Earth Syst. Dynam., 5, 383–397, 2014
www.earth-syst-dynam.net/5/383/2014/
doi:10.5194/esd-5-383-2014

Earth System Dynamics
Open Access

Mechanism for potential strengthening of Atlantic overturning prior to collapse

D. Ehlert[1,2,*] and A. Levermann[1,2]

[1]Potsdam Institute for Climate Impact Research, Potsdam, Germany
[2]Institute of Physics, Potsdam University, Potsdam, Germany
[*]now at: Simon Fraser University, Burnaby, British Columbia, Canada

Correspondence to: D. Ehlert (dehlert@sfu.ca)

Received: 22 November 2013 – Published in Earth Syst. Dynam. Discuss.: 2 January 2014
Revised: 5 September 2014 – Accepted: 21 September 2014 – Published: 6 November 2014

Abstract. The Atlantic meridional overturning circulation (AMOC) carries large amounts of heat into the North Atlantic influencing climate regionally as well as globally. Palaeo-records and simulations with comprehensive climate models suggest that the positive salt-advection feedback may yield a threshold behaviour of the system. That is to say that beyond a certain amount of freshwater flux into the North Atlantic, no meridional overturning circulation can be sustained. Concepts of monitoring the AMOC and identifying its vicinity to the threshold rely on the fact that the volume flux defining the AMOC will be reduced when approaching the threshold. Here we advance conceptual models that have been used in a paradigmatic way to understand the AMOC, by introducing a density-dependent parameterization for the Southern Ocean eddies. This additional degree of freedom uncovers a mechanism by which the AMOC can increase with additional freshwater flux into the North Atlantic, before it reaches the threshold and collapses: an AMOC that is mainly wind-driven will have a constant upwelling as long as the Southern Ocean winds do not change significantly. The downward transport of tracers occurs either in the northern sinking regions or through Southern Ocean eddies. If freshwater is transported, either atmospherically or via horizontal gyres, from the low to high latitudes, this would reduce the eddy transport and by continuity increase the northern sinking which defines the AMOC until a threshold is reached at which the AMOC cannot be sustained. If dominant in the real ocean this mechanism would have significant consequences for monitoring the AMOC.

1 Introduction

The Atlantic meridional overturning circulation (AMOC) is being considered as one of the tipping elements of the climate system (Lenton et al., 2008). While the definition by Lenton et al. (2008) is based on the idea that tipping elements respond strongly to a small perturbation, the AMOC might also be a tipping element in the dynamic sense of the word (Levermann et al., 2012). That is to say that a small external perturbation induces a self-amplification feedback by which the circulation enters a qualitatively different state. This self-amplification is due to the salt-advection feedback (Stommel, 1961; Rahmstorf, 1996) and has been found in a number of comprehensive ocean as well as coupled climate models (Manabe and Stouffer, 1993; Rahmstorf et al., 2005; Stouffer et al., 2006b; Hawkins et al., 2011). A cessation of the AMOC would have far-reaching implications for global climate (Vellinga and Wood, 2002) which include (1) a strong reduction of northern hemispheric air and ocean temperatures (Manabe and Stouffer, 1988; Mignot et al., 2007), (2) a reduction in European precipitation and (3) its wind pattern (Laurian et al., 2009), (4) a dynamic sea level increase in the North Atlantic (Levermann et al., 2005; Yin et al., 2009), (5) a perturbation of the Atlantic ecosystem (Schmittner, 2005; Kuhlbrodt et al., 2009), (6) a southward shift in the tropical rain belt and associated impacts on vegetation (Stouffer et al., 2006a) and (7) a perturbation of the Asian monsoon system (Goswami et al., 2006).

Conceptual models to capture the basic aspect of a meridional overturning circulation can be divided into models in which the overturning strength is determined by the meridional density difference in the Atlantic (Stommel, 1961; Rahmstorf, 1996) and those in which its strength is linked to the vertical density structure (Gnanadesikan, 1999). Stommel's model captures the salt-advection feedback in a pure form by resolving only the advection of the active tracers in two fixed-size boxes representing the northern downwelling and southern upwelling regions. The overturning strength is assumed to be proportional to the meridional density difference which was found to be valid in a number of ocean and climate models (e.g. Griesel and Morales-Maqueda, 2006; Rahmstorf, 1996; Schewe and Levermann, 2010). The Stommel model is however missing a representation of the energy-providing processes for the overturning, such as the Drake Passage effect and low-latitudinal mixing (Kuhlbrodt et al., 2007) as well as the influence of the Southern Ocean eddy circulation.

These processes are captured in a conceptual way by the model of Gnanadesikan (1999) which links the overturning to the vertical density profile as represented by the pycnocline depth but treats meridional density differences as a constant. It was shown that this kind of model is not consistent with the fact that the meridional density gradient indeed changes with changing overturning in a number of different climatic conditions (Levermann and Griesel, 2004; Griesel and Morales-Maqueda, 2006). By construction it does not capture the salt-advection feedback and can thereby not be used to study the possibility of a threshold behaviour of the overturning.

There have been a number of attempts to combine these two approaches and thereby to comprise the horizontal tracer-advection with the vertical one (Marzeion and Drange, 2006; Johnson et al., 2007; Fürst and Levermann, 2012).

Here we advance the simplest of the suggested models (Fürst and Levermann, 2012) by introducing an additional parameterization for the Southern Ocean eddy flux. As found in a comprehensive coarse resolution ocean model (Levermann and Fürst, 2010) the horizontal scale of the Southern upwelling region can change and neglecting this change leads to a misrepresentation of the circulation within the Gnanadesikan (1999) framework. We attempt to complement the conceptual model in order to correct for this shortcoming. To this end we add a variable, meridional density difference in the southern Atlantic ocean in the scaling of the eddy-induced return flow. As will be shown, this allows for a qualitatively different response of the AMOC under freshwater forcing compared to earlier studies: a growth of the northern deep water formation with increasing freshwater flux from low to high northern latitudes within the Atlantic before the threshold is reached and no AMOC in the modelled sense can be sustained. The threshold behaviour found here is consistent with the salt-advection feedback in the sense of a net-salinity transport from lower latitudes into the northern Atlantic by the overturning as suggested by Rahmstorf (1996). This threshold behaviour has been shown in box models and complex climate models (Huisman et al., 2010; Weaver et al., 2012) but also in observations (Bryden et al., 2011).

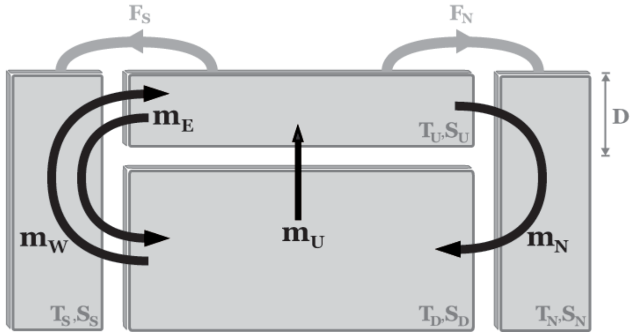

Figure 1. Schematic of the conceptual model as suggested in Fürst and Levermann (2012) and used here. The depth of the pycnocline D is determined by the balance between the northern deep water formation m_N, the upwelling in the low-latitudes m_U in response to downward mixing, the Ekman upwelling m_W and the eddy-induced return flow m_E. Salinity is advected along with these transport processes and determines together with a fixed temperature distribution the horizontal density differences. The differences are between low-latitudinal box and northern box, $\Delta\rho$, and low-latitudinal and southern box, $\Delta\rho_{SO}$, respectively. The density differences, in turn, determine the northern sinking, $m_N \propto D^2 \Delta\rho$, and the eddy-induced return flow, $m_E \propto D\,\Delta\rho_{SO}$.

This paper is structured as followed: firstly we describe the parameterization of the transport processes, pycnocline dynamics and salinity dynamics, i.e. horizontal density distribution (Sect. 2). The transport processes include the two fundamental driving mechanisms (Kuhlbrodt et al., 2007) which are low-latitudinal upwelling (Munk, 1966; Munk and Wunsch, 1998; Huang, 1999; Wunsch and Ferrari, 2004) and wind-driven upwelling in southern latitudes (Toggweiler and Samuels, 1995, 1998). In order to examine the behaviour of the model we derived governing equations for the two driving mechanisms separately as well as for the full case. The threshold behaviour, as described by Stommel (1961), is caused by the salinity advection. For simplicity we keep the temperatures fixed throughout the paper (Sect. 3). Section 4 discusses the change in the AMOC with increasing freshwater flux into the North Atlantic for the wind-driven case and the full case. Also, Sect. 5 discusses the behaviour of the AMOC under freshwater forcing, but for simulations using a climate model of intermediate complexity. We present our conclusions in Sect. 6.

2 Model description

We use a standard inter-hemispheric model with four varying boxes (Fig. 1): (1) a northern box representing the northern

North Atlantic with deep water formation, (2) an upper low-latitudinal box, (3) a deeper low-latitudinal box below the pycnocline and (4) a southern box with southern upwelling and eddy return flow (Gnanadesikan, 1999). The northern and southern boxes are fixed in volume while the low-latitudinal boxes vary in size according to the dynamically computed pycnocline depth. The four meridional tracer transport processes between the boxes control the horizontal and vertical density structure on the one hand and they control the overturning on the other hand. The density structure, in turn, determines the transport processes. Changes in the vertical density structure are described by variations in the pycnocline depth. The horizontal density structure is expressed by a southern and a northern meridional density difference. They depend on the dynamics of the active tracers, temperature, T, and salinity, S. For simplicity we assume a linear density function $\Delta\rho = \rho_0(\beta_S \Delta S - \alpha_T \Delta T)$ (Stommel, 1961). In order to capture the main feedback for a threshold behaviour while keeping the model legible, we include salinity advection and neglect changes in temperature. The simplification is further justified because the temperature in the upper layers is strongly coupled to atmospheric temperature which is to the first order determined by the solar insulation. We thus assume the ocean temperature in the upper layers to be constant. The high-latitudinal boxes represent strong outcropping regions which homogenizes the water column and extends the argument to depth. In steady state, the fourth box, deeper low-latitude ocean, is determined by the three other boxes. That means the approximation is valid for the whole model in equilibrium and temperature is used as an external parameter. The base of our work is the model by Fürst and Levermann (2012). We use the same parameterizations for the northern deep water formation and the upwelling processes. For the eddy return flow we introduce a different scaling by implementing southern meridional density difference which has strong influences on the behaviour of the model (Sects. 3 and 4).

2.1 Tracer transport processes

Different scaling for the deep water formation (as summarized in Fürst and Levermann, 2012) have been suggested. Here we use a parameterization suggested by Marotzke (1997) and apply a β-plane-approximation to it. The resulting northern sinking scales linearly with the meridional density difference and quadratically with the pycnocline depth following geostrophic balance and vertical integration.

$$m_\mathrm{N} = \frac{Cg}{\beta_N L_y^N} \frac{\Delta\rho}{\rho_0} D^2 = C_N \Delta\rho D^2 \tag{1}$$

Because all values are external parameters (Table 1) except the meridional density difference $\Delta\rho = \rho_N - \rho_U$ and the pycnocline depth D, the parameters are comprised into one constant C_N. In contrast to previous approaches (e.g. Rahmstorf, 1996) the meridional density difference does not span the whole Atlantic but instead is taken between low and high northern latitudes in accordance with the geostrophic balance between the meridional density difference and the North Atlantic Current.

The low-latitudinal upwelling follows a vertical advection-diffusion balance (Munk and Wunsch, 1998). That is to say, downward turbulent heat flux is balanced by upward advection. This balance with a constant diffusion coefficient for the full upwelling region yields an inverse proportionality between upward volume transport and pycnocline depth. Again all external parameters are expressed by one constant C_U to obtain

$$m_\mathrm{U} = BL_U \frac{\kappa}{D} = \frac{C_U}{D}. \tag{2}$$

The southern upwelling term is considered to be independent of the pycnocline depth and results from the so-called Drake Passage effect (Toggweiler and Samuels, 1995):

$$m_\mathrm{W} = B \frac{\tau_\mathrm{Dr}}{|f_\mathrm{Dr}|\rho_0} = C_W. \tag{3}$$

The eddy return flow is parameterized following Gent and McWilliams (1990), which yields a tracer transport proportional to the slope of the outcropping isopycnals. In the formulation of Gnanadesikan (1999) this is represented by a linear dependence on the pycnocline depth divided by a horizontal scale for the outcropping region which is taken to be constant. The assumption of a constant horizontal scale for the outcropping region is not consistent with freshwater hosing experiments in a comprehensive though coarse resolution ocean model (Levermann and Fürst, 2010). Levermann and Fürst (2010) show that the eddy return flow is proportional to pycnocline depth over a variable horizontal scale of the outcropping. Here we attempt to capture variations in the meridional horizontal length scale of the outcropping region by the meridional density difference between the low-latitude ocean and the Southern Ocean, $\Delta\rho_\mathrm{SO} = \rho_S - \rho_U$. We thus use the parameterization

$$m_\mathrm{E} = BA_\mathrm{GM} \frac{\Delta\rho_\mathrm{SO}}{\rho_0} \frac{D}{H} = C_\mathrm{E} \Delta\rho_\mathrm{SO} D. \tag{4}$$

As before, all quantities except D and $\Delta\rho_\mathrm{SO}$ are external parameters and compressed into one constant C_E.

2.2 Pycnocline and salinity dynamics

The temporal evolution of the pycnocline is determined by the tracer transport equation following Marzeion and Drange (2006).

$$BL_U \frac{\mathrm{d}D}{\mathrm{d}t} = m_\mathrm{U} + m_\mathrm{W} - m_\mathrm{E} - m_\mathrm{N} \tag{5}$$

Salinity equations for each box are derived from the advection in and out of the box, conserving salinity, as well as

Table 1. Physical parameters for used for the model.

Parameter	Value	Unit	Description
			Geometry
H	4×10^3	m	Average depth of the Atlantic Ocean basin
B	1×10^7	m	Average width of the Atlantic Ocean
L_N	3.34×10^6	m	Meridional extend of the northern box
L_U	8.90×10^6	m	Meridional extend of the tropical box
L_S	3.34×10^6	m	Meridional extend of the southern box
			Stratification
ρ_0	1027	$\mathrm{kg\,m^{-3}}$	Average density of the Atlantic Ocean
S_0	35	psu	Average salinity of the Atlantic ocean
L_y^N	1.5×10^6	m	Meridional extent of the northern outcropping
A_{GM}	1×10^6	$\mathrm{m^2\,s^{-1}}$	Thickness diffusivity
κ	4×10^{-5}	$\mathrm{m^2\,s^{-1}}$	Background vertical diffusivity
α_{T}	2.1×10^{-4}	$1\,^\circ\mathrm{C}^{-1}$	Thermal coefficient for isobars
α		$\mathrm{kg\,(m^3\,^\circ C)^{-1}}$	Product of ρ_0 and α_{T}
β_S	8×10^{-4}	$1\,\mathrm{psu}^{-1}$	Haline coefficient for isobars
β		$\mathrm{kg\,(m^3\,psu)^{-1}}$	Product of ρ_0 and β_S
C	0.1	–	Constant accounting for geometry and stratification
			External forcing
β_N	2×10^{-11}	$1\,\mathrm{ms}^{-1}$	Coefficient for β-plane approximation in the North Atlantic
f_{Dr}	-7.5×10^{-5}	$1\,\mathrm{s}^{-1}$	Coriolis parameter in the Drake Passage
τ_{Dr}	0.1	$\mathrm{N\,m^{-2}}$	Average zonal wind stress in the Drake Passage
F_N	0.1	Sv	Northern meridional atmospheric freshwater transport
F_S	0.1	Sv	Southern meridional atmospheric freshwater transport
T_N	5.0	°C	Temperature of the northern box
T_U	12.5	°C	Temperature of the tropical surface box
T_S	7.0	°C	Temperature of the southern box

the surface fluxes, F_N and F_S which represent atmospheric freshwater transport as well as the horizontal gyre transport. The advection scheme follows the arrows shown in Fig. 1. In computing the temporal changes in total salinity the changes in volume due to the pycnocline dynamics need to be accounted for.

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_U S_U) = m_{\mathrm{U}} S_D + m_{\mathrm{W}} S_S - S_U(m_{\mathrm{N}} + m_{\mathrm{W}}) + S_0(F_N + F_S) \tag{6a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_N S_N) = m_{\mathrm{N}}(S_U - S_N) - S_0 F_N \tag{6b}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_S S_S) = m_{\mathrm{W}}(S_D - S_S) + m_{\mathrm{E}}(S_U - S_S) - F_S S_0 \tag{6c}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(V_D S_D) = m_{\mathrm{N}} S_N + m_{\mathrm{E}} S_S - S_D(m_{\mathrm{U}} + m_{\mathrm{W}}). \tag{6d}$$

With the finite difference method applied to Eqs. (1)–(6), we made numerical simulations which reached in equilibrium the values shown in Table 2 with the parameters given in Table 1.

Table 2. Numerical solution of the model by applying the finite difference method on Eqs. (1)–(6). Equilibrium state is obtained after 2000 years with a time step of 14 days and the starting conditions: salinities set to 35 psu and the pycnocline depth set to 500 m.

Name		Value
Salinities	S_N	35.04 psu
	S_U	35.24 psu
	S_D	35.02 psu
	S_S	34.79 psu
Tracer transports	m_{U}	17.5 Sv
	m_{U}	5.8 Sv
	m_{W}	13.0 Sv
	m_{E}	1.2 Sv
Meridional density	$\Delta\rho$	$1.45\,\mathrm{kg\,m^{-3}}$
differences	$\Delta\rho_{\mathrm{SO}}$	$0.82\,\mathrm{kg\,m^{-3}}$
Pycnocline depth	D	615 m

3 Governing equation

Here we derive an equation for the steady-state solution of Eqs. (1)–(6) by comprising them into one equation of the oceanic pycnocline, D. We derive governing equations for the full case as well as for the analytically solvable cases of a purely mixing- and a purely wind-driven cases. The model is limited to positive and real solutions for the pycnocline (see Fig. 2) as well as for non-negative tracer transport values. A parameter combination that does not allow for a solution of this kind is thereby inconsistent with an overturning circulation as represented by this model. We denote a parameter region for which no such a physical solution exists as an "AMOC-off-state-region". As in the earlier version of the model (Fürst and Levermann, 2012) we find a threshold behaviour with respect to an increase of the freshwater flux, F_N, for all three cases. The focus of this study is not to show the existence of such a threshold of all parameter values. But, it is to present a mechanism by which the overturning can increase between steady states under different freshwater forcings before the threshold is reached and no AMOC can be sustained.

3.1 Full case

In the full case the governing equation is a polynomial of the 10th order in the pycnocline depth (Appendix A1, Eq. A7). Thus solutions can only be found numerically. Of the 10 mathematical roots, two are positive and real but of two adjacent solutions only one can be stable. Numerical solutions were obtained in two ways. First by finding the roots of the polynomial (Appendix A1, Eq. A7) and second by time forward integration of the original set of Eqs. (1)–(6) with different initial conditions. The time integration naturally selects the stable solutions. Though this is not a proof by any means, we feel confident to say that the solution with $D = 616$ m is the stable of the two physical solutions (Fig. 3a). The corresponding tracer transport values are provided in Fig. 3b. The northern sinking decreases with increasing freshwater forcing for the parameter set of Table 1. The equation for the northern sinking as it results from the scaling (Eq. 1) and the salinity equations:

$$m_{\mathrm{N}} = -\frac{1}{2} C_N D^2 \alpha \Delta T \pm \sqrt{\frac{1}{4} C_N^2 D^4 \alpha^2 \Delta T^2 - C_N D^2 \beta F_N S_0} \quad (7)$$

was also valid in the earlier version of the model (Fürst and Levermann, 2012). Rahmstorf (1996) provides a similar solution for the northern deep water formation with k as proportionality factor between the northern sinking and the north–south density difference:

$$m_{\mathrm{N}} = -\frac{1}{2} k \alpha \left(T_{\mathrm{N}} - T_S\right) \pm \sqrt{\frac{1}{4} k^2 \alpha^2 \left(T_{\mathrm{N}} - T_S\right)^2 + k \beta F_S S_0}. \quad (8)$$

In these earlier models only positive roots of the solution yield stable equilibria. That differs from our model where for certain amounts of freshwater forcing the negative sign of the root in Eq. (7) (respectively Eq. 8) needs to be considered, as for example in the wind-driven case discussed below.

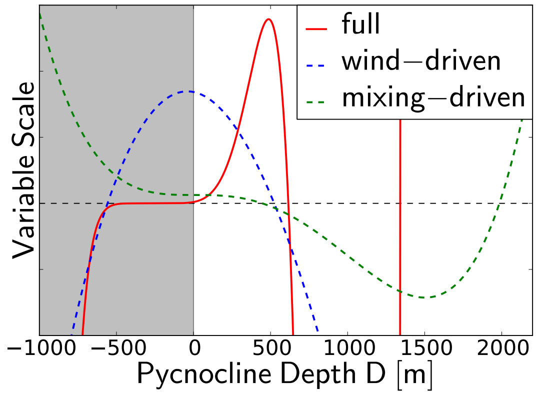

Figure 2. Trend of the governing equation for the full case (red line), the wind-driven case ($m_{\mathrm{U}} = 0$, blue line) and the mixing-driven case ($m_{\mathrm{W}} = m_{\mathrm{E}} = 0$, green line). The intersections with zero (black dashed line) are solutions of the polynomial but those in the grey shadowed area correspond to a negative pycnocline depth. Therefore they are not physical. All three cases have two positive solutions, a lower stable, physical one D and a higher unstable or non-physical one $\hat{D}$. In the wind-driven case the non-physical solution is out of range of the pycnocline but this solution is shown in Fig. 5. For the full case the solutions are $D = 616$ m and $\hat{D} = 1342$ m, for the wind-driven case they are $D = 523$ m and $\hat{D} = 6190$ m and for the mixing-driven case they are $D = 446$ m and $\hat{D} = 1985$ m.

The threshold of the overturning is reached when the eddy return flow becomes negative (Fig. 3b, grey shaded area) because the parameterization of the eddy return flow, i.e. the model itself, is only valid for positive values. Therefore, the shift to negative eddy return flow is interpreted as a point of instability of the circulation pattern described by the model. Physically, reaching the threshold means that there is no outcropping of isopycnals in the Southern Ocean anymore. Thus the eddy return flow does not follow the physics that is described by the baroclinic instability and thereby it does not follow the parameterization by Gent and McWilliams (1990). This also establishes a qualitatively different circulation pattern.

It should be noted that also negative freshwater forcing was applied, which might not be applicable with surface transport. However, the threshold freshwater forcing is in the positive range. This is also true for the mixing- and wind-driven case.

Besides freshwater forcing from lower latitudes into the northern box only, experiments were performed with freshwater forcing from the lower latitudes into the southern box and from lower latitudes into both southern and northern box. All experiments showed the same behaviour in the overturning as all experiments affect the meridional density differences in the same way.

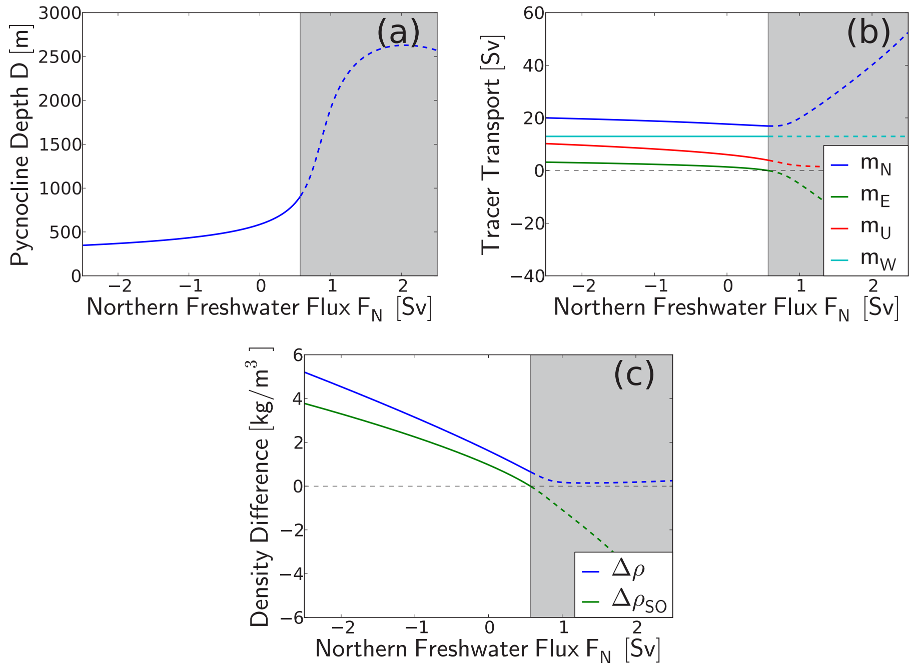

Figure 3. In steady state only one real stable solution of the governing equation of the full case exists. This solution increases under freshwater forcing **(a)**. The tracer transport processes show different behaviours **(b)**. The eddy return flow m_E decreases (**b**, green line) until it becomes negative and the break down of circulation is reached (grey shaded area). Also the density difference between the southern box and the low-latitudinal box, $\Delta\rho_{SO}$, crosses zero at the threshold level (**c**, green line). See also Fig. 1 for explanation of the tracer transports m and the density differences.

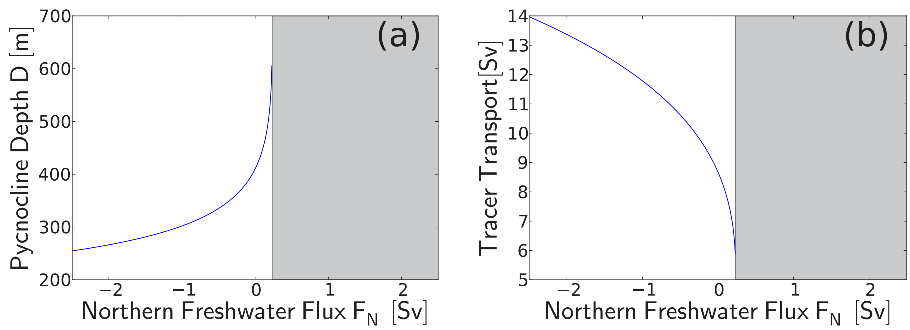

Figure 4. In steady state the governing equation for the mixing-driven case has one real, stable solution **(a)** until a threshold level is reached (grey shaded area). Thereafter, no real solution exists. The tracer transports are upwelling in the mid latitudes and northern sinking which balance each other ($m_N = m_U$) and decrease under increasing freshwater forcing **(b)**.

3.2 Mixing-driven case

The purely mixing-driven case is defined by $C_E = C_W = 0$. In this case the pycnocline dynamics in steady state (Eq. 5) reduces to $m_N = m_U = C_U/D$. As the eddy return flow is eliminated from the equation, this case has not changed compared to the model of Fürst and Levermann (2012): the governing equation is a polynomial of the fourth order in pycnocline depth and has one physical solution which decreases with increasing freshwater forcing (Fig. 4a). The overturning decreases until a threshold level (Fig. 4b) which is reached when the pycnocline and therefore the tracer transport processes become complex. The critical northern freshwater flux can be calculated by zero-crossing of the discriminant of the polynomial.

$$F^{\mathrm{crit}}_{N,\mathrm{mixing}} = \frac{3\,(2C_N\rho_0)^{1/3}\,C_U^{2/3}\alpha^{4/3}}{8\beta S_0}|\Delta T|^{4/3} \qquad (9)$$

3.3 Wind-driven case

The purely wind-driven circulation is defined by $C_U = 0$. Thus the tracer–transport balance in steady state (Eq. 5) reduces to $m_N = m_W - m_E$ into which the eddy return flow and the northern sinking are included as functions of the pycnocline depth and external parameters of Table 1 (see Appendix A for a detailed derivation). For the northern sinking the northern salinity difference is calculated via the salinity balance of North Atlantic (Eq. 6b) and inserted into the scaling of the northern sinking (Eq. 1), similarly for the eddy return flow by using the Southern Ocean salinity balance (Eq. 6c). The emerging governing equation is a third order polynomial of the pycnocline depth D which we solve analytically.

$$
\begin{aligned}
& D^3 C_{\mathrm{E}} C_N \alpha \Delta T \left[\frac{\beta S_0 (F_N + F_S)}{C_W} + \alpha \Delta T_{\mathrm{SO}} \right] \\
& + D^2 \left[C_N F_N S_0 \beta + C_N C_W \alpha \Delta T + \frac{C_{\mathrm{E}}^2}{C_W^2} (S_0 \beta (F_N + F_S) \right. \\
& \left. + C_W \alpha \Delta T_{\mathrm{SO}})^2 \right] + D 2 C_{\mathrm{E}} [\beta S_0 (F_N + F_S) + C_W \alpha \Delta T_{\mathrm{SO}}] \\
& + C_W^2 = 0
\end{aligned}
$$

The solutions depend on the sign of the discriminant of the polynomial $d = (q/2)^2 + (p/3)^3$ with p and q defined as follows:

$$
\begin{aligned}
\frac{q}{2} = & \frac{1}{2} \left(\frac{C_N F_N S_0 \beta + C_N C_W \alpha \Delta T + C_{\mathrm{E}}^2 A^2}{3 C_{\mathrm{E}} \alpha \Delta T A} \right)^3 \\
& - \frac{C_W F_N S_0 \beta + C_W^2 \alpha \Delta T}{3 C_{\mathrm{E}} C_N \alpha^2 \Delta T^2 A} - \frac{C_{\mathrm{E}} C_W A}{3 C_N^2 \alpha^2 \Delta T^2} \\
& + \frac{C_W^2}{2 C_{\mathrm{E}} C_N \alpha \Delta T A} \\
\frac{p}{3} = & \frac{6 C_W C_N \alpha \Delta T - 1}{9 C_N^2 \alpha^2 \Delta T^2} - \frac{F_N S_0 \beta + C_W \alpha \Delta T}{9 C_{\mathrm{E}}^2 C_N^2 \alpha^2 \Delta T^2 A^2}
\end{aligned}
$$

with $A = \left(\frac{S_0 \beta}{C_W} (F_N + F_S) + \alpha \Delta T_{\mathrm{SO}} \right)$.

A polynomial of the third order has either one root (Appendix A2, Eq. A2) if the discriminant is positive, or three roots (Appendix A2, Eq. A11) if the discriminant is negative which is the case for the parameters of Table 1 near the threshold (Fig. 5). Only one of the three mathematical roots is a physical solution of equilibrium state of the model because one root is negative (Fig. 5, solution 1) and the other solution has a negative northern sinking and the pycnocline values are out of range of the ocean depth (Fig. 5, solution 0). No physical solution exists, when the eddy return flow becomes negative. At this threshold the discriminant of the governing equation has a negative pole which can be used to calculate the critical freshwater flux. In the following we describe a more straightforward way to give dependencies of the critical freshwater flux. Assuming steady state for the salinity balance of the upper low-latitudinal box (Eq. 6a equal to zero, with $m_{\mathrm{U}} = 0$) and for the tracer transport balance ($m_{\mathrm{E}} + m_{\mathrm{N}} = m_{\mathrm{W}} = C_W$), the salinity difference between the Southern Ocean and the upper low-latitudes emerges:

$$
\Delta S_{\mathrm{SO}} = S_S - S_U = -\frac{S_0}{C_W} (F_N + F_S) .
$$

The salinity difference contains no variables. As the temperature dynamics are not considered in this model, the horizontal density difference between these two boxes is constant for a fixed set of parameters.

$$
\begin{aligned}
\Delta \rho_{\mathrm{SO}} &= \beta \Delta S_{\mathrm{SO}} - \alpha \Delta T_{\mathrm{SO}} \\
&= -\beta \frac{S_0}{C_W} (F_N + F_S) - \alpha \Delta T_{\mathrm{SO}} \qquad (10)
\end{aligned}
$$

The critical eddy return flow is equal to zero. Using the definition of the flow (Eq. 4) and the fact that the critical pycnocline depth is far in the positive range, Eq. (10) can be set to zero at the threshold level. The critical freshwater flow becomes:

$$
F_{N,\mathrm{wind}}^{\mathrm{crit}} = -\frac{\alpha \Delta T_{\mathrm{SO}} C_W}{S_0 \beta} - F_S .
$$

The critical northern freshwater flow depends linearly on the southern temperature difference and on the southern wind stress (via C_W) and a higher southern freshwater flux would lower the critical northern freshwater flow. Please note that this is a significant difference to previous approachers (Fürst and Levermann, 2012; Rahmstorf, 1996), where the critical freshwater flow is in first or higher order (Eq. 9) sensitive to the northern temperature difference which has no influence onto the critical freshwater flux in this case.

4 Freshwater-induced AMOC strengthening

The introduction of the southern density difference as a variable changing the eddy return flow results in a mechanism that has rarely been reported before: an increasing overturning under northern freshwater forcing prior to a threshold in freshwater beyond no AMOC, as described here, can be sustained. Cimatoribus et al. (2014) found a similar behaviour in a different box model, but under freshwater forcing from the southern into the northern Atlantic. The mechanism in the model presented here is simple: a freshwater flux from low latitudes into the high northern latitudes reduces the eddy return flow. If this reduction is not compensated by a reduction in mixing-driven upwelling (as for example in a mainly wind-driven AMOC), then due to continuity the northern sinking has to increase since the Southern Ocean upwelling is constant. Furthermore, it should be noted that this result depends on the assumption that the northern sinking, m_{N}, is a function of the square of the pycnocline depth and the meridional density difference (see Eq. 1). Consequently, only solutions of the pycnocline depth for the governing equation (see Eqs. A7 and A9) under freshwater forcing which fulfill the parameterization of the northern sinking (see Eq. 1) are allowed. In general, the changes in the meridional density differences are the main driver for changes in the northern sinking and the eddy return flow, i.e. the drivers for the

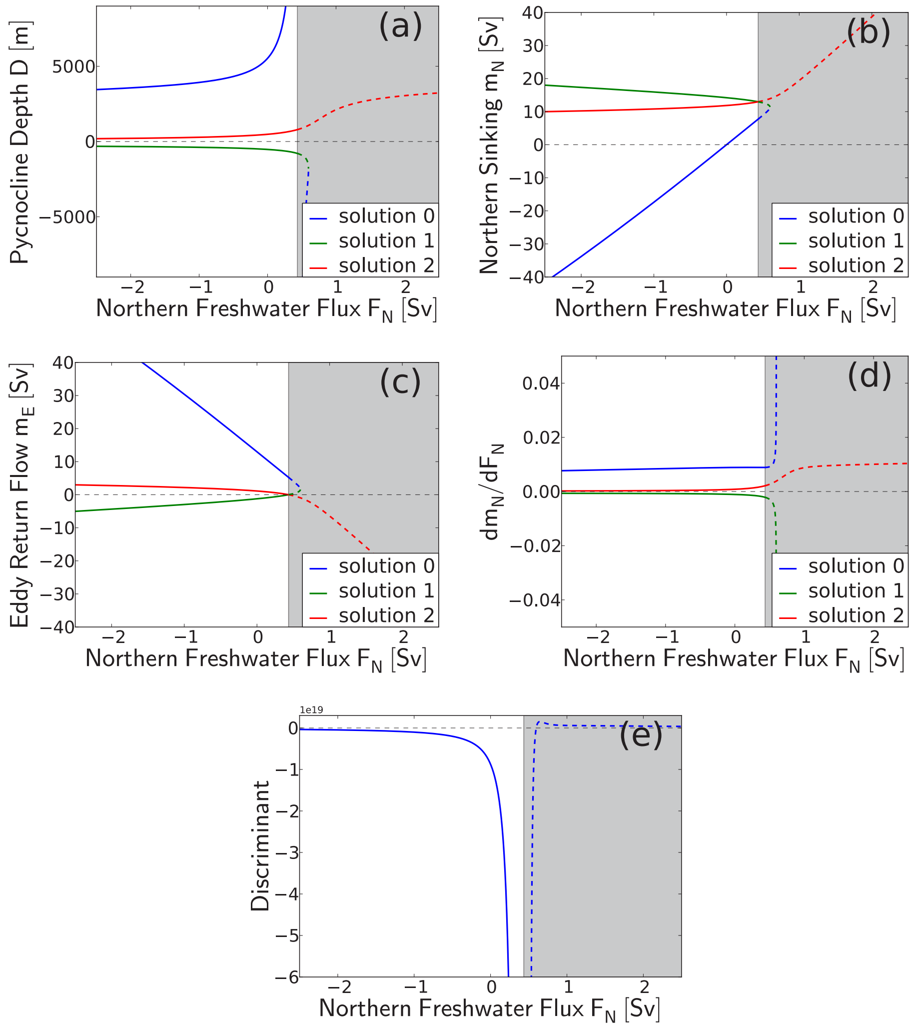

Figure 5. In steady state only one physical solution of governing equation for the wind-driven case exists. There are three real solutions before the circulation breaks down (**a**, white area) because the discriminant is negative **(e)**. The physical branch is solution 2 (red line). The threshold (grey shaded area) is reached when the eddy return flow becomes negative (**c**, red line) and the discriminant of the governing equation has a negative pole **(e)**. The zero crossing of the discriminant, which was in the parent model (Fürst and Levermann, 2012) the indicator for a cessation of the circulation, does not appear within the range applicability of our model. Within that range the northern sinking always increases (**b**, red line) and its derivative is positive (**d**, red line).

freshwater induced strengthening of the AMOC. Changes in the vertical density differences, implemented here as changes in the pycnocline depth, stabilize the overturning circulation. The mechanism of an AMOC strengthening under freshwater forcing is always dominant in the wind-driven case which we will proof at the end of this section. In the full case the mechanism takes no effect for the parameters of Table 1 but it emerges if the Southern Ocean temperature difference is changed in such a way as to make the mixing less relevant (Fig. 6).

4.1 Full case

In order to gain a better understanding of this behaviour, the tracer transport processes balance in steady state (Eq. 5 equal to zero) is differentiated with respect to the northern freshwater flux. That gives an equation for the derivative of northern sinking:

$$\frac{\mathrm{d}m_{\mathrm{N}}}{\mathrm{d}F_N} = -\frac{\mathrm{d}m_{\mathrm{E}}}{\mathrm{d}F_N} + \frac{\mathrm{d}m_{\mathrm{U}}}{\mathrm{d}F_N}. \tag{11}$$

Using the parameterizations of the eddy return flow (Eq. 4) and low-latitudinal upwelling (Eq. 2), Eq. (11) yields

$$\frac{\mathrm{d}m_{\mathrm{N}}}{\mathrm{d}F_N} = -\left(\frac{C_U}{D^2} + C_{\mathrm{E}}\Delta\rho_{\mathrm{SO}}\right)\frac{\partial D}{\partial F_N} - C_{\mathrm{E}}D\frac{\partial \Delta\rho_{\mathrm{SO}}}{\partial F_N}.$$

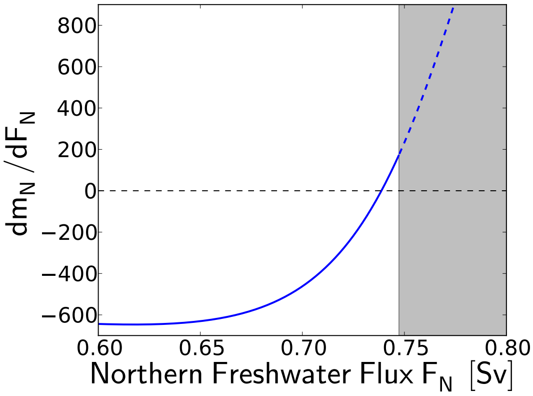

Figure 6. The derivative of the northern sinking with respect to freshwater forcing in the full case. The derivative is positive before the circulation collapses (white area). This behaviour is caused by a change in the Southern Ocean temperature from $T_S = 7\,°C$ to $T_S = 5\,°C$.

The polynomial consists of two terms of opposing sign: the first term on the left depends on the change of pycnocline depth (representing the vertical density structure) with increasing freshwater flux. Since this derivative, $\frac{\partial D}{\partial F_N}$, is generally positive the full term is negative. The second term is positive since the horizontal density difference in the Southern Ocean declines when F_N is increased. The sign of the derivative of the northern sinking is determined by the ratio between the two terms. Thus strongly increasing pycnocline depth, i.e. strong positive changes in vertical density structure, shift the overturning to a deceasing threshold behaviour. If the southern meridional density difference decreases stronger (in absolute values), then the overturning rises under freshwater forcing. The crucial point is that the absolute value of pycnocline depth is present in the term with the derivative of southern meridional density difference. That means rising pycnocline depth also amplifies the term that depends on horizontal density structure and vice versa for the meridional density difference. A stronger statement can be derived for the purely wind-driven case.

4.2 Wind-driven case

Upwelling in the lower latitudes amplifies the decreasing of northern sinking with increasing freshwater flow. Therefore, the wind-driven case provides a better example and a clearer image. Without low-latitudinal upwelling the derivative of northern sinking (Eq. 11) equals the negative derivative of the eddy return flow ($\mathrm{d}m_\mathrm{N}/\mathrm{d}F_N = -\mathrm{d}m_\mathrm{E}/\mathrm{d}F_N$). From the scaling of the eddy return flow (Eq. 4) and the derivative of the southern horizontal density difference (Eq. 10) the derivative of the northern sinking emerges.

$$\frac{\mathrm{d}m_\mathrm{N}}{\mathrm{d}F_N} = -C_\mathrm{E}\Delta\rho_\mathrm{SO}\frac{\partial D}{\partial F_N} - C_\mathrm{E}D\frac{\partial \Delta\rho_\mathrm{SO}}{\partial F_N}$$
$$= C_\mathrm{E}\left(\frac{S_0\beta}{C_W}(F_N+F_S)+\alpha\Delta T_\mathrm{SO}\right)\frac{\partial D}{\partial F_N} + C_\mathrm{E}D\frac{S_0\beta}{C_W}$$

Now, solely the term depending on the negative southern density difference could diminish the derivative. For the values given in Table 1, $\frac{\partial D}{\partial F_N} \simeq \frac{100\,\mathrm{m}}{0.1\,\mathrm{Sv}}$, and $D \simeq 1000\,\mathrm{m}$, the derivative is far in the positive range ($\frac{\partial m_\mathrm{N}}{\partial F_N} \simeq 5000$). In order to calculate the critical derivative, we use again the fact that the southern density difference equals zero at the threshold.

$$\left(\frac{\mathrm{d}m_\mathrm{N}}{\mathrm{d}F_N}\right)_\mathrm{crit} = C_\mathrm{E}D_\mathrm{crit}\frac{S_0\beta}{C_W} > 0$$

The emerging critical derivative depends only on positive constants and the positive critical pycnocline depth, i.e. the overturning always increases close to the threshold. This result is not surprising in light of the heuristic explanation given above, but it is not trivial due to the still complex vertical and horizontal density dynamics.

5 Climate model experiments

In order to investigate the possibility of the occurrence of a freshwater induced AMOC strengthening in a more complex climate model experiments were performed with the University of Victoria Earth System Climate Model, version 2.9 (UVic ESCM). UVic ESCM 2.9 is a model of intermediate complexity, with a simple 1-dimensional atmosphere but a 3-dimensional dynamic ocean (Weaver et al., 2001; Eby et al., 2009). The model is forced with a constant freshwater flux, ranging from 0.025 to 0.2 Sv, and run to equilibrium over 4300 years. Freshwater was transferred from the southern Atlantic (10 to 30° S) into the northern Atlantic (10 to 30° N) in all simulations. The maximum overturning, averaged over 1000 years, increases for a freshwater forcing of 0.075 Sv before it declines at 0.1 Sv (see Fig. 7). It ceases under a freshwater forcing of 0.16 Sv or higher (see Fig. 7). The AMOC strengthening is less pronounced compared to the box model behaviour. However, due to the strong differences between the box model and UVic ESCM slightly different behaviours can be expected. Especially differences in the freshwater forcing and the lack of thermal feedbacks in the box model lead to different overturning behaviour between the models. In the simulations using UVic ESCM the freshwater forcing is applied via a constant surface salinity flux because the simulations are time dependent. In the box model, a set amount of salinity flux is applied to one whole box (negative in the northern box and positive in the lower box) because instantaneous equilibrium is assumed. Also, a comparison of the location of the freshwater forcing is not straight forward because the latitudinal extension of the boxes is not specified due to the conceptual nature of the

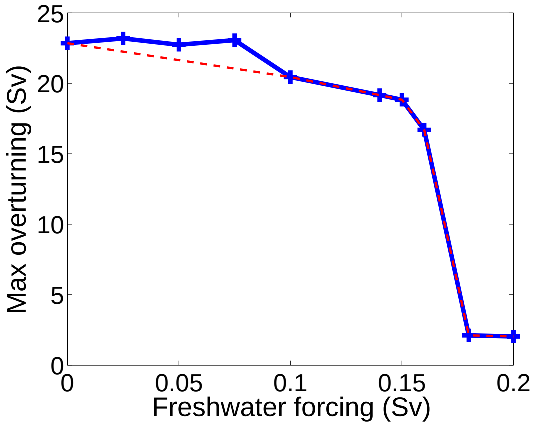

Figure 7. Maximum overturning for freshwater experiments with UVic ESCM 2.9. Freshwater was taken out at 10 to 30° S and dumped into the Atlantic at 10 to 30° N prescribing different constant amounts of freshwater. For each simulation the maximum overturning is averaged over the last 1000 years of the simulation (blue curve). Each blue cross corresponds to one equilibrium simulation. If the model would not show an AMOC strengthening, the modelled maximum overturning would be expected to follow the red curve.

box model. Furthermore, the southern ocean is not well represented in complex climate models, especially eddy flows. These experiments show that an increase of the AMOC under freshwater forcing is a possible behaviour of the overturning. However, further experiments would be needed to investigate the robustness of this behaviour.

6 Conclusion and discussion

The conceptual model of the Atlantic overturning presented here builds on a previous model (Fürst and Levermann, 2012) and advances the model by the introduction of a dynamic southern ocean density difference for the eddy return flow as imposed by comparison with comprehensive ocean model results (Levermann and Fürst, 2010). As a first result the model reproduces the qualitative result that a threshold behaviour is a robust feature that is independent of the driving mechanism, i.e. it is present in mixing a wind-driven as well as in a combined case. The regime of existence of a solution for the overturning for a specific parameter combination is defined by the simultaneous compliance of a number of conditions, e.g. positive volume fluxes and pycnocline depth. In the presented model the threshold is generally reached when the eddy return flow becomes negative. Similar to the predecessor of the model also here the threshold is associated with the salt-advection feedback. As suggested by Rahmstorf (1996), a threshold thus only exists when the salinity in the low-latitude box is higher than in the northern box. This is the case here (see Table 2). Whether the real ocean is in a bistable regime and thereby exhibits a threshold behaviour is of yet unclear. According to a diagnostic by Rahmstorf (1996), an overturning is bistable if the overturning carries a net salinity transport at 35° N. This diagnostic was confirmed to be valid in a comprehensive climate model (Dijkstra, 2007) and is discussed in depth by Hofmann and Rahmstorf (2009). Following this diagnostic most climate models do not show a threshold behaviour in earlier studies (Drijfhout et al., 2011). However, in a more recent model in-comparison study the majority of climate models do show a threshold behaviour (Weaver et al., 2012). Also observational data indicates that the real ocean is in a bistable regime (Bryden et al., 2011), i.e. the current circulation pattern could change after reaching a threshold. It should be noted that the model presented here does not capture an "off-state" of the circulation, i.e. describing a circulation pattern after the threshold in freshwater forcing has been crossed. There are models showing an inverse circulation, which is sometimes associated with the Antarctic Bottom Water filling up the Atlantic (Rahmstorf et al., 2005) and other models show a seemingly stagnant ocean (Stouffer et al., 2006b). Neither of these patterns would be properly captured by the physics that is incorporated in the conceptual model presented here. Therefore, a bistable situation can not be described but rather a threshold behaviour. This threshold behaviour shows that beyond a certain freshwater flux the circulation in the Atlantic cannot be captured by the conceptual model and is thereby not a classic overturning circulation as presently observed.

The main result is the observation that the overturning can increase prior to its collapse in response to a freshwater flux from low latitudes to high northern latitudes. Previous models including the base models (Johnson et al., 2007; Marzeion and Drange, 2006; Fürst and Levermann, 2012) show the opposite behaviour, similar to the bifurcation in the initial model of Stommel (1961). The emergence of the effect depends on the inclusion of Southern Ocean winds as a driving mechanism for the overturning and the inclusion of a dynamic southern ocean horizontal density difference. It thus does not include in the mixing-driven case. Thus our model has opposite behaviour prior to reaching the threshold depending on whether the circulation is wind or mixing driven.

This has strong implications for potential monitoring systems that aim to detect the vicinity to the threshold. Methods that depend on the decline of the overturning prior to the threshold for example in order to detect an increase in variability might not be suitable in a situation (Lenton, 2011; Scheffer et al., 2009) in which the presented mechanism is relevant. However, applicability of these findings for monitoring purposes are limited as the presented results refer to a system in equilibrium, and not a time dependent state as we see under current global warming.

Whether the mechanism described here is dominant in the real ocean is beyond the scope of this paper. This study presents the physical processes which need to be investigated with comprehensive quantitative models and verified against observation in order to assess its relevance. Though a large number of so-called water hosing experiments have been carried out (e.g. Manabe and Stouffer, 1995; Rahmstorf et al., 2005; Stouffer et al., 2007), few studies have focussed on freshwater transport from low to high latitudes. We were able to show a strengthening of the AMOC under freshwater forcing prior to a decline of the overturning by prescribing different amounts of constant freshwater transport from low latitudes in the southern Atlantic into the northern Atlantic. However, this behaviour is not strongly pronounced. Thus further experiments are needed in order to find whether the mechanism is indeed relevant for the real ocean.

Appendix A: Analytical calculations

A1 Full case

The salinities are exchanged by salinity differences between the boxes except the salinity of the northern box. The new salinity variables are defined as $\Delta S = S_N - S_U$, $\Delta S_D = S_N - S_D$, $\Delta S_{\mathrm{SO}} = S_S - S_U$, and S_N. The salinity balance of the northern box gives for the northern salinity difference:

$$\Delta S = -\frac{S_0 F_N}{m_N}. \tag{A1}$$

The scaling of the northern sinking (Eq. 1) with the linearly scaling of the meridional density difference $\Delta\rho = \beta\,\Delta S - \alpha\,\Delta T$ and Eq. (A1) yields a quadratic polynomial of m_{N}.

$$0 = m_{\mathrm{N}}^2 + m_{\mathrm{N}} C_N D^2 \alpha \Delta T + C_N D^2 \beta F_N S_0 \tag{A2}$$

It yields the following solution:

$$m_{\mathrm{N}} = -\frac{1}{2} C_N D^2 \alpha \Delta T \pm \sqrt{\frac{1}{4} C_N^2 D^4 \alpha^2 \Delta T^2 - C_N D^2 \beta F_N S_0}. \tag{A3}$$

The salinity balance of the upper box can be used to calculate ΔS_D:

$$\Delta S_D = \frac{m_{\mathrm{W}}}{m_{\mathrm{U}}} \Delta S_{\mathrm{SO}} + \Delta S + \frac{S_0}{m_{\mathrm{U}}} (F_N + F_S). \tag{A4}$$

The salinity balance of the southern box combined with Eq. (A4) results into an equation for the southern salinity difference.

$$\Delta S_{\mathrm{SO}} = -S_0 \frac{m_{\mathrm{W}} (F_N + F_S) + m_{\mathrm{U}} F_S}{m_{\mathrm{W}}^2 + m_{\mathrm{W}} m_{\mathrm{U}} + m_{\mathrm{E}} m_{\mathrm{U}}} \tag{A5}$$

The scaling of the eddy return flow (Eq. 4), the linear density function for southern meridional density difference ($\Delta\rho_{\mathrm{SO}} = \beta\,\Delta S_{\mathrm{SO}} - \alpha\,\Delta T_{\mathrm{SO}}$), and Eq. (A5) can be collapsed into a quadratic equation for m_{E}.

$$0 = m_{\mathrm{E}} + C_{\mathrm{E}} D \beta S_0 \frac{m_{\mathrm{W}} (F_N + F_S) + m_{\mathrm{U}} F_S}{m_{\mathrm{W}}^2 + m_{\mathrm{W}} m_{\mathrm{U}} + m_{\mathrm{E}} m_{\mathrm{U}}} + C_{\mathrm{E}} D \alpha \Delta T_{\mathrm{SO}} \tag{A6}$$

It yields the following solution:

$$m_{\mathrm{E}} = -\frac{1}{2}\left(\frac{m_{\mathrm{W}}^2}{m_{\mathrm{U}}} + m_{\mathrm{W}} + C_{\mathrm{E}} D \alpha \Delta T_{\mathrm{SO}}\right) + \sqrt{\begin{array}{l} \frac{1}{4}\left(\frac{m_{\mathrm{W}}^2}{m_{\mathrm{U}}} + m_{\mathrm{W}} + C_{\mathrm{E}} D \alpha \Delta T_{\mathrm{SO}}\right)^2 - C_{\mathrm{E}} D \beta S_0 \left(\frac{m_{\mathrm{W}}}{m_{\mathrm{U}}} (F_N + F_S) + F_S\right) \\ -C_{\mathrm{E}} D \alpha \Delta T_{\mathrm{SO}} \left(\frac{m_{\mathrm{W}}^2}{m_{\mathrm{U}}} + m_{\mathrm{W}}\right) \end{array}}.$$

The governing equation of the pycnocline depth emerges by using Eq. (A6) and replacing the eddy return flow by $m_{\mathrm{E}} = m_{\mathrm{U}} + m_{\mathrm{W}} - m_{\mathrm{N}}$, m_{N}^2 by Eq. (A2), and the upwelling transport processes, m_{U} and m_{W}, by their scaling (Eqs. 2 and 3).

$$\begin{aligned}
& D^{10} C_{\mathrm{E}} C_U C_W C_N^2 \alpha^2 \Delta T^2 [S_0 \beta (F_N + F_S) + C_W \alpha \Delta T_{\mathrm{SO}}] \\
& + D^9 C_N \alpha \Delta T \Big[C_{\mathrm{E}} C_W \Big(C_W^2 + C_{\mathrm{E}} C_U \alpha \Delta T_{\mathrm{SO}} \Big) \\
& (F_N S_0 \beta + F_S S_0 \beta + C_W \alpha \Delta T_{\mathrm{SO}}) \\
& + C_U C_N \Big(C_W^2 F_N S_0 \beta + C_W^3 \alpha \Delta T + 2 C_{\mathrm{E}} C_U C_W \alpha^2 \Delta T \Delta T_{\mathrm{SO}} \\
& + C_{\mathrm{E}} C_U S_0 \alpha \beta (F_S \Delta T + F_N \Delta T_{\mathrm{SO}}))] \\
& + D^8 \Big[C_{\mathrm{E}}^2 C_W^2 (F_N S_0 \beta + F_S S_0 \beta + C_W \alpha \Delta T_{\mathrm{SO}})^2 \\
& + C_N^2 C_U^2 \Big(F_N^2 S_0^2 \beta^2 + 3 C_W F_N S_0 \alpha \beta \Delta T + \alpha^2 \Delta T^2 \\
& \Big(3 C_W^2 + C_{\mathrm{E}} C_U \alpha \Delta T_{\mathrm{SO}} \Big) \Big) + C_N \Big(C_W^4 F_N S_0 \beta + C_W^5 \alpha \Delta T \\
& + C_{\mathrm{E}} C_U C_W^2 S_0 \alpha \beta \Delta T (3 F_N + 4 F_S) + 6 C_{\mathrm{E}} C_U C_W^3 \alpha^2 \Delta T \Delta T_{\mathrm{SO}} \\
& + C_{\mathrm{E}}^2 C_U^2 S_0 \alpha^2 \beta \Delta T_{\mathrm{SO}} (F_S \Delta T + F_N \Delta T_{\mathrm{SO}}) + 2 C_{\mathrm{E}} C_U C_W \\
& \Big(-F_N^2 S_0^2 \beta^2 - F_N F_S S_0^2 \beta^2 + C_{\mathrm{E}} C_U \alpha^3 \Delta T \Delta T_{\mathrm{SO}}^2 \Big) \Big) \Big] \\
& + D^7 \Big[C_N 2 C_U^3 \alpha \Delta T (2 F_N S_0 \beta + 3 C_W \alpha \Delta T) + 2 C_{\mathrm{E}} C_W \\
& (F_N S_0 \beta + F_S S_0 \beta + C_W \alpha \Delta T_{\mathrm{SO}}) \Big(C_W^3 + C_{\mathrm{E}} C_U F_S S_0 \beta \Big) \\
& + 2 C_{\mathrm{E}} C_U C_W \alpha \Delta T_{\mathrm{SO}} + C_N C_U \Big(4 C_W^3 F_N S_0 \beta + 6 C_W^4 \alpha \Delta T \\
& + 12 C_{\mathrm{E}} C_U C_W^2 \alpha^2 \Delta T \Delta T_{\mathrm{SO}} + C_{\mathrm{E}} C_U \Big(-2 F_N F_S S_0^2 \beta^2 + C_{\mathrm{E}} C_U \alpha^3 \Big) \Big) \Big] \\
& \Delta T \Delta T_{\mathrm{SO}}^2 \Big) + C_{\mathrm{E}} C_U C_W S_0 \alpha \beta (5 F_S \Delta T + 2 F_N (\Delta T + \Delta T_{\mathrm{SO}}) \\
& + D^6 \Big[C_W^6 + 2 C_U C_W^3 (C_{\mathrm{E}} S_0 \beta (3 F_N + 4 F_S) + 7 C_N C_U \alpha \Delta T) \\
& + 10 C_{\mathrm{E}} C_U C_W^4 \alpha \Delta T_{\mathrm{SO}} + 2 C_{\mathrm{E}} C_U^2 C_W \alpha \Delta T_{\mathrm{SO}} (C_{\mathrm{E}} S_0 \beta \\
& (F_N + 3 F_S) + 5 C_N C_U \alpha \Delta T) + C_U^2 C_W^2 (7 C_N F_N S_0 \beta \\
& + 6 C_{\mathrm{E}}^2 \alpha^2 \Delta T_{\mathrm{SO}}^2 \Big) + C_U^2 \Big(C_{\mathrm{E}}^2 F_S^2 S_0^2 \beta^2 + C_N^2 C_U^2 \alpha^2 \Delta T^2 \\
& + 2 C_{\mathrm{E}} C_N C_U S_0 \alpha \beta (F_S \Delta T + F_N \Delta T_{\mathrm{SO}}))] + D^5 C_U \\
& \Big[6 C_W^5 + 2 C_U C_W^2 (3 C_{\mathrm{E}} S_0 \beta (F_N + 2 F_S) + 8 C_N C_U \alpha \Delta T) \\
& + 20 C_{\mathrm{E}} C_U C_W^3 \alpha \Delta T_{\mathrm{SO}} + C_{\mathrm{E}} C_U^2 \alpha \Delta T_{\mathrm{SO}} (2 C_{\mathrm{E}} F_S S_0 \beta \\
& + 3 C_N C_U \alpha \Delta T) + 2 C_U^2 C_W \Big(3 C_N F_N S_0 \beta + 2 C_{\mathrm{E}}^2 \alpha^2 \Delta T_{\mathrm{SO}}^2 \Big) \Big] \\
& + D^4 C_U^2 \Big[15 C_W^4 + C_U C_W (2 C_{\mathrm{E}} S_0 \beta (F_N + 4 F_S) + 9 C_N C_U \\
& \alpha \Delta T) + 20 C_{\mathrm{E}} C_U C_W^2 \alpha \Delta T_{\mathrm{SO}} + C_U^2 \Big(2 C_N F_N S_0 \beta + C_{\mathrm{E}}^2 \alpha^2 \\
& \Delta T_{\mathrm{SO}}^2 \Big) \Big] + D^3 2 C_U^3 \Big[10 C_W^3 + C_U C_{\mathrm{E}} F_S S_0 \beta + C_N C_U^2 \alpha \Delta T \\
& + 5 C_{\mathrm{E}} C_U C_W \alpha \Delta T_{\mathrm{SO}}] + D^2 C_U^4 \Big[15 C_W^2 + 2 C_{\mathrm{E}} C_U \alpha \Delta T_{\mathrm{SO}} \Big] \\
& + D 6 C_U^5 C_W + C_U^6 = 0
\end{aligned} \tag{A7}$$

A2 Wind-driven case

For a wind-driven overturning, the upwelling in the lower latitudes equals zero by setting C_U to zero. Thus the tracer transport balance in steady state (Eq. 5 equal to zero) reduces to $m_\mathrm{W} = m_\mathrm{N} + m_\mathrm{E}$. Differences in salinity are defined as in the full problem and the salinity balance in the northern box is the same as in the full problem. Therefore Eqs. (A1)–(A3) are valid. Using the salinity balance of the southern box, in this case the southern salinity difference reduces to the following:

$$\Delta S_\mathrm{SO} = -\frac{S_0\,(F_N + F_S)}{m_\mathrm{W}}.$$

The eddy return flow is then represented by the following equation:

$$m_\mathrm{E} = -C_\mathrm{E} D\beta \frac{S_0\,(F_N + F_S)}{C_W} - \alpha\Delta T_\mathrm{SO} C_\mathrm{E} D. \tag{A8}$$

By replacing the northern sinking with Eq. (A3) and the eddy return flow with Eq. (A8) in the tracer transport balance, the governing equation of the pycnocline depth emerges.

$$\begin{aligned} &D^3 C_\mathrm{E} C_N \alpha\Delta T \left[\frac{\beta S_0\,(F_N+F_S)}{C_W} + \alpha\Delta T_\mathrm{SO}\right] \\ &+ D^2 \left[C_N F_N S_0 \beta + C_N C_W \alpha\Delta T + \frac{C_\mathrm{E}^2}{C_W^2}\,(S_0\beta\,(F_N+F_S) \right.\\ &\left. + C_W \alpha\Delta T_\mathrm{SO})^2\right] + D2C_\mathrm{E}\,[\beta S_0\,(F_N+F_S) + C_W\alpha\Delta T_\mathrm{SO}] \\ &+ C_W^2 = 0 \end{aligned} \tag{A9}$$

The solutions of the polynomial depend on the sign of the discriminant $d = (q/2)^2 + (p/3)^3$ with p and q defined as follows:

$$\begin{aligned} \frac{q}{2} &= \frac{1}{2}\left(\frac{C_N F_N S_0\beta + C_N C_W \alpha\Delta T + C_\mathrm{E}^2 A^2}{3C_\mathrm{E}\alpha\Delta T A}\right)^3 \\ &- \frac{C_W F_N S_0 \beta + C_W^2 \alpha\Delta T}{3C_\mathrm{E} C_N \alpha^2 \Delta T^2 A} - \frac{C_\mathrm{E} C_W A}{3C_N^2 \alpha^2 \Delta T^2} \\ &+ \frac{C_W^2}{2C_\mathrm{E} C_N \alpha\Delta T A} \\ \frac{p}{3} &= \frac{6C_W C_N \alpha\Delta T - 1}{9C_N^2 \alpha^2 \Delta T^2} - \frac{F_N S_0 \beta + C_W \alpha\Delta T}{9C_\mathrm{E}^2 C_N^2 \alpha^2 \Delta T^2 A^2} \end{aligned}$$

with $A = \left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)$.

If the discriminant is positive the governing equation has one real solution.

$$D = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}} \tag{A10}$$

For a negative discriminant there are three real solutions.

$$\begin{aligned} D_1 =& 2\sqrt{-\frac{p}{3}}\cos\left(\frac{1}{3}\arccos\left(-\frac{3q}{2p\sqrt{-\frac{p}{3}}}\right)\right) \\ &- \frac{C_N F_N S_0\beta + C_N C_W \alpha\Delta T}{3C_\mathrm{E} C_N \alpha\Delta T\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)} \\ &+ \frac{C_\mathrm{E}^2\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)^2}{3C_\mathrm{E} C_N \alpha\Delta T\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)} \\ &- \frac{C_N F_N S_0\beta + C_N C_W \alpha\Delta T + C_\mathrm{E}^2\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)^2}{3C_\mathrm{E} C_N \alpha\Delta T\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)} \\ D_2 =& 2\sqrt{-\frac{p}{3}}\cos\left(\frac{1}{3}\arccos\left(-\frac{3q}{2p\sqrt{-\frac{p}{3}}}\right) + \frac{2}{3}\pi\right) \\ &- \frac{C_N F_N S_0\beta + C_N C_W \alpha\Delta T}{3C_\mathrm{E} C_N \alpha\Delta T\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)} \\ &+ \frac{C_\mathrm{E}^2\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)^2}{3C_\mathrm{E} C_N \alpha\Delta T\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)} \\ &- \frac{C_N F_N S_0\beta + C_N C_W \alpha\Delta T + C_\mathrm{E}^2\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)^2}{3C_\mathrm{E} C_N \alpha\Delta T\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)} \\ D_3 =& 2\sqrt{-\frac{p}{3}}\cos\left(\frac{1}{3}\arccos\left(-\frac{3q}{2p\sqrt{-\frac{p}{3}}}\right) + \frac{4}{3}\pi\right) \\ &- \frac{C_N F_N S_0\beta + C_N C_W \alpha\Delta T}{3C_\mathrm{E} C_N \alpha\Delta T\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)} \\ &+ \frac{C_\mathrm{E}^2\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)^2}{3C_\mathrm{E} C_N \alpha\Delta T\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)} \\ &- \frac{C_N F_N S_0\beta + C_N C_W \alpha\Delta T + C_\mathrm{E}^2\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)^2}{3C_\mathrm{E} C_N \alpha\Delta T\left(\frac{S_0\beta}{C_W}\,(F_N+F_S) + \alpha\Delta T_\mathrm{SO}\right)} \end{aligned} \tag{A11}$$

Acknowledgements. D. Ehlert would like to thank M. Eby for the support in setting up simulations with UVic ESCM. The research leading to these results has received funding from the European Union Seventh Framework Program FP7/2007-2013 under grant agreement No. 603864.

Edited by: H. A. Dijkstra

References

Bryden, H. L., King, B. A., and McCarthy, G. D.: South Atlantic overturning circulation at 24° S, J. Mar. Res., 69, 38–55, 2011.

Cimatoribus, A. A., Drijfhout, S. S., and Dijkstra, H. A.: Meridional overturning circulation: stability and ocean feedbacks in a box model, Clim. Dynam., 42, 311–328, 2014.

Dijkstra, H. A.: Characterization of the multiple equilibria regime in a global ocean model, Tellus A, 59, 695–705, 2007.

Drijfhout, S. S., Weber, S. L., and van der Swaluw, E.: The stability of the MOC as diagnosed from model projections for pre-industrial, present and future climates, Clim. Dynam., 37, 1575–1586, doi:10.1007/s00382-010-0930-z, 2011.

Eby, M., Zickfeld, K., Montenegro, A., Archer, D., Meissner, K. J., and Weaver, A. J.: Lifetime of Anthropogenic Climate Change: Millennial Time Scales of Potential CO_2 and Surface Temperature Perturbations, J. Climate, 22, 2501–2511, doi:10.1175/2008JCLI2554.1, 2009.

Fürst, J. and Levermann, A.: Minimal model of a wind- and mixing-driven overturning – threshold behaviour for both driving mechanisms, Clim. Dynam., 38, 239–260, doi:10.1007/s00382-011-1003-7, 2012.

Gent, P. R. and McWilliams, J. C.: Isopycnal Mixing in Ocean Circulation Models, J. Phys. Oceanogr., 20, 150–155, 1990.

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

Goswami, B. N., Madhusoodanan, M. S., Neema, C. P., and Sengupta, D.: A physical mechanism for North Atlantic SST influence on the Indian summer monsoon, Geophys. Res. Lett., 33, L02706, doi:10.1007/s00382-011-1003-7, 2006.

Griesel, A. and Morales-Maqueda, M. A.: The relation of meridional pressure gradients to North Atlantic Deep Water volume transport in an OGCM, Clim. Dynam., 26, 781–799, 2006.

Hawkins, E., Smith, R. S., Allison, L. C., Gregory, J. M., Woollings, T. J., Pohlmann, H., and De Cuevas, B.: Bistability of the Atlantic overturning circulation in a global climate model and links to ocean freshwater transport, Geophys. Res. Lett., 38, 1–6, 2011.

Hofmann, M. and Rahmstorf, S.: On the stability of the Atlantic meridional overturning circulation, P. Natl. Acad. Sci., 106, 20584–20589, 2009.

Huang, R. X.: Mixing and Energetics of the Oceanic Thermohaline Circulation, J. Phys. Oceanogr., 29, 727–746, 1999.

Huisman, S. E., den Toom, M., Dijkstra, H. A., and Drijfhout, S.: An Indicator of the Multiple Equilibria Regime of the Atlantic Meridional Overturning Circulation, J. Phys. Oceanogr., 40, 551–567, 2010.

Johnson, H. L., Marshall, D. P., and Sproson, D. A. J.: Reconciling theories of a mechanically-driven meridional overturning circulation with thermohaline forcing and multiple equilibria, Clim. Dynam., 29, 821–836, 2007.

Kuhlbrodt, T., Griesel, A., Montoya, M., Levermann, A., Hofmann, M., and Rahmstorf, S.: On the driving processes of the Atlantic meridional overturning circulation, Rev. Geophys., 45, RG2001, doi:10.1029/2004RG000166, 2007.

Kuhlbrodt, T., Rahmstorf, S., Zickfeld, K., Vikebo, F. B., Sundby, S., Hofmann, M., Link, P. M., Bondeau, A., Cramer, W., and Jaeger, C.: An Integrated Assessment of changes in the thermohaline circulation, Climatic Change, 96, 489–537, 2009.

Laurian, A., Drijfhout, S. S., Hazeleger, W., and Hurk, B.: Response of the Western European climate to a collapse of the thermohaline circulation, Clim. Dynam., 34, 689–697, 2009.

Lenton, T. M.: Early warning of climate tipping points, Nat. Clim. Change, 1, 201–209, 2011.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., and Schellnhuber, H. J.: Tipping elements in the Earth's climate system, P. Natl. Acad. Sci., 105, 1786–1793, 2008.

Levermann, A. and Fürst, J.: Atlantic pycnocline theory scrutinized using a coupled climate model, Geophys. Res. Lett., 37, L14602, doi:10.1029/2010GL044180, 2010.

Levermann, A. and Griesel, A.: Solution of a Model for the Oceanic Pycnocline Depth: Scaling of Overturning Strength and Meridional Pressure Difference, Geophys. Res. Lett., 31, L17302, doi:10.1029/2004GL020678, 2004.

Levermann, A., Griesel, A., Hofmann, M., Montoya, M., and Rahmstorf, S.: Dynamic sea level changes following changes in the thermohaline circulation, Clim. Dynam., 24, 347–354, 2005.

Levermann, A., Bamber, J., Drijfhout, S., Ganopolski, A., Haeberli, W., Harris, N., Huss, M., Krueger, K., Lenton, T., Lindsay, R., Notz, D., Wadhams, P., and Weber, S.: Potential climatic transitions with profound impact on Europe, Climatic Change, 110, 845–878, 2012.

Manabe, S. and Stouffer, R. J.: Two stable equilibria of a coupled ocean–atmosphere model, J. Climate, 1, 841–866, 1988.

Manabe, S. and Stouffer, R. J.: Century-scale effects of increased atmospheric CO_2 on the ocean–atmosphere system, Nature, 364, 215–218, 1993.

Manabe, S. and Stouffer, R. J.: Simulation of abrupt climate change induced by freshwater input to the North Atlantic Ocean, Nature, 378, 165–167, 1995.

Marotzke, J.: Boundary Mixing and the Dynamics of Three-Dimensional Thermohaline Circulations, J. Phys. Oceanogr., 27, 1713–28, 1997.

Marzeion, B. and Drange, H.: Diapycnal mixing in a conceptual model of the Atlantic Meridional Overturning Circulation, Deep-Sea Res. Pt. II, 53, 226–238, 2006.

Mignot, J., Ganopolski, A., and Levermann, A.: Atlantic subsurface temperatures: response to a shut-down of the overturning circulation and consequences for its recovery, J. Climate, 20, 4884–4898, 2007.

Munk, W.: Abyssal recipes, Deep-Sea Res., 13, 707–730, 1966.

Munk, W. and Wunsch, C.: Abyssal recipes II, Deep-Sea Res. Pt. I, 45, 1977–2010, 1998.

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

Rahmstorf, S., Crucifix, M., Ganopolski, A., Goosse, H., Kamenkovich, I., Knutti, R., Lohmann, G., Marsh, B., Mysak, L. A., Wang, Z., and Weaver, A.: Thermohaline circulation hysteresis: A model intercomparison, Geophys. Res. Lett., 32, L23605, doi:10.1029/2005GL023655, 2005.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M., and Sugihara, G.: Early-warning signals for critical transitions, Nature, 461, 53–59, 2009.

Schewe, J. and Levermann, A.: The role of meridional density differences for a wind-driven overturning circulation, Clim. Dynam., 34, 547–556, 2010.

Schmittner, A.: Decline of the marine ecosystem caused by a reduction in the Atlantic overturning circulation, Nature, 434, 628–633, 2005.

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

Stouffer, R. J., Russell, J., and Spelman, M. J.: Importance of oceanic heat uptake in transient climate change, Geophys. Res. Lett., 33, 1–5, 2006a.

Stouffer, R. J., Yin, J., Gregory, J. M., Dixon, K. W., Spelman, M. J., Hurlin, W., Weaver, A. J., Eby, M., Flato, G. M., Hasumi, H., Hu, A., Jungclaus, J. H., Kamenkovich, I. V., Levermann, A., Montoya, M., Murakami, S., Nawrath, S., Oka, A., Peltier, W. R., Robitaille, D. Y., Sokolov, A. P., Vettoretti, G., and Weber, S. L.: Investigating the Causes of the Response of the Thermohaline Circulation to Past and Future Climate Changes, J. Climate, 19, 1365–1387, 2006b.

Stouffer, R. J., Seidov, B., and Haupt, B. J.: Climate Response to External Sources of Freshwater: North Atlantic versus the Southern Ocean, J. Climate, 20, 436–448, 2007.

Toggweiler, J. R. and Samuels, B.: Effect of Drake Passage on the global thermohaline circulation, Deep-Sea Res., 42, 477–500, 1995.

Toggweiler, J. R. and Samuels, B.: On the ocean's large scale circulation in the limit of no vertical mixing, J. Phys. Oceanogr., 28, 1832–1852, 1998.

Vellinga, M. and Wood, R. A.: Global climatic impacts of a collapse of the Atlantic thermohaline circulation, Climatic Change, 54, 251–267, 2002.

Weaver, A. J., Eby, M., Wiebe, E. C., Bitz, C. M., Duffy, P. B., Ewen, T. L., Fanning, A. F., Holland, M. M., MacFadyen, A., Matthews, H. D., Meissner, K. J., Saenko, O., Schmittner, A., Wang, H., and Yoshimori, M.: The UVic earth system climate model: Model description, climatology, and applications to past, present and future climates, Atmos.-Ocean, 39, 361–428, doi:10.1080/07055900.2001.9649686, 2001.

Weaver, A. J., Sedláček, J., Eby, M., Alexander, K., Crespin, E., Fichefet, T., Philippon-Berthier, G. L., Joos, F., Kawamiya, M., Matsumoto, K., Steinacher, M., Tachiiri, K., Tokos, K., Yoshimori, M., and Zickfeld, K.: Stability of the Atlantic meridional overturning circulation: A model intercomparison, Geophys. Res. Lett., 39, L20709, doi:10.1029/2012GL053763, 2012.

Wunsch, C. and Ferrari, R.: Vertical mixing, energy and the general circulation of the oceans, Annu. Rev. Fluid Mech., 36, 281–314, 2004.

Yin, J., Schlesinger, M. E., and Stouffer, R. J.: Model projections of rapid sea-level rise on the northeast coast of the United States, Nat. Geosci., 2, 262–266, 2009.