

A minimal model for wind- and mixing-driven overturning: threshold behavior for both driving mechanisms

Johannes J. Fürst · Anders Levermann

Received: 3 July 2010 / Accepted: 20 January 2011 / Published online: 12 February 2011
© Springer-Verlag 2011

Abstract We present a minimal conceptual model for the Atlantic meridional overturning circulation which incorporates the advection of salinity and the basic dynamics of the oceanic pycnocline. Four tracer transport processes following Gnanadesikan in Science 283(5410):2077–2079, (1999) allow for a dynamical adjustment of the oceanic pycnocline which defines the vertical extent of a mid-latitude box. At the same time the model captures the salt-advection feedback (Stommel in Tellus 13(2):224–230, (1961)). Due to its simplicity the model can be solved analytically in the purely wind- and purely mixing-driven cases. We find the possibility of abrupt transition in response to surface freshwater forcing in both cases even though the circulations are very different in physics and geometry. This analytical approach also provides expressions for the critical freshwater input marking the change in the dynamics of the system. Our analysis shows that including the pycnocline dynamics in a salt-advection model causes a decrease in the freshwater sensitivity of its northern sinking up to a threshold at which the circulation breaks down. Compared to previous studies the model is restricted to the essential ingredients. Still, it exhibits a rich behavior which reaches beyond the scope of this study and might be used as a

paradigm for the qualitative behaviour of the Atlantic overturning in the discussion of driving mechanisms.

Keywords Meridional overturning circulation · Northern sinking · Critical freshwater threshold · Overturning sensitivity · Conceptual model · Stability · Atlantic meridional overturning circulation · Pycnocline depth · Driving mechanism

1 Introduction

The Atlantic meridional overturning circulation (AMOC), a crucial branch of the global oceanic circulation system, transports large amounts of heat towards high latitudes. A cessation of this circulation would reduce temperatures regionally in the Nordic Seas by up to 8°C (Manabe and Stouffer 1988; Clark et al. 2002) with strong climatic implications world-wide (Laurian et al. 2009). These include changes in precipitation patterns (Vellinga and Wood Schmittner 2002), Atlantic ecosystems (Schmittner 2005; Kuhlbrodt et al. 2009), sea level distribution (Levermann et al. 2005; Yin et al. 2009), European climate (Laurian et al. 2010), the El Niño Southern Oscillation (Timmermann et al. 2005) and Asian monsoon systems (Goswami et al. 2006).

Since the initial study by Stommel (1961) the discussion about past and future variations of the AMOC is linked to the existence of multiple stable equilibria of the circulation. Across a large spectrum of climate models, existence of multiple states has been observed in conceptual models (Stommel 1961; Johnson et al. 2007; Guan and Huang 2008), ocean circulation models with idealised geometry (Marotzke et al. 1988; Marotzke and Willebrand 1991; Thual and McWilliams 1992; Rahmstorf 1995b), various

J. J. Fürst · A. Levermann
Earth System Analysis, Potsdam Institute for Climate Impact
Research and Institute of Physics, Potsdam University, Potsdam,
Germany
e-mail: anders.levermann@pik-potsdam.de

Present Address:
J. J. Fürst (✉)
Department of Geography and Earth System Sciences,
Faculty of Sciences, Vrije Universiteit Brussel, Pleinlaan 2,
Brussels, Belgium
e-mail: Johannes.Fuerst@vub.ac.be

Earth system Models of Intermediate Complexity (EMICs) (Manabe and Stouffer 1988; Rahmstorf et al. 2005; Yin and Stouffer 2007; Ashkenazy and Tziperman 2007) as well as uncoupled oceanic general circulation models (Rahmstorf 1996). As state-of-the-art coupled climate models are too computationally demanding to explore the full stability range of their circulation, no multi-stability under present day boundary conditions has yet been observed (Stouffer et al. 2006). Also some models of intermediate complexity are reported to lack multi-stability of the AMOC (Prange et al. 2003; Nof et al. 2007). In these studies, it has been speculated that extensive diapycnal mixing might be the reason for multi-stability. Recently Hofmann and Rahmstorf (2009), showed that multi-stability is possible for a wind-driven overturning. They attributed the existence of multiple stable states to the Atlantic salinity distribution.

The dispute about the physical mechanism providing the necessary energy to sustain an overturning circulation (Kuhlbrodt et al. 2007) is thus a crucial aspect in the stability analysis of the AMOC. The two main candidates for these so-called driving mechanisms are diapycnal mixing (Jeffreys 1925; Munk and Wunsch 1998; Park 1999) and Southern Ocean wind divergence (Toggweiler and Samuels 1998; Gnanadesikan et al. 2005). In addition to characterising the driving mechanisms of the AMOC, other processes need to be considered for its stability analysis. While surface fluxes of freshwater and heat alone can not sustain a deep overturning circulation (Sandström 1916; Kuhlbrodt 2010), they are important to set the density structure of the ocean. This density structure determines how much of the available energy is indeed directed into a basin wide overturning circulation (Schewe and Levermann 2010). We combine thus the two main driving mechanisms of the AMOC with two limiting processes not providing net energy to the system, but shaping its spatial pattern following Gnanadesikan (1999). These four processes are complemented by the advection of salinity and thereby a dynamical equation for the meridional density gradient. This is substantial since Levermann and Griesel (2004) showed that some variations in the Atlantic overturning are not captured in Gnanadesikan's model. So far an analytically solvable model that comprises both driving mechanisms of the overturning is missing. In contrast to Johnson et al. (2007) who suggested a similar model, their focus was to study an inherent oscillation between on- and off-state of this circulation. Here we aim to provide a minimal model that allows to examine the on-state AMOC stability in a wind- and mixing driven case.

Analytical solutions are derived for the purely wind- and the purely mixing-driven circulation cases. Our results reveal a threshold behaviour with respect to surface freshwater forcing that is independent of the mechanism

powering the AMOC. In agreement with the results in Park (1999), the sensitivity of the overturning circulation to freshwater fluxes in the North Atlantic is reduced compared to Stommel (1961). This results from a compensating effect of the pycnocline dynamics that stabilises the overturning.

The paper is structured as follows: the model design and its idealised components are presented in Sect. 2. In this context the necessity of tracer advection (see Stommel (1961)) in the approach of Gnanadesikan (1999) as already proposed by Levermann et al. (2005) and Levermann and Fürst (2010) is emphasised. The main results are introduced in Sect. 3, where the model is analysed for three instructive cases. This is followed by an analysis of the freshwater sensitivity of the northern sinking (Sect. 4) A validation of our conceptual approach is conducted in Sect. 5 using the model of intermediate complexity CLIMBER-3 α (Montoya et al. 2005) for a qualitative intercomparison. We conclude in Sect. 6.

2 Model description

For a minimal model that comprises wind- and mixing-induced overturning we propose a standard interhemispheric geometry as illustrated in Fig. 1. It combines the four basic meridional tracer transport processes associated with the overturning circulation (Gnanadesikan 1999) and thereby describes changes in the meridional density structure. Since we find that changes in heat advection represent a second order effect compared to salt advection, we keep oceanic temperatures fixed. This enables analytic solutions in a number of cases and captures significant atmospheric and oceanic feedbacks for the overturning circulation. In principle the model can be easily generalized to account for the advection of temperature. As surface boundary conditions for salinity we apply constant freshwater fluxes with global zero mean. This is represented by two freshwater bridges from the upper mid-latitude box (subscript U) to the southern F_S and northern F_N box denoted by subscripts S and N, respectively.

2.1 Pycnocline dynamics

The representation of the dynamics of the oceanic pycnocline follows (Gnanadesikan 1999). Here we assume that the oceanic pycnocline depth D is represented by the vertical extent of the mid-latitude box. Its time evolution is given by four tracer transport processes

$$BL_U \cdot \frac{\partial D}{\partial t} = m_W + m_U - m_E - m_N. \quad (1)$$

Here B is the average, zonal extent of the Atlantic ocean basin, L_U is the meridional extent of the tropical boxes

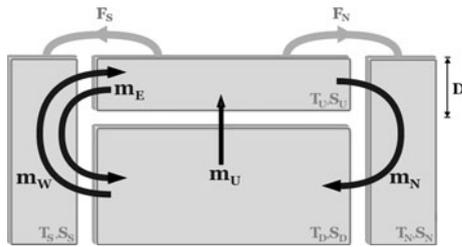


Fig. 1 Schematic depiction of the conceptual model. The depth of the pycnocline D is determined by the balance between northern deep water formation m_N , mixing driven upwelling in the low latitudes m_U , Ekman upwelling m_W and eddy-induced return flow m_E . Salinity is advected along with these transport processes and determines together with a fixed temperature distribution, the density difference between the northern and upper low-latitude box $\Delta\rho$

(between latitudes $\pm 30^\circ$). Using Ekman boundary layer theory, the wind driven volume transport is predicted via a scale analysis of the equation of motion.

$$m_W = B \cdot \frac{\tau_{Dr}}{|f_{Dr}| \rho_0} = C_W \tag{2}$$

The mean Coriolis parameter in the Drake passage is f_{Dr} while ρ_0 denotes the ocean average density. Wind stress feedbacks though possibly relevant (Fyfe et al. 2007; Toggweiler and Russell 2008) will not be captured by our model. Since there is no dependency on the oceans density stratification we substitute this flux by a constant $C_W = m_W$ for simplicity in later calculations. The eddy return flow m_E is parameterised following Gent and McWilliams (1990) using a thickness diffusivity A_{GM}

$$m_E = B \cdot \frac{A_{GM} D}{L_y^S} = C_E \cdot D \tag{3}$$

This flux is proportional to the meridional slope of the isopycnals which is approximated by the ratio of the meridional extent of outcropping L_y^S and the pycnocline depth D . The constants are again comprised within C_E to enhance legibility. Note that Levermann and Fürst (2010) recently showed that for capturing the AMOC changes under a CO2 increase scenario L_y^S would need to be varied. For these kind of changes in the geometry of the flow apart from changes in D , additional equations would be needed. Here we keep L_y^S constant. The mixing-driven low-latitudinal upwelling is described by a vertical advection-diffusion balance. Assuming an exponential density profile in the vertical yields

$$m_U = B \cdot L_U \cdot \frac{\kappa}{D} = \frac{C_U}{D} \tag{4}$$

where κ is the vertical diffusivity. This approach is frequently applied for the mixing induced upwelling in the ocean interior (e.g. Munk and Wunsch (1998) and references in Kuhlbrodt et al. (2007)).

Parameterisations of northern sinking have been derived following numerous approaches summarised in Appendix 1 (Robinson 1960; Marotzke 1997; Gnanadesikan 1999; Johnson and Marshall 2002; Guan and Huang 2008). In our model, the scaling is adopted from Marotzke (1997) who assumes a generic density distribution in the Atlantic. This allows to substitute the zonal density difference in the geostrophic equation with a meridional one (Marotzke 1997). A β -plane approximation finally gives

$$m_N = C \frac{g}{\beta_N L_y^N} \frac{\Delta\rho}{\rho_0} \cdot D^2 \equiv C_N \Delta\rho \cdot D^2 \tag{5}$$

The constant C is given by the current geometry and characteristics of the density distribution. Again all quantities except for D and $\Delta\rho$ are comprised in a constant C_N . We assume that the relevant density difference for the northern sinking is to be taken between the northern and the upper low-latitudinal boxes $\Delta\rho \equiv \rho_N - \rho_U$. This choice represents the most direct interpretation of the assumptions entering the derivation by Marotzke (1997). In its final form eq. (5) can furthermore be motivated in a more heuristic way: Currently major northern sinking occurs in the Nordic Seas. Thus its volume transport is mainly defined by the North Atlantic Current crossing the ocean basin from West to East. This flow is geostrophically balanced by a meridional difference in sea surface elevation which is observed in ocean models of varying resolution (Levermann et al. 2005; Landerer et al. 2007; Vellinga and Wood 2007; Schlesinger et al. 2006). Due to the existence of a level of no motion the sea surface elevation difference must be counteracted by a density difference in the upper levels (e.g. Griesel and Morales-Maqueda (2006)). This density difference is apparent in oceanic reanalysis data (Levitus 1982) though the scaling of the overturning circulation needs to be taken from the geostrophic argument and can not be directly observed due to lack of data. The quadratic dependence on the pycnocline depth originates from a vertical integration of the scaled geostrophic balance which is linear in D by use of the hydrostatic equation.

2.2 Salinity dynamics

In order to capture the salt-advection feedback deemed responsible for a possible multistability of the overturning circulation (Stommel 1961; Rahmstorf 1996), salinity advection is incorporated. Salinity changes are then linked to the pycnocline dynamics of Sect. 2.1 through the meridional density difference $\Delta\rho \equiv \rho_N - \rho_U$. For simplicity a linear equation of state $\Delta\rho = \rho_0(\beta_S \Delta S - \alpha_T \Delta\theta)$ is assumed, where α_T and β_S are the thermal and haline expansion coefficients. The time evolution for salinity then reads

$$\begin{aligned} \frac{\partial}{\partial t}(V_N S_N) &= m_N(S_U - S_N) - S_0 F_N \\ \frac{\partial}{\partial t}(V_U S_U) &= m_W \cdot S_S + m_U \cdot S_D - (m_N + m_E)S_U \\ &\quad + S_0(F_N + F_S) \\ \frac{\partial}{\partial t}(V_D S_D) &= m_N \cdot S_N + m_E \cdot S_S - (m_U + m_W)S_D \\ \frac{\partial}{\partial t}(V_S S_S) &= m_W(S_D - S_S) + m_E(S_U - S_S) - S_0 F_S. \end{aligned} \quad (6)$$

The volumes of the different boxes are computed via $V_N = B \cdot H \cdot L_N$, $V_U = B \cdot D \cdot L_U$, $V_D = B \cdot (H - D) \cdot L_U$, $V_S = B \cdot H \cdot L_S$, where H is the depth of the ocean.

2.3 Model equilibrium

Basing our model on eqs. (1) and (6) with parameters chosen from Table 1 we can numerically determine the equilibrium solution. The resulting circulation found after 40, 000 model years (Table 2) shows a mainly wind-driven northern sinking ($m_W = 13.0$ Sv) with some contribution of low-latitude upwelling ($m_U = 5.8$ Sv). The salinity distribution shows a southward salinity gradient. This means in the notation of Rahmstorf (1996) we are in a purely thermal state with salinity reducing northern

sinking. This is due to the positive northern freshwater bridge which will become evident in Sect. 3.

2.4 Parental models

Let us shortly recap how the two parental models emerge from the current one. In Gnanadesikan (1999) the north-south density difference $\Delta\rho$ is a constant and thus independent of the vertical density structure represented by the pycnocline D . Equation (1) is the same as the one used by Gnanadesikan (1999). However, the prognostic salinity eq. (6) need to be omitted. By prescribing a constant $\Delta\rho$, Gnanadesikan (1999) theory was able to explain the strong influence of surface boundary conditions for salinity and temperature on the overturning rate. In order to estimate the influence of changes in $\Delta\rho$ on the northern sinking, we consider the derivative

$$\frac{\partial m_N}{\partial \Delta\rho} = C_N D^2 + 2C_N \Delta\rho D \frac{\partial D}{\partial \Delta\rho}. \quad (7)$$

For the parameter set of Table 1, $\partial m_N / \partial \Delta\rho$ varies between $0.25 \text{ Sv} / (0.1 \text{ kg/m}^3)$ at a density difference of $\Delta\rho = 1.5 \text{ kg/m}^3$ and $3.03 \text{ Sv} / (0.1 \text{ kg/m}^3)$ at $\Delta\rho = 0.1 \text{ kg/m}^3$. Figure 2a depicts the relative change of m_N as a response to an increase of $\Delta\rho$ by 20%. Though this deviation of m_N

Table 1 Physical parameters used in our conceptual model

Variable name	Value	Unit	Description
Geometry			
H	$4 \cdot 10^3$	m	Average depth of the Atlantic ocean basin
B	$1 \cdot 10^7$	m	Average width of the Atlantic ocean basin
L_N	$3.34 \cdot 10^6$	m	Meridional extent of the northern box
L_U	$8.90 \cdot 10^6$	m	Meridional extent of the tropical box
L_S	$3.34 \cdot 10^6$	m	Meridional extent of the southern box
Stratification			
ρ_0	1027	kg/m^3	Average density of the Atlantic ocean
S_0	35	psu	Average salinity in the Atlantic ocean
L_y^S	$1.5 \cdot 10^6$	m	Meridional extent of the southern outcropping
L_y^N	$1.5 \cdot 10^6$	m	Meridional extent of the northern outcropping
A_{GM}	10^3	m^2/s	Thickness diffusivity
κ	$4 \cdot 10^{-5}$	m^2/s	Background vertical diffusivity
α_T	$2.1 \cdot 10^{-4}$	$1/^\circ\text{C}$	Thermal expansion coefficient for isobars
β_S	$8 \cdot 10^{-4}$	$1/\text{psu}$	Haline expansion coefficient for isobars
C	0.1	–	Constant accounting for geometry and stratification
External forcing			
β_N	$2 \cdot 10^{-11}$	$1/(\text{ms})$	Coefficient for β -plane approximation in the North Atlantic
f_{Dr}	$-7.5 \cdot 10^{-5}$	$1/\text{s}$	Coriolis parameter in the Drake Passage
τ_{Dr}	1.0	N/m^2	Average zonal wind stress in the Drake Passage
F_N	0.2	Sv	Northern meridional atmospheric freshwater transport
θ_N	5.0	$^\circ\text{C}$	Temperature of the Northern box
θ_U	12.5	$^\circ\text{C}$	Temperature of the tropical surface box
θ_S	5.0	$^\circ\text{C}$	Temperature of the southern box

Table 2 Equilibrium state obtained after 40,000 model years for eqs. (1) and (6) with parameters from Table 1

Variable name	Value	Unit	Variable name	Value	Unit
S_N	34.92	psu	m_N	14.7	Sv
S_U	35.39	psu	m_U	5.8	Sv
S_D	34.94	psu	m_W	13.0	Sv
S_S	35.05	psu	m_E	4.1	Sv
D	613	m	T_D	5.0	°C
$\Delta\rho$	1.23	kg/m ³			

The northern sinking is mainly fed by the Southern Ocean Ekman transport with some contribution from low-latitude upwelling. In the notation of (Rahmstorf 1996), the state is purely thermal, since the surface freshwater fluxes were chosen such that salinity reduces northern sinking

does not exceed 10% away from the singularity for $\Delta\rho$ (Fig. 2a), the significance of the density difference lies in its impact on the stability behaviour of the system.

Stommel (1961)'s original model is restricted to the advection of salt. Rahmstorf (1996) showed that no conceptual difference emerges if temperature advection is included together with surface restoring. In these models the size of the boxes is prescribed, which can be interpreted as a fixed pycnocline depth D . Consequently eq. (1) is omitted. The Stommel (1961) equation is then quickly derived by substituting $\Delta S = -S_0 F_N / m_N$ (eq. (10)) in the scaling for the northern sinking (eq. (5)). The resulting quadratic equation in m_N represents the bistability of the Atlantic overturning circulation found in a number of coarse resolution models (e.g. Manabe and Stouffer (1988);

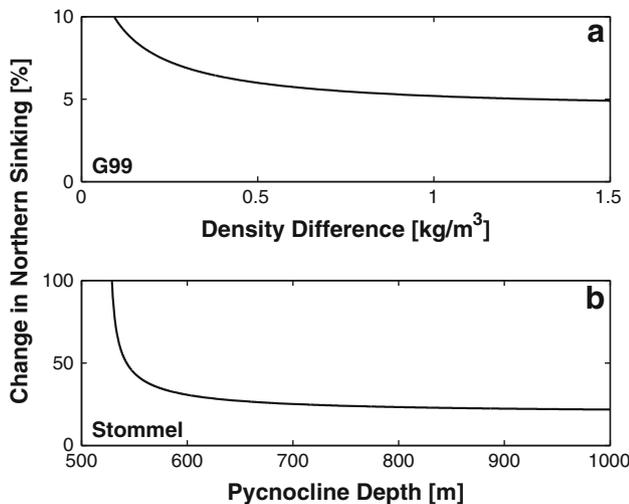


Fig. 2 Relative response of northern sinking due to an added density deviation of 20% in the [8] limit (a) and due to a correspondent 10% change in the pycnocline depth D in the [45] limit (b). These percentual changes differ since the northern sinking is linear proportional to $\Delta\rho$ but quadratic in D . For our parameter set, the Stommel case has no real solution when the pycnocline depth falls below 525 m (see eq. (8))

Rahmstorf (1995a); Rahmstorf et al. (2005)). In the present formulation the relevant solution reads

$$m_N = \frac{C_N D^2 \rho_0 \alpha_T |\Delta\theta|}{2} \left(1 + \sqrt{1 - \frac{4\beta_S S_0 F_N}{C_N \rho_0 D^2 \alpha_T^2 \Delta\theta^2}} \right). \quad (8)$$

In contrast to Rahmstorf (1996) the relevant density difference for the northern sinking is chosen to be the one between the northern and the low-latitude box. Consequently, the relevant freshwater bridge for the bistability of m_N is the one in the North not in the South. Note that the Stommel model does not capture any process in the Southern Ocean. The influence of density stratification changes in Stommel's approach can be inferred from the derivative

$$\frac{\partial m_N}{\partial D} = 2C_N \Delta\rho D + C_N D^2 \frac{\partial \Delta\rho}{\partial D}. \quad (9)$$

This derivative is positive and varies from 8.1 Sv/(100 m) at $D = 550$ m up to 10.4 Sv/(100 m) for large $D = 1,000$ m. The relative response of m_N for a 10% increase in the pycnocline depth reveals that this effect is almost one order of magnitude higher than in Gnanadesikan's limit for analog changes in $\Delta\rho$ (cf. Fig. 2a and b). Such changes in D therefore have an impact on the overturning rate m_N which is of the same order of magnitude than m_N itself. Consequently, since both variables D and $\Delta\rho$ are linked, neglecting their mutual dependence confines the physical applicability of both parental models.

3 Governing equation for equilibrium

Using the salinity balance eq. (6) in steady state we eliminate the meridional density difference $\Delta\rho$ from eq. (1) in order to obtain a governing equation for the oceanic pycnocline that allows for a salt-advection feedback. The salinity balance of the northern box yields

$$\Delta S = -\frac{S_0 F_N}{m_N} = -\frac{S_0 F_N}{C_U/D + C_W - C_{ED}} \quad (10)$$

which links the salinity difference $\Delta S \equiv S_N - S_U$ to the pycnocline depth. Substitution in eq. (5) in combination with eq. (1) yields the full governing equation of the model expressed in D .

$$\begin{aligned} & -D^5 \cdot C_N \rho_0 C_E \alpha_T \Delta\theta \\ & + D^4 \cdot (C_E^2 + C_N \rho_0 C_W \alpha_T \Delta\theta + C_N \rho_0 \beta_S S_0 F_N) \\ & + D^3 \cdot (C_N \rho_0 C_U \alpha_T \Delta\theta - 2C_E C_W) \\ & + D^2 \cdot (C_W^2 - 2C_U C_E) \\ & + D \cdot 2C_U C_W \\ & + C_U^2 = 0 \end{aligned} \quad (11)$$

Since the temperature difference between low-latitudes and high northern latitudes $\Delta\theta \equiv \theta_N - \theta_U$ will be negative for

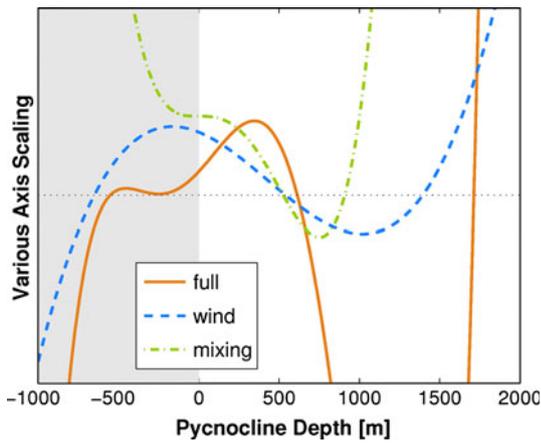


Fig. 3 Governing polynomials for each subcase of the model, where the equilibrium state’s D is determined by the zero transitions of the graphs. The horizontal dotted line marks the zero, while the grey shaded area indicates negative D without physical relevance. The polynomial of degree five (orange line) represents the entire model with its standard parameters. Only two of its roots are positive and thus have a physical meaning. The smaller one at $D = 613$ m denotes the stable state while the larger root, at $\tilde{D} = 1713$ m, must be an unstable solution. The polynomial of degree three (blue line), representing the wind-driven case, has also two relevant solutions $D = 542$ m and $\tilde{D} = 1394$ m. In addition, the mixing-driven case is depicted by the fourth order polynomial (green line). Only two real solutions are detected, a stable one at $D = 526$ m and an unstable one at $\tilde{D} = 909$ m

any realistic situation, our analysis of the equation will be restricted to $\Delta\theta < 0$.

The governing polynomial has five mathematical roots for D (see Fig. 3, orange line), each representing an equilibrium state of our model. Since negative or imaginary pycnocline depths do not have an interpretation in our model set-up, only the positive roots are of interest. Among this physical solutions, some might be unstable under the time-dependent dynamics of the model. We will not explicitly compute the Lyapunov-exponents of the system, but let the numerical integration determine the stability. The parameter choice of Table 1 yields a stable state with a pycnocline depth of $D = 613$ m (see Table 2). Since adjacent solutions cannot share the same stability properties the other zero transition in Fig. 3 with $D = 1,713$ m represents an unstable solution.

A more robust argument for the stability properties is obtained by the sign of the governing polynomial. For positive D (and as long as $D < (C_W + \sqrt{C_W^2 + 4 \cdot C_E C_U}) / (2C_E) = 2,191$ m), the polynomial is proportional to the time derivative of $\partial D / \partial t$. Thus an initial value of $D = 0$ m would increase because the polynomial is positive. But when the first root is exceeded at $D = 613$ m, the polynomial and thus $\partial D / \partial t$ become negative and D decreases. Consequently this root represents a stable steady state. The corresponding solutions for the salinity equations are separately presented in Appendix 4.

3.1 Mixing-driven overturning

First consider a purely mixing-driven case, where SO Ekman transport and eddy return flow are neglected¹ $C_E = 0$ and $C^W = 0$. A basically similar setup was already suggested in Park (1999). The governing eq. (1) reduces here to $m_N = m_U$ providing a simple scaling relation for the northern sinking.

$$m_N = (C_U^2 C_N \Delta\rho)^{1/3} \tag{12}$$

Since C_U is linear in the vertical mixing coefficient κ , the classical scaling $m_N \sim \kappa^{2/3} \Delta\rho^{1/3}$ introduced by Robinson (1960), Bryan (1987) and Park (1999) is reproduced (see Fig. 4f). But an important difference is the existence of a minimal $\Delta\rho^*$ beyond which no physical solution exists. This feature is derived from the governing eq. (11) which reduces to a fourth order polynomial in the mixing limit

$$C_N \rho_0 \beta_S S_0 F_N \cdot D^4 + C_N \rho_0 C_U \alpha_T \Delta\theta \cdot D^3 + C_U^2 = 0. \tag{13}$$

The functional form of the left hand side is depicted in Fig. 3 (green dashed-dotted line) together with the general case. Using $D = C_U / m_N$, eq. (13) can be rewritten in terms of the volume transport

$$m_N^4 - C_N \rho_0 C_U^2 \alpha_T |\Delta\theta| \cdot m_N + C_N \rho_0 C_U^2 \beta_S S_0 F_N = 0. \tag{14}$$

Both equations can be solved analytically. We omit the complicated functional form here and rather provide expressions for conceptually interesting characteristics of the solution.

As shown in Fig. 4 no real positive solution for the oceanic pycnocline exists for northern freshwater fluxes beyond a critical value F_M^* . This flux is defined by the zero transition of the determinant (Fig. 4a, vertical light black line) which is defined by the polynomial in eq. (13). The discriminant is presented in Appendix 2.1 and its root yields an equation for the critical freshwater flux

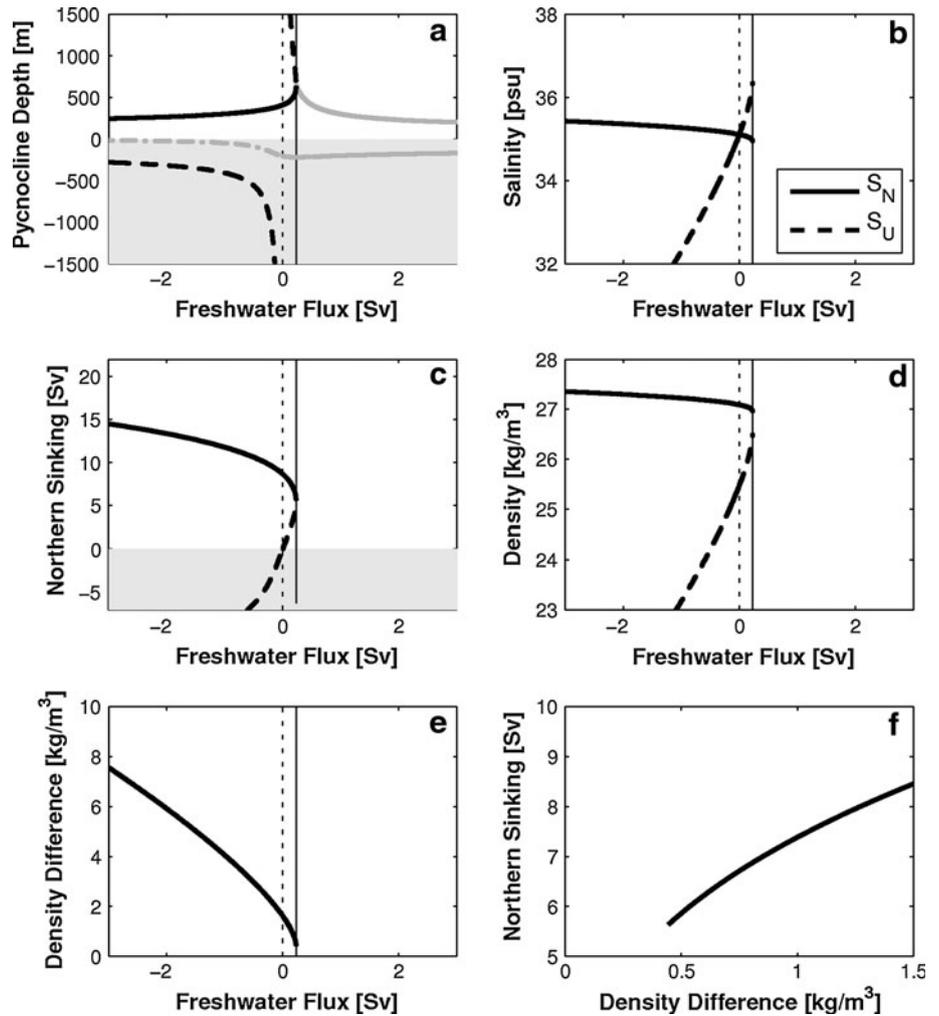
$$F_M^* = \frac{3(2C_N \rho_0)^{1/3} C_U^{2/3} \alpha_T^{4/3}}{8\beta_S S_0} |\Delta\theta|^{4/3}. \tag{15}$$

In contrast to the Rahmstorf (1996) model which yields a quadratic dependence of the critical freshwater flux on the north–south temperature difference $F^* = k\alpha_T^2 \Delta\theta^2 / (4\beta_S S_0)$, the pycnocline dynamics in our model reduce this sensitivity.² The derivative $\partial D / \partial F_N$ is infinite at F_M^* for the stable physical solution (see Appendix 2.1 and Fig. 4a),

¹ Since the southern box is now disconnected from the other basins, the corresponding salinity eq. (6) requires zero southern freshwater flux $F_S = 0$ in order to obtain an equilibrium solution. Obviously this does not affect the pycnocline depth D (cp. eq. 11), nor the volume transports. It only has an impact on the various box salinities.

² Here k is a positive constant which might depend on a prescribed pycnocline depth but not on $\Delta\theta$. In our model, k is quadratic in the pycnocline depth.

Fig. 4 Purely mixing-driven case: dependence of the equilibrium solution on the northern freshwater input F_N . Grey shaded areas indicate negative values either for D or m_N . The stable branch for the pycnocline depth (a, black heavy line) is in correspondence with the solution in Fig. 3. The unstable branch (a, dashed black heavy line) shows a pole with a change in sign at no freshwater flux. This panel also shows the real part of imaginary solutions (a, dark grey lines) to give an impression of the distribution of the solutions. Surpassing a certain freshwater flux, no physically meaningful solution can be obtained. This point is marked by the change of sign in the determinant (a-e, black light line). In panel (b), the behaviour of the northern sinking (stable and unstable branch) is depicted, clarifying the abrupt change from one regime to the other. The other panels (c-f) give an overview of the characteristics of the stable solution. Remarkable are the identity $S_N = S_D$ (b), the existence of a minimal density difference and the scaling of the northern sinking with $\Delta\rho^{1/3}$ (f)



which provides an additional equation to determine further properties of the critical point.

$$D_M^* = \left(\frac{4C_U}{C_N \rho_0 \alpha_T |\Delta\theta|} \right)^{1/3} \tag{16}$$

$$(m_N)_M^* = \left(\frac{C_U^2 C_N \rho_0 \alpha_T}{4} |\Delta\theta| \right)^{1/3} \tag{17}$$

Equations (12) and (17) give the critical density difference

$$\Delta\rho_M^* = \frac{\rho_0 \alpha_T}{4} |\Delta\theta| = 0.40 \frac{\text{kg}}{\text{m}^3}. \tag{18}$$

Despite the different scaling of F_M^* in Rahmstorf (1996), the critical density difference scales linear with $\Delta\theta$ in both models. The decline of D with increasing $\Delta\theta$ results in a weaker dependence of the critical northern sinking on $\Delta\theta$ compared to the linear dependence in Rahmstorf (1996).

A qualitative difference to the wind-driven case (Subsect. 3.2) emerges in the limit of highly negative northern freshwater fluxes. Dividing the governing eq. (13) by F_N and taking the limit $F_N \rightarrow -\infty$ shows that the overturning

circulation grows without bounds. The reasons are the identity $m_N = m_U = C_U/D$ in the mixing case and the fact that an infinite freshwater flux causes the pycnocline depth to vanish. This is not the case for a purely wind-driven overturning.

3.2 Wind-driven overturning

Next, we consider the purely wind-driven case, $C_U = 0$. In this limit, eq. (1) reduces to a quadratic equation in D and provides a relation between the pycnocline depth and the density difference. The only physical solution is

$$D = \frac{C_E}{2C_N \Delta\rho} \cdot \left(\sqrt{1 + \frac{4C_N C_W \Delta\rho}{C_E^2}} - 1 \right).$$

Insertion into eq. (5) yields

$$m_N = C_W - \frac{C_E^2}{2C_N \Delta\rho} \left(\sqrt{1 + \frac{4C_N C_W \Delta\rho}{C_E^2}} - 1 \right). \tag{19}$$

The relation between m_N and the density difference is very different from the mixing case. While no power law

exists for the entire range of $\Delta\rho$, the northern sinking approaches the (constant) strength of the southern ocean upwelling, $m_W = C_W$, for increasing density difference $\Delta\rho$. On the other hand, for a vanishing density difference, the northern sinking tends to the unphysical limit $m_N \rightarrow -\infty$ (see eq. (19)). The crucial question is if the variable $\Delta\rho$ can indeed become arbitrarily small in the wind-driven case. For this, set $C_U = 0$ in the full governing eq. (11) to obtain the complete equilibrium dynamics.

$$\begin{aligned}
 &+ D^3 \cdot C_N \rho_0 C_E \alpha_T \Delta\theta \\
 &+ D^2 \cdot (C_E^2 + C_N \rho_0 C_W \alpha_T \Delta\theta + C_N \rho_0 \beta_S S_0 F_N) \\
 &- D \cdot 2C_E C_W + C_W^2 = 0
 \end{aligned} \tag{20}$$

Using $D = (C_W - m_N)/C_E$ we can transform this equation into an expression for the northern sinking

$$\begin{aligned}
 &+ m_N^3 \cdot C_N \rho_0 C_W^2 \beta_S S_0 F_N \\
 &+ m_N^2 \cdot (C_E^2 - 2C_N \rho_0 C_W \alpha_T \Delta\theta + C_N \rho_0 \beta_S S_0 F_N) \\
 &+ m_N \cdot C_N C_W \rho_0 (C_W \alpha_T \Delta\theta - 2\beta_S S_0 F_N) \\
 &+ C_N C_W \rho_0 \alpha_T \Delta\theta = 0.
 \end{aligned} \tag{21}$$

Both equations for the wind-driven case show a third order polynomial which can be solved analytically. As in the mixing case a stable physical solution exists up to a critical threshold of the northern freshwater flux $F_N < F_W^*$ (see Fig. 5a–e, light vertical line). This critical freshwater flux F_W^* is determined by the only real root of the correspondent discriminant which shows a third order in F_N (cp. Appendix 2.2). Since the analytic solution is complicated it is only depicted in Fig. 5. As an alternative to analysing the full solutions, we focus on the sensitivity of F_W^* on the North-South temperature difference retrieved by the derivative of the discriminant with respect to $\Delta\theta$ (see Appendix 2.2). This derivative can suitably be approximated for realistic temperature differences from the limit $\Delta\theta \rightarrow -\infty$

$$\begin{aligned}
 \frac{\partial F_W^*}{\partial \Delta\theta} &\approx \lim_{\Delta\theta \rightarrow -\infty} \frac{\partial F_W^*}{\partial \Delta\theta} = -\frac{\alpha_T C_W}{\beta_S S_0} \\
 \text{for } |\Delta\theta| &\gg \frac{C_E^2}{C_N \rho_0 C_W \alpha_T},
 \end{aligned} \tag{22}$$

which is given as a slope in addition to the full dependence in Fig. 6 (dashed blue line). This finding comprises that the function $F_W^*(\Delta\theta)$ becomes quasi-linear for large temperature differences. Figure 6 indicates that it also is a good approximation for realistic $\Delta\theta$. Moreover, it is shown that $-\alpha_T C_W / (\beta_S S_0) \cdot |\Delta\theta|$ is an upper constraint for the actual critical freshwater flux F_W^* for realistic temperature differences (see Appendix 2.2).

In the present case, the limit $F_N \rightarrow -\infty$ causes the pycnocline depth D and consequently the eddy return flow m_E to vanish (see eq. (20)). Thus, in contrast to the mixing case where northern sinking diverges, here the northern sinking approaches the constant southern upwelling m_W .

3.3 Full problem

Though no complete analytic solution can be obtained for the full model (eq. (11)), some analytic insight can be gained. A formal expansion of the steady state pycnocline dynamics with respect to the parameter set (C_E, C_U) around the purely wind-driven case $m_N = m_W$, i.e. $(C_E, C_U) = (0, 0)$, yields

$$m_N^{(1)} = C_W - \sqrt{\frac{C_W}{C_N \Delta\rho_0}} \cdot C_E + \sqrt{\frac{C_N \Delta\rho_0}{C_W}} \cdot C_U \tag{23}$$

with

$$\Delta\rho_0 \equiv \Delta\rho|_{(0,0)} = -\rho_0 \left(\beta_S \frac{S_0 F_N}{C_W} + \alpha_T \Delta\theta \right) \tag{24}$$

Each of the three terms in (23) can be understood in light of the former limits of purely wind-driven and mixing-driven circulations. The first term is the Southern Ocean upwelling, a constant contributor balancing the northern sinking. It is reduced by the eddy return flow represented by the second term. An additional contribution emerges through the low-latitude upwelling of the third term. This approximation holds reasonably well for the parameter set of Table 1 for a realistic range of density differences (Fig. 7f, light line).

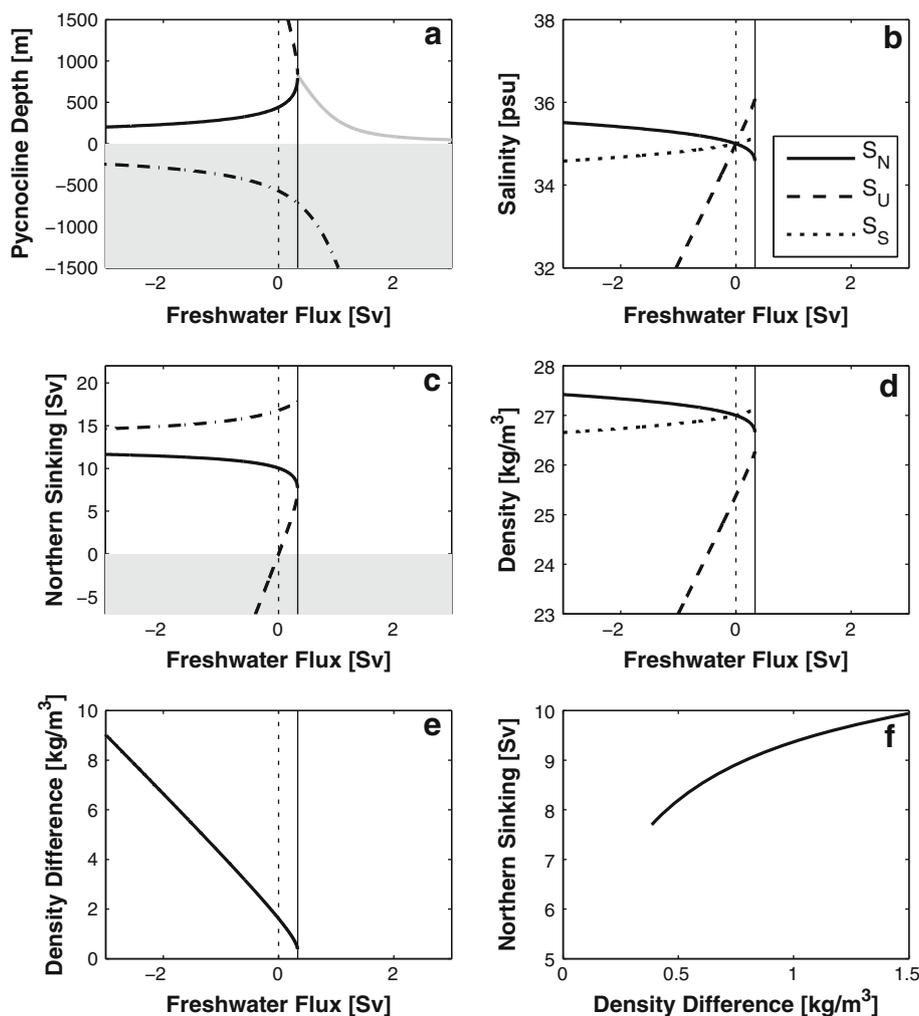
For the full problem, an analytic treatment of the critical freshwater flux F_F^* is not possible since the discriminant Υ_F (eq. (60)) is a fifth order polynomial in F_N (cf. Appendix 2.3). However, the intermediate value theorem states, that a fifth order polynomial has at least one real root. For a physical choice of parameters (positive C_N, C_E, C_W, C_U), the full problem therefore always exhibits a critical freshwater flux F_F^* . This value can now be estimated by using the linearised model eq. (23). The discriminant of this second order polynomial in D provides a second order polynomial to be solved for $F_{F(appr)}^*$. One of the two solutions is physically interesting and yields the approximated critical freshwater flux

$$F_{F(appr)}^* = -\frac{C_W}{\beta_S S_0} \cdot \left(\frac{\Delta\rho_0^*}{\rho_0} + \alpha_T \Delta\theta \right) \tag{25}$$

$$\begin{aligned}
 \Delta\rho_0^* &= \left(\frac{-C_W + \sqrt{C_W^2 + 4 \cdot C_E C_U}}{2 \cdot \sqrt{\frac{C_N}{C_W}} C_U} \right)^2 \\
 &= 0.29 \frac{\text{kg}}{\text{m}^3}
 \end{aligned} \tag{26}$$

This relation first of all confirms that linearity of the critical freshwater flux in $\Delta\theta$ is not merely restricted to the wind driven case, but also serves well to approximate the full problem. In both cases we find the same proportionality constant. Moreover, this approximation also provides an estimate for the offset of this linear relation. This offset is

Fig. 5 Dependence of the equilibrium solution for the wind-driven case on the northern freshwater input F_N . Grey shaded areas indicate negative values either for the pycnocline or the northern sinking. The stable branch for the pycnocline depth (a, black heavy line) is in correspondence with the solution in Fig. 3. The negative solutions for D are assumed to have no physical relevance. Nevertheless another positive unstable (a, dashed black heavy line) and the real part of an imaginary branch (a, dark grey line) are depicted to give an impression of the structure of the solutions. Exceeding a specific freshwater flux F_N , no physical meaningful solution can be found. This point is marked by the change of sign in the determinant (a-e, vertical black light line). In panel (c), the behaviour of the northern sinking (stable and unstable branch) is depicted, clarifying the abrupt change from one regime to the other. A solely wind-driven overturning imposes an upper bound on the northern sinking. The other panels (b-f) give an overview of the characteristics of the stable solution and (f) exhibits the existence of a minimal density difference $\Delta\rho_W^*$



proportional to the critical $\Delta\rho^*$, which itself is in this first-order approximation independent of the meridional temperature gradient and totally determined by the model parameters. Equation (25) captures the dependency of F_F^* on $\Delta\theta$ reasonably (Fig. 6). Another interesting detail is that the critical freshwater flux of the full problem F_F^* exceeds the ones from the wind- and the mixing-driven cases. This gives rise to a discussion for the freshwater sensitivity of the model.

4 Overturning sensitivity to freshwater

The derivative $\partial m_N / \partial F_N$ gives a mathematical measure for the sensitivity of the northern sinking m_N to changes in the intensity of the northern freshwater flux F_N . This derivative can be determined for the Stommel (1961) model and for all our subcases but not for the approach of Gnanadesikan (1999). F_N is here implicitly included via the parameter $\Delta\rho$ and one would need an extra equation to link them. However, for the mixing- and wind-driven case as well as for the full problem, the derivative $\partial m_N / \partial F_N$ is a function with a

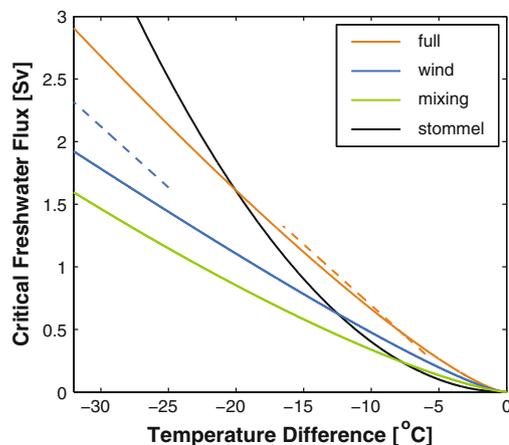
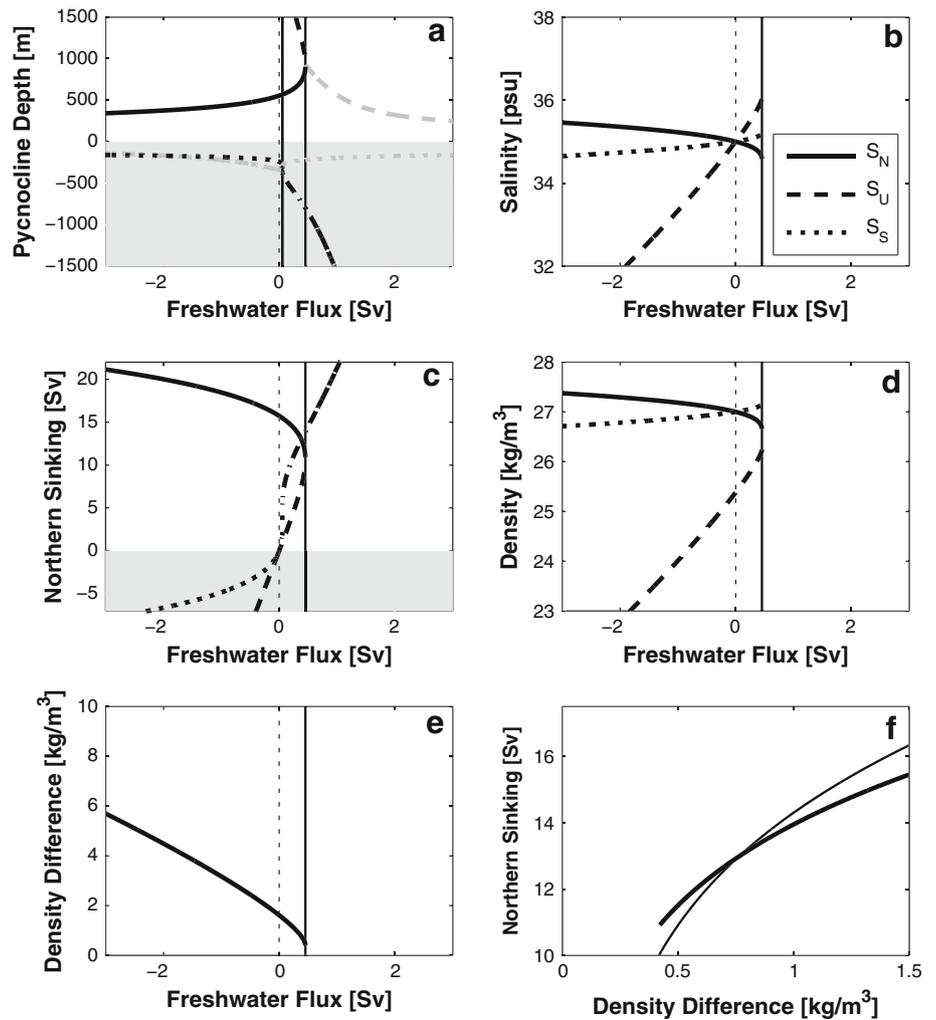


Fig. 6 Critical freshwater flux F^* as a function of the meridional temperature difference $\Delta\theta$ for the mixing-driven, wind-driven and full problem. In the wind-driven case the sensitivity on $\Delta\theta$ can be conveniently approximated by the value of the derivative $\partial F_W^* / \partial \Delta\theta$ at minus infinity (blue dashed). This slope is again retrieved by approximating the full problem (blue dashed), but in addition the offset can be determined (eq. (25)). The $\Delta\theta$ dependence of Stommel's F^* with a prescribed pycnocline depth (chosen according to Fig. 8) is quadratic and thus most pronounced

Fig. 7 Dependence of the equilibrium solution for the full problem on the northern freshwater input F_N . Grey shaded areas indicate negative values either for the pycnocline or the northern sinking. The stable branch for the pycnocline depth (a, black heavy line) is in correspondence with the solution in Fig. 3. The negative solutions for D is assumed to have no physical relevance. To give an impression of how the solutions are distributed in the phase space, panel (a) also shows a positive unstable (dashed black heavy line) and the real part of imaginary branches (dark grey lines). Surpassing a specific freshwater flux F_N , no physical meaningful solution can be found. This point is marked by the change of sign in the determinant (a, black light line). In the middle left panel, the behaviour of the northern sinking (stable and unstable branch) is depicted, clarifying the abrupt change from one regime to the other. Beside the other characteristics (b-e), the complete solution (f, heavy line) shows a scaling whose major shape can be adequately described via a Taylor expansion (f, light black line)



pole of order one in F_N (see Appendix 3). In the *Stommel* model, the derivative is obtained from eq. (8) showing a pole of order $\frac{1}{2}$. In order to determine which model has the highest sensitivity to freshwater input in the North Atlantic, it is necessary to align the positions of the respective poles. The pycnocline depth D_S is a parameter in the *Stommel* (1961) model, we choose it such that the pole of the *Stommel* model is at the same position as the pole of the mixing-driven case (Fig. 8). This is motivated by the resemblance of the circulation described in the *Stommel* model and in our mixing-driven case. Both exhibit only one mixing-driven circulation cell connecting the various boxes. This approach (cf. Appendix 3) yields

$$D_S = \sqrt{3} \cdot \left(-\frac{C_U}{2C_N \rho_0 \alpha_T \Delta \theta} \right)^{1/3} \quad (27)$$

For our parameters, the right hand side has a value of 564.1 m. Choosing a smaller D_S , shifts the pole of the *Stommel* model F_S^* to a lower position than the one for the mixing-driven case F_M^* . Studying the wind-driven case and the full

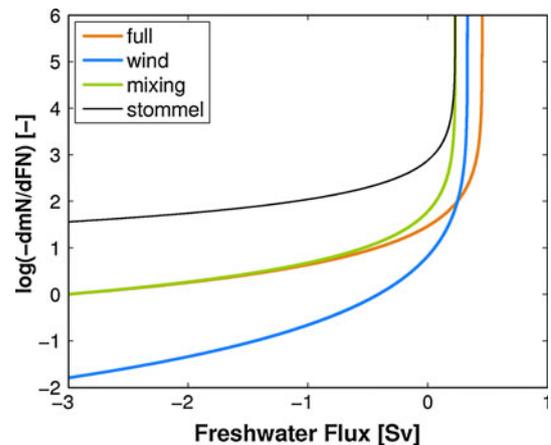
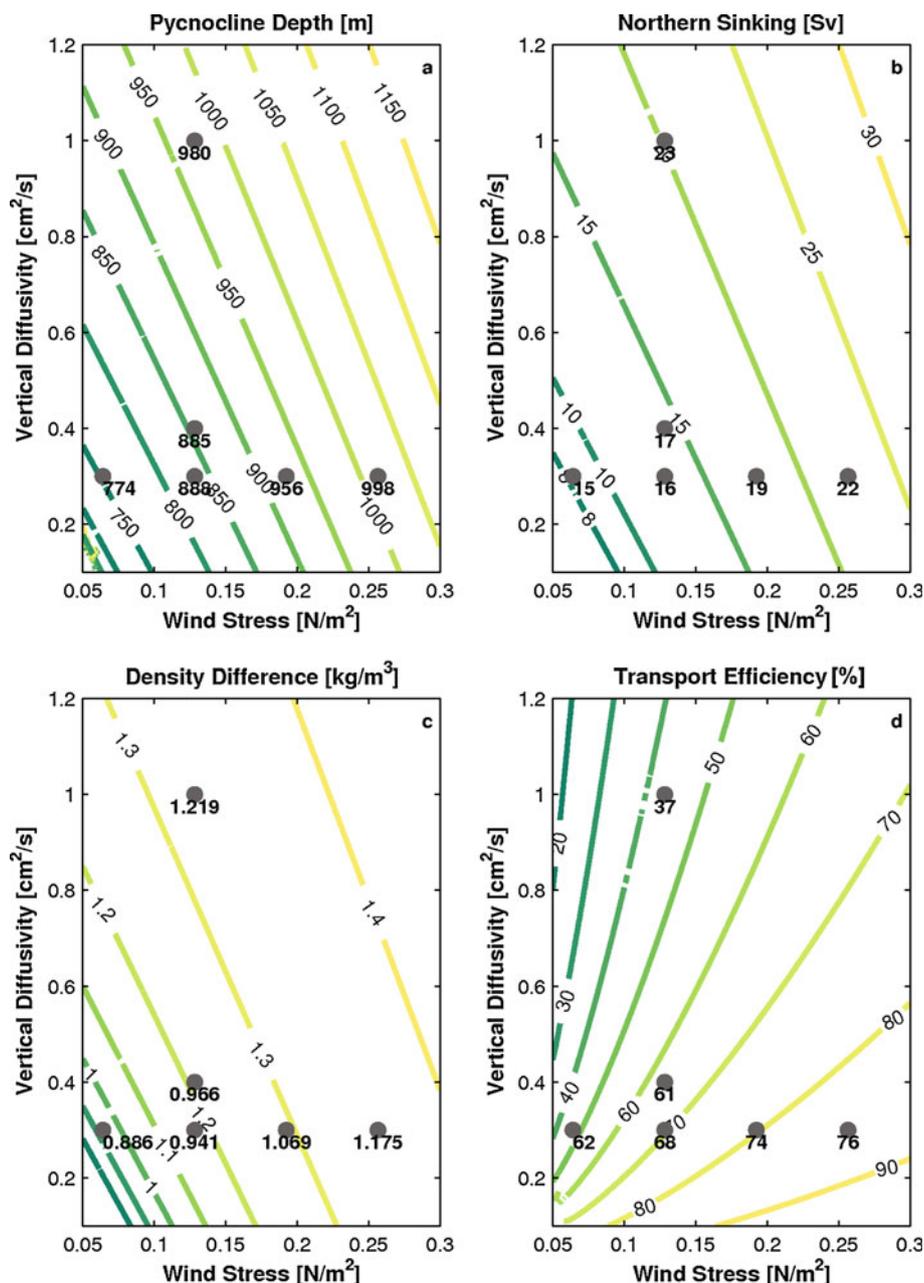


Fig. 8 Sensitivity of northern sinking to changes in surface freshwater flux: In the *Stommel* model (black line) the sensitivity of m_N to changes in freshwater flux F_N is higher than in our model with varying pycnocline. Thus independent of the physical driving process, the pycnocline stabilises the overturning circulation up to the critical threshold where no solution exists. In addition, considering all the subcases of our model, the mixing-driven case is most sensitive

Fig. 9 Solutions to the governing eq. (11) as a function of the Southern Ocean wind stress and the vertical diffusivity. Results from CLIMBER-3 α are superimposed as • symbols. The panels depict (a) the pycnocline depth, (b) the Northern sinking m_N , (c) the meridional density gradient and (d) the transport efficiency. This quantity determines the fraction of the Northern sinking supplied by the Southern Ocean $(m_W - m_E)/m_N$



problem, we find that considering some additional constraints on the parameter space, their respective poles F_W^* and F_F^* are located at higher positions than that of the mixing-driven case F_M^* (cp. Fig. 6 and app. 9). In fact these new constraints hardly restrict a physical parameter choice. For our set of parameters, the constraint for the wind-driven case reads implicitly $D \leq 913$ m which holds for the entire stable, physical solution branch (cp. Fig. 5). In the full problem, the implicit constraint includes an additional lower limit $351 \text{ m} \leq D \leq 1,329$ m, which is violated but only in the non-physical case of a strong, inverse northern freshwater flux (cp. Fig. 7)

Given the found sequence $F_S^* \leq F_M^* < F_W^*$ and $F_M^* < F_F^*$, we now focus on the freshwater sensitivity of

m_N . It is possible to show (Appendix 3) that the Stommel (1961) model exhibits a higher sensitivity compared to the mixing case (also see Fig. 8), as long as D_S fulfills a slightly more stringent constraint. Accounting for a small corrections term (see Appendix 3), the constraint of eq. (27) lowers slightly for our parameters to $D_S \leq 558.4$ m. This reduces F_S^* by merely $4.6 \cdot 10^{-3} \text{ Sv}$. The new constraint is therefore well approximated by the more intuitive one of eq. (27). However, even if F_S^* slightly surpasses this constraint (same order of magnitude 10^{-3} Sv), the sensitivity of the Stommel (1961) model would exceed that of the mixing-driven case below a freshwater input in a close vicinity of F_M^* (Fig. 8 and Appendix 3). In order to mutually compare

the freshwater sensitivities of the different cases in our model, the implicit expressions of the derivatives $\hat{\partial}m_N/\hat{\partial}F_N$ are used. Appendix 3 reveals that some additional parameter constraints grant that the freshwater sensitivity of m_N in the mixing-driven case is higher than that for the wind-driven case and the full problem (also see Fig. 8). These new constraints are again not violated by our parameter set. Thus, independent of the predominant driving mechanism, the dynamics of the model pycnocline stabilises the *Stommel* overturning up to the critical threshold. The model can even bear a lower density difference ($\Delta\rho_M^* = 0.40 \text{ kg/m}^3$ and $\Delta\rho_F^* \approx 0.29 \text{ kg/m}^3$). In correspondence with the finding that the wind-driven overturning is limited by C_W , this case shows the lowest freshwater sensitivity below a small positive F_N .

5 Comparison with comprehensive ocean model

For a brief validation of the qualitative behaviour of our conceptual approach, experiments with a model of intermediate complexity were carried out varying vertical diffusivity and SO wind forcing. All results are based on simulation with CLIMBER-3 α , described by Montoya et al. (2005). It includes modules describing the atmosphere, land-surface scheme as well as sea-ice. The three-dimensional oceanic component (MOM-3) has a horizontal resolution of $3.75^\circ \times 3.75^\circ$ and 24 non-uniformly spaced levels covering the vertical extent.

The first set of steady state experiments investigates the influence of vertical background diffusivity in the ocean, analogous to (Mignot et al. 2006). Three experiments with vertical diffusivity of 0.3, 0.4 to $1.0 \cdot 10^{-4} \text{ m}^2/\text{s}$ were conducted. The second set of experiments follows Schewe and Levermann (2010) and analyses the influence of the zonal wind stress in the Drake Passage on the MOC. An amplification of the zonal wind field was applied in a latitudinal band between 71.25°S and 30°S with factors of $\alpha = 0.5, 1.0, 1.5$ and 2. Both experiments are closely linked to one of the two upwelling mechanisms powering the AMOC. The wind experiments directly affect the rate of Ekman pumping in the SO, while changes in κ exert control on the low-latitude upwelling. The values for $D, \Delta\rho, m_N$ and $m_W - m_E$ are determined as described in Levermann and Fürst (2010).

For appropriate parameters, our conceptual model captures the qualitative response of CLIMBER-3 α to changes in the magnitude of the two driving mechanisms (Fig. 9). This is not trivial since CLIMBER-3 α allows for many more complex feedback mechanisms than the conceptual model. Since parameter sensitivity is strongly dependent on the used model, an other ocean general circulation model (GFDL Modular Ocean Model, Version 3.0) is consulted.

With this model a similar parameter scan was conducted and already presented in Gnanadesikan (1999). The general response is in agreement with our results. However the ocean model shows higher variations in the pycnocline depth and overturning. This confirms the choice for the central transport processes to be feasible and supports that the dynamics of the AMOC is well described by variations in both the pycnocline depth and the meridional density gradient.

6 Discussion and conclusion

In this study we address the question on whether and to what extent the stability properties of the AMOC depends on its driving processes that are associated with the upwelling branches of the overturning (Kuhlbrodt et al. 2007). At the moment two mechanisms are under discussion: upwelling in the low latitudes induced by turbulent mixing across isopycnals and an ascent of water masses in the latitudinal band of the Drake Passage due to diverging westerly winds. We present a conceptual model which includes both processes in addition to the salt-advection feedback considered at the heart of an AMOC instability. The strength of our model lies in the possibility of studying qualitative differences between a mixing- or a wind-driven overturning.

First and foremost, considering the conceptual model to be in steady state, an analytic description is found for the wind- and for the mixing-driven case (see Sect. 3) In the mixing-driven case, it reproduces the classical scaling of the northern sinking with $\kappa^{2/3}\Delta\rho^{1/3}$ introduced by Bryan (1987). Set by the SO winds, the purely wind-driven overturning imposes an upper bound for the northern sinking. For an overturning circulation which is powered by both driving mechanisms, a corresponding approximation of the northern sinking is found. This scaling relation (see eq. 23) provides an instructive equation for the respective influences of the two driving mechanisms and the SO eddy transport. This comprehensive case and the purely wind-driven one exhibit no simple power law for the entire range of $\Delta\rho$.

One of the main results is the existence of a critical threshold beyond which no AMOC can be sustained, i.e. no physical solution exists in our model. The existence of an off-states for a wind-driven overturning was already suggested by Johnson et al. (2007) in a similar but slightly more comprehensive model than presented here. They computed the off-state by setting the northern sinking to zero. It should however be noted that their model as well as the one presented here are designed for a situation with a functioning overturning. While it is possible to determine the point at which solutions cease to exist, it is not obvious

that these models can be used to compute the off-state in any realistic way. In a G99 set-up an off-state requires that SO upwelling is compensated by eddy return flow in the Southern Ocean. While it is clear that the tracer budget of heat and salinity can be closed in this fashion, it is not obvious whether the same holds for the momentum balance. For this, eddies would need to transport significant amount of momentum and it is questionable that such a flow is well described by the diffusion equation of Gent and McWilliams (1990) with one constant coefficient.

Our conceptual model makes it possible to explore the existence and the position of the critical freshwater threshold. The dependences of this critical freshwater flux are crucially dependent on the involved transport processes and can be expressed as a function of the meridional temperature difference. In our model the relevant density and thus temperature differences are taken between low and high latitudes. Due to polar amplification this temperature difference is likely to decrease under future warming (e.g. Cai and Lu (2007)). We find that the sensitivity of the critical freshwater flux to the meridional temperature difference is reduced from a quadratic dependence in a *Stommel* model (Rahmstorf 1996) to one of the power $4/3$ in the mixing-driven case and to 1 when SO wind forcing is included. Physically this means that the dynamics of the pycnocline depth causes the overturning to be more robust under atmospheric temperature forcing.

Concerning freshwater forcing, climate models of intermediate complexity show a large spread in sensitivity and hysteresis position (Rahmstorf et al. 2005). In our conceptual model, such differences can be associated with the dominant driving mechanism. An overturning partially or exclusively powered by SO winds is able to bear higher freshwater fluxes than a purely mixing-driven circulation (with parameters chosen from Table 1 one finds $F_M^* = 0.23$ Sv compared to $F_F^* \approx 0.60$ Sv). The sensitivity of the overturning to changes in freshwater fluxes in the North Atlantic below the critical threshold also depends on the main driving mechanism (Sect. 4) Setting the pycnocline depth in the *Stommel* model to a certain value D_S allows the sensitivity comparison to our model. This is done by setting the critical freshwater input of the Stommel (1961) model to the same position as the threshold of the mixing-driven case. Under this parameter constraint, the freshwater sensitivity of the northern sinking is less pronounced in the mixing-driven case than in a *Stommel* model. Under further parameter constraints, we were able to show that a mixing-driven overturning is more sensitive to freshwater perturbations than an AMOC driven by SO winds. One can thus conclude that the pycnocline dynamics stabilises the northern sinking under changes in the northern freshwater flux. For the mixing-driven case this was already proposed

by Park (1999) but the sensitivity reduces further for an overturning with SO upwelling. Less freshwater sensitivity might indeed pose a problem for AMOC monitoring since the threshold is not easily detected by a significant, preceding slow-down (cp. Figs. 4, 5 and 7). It should be noted that the model presented here was designed as a minimal model that captures the salt-advection feedback in combination with a representation of both AMOC driving mechanisms. Levermann and Fürst (2010) recently showed that in order to capture the behaviour of the coupled climate model CLIMBER-3 α under global warming an additional dynamical equation for the geometry of isopycnal out-cropping in the SO is necessary.

Acknowledgments JJF thank his colleagues Tore Hattermann, Daria Schönemann and Jacob Schewe for fruitful discussions that improved the quality of the presented work. This work also profited from the constructive suggestions of the reviewers whom the authors want to express their gratitude.

Appendix 1: Scaling of the northern sinking

This section briefly presents five different approaches to formulate scaling laws for the northern sinking. For each approach, a short derivation is given, which identifies the main assumptions and discusses their validity. In this context, the difference between two vertical scales is emphasised, the pycnocline depth D and the level of no motion Λ . The first description from *Robinson* (1960), also called the classical scaling, assumes the meridional overturning circulation to be in geostrophic balance. The vertical derivative of the momentum equation in the steady state yields

$$f \frac{\partial v}{\partial z} = -g \cdot \frac{\partial \rho}{\partial x}, \quad (28)$$

when the hydrostatic equation is applied. Here v denotes the meridional velocity, ρ is the ocean density field and g is the gravitational constant.

Two scales, one for the meridional velocity field V and one for the characteristic depth Λ of the vertical profile of horizontal velocities, are introduced. The second is identified with the level of no motion, where the mean meridional velocities vanish. Additionally using a scale for the zonal density gradient $\Delta_x \rho$ occurring over a length scale L_x gives

$$V = \frac{g}{f} \frac{\Delta_x \rho}{\rho_0} \cdot \frac{\Lambda}{L_x}. \quad (29)$$

In order to transform the zonal density gradient into a meridional one $\Delta_y \rho$, the zonal and meridional velocity scales are linked. Albeit a radical generalisation, a constant ratio $V = C_{cl} U$ is assumed.

$$V = C_{cl} \frac{g \Delta_y \rho}{f \rho_0} \cdot \frac{\Lambda}{L_y} \tag{30}$$

Since we seek an expression for the Atlantic overturning, an integration over the zonal extent L_m and the vertical extent Λ of the flow is conducted. This gives the classical scaling for the northern sinking

$$m_N = V \cdot \Lambda \cdot L_m = C_{cl} \frac{g \Delta_y \rho}{f \rho_0} \frac{L_m}{L_y} \cdot \Lambda^2, \tag{31}$$

which is proportional to the meridional density gradient, to the ratio between zonal and meridional length scales L_m/L_y and to the square of the level of no motion Λ . The theoretical basis for this estimate constrains its spatial applicability to (1) the geostrophic assumption which is not valid near continental boundaries and to (2) the ad-hoc transformation from zonal to meridional density gradients. *Gnanadesikan (1999)* suggests that a western boundary current exerts control on the northern sinking. Neglecting the velocity component perpendicular to the boundary, a relation between the meridional pressure gradient and the meridional velocity is derived

$$\frac{1}{\rho_0} \frac{\partial p}{\partial y} = v \nabla^2 v = v \frac{\partial^2 v}{\partial x^2}, \tag{32}$$

where ∇^2 is the Laplacian and v is the dynamic viscosity in the boundary current. For the second equality it is assumed that the zonal change in meridional velocity exceeds the changes in vertical and meridional direction by several orders of magnitude (*Montoya et al. 2005*).

A scale analysis analogue to the previous paragraph provides

$$v \frac{V}{L_m^2} = C_{G99} \frac{1}{\rho_0} \frac{\Delta_y p}{L_y} = -C_{G99} g \frac{\Delta_y \rho}{\rho_0} \frac{D}{L_y}, \tag{33}$$

where the constant C_{G99} accounts for any effects of geometry and boundary layer structure. For the second equality, where the hydrostatic equation is employed, a new vertical scale height D is introduced. This height represents the density stratification of the ocean and is referred to as the pycnocline depth D .

Integration yields another scaling law for the northern sinking

$$m_N = V \cdot \Lambda \cdot L_m = C_{G99} \frac{g L_m^2 \Delta_y \rho}{v \rho_0} \frac{L_m}{L_y} D \Lambda. \tag{34}$$

In analogy with the classical scaling law, we again find proportionality to a meridional density difference. However the relation between the Coriolis frequency f is replaced by the inverse of the zonal viscosity time scale in the western boundary current L_m^2/v . Another even more important change is the proportionality to the product of the pycnocline depth D and the level of no motion Λ . This

approach is drawn from the assumption that a meridional pressure gradient causes a frictional western boundary current which limits the deep water formation in the Nordic Seas. In contrast to this, *Johnson and Marshall (2002)* base their scaling of the northern sinking on an ocean model of reduced gravity. It is built up by a surface layer of depth h and an infinitely deep and motionless lower layer of fixed density. In this set-up, the level of no motion is implicitly equal to the pycnocline depth. Assuming a geostrophic flow in the interior of the surface ocean basin to provide the water needed for the northern sinking m_N , the meridional velocity becomes

$$v = -\frac{g \Delta_z \rho}{f \rho_0} \cdot \frac{\partial h}{\partial x}, \tag{35}$$

where $\Delta_z \rho$ is the vertical density difference between the two layers. Since the depth h is a function of x , a zonal integration from the western to the eastern boundary leads to

$$m_N = \int_{x_w}^{x_e} h v dx = \frac{g \Delta_z \rho}{2f \rho_0} \cdot (h_E^2 - h_W^2) \tag{36}$$

$$= \frac{g \Delta_z \rho}{2f \rho_0} \cdot \tilde{D}^2, \tag{37}$$

where h_E, h_W are the layer depth at the eastern and western boundary, respectively. The third step implies Johnson's redefinition of the pycnocline depth \tilde{D} (*Johnson et al. 2007*).

The main difference to previous scalings is a dependence on a density gradient $\Delta_z \rho$ in the vertical direction. In addition this approach sees the reason for the geostrophic flow in a zonal tilt of the pycnocline depth. However, it is argued that an outcropping of the pycnocline occurs at the western boundary, while at the eastern boundary the pycnocline is equal to D (*Johnson et al. 2007*). A fundamentally different approach is provided by *Guan and Huang (2008)* who introduce an energy constraint, instead of the well known buoyancy constraint. The idea is that the energy supply is used for diapycnal mixing, which is described by a vertical advection-diffusion balance. Using a scale for the vertical density difference $\Delta_z \rho$, the scale of its vertical change $\Delta_z(\Delta_z \rho)$ and another for the pycnocline depth D , the equation can be rewritten for a constant diapycnal diffusivity κ

$$w \Delta_z \rho = \frac{\kappa}{D} \Delta_z(\Delta_z \rho). \tag{38}$$

Assuming an exponential density profile, the proportionality $\Delta_z(\Delta_z \rho) \sim \Delta_z \rho$ becomes valid.

The gravitational potential energy (GPE) in a two-layer box model, with a vertical density difference $\Delta_z \rho$, increases due to vertical mixing with a rate of $-g \kappa \Delta_z \rho$ (per unit

area). Thus, meridional and zonal integration over a range of respectively L_m and B yields

$$E_{pot} = -g\Delta_z\rho\kappa L_m B = -g\Delta_z\rho\omega D L_m B. \tag{39}$$

Rearranging the equation and integrating over the same horizontal plane, a new scaling of the northern sinking arises

$$m_N = -\frac{E_{pot}}{g\Delta_z\rho D}. \tag{40}$$

It is obvious that this approach diametrically opposes the ones above, because the northern sinking m_N is now inversely proportional to the vertical density gradient $\Delta_z\rho$ and D . This discrepancy is caused by its fundamentally different assumptions. Unlike in the others, an energy source for the circulation is included, which maintains diapycnal diffusion and therefore produces available potential energy E_{pot} for the northern sinking. *Marotzke (1997)* introduces a scaling law similar to the classical approach. The difference lies in a convincing transformation from a zonal density gradient into a meridional one. *Marotzke* links them by using several assumptions about the density stratification of the ocean:

1. The density of the ocean surface is solely a linear function of latitude and the properties of the deep ocean are given by the surface water of highest density.
2. The density of the ocean surface is solely a linear function of latitude and the properties of the deep ocean are given by the surface water of highest density.
3. The occurrence of Kelvin and Rossby waves in equatorial regions eliminates all zonal isopycnal slopes except for the western boundary current.
4. At the eastern boundary, a well mixed surface layer down to a fixed depth z_ρ is assumed. This depth is prescribed in equatorial regions up to a specific latitude, where it becomes zero. This means that the isopycnal that separates the surface layer from the abyssal ocean, outcrops at a defined latitude.

This idealised stratification provides a linear relation between zonal and meridional density differences

$$\Delta_x\rho(y, z_\rho) = \Delta_y\rho(y, z_\rho)(1 - y/L_y) \cdot y/L_y, \tag{41}$$

calculated at a specific depth z_ρ , with L_y being the meridional extent of the basin. The density differences $\Delta_x\rho, \Delta_y\rho$ are determined at opposite edges of the North Atlantic basin, respectively in meridional and zonal directions. Since the dependence on latitude is not the main focus here, the latitudinal maximum for the equation is used, which is attained at $y = \frac{1}{2}L_y$. Inserting in the meridional geostrophic equation and applying a scale analysis yields

$$\frac{V}{\Lambda} = C_{Ma} \frac{g \Delta_y \rho}{f \rho_0} \frac{1}{L_x}, \tag{42}$$

where C_{Ma} is a constant accounting for geometry.

Integrating twice in vertical direction from the surface to the level of no motion gives

$$m_N = C_{Ma} \frac{g \Delta_y \rho L_m}{f \rho_0 L_x} \cdot \Lambda^2. \tag{43}$$

In correspondence to the classical scaling law, this one also predicts a proportionality of the northern sinking to a meridional density difference $\Delta_y\rho$ of the ocean surface layer, and to the square of the level of no motion Λ . Apart from geostrophy, it is the four assumptions from above which should be evaluated to judge the validity of this scaling. *Marotzke* argues that the first three have already been used successfully and therefore are generally accepted. The last one for the eastern boundary layer is based on model observations (*Marotzke 1997*).

In this work, the decision fell on the approach of *Marotzke (1997)*, because it combines the idea of a geostrophic current and a boundary layer theory. In this way, the zonal pressure gradient is convincingly converted into a meridional one. Note that although the level of no motion Λ and the pycnocline depth D are physically different, these two scales cannot be seen as independent from each other. The level of no motion definitely separates two water bodies whose dynamics brings water from spatially separated areas. But this also creates a significant difference in the salinity and temperature characteristics, which gives rise to strong stratification.

Appendix 2: Discriminants

This section deals with the derivation of the discriminants for the various model subcases. They contain all information needed to characterise the transition of the model between different dynamic regimes.

A polynomial p_n of degree n in one variable $x \in \mathcal{C}$ is described via its roots

$$p_n(x) = a_n \prod_{i=1}^n (x - \alpha_i) = \sum_{i=0}^n a_i \cdot x^i, \tag{44}$$

where $\alpha_1, \alpha_2, \dots, \alpha_n$ are the roots of p_n and a_n is the coefficient of the highest order term. One general form to determine its discriminant is

$$Y[p_n] := n^n (-1)^{\frac{n(n-1)}{2}} a_n^{2n-2} \prod_{i < j} (\alpha_i - \alpha_j)^2. \tag{45}$$

Given a concrete polynomial $p_n(x)$ with $a_n \neq 0$, then $Y[p_n] = 0$ if and only if p_n has a double root [Mathematics:

Theory & Applications, *Discriminants, Resultants and multidimensional Determinants*, p.404,6]. This can directly be deduced by applying the definition of the discriminant. It implies that a change in the amount of real roots is indicated by the roots of the discriminant, which are a function of the coefficients a_i . If the sign of the discriminant $\Upsilon[p_n]$ changes, it has drastic implications for the roots of the polynomial p_n . This is best illustrated by a quadratic polynomial p_2 , whose roots are found at

$$\alpha_{1/2} = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2a_0}}{2a_2}. \tag{46}$$

The discriminant is the negative term within the square root $\Upsilon[p_2] = 4a_2a_0 - a_1^2$ and its sign determines whether the two roots are imaginary or real. For cubic polynomials p_3 , $\Upsilon[p_3] < 0$ signifies three real roots, $\Upsilon[p_3] = 0$ one real root and $\Upsilon[p_3] > 0$ two imaginary and one real root. For polynomials of degree $n \geq 4$, the connection between the sign of the discriminant and the characteristics of the roots becomes more elaborate and will not be used later.

In general, we assume that all volume flux constants C_N, C_U, C_W, C_E , the mean density ρ , the average salinity S_0 , the expansion coefficients α_T, β_S , the freshwater bridges F_N, F_S and the pycnocline depth D are positive. In addition the pole to equator temperature difference $\Delta\theta$ can only be negative. This defines the physical parameter space, which we are going to adopt in the following discussion.

2.1 Mixing-driven case

Instead of explicitly calculating the discriminant of the governing polynomial of the mixing-driven case (eq. (13)), a mathematical tool from number theory has been used (Leutbecher 1996, p.225), since it provides a simple way to compute it. A so-called resolvent function is determined by the original polynomial of degree four. On one hand, this new function is a polynomial whose order is reduced by one and its discriminant equals that of the governing polynomial (eq. (13)). It allows thus to calculate the discriminant via a polynomial of degree 3, which yields

$$\Upsilon_M = 256C_N^3\rho_0^3\beta_S^3S_0^3F_N^3C_U^6 - 27C_N^4\rho_0^4C_U^8\alpha_T^4\Delta\theta^4. \tag{47}$$

Its zero transitions determine a critical value, where the dynamic of the model changes. The root of the discriminant reads, with respect to the freshwater flux,

$$F_M^* = \frac{3(2C_N\rho_0)^{1/3}C_U^{2/3}\alpha_T^{4/3}}{8\beta_S S_0} |\Delta\theta|^{4/3}. \tag{48}$$

In order to compute the critical pycnocline depth and the corresponding northern sinking, an additional equation for

the critical value is needed. A lemma from Galois theory states that the discriminant of a classical polynomial $p_n(x)$ in one variable x can be determined via the first derivative of the polynomial $p_n'(x)$. We have

$$\Upsilon[p] = n^n a_n^{n-1} \prod_{\gamma: p_n'(\gamma)=0} p_n(\gamma), \tag{49}$$

where n is the maximal order of the polynomial and a_n the coefficient of the highest order term (Gelfand et al. 1994, p.404). If the discriminant is zero for a critical choice of coefficients, at least one root of the polynomial p_n coincides with one root of its first derivative p_n' .

The derivative of the governing polynomial of the mixing case (eq. (13)) with respect to F_N yields

$$\frac{\partial D}{\partial F_N} = -\frac{\beta_S S_0 D^4}{4\beta_S S_0 F_N D^3 + 3\alpha_T \Delta\theta C_U D^2}, \tag{50}$$

while the denominator is the derivative of the same polynomial with respect to D . Using the lemma, the discriminant vanishes at the critical point F_M^* and, consequently, the first D -derivative of the polynomial shows a root. Since this derivative appears in the denominator of $\partial D/\partial F_N$, it diverges at F_M^* . In addition, there is only one real value for F_N where the denominator crosses zero and thus the derivative $\partial D/\partial F_N$ is positive for physical D as long as the $F_N < F_M^*$. Finally, by inserting the critical value F_M^* into the derivative of the polynomial with respect to D , an additional equation is obtained that permits the determination of the critical pycnocline depth D_M^*

$$D_M^* = \left(\frac{4C_U}{C_N\rho_0\alpha_T|\Delta\theta|} \right)^{1/3}. \tag{51}$$

The relation $m_N = \frac{C_U}{D}$ provides the key to find the critical northern sinking.

2.2 Wind-driven case

The discriminant for the governing equation of the wind-driven case (eq. (20)) is calculated in a straightforward manner (Gelfand et al. 1994, p.405)

$$\begin{aligned} \Upsilon_W = & 4\{C_N^3\rho_0^3C_W^5\alpha_T^3\Delta\theta^3 \\ & + \left(-\frac{1}{4}C_N^2\rho_0^2C_E^2C_W^4 + 3C_N^3\rho_0^3C_W\beta_S S_0 F_N\right)\alpha_T^2\Delta\theta^2 \\ & + (3C_N^3\rho_0^3C_W\beta_S^2 S_0^2 F_N^2 - 5C_N^2\rho_0^2C_E^2C_W\beta_S S_0 F_N)\alpha_T\Delta\theta \\ & + C_N\rho_0 C_E^4 C_W^2 \beta_S S_0 F_N + 2C_N^2\rho_0^2 C_E^2 C_W^2 \beta_S^2 S_0^2 F_N^2 \\ & + C_N^3\rho_0^3 C_W^2 \beta_S^3 S_0^3 F_N^3\} \end{aligned} \tag{52}$$

This is a third order polynomial in F_N and, in general, there are three roots that determine the critical freshwater input. For our parameter set, only one real root can be

found, which we refer to as F_W^* . In general, F_W^* denotes the largest possible critical freshwater input. Although its solution can be analytically determined, it shows a lack of lucidity and therefore an approximation is presented. It is possible to deduce that the derivative of F_W^* with respect to $\Delta\theta$ is constant in the limit $\Delta\theta \rightarrow -\infty$. To show this, the discriminant Υ_W is divided by $\Delta\theta^3$. Considering both the temperature difference limit and that Υ_W vanishes at the critical point, yield that only a linear term remains. Without the offset, the linear approximation reads

$$(F_W^*)_{appr} \approx -\frac{C_W}{\beta_S S_0} \alpha_T \Delta\theta. \tag{53}$$

It is the temperature difference between the North Atlantic and the equator in combination with the SO winds that create a non-zero critical freshwater flux. The linear estimate for F_W^* approximates the slope of the analytic solution fairly well as long as

$$|\Delta\theta| \gg \frac{C_E^2}{C_N \rho_0 C_W \alpha_T} = 0.5^\circ C. \tag{54}$$

This expression is deduced from $\partial F_W^* / \partial \Delta\theta$, which is obtained by the $\Delta\theta$ -derivative of the discriminant at F_W^* . Knowing that the maximal dependence of F_W^* on $\Delta\theta$ is linear (fact of the approximation), a comparison of the terms in the resulting expression gives this restriction to the applicability of the linearisation.

Besides, using the same line of argument as in the previous section (Gelfand et al. 1994, p.404), the D -derivative of the governing polynomial of the wind-driven case (eq. (20)) is zero at F_W^* . This provides, on one hand, the information that $\partial D / \partial F_N$ is positive for $F_N < F_W^*$ (if there are more than one real roots for F_W^* choose the smallest), and on the other, it provides an additional equation to determine the critical pycnocline depth. Inserting the linear approximation for F_W^* yields

$$D_W^* \approx \frac{C_E^2 - \sqrt{C_E^4 - 6C_N \rho_0 C_E^2 C_W \alpha_T \Delta\theta}}{3C_N \rho_0 C_E \alpha_T \Delta\theta}. \tag{55}$$

The linear approximation without the offset $(F_W^*)_{appr}$ exhibits an additional feature: it serves as an upper boundary for F_W^* for negative $\Delta\theta$. The complete analytic solution for F_W^* as the root of Υ_W has the following structure

$$F_W^* = \xi \cdot \Delta\theta + \eta + \left\{ -\zeta(\Delta\theta) + \sqrt{\chi(\Delta\theta)} \right\}^{1/3} \tag{56}$$

Since ξ is the same constant which was already found in our approximation, the other terms must sum up to a negative value, if the approximation should serve as an upper boundary. The respective terms are calculated via

$$\begin{aligned} \eta &= -\frac{2}{3} \cdot \frac{C_E^2}{C_N \rho_0 \beta_S S_0} \\ \zeta &= \zeta_2 \cdot \Delta\theta^2 + \zeta_1 \cdot \Delta\theta - \zeta_0 \\ \chi &= \chi_4 \cdot \Delta\theta^4 - \chi_3 \cdot \Delta\theta^3 + \chi_2 \cdot \Delta\theta^2 - \chi_1 \cdot \Delta\theta, \end{aligned} \tag{57}$$

which implies that η is negative. Furthermore, χ is positive for all $\Delta\theta \leq 0$, because the constants χ_i are all positive and terms with odd exponents are without exception multiplied by -1 . Thus, if ζ would be positive then the sum of the two cube roots is negative, because one subtracts the root of the sum of ζ and $\sqrt{\chi}$ from the root of their difference. However, ζ is a polynomial of degree two with exclusively positive constants ζ_i with zeros at

$$\Delta\theta_{1/2} = \left(-\frac{10}{27} \pm \frac{2\sqrt{3}}{9} \right) \cdot \frac{C_E^2}{C_N \rho_0 C_W \alpha_T}. \tag{58}$$

For our parameters, this implies that below a value of $\Delta\theta = -0.4^\circ C$ our approximation provides an upper boundary for the critical freshwater flux. Observed ocean temperature differences between the North Atlantic and the equator $\Delta\theta$ differ considerably from zero and are likely on the order of several degrees.

2.3 Full problem

For the full problem, the discriminant for the normalised form of the governing eq. (11) is

$$\begin{aligned} \Upsilon_F &= a^4(-27d^4 + 144cd^2e - 128c^2e^2 - 192bde^2) \\ &+ 2a^3(-2c^3d^2 + 8c^4e - 40bc^2de - 3b^2d^2e \\ &- 18d^3e - 800be^3 + 9bcd^3 + 72b^2ce^2 + 80cde^2) \\ &+ a^2(-6c^2d^3 + 24c^3de - 27b^4e^2 - 50d^2e^2 \\ &+ 2000ce^3 - 4b^3d^3 + 18b^3cde + 144bd^4 \\ &- 746bcd^2e + 560bc^2e^2 + b^2c^2d^2 - 4b^2c^3e \\ &+ 1020b^2de^2) + a(24b^3d^2e \\ &- 630b^3ce^2 + 18bc^3d^2 - 72bc^4e + 160bd^3e \\ &- 2050bcde^2 - 80b^2cd^3 + 356b^2c^2de \\ &- 2250b^2e^3 - 192cd^4 + 1020c^2d^2e - 900c^3e^2 \\ &- 2500de^3) + (-27c^4d^2 + 256d^5 \\ &+ 108c^5e - 1600cd^3e + 108b^5e^2 + 2250c^2de^2 \\ &+ 256a^5e^3 + 3125e^4 + 16b^4d^3 - 72b^4cde \\ &+ 144bc^2d^3 - 630bc^3de + 2000bd^2e^2 - 3750bce^3 \\ &- 128b^2d^4 + 560b^2cd^2e + 825b^2c^2e^2 \\ &- 4b^3c^2d^2 + 16b^3c^3e - 900b^3de^2), \end{aligned} \tag{59}$$

with the following definitions

$$\begin{aligned}
 a &= -\frac{C_E^2 + C_N\rho_0\alpha_T\Delta\theta + C_N\rho_0\beta_S S_0 F_N}{C_N\rho_0 C_E\alpha_T\Delta\theta} \\
 b &= -\frac{C_N\rho_0 C_U\alpha_T\Delta\theta - 2C_E C_W}{C_N\rho_0 C_E\alpha_T\Delta\theta} \\
 c &= -\frac{C_W^2 - 2C_U C_E}{C_N\rho_0 C_E\alpha_T\Delta\theta} \\
 d &= -\frac{2C_U C_W}{C_N\rho_0 C_E\alpha_T\Delta\theta} \\
 e &= -\frac{C_U^2}{C_N\rho_0 C_E\alpha_T\Delta\theta}.
 \end{aligned} \tag{60}$$

Analogous to the previous cases, the derivative of the polynomial (eq. (11)) is zero for the critical freshwater input (Gelfand et al. 1994, p.404). Although this provides the additional equation to determine the critical pycnocline depth, the required F_F^* is not available in an analytic form (only one real F_F^* is observed for our parameter set), as the discriminant Υ_F is a polynomial of degree five in F_N . Moreover, such a polynomial has at least one real root.

Appendix 3: Sensitivity to F_N

The fundamental idea is to show that the *Stommel* model is more sensitive to a change in the freshwater flux F_N than our model. This is done by analysing the derivative of m_N with respect to F_N in the steady state. Again the set of parameters is chosen to be physical (see Appendix 2). The analysis, presented here, can be reduced to the following problem.

Given two functions

$$\begin{aligned}
 f, g &: \mathcal{R} \mapsto \mathcal{R} \\
 f(x) &= -\left(\frac{p}{q-x}\right)^{1/2} \quad g(x) = \frac{r}{x-s(x)},
 \end{aligned}$$

with $x \in \mathcal{R}$ and $p, q, r \in \mathcal{R}^+$. The real function s is referred to as the pole function.

The intersections x_i of these two functions $f(x_i) = g(x_i)$ are determined via the following equation

$$x_i = s(x_i) - \frac{r^2}{2p} \pm \frac{r}{p} \sqrt{p \cdot (q - s(x_i)) + \frac{r^2}{4}}.$$

If the expression within the root is negative, then there exist exclusively imaginary x_i . This implies that if the pole of function f at q has a smaller value than that of g at $s(x)$, and their difference is larger than $r^2/(4p)$, then the two functions do not cross for $x < q$. Assume $q < s(x) - r^2/(4p)$, then

$$g(x) \geq f(x) \quad \forall x \in \mathcal{R} : f(x) \leq 0,$$

The crucial question is which restriction for x is defined by the condition $q < s(x) - r^2/(4p)$.

The derivative $\partial m_N / \partial F_N$ is chosen to measure the sensitivity of the various subcases and they are associated with f and g (x stands for F_N). In the *Stommel* case, the analytic function for m_N is available (compare Sect. 3) and its derivative with respect to F_N is a function of the form f . On the other hand, we find for each subcase of our model an implicit form of the function $\partial m_N / \partial F_N$ which is represented by type g . Rearranging the specific volume flux balance (eq. (1)) gives an expression for m_N . The function $\partial m_N / \partial F_N$ is calculated using the F_N -derivative of the associated governing equation. The mathematical considerations reveal that the pole functions for all subcases of our model must exceed that of the *Stommel* case by more than $r^2/(4p)$, in order to show that the *Stommel* (1961) model is more sensitive to F_N .

$$F_S^{pole} = \frac{1}{4} \frac{\alpha_T \Delta\theta}{\beta_S S_0} C_N \rho_0 \alpha_T \Delta\theta D_S^2 \equiv F_S^* \tag{61}$$

$$F_M^{pole} = -\frac{3}{4} \frac{\alpha_T \Delta\theta}{\beta_S S_0} m_U \tag{62}$$

$$F_W^{pole} = -\frac{1}{2} \frac{\alpha_T \Delta\theta}{\beta_S S_0} (2m_W - 3m_E) + \frac{m_E(m_W - m_E)}{C_N \rho_0 \beta_S S_0 D^2} \tag{63}$$

$$\begin{aligned}
 F_F^{pole} &= -\frac{\alpha_T \Delta\theta}{\beta_S S_0} \left(m_W - \frac{5}{4} m_E + \frac{3}{4} m_U \right) \\
 &\quad - \frac{m_E^2 - \frac{3}{2} m_E m_W + \frac{1}{2} m_W^2 - m_E m_U + \frac{1}{2} m_U m_W}{C_N \rho_0 \beta_S S_0 D^2}.
 \end{aligned} \tag{64}$$

At first, the pole of the mixing-driven case F_M^{pole} is compared to the constant one of the *Stommel* case F_S^{pole} . Since the derivative of m_N with respect to F_N in the mixing-driven case is negative as long as F_N is smaller than F_M^* (cf. function g), m_N decreases strictly monotonic with F_N (see Fig. 4). For $F_N = F_M^{pole}$, this derivative diverges and, consequently, the smallest value for m_N is reached. This minimal value equals the critical freshwater input F_M^* (cf. Appendix 2) which provides an analytic expression. Note that solutions for $F_N > F_M^*$ are not physical (see Sect. 3.1). All in all, the position of the pole in the *Stommel* case F_S^{pole} has to undercut the minimal value for F_N^* corrected by the respective $r^2/(4p)$ in order to be more sensitive. Since $r^2/(4p)$ is in the order of 10^{-3} Sv, it will for the moment be neglected, to get a useful qualitative expression. We find an approximative upper boundary for *Stommel's* prescribed pycnocline depth D_S

$$D_S^{max} \approx \sqrt{3} \cdot \left(-\frac{C_U}{2C_N \rho_0 \alpha_T \Delta\theta} \right)^{1/3}, \tag{65}$$

which has, in our framework (see Table 1), a value of 564.1 m. The exact solution can also be analytically derived

giving constraints of $79.7 \text{ m} \leq D_S \leq 558.4 \text{ m} = D_S^{max}$. The lower analytic bound, is not physical because it causes a negative expression under the square root in eq. (8). The upper analytic bound is lower than our approximative value for all physical parameter sets and is in good agreement with the approximation. The analytic constraint is more stringent and therefore also guarantees that the Stommel (1961) model’s critical freshwater input F_S^{pole} does not exceed that of our mixing-driven case F_M^{pole} . As long as this criteria for the Stommel case is valid, it is sufficient to show that the poles for the wind-driven case F_W^{pole} and the full problem F_F^{pole} surpass the critical value of the mixing-driven case F_M^* . Then Stommel’s model would be more sensitive to a change in freshwater flux F_N than our conceptual framework.

Starting with the wind-driven case, the condition $F_W^{pole} \geq F_M^*$ yields a polynomial of degree two

$$0 \leq 6 \frac{\alpha_T \Delta \theta}{\beta_S S_0} C_E \cdot D^2 + \frac{C_E C_W}{C_N \rho_0 \beta_S S_0} + \left\{ -\frac{C_E^2}{C_N \rho_0 \beta_S S_0} + \frac{\alpha_T \Delta \theta}{\beta_S S_0} \left(3 \left(\frac{1}{4} C_U^2 C_N \rho_0 \alpha_T \Delta \theta \right)^{1/3} - 4 C_W \right) \right\} \cdot D, \tag{66}$$

whose two real roots have opposite signs for a physical choice of parameters. This conclusion can be made by comparing the signs of the different factors in the formula for the roots (not shown). The positive zero is an upper boundary to D in the wind-driven case $D_W^{max} = 913 \text{ m}$. Figure 5a illustrates that for our parameter set, the stable branch of the pycnocline depth does not reach this critical value. The found constraint is merely implicit because all parameters must be set to determine the variable D . It actually indicates a direct restriction on the parameter space. A more rigorous approach makes use of the discriminant Υ_W (eq. (53)). In order to check the sign of the discriminant, the freshwater flux F_N is substituted by the critical value of the mixing-driven case F_M^* . As long as it is negative, three solutions for the wind-driven case exist. This is associated to a freshwater input, which is smaller than the maximal critical value for F_W^* (see Appendix 2.2). The discriminant Υ_W basically reads

$$\Upsilon_W = \sigma_4 \cdot C_E^4 + \sigma_2 \cdot C_E^2 + \sigma_0, \tag{67}$$

while σ_i are parameter dependent constants and σ_4 is positive (cf. Appendix 2.2). Υ_W is a polynomial of degree four in C_E with even exponents which makes it axially symmetric around the ordinate. In general, four zeros are defined by

$$(C_E)_{1/2/3/4} = \pm (\varepsilon_1 \cdot (\varepsilon_2 \pm \sqrt{\varepsilon_3}))^{1/2}, \tag{68}$$

with

$$\begin{aligned} \varepsilon_1 &= \frac{C_N \rho_0}{2 \beta_S S_0 F_M^*} \\ \varepsilon_2 &= \frac{1}{4} C_W^2 \alpha_T^2 \Delta \theta^2 + 5 C_W \alpha_T \Delta \theta \beta_S S_0 F_M^* - 2 \beta_S^2 S_0^2 (F_M^*)^2 \\ \varepsilon_3 &= \frac{1}{16} C_W^4 \alpha_T^4 \Delta \theta^4 - \frac{3}{2} C_W^3 \alpha_T^3 \Delta \theta^3 \beta_S S_0 F_M^* \\ &\quad + C_W^2 \alpha_T^2 \Delta \theta^2 \beta_S^2 S_0^2 (F_M^*)^2 - 32 C_W \alpha_T \Delta \theta \beta_S^3 S_0^3 (F_M^*)^3. \end{aligned} \tag{69}$$

Excluding negative and imaginary values for C_E , only one positive zero remains, because ε_1 is positive and the difference between ε_2 and $\sqrt{\varepsilon_3}$ is negative. The second conclusion is based on the fact that from the first term in ε_2 something is subtracted, while to the first term in ε_3 something is added - as long as $\Delta \theta$ is negative. The graph of Υ_W has a lower boundary, because σ_4 is positive. Since for reasonable positive C_E only one zero exists, Υ_W is negative up to this root $C_E^{max} = 1.2 \cdot 10^{-2} \text{ Sv/m}$. This does not seem to be a severe criterion, because it lies one orders of magnitude above the used value (cf. Table 1). Under this condition, F_W^* is bigger than that of the mixing-driven case F_M^* . Since the function F_W^{pole} is strictly monotonic decreasing with F_N (seen from its derivative), F_W^{pole} is larger than the maximal F_W^* . In other words $F_M^* \leq F_W^* \leq F_W^{pole}$. This ensures that the wind-driven case is overall less sensitive to a change in freshwater than Stommel (1961). To claim that the wind-driven case is less sensitive than the mixing-driven case, it is necessary to compare their two derivatives $\partial m_N / \partial F_N$ directly. For that the analytic solutions of the pycnocline depth in the two cases are needed, which would give a rather complex mathematical constraint. Therefore the analytic form of this constraint is omitted here but it is in principle depicted in Fig. 8. This figure indicates that for our parameter set the derivative in the mixing-driven case is higher than the one in the wind-driven case.

For the full problem an analytic constraint for the parameter set is not possible but an implicit one for the variable D can be provided. A polynomial of degree four arises from the condition $F_F^{pole} \geq F_M^*$

$$\begin{aligned} 0 \leq & -D^4 \cdot 5 \mu C_E + \\ & + D^3 \cdot (4 \mu C_W - 4 F_M^* - 2 \nu C_E^2) \\ & + D^2 \cdot (3 \mu C_U + 3 \cdot \nu C_E C_W) \\ & + D \cdot (2 C_U C_E - C_W^2) \nu \\ & - \nu C_W C_U, \end{aligned} \tag{70}$$

with the positive constants

$$\mu = -\frac{\alpha_T \Delta \theta}{4 \beta_S S_0} \tag{71}$$

$$v = (2C_N\rho_0\beta_S S_0)^{-1}. \quad (72)$$

For our set of parameters, two positive real roots are found, along with a pair of negative roots. To ensure that the inequality holds, the positive roots are the upper and lower limit for the pycnocline depth of the full case $351\text{m} \leq D \leq 1,329\text{m}$. The upper constraint is out of the range of the observed values, while the lower one is identified with a large negative freshwater input in the North Atlantic (cf. Fig. 7). If these constraints for the variable D are fulfilled, the pole function of the full problem F_F^{pole} is smaller than the minimal one of the mixing-driven case F_M^* . For the full problem to be less sensitive than the mixing-driven case additional constraints must hold. They emanate from the direct comparison of their respective derivatives $\partial m_N / \partial F_N$, which include the full analytic solutions for D in both subcases. We omit the complex, analytic form of the constraint since we find that our parameter set does not violate it (cf. Fig. 8).

Altogether, it can be concluded that if the Stommel (1961) model uses a pycnocline depth of $D_S \leq 558.4\text{m}$, then its northern sinking is more sensitive to a change in F_N than in our model. In other words, if the critical freshwater input of the Stommel model is slightly lower (10^{-3} Sv) than that of the mixing-driven case, then our model is less sensitive to a change in freshwater flux - if constraints on the parameter choice are considered. Under further conditions, it is argued that the mixing-driven case is less sensitive than the wind-driven case or even the full problem. All these constraints are fulfilled for our physical set of parameters.

Appendix 4: Salinity equations

Here we present the equilibrium solutions for the box salinities with respect to the determined pycnocline depth D . All box salinities are given in reference to that of the northern box S_N . The system gets consistent if an average salinity for all boxes of S_0 is prescribed, which gives the additional equation to determine S_N

$$(V_N + V_U + V_D + V_S) \cdot S_0 = V_N \cdot S_N + V_U \cdot S_U + V_D \cdot S_D + V_S \cdot S_S. \quad (73)$$

It is furthermore notable, that a direct influence of the entire model geometry is only noticed in the absolute salinities. Their relative differences are only dependent on D , which is a function of the width B and the meridional extent of the low-latitude box L_U . The relative values of the salinities are derived from the system of equations for the equilibrium salinities (see eq. (6)). A brief overview for the different subcases is given.

4.1 Mixing driven case

$$S_U = S_N + \frac{F_N D}{C_U} S_0 \quad (74)$$

$$S_D = S_N \quad (75)$$

In this special case the southern ocean box is disconnected from the others. Its salinity S_S is therefore independent and is only determined by the average ocean salinity.

4.2 Wind driven case

$$S_U = S_N + \frac{F_N}{C_W - C_E D} \cdot S_0 \quad (76)$$

$$S_D = S_N + \frac{C_E D}{C_W} \left(\frac{F_N}{C_W - C_E D} - \frac{F_N + F_S}{C_W} \right) \cdot S_0 \quad (77)$$

$$S_S = S_N + \left(\frac{F_N}{C_W - C_E D} - \frac{F_N + F_S}{C_W} \right) \cdot S_0 \quad (78)$$

4.3 Full problem

$$S_U = S_N + \frac{F_N}{\frac{C_U}{D} + C_W - C_E D} \cdot S_0 \quad (79)$$

$$S_D = S_N + \frac{C_E D}{\frac{C_U}{D} + C_W - C_E D} \cdot \frac{C_E D F_N - \left(\frac{C_U}{D} + C_W - C_E D \right) F_S}{\frac{C_U}{D} C_W + C_W^2 + C_E C_U} \cdot S_0 \quad (80)$$

$$S_S = S_N + \frac{\frac{C_U}{D} + C_W}{\frac{C_U}{D} + C_W - C_E D} \cdot \frac{C_E D F_N - \left(\frac{C_U}{D} + C_W - C_E D \right) F_S}{\frac{C_U}{D} C_W + C_W^2 + C_E C_U} \cdot S_0 \quad (81)$$

References

- Ashkenazy Y, Tziperman E (2007) A wind-induced thermohaline circulation hysteresis and millennial variability regimes. *J Phys Oceanogr* 37:2446–2457
- Bryan F (1987) On the parameter sensitivity of primitive equation ocean general circulation models. *J Phys Oceanogr* 17(7):970–985
- Cai M, Lu J (2007) Dynamical greenhouse-plus feedback and polar warming amplification. Part II: meridional and vertical asymmetries of global warming. *Climate Dynamics* 29(4):375–391
- Clark PU, Pisias NG, Stocker TF, Weaver AJ (2002) The role of the thermohaline circulation in abrupt climate change. *Nat Biotechnol* 415(6874):863–869
- Fyfe JC, Saenko OA, Zickfeld K, Eby M, Weaver AJ (2007) The role of poleward intensifying winds on Southern Ocean warming. *J Climate* 20(21):5391–5400
- Gelfand IM, Kapranov MM, Zelevinsky AV (1994) *Mathematics: theory & applications: discriminants, resultants and multidimensional determinants*. Birkhäuser Boston, 675 Massachusetts Avenue, Cambridge, MA 02139

- Gent PR, McWilliams JC (1990) Isopycnal mixing in ocean circulation models. *J Phys Oceanogr* 20(1):150–155
- Gnanadesikan A (1999) A simple predictive model for the structure of the oceanic pycnocline. *Sci Agric* 283(5410):2077–2079
- Gnanadesikan A, Slater RD, Swathi PS, Vallis GK (2005) The energetics of ocean heat transport. *J Climate* 18(14):2604–2616
- Goswami BN, Madhusoodanan MS, Neema CP, Sengupta D (2006) A physical mechanism for north atlantic sst influence on the indian summer monsoon. *Geophys Res Lett* 33(L02706)
- Griesel A, Morales-Maqueda MA (2006) The relation of meridional pressure gradients to North Atlantic Deep Water volume transport in an OGCM. *Climate Dynamics* 26(7–8):781–799
- Guan YP, Huang RX (2008) Stommel's box model of thermocline circulation revisited—the role of mechanical energy supporting mixing and the wind-driven gyration. *J Phys Oceanogr* 38(4):909–917
- Hofmann M, Rahmstorf S (2009) On the stability of the atlantic meridional overturning circulation. *PNAS* 106(6):20584–20589
- Jeffreys H (1925) On fluid motions produced by differences of temperature and humidity. *QJR Meteorol Soc* 51:347–356
- Johnson HL, Marshall DP (2002) A theory of the surface atlantic response to thermohaline variability. *J Phys Oceanogr* 32(4):1121–1131
- Johnson HL, Marshall DP, Sproson DAJ (2007) Reconciling theories of a mechanically-driven meridional overturning circulation with thermohaline forcing and multiple equilibria. *Climate Dynamics* 29(7–8):821–836
- Kuhlbrodt T, Griesel A, Montoya M, Levermann A, Hofmann M, Rahmstorf S (2007) On the driving processes of the Atlantic meridional overturning circulation. *Rev Geophys* 45(1):RG2001
- Kuhlbrodt T, Rahmstorf S, Zickfeld K, Vikebo FB, Sundby S, Hofmann M, Link PM, Bondeau A, Cramer W, Jaeger C (2009) An integrated assessment of changes in the thermohaline circulation. *Climatic Change* 96(4):489–537
- Landerer FW, Jungclauss JH, Marotzke J (2007) Regional dynamic and steric sea level change in response to the IPCC-A1B scenario. *J Phys Oceanogr* 37(2):296–312
- Laurian A, Drijfhout SS, Hazeleger W, van Dorland R (2009) Global surface cooling: the atmospheric fast feedback response to a collapse of the thermohaline circulation. *Geophys Res Lett* 36(doi:10.1029/2009GL040938)
- Laurian A, Drijfhout SS, Hazeleger W, van den Hurk B (2010) Response of the Western European climate to a collapse of the thermohaline circulation. *Climate Dynamics* 34(5):689–697. doi: 10.1007/s00382-008-0513-4
- Leutbecher A (1996) Zahlentheorie. In: Eine Einführung in die Algebra. Springer, Berlin
- Levermann A, Fürst JJ (2010) Atlantic pycnocline theory scrutinized using a coupled climate model. *Geophys Res Lett* 37(L14602)
- Levermann A, Griesel A (2004) Solution of a model for the oceanic pycnocline depth: scaling of overturning strength and meridional pressure difference. *Geophys Res Lett* 31(L17302)
- Levermann A, Griesel A, Hofmann M, Montoya M, Rahmstorf S (2005) Dynamic sea level changes following changes in the thermohaline circulation. *Climate Dynamics* 24(4):347–354
- Levitus S (1982) Climatological Atlas of the World ocean. In: NOAA Professional Paper, vol 13, US department of commerce, NOAA, Washington DC
- Manabe S, Stouffer RJ (1988) Two stable equilibria of a coupled ocean-atmosphere model. *J Climate* 1:841–866
- Marotzke J (1997) Boundary mixing and the dynamics of three-dimensional thermohaline circulations. *J Phys Oceanogr* 27(8):1713–1728
- Marotzke J, Willebrand J (1991) Multiple equilibria of the global thermohaline circulation. *J Phys Oceanogr* 21(9):1372–1385
- Marotzke J, Welander P, Willebrand J (1988) Instability and multiple steady states in a meridional-plane model of the thermohaline circulation. *Tellus* 40A(2):162–172
- Mignot J, Levermann A, Griesel A (2006) A decomposition of the Atlantic meridional overturning circulation into physical components using its sensitivity to vertical diffusivity. *J Phys Oceanogr* 36:636–650
- Montoya M, Griesel A, Levermann A, Mignot J, Hofmann M, Ganopolski A, Rahmstorf S (2005) The earth system model of intermediate complexity CLIMBER-3 α . Part I: description and performance for present day conditions. *Climate Dynamics* 25(2–3):237–263
- Munk W, Wunsch C (1998) Abyssal recipes II. *Deep-Sea Res I* 45(12):1977–2010
- Nof D, Gorder SV, de Boer A (2007) Does the Atlantic meridional overturning cell really have more than one stable steady state. *Deep-Sea Res I* 54(11):2005–2021
- Park YG (1999) The stability of thermohaline circulation in a two-box model. *J Phys Oceanogr* 28(12):3101–3110
- Prange M, Lohmann G, Paul A (2003) Influence of vertical mixing on the thermohaline hysteresis: analyses of an OGCM. *J Phys Oceanogr* 33(8):1707–1721
- Rahmstorf S (1995) Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle. *Nat Biotechnol* 378(6553):145–149
- Rahmstorf S (1995) Multiple convection patterns and thermohaline flow in an idealized OGCM. *J Climate* 8:3028–3039
- Rahmstorf S (1996) On the freshwater forcing and transport of the Atlantic thermohaline circulation. *Climate Dynamics* 12(12):799–811
- Rahmstorf S, Crucifix M, Ganopolski A, Goosse H, Kamenkovich I, Knutti R, Lohmann G, Marsh B, Mysak LA, Wang Z, Weaver A (2005) Thermohaline circulation hysteresis: a model intercomparison. *Geophys Res Lett* 32(L23605)
- Robinson AR (1960) The general thermal circulation in equatorial regions. *Deep-Sea Res* 6(4):311–317
- Schewe J, Levermann A (2010) The role of meridional density differences for a wind-driven overturning circulation. *Climate Dynamics* 34:547–556
- Schlesinger M, Yin J, Yohe G, Andronova N, Malyshev S, Li B (2006) Assessing the risk of a collapse of the Atlantic thermohaline circulation. Cambridge University Press *Avoiding Dangerous Climate Change* (J. Schellnhuber)
- Schmittner A (2005) Decline of the marine ecosystem caused by a reduction in the Atlantic overturning circulation. *Nat Biotechnol* 434(7033):628–633
- Stommel H (1961) Thermohaline convection with two stable regimes of flow. *Tellus* 13(2):224–230
- Stouffer RJ, Yin J, Gregory JM, Dixon KW, Spelman MJ, Hurlin W, Weaver AJ, Eby M, Flato GM, Hasumi H, Hu A, Jungclauss JH, Kamenkovich IV, Levermann A, Montoya M, Murakami S, Nawrath S, Oka A, Peltier WR, Robitaille DY, Sokolov AP, Vettoretti G, Weber SL (2006) Investigating the causes of the response of the thermohaline circulation to past and future climate changes. *J Climate* 19:1365–1387
- Thual O, McWilliams JC (1992) The catastrophe structure of thermohaline convection in a tow-dimensional fluid model and a comparison with low-order box models. *Geophys Astrophys Fluid Dynamics* 64(1–4):67–95
- Timmermann A, An SI, Krebs U, Goosse H (2005) ENSO suppression due to weakening of the North Atlantic thermohaline circulation. *J Climate* 18(16):3122–3139
- Toggweiler JR, Russell J (2008) Ocean circulation in a warming climate 451(7176):286–288

- Toggweiler JR, Samuels B (1998) On the ocean's large scale circulation in the limit of no vertical mixing. *J Phys Oceanogr* 28(9):1832–1852
- Vellinga M, Wood RA (2002) Global climatic impacts of a collapse of the Atlantic thermohaline circulation. *Climatic Change* 54(3):251–267
- Vellinga M, Wood RA (2007) Impacts of thermohaline circulation shutdown in the twenty-first century. *Climatic Change* 91(1-2):43–63
- Yin J, Stouffer RJ (2007) Comparison of the stability of the Atlantic thermohaline circulation in two coupled atmosphere-ocean general circulation models. *J Climate* 20(17):4293–4315
- Yin J, Schlesinger M, Stouffer RJ (2009) Model projections of rapid sea-level rise on the northeast coast of the united states. *Nature Geosci* 2(4):262–266