemisphere overturning M_n and the advective eddy flux in the Southern Ocean M_{eddy} which returns a significant fraction of the Ekman flux to the Southern Ocean.

$$\frac{\partial V_{Low}}{\partial t} = -\frac{\partial V_{Deep}}{\partial t} = M_{upw} + M_{ek} - M_{eddy} - M_n \quad (1)$$

We use the same closures as Gnanadesikan (1999) for each of these fluxes. 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 this box A (set to 2x 10^{14} m^2) to cover the area of the tropics.

$$M_{upw} = \frac{K_v A}{D} \quad (2)$$

The Southern Ocean Ekman flux $M_{ek} = \frac{\tau_s L_x^S}{\rho f_s}$ where τ_s is a wind stress, ~~is the~~ L_x^S is the length around the Circumpolar region (set to 25000 km) and f_s is a Coriolis parameter. In our code 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_s^y where L_y^S (set to 800 km here) is the length scale over which the thermocline shallows in the Southern Ocean. Integrating around the Circumpolar region, using the same length scale as used for the Ekman flux- L_x^S, gives us the eddy flux

$$M_{eddy} = \frac{A_I D L_x^s}{L_y^s} \quad (3)$$

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

$$g' = g\,\frac{\{\rho(S_N, T_N, p=0) - \rho_L(S_L, T_L, p=0)\}}{\rho(S_N, T_N, p=0)} \quad (4)$$

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

$$M_n = \frac{g' D^2}{\epsilon} \quad (5)$$

w~~W~~here ϵ is a scaling parameter that incorporates a variety of geometric and physical controls on the overturning.

These four net mass fluxes are then used to advect temperature and salinity between the different boxes. Additionally, we add mixing fluxes between the surface southern box and the

deep ocean to simulate the formation of Antarctic deep waters, and mixing fluxes between the low and high latitude boxes in both hemispheres to simulate the along-isopycnal mixing of heat and salt due to mesocale eddy diffusion and gyre circulation. The latter fluxes are defined as

$$M_{LS,LN} = A_{Redi} * D * \frac{L_x^{s,n}}{L_y^{s,n}}$$

By 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. Finally, we include a vertical mixing flux in the Southern oceanSouthern Ocean of 15 Sv, to account for the fact that the deep ocean is affected by both Northern and Southern source waters. Below we vary a subset of these parameters (K_v, M_{ek}, A_{GM}, $A_{Redi}$and ϵ) in ways that mimic differences seen across modern climate models.

3. Sources of parameter uncertainty in climate models

a.) NH resistance to overturning

The relationship between overturning and NH density gradients may differ across models. Using the simplified coupled AGCM Climber-3a and a definition of *D* which is exactly twice the value we use here, Levermann and Furst (2010) found that a constant value of the resistance parameter $\epsilon \approx 2\text{x}10^{-4}$ s^{-1} produced a reasonable fit to a suite of simulations made under different conditions, but also that there was significant spread around this relationship. De Boer et al. (2007), using an idealized coupled model, 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 2000m) of 400m and an

NH overturning of 22.5 Sv. The difference between the average density within the top 400m of the Atlantic (26.3) and the North Atlantic Deep waters (27.8) allows us to solve for a resistance parameter of 1 x 10^{-4} s^{-1}. We then examined the impact of doubling and halving this value.

b.) Vertical mixing

It is generally recognized that the value ~~-~~ K_v = 0.1x 10^{-5} m^2/s measured by Ledwell et al., (1993) is typical of the low-latitude thermocline. Given the e-folding scale of 400m found in observed hydrography, this then gives an upwelling flux of 5 Sv, or about ¼ of Northern Hemisphere overturning. This implies that much of the transformation of light to dense water in the North Atlantic is 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 ~~then means that~~results in lower radiocarbon concentrations in the tropics ~~are much lower~~ than observed. However, the effective mixing coefficient in models may be higher than this due to implicit mixing from the advection scheme. We thus consider one larger value~~s~~ of the vertical mixing at 0.5 x 10^{-4} m^2/s 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 22.5 Sv, and a tropical upwelling of 5 Sv, our model would then demand that the sum of these two fluxes be 17.5 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 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, Stouffer and Dixon (2006) found a twofold range in the maximum wind stress over the Southern Ocean across 23 models that were part of the CMIP3 intercomparison. Swart and Fyfe (2012) showed that the CMIP5 models showed a smaller range of maximum wind stresses from (0.18 to 0.22 Pa) but with significant ranges in where this maximum was found, with one model placing it as far north at 45°S, one as far south at 53°S. The ensemble average of the CMIP5 models place the mean position at 50°S, about two degrees 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 the shifting the peak latitude equatorwards by 2 degrees results in a 7% bias in Ekman flux. Across the entire range of CMIP5 models it is closer to 25%. Variations in Ekman flux are much larger if one assumes that what matters is not the maximum wind stress, but the stress at the northern edge of Drake Passage (35 Sv in the ERA40 reanalysis). The wind stress at this latitude must be balanced by form drag with the bottom and thus must set the entire water column into motion, while winds further to the north can pile up water along the west coast of South America allowing for a near-surface geostrophic return flow. In the CMIP5 models, this flux can be as small as 6.5 Sv in the IPSL-CM5A models (Dufresne et al., 2013). Below, we consider three values for the Ekman flux, 20, 30 and 40 Sv, with the central value as the control case.

The Gent-McWilliams coefficient in the Southern Ocean also varies across models. In the GFDL ESM2M model (Dunne et al. 2012), the values are given by a closure based on the growth rate of baroclinic eddies in which the coefficient depends on the slope, but with a peak value of 800 m^2/s. In the HadGEM model (Johns et al., 2006) the peak value is 2000 m^2/s. In

the current version of the DOE E3SMv2 model it is 1800 m^2/s (Petersen et al., in prep.). Below we consider three values of A_{GM}, 500, 1000 and 2000 m^2/s, 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 coefficient might also be associated with changes in the meridional length scale L_y^s (Levermann and Furst, 2010), but currently no theory exists to predict such changes. Our results should thus be taken as illustrative of the kinds of changes 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. Given the values used here for L_y^s, a value of 1000 m^2/s with a pycnocline depth of 400 m will give an eddy return flux of 12.5 Sv, closing our balance when the Ekman flux is set to 30 Sv.

d.) Along-isopycnal mixing coefficient

As described in Gnanadesikan et al., (forthcoming) the lateral mixing coefficient A_{Redi} shows even more variability across models than does A_{GM}. Some modern climate models (for example GFDL's ESM2G model of Dunne et al¸2012 and NCAR CESM4 of Danabsasoglu et al., 2009) prescribe values of $A_{Redi} = A_{GM}$ with a spatially varying pattern that depends on the local baroclinicity. Others use constant values ranging from 500 m^2/s (the HadGEM model of Johns et al., 2006) to 2000 m^2/s (the CMCC model 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 m^2/s). 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 baseline value of A_{Redi}=1000 m^2/s, 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~~, and as~~ may be justified by the fact that A_{GM} involves advection of PV, which is not a passive tracer and thus may interact differently with turbulence.

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~~ing~~ with the vertical diffusion, then mov~~eing~~ to the Ekman flux and A_{GM} coefficients and then finish~~ing~~ with the A_{Redi} coefficient the efficiency parameter ϵ. As will be seen, this allows us to ~~moves~~contrast ~~from~~ the parameters where flux adjustment helps to ~~works best to find~~ the estimate critical stability with~~to~~ the ones where it can yield serious errors.

Formatted: Font: Italic

Formatted: Font: Italic, Subscript

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 above 0.8 Sv and corresponds to an overturning of 17.8 Sv. For the baseline case, the freshwater flux must be greater 0.9 Sv and the corresponding overturning is 21.2 Sv, while for the high mixing case the collapse occurs at a flux of 1.05 Sv and an overturning exceeding 33 Sv. The collapse is

associated with a deepening of the pycnocline (Fig 2b). 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 Ekman flux (or in the high mixing case, also carries an additional diffusive upwelling into the Southern Ocean). Such a two-state behavior of this type of model was described by Johnson et al. (2007).

However, if we look at the density difference at which the overturning collapses, it is essentially the same across all three simulations. This has interesting implications for what flux correction would do to our model. If we assume that the correct diffusive coefficient is 1 x 10^{-5} m^2/s, the observed density gradient of 1.5 kg/m^3 would correspond to a NH freshwater flux of ~0.57 Sv in the control simulation-about 0.33 Sv below the critical transition point. This would imply that an additional freshwater flux of 0.33 Sv would be required shut off the overturning. If our model has too much diffusion, but the correct freshwater flux, the density gradient would be slight too large (1.57 kg/m^3) and such a model would predict that the required additional freshwater flux to shut off the overturning would be 0.48 Sv. However, if we flux-correct the high mixing model to get the right density gradient, we would add an additional 0.1 Sv of freshwater flux, the marginal addition required to switch off the overturning would only be 0.38 Sv, and the density criterion to switch off the overturning would be exactly correct. 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 how stable the model is would be substantially improved. This can be seen by looking at the overturning plotted against the density gradient (Fig 3a) rather than the freshwater flux. Although the transition from northern overturning on to off occurs at a different value of freshwater flux

(transition occurs at a different point on the x-axis in Fig. 2a) it occurs at approximately the same value of density gradient (transition occurs at the same point on the x-axis in Fig. 3a).

b.) Ekman flux

A similar story holds for changing the Southern Ocean Ekman flux. As seen in Fig. 4a, the same picture of collapse holds as freshwater fluxes increase, with a concomitant deepening of the pycnocline and shift in the Southern Ocean circulation from net negative to near-zero/net positive. Similar to vertical mixing, increasing the magnitude of the Ekman flux both increases the overturning circulation and the freshwater flux required to shut it off.

Focussing 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, 14 Sv for a 20 Sv Ekman flux but 29 Sv for a 40 Sv Ekman flux - more than a factor of two. However, the required freshwater flux is 0.7 Sv in the low-mixing case and 1.0 Sv in the high-mixing case. Doubling the "pumping" of water from the Ekman flux thus requires only a 45% increase in freshwater flux to shut off the overturning.

If we were to assume the real world were like our central model, flux-adjusting the high Ekman flux model flux to match the true density gradient would give us a NH freshwater flux of around 0.65 Sv, again ~0.35 Sv below the transition point. For the low Ekman flux model, matching the density gradient gives us an Ekman flux of ~0.46 Sv, about 0.25 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. Note however, that the prediction is not exact; the flux-adjusted low Ekman flux model becomes unstable to positive

perturbations well before the "true" model does. This can be clearly seen by looking at Fig 3b, the transitions do not occur at exactly the same density, but shift from around 1.0 kg/m^3 in the low Ekman flux case to 1.2 kg/m^3 in the high ~~-~~mixing flux case.

c.) Gent-McWilliams eddy coefficient

A broadly similar picture to Fig. 4 is seen for the Gent-McWilliams coefficient (Fig. 5), where now 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 ~~and~~ a stronger NH overturning, requiring a higher freshwater flux to shut it off. For A_{GM} ~~AGM~~=2000 m^2/s the overturning collapses between freshwater fluxes of 0.7 and 0.75 Sv and a strength of about 14 Sv. These values are very similar to the M_{ek} ~~Mek~~=20 Sv case reported in the previous section. For A_{GM}=500 m^2/s the collapse occurs between freshwater fluxes of 0.95 and 1.0 Sv at a value of around 26 Sv. However, as was the case for the Ekman flux, the value of the critical density gradient is much more similar between the two extreme cases (Fig. 3c), occurring near a value of 1.2 kg/m^3 in the low mixing case and 1.0 kg/m^3 in the high mixing case.

The performance of flux adjustment at predicting the stability of the perturbed models when the bias is in A_{GM} is similar to that for M_{ek}. The high mixing case gets the right density for a freshwater flux of 0.47 Sv and the transition occurs at 0.7 Sv, just as it does in the M_{ek}=20 Sv run, so that a perturbation of about 0.23 Sv is required to shut off the overturning. This is still larger than the 0.13 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.

d.) Along-isopycnal diffusion coefficient

The along-isopycnal diffusion coefficient A_{Redi} has a weaker impact on overturning and its sensitivity than does A_{GM}, but its impacts are in the opposite direction. As shown in Fig. 6, increasing A_{Redi} from 500 to 2000 m^2/s results in the freshwater flux inducing the overturning to collapse to increase from 0.85 to 0.95 Sv and the value of overturning at which the collapse ~~to~~ occurs to decrease slightly from 21.5 to 21.1 Sv. This makes sense as high diffusion coefficients reduce the gradient in salinity, acting in the opposite direction as the freshwater flux. Along the branch of the solution where the NH overturning is on, the solutions for fluxes below about 0.85 Sv are essentially identical for the three values of A_{Redi}. As a result, as seen in Fig 3d, biases in A_{Redi} do not fundamentally change the relationship between density gradient and NH overturning when we have northern sinking. This means that if our true model had a density gradient of 1.5 kg/-m^{-3} but the wrong A_{Redi}, flux adjustment would do nothing to actually correct the stability biases introduced by this incorrect coefficient. Along the branch of the solution where the NH overturning is off, it appears that flux adjustment can make the solution at one value of A_{Redi} look like the solution at a higher value of A_{Redi}. In this case, however, such an adjustment will come at the price of making the model seem more stable than it actually is. Adjusting a solution along the blue line to a greater stability associated with the same flux but a higher mixing coefficient would require putting more freshwater into high latitudes, making it more difficult to recover from our “overturning off” state.

e.) NH efficiency of overturning

Finally, we turn to the 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 which collapses at lower values of freshwater flux (Fig. 7a). At ϵ=2 x 10^{-4} s^{-1} the overturning collapses at a value of around 18.4 Sv and a freshwater flux of 0.65 Sv, while at ϵ=2 x 10^{-4} s^{-1} it collapses at a value of around 24 Sv and a freshwater flux of 1.15 Sv, almost twice as high. The increased resistance results in a deep pycnocline depth (Fig. 7c) which then reduces the net transformation of dense to light water in the Southern Ocean (Fig. 7b). In the ϵ=2 x 10^{-4} s^{-1} case, the target density of 1.5 kg/m^3 is found at a value of 0.51 Sv, close to the 0.57 Sv in which it is found in the baseline case. Thus if we were to run the model with the "right" water flux but too high an ϵ we would conclude that we were too close to a collapse (with only a 0.08 Sv or 15% increase in water flux required rather than the 0.23 Sv increase found for our baseline case). Flux-adjusting the model would get us a model that required 0.14 Sv to shut off the overturning, which ~~moving~~ moves the model in the right direction, but only partly. Similarly, the model with low resistance produces the right density gradient at 0.62 Sv, requiring an additional flux of 0.43 Sv to shut off the overturning, rather than 0.48 Sv given the "right" water flux. Flux-adjustment in this case only compensates for a small fraction of the bias in stability. This can also be seen by looking at the relationship between overturning and density gradient (Fig. 7d) which clearly shows that the transition occurs at a substantially different point for the different values of ϵ.

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 due to 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 freshwater flux needed to shut off the overturning and the density gradient 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 is remarkably poorly known) can produce. Similarly, flux correction does not capture the changes in stability associated with biases in the parameter ϵ governing the relationship between large-scale pressure gradient and NH overturning.

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 which 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 which was found to qualitatively predict pathways of interbasin exchange. A key point of the Wu et al. (2017) paper that motivated this work was related to this question of whether the cold-water or warm water pathway represented the dominant supply for the NH overturning- as these different pathways have quite different implications for the salinity budget. We plan to investigate such questions further in future work.

References

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, doi:10.1002/jgrd.50153.

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.

Collins, W.D., C. M. Bitz, M. L. Blackmon, G. B. Bonan, C. S. Bretherton, J. A. Carton, P. Chang, S. C. Doney, J. J. Hack, T. B. Henderson, J. T. Kiehl, W. G. Large, D. S. McKenna, B. D. Santer, R. D. Smith (2006) , The Community Climate System Model version 3 (CCSM3). *J. Climate*, 19, 2122–2143.

de Boer, A., A. Gnanadesikan. N.E. Edwards and A. J. Watson, Meridional density gradients do not control the Atlantic Overturning Circulation, *J. Phys. Oceanogr.*,40, 368-380, doi: 10.1175/2009JPO4200.1, 2010.

Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nuñez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications. *Science*, **293**, 474, doi:10.1126/science.1060040

Gnanadesikan, A., 1999: A simple predictive model for the structure of the oceanic pycnocline, *Science.*, 283:2077-2079.

Gnanadesikan, A., 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*, GB4010, doi:10.1029/2003GB002097.

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Gnanadesikan, A., J.P. Dunne and R. Msadek (2014) Connecting Atlantic Temperature Variability and Biological Cycling in two Earth System Models, J. Mar. Sys.,133, 39-54.

Gnanadesikan, A., I. Koszalka, R. Abernathey and M.A. Pradal, (2018) Dispersion, diffusion and confusion: Lateral mixing in the ocean, rev. for Ann. Rev. Mar. Sci.

Johnson H.L., D.P. Marshall and D.P.J. Sproson (2007) Reconciling theories of a mechanically driven meridional overturning circulation with thermohaline forcing and multiple equilibria 29: 821-836. https://doi.org/10.1007/s00382-007-0262-9

Levermann, A., and J. J. Fürst (2010), Atlantic pycnocline theory scrutinized using a coupled climate model, *Geophys. Res. Lett.*, 37, L14602, doi: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. *Science Advances* 3.1.

Manabe S. and R.J. Stouffer, (1988) Two stable equilibria of a coupled ocean-atmosphere model, *J. Climate*, 1, 841-866.

Meijers A.J.S. (2014) The Southern Ocean in the Coupled Model Intercomparison Project phase 5. Phil. Trans. R. Soc. A 372: 20130296.http://dx.doi.org/10.1098/rsta.2013.0296.

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 climatic oscillation: vegetation change in Eastern North America during the Younger Dryas chronozone, *Quater. Science Reviews*, 1777-1791.

Stommel, H.M., 1961: Thermohaline circulation with two stables states of flow, *Tellus*, 2,224-229.

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Font: Italic

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Formatted: Font: (Default) Times New Roman

Weaver , A.J., J. Sedláček, M. Eby, K. Alexander, E. Crespin, T. Fichefet, G. Philippon-Berthier, F. Joos, M. Kawamiya, K. Matsumoto, M. Steinacher, K. Tachiiri, K. Tokos, M. Yoshimori, K. Zickfeld, 2012: Stability of the Atlantic meridional overturning circulation: A model intercomparison. *Geophys. Res. Lett*. 39, L20709.

S. G. Yeager, C. A. Shields, W. G. Large, J. J. Hack , The low-resolution CCSM3. *J. Climate*, 19, 2545–2566 (2006).

Zhang R and T. L. Delworth 2005 Simulated tropical response to a substantial weakening of the Atlantic thermohaline circulation, *J. Climate*, 18, 1853-1860.

Zhang, R., and T. L. Delworth, 2006: Impact of Atlantic multidecadal oscillations on India/Sahel rainfall and Atlantic hurricanes. *Geophys. Res. Lett*., 33, L17712, doi:10.1029/2006GL026267.

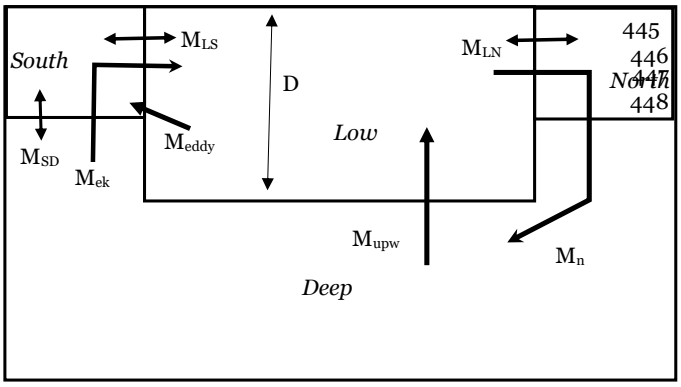
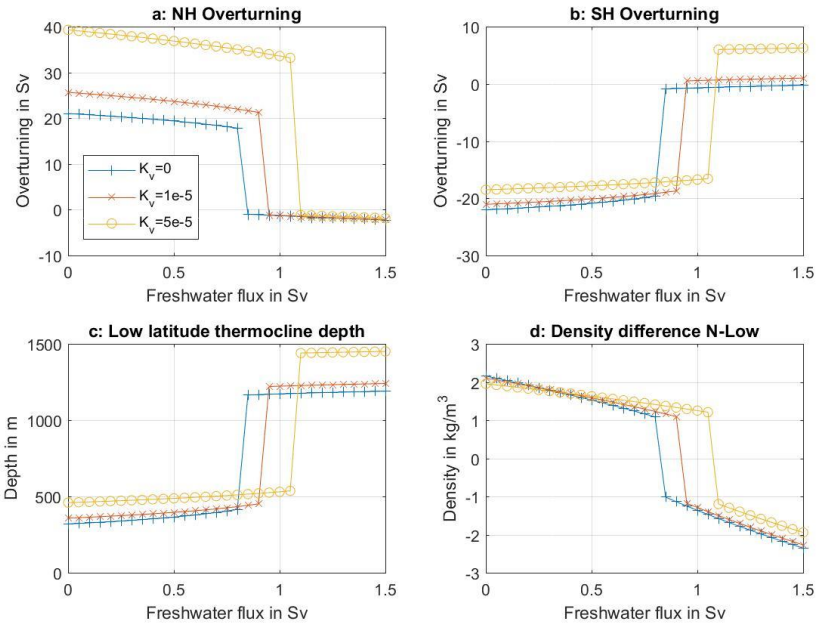


Figure 1: Schematic of the dynamical box model used in this paper. Italics denote the individual boxes. Arrows with a single point denote net advective fluxes. Double-headed arrows denote mixing fluxes.



458

459 **Figure 2:** Sensitivity of the model solution to the changing Northern Hemisphere freshwater flux
460 at different values of the vertical eddy mixing coefficient K_v that affects the flux M_{upw} in Fig. 1.
461 Blue lines with + symbols show the results for an K_v=0 m^2/s, red lines with x symbols show the
462 results for K_v=1x10^{-5} m^2/s, (default for the other simulations) and yellow lines with o symbols
463 show results for K_v=1x10^{-5} m^2/s. (A) NH overturning M_n. (B) SH overturning $(\frac{A_{GM}*D*L_x^S}{L_y^s} - \frac{\tau_s L_x^s}{\rho f_s})$,
464 positive values correspond to net transformation of light to dense water as for M_n. (C) Low-
465 latitude pycnocline depth. (D) Density difference between high-latitude NH surface box and low-
466 latitude surface box.

467

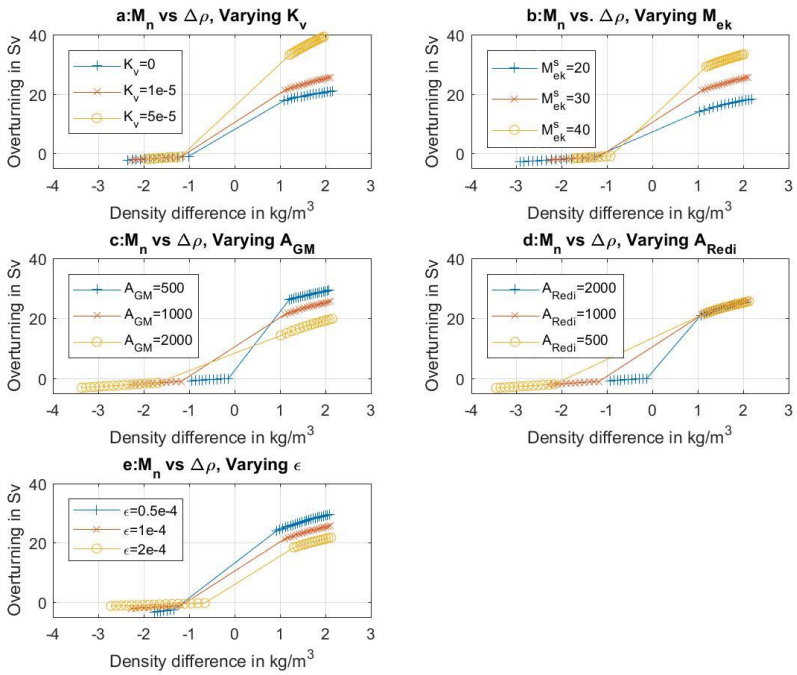
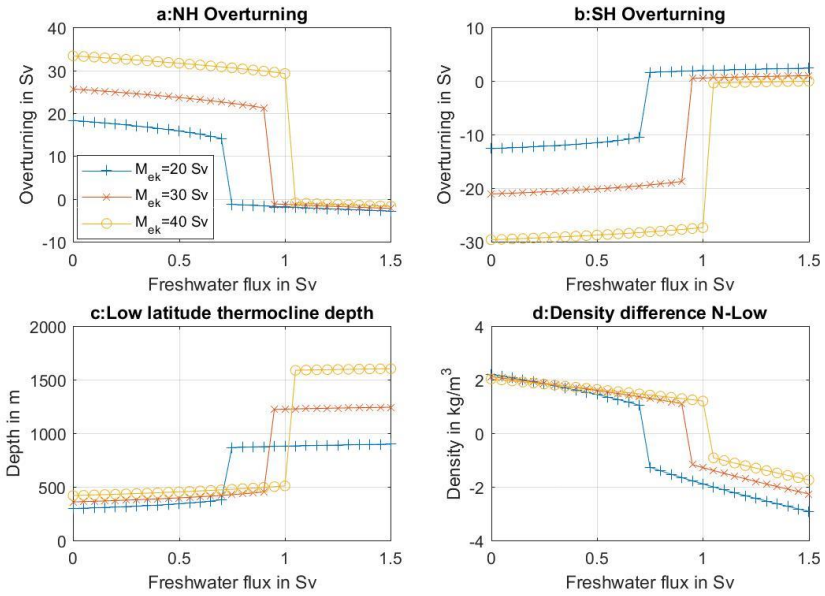


Figure 3: Relationship between density gradient and NH overturning across the different model suites. (a) Effect of varying vertical diffusion coefficient K_v. (b) Effect of varying Southern Ocean Ekman flux M_{ek}. (c) Effect of varying Gent-McWilliams coefficient affecting Southern Ocean eddy flux M_{eddy}. (d) Effect of varying lateral diffusion coefficient A_{Redi} affecting mixing fluxes M_{LN} and M_{LS}. (e) Effect of varying the resistance coefficient ϵ affecting the proportionality between depth integrated pressure gradient and NH overturning M_n.



476

477 **Figure 4:** Sensitivity of the model solution to the changing Northern Hemisphere freshwater flux
478 at different values of Southern Hemisphere Ekman flux. Blue lines with + symbols show the
479 results for an Ekman flux of 20 Sv, red lines with x symbols show the results for an Ekman flux
480 of 30 Sv (default for the other simulations) and yellow lines with o symbols show results for an
481 Ekman flux of 40 Sv. (A) NH overturning M_n. (B) SH overturning ($\frac{A_{GM}*D*L_X^S}{L_y^S} - \frac{\tau_s L_x^S}{\rho f_s}$), positive
482 values correspond to net transformation of light to dense water as for M_n. (C) Low-latitude
483 pycnocline depth. (D) Density difference between high-latitude NH surface box and low-latitude
484 surface box.

485

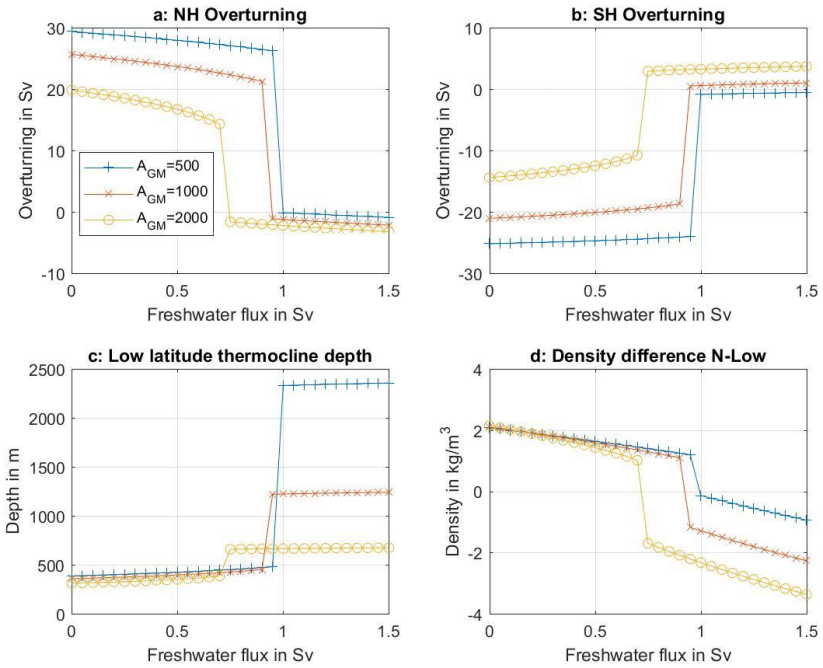
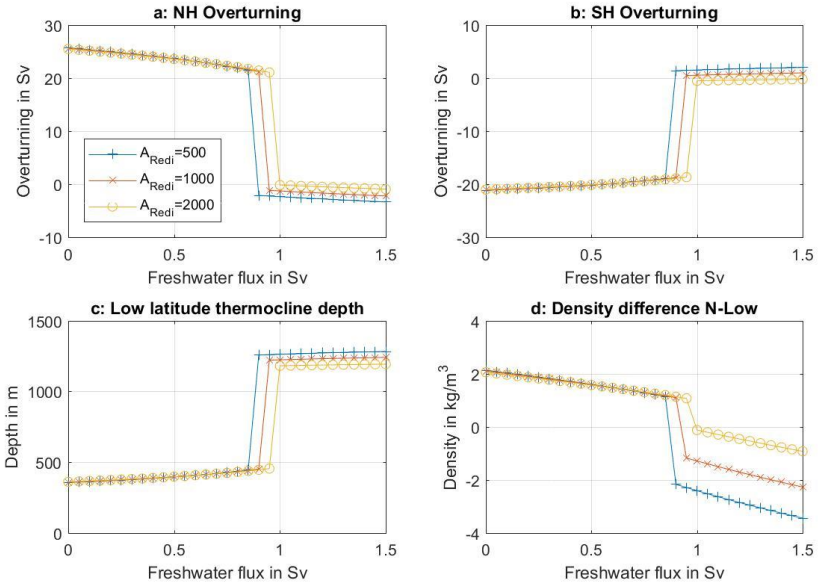


Figure 5: Sensitivity of the model solution to the changing Northern Hemisphere freshwater flux at different values of the Gent-McWilliams eddy mixing coefficient A_{GM} that affects the flux M_{eddy} in Fig. 1. Blue lines with + symbols show the results for an A_{GM}=500 m^2/s,- red lines with x symbols show the results for A_{GM}=1000 m^2/s,- (default for the other simulations) and yellow lines with o symbols show results for A_{GM}=2000 m^2/s. (A) NH overturning M_n. (B) SH overturning ($\frac{A_{GM}*D*L_x^S}{L_y^S} - \frac{\tau_s L_x^S}{\rho f_s}$), positive values correspond to net transformation of light to dense water as for M_n. (C) Low-latitude pycnocline depth. (D) Density difference between high-latitude NH surface box and low-latitude surface box.



496

497 **Figure 6:** Sensitivity of the model solution to the changing Northern Hemisphere freshwater flux
498 at different values of the lateral eddy mixing coefficient A_{Redi} that affects the flux $M_{LS,LN}$ in Fig.
499 1. Blue lines with + symbols show the results for an A_{Redi}=500 m^2/s,- red lines with x symbols
500 show the results for A_{Redi}=1000 m^2/s,- (default for the other simulations) and yellow lines with o
501 symbols show results for A_{Redi}=2000 m^2/s. (A) NH overturning M_n. (B) SH overturning
502 $(\frac{A_{GM}*D*L_x^S}{L_y^S} - \frac{\tau_S L_x^S}{\rho f_S})$, positive values correspond to net transformation of light to dense water as for
503 M_n. (C) Low-latitude pycnocline depth. (D) Density difference between high-latitude NH surface
504 box and low-latitude surface box.

505

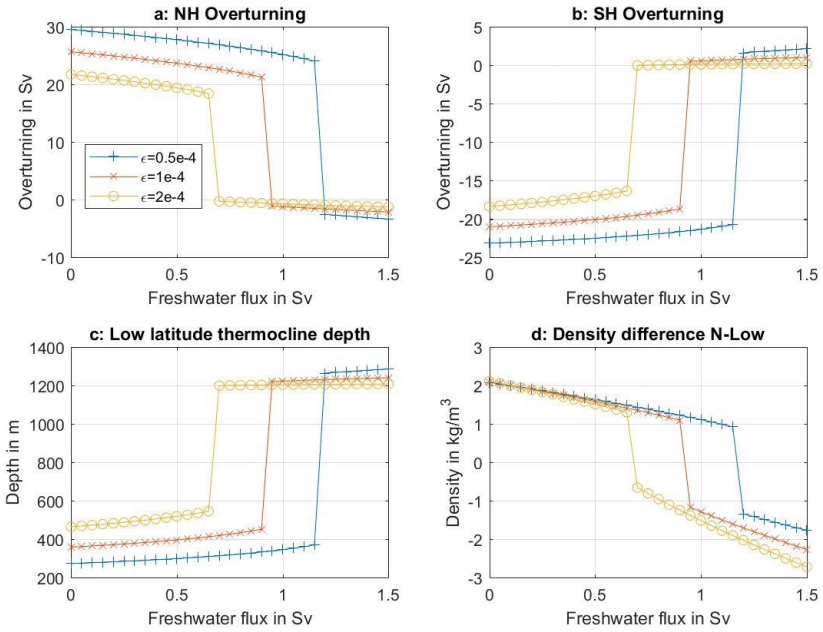


Figure 7: Sensitivity of the model solution to the changing Northern Hemisphere freshwater flux at different values of scaling coefficient ϵ which governs the conversion between the large-scale pressure gradients and the overturning. Blue lines with + symbols show the results for ϵ=5 x 10^{-5} s^{-1}, red lines with x symbols show the results for ϵ=1 x 10^{-4} s^{-1} (default for the other simulations) and yellow lines with o symbols show results for ϵ=1 x 10^{-4} s^{-1}. (A) NH overturning M_n. (B) SH overturning ($\frac{A_{GM} * D * L_x^S}{L_y^S} - \frac{\tau_s L_x^S}{\rho f_s}$), positive values correspond to net transformation of light to dense water as for M_n. (C) Low-latitude pycnocline depth. (D) Density difference between high-latitude NH surface box and low-latitude surface box.