The stability of the MOC as diagnosed from model projections for pre-industrial, present and future climates

**Sybren S. Drijfhout · Susanne L. Weber ·
Eric van der Swaluw**

Abstract The stability of the Atlantic meridional overturning circulation (MOC) is investigated for various climate scenario runs, using data from the CMIP3 archive of coupled atmosphere-ocean models. Apart from atmospheric feedbacks, the sign of the salt flux into the Atlantic basin that is carried by the MOC determines whether the MOC is in the single or multiple equilibria regime. This salt advection feedback is analyzed by diagnosing the freshwater and salt budgets for the combined Atlantic and Arctic basins. Consistent with the finding that almost all coupled climate models recover from hosing experiments, it is found that most models feature a negative salt advection feedback in their pre-industrial climate: freshwater perturbations are damped by this feedback, excluding the existence of a stable off-state for the MOC. All models feature enhanced evaporation over the Atlantic basin in future climates, but for a moderate increase in radiative forcing (B1 and 2 CO_2 scenarios), there is a decrease of the fresh water flux carried by the MOC into the Atlantic (the deficit is made up by increased fresh water transport by the gyre circulation). In this forcing regime the salt advection feedback becomes less negative: for three models from an ensemble of eight it is positive in a 2 CO_2 climate, while two models feature a positive feedback in the pre-industrial climate. For even warmer climates (A1B-equilibrium and 4 CO_2) the salt feedback becomes more negative (damping) again. It is shown that the decrease in northward fresh water transport at 34°S by the MOC (in B1-equilibrium and 2 CO_2) is due to a reduction of the inflow of intermediate waters relative to thermocline waters, associated with a robust shoaling of the MOC in future, warmer climates. In A1B and 4 CO_2 climates northward freshwater transport increases again. The MOC keeps shoaling, but both intermediate and thermocline water masses freshen.

1 Introduction

The Atlantic MOC affects the climate by its northward heat transport; about 1.5 PW at 25°N (Trenberth and Caron 2001). A shut-down of the MOC would have large climate impacts, with pronounced cooling over the North Atlantic and warming in the Southern Hemisphere (Vellinga and Wood 2002; Drijfhout 2010). Such a collapse is associated with the hypothesis that the MOC possesses multiple equilibria, which have been shown to exist in simple box models (Stommel 1961); Earth System Models of intermediate complexity (Rahmstorf et al. 2005); and state-of-the art coupled climate models (CCMs; Manabe and Stouffer 1988). This lead Broecker et al. (1985) to hypothesize that abrupt changes in the MOC are the main cause for the observed abrupt changes in past climates. Paleoclimatic proxy data of the last glacial cycle indeed show evidence of abrupt climate changes that are associated with major reorganizations of the Atlantic MOC (Clark et al. 2002). As a result, transitions between different states of the MOC are now commonly invoked to explain abrupt climate change. The absence of rapid climate changes after the Younger Dryas suggests that the Atlantic MOC has resided in a more stable regime over the last 10,000 years (Grootes et al. 1993). However, an intensified hydrological cycle, due to global warming, might

S. S. Drijfhout (✉) · S. L. Weber · E. van der Swaluw
Royal Netherlands Meteorological Institute (KNMI),
PO Box 210, 3730 AE De Bilt, The Netherlands
e-mail: drijfhou@knmi.nl

destabilize the MOC again (Broecker 1997; Rahmstorf 2000).

In simplified theoretical models, the existence of multiple equilibria for the MOC can be derived when a given set of atmospheric boundary conditions is assumed. It is not immediately clear that the multiple equilibria carry over to the coupled system, in which such boundary conditions become part of the solution. In almost all modern CCMs, only one equilibrium state for the MOC seems to prevail (Stouffer et al. 2006), putting the existence of multiple equilibria for the MOC in the coupled system into question. On the other hand, in Rahmstorf et al. (2005) all CCMs did show hysteresis behavior. Although the models analyzed in that paper were of intermediate complexity, some of the models showed a similar response of the Atlantic MOC to greenhouse gas forcing and freshwater hosing as more complete CCMs, suggesting that the later category of models should also possess hysteresis behavior. Furthermore, it should be realized that multiple equilibria do not exist over the whole regime of climate states. Hysteresis experiments (Rahmstorf 1996) and steady state calculations with continuation techniques (Dijkstra 2007) indicate that over large ranges in parameter space only one equilibrium exists, while the multiple equilibrium regime is associated with a more restricted range of parameter values.

It was argued by Rahmstorf (1996) that the transition between the single and multiple equilibria regime of the MOC is determined by the sign of the atmospheric freshwater transport across the southern border of the "Atlantic box". But the argument he used, involving a large-scale salt feedback, was indirect. It turned out to be more appropriate to use the salt flux divergence into the Atlantic carried by MOC itself as a control parameter, which in Rahmstorf's conceptual model was proportional to the atmospheric fresh water flux. This idea was validated in a coupled model of intermediate complexity by de Vries and Weber (2005) and Weber and Drijfhout (2007). Moreover, Dijkstra (2007) was able to show with continuation techniques that the saddle node bifurcation that marks the boundary between the single and multiple equilibria regime, occurs exactly where the salt transport by the overturning circulation into the Atlantic becomes zero. A further theoretical underpinning for this diagnostic was given by Huisman et al. (2010).

In these models salt export from the Atlantic by the MOC puts the MOC in the single equilibrium regime, whereas salt import puts it into the multiple equilibria regime. In the former case, a weakening of the MOC in response to a freshwater anomaly in the northern North Atlantic is followed by a gradual salinification of the basin, which eventually leads to the restart of deep convection (Weber and Drijfhout 2007). In the latter case, the basin remains capped by a freshwater halocline, preventing convection to restart. This large-scale salt advection feedback fully determines the stability of the MOC in ocean-only models. Atmospheric feedbacks may also have an effect on the stability of the MOC, but the sign of the atmospheric feedback is not evident. Without dynamical changes, the atmospheric response to colder SSTs is a weakened hydrological cycle and reduced net evaporation. This is a positive feedback for a MOC collapse, promoting the multiple equilibria regime. In addition, the dynamical atmospheric response consists of a southward shift of the intertropical convergence zone (ITCZ) and Hadley circulations (Vellinga and Wood 2002; Drijfhout 2010). This affects the efficiency of moisture transport across Central America. Reconstructions from proxy data point to a positive feedback, as westward moisture transport shifts to the south where it is blocked by the Andes (Leduc et al. 2007).

This response, however, is not recovered in many CCMs. In those models, the moisture transport from the Atlantic increases by enhanced evaporation over the North Atlantic subtropics (Lohmann 2003; Lenton et al. 2007; Yin and Stouffer 2007), which results in a negative feedback. Also, the response of ENSO to a collapse of the MOC may lead to evaporation changes over the Atlantic that consist of a negative feedback (Latif et al. 2000). On the whole, the role of atmospheric feedbacks on the stability of the MOC is still unclear. This was illustrated by Yin and Stouffer (2007), who compared the response to freshwater hosing in two GFDL versions; R30 and CM2.1. In CM2.1, the MOC recovers from hosing and Yin and Stouffer argue that the atmospheric feedback discussed above is essential in this recovery. In R30, the ITCZ response is much weaker and the associated atmospheric feedback is small. As a result, the MOC remains collapsed. But Yin and Stouffer (2007) also calculated the sign of the large-scale salt feedback in the two models, and for R30 it was positive, while for CM2.1 it was negative. Yin and Stouffer attribute the different stability of the two collapsed states completely to different atmospheric feedbacks. But also without any atmospheric feedback, in R30 the collapsed state would be stable, while in CM2.1 it would be unstable due to the sign of the ocean feedback. The role of atmospheric feedbacks in the recovery from hosing in CM2.1 therefore remains undetermined.

Given the present level of understanding it is here assumed that the large-scale salt feedback being positive is a necessary condition for the existence of multiple equilibria for the MOC. Without atmospheric feedbacks it is a sufficient condition for a multistable MOC. If atmospheric feedbacks tend to counteract the ocean's salt feedback, it is no longer a sufficient condition, but still a necessary condition, depending on how strong the atmospheric feedbacks

in reality are. In any case it is important to diagnose the sign and strength of the ocean feedback in climate models, because it gives a first indication of whether the simulated climate equilibria possess a multistable MOC, or not.

2 Method

The connection between the large-scale advective salt feedback and atmospheric freshwater forcing in a CCM was formulated by de Vries and Weber (2005), but a rigorous derivation was not given there. In this section, a derivation is presented of the *equivalent* freshwater budget, which follows from a combination of conservation of mass (volume) of water and of salt. For a Boussinesq ocean the vertically-integrated equation for continuity reads (Griffies et al. 2004):

$$\eta_{,t} = -\int_{-H}^{\eta} dz \nabla \cdot \mathbf{u} + q_w + \text{Res}_v \tag{1}$$

where H is the depth of the ocean, η is the free surface elevation, u is the horizontal velocity vector, q_w is the volume flux per unit time per unit horizontal area of fresh water crossing the sea surface. Res_v is a rest term. It is anticipated that the mass (volume) budget will be diagnosed in a discretised model, where interpolation errors may arise. Also, mass (water volume) is sometimes exchanged over land between adjacent seas. We note that for rigid-lid models ($\eta = 0$) this flux is zero. The vertically-integrated equation for salinity S reads

$$\int_{-H}^{\eta} dz S_{,t} + \eta_{,t}(H + \eta)^{-1} \int_{-H}^{\eta} dz S$$
$$= \int_{-H}^{\eta} dz \nabla \cdot (S\mathbf{u}) - \int_{-H}^{\eta} dz \nabla \cdot \mathbf{F} + \mathbf{S} + \text{Res}_s \tag{2}$$

where $\mathbf{F}$ is the flux from small-scale mixing; $\mathbf{S}$ is a source term (virtual salt flux) only for rigid-lid models, and Res_s is a rest term. Now, Eq. 1 is integrated over the entire Atlantic and Arctic from 34°S to Bering Strait.

$$V_{,t} = \int_{\text{BS}} v dx dz + \int_{34\text{S}} v dx dz - \text{EPR} + \text{Res}_V \tag{3}$$

The subscripts BS and 34S indicate integration over a zonal transect in the Bering Strait and near 34°S, at the latitude where the African continent terminates. Note that at Bering Strait, in- and outflow have the opposite sign of in- and outflow at 34°S, which explains the plus signs for both advective terms in Eq. 3. EPR is the net evaporation over

the basin and V is water volume. All terms in Eq. 3 can be expressed in Sverdrups ($Sv = 10^6$ m^3 s^{-1}). We abbreviate this equation to:

$$V_{,t} = T_{\text{BS}} + T_{34\text{S}} - \text{EPR} + \text{Res}_V \tag{4}$$

Next, Eq. 2 is integrated over the Atlantic and Arctic basin.

$$\int_{\text{Atl}} S_{,t} dV + V_{,t} S_{\text{Atl}} = \int_{\text{BS}} vS dx dz + \int_{34\text{S}} vS dx dz + \text{MIX}$$
$$+ \text{SOURCE} + \text{Res}_S \tag{5}$$

In this equation MIX denotes the contribution by small scale mixing. It may consist of a term due to isopycnal (horizontal) mixing and a term due to advection by eddies following Gent and McWilliams (1990). This term can not be diagnosed directly from model output archives. It will be treated as a rest term. SOURCE is only nonzero for rigid-lid models where EPR = 0 in Eq. 3. It contains the virtual salt flux proportional to EPR in such models. S_{Atl} is the average salinity over the Atlantic/Arctic basin.

The salt budget in Eq. 5 can be transformed into a freshwater volume budget, similar to Eq. 3, by dividing Eq. 5 by a reference salinity S_0:

$$S_0 = \int_{34\text{S}} S dx dz \bigg/ \int_{34\text{S}} dx dz \tag{6}$$

Then the salt budget becomes an *equivalent* freshwater budget in which all terms can be expressed in Sverdrups:

$$-\frac{1}{S_0} \int_{\text{Atl}} S_{,t} dV - V_{,t} \frac{S_{\text{Atl}}}{S_0} = -\frac{1}{S_0} \int_{\text{BS}} vS dx dz$$
$$-\frac{1}{S_0} \int_{34\text{S}} vS dx dz + \text{MIX}_{S_0} + \text{SOURCE}_{S_0} + \text{Res}_{S_0} \tag{7}$$

The final equivalent freshwater budget results from adding Eq. 4–7. But before the addition is made, it is instructive to rewrite Eq. 7. First, it is noted that $\text{Res}_S = S_{\text{cl}} \text{Res}_V$, where S_{cl} is the weighted average salinity of the cross land mixed water. Then

$$\text{Res}_{S_0} = -\frac{S_{cl}}{S_0} \text{Res}_V = -\left(1 + \frac{S_{cl} - S_0}{S_0}\right) \text{Res}_V$$
$$= -(1 + \epsilon_1) \text{Res}_V \tag{8}$$

Also, SOURCE = $S_{\text{ref}} \text{EPR}^*$, where EPR* is the equivalent water flux through the ocean's surface in rigid lid models. In most models S_{ref} is chosen to be 35.0 psu. Then

$$SOURCE_{S_0} = -\frac{S_{\text{ref}}}{S_0} \text{EPR}* = -\left(1 + \frac{S_{\text{ref}} - S_0}{S_0}\right) \text{EPR}*$$
$$= -(1 + \epsilon_2) \text{EPR} * \tag{9}$$

Finally,

$$-\frac{S_{Atl}}{S_0}V_{,t} = -\left(1+\frac{S_{Atl}-S_0}{S_0}\right)V_{,t} = -(1+\epsilon_3)V_{,t} \quad (10)$$

Now, ϵ_1, ϵ_2, $\epsilon_3 = O(0.01)$. Equation 7 can be simplified by grouping these small terms together and interprete them as mixing. They are added to the larger term MIX_{S_0}:

$$M_{mix} = MIX_{S_0} + \epsilon_3 V_{,t} - \epsilon_1 Res_V - \epsilon_2 EPR* \quad (11)$$

Then we can write Eq. 7 as

$$-M_{trend} - V_{,t} = M_{BS} + M_{34S} + M_{mix} - Res_V - EPR* \quad (12)$$

Here, M_{BS} and M_{34S} are the first two terms in the RHS of Eq. 7, respectively, and

$$M_{trend} = \frac{1}{S_0}\int_{Atl} S_{,t}dV \quad (13)$$

Comparing Eqs. 4 and 12, it becomes evident that for a free surface model (EPR* = 0) net evaporation causes the barotropic mass transport across Bering Strait and 34°S to differ, while salt conservation demands that the salt transport across the two sections is (nearly) equal. In other words, salt transport by the baroclinic circulation across 34°S is directly associated with the net evaporation over the Atlantic basin. This conclusion also holds for rigid-lid models. By rewriting the salt transport as an equivalent freshwater transport, the freshwater loss by evaporation has to be supplied by a net equivalent freshwater transport across 34°S by the baroclinic circulation. This relation follows from combining Eqs. 4 and 12. Then we obtain for both free surface and rigid-lid models (dropping the * from EPR*)

$$EPR = M_{34S} + T_{34S} + M_{\Delta BS} + M_{mix} + M_{trend} \quad (14)$$

where

$$M_{\Delta BS} = \int_{BS} vdxdz - \frac{1}{S_0}\int_{BS} vSdxdz \quad (15)$$

The terms in Eq. 15 are dominated by barotropic transports. As a result, $M_{\Delta BS}$ can be interpreted as the salinification of the Bering Strait inflow along its trajectory through the Arctic and Atlantic ocean.

The baroclinic transport across 34°S results from the sum of M_{34S} and T_{34S}. This becomes clear when M_{34S} is decomposed. First, the zonally averaged barotropic and baroclinic velocities are split. The barotropic part becomes

$$\widetilde{v} = \int_{34S} vdxdz \bigg/ \int_{34S} dxdz \quad (16)$$

and the baroclinic part

$$v^* = \langle v \rangle - \widetilde{v}, \quad (17)$$

where, for a generic function f, $\langle f \rangle = \int f\,dx/\int dx$ is the zonal mean, and $f' = f - \langle f \rangle$

M_{34S} in Eq. 14 is then split in three parts (for convenience we drop the subscript 34S from the integrals):

$$M_{34S} = -\frac{1}{S_0}\int vSdxdz = -\frac{1}{S_0}\int(\langle v \rangle + v')(\langle S \rangle + S')dxdz$$
$$= -\frac{1}{S_0}\left(\int \widetilde{v}\langle S \rangle dxdz + \int v^*\langle S \rangle dxdz + \int v'S'dxdz\right) \quad (18)$$

$\widetilde{v}$ is constant with respect to both x and z, and v^* and $\langle S \rangle$ are constant with respect to x. Equation 18 is rewritten as

$$M_{34S} = -\frac{\widetilde{v}}{S_0}\int \langle S \rangle dxdz - \frac{1}{S_0}\int \overline{v^*\langle S \rangle}dz - \frac{1}{S_0}\int \overline{v'S'}dz \quad (19)$$

where the overbar denotes zonal integration. Now it becomes clear why S_0 was chosen as a reference salinity. Namely,

$$-\frac{\widetilde{v}}{S_0}\int \langle S \rangle dxdz = -\frac{\widetilde{v}}{S_0}\int Sdxdz = -\frac{\widetilde{v}}{S_0}S_0\int dxdz$$
$$= -\widetilde{v}\int dxdz = -\int vdxdz = -T_{34S} \quad (20)$$

As a result, T_{34S} can be eliminated from Eq. 14, which is favorable because the barotropic transport term is badly constrained when interpolated data are used.

We make two further definitions:

$$M_{ov} = -\frac{1}{S_0}\int_{34S} \overline{v^*}\langle S \rangle dz \quad (21)$$

and

$$M_{az} = -\frac{1}{S_0}\int_{34S} \overline{v'S'}dz \quad (22)$$

Then, Eq. 19 can be written as

$$M_{34S} = -T_{34S} + M_{ov} + M_{az} \quad (23)$$

Part of the Atlantic MOC is associated with North Atlantic Deep Water and its lighter return flow, part with Antarctic Bottom Water and its deep return flow. Only the first part is relevant for the stability of the Atlantic MOC. So we split M_{ov} in a part associated with deep water formation in the Atlantic, M_{na}; and a part associated with deep water formation in the Southern Ocean, M_{so}. The boundary between the two is where the MOC becomes zero. After inserting Eq. 23 into Eq. (14) and splitting M_{ov} we obtain

$$EPR = M_{na} + M_{so} + M_{az} + M_{mix} + M_{\Delta BS} + M_{trend} \qquad (24)$$

Equation 24 is the final budget that we diagnose to evaluate M_{na}. This equation constitutes the large-scale salt advection feedback that is crucial in assessing the stability of the MOC. The budget expresses that net evaporation over the Atlantic basin drives salt export/freshwater import by the internal baroclinic circulation through its three components: overturning circulation, horizontal gyre circulation, and small-scale mixing. It also causes a salinification of the Bering Strait transport flowing through the Atlantic, and in case of imbalance, a salinity trend.

3 An illustrative example

Before a diagnosis of the data from the CMIP3 archive is presented, we illustrate the various budgets from Sect. 2 with GFDL CM2.1 (Delworth et al. 2006). In addition to the interpolated data, also the original data on the GFDL model's tripolar grid have been made available on their own website. This allows us to (almost) close the budgets discussed in Sect. 2.

First, the volume budget of Eq. 4 is discussed. We take the data from the pre-industrial control run (PIC), because this run contains the least drift, as the spin-up of the model is performed under PIC conditions. We start with defining a mask that displays the area of integration. For GFDL2.1 the mask is displayed in Fig. 1. It is evident that in the model the Hudson Bay, Baltic Sea, Mediterranean Sea and Black Sea are closed basins. Hence, cross land mixing occurs. Whether these basins are closed or open varies from model to model, and for each model the mask has to be constructed carefully, comparing the grids for velocity and tracer points. Also, the mask in the Arctic is quite peculiar, associated with the tripolar grid. Other models do not have
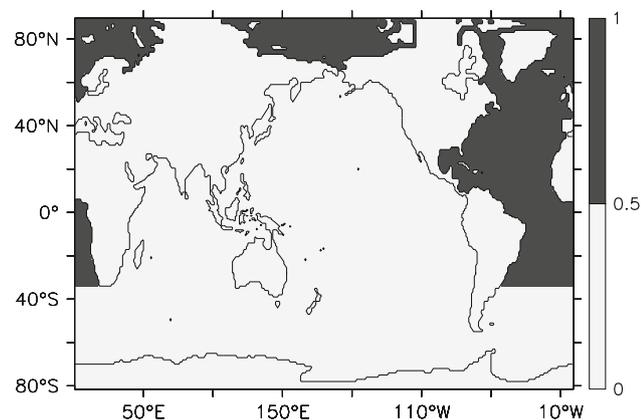
this feature. Volume changes are calculated from changes in sea-ice thickness and sea surface height; the fields *sit*, *sic* and *zos*. EPR is calculated from the *wfo* fields, dividing this number by a reference density of 1,035 kg m^{-3} (Griffies et al. 2004). The volume transport is estimated from the meridional velocity field *vo* at the relevant sections.

The resulting volume budget for GFDL2.1 is shown in Fig. 2a. Volume changes are negligible, <1 mSv. Averaged over periods of 30 years, the net exchange between water and sea-ice is small, although it is a dominant term over monthly time scales and shorter. The first order balance implies that about 60% over the Bering Strait inflow evaporates along its course through the Atlantic, so that the outflow at 34°S is only about 40% of the inflow. There is a small residue, associated with the contribution from cross-land mixing, of 10–20 mSv; 2% of the Bering Strait inflow.

The budget of the equivalent freshwater transported (scaled salt transport) follows from Eq. 12 and is shown in Fig. 2b. In GFDL2.1 velocity and tracer points are arranged according to a B-grid configuration (Griffies et al. 2004), and the same holds for the majority of the interpolated CMIP data. For consistency, the salt, or equivalent freshwater transports have to be calculated on velocity points with the correct discretization applied. Salinities are contained in the *so* fields. The equivalent freshwater transport across 34° and Bering Strait are almost equal. The equivalent freshwater transport is northward, which means that the salt transport is southward. There is salt import through Bering Strait and salt export across 34°S. The Bering Strait equivalent freshwater transport is slightly lower than its volume transport, reflecting the lower salinity at Bering Strait compared to 34°S (Eq. 7). But because at Bering Strait all the transport is accomplished by the barotropic flow, the magnitude of the equivalent freshwater and volume transport at Bering Strait is almost
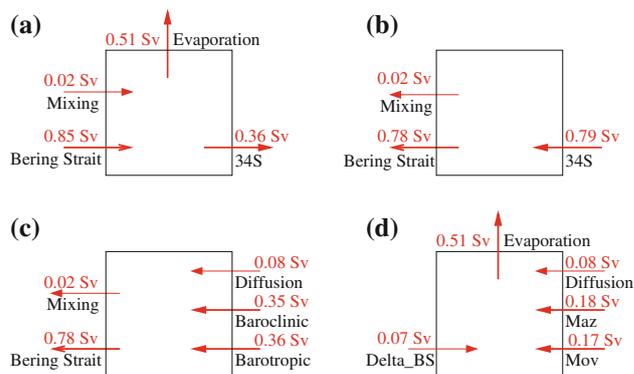

Fig. 1 The mask used to define the integration area of the combined Atlantic and Arctic Oceans in the GFDL models. The *black* contour denotes the land-sea mask of the model

Fig. 2 The volume and various (equivalent) freshwater budgets discussed in Sect. 3 for GFDL2.1. **a** The budget of Eq. 4, **b** and **c** the budget of Eq. 12. In the figure, "Mixing" corresponds to the Res terms in Eqs. 4 and 12. **d** The budget of Eq. 24, with M_{na} and M_{so} taken together as M_{ov}. In **c** and **d** the term "Diffusion" in the figure corresponds to M_{mix} in Eqs. 12 and 24

similar in magnitude, but with opposite sign. At 34° this is clearly not the case; the equivalent freshwater transport is much larger in magnitude than the volume transport. This is because the internal, baroclinic flow here accomplishes a significant amount of salt (equivalent freshwater) transport. The scaled trend in salinity is relatively small (<10 mSv), but much larger than the trend in volume. The equivalent freshwater trend is <1% of the Bering Strait volume flux.

The decomposition of the equivalent freshwater transport at 34°S (Eq. 23) is shown in Fig. 2c. The advective transport is for 50% accomplished by the barotropic flow, and for 50% by the baroclinic flow. Diffusive transport is about 10% of the advective transport.

The final balance (Eq. 24) is given in Fig. 2d. The largest terms that compensate the freshwater loss by net evaporation are the advective contributions by the baroclinic flow. Diffusive transport and salinification of the Bering Strait transport are about half as large as the advective terms. The equivalent freshwater trend is small, but non negligible. The overturning circulation exports salt and imports equivalent freshwater. As a result, the large-scale salt advection feedback is negative. It damps the effect of freshwater perturbations and indicates that the MOC is in the mono-stable regime. This is consistent with the analysis of Yin and Stouffer (2007). The model is apparently in almost equilibrium, but internal variability still might affect the estimates of Eq. 24.

To estimate the effect of internal variability we have calculated the budgets for three successive periods of 30 years, together spanning the last 90 years of the PIC run (year 410–500). The associated error (random error) is quite modest, defined as 50% of the difference between minimum and maximum values of each sample of three, see Table 1.

In addition, the interpolation error, relevant when diagnosing output from the CMIP archive, was estimated. Ideally, the interpolation error would have been estimated by comparing the budgets calculated on the original and interpolated grids respectively, and taking the difference. But not all the data that is needed to calculate the budget on the interpolated grid was available. As an alternative, we compared exact integrations, taking the topography of the tripolar grid into account, to integrations with standard Ferret routines, which also were used to perform the analysis of the interpolated data. More importantly, in the Ferret routines we used a bottom topography defined on velocity points after interpolating the original topography file which is defined on tracer points. Also in the CMIP archive the topography file is interpolated on velocity points. The interpolation error is then defined as the difference between those two calculations.

The interpolation errors are given in Table 1. It is evident that the largest error is associated with T_{34S}. Fortunately, this term is eliminated from the final balance. Errors

Table 1 Error estimate in equivalent freshwater budget

Quantity	Mean value	Random error	Interpolation	Total error
$V_{,t}$	0.000	0.000	0.01	0.01
T_{BS}	0.85	0.02	0.003	0.02
T_{34S}	−0.36	0.02	1.0	1.0
$-EPR$	−0.507	0.004	0.05	0.05
Res_V	0.015	0.002	0.01	0.01
$Q_{,t}$	−0.008	0.005	0.001	0.005
M_{BS}	−0.78	0.02	0.003	0.02
M_{34S}	0.71	0.02	1.0	1.0
M_{mix}	0.08	0.01	0.04	0.04
$-Res_V$	0.015	0.002	0.01	0.01
M_{34S}	0.71	0.02	1.0	1.0
M_{ov}	0.176	0.006	0.005	0.008
M_{az}	0.180	0.003	0.003	0.004
$-T_{34s}$	0.36	0.02	1.0	1.0
EPR	0.51	0.004	0.05	0.05
M_{na}	0.165	0.006	0.005	0.008
M_{so}	0.011	0.001	0.002	0.002
M_{az}	0.180	0.003	0.003	0.004
M_{mix}	0.08	0.01	0.04	0.04
$M_{\Delta BS}$	0.067	0.003	0.001	0.003
M_{trend}	0.008	0.005	0.001	0.005

The random error is defined as 50% of the difference between minimum and maximum values of each sample of three 30-year runs; the interpolation error is defined as the difference between exact calculations on the model grid and calculations with bottom topography interpolated on velocity points. The total error is the sum of random and interpolation errors

in estimating the barotropic transport at Bering Strait are much smaller than at 34°S. This is so, because the bottom is shallow there, and relatively flat. Also, the flow at Bering Strait is almost unidirectional. There is hardly any overturning or gyre component. This reduces the error in calculating net zonally-averaged barotropic transport at interpolated data. Moreover, the error in the final budget is further reduced, because volume and equivalent freshwater transport at Bering Strait nearly cancel and the error in the residual is smaller than in its individual components, since these errors are dependent. In the final budget of Eq. 24, the largest uncertainties are associated with estimates of EPR and M_{mix}. The estimate of M_{na} appears to be well constrained. Note that for GFDL2.1, the total error equals the random error due to internal variations, because the interpolation error does not apply to this model.

4 Multi-model analysis

To diagnose the equivalent freshwater budget from the CMIP3 archive the fields *so*, *vo*, *wfo*, *sic*, *sit*, and *zso* are

needed, as well as a file of the bottom topography. This demand already excludes more than half of the archived model runs. In addition, it was demanded that the runs are close enough to equilibrium, that is, Eq. 24 should not be dominated by M_{trend}. It was therefore demanded that the absolute value of M_{trend} is <0.05 Sv, which is 6 times larger than the values in GFDL2.1 PIC, but typically <10% of the EPR forcing. This implies that the only runs considered here are PIC; the Present Day Control or Committed Climate Change run (PDC/CCC); the $A1B$ scenario run after it has equilibrated to the forcing kept constant at year-2100 values; the same for the $B1$ scenario run; the two times CO_2 run in its equilibration phase; and the four times CO_2 run in its equilibration phase. Needless to say that not all model centers have run these scenarios to their equilibrium phase, or provided data of the last part of these runs. And even then, the drift was sometimes still unacceptably large. A second demand was that the global EPR is in approximate balance, that is, the zonal and meridional integral of EPR should be zero. In some models it is a few Sverdrups out of balance, making an estimate of the net evaporation over the Atlantic basin pointless. To exclude such runs, it was demanded that the global integral of EPR should be <0.05 Sv. In some cases, velocity and tracer grids were completely inconsistent with respect to land/sea mask and ocean floor depth, making an estimate of the salt advection impossible. These models were also discarded, but a few model runs remained available for diagnosis and they are summarized in Table 2.

Figure 3 shows the different terms of the equivalent freshwater budget of Eq. 24 for PIC. A similar analysis was also performed by Weber et al. (2007) for PMIP model-data, but the equivalent freshwater budget was less extensively analyzed in that study. Apart from CGCM3.1, all CCMs feature freshwater import by the overturning

circulation. In MIROC the total MOC imports freshwater, but the part associated with North Atlantic deep water formation exports freshwater. This means that in the other models freshwater perturbations that decrease deep water formation in the North Atlantic are damped by the salt advection feedback. These models are in the monostable regime and will always recover from hosing; the off-state is unstable. A positive M_{ov} was also a feature of most models diagnosed in Weber et al. (2007) (In that study M_{ov} was not split). Weber et al. (2007) also shows the PIC budget for HadCM3 and MIROC in their Fig. 1. The budget estimates are consistent with the ones shown in Fig. 3, but there are small differences. First, different timeslices have been chosen for averaging (often the same run is extended for a longer period in the CMIP archive), and also the diagnostic routines sometimes were updated, for instance, M_{ov} is now calculated on velocity-points, while in Weber et al. (2007) it was calculated on tracer points.

Two models from the present ensemble were also involved in the CMIP/PMIP hosing experiment with a 100 year 1.0 Sv perturbation (Stouffer et al. 2006), namely HadCM3 and GFDL2.1. Consistent with our analysis on the sign of the salt advection feedback, these models did recover from hosing. The present analysis suggests however, that CGCM3.1 probably has a stable off-state in PIC. But CGCM3.1 is also the model with the weakest MOC. Also MIROC may have a stable off-state, but the salt advection feedback in this model is quite weak and the existence of the off-state in this model will strongly depend on the atmospheric feedbacks, and possibly also to the response of Antarctic Bottom Water formation to a decrease in North Atlantic Deep Water formation.

The models disagree on the sign of M_{na} and the drift term, but the other terms are always positive to balance the freshwater loss by evaporation. The models disagree in the

Table 2 The global CCMs used in this study

Model	PIC	PDC/CCC	B1	A1B	2 CO₂	4 CO₂	Resolution	Max overturning
CGCM3.1-*T63*	2,210–2,235		2,220–2,240	2,255–2,265			256 × 192 × 29	07.2
CSIRO-mk3.0	2,210–2,250	2,060–2,090	2,270–2,300	2,180–2,190			192 × 189 × 31	19.9
CSIRO-mk3.5	2,330–2,360	2,055–2,065	2,240–2,300				192 × 189 × 31	20.0
ECHAM5/MPI-OM	1,860–1,880				2,055–2,075	2,225–2,235	360 × 180 × 40	21.8
GFDL-CM2.0	0440–0475	2,065–2,085	2,190–2,200	2,250–2,270	0260–0280	0265–0290	360 × 200 × 50	16.1
GFDL-CM2.1	0410–0500	2,075–2,085	2,275–2,300	2,260–2,290	0160–0175	0290–0300	360 × 200 × 50	24.1
UKMO-HadCM3	2,119–2,199	2,009–2,049	2,159–2,199	2,169–2,199			288 × 144 × 20	20.3
MIROC3.2-medres	2,749–2,799	2,080–2,100	2,260–2,300	2,250–2,300	0170–0195	0255–0275	256 × 192 × 33	19.5

The time-slices that were analyzed are shown for each scenario. The acronym *PIC* refers to pre-industrial control; *CCC* to committed climate change; *PDC* to present day control; *B1* and *A1B* to the respective SRES scenarios; and 2 CO_2 and 4 CO_2 to doubling and quadrupling CO_2 concentrations. Resolution is in latitude × longitude × depth. Maximum overturning is constrained to the Atlantic domain, north of 34°S, below a depth of 500 m, and for PIC. Units are in Sverdrups. Reference for CGCM is Kim et al. (2002); for CSIRO is Gordon et al. (2002); for ECHAM is Jungclaus et al. (2006); for GFDL is Delworth et al. (2006); for HadCM3 is Gordon et al. (2000); for MIROC is Hasumi et al. (2004)
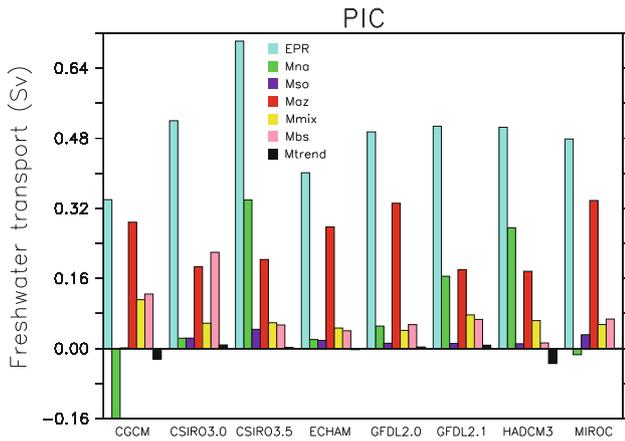
Fig. 3 The budget of Eq. 24 for the pre-industrial control run for the models shown in Table 2

amount of net evaporation over the Atlantic by more than a factor of two: 0.34 Sv (CGCM3.1) and 0.70 Sv (CSIRO3.5).

Figure 4 show the changes in all components of the freshwater budget for present day and future climates, compared to PIC. In general, there is a good correlation between greenhouse forcing and net evaporative freshwater loss over the Atlantic. The evaporative response in ECHAM5 is particularly strong, especially in the 4 CO_2 run where the increase in net evaporation is at least two times higher as in any other model. The robust signal of increased EPR is associated with an increase in freshwater import by the gyre, M_{az}, increased salinification of Bering Strait transport, $M_{\Delta BS}$, and a remaining salinity trend in the basin. The response in M_{mix} is less robust, possibly because the horizontal or along isopycnal diffusion and eddy-advection signals may counteract each other. Also, the response in M_{na} and M_{so} is model dependent and different for different scenarios.

Figure 4 shows that in three models (CGCM, HadCM3, MIROC) changes in M_{na} relative to the PIC value remain small for all scenarios. In MIROC all changes are negative, with a M_{na} minimum in the 2 CO_2 scenario. Also CGCM and HadCM3 feature a M_{na} minimum in a moderately forced scenario ($B1$), but those models feature positive changes in M_{na} for a stronger forced scenario ($A1B$). GFDL2.0 and ECHAM show for modest forcing (2 CO_2) a weakly negative change, but for stronger forced scenarios a larger, positive change. In GFDL2.1, CSIRO3.0 and CSIRO3.5 M_{na} is more sensitive to changes in forcing. These three models already show a large change in PDC. In CSIRO3.5 the change in M_{na} in the $B1$ scenario is less than in PDC; in CSIRO3.0 M_{na} attains a minimum in the $B1$ scenario; in GFDL2.1 the lowest value occurs in the 2 CO_2 climate. In these three models for each scenario the M_{na} change is negative, but for more strongly forced scenarios

($A1B$ and 4 CO_2) the change is less than in the moderately forced scenarios. In GFDL2.1 the response in M_{na} is larger than in any other model. For all six models with a PDC run analyzed the change in M_{na} between PDC and PIC is negative. For five models it becomes more negative in moderately forced scenarios like $B1$ and 2 CO_2 (the only exception is CSIRO3.5). For more strongly forced scenarios like $A1B$ and 4 CO_2 the change becomes less negative or even positive in eight cases, in only two cases it becomes slightly more negative with an insignificant change of 0.001 Sv.

In PIC only CGCM and MIROC feature a negative M_{na}. In future climates also CSIRO3.0 features negative M_{na} values, suggesting that it might possess a stable off-state in those climates (Fig. 5). The values for M_{na}, however, are so small in MIROC and CSIRO3.0 that the salt advection feedback might be offset by atmospheric feedbacks in those models.

5 Discussion and conclusions

In the weak and moderately forced scenarios (PDC, $B1$, 2 CO_2) we observe a reduction in M_{na} relative to PIC for all cases but one. At first sight, it may seem plausible that this decrease in M_{na}, or salt transport by the MOC, is completely due to the weakening of the MOC relative to its PIC value. However, in most cases the relative decrease in salt transport is much larger than the decrease in MOC, see Table 3. For instance, CSIRO3.0 features a sign change in M_{na}, while the MOC in CSIRO3.0 does not reverse.

To elucidate the cause for the observed changes in M_{na}, we determine whether the decrease in M_{na} is primarily caused by a change in circulation, or by a change in salt stratification. This is analyzed by calculating

$$\Delta M_{na} = -(V\Delta S + S\Delta V)/S_0 \tag{25}$$

where the terms in the RHS are defined as

$$V\Delta S = \int 0.5(\bar{v}_{cl1} + \bar{v}_{cl2})(\langle S \rangle_{cl2} - \langle S \rangle_{cl1})\,dz \tag{26}$$

with the subscripts $cl1$ and $cl2$ denoting two different climate states. The equation for $S\Delta V$ reads

$$S\Delta V = \int 0.5(\langle S \rangle_{cl1} + \langle S \rangle_{cl2})(\bar{v}_{cl2} - \bar{v}_{cl1})\,dz \tag{27}$$

Table 4 shows the ensemble-mean values for all scenarios. It should be noted that for each scenario the respective ensemble mean comprises different models (see Table 2), so the results should be treated with care, but some features are robust and can be deduced from this comparison. For instance, the change in circulation yields a decreasing salt transport by the MOC for increasing
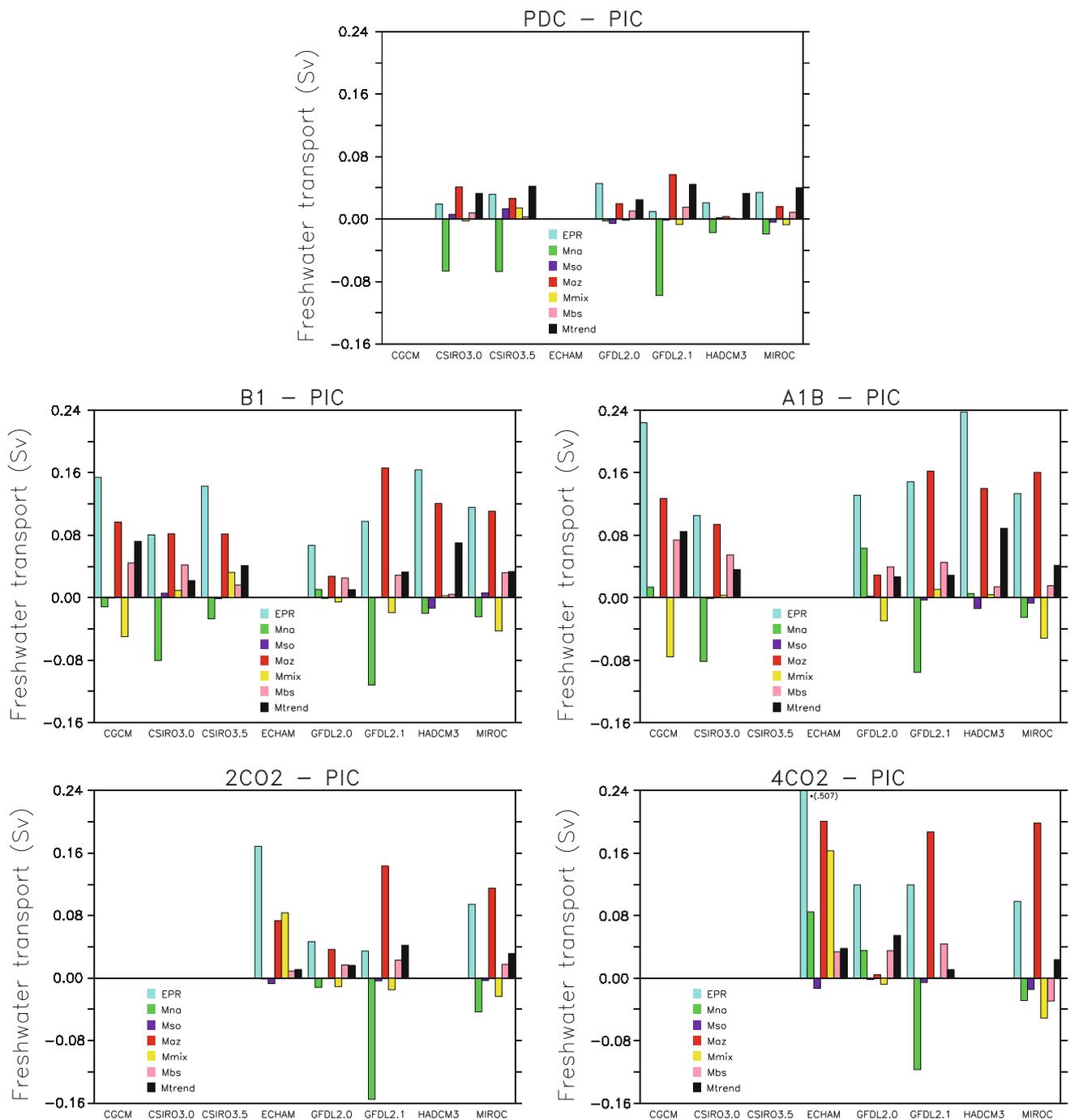
Fig. 4 Same as Fig. 3, but now for each scenario the change in equivalent freshwater budget relative to the pre-industrial control run is displayed. Note that the *blue bar* for EPR in ECHAM extends to 0.507 *Sv*

greenhouse forcing. Moreover, the relative change in $S\Delta V$, or salt transport by the MOC, is much larger than the change in MOC itself. To explain this feature, we investigate the change in circulation further by splitting the ocean in four layers: thermocline water, intermediate water, deep water, and bottom water. The boundaries between thermocline/intermediate, deep, and bottom water are defined by the maxima and minima in overturning

streamfunction at 34°S; the boundary between thermocline and intermediate water is defined by the maximum in vertical salinity gradient. The meridional transport for PIC in these four layers is shown in Fig. 6, which presents the ensemble mean value, together with the change between A1B and PIC, as A1B features the largest change in salt transport by the MOC due to circulation changes ($S\Delta V$). It is seen that the decrease in MOC is almost completely due
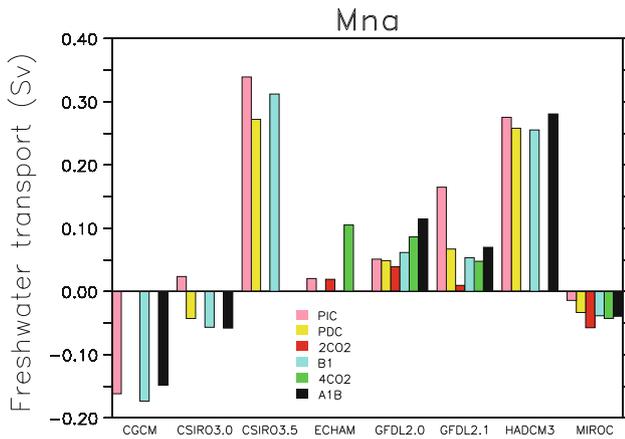
Fig. 5 M_{na} for the six scenarios for each CCM

to reduced inflow of intermediate water, with no significant decrease in inflow of thermocline water. This shift in the vertical enhances the import, or reduces the export of salt more than the decrease in MOC by itself would suggest. The decrease of the net inflow of intermediate waters relative to thermocline waters is due to a shoaling of the MOC in warmer climates. This is a robust feature among all models.

The change in $-V\Delta S/S_0$ is negative for PDC and $2\,CO_2$; positive for the other scenarios. This feature is associated with a salinification of thermocline and intermediate waters in PDC and $2\,CO_2$, in the other scenarios the opposite occurs and these waters freshen while deep water becomes more saline. The change in salinity may reflect a subtle change in the balance of freshwater forcing for the waters crossing 34°S. Locally, and just upstream of 34°S, evaporation increases, causing the surface waters to become more saline. Further upstream, in the tropics and poleward of 45°S, the surface waters gain more freshwater in future greenhouse scenarios. Without a complete Lagrangian analysis it will remain unclear how these changes affect the

salt budget at 34°S, but such an analysis is beyond the scope of this study.

Our finding that the majority of CCMs features an MOC importing freshwater, with M_{na} positive, is consistent with the observation that in future scenario runs (Schmittner et al. 2005; Gregory et al. 2005) these models do not show a shutdown of the MOC. However, in Weber et al. (2007) it was suggested that the value for M_{ov} might be biased in current CCMs. This was based on a inverse model study of Weijer et al. (1999), where $M_{ov} = -0.20\ Sv$; $M_{az} = 0.38Sv$; and EPR $= 0.25\ Sv$. According to these numbers the MOC is exporting freshwater from the Atlantic basin and already resides in the bistable regime. But after 1999 a lot of new ocean data has become available and the numbers of Weijer et al. might have to be revised. The possibility of a bias in CCMs, however, is further corroborated by the initial drift in freshwater transport in HadCM3 (Pardaens et al. 2003), when starting from Levitus data. Apparently, the freshwater transport by the ocean circulation that is in equilibrium with the observed large-scale density field, is much weaker than needed to compensate for the freshwater loss by evaporation in the HadCM3 model. In addition, de Vries and Weber (2005) suggested that coupled models feature a too low M_{az} associated with a too low east-west salinity contrast at 34°S in the South Atlantic. A more detailed discussion on possible biases is beyond the scope of this study; it should involve an analysis of more recent ocean and ocean reanalysis data. In our opinion, there is at the moment not enough evidence to definitely conclude on possible biases in M_{na} in CCMs, moreover, model estimates of M_{na} differ wildly, but on the other hand the presence of a significant bias can also not be excluded; the aforementioned studies at least suggest the possibility of such a bias.

Our main conclusions from this study are: (1) In six out of eight CCMs the salt advection feedback is damping freshwater anomalies, because the overturning circulation exports salt and imports freshwater, indicating that no

Table 3 Maximum overturning at the southern boundary of the Atlantic in Sv (left column below scenario header) and M_{na} in mSv (right column) for PIC

Model	PIC		PDC		2 CO_2		B1		A1B		4 CO_2	
CGCM3.1	5.4	−162					0.37	−0.07	0.41	0.08		
CSIRO-3.0	15.8	24	−0.06	−2.80			−0.13	−3.39	−0.17	−3.44		
CSIRO-3.5	16.1	339	−0.07	−0.20			−0.09	−0.08				
ECHAM5	17.1	21			−0.06	−0.06					−0.10	4.13
GFDL-2.0	15.1	51	0.01	−0.05	−0.03	−0.23	0.01	0.20	0.03	1.24	−0.02	0.69
GFDL-2.1	19.3	165	−0.09	−0.59	−0.21	−0.94	−0.16	−0.68	−0.19	−0.58	−0.23	−0.71
HadCM3	15.9	276	−0.01	−0.06			−0.08	−0.07	−0.09	0.02		
MIROC3.2	16.7	−14	−0.06	−1.35	−0.17	−3.06	−0.20	−1.72	−0.35	−1.79	−0.52	−2.02

For all other scenarios the fractional increase in maximum overturning (left) and M_{na} (right) with respect to PIC is displayed; negative numbers denote a decrease

Table 4 The ensemble mean budget of Eq. 25 for all scenarios with respect to PIC

Scenario	ΔM_{na}	$-S\Delta V/S_0$	$-V\Delta S/S_0$
PDC	-0.045	-0.012	-0.033
2 CO_2	-0.053	-0.025	-0.028
$B1$	-0.038	-0.040	0.002
4 CO_2	-0.006	-0.033	0.027
$A1B$	-0.020	-0.042	0.022

Units are in Sverdrups

Fig. 6 Ensemble mean northward transport, of thermocline (T), intermediate (I), deep (D), and bottom water (B) at the southern boundary of the Atlantic ocean. *Blue bars* denote PIC values, *red bars* the change between *A1B* and PIC

stable off-state exists in the pre-industrial climate in those models; (2) in future climates the Atlantic ocean becomes more evaporative, which must be compensated by an even stronger import of freshwater by the ocean circulation; (3) nevertheless, the fresh water import by the MOC decreases or even changes sign in moderately forced scenarios (*B1* and 2 CO_2) for ten out of eleven scenario runs, while the fresh water import by the gyre increases to close the budget. As a result, the damping effect of the salt advection feedback reduces, or changes to amplifying; (4) the reduced fresh water import associated with smaller M_{na} is mainly caused by a reduced inflow of intermediate water relative to thermocline waters in future climates; (5) when the climate warms further, the increased net evaporation can no longer be reconciled with a further decrease in M_{na}. For most models M_{na} is more negative in a 2 CO_2 than in a 4 CO_2 climate, and more negative in a *B1* scenario run than in an *A1B* scenario run. The only two exceptions show insignificant differences of 0.001 *Sv*.

References

Broecker WS (1997) Thermohaline circulation, the Achilles heel of our climate system: will man-made CO_2 upset the current balance? Science 278:1582–1588

Broecker WS, Peteet DM, Rind D (1985) Does the ocean-atmosphere system have more than one stable mode of operation? Nature 315:21–26

Clark PU, Pisias NG, Stocker TF, Weaver AJ (2002) The role of the thermohaline circulation in abrupt climate changes. Nature 415:863–869

Delworth TL, et al (2006) GFDL's CM2 global coupled climate models. Part I: formulation and simulation characteristics. J Clim 19:643–674

de Vries P, Weber SL (2005) The Atlantic freshwater budget as a diagnostic for the existence of a stable shut down of the meridional overturning circulation. Geophys Res Lett 32. doi: 10.1029/2004GL021450

Dijkstra HA (2007) Characterization of the multiple equilibria regime in a global ocean model. Tellus 59:695–705

Drijfhout SS (2010) The atmospheric response to a THC collapse: scaling relations for the Hadley circulation and the nonlinear response in a coupled climate model. J Clim 23:757–774

Gent P, McWilliams J (1990) Isopycnal mixing in ocean circulation models. J Phys Oceanogr 20:150–155

Gordon C, et al (2000) The simulation of SST, sea ice extent and ocean heat transport in a version of the Hadley centre coupled model without flux adjustments. Clim Dyn 16:147–168

Gordon HB, et al (2002) CSIRO Mk3 climate system model. CSIRO technical report 60, pp 130

Gregory JM, et al (2005) A model intercomparison of changes in the thermohaline circulation in response to increasing atmospheric CO_2 concentration. Geophys Res Lett 32. doi:10.1029/2005GL023209

Griffies SM, Harrison MJ, Pacanowski RC, Rosati A (2004) A technical guide tot MOM4, NOAA/GFDL, pp 337

Grootes PM, Stuiver M, White JWC, Johnsen SJ, Jouzel J (1993) Comparison of oxygen isotope records from the GISP2 and GRIP Greenland ice cores. Nature 366:552–554

Hasumi H, et al (2004) K-1 coupled model (MIROC) description. K1 technical report 1, pp 34

Huisman SE, den Toom M, Dijkstra HA, Drijfhout SS (2010) An indicator of the multiple equilibria regime of the Atlantic meridional overturning circulation. J Phys Oceanogr 40:551–567

Jungclaus JH, et al (2006) Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM. J Clim 19:3952–3972

Kim SJ, Flato GM, de Boer GJ, McPharlane NA (2002) A coupled climate model simulation of the last glacial maximum, part I: transient multi-decadal response. Clim Dyn 19:515–537

Latif ME, Roeckner E, Mikolajewicz U, Voss R (2000) Tropical stabilization of the thermohaline circulation in a greenhouse warming simulation. J Clim 13:1809–1813

Leduc G, Vidal L, Tachikawa K, Rostek F, Sonzogni C, Beaufort L, Bard E (2007) Moisture transport across Central America as a positive feedback on abrupt climatic changes. Nature 445:909–911

Lenton TM, et al (2007) Effects of atmospheric dynamics and ocean resolution on bi-stability of the thermohaline circulation examined using the Grid ENabled Integrated Earth system modelling (GENIE) framework. Clim Dyn 29:591–613

Lohmann G (2003) Atmospheric and oceanic freshwater transport during weak Atlantic overturning circulation. Tellus 55:438–449

Manabe S, Stouffer RJ (1988) Two stable equilibria of a coupled ocean-atmosphere model. J Clim 1:841–863

Pardaens AK, Banks HT, Gregory JM, Rowntree PR (2003) Freshwater transports in HadCM3. Clim Dyn 21:177–195

Rahmstorf S (1996) On the freshwater forcing and transport of the Atlantic thermohaline circulation. Clim Dyn 12:799–811

Rahmstorf S (2000) The thermohaline ocean circulation—a system with dangerous thresholds? Clim Change 46:247–256

Rahmstorf S, et al (2005) Thermohaline circulation hysteresis:a model intercomparison. Geophys Res Lett 32. doi:10.1029/2005GL023655

Schmittner A, Latif M, Schneider B (2005) Model projections of the North Atlantic thermohaline circulation for the 21st century assessed by observations. Geophys Res Lett 32. doi:10.1029/2005GL024368

Stommel H (1961) Thermohaline convection with two stable regimes of flow. Tellus 13:224–230

Stouffer RJ, et al (2006) Investigating the causes of the response of the thermohaline circulation to past and future climate changes. J Clim 19:1365–1387

Trenberth KE, Caron JM (2001) Estimates of meridional atmosphere and ocean heat transports. J Clim 14:3433–3443

Vellinga M, Wood RA (2002) Global climatic impacts of a collapse of the Atlantic thermohaline circulation. Clim Change 54:251–267

Weber SL, Drijfhout SS (2007) Stability of the Atlantic meridional overturning circulation in the last glacial maximum climate. Geophys Res Lett 34. doi:10.1029/2007GL031437

Weber SL, Drijfhout SS, Abe-Ouchi A, Crucifix M, Eby M, Ganopolski A, Murakami S, Otto-Bliesner B, Peltier WR (2007) The modern and glacial overturning circulation in the Atlantic ocean in PMIP coupled model simulations. Clim Past 3:51–64

Weijer W, de Ruijter WPM, Dijkstra HA, van Leeuwen PJ (1999) Impact of interbasin exchange on the Atlantic overturning circulation. J Phys Oceanogr 29:2266–2284

Yin J, Stouffer RJ (2007) Comparison of the stability of the Atlantic thermohaline circulation in two coupled atmosphere-ocean general circulation models. J Clim 20:4293–4315